Technical Progress Report of the Implementation Plan

for Probabilistic Ecological Assessments:

Aquatic Systems

Meeting Scheduled for April 6-7, 2000

March 15, 2000

Technical Progress Report of the Implementation Plan for Probabilistic Ecological Assessments: Aquatic Systems

Table of Contents

Technical Progress Report of the Implementation Plan for Probabilistic Ecological Assessments: Aquatic Systems

I. Introduction

The Environmental Fate and Effects Division (EFED) is proposing a basic outline of an approach to implementing probabilistic risk assessments for aquatic ecosystems. The implementation plan is based, as a starting point, on the aquatic Ecological Committee on FIFRA Risk Assessment Methods (ECOFRAM) Aquatic Draft Report (1999). The ECOFRAM workgroups are to be commended for their considerable effort and progress towards providing a framework for ecological probabilistic assessment. During the period since the June 1999 Peer Input Workshop, the implementation team has been reviewing the recommendations of the aquatic ECOFRAM workgroups and the comments of the peer reviewers from the Peer Input Workshop (1999), evaluating the Terrestrial ECOFRAM draft document for a comparison of approaches, conducting extensive outreach within and outside of EFED, and devising a preliminary plan for implementation of probabilistic assessments. We are seeking the input of the Scientific Advisory Panel at this stage of development of the implementation plan in order to obtain the opinion and advice of the SAP on the most feasible and useful approaches and direction to take in implementing probabilistic ecological assessment in the Office of Pesticide Programs (OPP)

The proposed implementation plan outlined here is intended to be flexible and focused on moving the Agency towards an improved method for conducting pesticide risk assessments based on considerations of the magnitude and probability of ecological effect. The implementation plan presents an approach to ecological risk assessment that proceeds through four levels of analysis or refinement, from simple deterministic quotients through a general probabilistic analysis, to an issue-specific probabilistic assessment in the highest levels. The plan is designed to permit immediate completion of the simple levels of deterministic and rudimentary probabilistic assessment, using tools and models currently available to Agency scientists. In the near future, it is anticipated that improved models and databases will be available to provide more comprehensive risk assessments that more effectively consider the variety of ecological systems under consideration and the improved state of science.

There are many implicit assumptions involving the validity of extrapolations that are not explicitly addressed in current risk assessment methodology. It is important to be aware that estimates of potential risk are not accounting for a number of factors which may profoundly affect the actual risk outcome. These factors include: (1) intraspecies sensitivity differences (e.g., between life-stages), (2) interspecies sensitivity differences, (3) extrapolations from laboratory effects to field effects, (4) extrapolations from data based on a single active ingredient and not the formulated product applied to the environment, (5) exposure estimates based on a single crop use instead of uses on multiple crops within a watershed, (6) effects from concurrent or sequential exposures to multiple chemicals, (7) indirect effects of exposure, (8) sublethal acute effects on fitness (behavior, immunocompetence, etc.), and (9) the effect of additional stressors (e.g., habitat loss and natural stressors).

In the approach outlined here, we have tried to explicitly account for interspecies sensitivity differences through the use of extrapolation factors in Level 1. The interspecies extrapolation factors may partially address additional uncertainties such as inter-laboratory variation, and some degree of intraspecies age variation in sensitivity, depending on the scale of data used in the extrapolation factor generation. At the higher levels of refinement, there are approaches outlined that begin to address population-level effects, behavioral effects, other sublethal effects related to specific modes of action, as well as actual field effects. We currently do not have tools or approaches to sufficiently address a number of indirect effects, and the effects of multiple chemical exposures and non-chemical stressors. We hope that the approach outlined here will serve as a stepping-stone to future approaches that will more comprehensively consider effects of pesticides on the environment.

The reviewer may note a change in terminology in this document for the degree of risk assessment evaluation, from the previously used term of “Tier” to the new term of “Level of Refinement.” This change in terminology was recommended by the Terrestrial ECOFRAM group and is intended to reflect a degree of flexibility between the levels of an assessment, as opposed to a rigid, step-wise evaluation. This limited flexibility is intended to permit regulators and registrants to maximize efforts in areas of the assessment that appear to be most significant and illustrative of the actual risk. This flexibility must, however, be balanced with the need to be consistent, predictable and transparent in the regulatory process. EFED intends to develop guidance on which tools would be most useful to select in the flexible third and fourth levels of refinement, given different types exposure/effects profiles.

The development of criteria, or triggers, for progression to higher levels of refinement requires both a strong foundation in science as well as close collaboration between risk assessors and risk managers. The EFED Implementation Team is interacting with risk managers to discuss the criteria for moving between the Levels of Refinement.

Following this introduction, the second part of this document details the proposed Levels of Refinement, their associated data requirements, tools, and potential issues for consideration at each Level. This discussion includes a comparison with current methodology and data requirements, as well as ECOFRAM recommendations. The third part of the document provides a synopsis, in table form, of the levels of refinement in terms of data requirements and approaches taken. The fourth part provides an example of an assessment conducted as described in this approach, through the second Level of Refinement. The fifth part of the document outlines research needs related to the improvement of information for probabilistic assessments. The sixth section outlines recommended changes to the guideline requirements to improve data quality. A summary of the Aquatic ECOFRAM Draft Report is given in Attachment 1. This attachment includes a brief comparison of some of the issues addressed in the Aquatic and Terrestrial ECOFRAM Draft Reports. A summary of the Peer Input Panel members’ comments may be found in Attachment 2.

II. Proposed Levels of Refinement

(Please see tables outlining each level of refinement for an overview.)

First Level of Refinement

Description

Level 1 assessments use simple, relatively conservative assumptions yielding deterministic risk quotients which are to be compared to established Levels of Concern. Level 1 is aimed at addressing the risk to aquatic species in vulnerable headwater environments resulting from a high exposure, edge-of-field scenario. This is a screening level designed to be protective, not predictive, and to identify compounds thought to present minimal risk to ecological systems.

Objectives

The objectives of this level of refinement are to identify products which are likely to present minimal environmental/ecological concern.

Exposure assessment

The Level 1 **exposure** approach provides an estimated 90 % upper bound point estimate of exposure for a vulnerable headwater environment based on conservative assumptions of environmental fate properties, processes, and environmental conditions. The GENEEC (GENeric Estimated Environmental Concentration) model is used by EFED as the current Tier 1 aquatic exposure estimation tool. Level 1 is intended to provide a coarse screen to sort out those pesticides highly likely to pose minimal aquatic risk, based on the data available for the analysis. The screen only considers a few basic chemical properties, the application rate, number, and methods. If application rates or methods vary substantially by crop, several crops may be considered.

GENEEC provides estimated environmental concentrations (EEC's) for a pesticide in an edge-of-field farm pond. GENEEC was specifically designed to mimic PRZM (Pesticide Root Zone Model) and EXAMS (Exposure Analysis Modeling System), the more complex computer models used in current Tier 2 assessments. The current selected site severity is based on approximately the 90th centile Tier 2 site for cotton, ranked by erosion potential. GENEEC is a single runoff event model which can account for spray drift from multiple applications. It is based on the assumption of 10 percent chemical runoff from the treated field. When coupled with the assumption of a 10-to-1 land-to-water ratio, GENEEC calculates a single runoff loading event equivalent to a single application of chemical directly to the 1 hectare farm pond. The 1 hectare pond is two meters deep and has no outlet Outputs provide a maximum peak, 96 hour average, 21-day and 56-day time-weighted average of water column concentrations. This model is generic in that it does not consider differences in climate, soils, topography, and crop.

The limitations of GENEEC have been discussed in the ECOFRAM report. The major ones include (1) degradates are not considered, (2) the foliar component of exposure is not considered, (3) there is no volatilization consideration, (4) only one runoff event is considered, (5) for a persistent chemical, GENEEC may not capture the effect of accumulation (i.e., the continuous modeling simulation approach in Level 2 may predict higher EEC's than Level 1 values), and (6) the model addresses only one scenario. Its primary virtues are its ease and speed of use, and the minimal data needed for its execution, which are important characteristics for a screening level method.

For reregistration assessments, monitoring data that is readily available (e.g. USGS National Water Quality Assessment Program (NAWQA) data) should also be considered at this point, as a form of order-of-magnitude ground-truthing for the exposure modeling output.

<u>Proposed Refinements of Level 1 Exposure Modeling</u>

Proposed refinements to GENEEC include an update of the current row crop and turf scenarios. The updated GENEEC will also include cranberry, forestry, rice, and right-of-way scenarios.

In the future, the Agency may need to develop an approach for estimating estuarine exposures. Currently, effects data for estuarine species are compared to GENEEC or PRZM/EXAMS modeling outputs for risk estimations.

As noted above, Level 1 is intended to function as a screen to identify products which are likely to present minimal environmental/ecological concern. GENEEC is a conservative model, however, there are some situations where the model may underestimate exposures. As noted above, GENEEC simulates a single runoff loading event which is equivalent to a single application of chemical directly to a 1 hectare farm pond. The model pond is two meters deep and has no outlet. In general this model appears to sufficiently estimate a high-end exposure scenario, in most situations. There is some concern, however, that the field-size-to-pond-volume ratio may not represent ratios found in small, sensitive aquatic areas. For example, areas such as small or ephemeral water bodies that amphibians may use for breeding could contain higher peak concentrations than that estimated based on GENEEC output. Small streams receiving multiple inputs from adjacent fields, smaller volume ponds receiving an equivalent level of runoff and/or drift as is estimated by the model, or ponds of equivalent size to the model pond that receive runoff from fields larger than ten hectares may have higher concentrations than GENEEC model estimates. Additionally, as noted above, for a persistent chemical, GENEEC may not capture the effect of accumulation over time (i.e., the continuous modeling simulation approach in Level 2 may predict higher EEC's than Level 1 values). The effect of degradates is also not considered. Furthermore, empirical data indicates that, occasionally, environmental levels actually exceed GENEEC values. In general, it is considered that GENEEC is a good, conservative estimate generator, since it is based essentially on direct application to water, and generally yields higher estimated concentrations than that found in monitoring studies. However, because of the concerns listed above, the sufficiency of GENEEC as an interim Level 1 screen to enable reasonably confident conclusions of minimal ecological concern, in a variety of ecological environments, is under continuing evaluation. For example, Level 1 approaches to estimate concentrations in water bodies with larger field-size-to-pond-volume ratios than that used in GENEEC may need to be considered.

Effects assessment

The new, initial effects assessment is a revised version of OPP's current Tier 1 assessment. Current data requirements for aquatic species are considered insufficient to conduct a probabilistic assessment. Therefore, the use of an extrapolation factor to address interspecific variation in sensitivity is included in the new approach. Additional effects data requirements may also be a necessary element of the new Level 1 effects assessment, in cases where exposure estimates are within an order of magnitude of effects estimates. These potential additional effects data requirements are discussed below.

Acute effects test species:

Current acute testing requirements include: two freshwater fish species, one freshwater invertebrate, and 1-7 aquatic plant species. If there is potential for estuarine exposure, one acute test each for an estuarine fish, arthropod (mysid shrimp) and mollusk (oyster) are currently required.

EFED recommends that acute testing requirements include:

1. Two freshwater fish tests (preferably bluegill and rainbow trout)
2. One aquatic invertebrate test (preferably *Daphnia magna*)
3. Algae and aquatic plant toxicity tests

If estuarine/marine exposures are possible, additional acute testing should include:

4. One estuarine fish test (preferably Atlantic silverside)
5. One estuarine mysid test (preferably *Americamysis bahia*)
6. One estuarine bivalve test (preferably Eastern oyster)

These requirements are all the same as currently required, but additional potential requirements are noted below.

Chronic effects test species:

For chronic considerations, a fish early life-stage test and an aquatic invertebrate life-cycle test for freshwater and estuarine/marine species, as appropriate, are, though conditional, also generally required for most outdoor use pesticides. Historically, the vast majority of tests requested to address chronic aquatic concerns were the aquatic invertebrate and fish early-life stage test, based on data in EFED's toxicity database (Montague, 1992). Unfortunately, the fish early life-stage test does not address many of reproductive endpoints that could be affected by chemical exposure, such as egg production. Hence, the effects of pesticides on fish reproduction are unknown for many pesticides. EFED would like to improve data collection on the chronic effects of pesticides on aquatic species.

EFED recommends the following chronic studies be conducted at Level One:

1. One or more fish early life-stage tests (preferably rainbow trout or fathead minnow)
2. One screening level fish reproduction test (preferably fathead minnow). To fulfill the objective of Level 1 to identify materials which present minimal risk, this additional test is

needed. The USEPA ORD Duluth laboratory has developed an abbreviated (21 day) fish reproductive toxicity test, which may be a good candidate Level 1 fish chronic toxicity test, pending complete validation against the current fish full life-cycle test.

3. One aquatic invertebrate life-cycle test (preferably *Daphnia magna*)

If estuarine/marine exposures are possible, additional chronic testing should include:

4. One estuarine fish test (preferably sheepshead minnow)
5. One aquatic invertebrate life-cycle test (preferably *Americamysis bahia*)

For the chronic data, a regression based EC_{10} was recommended by ECOFRAM as the regulatory endpoint, in place of an NOAEC. ECOFRAM recommended that a NOAEC should only be used when a study is technically acceptable, but the data do not support regression analysis. It was noted by some of the Peer Review Panel members of ECOFRAM that the use of a LOAEC or NOAEC as the regulatory endpoint has the unfortunate effect of rewarding poor experimental design and high control variability with increased NOAECs and LOAECs. Difficulties with the use of a regression-based approach include the fact that for many of the submitted fish full life-cycle tests, there may be only one effect level. The data frequently do not support the use of regression data. EFED is considering the use of a regression-based approach for the chronic tests, and would like to obtain the SAP's comments regarding this issue.

An overarching consideration for the effects data is the power of the tests to detect differences between doses in effect, given the small number of replicates required under current testing protocols. Revision of current testing protocols to provide more power within a given test through additional, reasonable replication, as well as the replication of entire tests would improve the overall risk assessment.

<u>Potential new data requirements</u>

When a small margin of safety exists (less than one order of magnitude) between expected concentrations of a chemical and its toxicity, then a more extensive data base is justified (Mayer and Ellersieck, 1988). The Agency is proposing that these new data would include additional invertebrate species in both acute and chronic tests, sediment toxicity testing (if the chemical partitions to sediment), rooted plant testing (freshwater and estuarine if appropriate), and amphibian testing (e.g., FETAX).

The minimal increase in invertebrate species tests reflects an attempt to provide some consideration of the range of potential responses. In assessing impact to aquatic invertebrates, risk assessments currently focus on *Daphnia magna*, a parthenogenetic, short-lived crustacean. However, it is a concern that one species by itself only represents a limited view for assessing the risk of toxicants to other invertebrates. A major issue is one of temporal variation including the following: 1) different and often variable life spans; 2) unpredictable lengths of life stages and different metamorphic stages; 3) indeterminate growth in some species; 4) significant differences in adipose stores. These variable should be considered for multispecies assemblage testing especially since impact to trophic levels will result in exposures of differing magnitudes affecting different species.

Vall *et al.*, (1997) noted that differences in invertebrate species sensitivity to toxicants can be substantial. Evaluating variation in sensitivity relative to toxicant mode of action, they divided chemicals into the following classes: non-polar narcotics, polar narcotics, reactive compounds and specifically acting compounds. Interspecies variations were analyzed for 35 chemicals, with data ranging from 12 to 62 species per compound. They found that non-polar and polar narcotic compounds demonstrated the smallest variation in sensitivity with levels of toxicity that are predictable on the basis of hydrophobicity. However, reactive and specifically-acting chemicals were much more toxic than predicted, with a large variation in species sensitivity. Therefore, the larger the interspecies variation the more asymmetric the toxicity distribution. A reliance on one species of invertebrate for defining uncertainties at the first Level of Refinement could result in an oversight regarding reactive and specifically-acting compounds and the risk they may present to ecosystems. Therefore, in order to better ascertain a range of toxicity distributions for representative species, we suggest that a minimal increase in the number of freshwater and marine/estuarine invertebrates tested at Level One may be warranted.

Examples of additional species for invertebrate acute toxicity tests would include stoneflies and amphipods. Mayer and Ellersieck (1988) recommended the inclusion of these species when the toxicity was within an order of magnitude of the expected environmental concentration, to better approximate toxicity for more sensitive species, based on their extensive review of acute toxicity data. Examples of additional species for chronic toxicity testing that could be used to begin to address concerns of chronic effects on sexually reproducing invertebrate species, would include copepods and chironomids. EFED seeks discussion and recommendation by the SAP on whether a minimal data requirement increase for aquatic invertebrates should be addressed at the first or second Level of Refinement.

EFED has also discussed the addition of a single acute amphibian test to reflect the need to begin to directly consider effects on amphibian species. This seems especially relevant and timely given the reports of international amphibian population declines. The amphibian assay under current consideration (FETAX: Frog Embryo Teratogenesis Assay- *Xenopus*) would provide information on developmental effects on aquatic vertebrates from acute pesticide exposures. An extended exposure and observation period FETAX test could also be considered as an interim measurement of chronic toxicity to amphibians.

The team seeks the SAP's opinion on the importance and relevancy of this additional data on invertebrates and amphibians, and, if this testing is deemed advisable, whether this data should be sought at the first or second Level of Refinement.

If a pesticide's properties (e.g. solubility, K_{OC}, K_D, half-life) indicate that it may partition to sediment and toxicity to water-column organisms has been observed, then the hazard to sediment-dwelling organisms may need to be evaluated. The Level 2 exposure model output (PRZM/EXAMS) would determine if toxicity testing on benthic species is required. Benthic species toxicity testing is further discussed under Level 2.

Plant toxicity test requirements were recently reviewed in a workshop conducted by the International Life Sciences Institute (ILSI), in process separate from ECOFRAM. The workshop participants, an international panel of plant experts, included members from EFED. A meeting with the SAP on the harmonization of aquatic plant toxicity tests with Canada, under the NAFTA agreement, is scheduled for later this year. In the interim, most of the basic probabilistic tools and methodologies developed for nontarget animals, as described in this implementation plan, may be applied to plants using currently required species.

Risk assessment

The risk assessment will be based on deterministic risk quotients. The acute risk quotient will be the ratio of model-estimated peak environmental concentration to the quotient of the LC_{50} or EC_{50} divided by an extrapolation factor. The chronic risk quotient will be the ratio of model-estimated *peak* environmental concentration to the quotient of the EC_{10} (or NOAEC) divided by an extrapolation factor.

The use of a peak EEC for comparison with chronic effects endpoints was supported by ECOFRAM. ECOFRAM noted that quantal effects such as mortality and hatching may reflect effects of short-term exposure at critical stages in the life-cycle, while continuous variables such as growth may generally reflect the effects of cumulative exposure. Since Level 1 is intended to be protective, they therefore recommended that the chronic risk quotient should be based on the peak EECs. Experimental data supports this approach. Nimmo *et al.* (1980) noted that a 24 hour exposure to a compound at one-half of the LC_{50} resulted in sterilization in mysids. Other authors have noted chronic effects resulting from short-term exposures (Barry *et al.*, 1997).

In the new Level One approach, extrapolation factors will be applied to the LC_{50} and EC_{10} values in order to begin to include consideration of interspecies sensitivity differences. The extrapolation factors will provide an estimate of potential hazard to the more sensitive species (e.g. 95% most sensitive) within each group. The use of extrapolation factors is intended to address uncertainty in species sensitivity between tested and untested species, without requiring a large number of toxicity tests on additional species at the initial levels of a risk assessment. The factors will be determined through a review of historical acute and chronic toxicity data, and will be determined for sensitivity differences within groups such as: freshwater fish, estuarine/marine fish, freshwater invertebrates, estuarine/marine invertebrates, amphibians, and aquatic plants. Also evaluated will be differences between pelagic and benthic species in each group. Additional issues that will be considered when the species sensitivity data are reevaluated include consideration of the need to separate species groups (e.g. freshwater vs. marine fish), or whether a common extrapolation factor can be derived due to lack of a large degree of variability between the groups. The extrapolation factor will be divided by a toxicity measurement for each group, thus the extrapolated LC_{50}s/EC_{10}s will be less than measured LC_{50}s/EC_{10}s. The extrapolated LC_{50}s/EC_{10}s will be the denominator in the risk quotients, and should help provide an estimate of potential hazard to the more sensitive species within each group (see part II of this document for an example). Critical in this consideration is a wise use of the extrapolation factor. Risk considerations must address the issue of whether keystone species in an ecosystem will fall outside of the protected 95% of species.

Weak points in the proposed Level 1 process include the following points: Evaluating the effect of degradates has not been adequately developed, the effect of persistent chemicals cannot be fully addressed with the Level 1 exposure model, GENEEC has never been formally validated, and a potential exists for underestimating of exposure in small, sensitive water bodies.

Differences from the current approach that agree with ECOFRAM draft report recommendations include:

1. Reporting and using acute data at 24 hour intervals, collecting sufficient data to estimate the slope of the dose-response relationship, and conducting time-to-event analysis, were recommended. The time-to-event analyses are relevant for a fuller evaluation of potential short duration exposures in streams. ECOFRAM recommended time-to-event testing since it "generally provides a better estimate of acute LC_{50} values than conventional probit analysis."
2. ECOFRAM recommended adding invertebrate chronic toxicity testing and fish early life-stage tests at Tier 1. EFED concurs.
3. ECOFRAM recommended using a regression based EC_{10} for chronic risk assessment instead of an Analysis of Variance-generated NOAEC. The use of a regression- based approach for toxicity is an important consideration. EFED seeks the view of the SAP on this matter.

Recommended differences from the ECOFRAM draft report include:

1. A screening level fish reproduction test is recommended at Level 1. The ECOFRAM draft report stated that objectives of Tier 1 were to provide an assessment of whether acute or chronic concentrations may be of concern. However, the fish full life-cycle test was not included by ECOFRAM until Tier 3, so this objective was not fully achievable for aquatic vertebrates using the ECOFRAM-outlined process. Current information from ECOFRAM members indicates that this point may be being reconsidered by that team, and that additional information on chronic effects on fish may be recommended at a lower tier in their final report.
2. Inclusion of additional invertebrate species for selected tests is under consideration. This concern is based on the large differences in invertebrate life histories, which may affect the potential to underestimate significant effects when basing an assessment solely on one invertebrate species. Regarding chronic testing of invertebrates, *Daphnia* primarily reproduce via parthogenesis, except under conditions of stress, when sexual reproduction may occur. Chronic toxicity testing with a primarily parthenogenetic species seems counterintuitive because potential reproductive effects on male invertebrates may not be detected. Secondarily, the study of effects on an invertebrate with a prolific, rapid reproductive life-history may not be the best means to characterize reproductive effects on dissimilar invertebrates, especially regarding reactive and specifically-acting chemicals, as

noted above. Comments and recommendations by the SAP on this issue would be welcomed.

3. Amphibian testing via FETAX (Frog Embryo Teratogenesis Assay-*Xenopus*) is under consideration. Comments and recommendations by the SAP are sought on this issue.

Second Level of Refinement

Description

For chemicals not demonstrated to pose minimal risk, and when risk managers and risk assessors agree that not enough information is available to make mitigation decisions, the risk assessment will move to the second level of refinement. Level 2 will yield a probabilistic assessment aimed at addressing the risk to aquatic species in vulnerable headwater environments for the crop of interest through the use of PRZM/EXAMS modeling.

Objectives

The objectives of Level 2 are to characterize temporal variations in risk to headwater/static ecosystems, to better qualify risk when more data are used, to provide basic probabilistic expressions of risk to potential problems indicated in Level 1, to consider which regional/cropping patterns or effects may merit additional refinement at Level 3, and/or to permit preliminary evaluation of mitigation options.

Exposure assessment

The exposure assessment at Level 2 is intended to provide initial probabilistic estimates of exposure for vulnerable headwater environments across a wide range of geographical conditions as appropriate for product use. Currently, Tier 2 EEC assessments use a single site which represents a high-end exposure scenario from pesticide use on a particular crop or non-crop use site for multiple years (typically 36 years). In some cases, multiple exposure model runs are conducted to address different scenarios. The scenario or scenarios chosen are best professional judgement sites expected to produce runoff greater than would be expected at 90% of the sites for a given crop/use.

Currently, PRZM and EXAMS are linked for a refined estimation of pesticide concentration in surface water environment. PRZM 3.12 simulates the runoff and erosion from an agricultural field and EXAMS 2.97.5 simulates the fate in a receiving water body. The water body simulated is a static pond, adjacent to the crop of interest. The model yields an output of annual maxima distributions of peak, 96 hour, 21 day, 60 day, 90 day and yearly intervals.

Proposed Refinements of Level 2 Exposure Modeling

Currently, a PRZM/EXAMS shell is under development. The shell is designed to be a user-friendly input/output program around the EPA PRZM3 and EXAMS linked models. The shell will facilitate Level 2 exposure modeling.

The current PRZM/EXAMS model predicts the EEC's at 1-in-10 year probability based on the temporal consideration of continuous 36 years simulation from 1948 to 1983. To refine the temporal variability estimates, the simulation period will, in the future, be increased to 1998, allowing a continuous simulation of more than 50 years. For the EXAMS standard farm pond scenario, the current off-target drift loadings to the receiving farm pond are assumed to be a constant. Future EXAMS refinements will include the drift predictions from AgDRIFT considering the impact of buffer restrictions and the regional effects of wind speed and direction.

A second approach to refining Level 2 exposure assessments would be to try to address the uncertainties associated with model input parameters by performing Monte Carlo simulations with PRZM. In a recent SETAC meeting, Carbone, *et al.* (1999) presented an approach of linking the Monte Carlo sampling software Crystal Ball (Decisioneering, Inc) with PRZM 3.12, in order to address the uncertainties associated PRZM 3.12 input parameters. An initial sensitivity analysis was performed to identify the most sensitive parameters by defining the distributions of these input parameters, and employing a Monte Carlo simulation with the Crystal Ball Pro software. This approach demonstrates the feasibility of linking Crystal Ball and PRZM 3.12 as an analysis tool. EFED will consider this approach in the future, in addition to continuing to investigate the current built-in Monte Carlo shell in PRZM 3.12 to determine the best approach to address the uncertainties associated with the input parameters.

The Agency seeks the SAP's recommendation regarding the usefulness of including Monte Carlo simulations addressing variability in input parameters in Level Two exposure refinements. An issue of concern is resource utilization at early levels of refinement. That is, we are seeking comments on whether it is considered that this level of sophistication and complexity in the modeling scenario would be most useful at this or a higher level of refinement.

To further address the uncertainties associated with the environmental fate data, EFED proposes to modify the FIFRA part 158 guidance, as outlined in part six of this document.

Available monitoring data (e.g. USGS National Water Quality Assessment Program (NAWQA) data) should be considered at this point, relative to exposure modeling output, to ascertain if monitoring information indicates a problem not predicted by modeling information.
The need for additional environmental fate data acquisition and analysis (lab or field), or additional fate and transport studies, may then be indicated.

Differences in Level 2 exposure assessment from ECOFRAM recommendations:

1. As an interim process ECOFRAM recommended the use of MUSCRAT (Multiple Scenario Risk Assessment Tool) for a wider range of scenarios than currently considered at Tier 2. (A description of MUSCRAT is given in the Level 3 exposure discussion). ECOFRAM also recommended addressing lotic as well as lentic water body considerations. EFED recommends a standard Level 2 use of PRZM/EXAMS modeling, essentially as is currently applied with minor changes as noted above. The primary reason is EPA's desire to move immediately toward implementation of probabilistic assessment;

the use of PRZM/EXAMS would permit a more rapid and feasible initiation of rudimentary probabilistic assessment. The current MUSCRAT version is under beta testing and has not been officially approved by EPA/OPP/EFED via the EPA/ORD laboratories. Thus, the development of MUSCRAT in accordance with ECOFRAM's recommendations could substantially delay the implementation process.

2. The use of Monte Carlo analysis to address variability in specific input parameters is under consideration. ECOFRAM did not address this.
3. ECOFRAM recommended that a newly-developed tool, RADAR (Risk Assessment Tool to Evaluate Duration and Recovery), be finalized for use at Tier 2. RADAR is essentially a postprocessor for EXAMS, which yields information on exposure event magnitude and duration, which can be compared to a concentration or response threshold for use in population recovery estimates following pulsed exposure. EFED recommends moving this tool to Level 3, in order to maintain Level 2 as an initial probabilistic assessment without the need for elaborate refinement and concomitant resource burden.

Effects assessment

The Level 2 effects assessment includes current toxicity tests, as listed in Level 1 and those conditionally added, with the inclusion of the slopes of the dose-response curves. No additional toxicity tests are included at this level, if the additional invertebrate, amphibian, and benthic tests are conducted at Level 1. It is proposed that the Level 2 analysis may address all taxonomic groups. ECOFRAM proposed that Tier 2 assessments be focused only on taxa or scenarios where LOC's are exceeded based on directly tested species in Tier 1.

The full dose-response curves will be examined for each species and group/taxa indicated to be of concern. The group dose-response curve will be an extrapolated curve, created in order to provide an estimate of the full dose-response relationship of sensitive species. EFED is considering the options for creation of and validity of extrapolated dose-response curves, and seeks the SAP's advice on this issue. An issue of concern is which slope of the dose-response curve to extrapolate. The mean or geometric mean slope of the tested species could be used as the standard slope, to which an extrapolation factor could be added, or other approaches could be taken. EFED would welcome other recommendations on an appropriate dose-response function to protect sensitive species within each group. The output would contain both the extrapolated dose-response curves and the curves for the species actually tested, all in one graph where feasible, for ease of comparison.

Additional conditional tests discussed above under Level 1:

1. Additional acute and chronic invertebrate toxicity tests
2. Amphibian acute and chronic tests

If chemical has properties indicating partitioning to sediment, sediment-associated exposure is predicted based on model estimates, and toxicity to water-column species has been observed, additional tests should include acute and chronic benthic tests:

1. Benthic invertebrate species tests

2. Benthic fish tests
3. Potentially, rooted plant testing.

EFED would like to obtain the SAP's recommendation on sediment toxicity testing, particularly regarding appropriate level of refinement for considering and conducting these tests.

Risk assessment

The Level 2 evaluation process yields a rudimentary probabilistic assessment. Joint probability curves were suggested by ECOFRAM. This approach combines the exposure estimates describing the probability of the environmental concentration exceeding a given level with the estimates of biological effects observed at various concentrations. The joint probability curves produced from these estimates would describe the probability of the environmental concentration exceeding a concentration which would result an estimated magnitude of detrimental biological effect, such as mortality or reproductive impairment, for a given scenario(s). The dose-response curve for an individual species can be considered to represent the percent probability of an individual being affected, or the percent of the population that will be affected. (An example of a Level 2 risk assessment is given in part four of this document.) The assessment would then proceed to outline the consequences of the potential effects.

Level 2 shortcomings include the lack of chemical-specific species sensitivity information since generic extrapolation factors are used. In addition, the exposure estimates will be representative of high end-exposure scenarios only, and so the probabilistic aspect of the assessment will be limited to an estimate of the magnitude and probability of effects only in expected high-end exposure scenarios. However, the Implementation Team believes that this approach will be useful in that additional resources will not be spent on chemicals that do not indicate risk at a high-end site, and so will not be likely to pose a risk in more common areas of use. Uncertainties will need to be identified and described in the joint probability approach. These uncertainties were not fully discussed in the ECOFRAM document.

Differences from ECOFRAM draft document approach.

1. EFED proposes the Level 2 analysis may include all groups. EFED believes that excluding potentially sensitive groups from consideration, prior to examining the full dose-response relationship, may be premature. The data will be analyzed through the use of dose-response curves containing both the extrapolated dose-response curves and the curves for the species actually tested, all in one graph. ECOFRAM proposed that Tier 2 assessments be focused only on taxa or scenarios where LOC's are exceeded in Tier 1, based only on species that were directly tested.
2. ECOFRAM also recommended extrapolating the toxicity data to hypothetical population effects estimates. For example, the effects of a given exposure for a species with high reproductive rates and a short lifespan (e.g. *Daphnia*) could be compared to the effects on a relatively low reproductive rate, long-lived species (e.g. trout), using time-to-recovery estimates versus concentration estimates. EFED recommends that initial, generic estimates of population-level effects be undertaken at Level 3, as described below. The team believes that Level 2 should be focused on initial probabilistic estimates of mortality or

chronic effects, and that using additional tools and models at Level 2 would unnecessarily slow down and complicate the evaluation.

3. As an interim process ECOFRAM recommended the use of MUSCRAT (Multiple Scenario Risk Assessment Tool) for a wider range of scenarios than currently considered at Tier 2. ECOFRAM also recommended addressing lotic as well as lentic body considerations. EFED recommends a standard Level 2 use of PRZM/EXAMS modeling, as is currently applied with small changes as noted above to maintain some consistency with current Tier 2 exposure modeling estimates in order to permit a more rapid and feasible initiation of rudimentary probabilistic assessment, and because the current MUSCRAT version is under beta testing and has not been officially approved by EPA/OPP/EFED via the EPA/ORD laboratories.
4. The use of Monte Carlo analysis to address variability in specific input parameters is under consideration. ECOFRAM did not address this in Level 2.
5. ECOFRAM recommended that a newly-developed tool, RADAR (Risk Assessment Tool to Evaluate Duration and Recovery), be finalized for use at Tier 2. EFED recommends moving this tool to Level 3, in order to maintain Level 2 as an initial probabilistic assessment without the need for elaborate refinement and concomitant resource burden.
6. The potential toxicity of sediment-associated pesticides to benthic species is more fully considered at Level 2 in the proposed implementation plan than in the ECOFRAM Draft Report. ECOFRAM recommended direct comparisons of estimated pore water concentrations to toxicity data for pelagic invertebrate species in Level 2, to calculate potential risk to benthic species. These risk estimates would determine the Level 3 need for benthic species testing. The EFED plan conditionally involves Level 1 or 2 benthic species testing. If a pesticide's properties indicate that it may partition to sediment and *some* toxicity to water-column organisms has been observed, then the Level 2 model output (PRZM/EXAMS) would determine if toxicity testing on benthic species is necessary in Level 2.

Third Level of Refinement

<u>Description</u>

If the output from Level 2 indicates that the chemical still seems to pose an unacceptable risk, or that the uncertainty of the risk estimates are high, and/or we do not have sufficient information to make informed risk management decisions, the assessment would proceed to Level 3. Level 3 involves further probabilistic refinement of the Level 2 risk assessment, using additional data and/or modeling. Level 3 focuses on sites or taxa where the risk is estimated or reported to be high, aspects of the assessment where more information is needed, or on the impacts of potential mitigation.

<u>Objectives</u>

The objectives of Level 3 are to reduce the uncertainty in the Level 2 assessment, and to better understand the nature of the risk and/or the effectiveness of potential mitigation options.

Exposure assessment

Because Level 3 begins to focus on specific problem areas, the Level 3 exposure assessment would be unique to each chemical. Possible new exposure data or analyses were proposed by ECOFRAM based on a "toolbox" approach. Based on best professional judgement and the results of sensitivity analyses regarding the "risk drivers", the assessor would select from the toolbox the tool most appropriate to fulfill the needs of the risk assessment. EFED concurs that this toolbox approach is useful. EFED intends to develop guidance on which tools would be most useful to select in the third and fourth levels of refinement, given different types exposure/effects profiles. This guidance is necessary to permit a consistent, predictable and transparent regulatory process.

Tools suggested for Level 3 include:

1. Refined exposure modeling to reflect spatio-temporal differences in exposure. Within the Level 3 refined modeling approach there are two possible options for adjustment of the current basic modeling approach.

   a. The first approach involves the use of MUSCRAT (Multiple Scenarios Risk Assessment Tool). Instead of focusing on one site, MUSCRAT will be used to simulate multiple sites with one region. This approach will provide a better representation of spatial variability in exposures. MUSCRAT is a Windows-based application program developed to standardize and automate current Tier 3 ecological aquatic exposure assessment which links chemical, crop, soil , and climate data bases. It facilitates the creation of PRZM-3 and EXAMS II input files, provides batch processing for multiple model simulation runs, and performs statistical analyses on predicted exposure concentrations. MUSCRAT divides the forty-eight contiguous states into eleven regions, with up to twenty-five scenarios within each region. The current version is under beta testing and has not been officially approved by EPA/OPP/EFED via the EPA/ORD laboratories. The updated MUSCRAT model will incorporate twenty-five potential crops. The shorter-term refinements to PRZM/EXAMS outlined in Level 2, such as the increase in the weather data to 50 years and AgDRIFT considerations, would be reflected in a refined MUSCRAT product at Level 3. The future development of a regional multi-water body analysis tool, as outlined under the Level 2 exposure assessment section, would be expected to eventually replace MUSCRAT modeling.

   b. The second approach would be to try to address the uncertainties associated with each model input parameter by performing Monte Carlo simulation with PRZM, at Level 3, instead of Level 2, as discussed above.

2. A refinement in the data analyses is also being suggested. In addition to providing the different percentile exposure values based on multiple-year (36 or 50 years) simulations, the new approach will focus on a single event exposure and also consider the possible recovery with the exposure duration. A post processor tool (Risk Assessment Tool to

Evaluate Duration and Recovery - RADAR) will be used to analyze the detailed daily output series predicted by EXAMS to examine the data in terms of events during which the concentration exceeds a threshold (such as an LC_{50} or EC_{10}). For each event, RADAR determines the maximum concentration, the average concentration, the duration of exposure, and the time before the next event occurs. This aspect of exposure characterization is of value in interpreting population recovery potential and design of pulse-dose toxicity studies where appropriate.

3. Additional environmental fate data acquisition and analysis (lab or field). For example, testing with more soils, or examining foliar degradation rates, if sensitivity analyses indicate that these parameters are important contributors to uncertainty.
4. Monitoring data that is available (e.g. USGS National Water Quality Assessment Program (NAWQA) data) should be evaluated.

Ideally, EFED agrees with ECOFRAM's vision of including automated, standardized scenarios for each type of water body and a range of soil and weather conditions which would automatically be generated for the appropriate use pattern. The scenarios that would be ideally included would be selected on a watershed basis, possibly using GIS, and by region. Lentic and lotic systems would be analyzed, with different appropriate species and hydraulic residence times, a range of soil and climates for each use. The modeling tool used would optimally be modular with respect to databases and algorithms for flexibility. In moving towards this aim, EFED has been working with exposure modeling experts in ORD to outline a plan for developing a regional approach to exposure modeling. Under consideration is the development of a series of regional exposure scenarios, each containing a pond, small stream, ephemeral water bodies, fields, and buffer areas. The exposure scenario (e.g. pond size and configuration) would be adjusted regionally based on GIS and other data. For example, for ponds where surface runoff is the main source of water, the contributing drainage area must be large enough to maintain water in the pond during the droughts. However, the drainage area should not be so large that expensive overflow structures are needed to bypass excess runoff during storms. Currently, in the Tier 2 approach, a drainage area to water area is assumed a constant of 10-to-1 for all regions. The receiving farm pond is assumed a rectangular pond (approximately 157 m x 64 m) and a depth of 2 m. One of the refinements proposed is to create a pond database to capture the differences between ponds in different regions. Additional refinements would include simulation of a flow-varying pond instead of the current steady-state, no-outflow pond, to give a better representation of dilution effects on the pesticide loading. Considerations of the relative effects of a pesticide on different water bodies in various use regions could then be more fully evaluated.

Whereas smaller refinements to PRZM/EXAMS, such as expanding the period for the weather files to 50 years and implementing AgDrift predictions, could be implemented in the shorter-term, the development of multi-system regional approach will be a longer-term refinement goal for EFED.

<u>Effects assessment</u>

The Level 3 effects assessment would also be unique to each chemical, with consideration given to species or issues where negative outcomes are expected or additional information is required. The Level 3 effects “toolbox” outlined by ECOFRAM includes several possible new effects data tests or analyses. The tool selected would again be determined by best professional judgement and/or sensitivity analyses. The tool(s) selected would be based on gathering more information regarding the major sources of uncertainty in the risk assessment. EFED concurred on the majority of the Level 3 tools ECOFRAM recommended.

1. Species sensitivity distributions for acute toxicity. Often these additional tests would involve the most sensitive taxonomic groups as indicated in Level 2. The number of species recommended for additional testing is under consideration by EFED. A minimum of four to six additional species may be needed to establish a sufficient level of understanding of the variability in sensitivities beyond that provided in the generic extrapolation factors used in Level 1 and 2. Recommendations by the SAP on the number of additional species tests needed are sought.
2. Additional chronic toxicity tests (e.g. fish full life-cycle, more invertebrate chronic testing) may be needed if chronic risk were indicated in the Level 1 aquatic chronic tests, if prolonged or repeated exposures are expected to occur, if bioconcentration is expected, if the mode of action suggests potential reproductive effects, or if reproductive effects were detected in avian or mammalian chronic testing. The Level 1 abbreviated fish reproductive test and the fish early life-history test only provide a rapid, basic screening for aquatic vertebrate chronic testing. Therefore, if potential chronic effects were indicated by any of the factors listed above, regardless of the results of the abbreviated fish chronic test and fish early life-stage test, the full fish life-cycle test would be needed to establish a more definitive answer. The current fish full life-cycle protocol may need to be revised to provide a more robust test.
3. Analyses of time-varying or repeated exposures. If it is predicted that a water body will receive multiple inputs of a pesticide, or that pesticide concentrations will vary significantly over time relative to the duration of the toxicity tests, then the effects of time-varying or repeated exposures may be examined. These tests may be particularly useful to evaluate the time-course of chronic effects when initial assessments indicate a detrimental effect. RADAR exposure modeling output would aid in the design these studies by providing information on the expected exposures sensitive species would experience in the environment.
4. Population and Community Considerations. Assessments at the population and community-level of biological organization in aquatic systems are now somewhat feasible given the refinement and availability of suitable models. However, it is felt these models at present would be primarily for illustrative purposes than for predictive ones. It is contemplated that one or perhaps more generic population models will be employed at Level 3 to exemplify how impacts to individuals (i.e., mortality, reproduction impairment, etc.) may affect population abundance, production, and persistence. RAMAS models and the PondFX (Oregon State University) population model element will be explored as appropriate candidate models. AQUATOX is a promising Agency supported (OPPT and

Office of Water) community-level model which has stochastic features embedded. To what extent AQUATOX can be adapted for use in probabilistic assessments of pesticides will be investigated. Again, it is contemplated that a generic community model will be used to exemplify community-level consequences from a pesticide exposure.

Suter (1993) argues that the abundance and persistence of populations of organisms are more relevant as endpoints for assessment than are responses of individual organisms observed in controlled laboratory experiments. It seems to follow that community-level assessments would be useful to characterize the further consequences of a population impact. Sources of uncertainty for population analysis include (1) environmental variability in time and space, (2) variations in sensitivity among individuals and their various life stages, (3) stochastic birth and death processes, and (4) the lack of understanding of population dynamics.

5. Evaluation of sublethal effects in a species of concern. When data on the mode of action indicates that sublethal effects may lead to a deleterious effect on a population, additional laboratory testing to better evaluate the effect may be conducted. Examples of this may include testing of immunocompetence compromise, or endocrine disruption. These data would help characterize the possible influence a mode of action has expected effects and their full ecological consequence.
6. Focal species toxicity testing. If the assessment indicates that a particular species or group of species may be adversely affected, and the uncertainty remains high regarding the magnitude of adverse effect, focal species toxicity testing may prove useful.
7. Geographical Information Systems. The use of GIS to identify what proportion of a particular species and habitat are associated with a given pesticide use may assist in probabilistic characterizations of risk. GIS considerations could be tied to focal species testing where a particular concern is indicated.

<u>Risk assessment</u>
Level 3 Risk assessments could include joint probability curves comparing toxicity data to MUSCRAT output, or evaluating species sensitivity acute and/or chronic curves to exposure modeling output. RADAR modeling output may be evaluated with respect to biological effects expected with time-varying exposure. Population and community modeling output could be compared to exposure modeling output from a region(s) relevant to the populations or communities examined.

A potential problem with Level 3 is the lack of standardization which may lead to inconsistencies in risk assessments among chemicals and between individual assessors. EFED intends to develop guidance on which tools would be most useful to select in the third and fourth levels of refinement, given different types exposure/effects profiles, to permit a more consistent, predictable and transparent regulatory process. Other issues that need to be addressed include a full examination of the MUSCRAT model input and processing of output. The current version is under beta testing and has not been officially approved by EPA/OPP/EFED via the EPA/ORD laboratories. EFED will also need to examine in detail the population and community models under consideration to select the best candidate.

Differences from the ECOFRAM Draft report approach:

1. As an interim process ECOFRAM recommended the use of MUSCRAT (Multiple Scenario Risk Assessment Tool) at Tier 2. EFED recommends using MUSCRAT at Level 3.
2. ECOFRAM mentioned the Level 3 use of Monte Carlo analysis to estimate the range of exposure associated with distributions of environmental fate parameters. As noted above, EFED is considering whether this Monte Carlo analysis should be employed in Level 2 or 3, and requests the SAP's recommendation
3. ECOFRAM recommended that a newly-developed tool, RADAR (Risk Assessment Tool to Evaluate Duration and Recovery), be finalized for use at Tier 2. As noted above, EFED recommends moving this tool to Level 3, in order to maintain Level 2 as an initial probabilistic assessment without the need for elaborate refinement and concomitant resource burden.
4. Additional non-guideline fate and transport studies, such as "fate-o-cosm", field runoff studies (small scale and/or field scale), were recommended by ECOFRAM at Tier 3. EFED's plan addresses these additional fate and transport studies at Level 4, including regional watershed scale monitoring studies.
5. As noted above, ECOFRAM also recommended extrapolating the toxicity data to preliminary hypothetical population effects estimates in Tier 2. EFED recommends that initial, generic estimates of population-level effects be undertaken at Level 3.
6. EFED is considering the need to include the evaluation of sublethal effects at Level 3, when data on the mode of action indicates that sublethal effects may lead to a deleterious effect on a population. These data would help characterize the possible influence a mode of action has expected effects and their full ecological consequence. EFED is requesting that the SAP make a recommendation regarding the importance and usefulness of requesting this data in limited situations.
7. Focal species toxicity testing is also under consideration by EFED. If the assessment indicates that a particular species or group of species may be adversely affected, and the uncertainty remains high regarding the magnitude of adverse effect, focal species toxicity testing may prove useful. EFED seeks the SAP's recommendation on this issue.
8. ECOFRAM recommended conditional triggering of benthic species sediment toxicity tests at Tier 3. EFED is considering conditional triggering of benthic species sediment toxicity testing at Level 2, as noted above.

Fourth Level of Refinement

Description

If Level 3 still indicates unacceptable risk or additional issues, and/or there is still insufficient information to make informed risk management decisions, the assessment would proceed to Level 4, on a case-by case basis. Level 4 would yield a probabilistic assessment which is tightly focused on a specific question such as a specific use or scenario, and generally may involve broad field monitoring programs.

Objectives

The objectives of Level 4 are to validate Level 3 predictions and to further reduce uncertainties.

Exposure assessment

The Level 4 exposure assessment would be unique to each chemical. It would yield a probabilistic assessment which is tightly focused on a specific question such as use or scenario, and generally may involve highly refined or watershed-scale modeling or broad experimental, field monitoring or mitigation validation programs. Tool selection would be based on the need for additional information regarding significant sources of uncertainty.

The Level 4 "toolbox" would include:

1. Highly refined watershed evaluations and modeling addressing landscape issues. For example, the basin scale SWAT (Soil and Water Assessment Tool) could be used to investigate the impact of a specific chemical use on a specific crop at certain acreage within the watershed.
2. Additional non-guideline fate and transport studies to better represent important processes. Examples may include "fate-o-cosms" (a termed coined by ECOFRAM to represent small scale fate studies similar in principle to a biological microcosm), field runoff studies (small scale and/or field scale), and field monitoring studies. This may also include regional watershed scale monitoring studies.
3. Detailed investigation of the efficacy of mitigation options. For example, monitoring studies on the effectiveness of specific mitigation efforts in reducing pond levels could be undertaken. A mitigation task force, similar to the Spray Drift Task Force, has been suggested by a peer reviewer of the ECOFRAM document, as well as other scientists, to examine the efficacy of mitigation measures in general. The implementation team members concur that this would be a useful approach. Until the results of a task force for mitigation become known, the chemical-specific evaluation of mitigation options would be useful in cases where it is indicated.
4. For new or limited-use chemicals, benchmark modeling/monitoring relative to existing chemical data. Comparative modeling would be conducted for a new or limited-use chemical that has similar environmental fate and application characteristics to a chemical currently used. This information would be evaluated in combination with monitoring data available on the old, "benchmark" chemical. In this way, a more complete evaluation of the potential concentrations of the new chemical in the environment may be made.

Effects assessment

A Level 4 effects assessment would, again, be specific for a given chemical. Effects tools in the Level 4 toolbox would include:

1. Micro/mesocosm studies.
2. More in-depth population level or community analysis of effects, on a species and

environment-specific basis.
3. Behavioral tests for toxicity.
4. Pharmaco/toxicokinetic modeling. These models would provide information on the disposition of a chemical in an organism, which would be used to predict a dose to target tissues and/or yield body burden information. This approach would be used to compare to critical body residues to exposure information, in order to reduce the uncertainty in conventional dose-response analysis.
5. Field testing/monitoring.

<u>Risk assessment</u>
A Level 4 Risk assessment would refine Level 2 or 3 predictions based on new data.

Differences from the ECOFRAM Draft report approach:
1. The majority of Level 4 exposure tools were the same as that recommended by ECOFRAM. However, additional non-guideline fate and transport studies, such as "fate-o-cosms" and field runoff studies were moved to Level 4, from ECOFRAM's recommended Tier 3. Field monitoring studies, including watershed or regional studies, were grouped with these additional fate and transport studies in EFED's Level 4 plan.
2. Field *effects* testing and/or monitoring was included as a Level 4 option in EFED's plan, but was not fully addressed in ECOFRAM's draft document.

Next Steps in Implementing Probabilistic Ecological Assessments

Following the consultation with the SAP and consideration of recommendations, the next steps EFED envisions taking toward implementing probabilistic assessments include:
1. Development of guidance for movement to higher tiers in collaboration with risk managers
2. Development of extrapolation factors for species sensitivity differences
3. Further analysis and development of Level 3 and 4 tools
4. Revision and addition of guideline tests as needed
5. Development of guidance for appropriate selection of Level 3 and 4 tools
6. Progress on exposure model improvements outlined
7. Further analysis of use of distributions of fate input parameters in Monte Carlo analyses
8. Further analysis and development of GIS tools

III. Outline of the Levels of Refinement for Aquatic Assessments

Level of Refinement	Exposure Approach	Effects Approach	Risk Assessment
1 Modified Deterministic Screen A screening level designed to be protective, not predictive, and to identify compounds thought to present minimal risk to ecological systems Vulnerable headwater environment, high exposure scenario, edge-of-field scenario	Interim Exposure Modeling method: GENEEC, No degradates Monitoring information evaluated for currently registered chemicals.	Data Requirements: Current tests plus potential additional species, conditionally. **Extrapolation factor used** **ACUTE (*):** 1. Two freshwater fish 2. Aquatic invertebrate 3. Aquatic algae/plants **Estuarine/marine** add: 5.One fish 6. Two invertebrates **CHRONIC (*):** 1. Fish early life-stage 2. Screening level fish reproduction 3. Aquatic invertebrate **Estuarine/marine** add: 4. One fish 5. One invertebrate *** Potential conditional additions at Level 1**: If exposure within 1 magnitude of effect: 1. Additional invertebrate species testing 2. Amphibian testing If chemical properties indicate partitioning to sediment: 1. Benthic fish 2. Benthic invertebrates 3. Benthic plants, potentially	Deterministic Risk Quotients Acute: Peak EEC/ (LC_{50},EC_{50}/Extrapolation factor) Chronic: Peak EEC/(EC_{10} or NOAEC/Extrapolation Factor)

Level of Refinement	Exposure Approach	Effects Approach	Risk Assessment
2 Preliminary Probabilistic Assessment Vulnerable headwater environments, edge-of-field scenarios	1. Improved exposure modeling with PRZM3/EXAMS for single or multiple sites (in future uses additional weather data and improved drift modeling) 2. Future potential addition of Monte Carlo using input parameters 3. Monitoring information evaluated for currently registered chemicals.	Level 1 tests with dose-response relationship examined: **No additional toxicity tests if conditional tests are conducted in Level 1** Conditional, based on EEC being within order of magnitude of toxicity information: 1. Additional invertebrate species testing 2. Amphibian testing If chemical properties indicate partitioning to sediment, indicate toxicity based on pelagic species, and model estimates indicate sediment-association, add: 1. Benthic fish 2. Benthic invertebrates 3. Benthic plants, potentially	Probabilistic Joint probability curves: Addresses magnitude and probability of effect (e.g. Gives probability of exceeding concentration vs mortality (etc) at that concentration)

Level of Refinement	Exposure Approach	Effects Approach	Risk Assessment
3 Refined Probabilistic Assessment Similar to Level 2 but refined with additional data/modeling Regional analyses or focus on taxa/scenarios where need more information	Unique to each chemical <u>Possible new data/analyses</u>: 1. Refined modeling a. MUSCRAT b. Monte Carlo Input parameters if not done in Level 2 2. RADAR 3. Additional fate data (lab or field, e.g. more soils) 4. Monitoring information evaluated for currently registered chemicals <u>Ideally:</u> Automated standardized scenarios-Lentic and lotic systems, hydraulic residence times, range of soil and climate for each use, Watershed basis using GIS and by region	Unique to each chemical <u>Possible new data/analyses</u>: 1. Additional acute testing to refine species sensitivity estimates 2. Additional chronic testing to refine species sensitivity estimates. 3. Pulsed testing: Analyses of time-varying or repeated exposures 4. Generic population/community evaluation 5. Sublethal effects for specific mode of action 6. Focal species testing 7. GIS use to identify particular species and habitats of concern	Joint Probability Curves Examples: Species sensitivity vs. probability of exceedence RADAR vs. effect following time-varying/pulse exposure

Level of Refinement	Exposure Approach	Effects Approach	Risk Assessment
4 Most Refined Probabilistic Assessment Case by case, tightly focused on specific question	Unique to each chemical 1. Highly refined watershed evaluations and modeling (e.g. SWAT) 2. Fate and Transport Studies (Field monitoring, Fate-o-cosm, small scale runoff, regional monitoring) 3. Studies of mitigation effectiveness, for new chemicals 4. For new chemicals, benchmark modeling/monitoring relative to existing chemical data	Unique to each chemical 1. Micro/mesocosms 2. Population/community level analysis- species and environment specific 3. Behavioral tests 4. Pharmaco/toxico kinetic models 5. Field testing/monitoring	Refine or validate Level 2 and/or 3 predictions

IV. Example of Assessment Output at Refinement Levels 1 and 2

Level 1: Acute Risk Example

Level 1 Exposure Input and Output Example

The GENEEC (GENeric Estimated Environmental Concentration) model is used by the EFED as the Tier 1 aquatic exposure estimation tool. GENEEC is a single runoff event model, which can account for spray drift from multiple applications. Outputs provide a maximum peak, a 4-day average, a 21-day, and a 56-day average. GENEEC estimates field runoff and spray drift from a 10 hectare field into a one hectare by two meter deep farm pond with no outlet. GENEEC is generic in that it does not consider differences in climate, soils, topography, and crop.

The testing compound ABC to be used on peanuts. It is applied at a rate of 2.0 lb a.i./ac. The product and fate chemistry information is tabulated below.

Parameter	Value
Water Solubility	2 mg/L (ppm)
Adsorption Coefficient	6070 ml/g o.c.
Soil Aerobic Metabolism Half-life	180 days
Hydrolysis Half-life	Stable
Photolysis Half-life	Stable
Aquatic Aerobic Metabolism Half-life	Stable

The GENEEC output is shown below:

```
RUN No.   1 FOR ABC use on peanuts INPUT VALUES
--------------------------------------------------------------------
 RATE (#/AC)    APPLICATIONS   SOIL   SOLUBILITY   % SPRAY INCORP
  ONE(MULT)     NO.-INTERVAL   KOC       (PPM)       DRIFT  DEPTH(IN)
--------------------------------------------------------------------
2.000(  2.000)   1    1         6070.0    2.0          1.0      0

FIELD AND STANDARD POND HALFLIFE VALUES (DAYS)
--------------------------------------------------------------------
METABOLIC  DAYS UNTIL  HYDROLYSIS   PHOTOLYSIS   METABOLIC  COMBINED
 (FIELD)   RAIN/RUNOFF   (POND)     (POND-EFF)    (POND)     (POND)
--------------------------------------------------------------------
 180.00        2           N/A      .00- .00         .00       ****
```

```
GENERIC EECs (IN PPB)
--------------------------------------------------------
    PEAK      AVERAGE 4     AVERAGE 21    AVERAGE 56
    GEEC      DAY GEEC       DAY GEEC      DAY GEEC
--------------------------------------------------------
    7.37       6.46          3.72         2.40
```

Level 1 Effects Output Example
$LC50_A$=27 ppb
$LC50_B$= 6 ppb
$LC50_{combined}$ (Use geometric mean for example) = 13 ppb
Extrapolation Factor=5 (example only)

Level 1 Risk Assessment Example
RQ= EEC/($LD50_{comnbined}$/EF)
RQ= 7 ppb/(13 ppb/5)= 2.7
Acute High Risk LOC=0.5
RQ>LOC. Further analysis shown under Level 2 below.

Level 2: Acute Risk Example
A Level 2 exposure assessment would yield information on the probability of environmental concentrations exceeding a given level, while the Level 2 effects assessment would yield information on the percent effects (e.g. mortality, reduced hatching) at various concentrations. The joint probability function, produced through a joint analysis of the exposure and effects distribution, would provide information on the proportion of the population expected to be affected and the probability of that occurring. For example, the probability that a given percent or more of fish dying, based on individual sensitivities, could be estimated for various pesticide exposure levels resulting from an application on a given crop.

Level 2 Exposure Input and Output Example
Currently, PRZM and EXAMS are linked for a refined estimation of pesticide concentration in surface water environment. PRZM 3.12 simulates the runoff and erosion from an agricultural field, and EXAMS 2.97.5 simulates the fate in a receiving water body.

The example uses the same use scheme as the previous example. In addition to the rate and environmental fate information, PRZM/EXAMS considers soil information as listed below, as well as the site specific weather data.

Parameter	Information
Site	Major Land Resource Area (MLRA) 153A - Atlantic Coast Flatwoods
Soil	Tifton Loamy Sand (Hydraulic Soil Group: C)

ABC Use on Peanuts
WATER COLUMN DISSOLVED CONCENTRATION (PPB)

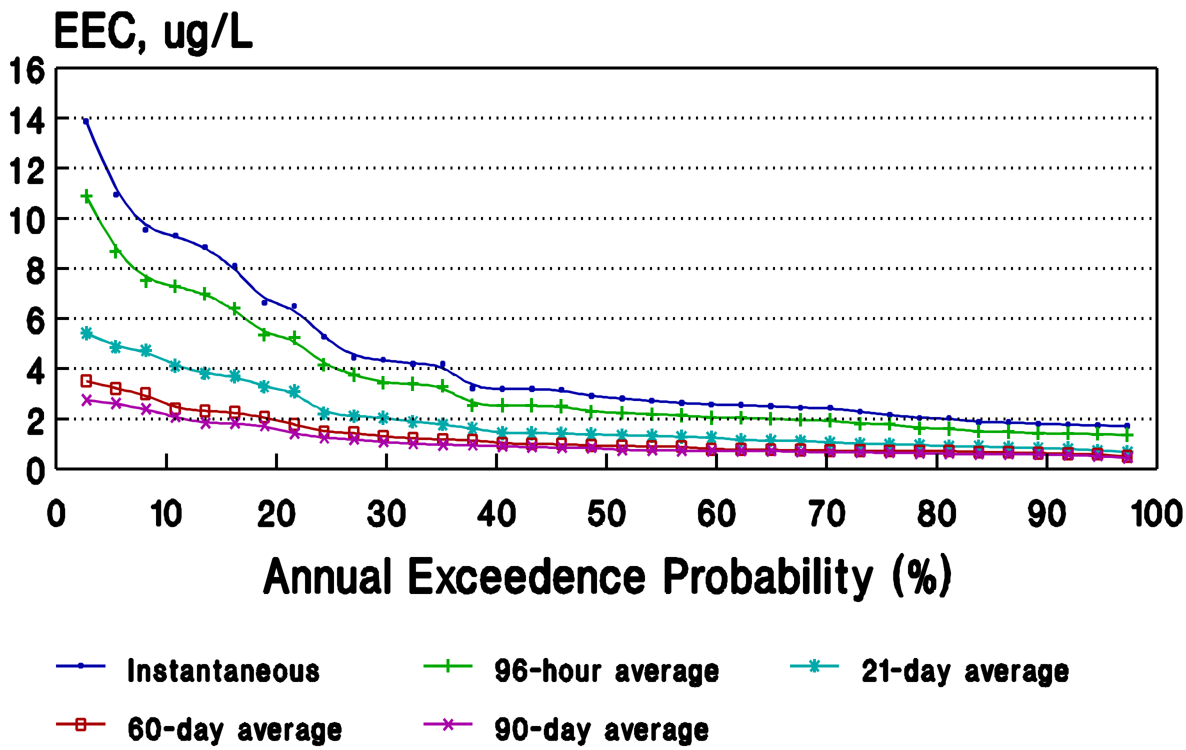
YEAR	PEAK	96 HOUR	21 DAY	60 DAY	90 DAY	YEARLY
1948	1.792	1.416	.734	.589	.513	.225
1949	2.150	1.796	1.006	.722	.648	.305
1950	10.940	8.675	4.862	3.196	2.613	.961
1951	3.192	2.527	1.378	.899	.752	.376
1952	2.017	1.609	.902	.658	.575	.260
1953	1.769	1.413	.811	.672	.574	.268
1954	2.028	1.619	.857	.715	.576	.228
1955	1.857	1.489	1.054	.778	.722	.342
1956	3.149	2.499	1.612	1.181	.945	.356
1957	2.809	2.226	1.342	.728	.605	.296
1958	2.435	1.946	1.115	.774	.687	.384
1959	2.281	1.799	.902	.723	.655	.300
1960	2.631	2.183	1.319	.753	.728	.349
1961	6.624	5.231	3.277	1.777	1.406	.542
1962	9.306	7.294	4.106	2.313	1.810	.665
1963	4.188	3.391	2.108	1.439	1.240	.484
1964	3.205	2.528	1.278	1.019	.900	.382
1965	6.492	5.343	3.089	2.061	1.719	.646
1966	13.870	10.880	5.414	3.514	2.742	.967
1967	2.503	2.006	1.145	.932	.874	.377
1968	2.898	2.290	1.268	.759	.617	.267
1969	8.844	6.983	3.806	2.398	2.060	.749
1970	1.859	1.479	.823	.599	.556	.281
1971	1.732	1.397	.954	.721	.718	.328
1972	5.273	4.150	2.042	1.149	.960	.383
1973	9.538	7.497	4.733	3.001	2.378	.894
1974	4.182	3.297	1.880	1.297	1.063	.445
1975	2.715	2.138	1.140	.989	.839	.357
1976	8.094	6.403	3.702	2.255	1.828	.679
1977	2.546	2.034	1.423	.991	.838	.363
1978	2.566	2.040	1.418	.896	.748	.316
1979	4.353	3.423	1.777	1.214	1.008	.439
1980	1.704	1.349	.680	.497	.437	.192
1981	3.185	2.537	1.433	.914	.711	.359
1982	4.439	3.750	2.206	1.466	1.184	.512
1983	2.442	1.938	.989	.628	.664	.298
upper 10th percentile	9.376	7.355	4.294	2.579	2.155	.792

MEAN OF ANNUAL VALUES = .433

STANDARD DEVIATION OF ANNUAL VALUES = .204

UPPER 90% CONFIDENCE LIMIT ON MEAN = .483

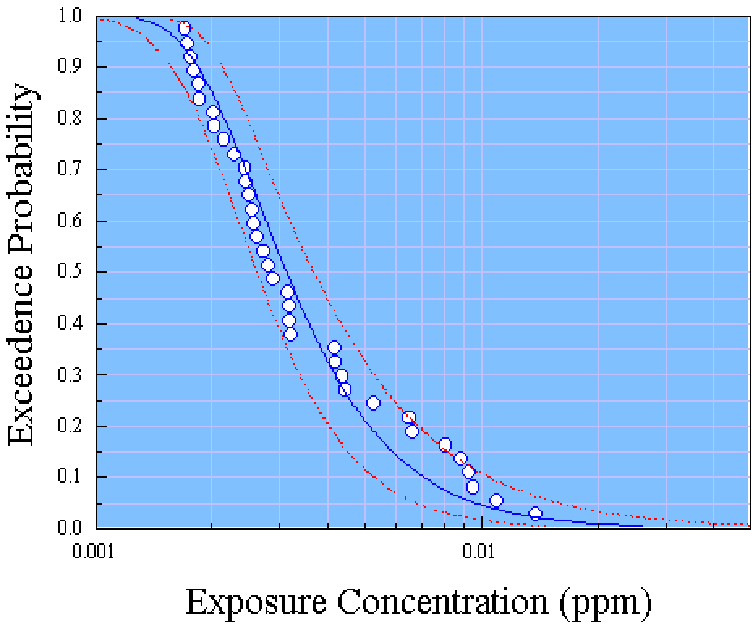
EEC Plot- ABC Use on Peanuts



Tifton Loamy Sand (HSG: C)
2 Ground Applications @ 2.0 lb a.i./ac

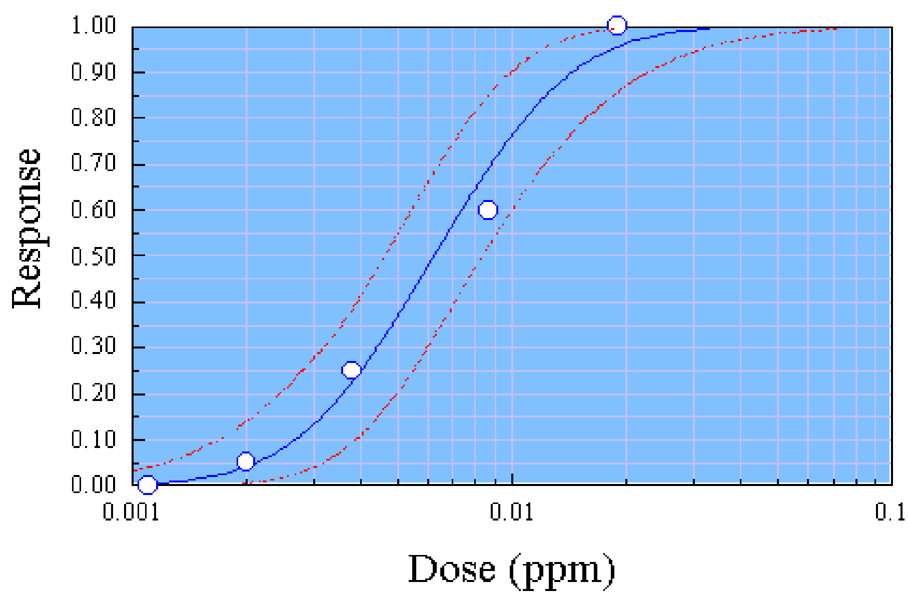
Level 2 Exposure Assessment Example

The probability of peak concentrations being exceeded in the environment is given below, based on 36 years of exposure modeling data. The Y-axis shows the probability that the environmental concentrations exceed the X-axis values. For example, the graph indicates that there is approximately a 30% probability (95% confidence bounds of 20-45%) that the maximum estimated environmental concentration will exceed concentration of 0.004 ppm (4 ppb).

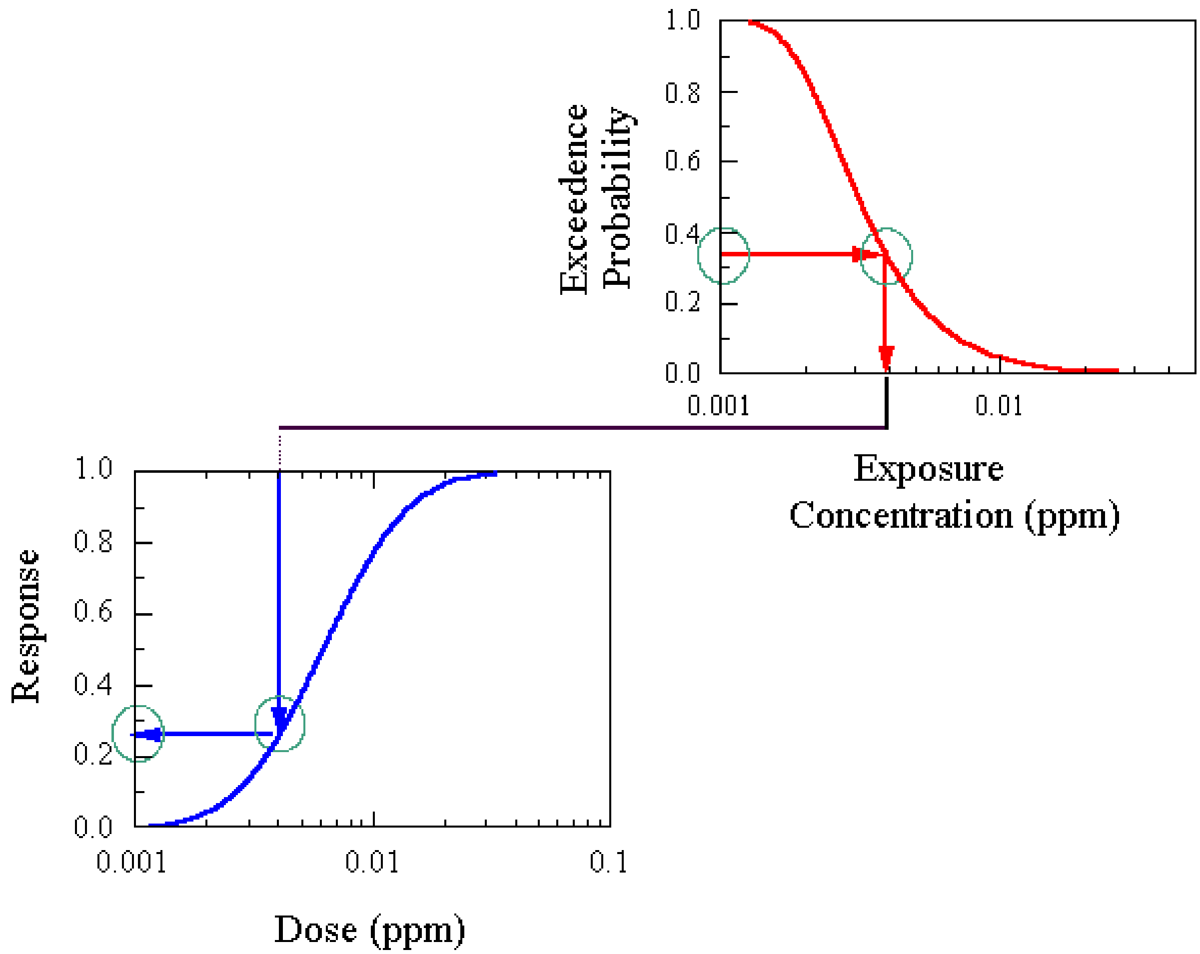


Level 2 Effects Data Example
$LC50_A$=27 ppb, slope = 3.3
$LC50_B$=6 ppb, slope = 3.7
GM LC_{50}= 200 ppb, GM slope= 3.5

Level 2 Effects Assessment Example for Most Sensitive Species
An example of graphical data expressing a dose-response curve for the most sensitive species in this example is given below. The graph indicates the percent mortality after acute exposure, and the 95% confidence bounds. An example of information gleaned from the graph is that the dose that would result in mortality of 25% of the fish is approximately 0.004 ppm (=4 ppb), with 95% confidence limits of 10 to 45% mortality at that exposure level.

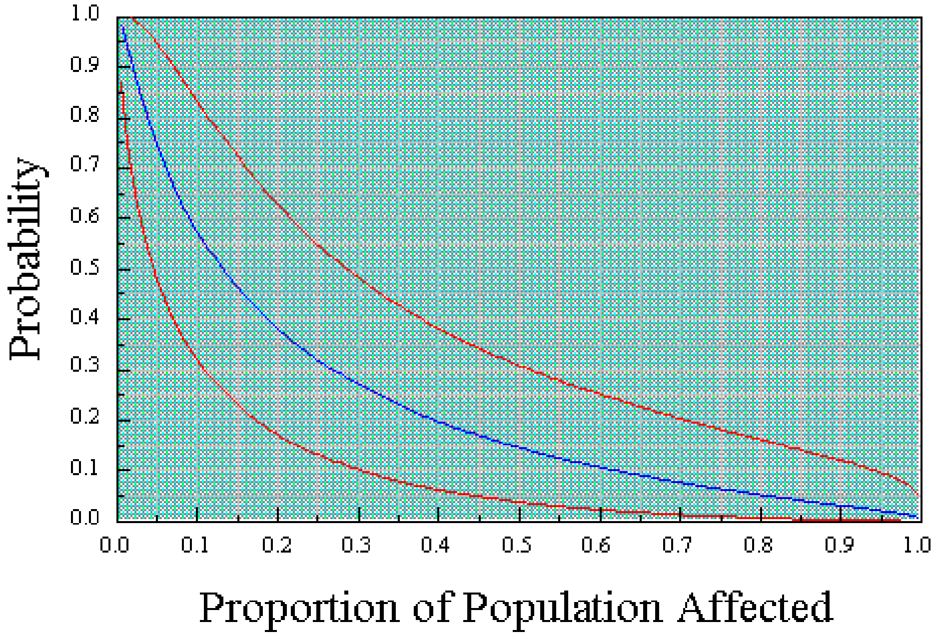


Level 2 Exposure and Effects graphs: Simplified for comparison

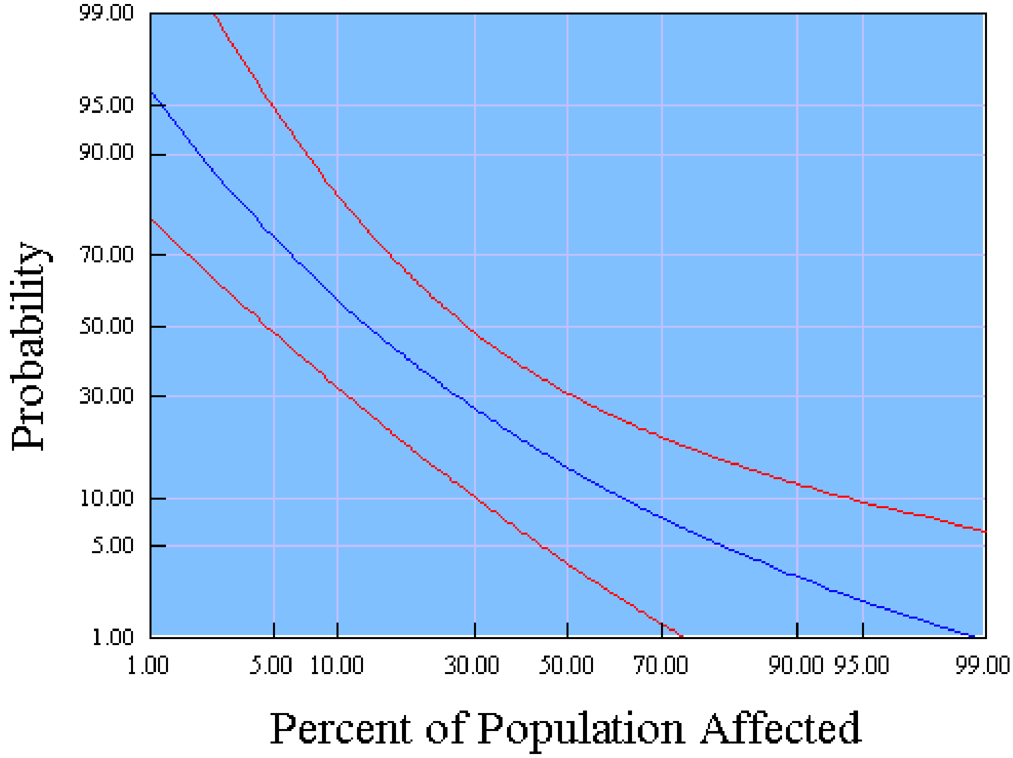


Level 2 Risk Assessment Example : Joint Probability Function, a Probabilistic Expression of Risk
The X-axis is the proportion of the population affected (in this case the risk is mortality) and the Y-axis is the probability of that occurring. For example, the probability that 25% or more of the population is affected is approximately 32%, with 95% confidence bounds of approximately 13% and 55%.
The probability that 50% or more of the population is affected (risk = 0.5) is roughly between 5% and 29% with a best estimate of 13%. Also,

	low estimate	best estimate	high estimate
probability Risk >= 10%	0.30	0.62	0.94
probability Risk >= 50%	0.05	0.13	0.29
probability Risk >= 90%	0.00	0.03	0.12



The axes of the following graph are in probability scales, for ease of viewing effects in the tails of the distribution.



V. Aquatic Exposure and Effects Research Needs to Support Probabilistic Risk Assessments

Exposure Research Needs
1. Exposure simulation models for pesticides: a. Complete field-truthing/validating of PRZM/EXAMS for pesticides b. Upgrading Monte-Carlo shell for PRZM c. Refine link to EXAMS, particularly with respect to Monte Carlo PRZM outputs. d. Refine EXAMS model to permit dynamic depths e. Develop revised approach for adding PRZM-simulated bound fraction of pesticide into EXAMS waterbody
2. ECOFRAM recommended the development of a suite of modeling tools that utilizes modern and well documented coding and offers the capability of adding “modules” as improved approaches and/or algorithms are developed. Ideally, the model suite should handle leaching, runoff and foliar degradative and dissipation processes with distributions of inputs. The same model (driven by different scenarios and "complexity levels") should handle at least Tiers 1, 2 and 3. The objective of including as many processes as possible in one model is to approach the goal of accounting for mass balance. Additionally, the models should be carefully integrated to facilitate automated use and they should automatically produce probabilistic reports in a format that fully documents input parameters and output appropriate to the tier of modeling being conducted.
3. Develop method/model to estimate combined terrestrial and aquatic exposures for amphibians, aquatic birds, mammals.
4. Characterize efficiency of mitigation options with field measurements (e.g. concentrations within and beyond buffer strips).
Plant uptake/dissipation a. Conduct research on foliar dissipation/washoff of pesticides b. Investigate plant interception of pesticides c. Conduct research on uptake from soils into plants
6. Sorption and Degradation: Investigate horizontal and vertical variability in pesticide sorption and degradation associated with soil and aquifer materials
7. Sorption: Evaluate mechanisms of pesticide extraction into runoff
8. Sorption: Examine the effects of adsorption kinetics on runoff and leaching
9. Sorption: Evaluate OPP guidance on sorption/desorption studies with particular emphasis on centrifugation effects, soil:water ratios, kinetic analysis of sorption
10. Degradation: Investigate magnitude of uncertainty introduced into model estimates through the use of soil metabolism information as a surrogate for aquatic sediment metabolism
11. Degradation: Examine the effects of water content, aeration, and microbial activity on chemical degradation through literature research and experimentation.

Exposure Research Needs
12. Degradation: Compilation of literature data on the extent to which oxygen, organic carbon and alternate electron acceptors are transported to subsoils in infiltrating water and an assessment of how this relates to subsoil degradation
13. Develop additional data to support model developments for the occurrence of multiple soils within a watershed/field
14. Compile species range maps, for minor use crop assessments
15. Data on variation of residues within a water body is needed to improve surface water models
16. Conduct field research on streams and other water bodies impacted by urban runoff
17. Research fate characteristics of chiral compounds
18. Assess and characterize paths of bioavailability of pesticides and methodology to predict whether classes of pesticides will move/magnify through the food chain

Effects Research Needs
1. Evaluate additional contributions to uncertainty in probabilistic risk assessments of extrapolations such as lab to field extrapolation, active ingredient to formulation, individual pesticide to mixture exposures, individual to population: commonly used endpoints to population endpoints (lethality, growth reduction to pop growth and stability)
2. Guidelines tests are needed for Time-to-Event toxicity tests. Compile and evaluate literature data on Time-to-Event toxicity tests, and develop protocol for Time-to-Event testing
3. Guidelines tests are needed for laboratory toxicity tests with time-varying/pulsed exposures . Compile and evaluate literature data, and develop test design/protocol
4. Validate population model methods
5. Validate community model methods
6. Develop toxicity tests with amphibians: acute and chronic
7. Develop standardized, easy-to-use biomarkers/sublethal toxicity tests for registration purposes. Examine if *in vitro* assays are viable alternatives to, or useful supporting evidence for, *in vivo* testing for assessing pesticide risk.
8. Evaluate population changes in gene frequency in toxicant-stressed environments

VI. Summary of Recommendations for Alterations and Additions to Guideline Tests for Aquatic Assessments

Changes to Exposure Guidelines Recommended by ECOFRAM:

1.) Add multiple soils for soil aerobic metabolism guidelines
2.) Alter metabolism guidelines in order to obtain rate constants for degradate formation and decline
3.) Redesign aquatic metabolism studies to separate degradation in the water column and sediment
4.) Alter hydrolysis guidelines to obtain hydrolysis rate constants as function of temperature
5.) Alter photolysis guidelines so that quantum yields can be determined
6.) Redesign anaerobic soil metabolism studies to focus on degradation in subsoil horizon aquifers, not flooded soils
7.) Enhance batch equilibrium study design
8.) Add foliar dissipation and washoff guideline requirements for foliar pesticides (or any spray application as suggested by EFED scientists)
9.) Add guideline for potential of uptake from soils into plants

Also recommended by EFED scientists:

10.) Add guideline test for subsoil and aquifer material aerobic metabolism
11.) Add guideline test for sorption to aquifer material
12.) Write Standard Evaluation Procedures for anaerobic and aerobic aquatic metabolism studies

Changes to Effects Guidelines:

1.) Add guideline tests for additional invertebrate test species
2.) Add guideline tests for amphibian test species
3.) Guidelines tests are needed for Time-to-Event toxicity tests. Compile and evaluate literature data on Time-to-Event toxicity tests, and develop protocol for Time-to-Event testing.
4.) Guidelines tests are needed for laboratory toxicity tests with time-varying/pulsed exposures. Compile and evaluate literature data, and develop test design/protocol.
5.) Review and revise acute and chronic tests as needed to improve statistical power.
6.) Develop fish benthic toxicity test.

VII. References

Barry, M.J., K. O'Halloran, D.C. Logan, J.T. Ahokas, D.A. Holdway. 1995. Sublethal Effects of Esfenvalerate Pulse-Exposure on Spawning and Non-Spawning rainbowfish (*Melanotaenia fluviatilis*). Archives of Environmental Contamination and Toxicology 28(4): 459-463.

Carbone, J.P., P. Havens, W. Warren-Hicks. 1999. A Critical Evaluation of PRZM 3.12 Estimated Environmental Concentrations Accounting for the Uncertainty Associated with Measured Environmental Fate Data and Model Inputs. Abstract. Society of Environmental Toxicology and Chemistry 20th Annual Meeting, November 1999, Philadelphia, PA.

Ecological Committee on FIFRA Risk Assessment Methods (ECOFRAM) Aquatic Draft Report 1999. http://www.epa.gov/oppefed1/ecorisk/aquareport.pdf

ECOFRAM Peer Input Workshop Comments. 1999. Various authors. http://www.epa.gov/oppefed1/ecorisk/index.htm

Mayer, F.L. and M.R. Ellersieck. 1988. Manual of Acute Toxicity: Interpretation and Database for 410 Chemicals and 66 Species of Freshwater Animals. US Fish and Wildlife Service Resource Publication 160.

Montague, B. 1992. Pesticide Ecotoxicity Database. EPA Office Pesticide Programs, Environmental Fate and Effects Division.

Nimmo, D.R., T.L. Hammaker, J.C. Moore, R.A. Wood.1980. Acute and Chronic effects of Dimilin on survival and Reproduction of Mysidopsis bahia. Pp. 366-376. In: Aquatic Toxicology, ASTM 707. J.G. Eaton *at al.*, (eds.) American Society for testing and materials, Philadelphia, PA.

Suter, G.W. II. 1993. Ecological Risk Assessment. Lewis Publishers, Chelsea, MI. 538 pp.

Vall, M. J.T. van der Wal, J. Hermans, J. Hoekstra. 1997. Pattern Analysis of the Variation in the Sensitivity of Aquatic Species to Toxicants. Chemosphere 35(6), 1291-1309.

Attachment 1
Summary of the Aquatic ECOFRAM Draft Report

The ECOFRAM Aquatic draft report presents a probabilistic approach to ecological risk assessment that proceeds through a four-tiered analysis, from simple deterministic quotients through a general probabilistic analysis, to an issue-specific probabilistic assessment at the highest tiers.

Tier 1

Description

Tier 1 uses simple, relatively conservative assumptions yielding a deterministic risk quotients to be compared to established levels of concern. This tier is designed to be protective, not predictive. Tier 1 is aimed at addressing the risk to aquatic species in vulnerable headwater environments resulting from a high exposure, edge-of-field scenario.

Objectives

The stated objectives of this tier are to identify products which appear to present minimal environmental/ecological concerns, to focus higher tiered assessments on combinations of use patterns and taxa most likely to be of concern, and to prioritize use patterns in terms of potential environmental exposures, provide an assessment of whether acute or chronic concentrations may be of concern, and to determine the need to consider sediment toxicity impacts.

Exposure assessment

The recommended Tier 1 exposure assessment is primarily based on GENEEC, as an interim measure. The output provides estimated peak concentrations, 96 hour, 21 day and 60 day time-weighted average water column concentrations. The current selected site severity is based on approximately the 90th percentile Tier 2 site for cotton, ranked by erosion potential. Proposed for the future is, ideally, an exposure tool with multiple scenarios covering a wide range crops, use patterns, multiple water body types, all of which are adjacent to a treated area. The scenario severity is also proposed to be set relative to Tier 2 (e.g. 90th percentile of 10 year return frequency of exposure). The output would ideally yield water-column instantaneous and time interval concentrations for static ponds and flowing waters, with an estimate of variability. Also included in the ideal output would be sediment and pore water concentrations. Degradates were not included in the process.

Effects assessment

The Tier 1 effects assessment is essentially OPP's current Tier 1 assessment, involving LC_{50} estimations for 4 to 7 species. Acute testing would include: one invertebrate, one warm water and one cold water fish, one or more species of algae or higher aquatic plant. If there is potential for estuarine exposure, one acute test each for an estuarine fish, arthropod, mollusk are recommended. It was stated that, usually, freshwater species could be used as surrogates for estuarine/marine species. In a slight change from current policy, ECOFRAM recommended that a chronic freshwater or marine invertebrate test, and one or more early life-stage test with fish be added as a mandatory data requirement at tier 1. ECOFRAM also recommended that acute data be collected in 24 hour intervals, at least, and that the dose-response slope be calculated for each time for use in Tier 2 evaluations. Time-to-event testing was recommended as a possible additional change of current methodology, since it "generally provides a better estimate of acute LC_{50} values than conventional probit analysis." For the chronic data, a regression based EC_{10} was recommended as the regulatory endpoint in place of an NOAEC. It was stated that a NOAEC should only be used if a study is technically acceptable, but the data do not support regression analysis.

Risk assessment

The risk assessment will be based on deterministic risk quotients. The acute risk quotient is the ratio of model-estimated peak environmental concentration to LC_{50} or EC_{50}, based on the most sensitive species within a taxa. If chronic data were not available for the species that was most sensitive in the acute tests, ECOFRAM recommends that an acute-to-chronic ratio derived for another species be used to estimate the chronic endpoint for the most sensitive species. The chronic risk quotient is the ratio of model-estimated *peak* environmental concentration to EC_x or NOAEC, instead of a time-weighted average concentration. ECOFRAM noted that quantal effects such as mortality and hatching may reflect effects of short-term exposure at critical stages in the life-cycle, while continuous variables such as growth may generally reflect the effects of cumulative exposure. Since Level 1 is intended to be protective, they therefore recommended that the chronic risk quotient should be based on the peak EECs.

The calculated risk quotients are compared to current risk criteria (Levels of Concern). If risk criteria are not exceeded, it is concluded that minimal aquatic risk is posed from the proposed use, and the risk assessment process is considered complete. If the risk quotients exceed the risk criteria, the risk assessment proceeds to Tier 2, but only for the taxa or scenarios that are indicated to be of concern.

Tier 2

Description

If the output of Tier 1 suggests that there may be a concern with the chemical under review, the risk assessment will move to Tier 2. Tier 2 yields a preliminary probabilistic risk assessment using complete dose-response relationships derived in Tier 1, aimed at addressing the risk to aquatic species in vulnerable headwater environments, including regional and national edge-of-field scenarios. ECOFRAM proposed that Tier 2 assessments be focused *only* on taxa or scenarios where LOC's are exceeded in Tier 1. Tier 2 thus refines the Tier 1 exposure assessment only for selected situations.

Objectives

The objectives of Tier 2 are to characterize spatio-temporal variations in risk to headwater/static ecosystems, to confirm that risk predicted in Tier 1 still applies when physico-chemical processes and environmental fate parameters are more fully expressed, to provide probabilistic expressions of risk to potential problems indicated in Tier 1, to extend the interpretation of potential effects through simple, generic population level analysis, to consider which regional/cropping patterns or effects may merit additional refinement at Tier 3, to permit preliminary evaluation of mitigation options, and to provide guidance on which Tier 3 approaches may be most useful for refining the assessment.

Exposure assessment

The exposure assessment at this tier is intended to characterize spatio-temporal changes in risk to headwater/static ecosystems. As an interim process ECOFRAM recommends the continued use of

PRZM3/EXAMS for multiple sites using MUSCRAT (Multiple Scenario Risk Assessment Tool). MUSCRAT, an automated processor linking PRZM and EXAMS, is a currently available tool to accomplish the multiple site evaluation aims of this tier. The water body simulated is a static pond, adjacent to the crop of interest. ECOFRAM noted that numerous improvements to PRZM/EXAMS and MUSCRAT are needed to improve interim Tier 2 modeling.

ECOFRAM also recommends that a newly-developed tool, RADAR (Risk Assessment Tool to Evaluate Duration and Recovery), be finalized for use at Tier 2. RADAR is essentially used as a postprocessor for EXAMS, and yields information on exposure event magnitude and duration, which can be compared to a concentration or response threshold for use in population recovery estimates following pulsed exposure.

Ideally, ECOFRAM would envision including automated, standardized scenarios for each water body, a range of soil and weather conditions, which would automatically be generated for the appropriate use pattern. The scenarios that would be ideally included would be selected on a watershed basis, possibly using GIS, and grouped by region. Each region would contain lentic and lotic systems capable of addressing different appropriate species and hydraulic residence times. A range of soil and climates would automatically be generated for each use. The modeling tool used would optimally be modular with respect to databases and algorithms for flexibility. Three levels of output were suggested. (1) Cumulative area-weighted probability curves, nationally and by region. The output would be distributions of peak, 24 hour, 48 hour, 96 hour, 21 day, 60 day, 90 day durations for monthly, seasonal and annual maximum series. Ideally, thematic maps where soil/climate combinations likely to occur at different risk endpoint levels would be included in this output level.(2) Frequency distributions produced for each individual scenario, and (3) Summary information on mass loadings (drift, runoff, etc.) for the scenarios closest to the assessment endpoint criteria would be included to assist in mitigation considerations.

Effects assessment

The Tier 2 effects assessment consists of current toxicity tests, as listed in Tier 1, with the inclusion of the slopes of the dose-response curves. No additional toxicity tests were included. ECOFRAM also recommended extrapolating the toxicity data to population effects estimates. This population-level analysis will only involve a comparison of hypothetical populations representing common life-history strategies. For example, the effects of a given exposure for a species with high reproductive rates and a short lifespan (*e.g. Daphnia*) could be compared to the effects on a relatively low reproductive rate, long-lived species (e.g. trout), using time-to-recovery estimates versus concentration estimates. This level of analysis will provide only rough population estimates of widely differing species.

Risk assessment

The Tier 2 risk assessment yields a probabilistic risk assessment. Joint probability curves are produced. This approach combines the exposure estimates, describing the probability of the environmental concentration exceeding a given level, with the estimates of biological effects observed various concentrations. The joint probability curves produced from these estimates would describe the probability of the environmental concentration exceeding a level which would result in given magnitude of negative biological effect, such as mortality or reproductive impairment, for a given scenario. The curves are only produced for scenarios/species that exceed acceptable risk levels in Tier 1, and may be based on maximum or time-weighted EECs. The dose-response curve can be considered to represent the percent probability of an individual being affected or the percent of the population that will be affected. For an exposure distribution modeled, the output might be that 90 percent of the time (1 in 10 year returns), the annual maximum concentration at each of 25 sites for particular use pattern in particular region will be less than a given level. This could be compared to the mortality versus concentration curve for *Daphnia*. This would yield output statements such as 'in 20 percent of the crop sites in a given region, a concentration causing 60% of the *Daphnia* to die will be exceeded, on average, one year in ten, after an application of chemical X.' The assessment would then proceed to outline the consequences of the potential effects. For mitigation purposes, the exposure models could be rerun with varying buffer sizes, and compared to resultant estimated effect.

At this tier, the exposure output from RADAR, addressing the exposure magnitude, duration, and length of recovery time intervals, could be compared to hypothetical population model estimates, to attempt to describe in a generic sense exposures a population level effects.

Because Tier 2 only addresses taxa and scenarios that were shown to be problematic in Tier 1, probabilistic risk assessments will only be conducted for a minority of taxa and scenarios.

Tier 3

Description

If the output from Tier 2 still suggests that the chemical still poses an unacceptable risk, or that the uncertainty of the risk estimates are high, the assessment proceeds to Tier 3. Tier 3 involves further probabilistic refinement of the Tier 2 risk assessment, using additional data or modeling. Tier 3 focuses on sites where the risk is estimated to be high, aspects of the assessment where more information is needed, or on the effects of potential mitigation. ECOFRAM states that the approach at Tier 3 involves determining what the risk drivers are, and where in Tier 2 "conservative simplifying assumptions resulted in unrealistic output." The workgroup also noted that many compounds that reach Tier 2 will proceed to Tier 3, and that many registrants will have completed work to this tier prior to EPA evaluation of the information.

Objectives

The objectives of Tier 3 are to reduce the uncertainty in the Tier 2 assessment, and to better understand the nature of the risk and/or the potential mitigation options.

Exposure assessment

Because Tier 3 focuses on specific problem areas, the Tier 3 exposure assessment would be unique to each chemical, and, according to ECOFRAM will "often be focused on impacts of mitigation alternatives." Possible new exposure data or analyses were proposed based on a "toolbox" approach. Based on best professional judgement, the assessor would select from the "toolbox" the "tool" most appropriate to fulfill the needs of the risk assessment. The Tier 3 exposure tools include:

9. Refinements of Tier 2 modeling to help define uncertainty and evaluate exposure output sensitivity to various parameters. Examples include more sophisticated modeling such as

RICEWQ for a specific scenario, use of 'typical case' parameters as opposed to more conservative assumptions, examination of region-specific differences in environment or uses for regional label modifications; evaluation of more mitigation options; comparison of modeling to monitoring for closely related compounds; AgDrift considerations to reduce drift input, Monte Carlo analyses to estimate the range of exposure associated with distributions of environmental fate parameters, comparison of modeling output with existing monitoring data for closely related chemicals.

10. Additional environmental fate data acquisition and analysis (lab or field). For example, testing with more soils, or examining foliar degradation rates.

11. Fate and Transport studies to better represent important processes. Examples include "Fate-o-cosms", small scale runoff studies, field soil metabolism studies.

12. Landscape configuration issues for model scenarios. Examples include GIS information for percent crop information, soil/slope distributions for the crop of interest in areas of concern, or proximity to water issues.

<u>Effects assessment</u>
The Tier 3 effects assessment would also be unique to each chemical, with consideration given to species or issues where negative outcomes are expected or additional information is required. As outlined by ECOFRAM, the Tier 2 effects "toolbox" includes several possible new effects data tests or analyses including:

1. Analyses of time-varying or repeated exposures. If it is predicted that a water body will receive multiple inputs of a pesticide, or that pesticide concentrations will vary greatly over time relative to the duration of the toxicity tests, then the effects of time-varying or repeated exposures should be examined. RADAR output would aid in the design these studies by providing information on expected exposures sensitive groups would be expected to experience in the environment.

2. Species sensitivity distributions for acute toxicity. More species would be tested if concern remained regarding acute toxicity, or if substantial variability in sensitivity is demonstrated or expected based on a pesticide's mode of action. Often these additional tests would involve testing the most sensitive taxonomic groups to examine interspecies sensitivity differences.

3. Additional chronic toxicity tests (e.g. fish full life-cycle, more invertebrate chronic testing) only if chronic risk was found at lower tiers, or if prolonged/repeated exposures are expected to occur, if bioconcentration is expected, or if the mode of action suggests potential reproductive effects. In the ECOFRAM-recommended approach, as outlined in their draft report, this is the first inclusion of fish full-life-cycle studies, thus, there would be no information on most fish reproductive effects prior to this tier.

4. Sediment toxicity testing. Based on the results of the acute and chronic risk estimates using

pore water concentrations from PRZM compared to invertebrate (pelagic species) toxicity testing from Tier 1 or 2, then acute or chronic testing with benthic invertebrates may be suggested.

Risk assessment

Tier 3 risk assessment could include joint probability curves comparing RADAR modeling output with biological effects associated with time-varying or repeated exposure, or the exposure exceedence estimations for species with different sensitivities.

Tier 4

Description

If unacceptable risk or additional issues were still indicated following Tier 3 analyses, the assessment would proceed to Tier 4, on a case-by case basis. Tier 4 would yield a probabilistic assessment which is tightly focused on a specific question such as a specific use or scenario, and generally may involve broad experimental, field monitoring or mitigation validation programs.

Objectives

The objectives of Tier 4 are to validate Tier 3 predictions and to further reduce uncertainties.

Exposure assessment

The Tier 4 exposure assessment would, again, be unique to each chemical. Examples of proposed Tier 4 tools include:

1. Widespread monitoring of water, for reregistration of chemicals. ECOFRAM suggests that monitoring results "should not be given undue emphasis" and that the "the workgroup members feel that compound specific monitoring should only be set in place when relatively sophisticated modeling has indicated key issues and regional differences."

2. Detailed investigation of the efficacy of mitigation options, especially for new chemicals. For example, monitoring studies on the impact of various mitigation approaches, in well-characterized sites.

3. Highly refined watershed evaluations and modeling addressing landscape issues.

4. For new chemicals, benchmark modeling relative to existing chemical data. Comparative modeling would be conducted for a new chemical and an old, "benchmark" chemical with similar environmental fate and application characteristics. Monitoring information for the benchmark chemical could then provide surrogate field information on potential environmental levels of the new chemical.

Effects assessment

A Tier 4 effects assessment would, again, be specific for a given chemical. Effects tools in the Tier 4 toolbox would include:

1. Population level analysis of effects, on a species and environment-specific basis.

2. Pharmaco/toxicokinetic modeling. These models would provide information on the disposition of a chemical in an organism, which would be used to predict a dose to target tissues and/or yield body burden information. This approach would be used to compare to critical body residues to exposure information, in order to reduce the uncertainty in conventional dose-response analysis.

3. Behavioral tests for toxicity.

4. Micro/mesocosms.

Risk assessment

A Tier 4 risk assessment would refine Tier 2 or 3 predictions based on new data. For example, Tier 3 predictions of time-varying exposure and resultant effect could be further evaluated through the use of exposure monitoring data and/or mesocosm studies.

Brief Comparison of Aquatic and Terrestrial Draft Documents

1. Levels of Refinement (Terrestrial) versus Tiers (Aquatic)

The terrestrial draft document states that the Levels of Refinement are intended to be used in a flexible manner. Specifically,
- in a completed assessment some components may have been refined to a higher level than others
- there is no requirement to refine every component to one level before proceeding to the next

The terrestrial workgroup also stated that it did not regard the assignment of methods to levels as definitive, and that further development and experience might suggest modifications.

The aquatic draft document stated that the tiered approach provides necessary structure and organization, defines a progression for refined assessments, but that the "separation between tiers is not intended to be rigid." (their emphasis). They noted that all relevant data should be used and that valid effects and exposure comparisons may cross tier boundaries. For example, they noted that higher tiered effects data, if available, may be compared with exposure estimates generated at lower tiers (or vice versa). As the assessment is refined, the assessor may identify the most appropriate tool or tools from the toolbox for each tier. For the higher tiers, tools are suggested but "there is no set process or required studies."

In both cases the assessment begins with conservative assumptions and is refined by moving towards more realistic estimates or chemical-specific issues. The refinement process is intended to ensure that appropriate resources are expended on pesticides/issues of concern. In both the terrestrial and aquatic documents, it is noted that as the process proceeds through refinements the measurement endpoints may change as the risk assessment proceeds through higher levels, but the assessment endpoints remain unchanged.

2. Exposure Model Status and Development.

The Aquatic ECOFRAM workgroup had the benefit of using, as a starting point, currently existing exposure simulation models, which are much used in the current pesticide regulatory approach. The aquatic exposure modeling was focused on exposure from the water column; sediment exposures, and particularly potential dietary exposures, were not as fully considered. Future development of new, more refined and complex aquatic exposure models, which could more fully address watershed and regional considerations, were suggested by ECOFRAM

The terrestrial group had the difficult task of developing a new theoretical exposure model. The theoretical model they developed addressed a number of routes of exposure, including diet, preening, inhalation, and dermal exposures. Currently, well-developed terrestrial exposure simulation models are not available and/or used in a pesticide regulatory context.

3. Differences/Similarities Between the Terrestrial and Aquatic Tiers and Data Requirements

A. Level/Tier 1: Both are deterministic

i. Extrapolation factors

The terrestrial group proposed the use of a conservative uncertainty factor to be applied to toxicity data due to its limited nature at this level. In the terrestrial approach, the single avian LD_{50} dose-response is multiplied by an extrapolation factor to account for variation in species sensitivity. In the aquatic approach no extrapolation factor is applied when taxa have single measured LC_{50} values.

ii. Chronic Data requirements

The terrestrial group has recommended 2 reproductive toxicity tests for birds at Level 1. The aquatic group does not recommend reproductive toxicity tests for vertebrates until Tier 3.

B. Level /Tier 2: Both are probabilistic

a. Increased data requirements in terrestrial approach

The terrestrial group has noted that uncertainty is decreased at Level 2 through the collection of additional toxicity data. An additional 1 to 2 acute oral toxicity tests are recommended at this level. The aquatic group does not recommend any additional effects data collection over that required in Tier 1; the dose-response curve for the Tier 1 data is simply now included as part of the analysis.

C. Level/Tier 3

a. Behavioral toxicity tests

The terrestrial group recommends the avoidance behavior be incorporated into the analyses at this level. Additional behavioral effects of exposure are not recommended at this time. For the aquatic group behavioral toxicity tests are included in the Tier 4 toolbox.

IV. Level/ Tier 4

Both groups recommend including landscape analyses, cropping and regional specific issues and field studies at this level.

Attachment 2
Summary of Peer Input Comments on Aquatic Report

The Peer Input Workshop for the Aquatic ECOFRAM Draft Report yielded approximately 100 pages of comments. All of the aquatic report reviewers concluded that the basic approach taken by ECOFRAM was scientifically sound. Beyond that, there were a number of similar comments that the majority of the reviewers made, and numerous specific comments and suggestions made by individual reviewers. General comments made by a number of the reviewers included the following:

General issues

1. The need for validation of all models was noted by most reviewers. Particularly noted by several reviewers was the need to consider the effect of propagation of errors when using or combining multiple-input models. Several reviewers noted the need for validation or field-truthing of the exposure simulation models in particular, due to the reliance of the approach on the ability of exposure models to simulate actual field levels of pesticides. A number of reviewers stated that assumptions of conservatism of the models should be proven. One reviewer noted that, to a large extent, model error was not addressed in the document.

2. The majority of reviewers noted the importance of describing uncertainty at each tier; several noted the need to describe confidence bounds on joint probability curves in particular. Conducting sensitivity analyses to establish the relative importance of additional tests was suggested.

3. Several of reviewers pointed out the fact that this approach does not address the effects of multiple stressors or multiple chemicals, nor does it fully address indirect effects of exposure to a single pesticide.

4. A number of reviewers described the need to train personnel in probabilistic risk assessment methods as crucial to the success of the proposed approach.

5. Several comments addressed the need for case studies using historical data to examine the viability and accuracy of the proposed process relative to the current one.

6. The creation of a webpage which would contain information on conducting probabilistic risk assessments was endorsed by several reviewers. The webpage would house items such as links to data requirements and currently accepted exposure models.

7. Many reviewers noted that a some of the tools described need further development prior to use.

8. The increase in time and resources required to conduct a probabilistic risk assessment was noted as a concern by several reviewers.

9. Several reviewers noted an over-reliance on exposure modeling and minimization of monitoring data and field studies in the ECOFRAM approach. Most of these reviewers advocated the judicious use of monitoring data, in collaboration with modeling output.

10. The effects of scale of the assessment was noted by several reviewers as a critical factor. A number of reviewers were concerned that ecological effects in small, sensitive areas may be overlooked in regional assessments. Some reviewers noted that the use of PRZM/EXAMS to evaluate basin-scale processes may not be realistic.

11. A few of the reviewers cited the need to consider groundwater in addition to surface water in evaluating potential sources of exposure.

Specific comments related to methods and data requirements

1. Approximately half of the reviewers recommended additional effects data requirements be included in the approach; this additional information would be necessary to obtain a reasonably well-characterized percentile effect function. One reviewer stated the belief that the current effects data requirements were sufficient for the lowest tier assessments. One reviewer recommended reducing current exposure data requirements.

2. A few reviewers were concerned that considering only the most sensitive taxa beyond tier 1 was too restrictive. One reviewer noted that they would be uncomfortable stating minimal concern based solely on tier 1 screening, even if it is conservative.

3. Time-to-event and pulsed toxicity tests were endorsed by approximately half of the reviewers. The remaining reviewers did not comment on these testing approaches. The use of the RADAR modeling tool (Risk Assessment tool to evaluate Duration And Recovery) was endorsed in this context.

4. Some of reviewers agreed with the ECOFRAM recommendation of using a regression-based approach to calculate chronic effects (e.g. use of an EC_x), in lieu of the current hypothesis testing approach. The effect of rewarding poor experimental design and high control variability by using the hypothesis testing approach was cited as one of the reasons. Guidance for improving the designs of chronic tests to improve data gathered under the regression approach was recommended by a reviewer. One reviewer noted that comparing peak estimated environmental concentrations to a chronic endpoint would only be appropriate at tier 1, and that at higher tiers the chronic toxicity data should be compared to time-weighted average concentrations.

5. Several reviewers expressed the opinion that aquatic population models are not in a sufficient state of development to permit their routine usage for regulatory decisions. The life table and logistic approaches were cited by one reviewer as currently the most feasible approaches for regulatory needs.

6. A few reviewers stated that there was a continuing need to conduct small scale field experiments and/or limited biological monitoring to verify the accuracy of risk predictions, at least on a limited scale.

7. A few reviewers questioned the reliance on functional redundancy in the ecosystem. The assumption that protecting 90% of the species is sufficient to protect an entire ecosystem was also called into question. The consideration of effects on keystone species within an ecosystem was noted as critical.

8. Some of the reviewers noted that the choice of a single value to represent variable fate data in model inputs was problematic. Several recommended using the distribution of fate parameters available as model inputs, where relevant. Sensitivity analyses were recommended in conducting this evaluation.

9. A number of reviewers approved of the suggested use of GIS data and landscape considerations in refining exposure estimates.