ESA STAKEHOLDER WORKSHOP (JUNE 29-30, 2016):

Breakout Sessions: Improving Aquatic Modeling

<u>Breakout Session AQUATIC 1: Changes to Conceptual Models and Mathematical Approaches Incorporated into Bins 3 and 4 (Flowing Waters):</u>

In the draft Biological Evaluations (BEs), effect determinations are made at the individual scale of biological organization. Consequently, the goal is to accurately predict maximum pesticide concentrations that may occur in different aquatic habitats utilized by listed species and are spatially and temporally relevant to the listed species. The modeling approach presented in the draft BEs leveraged EPA's current generic aquatic modeling approach by using the Pesticide in Water Calculator (PWC) shell, a combination of field-scale models (PRZM5/VVWM), to generate estimated exposure concentrations (EECs) for three generic flowing water bins of varying volumes and flow rates (Bins 2, 3, and 4). The Bin 2 estimates are intended to represent lower-flow habitats, such as first-order streams. When considered in relation to field-scale monitoring data, such as those obtained from edge-of-field (EOF) studies, model results should provide confidence in EECs for this bin. There is expected to be less confidence in applying this approach for deriving estimates for Bins 3 and 4, because processes that affect larger-scale concentration dynamics (*e.g*., longitudinal dispersion) are not accounted for. The EECs derived for these higher-flow habitats in the draft BEs are extremely high and seem to defy both professional judgement and typical patterns seen in contaminant monitoring data.

In the context of watershed hydrodynamics, the three flowing bins represent aquatic habitats which would ideally be representative, for example, of streams that are sequentially connected within a watershed. While runoff and drift from a field adjacent to a Bin 3 and/or 4 waterbody can directly contribute loading, the EECs generated from these types of events are being characterized with Bin 2 EECs, as these EECs may be reflective of concentrations occurring before complete mixing within the Bin 3 and/or 4 waterbody had occurred. Initial modeling generated Bin 3 and 4 EECs that exceed those generated for Bin 2, which runs counter to expectations based on standard transport dynamics, *e.g.,* dispersive dampening of chemographic peak maxima as a pulse of contaminant moves downstream. Given the apparently unreasonably high EECs for Bins 3 and 4, a qualitative approach was considered in the draft BEs for use in assessing these bins. The approach relied on monitoring data to demonstrate a downward trend in the magnitude of peak exposures. Consistent with published studies showing a reduction in exposures as one moves down a watershed network, the approach showed a 5-fold reduction in exposure from Bin 3-like streams and a 10-fold reduction from Bin 3-like streams to Bin 4-like streams. The draft BE also applied a qualitative comparison of volumes and flowrates to suggest a reasonably conservative magnitude of exposure expected in Bins 3 and 4 as a separate line of evidence.

Charge Questions:

1. EPA explored several factors in using the PWC, including incorporation of a baseflow and use of the daily average instead of the instantaneous peak EEC. What are the strengths and weaknesses of these modifications? Are there other modifications that can be made and what are their strengths and weaknesses?
2. How appropriate are the methods used in the draft BEs to develop field/watershed sizes and waterbody lengths for these Bins? What reasonable alternatives could be used to model

watershed processes that allow for accurate estimation of possible exposure concentrations (including the maximum) in these flowing bins based on product labeling?

3. For the bins (3 and 4) that represent larger flowing systems, what ways of incorporating the effects of dispersive mixing and/or peak desynchronization into concentration estimates are reasonable?
4. What are the strengths and weaknesses of alternative mechanistic or regression-based watershed models such as the Soil and Watershed Assessment Tool (SWAT), the Hydrological Simulation Program-Fortran (HSPF) and the Watershed Regressions for Pesticides (WARP) for simulating aquatic pesticide concentrations at the temporal resolution and national scales required for ESA assessment? Are there other watershed models that should be considered?
5. What is the desired and appropriate spatial scale for EECs for Bins 3 and 4? Specific PWC EECs were developed for HUC2 regions. Can or should the EECs for Bins 3 and 4 be at a finer spatial scale given a nationwide consultation?

Breakout Session AQUATIC 2: Evaluating Watershed Model Results:

In the Draft BEs, EPA employed an approach for flowing waters in an effort to approximate watershed processes. Regardless of the model employed, the EECs from any model need to be conservative (*i.e,* protective of the species of concern) and scientifically defensible in order to be used for risk assessment purposes. Typically, for EPA's use of PRZM5/VVWM as a field-scale model for vulnerable waters (*e.g.*, headwater streams), this would be done by comparing model outputs to field monitoring data (*i.e.*, edge of field runoff studies), where pesticide monitoring data is associated with pesticide-applications under well-described conditions (*i.e.*, application rates, field characteristics, water characteristics, and meteorological conditions). However, for watershed modeling, which aggregates exposure across a larger area, field-scale monitoring data, and the associated well-described conditions for all locations in the watershed, can be extremely difficult to obtain and, as a watershed model aggregates exposure, it may not be necessary.

Available literature documents have evaluated watershed models, including the NAS-recommended model SWAT, using general and targeted watershed monitoring data that is focused on known high pesticide-use areas, provided the data are collected at a high enough frequency to adequately capture the peak exposure concentration along with variations in concentration in the receiving stream. Unlike field monitoring data, general monitoring data (*i.e.*, sometimes described as ambient monitoring data) often lacks background information on application rates and field conditions and can be problematic when used for comparisons to model-generated EECs. They may, however, provide a lower bound for model-generated EECs. Targeted watershed monitoring (*e.g.*, studies at a watershed scale that are targeted to areas of known high pesticide use, with a sampling frequency targeted to the timing of use and subsequent runoff events) has been proposed as a means to provide more than a lower bound, especially when such monitoring spans multiple years and can be tied to factors that drive pesticide transport from field to water bodies. Such data are used to complement the results from modeling, not as a substitute for modeling.

In the Exposure chapter of the 2013 NAS report[1], the NAS noted that "If pesticides are to be used without jeopardizing the survival of listed species and their habitats, the estimated environmental concentrations (EECs) to which the organisms and their habitats will be exposed need to be determined. Chemical fate and transport models are the chief tools used to accomplish that task." (p. 49) The NAS further went on to describe a stepwise approach to fate and transport modeling, commenting on the use of various models such as AgDRIFT, PRZM, and EXAMS (p. 52-54). The NAS then cautioned that "in evaluating models, general monitoring data and field studies need to be distinguished. General monitoring studies provide information on pesticide concentrations in surface water or ground water on the basis of monitoring of specific locations at specific times. The monitoring reports, however, are not associated with specific applications of pesticides under well-described conditions, such as application rate, field characteristics, water characteristics, and meteorological conditions. General monitoring data cannot be used to estimate pesticide concentrations after a pesticide application or to evaluate the performance of fate and transport models." (p. 54) Though not as abundant as general monitoring data, field-scale monitoring studies are available for many pesticides, including the three OPs. However, monitoring data with this type of supporting information are generally lacking at the watershed scale.

[1] National Academy of Sciences. 2013. Assessing Risks to Endangered and Threatened Species from Pesticides. The National Academies Press. Washington, DC.

Additionally, the general monitoring data, specifically at the watershed scale, sometimes include data sets which are spatially and temporally targeted to varying degrees with pesticide applications. Lastly, the NAS noted that "pesticide fate and transport models do not provide information on the watershed scale; they are intended only to predict pesticide concentrations in bodies of water at the edge of a field on which a pesticide was applied." (p. 54) The NAS also noted that "different hydrodynamic models are required to predict how pesticide loadings immediately below a field are propagated through a watershed or how inputs from multiple fields (or multiple applications) aggregate throughout a watershed." The NAS report did not provide additional discussion on the monitoring data requirements (*e.g.,* metadata such as use rates, location, and timing) needed to evaluate watershed models.

Given the distinctions above between field-scale and watershed-scale models, the question arises "how does one evaluate the results generated from a watershed model?" EPA is proposing to use of the following multiple lines of evidence to evaluate the range of scientifically-defensible EECs for each flowing bin: consideration of available edge-of-field monitoring data and edge-of-field modeled estimates from PRZM5; incorporation of results from multiple watershed models, as appropriate; and consideration of statistical approaches to estimate confidence bounds around general monitoring data that were collected at a greater than a daily time step (*i.e.*, SEAWAVE Q and bias factors).

Charge Questions:

1. In what ways are a "multiple lines of evidence" approach appropriate for evaluating the results from a watershed model? What would be the "lines of evidence" and sources of information?
2. How can different types of monitoring data be distinguished? What metadata requirements (*e.g.,* use info, sample frequency, etc.) can be used to distinguish types of monitoring data?
3. What roles can the various types of monitoring data play in the evaluation of results from a watershed model (*e.g.*, general monitoring doesn't predict maximum but has other roles)?
4. What other approaches are available for evaluating results from watershed models?
5. To what extent can we rely on historical monitoring data when product labeling has changed and application-specific information is lacking?
6. Are there new or different types of monitoring that could be employed to further our understanding of aquatic modeling estimates?