

Pesticide Risk Mitigation Engine

Aquatic Risk Indices

White Paper

11/09/2010

Aquatic Invertebrate Risk Index..... 3

 Background 3

 Data Sources 3

 Index Structure 4

 Use Pattern Adjustment Factors..... 8

 Literature Cited 8

Algal Risk Index 10

 Background 10

 Data Sources 10

 Index Structure 11

 Use Pattern Adjustment Factors..... 14

 Literature Cited 14

Fish Chronic Risk Index..... 16

 Background 16

 Data Sources 16

 Index Structure 20

 Literature Cited 21

Appendix A: Peer Review Comments 22

Aquatic Invertebrates Risk Index

Background

Most of the regulatory risk assessments carried out on the effects of toxicants in the aquatic environment are based on extrapolation methods that employ single species laboratory data on *Daphnia* or other indicator species, with survival (most frequently), growth and reproduction being the usual endpoints. Occasionally, simulated (mesocosms) or actual field tests are carried out in addition to the single species data. This allows for a verification of two large uncertainties in the risk assessment: 1) How representative the single species tests are to a more complex and diverse aquatic community; and 2) How representative impacts seen in a laboratory setting are to impacts in a pond or simulated pond environment where physical and biological conditions can modify a pesticide's toxic action.

The proposed PRiME index will rely on three distinct sources of information:

1. An estimate of the peak concentrations of pesticides from both drift and runoff expected into receiving water bodies (this will be covered in detail in another white paper)
2. Laboratory acute toxicity data for crustacean species
3. An analysis of the existing corpus of small pond or mesocosms studies where the response of the crustacean community to known pesticide inputs has been quantified

The index will represent the acute risk that a pesticide represents to the crustacean community of receiving waters.

Data Sources

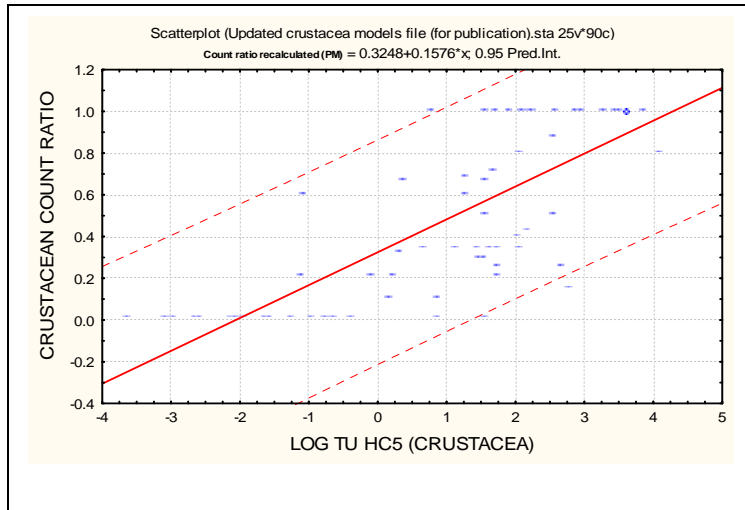
All components of the aquatic index require a representative toxicity endpoint, e.g., the level where toxic effects are expected to occur. Due to general availability of data from a broad range of species, especially for older pesticides, a species sensitivity distribution approach is proposed (described in detail in Whiteside *et al.* 2006, 2008). A series of steps are used to derive separate Hazardous Concentration values (HC₅; i.e. those at the 5% tail of a cumulative frequency distribution). This is carried out here for crustacean; values are to be generated also for fish, insects and algae. Where inadequate sample size does not permit use of a standard species sensitivity distribution approach, the available data are used to estimate the mean of the distribution and calculated inter-species generic variances (Whiteside 2008) are used to values comparable to those generated with a full species sensitivity distribution (Van Vlaardingen *et al.* 2004).

A preliminary analysis and modeling of aquatic pond and mesocosm studies has been carried out and was reported in Singh (2007), and further refined in Mineau *et al.* (2009). A large literature review identified 60 studies representing 184 experiments on 33 pesticides – although not all of these studies were suitable for this index. Only studies with sound experimental designs and with quantification of effects (either explicitly stated or interpreted from figures provided) were retained. In the final analysis, sixty-nine independent water bodies contributed to the models. The data points were extracted from 28 studies examining 21 pesticides: one fungicide, five herbicides and 15 insecticides. By and large, Cladocera, Calanoida and Cyclopoida were the most studied crustaceans with pesticide effect data for these zooplankton orders found in 89% of studies. The effects of pesticides on Amphipoda and Ostracoda were followed in 30% of the studies.

An estimation of the likely drift and runoff of pesticide from single-field treatments will be described in a separate document.

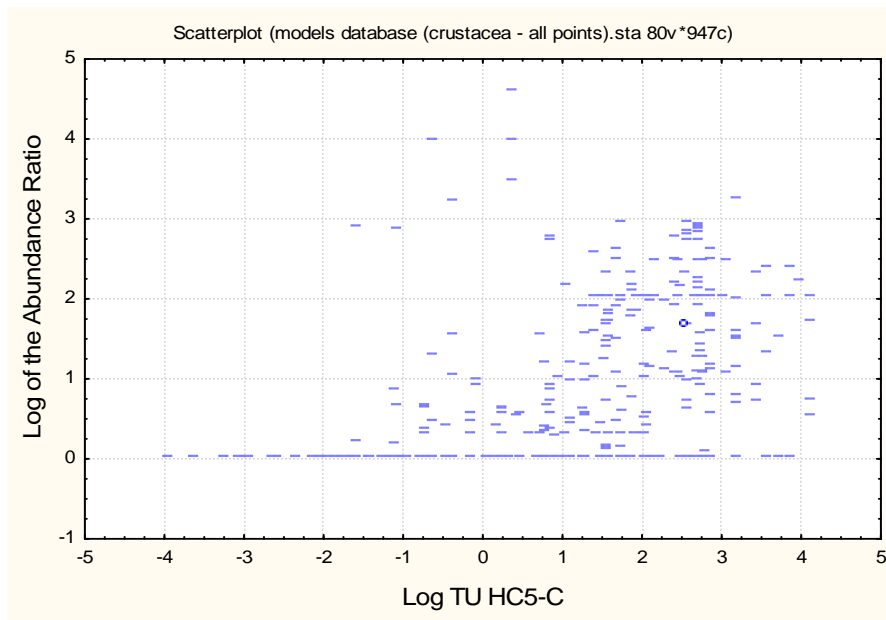
Index Structure

A simple risk quotient (RQ) can be constructed to relate predicted end of field concentrations to toxicological endpoints. An RQ provides an estimate of whether concentrations predicted to be lethal to various taxa are exceeded. It does not provide a strong basis to assess the magnitude of the impact or determine whether this impact is biologically meaningful. Rather, our approach will be to ‘calibrate’ RQs by constructing models relating total abundance or biomass of selected taxa (expressed as a ratio of treated to control pond) or the proportion of impacted species (the proportion showing significant declines) to toxicity/water concentration variables. For example, the following figure shows the regression between the log number of toxicity units (TUs) calculated with the crustacean HC₅ value against the crustacean species count ratio (the proportion of significantly affected species – or families in some taxa) for a sample of 69 ponds and mesocosms representing 21 pesticides. 95% prediction intervals are shown. Only pond and mesocosms without fish are included here as it has been shown that the presence of fish can complicate the interpretation of the results.

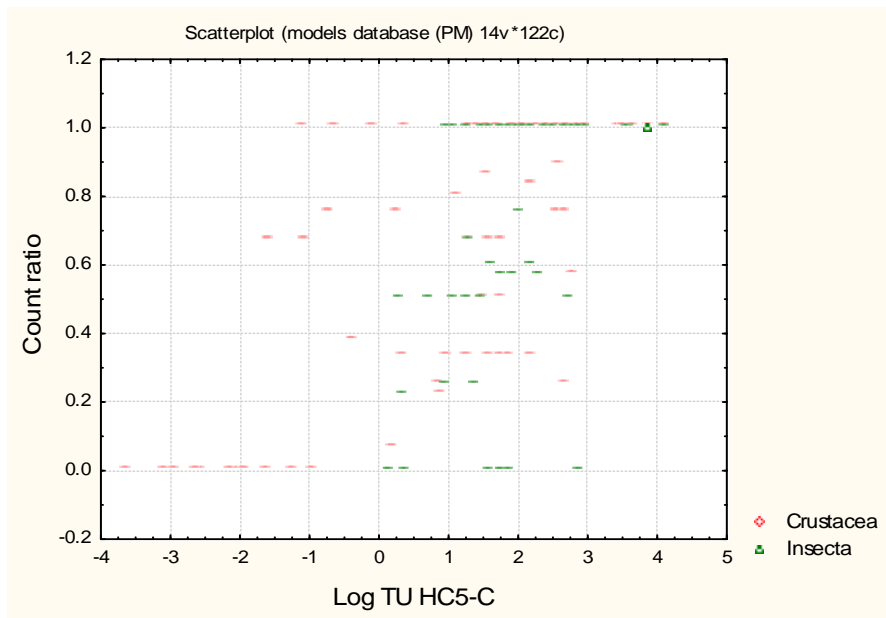


It should be noted that the number of TUs delivered to the ponds or mesocosms can be calculated directly because peak water concentrations are measured in all of the studies. In the PRiME index, this step will depend on a good estimation of peak water concentrations. A risk quotient based on water concentrations is the only approach possible if we are to consider all possible routes of exposure to the aquatic environment, notably drift and surface runoff. Models exist to estimate both of these inputs separately, but their separate contributions must be added and the result expressed as a water concentration if they are both to be considered in our scores.

The following plot shows all abundance changes recorded for all measured taxonomic groups in all pond and mesocosm studies. The points are not strictly independent but this allows visualizing all the available data. Two observations can be made: 1) the abundance data show a similar step function as shown with the count ratios above. Large abundance changes and an increase in the proportion of affected species begin around 0.1 to 1 Toxic Units (-1 to 0 on log scale) based on the crustacean HC5 values. By 10 Toxic Units or more (1 on log scale), disturbances to the aquatic systems are typically serious; 2) abundance changes are substantial – commonly higher than 10 fold and as high as 1000 fold – when the studies record that some species have been significantly affected. This is an indication of the typically low power of mesocosm studies to detect significant effects on account of the high natural variability of those systems.

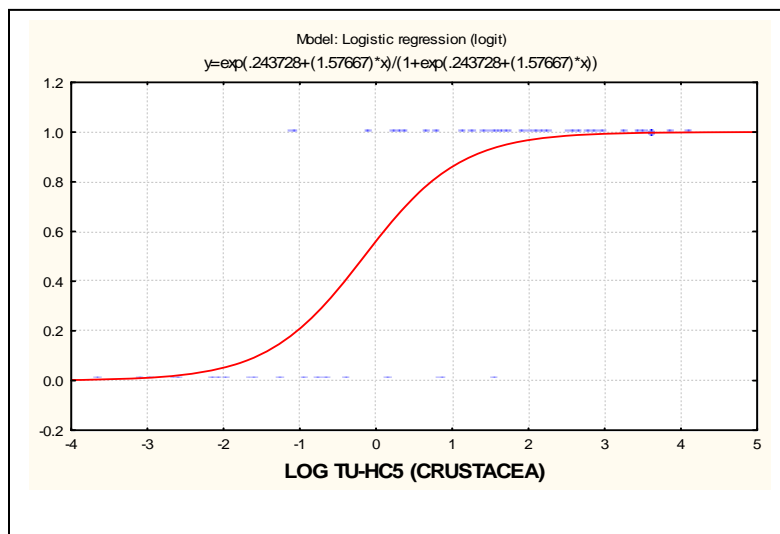


Fewer studies are available on aquatic insect species. In order to see whether a score based on crustacean data would also be protective of insects, the crustacean species count ratio and the equivalent count ratios for insect (on a family basis) are both plotted against the crustacean TU index.



The impact on insects is not well predicted by the number of crustacean TUs – emphasizing the need to develop insect-specific toxicity measures. However, it appears that a scheme based on crustacean TU values and sufficiently protective of crustacea will *de facto* protect insects. We can therefore carry on with the development of a crustacean-based score with the knowledge that it will be protective of most aquatic invertebrates. (However, there has been relatively little testing of some insect larval groups potentially very sensitive to pesticides. As our understanding of trait-based toxicological sensitivity increases, it may be necessary to revise this position.)

In order to work with the apparent non-linearity of the data and to transform the score into a probability of adverse outcome in line with the avian and mammalian scores, we defined a count ratio of 0.1 (10% of species being significantly affected by treatment) as an adverse outcome. Realistically, the distribution of points in the dataset assembled does not allow us to distinguish between 10% and 20% of species being significantly affected. As more field data are made available, it will be possible to refine these estimates. The count ratio data can therefore be re-plotted as follows:



The final algorithm proposed for our aquatic invertebrate indicator gives P as the probability that an application will give rise to an undesirable outcome defined as 10% of crustacean species being significantly affected:

$$p = \left(\frac{e^{a+bx}}{1 + e^{a+bx}} \right)$$

...where x is the number of TUs (Toxic Units) expressed as:

- Peak pesticide concentration / Crustacean HC₅
- a = 0.243728
- b = 1.57667

Use Pattern Adjustment Factors

Unlike the terrestrial indicators, UPAFs are not used to change the above algorithm or modify the probability of impact once peak water concentrations have been measured. Any mitigation of runoff or drift will be introduced in the water concentration modeling phase described elsewhere.

Literature cited

- Mineau, P., T. Dawson, M. Whiteside, C. Morrison, K. Harding, L. Singh, T. Längle, and D.A.R. McQueen. 2009. Environmental Risk-based Standards for Pesticide Use in Canada. National Agri-Environmental Standards Initiative Synthesis Report. Environment Canada, 94pp.
- Singh, L. 2007. Developing Risk-based Rankings for Pesticides in Support of Standard Development at Environment Canada: Predicting Effects of Pesticides in Freshwater Aquatic Ecosystems – Using Field Data Validation. National Agri-Environmental Standards Initiative Technical Series Report No. 3-31. 207 p.

- Van Vlaardingen, P.L.A., T.P. Traas, A.M. Wintersen and T. Aldenberg. 2004. ETX2.0. A program to calculate hazardous concentrations and fraction affected, based on normally-distributed toxicity data. RIVM report (and software) 601501028/2004.
- Whiteside, M., P. Mineau, C. Morrison and K. Harding. 2006. Developing risk-based rankings for pesticides in support of standard development at Environment Canada: Setting pesticide priorities for aquatic ideal performance standards and ranking the aquatic hazard of in-use pesticides in Canada. National Agri-Environmental Standards Initiative Technical Series Report No. 2-42, Environment Canada, 173 pp.
- Whiteside, M., P. Mineau, C. Morrison, and L.D. Knopper. 2008. Comparison of a score-based approach with risk based ranking of in-use agricultural pesticides in Canada to aquatic receptors. *Integrated Environmental Assessment and Management* 4(2):215-236.

Algal Risk Index

Background

Most of the regulatory risk assessments carried out on the effects of toxicants in the aquatic environment are based on extrapolation methods that employ single species laboratory data on various algal indicator species, with growth and reproduction being the usual endpoints. Occasionally, simulated (mesocosms) or actual field tests are carried out in addition to the single species data. This allows for a verification of two large uncertainties in the risk assessment: 1) How representative the single species tests are to a more complex and diverse algal community; and 2) How representative impacts seen in a laboratory setting are to impacts in a pond or simulated pond environment where physical and biological conditions can modify a pesticide's toxic action.

The proposed PRiME index will rely on three distinct sources of information:

1. An estimate of the peak concentrations of pesticides from both drift and runoff expected into receiving water bodies (this will be covered in detail in another white paper)
2. Laboratory acute toxicity data for algal species
3. An analysis of the existing corpus of small pond or mesocosms studies where the response of the algal community to known pesticide inputs has been quantified

The index will represent the acute risk that a pesticide represents to the algal community of receiving waters.

Data Sources

All components of the aquatic index require a representative toxicity endpoint, e.g., the level where toxic effects are expected to occur. Due to general availability of data from a broad range of species, especially for older pesticides, a species sensitivity distribution approach is proposed (described in detail in Whiteside *et al.* 2006, 2008). A series of steps are used to derive separate Hazardous Concentration values (HC₅; i.e. those at the 5% tail of a cumulative frequency distribution). This is carried out here for algae; values are to be generated also for fish and crustacea. Where inadequate sample size does not permit use of a standard species sensitivity distribution approach, the available data are used to estimate the mean of the distribution and calculated inter-species generic variances (Whiteside 2008) are used to values comparable to those generated with a full species sensitivity distribution (Van Vlaardingen *et al.* 2004).

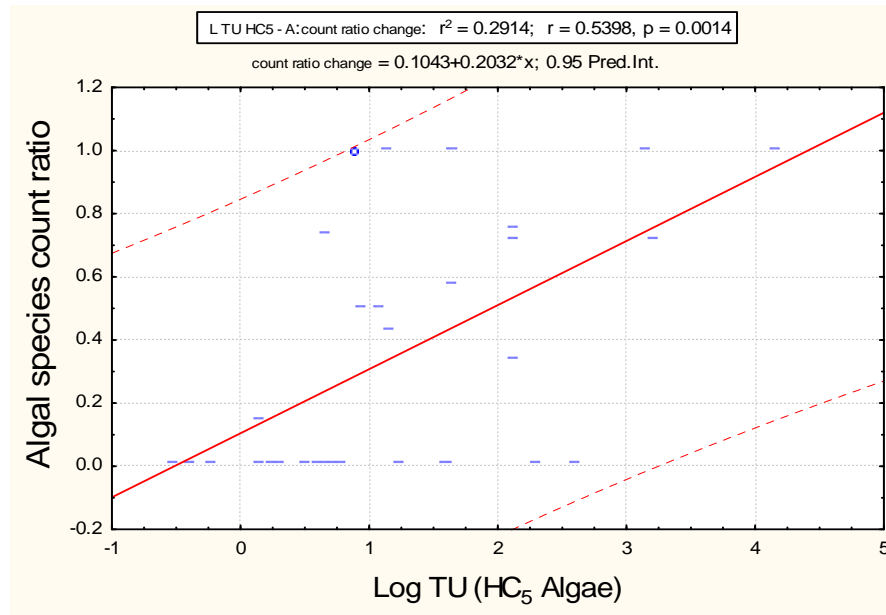
A preliminary analysis and modeling of aquatic pond and mesocosm studies has been carried out and was reported in Singh (2007), and further refined in Mineau *et al.* (2009). A large literature review identified 60 studies representing 184 experiments on 33 pesticides – although not all of these studies were suitable for this index. Only studies with sound experimental designs and with quantification of effects (either explicitly stated or interpreted from figures provided) were retained.

An estimation of the likely drift and runoff of pesticide from single-field treatments will be described in a separate document.

Index Structure

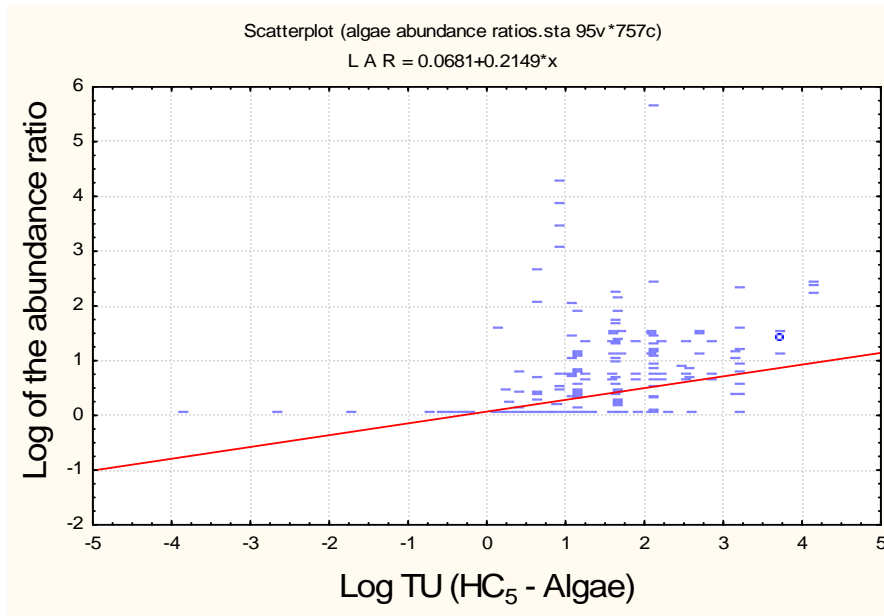
A simple risk quotient (RQ) can be constructed to relate predicted end of field concentrations to toxicological endpoints. An RQ provides an estimate of whether concentrations predicted to be lethal to various taxa are exceeded. It does not provide a strong basis to assess the magnitude of the impact or determine whether this impact is biologically meaningful. Rather, our approach will be to ‘calibrate’ RQs by constructing models relating the proportion of impacted species (the proportion showing significant declines) to toxicity/water concentration variables.

For example, the following figure shows the regression between the log number of toxicity units (TUs) calculated with the algal HC₅ value against the algal species count ratio (the proportion of significantly affected species – or families in some taxa) for a sample of 32 pond and mesocosm studies representing 9 herbicides. Only studies with herbicides were retained because often cause massive increases of phytoplankton in response to effects on grazer communities, i.e. crustacean or aquatic insects primarily. 95% prediction intervals are shown.

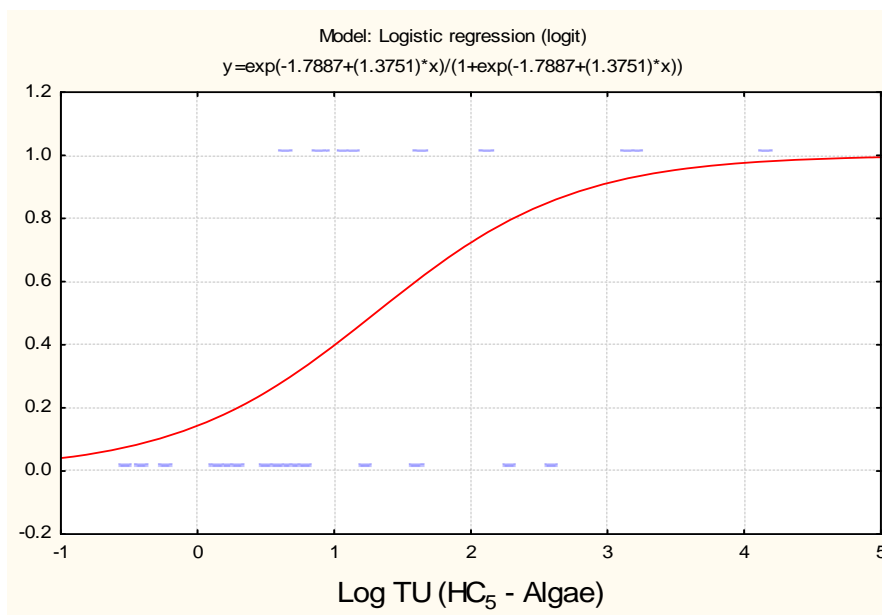


It should be noted that the number of TUs delivered to the ponds or mesocosms can be calculated directly because peak water concentrations are measured in all of the studies. In our tool, this step will depend on a good estimation of peak water concentrations. A risk quotient based on water concentrations is the only approach possible if we are to consider all possible routes of exposure to the aquatic environment, notably drift and surface runoff. Models exist to estimate both of these inputs separately, but their separate contributions must be added and the result expressed as a water concentration if they are both to be considered in our scores.

The following plot shows all abundance changes recorded for all measured taxonomic groups in all pond and mesocosm studies. The points are not strictly independent but this allows visualizing all the available data. Two observations can be made: 1) the abundance data show a similar step function as shown with the count ratios above. Large abundance changes and an increase in the proportion of affected species begin around 1 to 10 Toxic Units (0 to 1 on log scale) based on the algal HC5 values. Abundance changes in affected taxa can be substantial – commonly higher than 10 fold and as high as 100,000 fold. This is an indication of the typically low power of mesocosm studies to detect significant effects on account of the high natural variability of those systems.



In order to work with the apparent non-linearity of the data and to transform the score into a probability of adverse outcome in line with the avian and mammalian scores, we defined a count ratio of 0.2 (20% of species being significantly affected by treatment) as an adverse outcome. Realistically, the distribution of points in the dataset assembled does not allow us to distinguish between 20% and 30% of species being significantly affected. As more field data are made available, it will be possible to refine these estimates. The count ratio data can therefore be re-plotted as follows:



The final algorithm proposed for the algal indicator gives P as the probability that an application will give rise to an undesirable outcome defined as 10% of algal species being significantly affected:

$$p = \left(\frac{e^{a+bx}}{1 + e^{a+bx}} \right)$$

...where x is the number of TUs (Toxic Units) expressed as:

- Peak pesticide concentration / Algal HC₅
- a = -1.7887
- b = 1.3751

Use Pattern Adjustment Factors

Unlike the terrestrial indicators, UPAFs are not used to change the above algorithm or modify the probability of impact once peak water concentrations have been measured. Any mitigation of runoff or drift will be introduced in the water concentration modeling phase.

Literature cited

Mineau, P., T. Dawson, M. Whiteside, C. Morrison, K. Harding, L. Singh, T. Längle, and D.A.R. McQueen. 2009. Environmental Risk-based Standards for Pesticide Use in Canada. National Agri-Environmental Standards Initiative Synthesis Report. Environment Canada, 94pp.

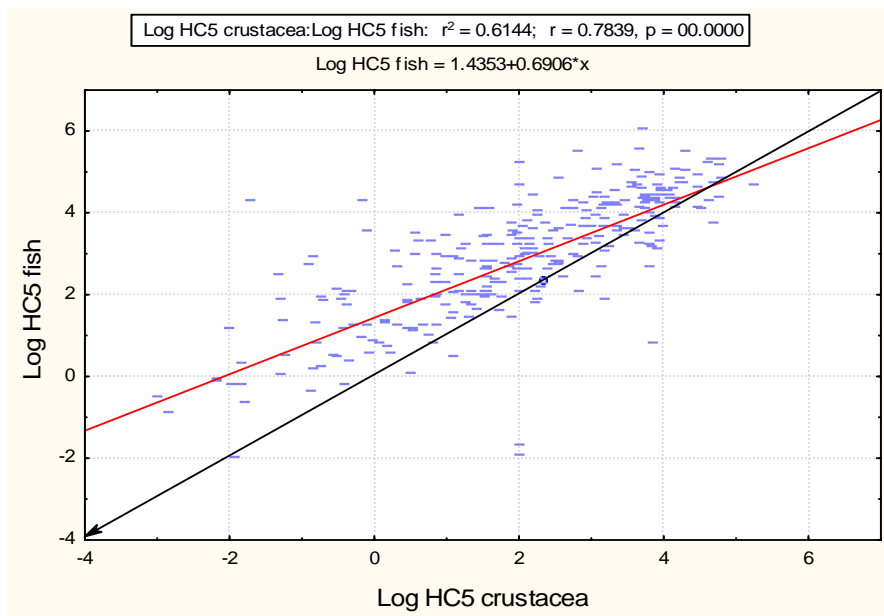
- Singh, L. 2007. Developing Risk-based Rankings for Pesticides in Support of Standard Development at Environment Canada: Predicting Effects of Pesticides in Freshwater Aquatic Ecosystems – Using Field Data Validation. National Agri-Environmental Standards Initiative Technical Series Report No. 3-31. 207 p.
- Van Vlaardingen, P.L.A., T.P. Traas, A.M. Wintersen and T. Aldenberg. 2004. ETX2.0. A program to calculate hazardous concentrations and fraction affected, based on normally-distributed toxicity data. RIVM report (and software) 601501028/2004.
- Whiteside, M., P. Mineau, C. Morrison and K. Harding. 2006. Developing risk-based rankings for pesticides in support of standard development at Environment Canada: Setting pesticide priorities for aquatic ideal performance standards and ranking the aquatic hazard of in-use pesticides in Canada. National Agri-Environmental Standards Initiative Technical Series Report No. 2-42, Environment Canada, 173 pp.
- Whiteside, M., P. Mineau, C. Morrison, and L.D. Knopper. 2008. Comparison of a score-based approach with risk based ranking of in-use agricultural pesticides in Canada to aquatic receptors. *Integrated Environmental Assessment and Management* 4(2):215-236.

Fish Chronic Risk Index

Background

Fish are poorly covered in the pond and mesocosms studies that form the basis of the aquatic invertebrate and algal indices. Indeed, because of problems encountered in mesocosm studies when fish are present, studies with fish were not used in deriving the other aquatic indices and are not reliable for deriving a fish index.

From an acute toxicity point of view, we believe that fish will be sufficiently protected provided we provide sufficient protection to aquatic invertebrates. The following plot shows the relative fish and crustacean HC₅ values (the 5% tail of a cumulative frequency distribution of LC₅₀ values in ppb for the entire taxon) for a large number of pesticides, both currently registered and historical. In most cases, an acute index based on the protection of crustacea should be reasonably protective of fish.



We propose to make the fish index one of chronic toxicity. This will provide a parallel to the terrestrial indices where we have developed avian acute and avian chronic indices; the latter designed to flag compounds that have a long half life in the environment.

Data sources

MATC values – the geometric midpoint between lifecycle NOAEC and LOAEC values – are the desired endpoints. These values are either obtained directly from a fish full life cycle test or, more often from an early life test, a shorter test which has become the norm for pesticide testing. It has been suggested that these two tests are functionally equivalent and give MATC

values within a factor of two of one another (McKim 2005). However, this conclusion differs dramatically from the earlier opinion of Suter and colleagues (1987) who found that fecundity (which can only be measured from a full life cycle test), and not early life stage survival was the most sensitive effect. Suter and colleagues had argued that a full life cycle test was needed. McKim (op. cit.) does not cite that earlier study by Suter et al. so it is difficult to assess these relative claims.

When either is unavailable, it is customary to estimate the MATC from an acute to chronic extrapolation. For example, the acute to chronic extrapolation used by NRCS in the WIN PST program is the relation between the EC₂₅ hatching rate and the LC₅₀ obtained from Barnthouse, Suter and Rosen (1990) based on a sample of 31 studies.

$$\text{Log (Hatch EC}_{25} \text{ in ppb)} = 1.1 * (\text{log LC}_{50} \text{ in ppb)} - 1.2$$

However, use of this formula alone to approximate a measured MATC does introduce additional uncertainty because measured MATCs might reflect a number of different endpoints. Barnthouse and colleagues give three more extrapolations:

$$\text{Log (Parental LC}_{25} \text{ in ppb)} = 0.87 * (\text{log LC}_{50} \text{ in ppb)} - 0.87 \text{ (N = 28)}$$

$$\text{Log (Larval LC}_{25} \text{ in ppb)} = 1.0 * (\text{log LC}_{50} \text{ in ppb)} - 0.89 \text{ (N = 89)}$$

$$\text{Log (EC}_{25} \text{ eggs per surviving fish in ppb)} = 1.1 * (\text{log LC}_{50} \text{ in ppb)} - 1.89 \text{ (N = 42)}$$

As indicated by Suter et al. (1987), early life survival is not necessarily the most sensitive endpoint of a full life-cycle fish test. Fecundity was typically affected at lower concentrations. Therefore, an alternative acute-chronic regression that would seem to be more protective of fish reproduction is the relationship between LC₅₀ and a measured MATC from the available full life cycle tests (N=55).

$$\text{Log (MATC in ppb)} = 0.90 * (\text{log LC}_{50} \text{ in ppb)} - 1.16$$

Because all these regressions included a wide range of contaminants – not just pesticides – I ran the analysis using the NRCS database of measured full life cycle tests (N=17) and early life tests (N=157) combined. These were matched where possible to LC₅₀ data compiled from a number of sources including USEPA one liner data, Ecotox, Pesticide Manual and European registration data.

In choosing from available LC₅₀ data, the following rules were followed:

- Only technical a.i. data were used
- Exact species match was required
- 96 hr. data were used where available
- Flow through data were used in preference to static tests
- Obvious outlier values among replicate studies within a given species were excluded
- A geometric mean of suitable replicate values was retained

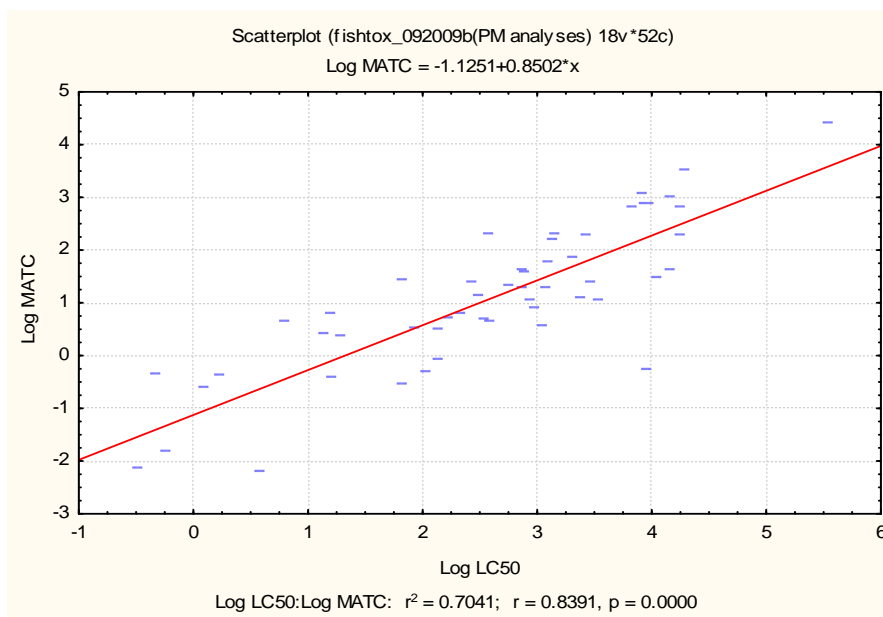
A total of 52 pairs of data points were available for the regression.

In addition, HC₅ values based on SSDs were also tabulated. These are the values given in Whiteside et al. (2008). Detailed methods are given therein.

The regression equation (see plot below) for the NRCS pesticide dataset (download obtained sept 09) is:

$$\text{Log (MATC in ppb)} = 0.85 (\text{log LC}_{50} \text{ in ppb}) - 1.125$$

... which is very close to the regression obtained by Suter and colleagues for their sample of full life cycle studies but with a broader range of chemicals. The regression is shown below.

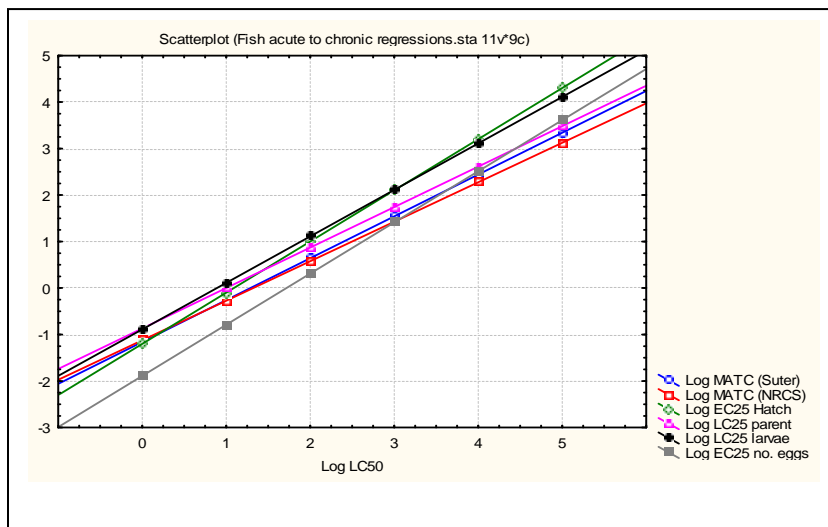


We recognize that this last regression (labeled Log MATC NRCS on plot below) may contain more uncertainty than the careful comparisons of Suter and colleagues. The match of static and

flow through tests may not always be exact; the time interval, although 96h in the vast majority of cases did vary in a few cases. Nevertheless, the similarity between all the regressions is encouraging.

All of these regressions were plotted against a range of fictitious LC_{50} values from 1 to 100,000 ppb. This plot is shown below. The important observation is that the slopes of the different regression equations differ; therefore the curves are not parallel and intersect one another at different values of the LC_{50} . This means that no one algorithm gives the consistently most protective estimate of the MATC across the whole range of possible LC_{50} values.

At lower LC_{50} values, the algorithm based on the egg production EC_{25} endpoint gives approximated MATC values that are much lower than any of the other regressions (as already concluded by Suter and colleagues). The curves appear reasonably close on the plot but it is important to recall that this is a log-log plot. At either end of the LC_{50} continuum modeled here, there is an approximate 10 fold difference in MATC estimates.



One reason for wanting to pick a protective estimate of the MATC is that Barnthouse, Suter and colleagues (1987) warn that problems inherent to the concept of MATC (chiefly its dependence on study design and dose levels chosen) mean that the MATC *per se* is not necessarily protective of fish populations and could be higher than an EC_{10} population effect value.

Also, in keeping with principles established for all the PRIME indices, it is important to choose an estimation method that carries the least amount of bias. A common strategy is to estimate the MATC from the lowest 96 hr. LC_{50} available. This introduces a severe bias in that more heavily tested pesticides (usually older products) are likely to have had a test carried out with a

sensitive species. In contrast, experimentally-determined MATC values are typically obtained for a single species only – which is unlikely to be the most sensitive species.

In order to provide a protective MATC and, more importantly, reduce any bias associated with unequal testing among different pesticides, the following strategy is proposed:

- As with many of the other PRiME indices, HC₅ values – the lower 5% tail of a species sensitivity distribution for all fish data – will be calculated (Van Vlaardingen et al. 2004). Where needed, small sample approaches will be used as described in Whiteside et al. (2008). In theory, this provides the best unbiased indication of the LC₅₀ value for a sensitive species in receiving waters.
- All available regressions given above will be run from the HC₅ value. The lowest MATC estimate obtained will be retained.
- If available, measured MATC values will be compared to the estimated MATC values. Measured MATC values will be used only if lower than the lowest estimated value.

Index structure

We propose to use this minimum MATC as the relevant endpoint of the index and to assess the amount of time that residues in the aquatic environment are above the MATC.

The final score will be expressed as a proportion of the reproduction period where MATC values are exceeded. This is analogous to the derivation of the avian chronic index. We propose that a 1 month time interval would be a suitable approximation for the typical fish breeding season and the denominator for this index. This is the approximate duration of fish life-cycle tests. At this point, we do not propose to match specific fish breeding phenology with dates of pesticide applications. The wide range in fish breeding seasons makes it difficult to do this. Therefore, the index will, by necessity, represent the theoretical risk to fish species whose breeding seasons coincide with the period of pesticide application.

The score will therefore be the number of days that water concentration exceeds the MATC expressed as a proportion of a 30 day interval. In other words, the worst possible score of 1 would be a pesticide that causes a reproductive threshold to be exceeded for the entire 30 day period.

Literature cited

Barnthouse, L.W., G.W. Suter II and A.E. Rosen. 1990. Risks of toxic contaminants to exploited fish populations: Influence of life history, data uncertainty and exploitation intensity. *Env. Toxicol. Chem.* 9:297-311.

McKim, J.M. 1995. Appendix B: Early life stage toxicity tests. In: G.M. Rand. (Ed.) *Fundamentals of Aquatic Toxicology (Second Edition)*, Taylor and Francis, North Palm Beach, Florida, pp. 974-1011.

Suter, G.W., A.E. Rosen, E. Linder, and D.F. Parkhurst. 1987. Endpoints for responses of fish to chronic toxic exposures. *Env. Toxicol. Chem.* 6:793-809.

Van Vlaardingen, P.L.A., T.P. Traas, A.M. Wintersen and T. Aldenberg. 2004. ETX2.0. A program to calculate hazardous concentrations and fraction affected, based on normally-distributed toxicity data. RIVM report (and software) 601501028/2004.

Whiteside, M., P. Mineau, C. Morrison, and L.D. Knopper. 2008. Comparison of a score-based approach with risk based ranking of in-use agricultural pesticides in Canada to aquatic receptors. *Integrated Environmental Assessment and Management* 4(2):215-236.

Appendix A: Peer Review Comments

This white paper was reviewed by the following independent experts. Below are their comments, listed anonymously, along with the author's responses.

- **Katherine Palmquist**, senior scientist, Exponent
- **John Stark**, professor of ecotoxicology, Washington State University

General comments:

- It is my opinion that PRiME will be a useful tool for the relative ranking of pesticide risk that can be used by growers, farm advisors food producers, IPM teams and others to make decisions on which pesticides to use.
- The PRiME model appears to be a simple, useful tool to predict the risk of agriculturally-applied pesticides, and the fact that it provides guidance in terms of mitigation further enhances the usefulness of the model application.

Detailed comments and responses:

Comment 1: For the aquatic invertebrate risk index, crustaceans were used as the species of concern. It is not listed which crustaceans were used in this analysis; I presume that Daphnids, grass shrimp, and amphipods make up the majority of the organisms used in the program. More detail on the species evaluated would be helpful.

Response: The complete list of tested species would be a very large list, each active ingredient potentially having different species represented in the species sensitivity distribution. It could, however, easily be compiled if needed. A summary statement was prepared for publication of the crustacean model (the paper is about to be sent for review). The following summary statement is taken from that draft publication and will be added to the white paper:

“By and large, the Cladocera, Calanoida and Cyclopoida were the most studied crustaceans with pesticide effect data for these zooplankton orders found in 89% of studies. The effects of pesticides on Amphipoda and Ostracoda were followed in 30% of the studies.” I will also add more information on the size of the dataset.

Comment 2: The choice of crustaceans as the key organisms of concern appears to me to be valid as this group of organisms is usually quite sensitive to pesticides, particularly insecticides. Thus, protecting crustaceans should provide protection for other aquatic invertebrates. In fact, it is shown in the fish index that crustaceans can be used to protect fish. However, one concern is that there is far less aquatic insect-pesticide data in the literature than for crustaceans and that these species, especially the larval forms of mayflies, stoneflies and dragon flies may be particularly sensitive to pesticides, maybe more so than crustaceans.

Response: We agree. It was our intent to include both groups but, as pointed out, insect data is much more sparse. The data we do have does not suggest that insects will be sensitive but, the point is taken that the highly sensitive groups are not tested often.

Comment 3: A series of papers in the past promoted an approach to protect a percentage of species using a species-sensitivity index (see below for some of these). The approach by Solomon et al. had an associated probability of effect but relied heavily on the acute LC50. It was never widely adopted but did make it into an EPA document on ecological risk assessment. The question is: why is the PRiME approach better than the one proposed by Solomon et al. or the others listed above?

Response: Solomon and others make the critical (untested) assumption that laboratory data are adequate to represent field impacts. We simply test that assumption with the pond and mesocosms data.

Comment 4: I think that the incorporation of mesocosm data may be helpful but one could argue that the mesocosm data is a mess and was very hard to interpret. That is why it was abandoned by EPA. Because the fish index relies on reproduction data I think it has an advantage over acute mortality data.

Response: We have been able to extract workable models from these data. Others (e.g. Brock Maltby etc...) have similarly used mesocosms data successfully. The advantage of incorporating the fish index is that it considers pesticide persistence and chronic effects. As such, the acute crustacean and reproductive fish indices complement each other.

Comment 5: Given the number of potential end-users and the diversity of applications for the PRiME model, I would caution against the unintended overestimation of risk, especially in the case of pesticides with limited toxicological data sets. To this end, would it be possible to weight or rank the certainty of the risk estimate based on either the quantity or the variability of the toxicological data?

Response: That would add a great deal of complexity to the output and make the output very difficult to interpret by the user. The methods used to derive the species sensitivity distributions have been chosen to produce the least bias possible relative to data availability.

Comment 6: I am concerned that the reliability of the original data set may become lost among all the model computations. Also, a measure of confidence applied to the “ranks” of potentially used pesticides would inform the relative risk ranking and should improve the subsequent selection process by the applicator.

Response: In principle, I agree. However, incorporating all known sources of variability in the model (from the conduct of every toxicity test in the laboratory all the way to curve fitting in the models) is technically very difficult or impossible. This is why it has

never been done by any jurisdiction anywhere. Also, assuming we could do it, it would make interpretation very complex.

Comment 7: The SSD methodology included combining marine/estuarine and freshwater species on the basis that the sensitivities of the two groups were statistically indistinguishable. However, marine/estuarine species are usually more sensitive than freshwater organisms, and therefore an SSD comprised of both taxa subsets will be skewed versus an SSD containing only one taxa set. In the case of freshwater organisms, the calculated HC₅ will be pushed lower, while for marine/estuarine organisms, the HC₅ value derived from the entire data set will be higher than that which would be derived using only the marine/estuarine toxicity values. Given the disparate physio-chemical properties of the two environments (impacting fate and degradation) and the different biological and physiological attributes of the resident organisms, separating the toxicological data into two different SSDs would be more appropriate. Not only would this improve the power of the model, but it would also produce separate, unique risk rankings for both potentially impacted freshwater and saltwater systems.

Response: It was shown (Maltby) that any freshwater/marine skew is as a result of a larger proportion of sensitive taxonomic groups in the marine environment. However, those taxonomic groups are also present in freshwater systems, just not tested as commonly. This is why we opted to include them in an overall SSD. This is an interesting idea we might be able to incorporate in a future version of the software; however, it is unlikely there will be insufficient marine data to develop a separate SSD.

Comment 8: Developers of the SSD methodology expressed concern that life stage at time of exposure may have introduced additional variability into the model. However, it is unlikely that significant differences in sensitivity would be observed in this case, especially in the case of crustacean and insect toxicity tests, due to the fact that handling test organisms would require that animals be of “workable” size and robustness. A more plausible source of variability would be the origin of the test animals. Field-collected organisms can exhibit increased tolerance for pesticides when compared to laboratory-collected organisms of the same species. While *Daphnia* sp. are rarely collected from field sites, insect and fish toxicity tests have been more frequently conducted with organisms not obtained from laboratory stock. A better understanding of test organisms origins may help explain some of the variability.

Response: Agreed. However, source of test animals is seldom available in the test data. Trait-based strategies for choosing test species have also become popular (Van Den Brink, Baird etc...). However, this science is not yet advanced enough to be incorporated in this version of PRiME.

Comment 9: Given the number of studies incorporated into the model, it is likely that many of the relied-upon documents utilized only nominal concentrations (e.g. did not chemically verify reported concentrations). This probably is a significant source of model variability and one that is unlikely to be easily resolved. However, given that purity of the test compounds is frequently reported, it may be useful to weight the resulting LC/EC₅₀ by the reported purity or by an

estimated purity, where absent (e.g. model $LC_{50} = \text{reported } LC_{50} \times 0.9$, where the purity of the test substance is 90%).

Response: Unfortunately, this information is also often missing. We did make an attempt to exclude data from compounds of low purity (see Whiteside et al. 2008).

Comment 10: For algae, the data set likely suffers from the limited number of algae species utilized in toxicity tests; as a consequence, the mesocosm effects observed are likely to be significantly different than those predicted by single-species toxicity studies. This may be one of the reasons why the current algae model explains so little of the inherent variability ($r^2 = 0.29$), and the EPA is currently exploring new data methodologies and research objective to remedy the limited algal and macrophyte species used in toxicity studies (http://www.epa.gov/oppefed1/cwa_fifra_effects_methodology/presentations/community_aquatic_plant.html#title).

Response: Agreed. This is clearly a weak area. There is also the possibility of introducing a terrestrial plant index if desired. This will be fodder for future versions of PRiME.

Comment 11: The modes-of-action of several important herbicides may be slow acting, e.g. through the inhibition of gibberellins or carotenoid synthesis. In the case of these pesticides, a long time period may pass between exposure and exhibition of toxic effects. This is in contrast to insecticides, which are often designed to be fast-acting in efforts to immediately prevent the proliferation of pest arthropods. [The notable exceptions here are juvenile hormone analogs, the effects of which may only emerge long after exposure. However, effects of JH-type pesticides do appear to occur in closer proximity to exposure than do effects of GSIs or CSIs]. Again, some of the high unexplained variability within the algae model may result from the fact that short-term or acute toxicity studies may not appropriately predict the field effects of slow-acting herbicides.

Response: Agreed.