Integration of Risk and Ecology into Water Quality Criteria for the Protection of Aquatic Life

By Rachel Workin

Dr. David Hinton, Advisor

Dr. Kevin Kwok, Advisor

May 2012

Masters project submitted in partial fulfillment of the

requirements for the Master of Environmental Management degree in

the Nicholas School of the Environment of

Duke University

2012

ABSTRACT

The demand for increased consideration of ecology has been an important driver of developments in ecotoxicology, a field that originally relied much more heavily on the principles of human toxicology. However, translating these developments into policy-level applications has been hindered by a lack of formalized acceptance and guidance. This study analyzes the potential for applying advances in ecotoxicology and risk assessment to the United States Environmental Protection Agency's process for developing water quality criteria for the protection of aquatic life. The current status of ecology in this process is assessed and suggestions for further integration of ecological principles are provided. By increasing consideration of ecology in the criteria development process, water quality criteria will be more realistic and achieve optimum protection of aquatic ecosystems.

TABLE OF CONTENTS

LIST OF ACRONYMS

ACR- Acute-Chronic Ratio

ALC- Aquatic Life Criteria

CCC- Criterion Chronic Concentration

CD- Criteria Development

CEC- Chemical of Emerging Concern

CMC- Criterion Maximum Concentration

CWA- Clean Water Act

EDC- Endocrine Disrupting Compound

ELS- Early Life Stage

EQS- Environmental Quality Standard

ERA- Ecological Risk Assessment

EtRA- Ecotoxicologial Risk Assessment

FAV- Final Acute Value

FRV- Final Residue Value

GMAV- Genus Mean Acute Value

NPDES- National Pollution Discharge Elimination System

SABS- Suspended and Bedded Sediments

SMAC- Species Mean Acute Value

SSD- Species Sensitivity Distribution

USEPA- United States Environmental Protection Agency

WoE- Weight of Evidence

WQC- Water Quality Criteria

WQS- Water Quality Standard

INTRODUCTION

The demand for increased consideration of ecology has been one of the most important drivers of developments in the fields of ecotoxicology and, consequently, ecological risk assessment (ERA) (Calow and Forbes 2003; Chapman 2002). From the initial form of standardized, single-species laboratory tests, ecotoxicology has evolved to include more complex testing conditions that better approximate real-world effects at ecologically-relevant levels of concern (Relyea and Hoverman 2006). Consequentially, a paradigm shift has occurred within ERAs, which are created based on the results of ecotoxicological studies, to better integrate this new type of data. (van den Brink 2008).

These developments in ecotoxicology enhanced the methods for performing ERAs into a process hereafter referred to as ecotoxicological risk assessment (EtRA). A natural outgrowth of these developments in ecotoxicology and EtRA is into the field of criteria development (CD, also known as criteria assessment), the regulatory process for determining an allowable concentration of a stressor in a given medium (Suter and Cormier 2008a). While there are multiple examples of criteria, this project will focus on the numerical water quality criteria (WQC) that are developed for certain stressors by the United States Environmental Protection Agency (USEPA) to protect aquatic life. These aquatic life criteria (ALC) are required under section 304 of the Clean Water Act (CWA) and then modified and implemented by the States and Tribes under section 303(c) as well as on a site-specific basis by environmental managers (USEPA 1994).

Risk in CD

Any criterion that allows for a concentration of a stressor above zero is inherently risk-based. Risk, which is the probability for harmful effects to an ecological system from an exposure to a stressor, is a function of both the exposure to and toxicity of the stressor (USEPA 2011b). The WQC discussed in this project are set at a concentration that achieves the goal of protecting aquatic life and their uses (Stephen et al. 1985) . Operationally, this is the concentration that, under most circumstances, is protective of 1) important species and 2) fish, benthic, and invertebrate assemblages, in which protection of an assemblage is defined as 95% of a group of diverse genera based on a sensitivity distribution. Since these WQC allow for some amount of the stressor in the aquatic system, potential for harm to the ecological system by the stressor (risk) must be evaluated.

Due to this shared evaluation of potential for harm in ERAs and CDs, there are similarities in their methodologies, especially as related to the exposure-response model (Suter and Cormier 2008a). An exposure-response model is typically created based on single-species laboratory tests that equate an ambient concentration of a stressor with endpoints regarding mortality, growth, and reproduction. In an ERA, an exposure-response model is used to estimate effects caused by a certain level of exposure of an ongoing or proposed source of a stressor. In a CD, a desired environmental goal is set, an effects level that constitutes the goal is characterized, and then an exposure-response model is used to determine the level of exposure from a stressor that maintains the chosen effects level (see Figure 1). Ideally, both incorporate uncertainty and multiple lines of evidence, as will be discussed later in this paper.

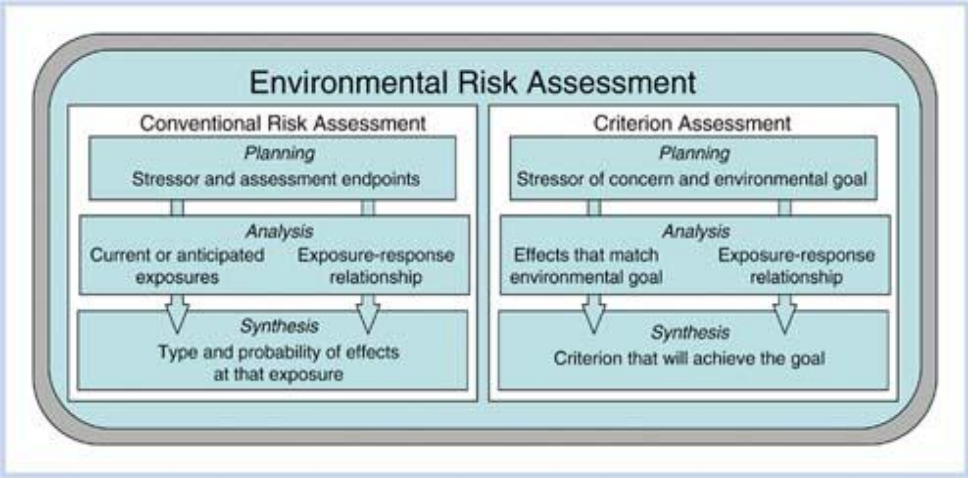

Figure 1. A comparison of frameworks for ERA and CD (Suter and Cormier 2008a).

In addition to the similarity in their derivations, ERA and CD are also linked in that one often informs or prompts another. For example, an ERA performed at a Superfund site may indicate that an ALC is needed. Or, an ERA conducted during the registration of a chemical may later be used to inform an ALC CD (USEPA 2006). Finally, an ALC can prompt a stressor-initiated risk assessment or be used to set a remediation goal.

Need for EtRA

The centrality of the exposure-response model to ERA and CD is a consequence of the outgrowth of ecotoxicology from the principles of human toxicology (Calow and Forbes 2003). The use of the exposure-response model is appropriate for conducting human health risk assessments, where the goal is the protection of individuals. However, while there is undeniably a place for single-species laboratory tests within ecotoxicology, there are deficiencies that arise

when this method is applied to ecological rather than human receptors. Four of the primary problems with the exposure-response model in the context of ERA will be discussed here:

1) Species Tested- While human risk assessments use laboratory data from multiple species to predict effects on one species, ERAs must extrapolate from a few species to many species with varying levels of sensitivity. Therefore, ERA must be highly aware of the ecological applicability of the test species as compared to the species the ERA is trying to protect (Calow and Forbes 2003; Breitholtz et al. 2006; Forbes et al. 2008; De Lange et al. 2010).

2) Spatial and Temporal Variability- The variability in aquatic systems and impacts of biotic and abiotic factors means that a constant dose applied in a laboratory system may not be representative of the exposure profile experienced in the wild. Furthermore, there may be inherent spatial and temporal variations within the ecological receptors (De Lange et al. 2010; Brock et al. 2008).

3) Level of Concern- The exposure-response model analyzes effects at the individual level, which is the focus of human toxicology. However, most ecologists agree that studying higher organizational levels, such as populations or communities, are more relevant to ecological health (Breitholtz et al. 2006; Dalkvist et al. 2009). From an ecological standpoint, the survival of individual organisms is not as important as overall impacts on genetically diverse populations. Furthermore, focus on the individual ignores the impact of indirect effects from stressors from altered ecological

interactions, such as predator-prey dynamics, or habitat modification (Rohr et al. 2006; Preston 2002). This can lead to an evaluation of inappropriate endpoints and omission of important effects. Finally, focusing on protecting individuals does not consider that communities are resilient and capable of absorbing certain amounts of stress (Chapman 2002).

4) Path of toxicity- The exposure-response model is best suited for evaluating chemical stressors with a direct route of exposure from the ambient environment and traditional, direct effects such as lethality or reproductive or growth impacts. However, there is an increasing realization of the importance of non-traditional effects such as endocrine disruption and behavior modification as well as non-traditional exposure routes such as through the diet. There is also an increasing realization of the impact of non-traditional stressors, such as sediments and warming that do not fit this model (Chapman 2002). This model is also hard to apply to evaluate mixtures of chemicals commonly present *in vivo*. (Wiegers et al. 1998).

Some of these deficiencies were solved during the initial iterations of ERA, whereas others were mentioned but could not be addressed due to a lack of the necessary formative ecotoxicological data (USEPA 1998). However, new methods, which will be discussed in further detail below, have been developed to further integrate the missing ecological principles- especially those from population, community, and stress ecology- into ERA (van den Brink 2008). While there is not yet an officially recognized framework for these methods, the shared intents will be represented in this project by the term EtRA. EtRA is not separate from ERA, but is characterized its use of ecotoxicological data involving ecologically relevant test species,

variable and realistic exposure conditions, population and community level analysis, and consideration of indirect paths of toxicity. The process of developing the EtRA paradigm began over 20 years ago with more improvements being put forth continuously (Cairns Jr. 1988). Yet even with a strong body of work supporting developments in ecotoxicology and EtRA, risk assessors and scientists lament that these improvements have been ignored in policy and management decisions (Suter and Cormier 2008b).

Need for Inclusion of EtRA in CD

As ERA has evolved to better integrate ecological theory, ALC CD is just beginning to acknowledge its intrinsic risk-based nature (Suter and Cormier 2008a). New ALC CDs are increasingly relying on the risk-based languages and principles of ERA. However, it is important that developments in ALC CD allow for and rely on EtRA and newer more environmentally relevant/realistic ecotoxicological data in order to produce the optimum criteria. This is necessary not only to maintain ease in linking the two processes and transparency in the CD process, but also to fulfill the Section 304 mandate that WQC reflect the latest environmental knowledge (Stephen et al. 1985). In 2003, the USEPA even released a draft strategy for proposed revisions to the aquatic life WQCs emphasizing a need for revisions that include the best-available, risk-based methods and procedures (USEPA 2003a) . However, while acknowledged, these revisions have not been developed.

PROJECT GOALS

The demand for more ecology in ecotoxicology has been crucial to the field and the development of ERA; however, there has not been a discussion of the possible opportunities for applying these advancements to ALC CDs. In this project, I aim to analyze the ALC CD process for the potential to integrate ecology. Due to the similarities between ERA and CD, integration of the ERA framework will also be evaluated for realizing the goal of increased ecology. Newer chemical specific ALC CDs will be compared to the original methodology to identify where integration has already occurred, continuing deficiencies will be identified, suggestions for incorporating more developments will be provided, and impacts for environmental managers will be discussed. This analysis is necessary for the USEPA to meet its goal of producing transparent and scientifically defensible CDs and to ensure that ecotoxicological and EtRA developments have actual policy implications. This analysis is also significant as many countries are reviewing their WQC and may look to the USEPA guidelines for direction (Leung 2011). Most importantly, this analysis aims to ensure that CD for aquatic life WQC evolves from a prescriptive, exercise of numbers into an ecologically relevant process that best meets environmental goals and ensures protection against all stressors.

BACKGROUND

Aquatic Life Water Quality Criteria (AL WQC)

WQC are a form of environmental quality standard (EQS) developed to protect endpoints such as human health, ecosystem integrity and commercial uses from pollution in the water; ALCs specifically protect aquatic life. Based on the assumption that there is a causal dose-response relationship between agents and effects, WQC serve as an economically beneficial alternative to a complete prohibition of chemical discharge.

International ALCs

Internationally, there are a variety of ALCs that, while differing in vocabulary, have similar goals. The European Commission applies the Water Framework Directive to achieve "good ecological status" (European Commission 2012). This framework undertakes a community-based approach to incorporate natural ecological and physical variability. Operationally, "good ecological status" allows for a slight deviation from a reference community and implies compliance with quality standards used through Europe (European Commission 2012). Japan relies on an EQS derived from ecotoxicological studies supporting "the conservation of the living environment" including fish and shellfish as well as the other invertebrates and algae which are their food (Yamakazi 2011). China is currently developing a water quality CD process utilizing the 95% sensitivity distribution of Stephen et al. (Wang et al. 2011).

There has been some momentum to establish international ALCs. For example, the United Nations Water's initiative to establish "international" water quality guidelines for aquatic ecosystems that can be applied globally based on the guidelines of Australia, New Zealand, Canada, South Africa, Unites States, and European Union(Yillia et al. 2011).

Aquatic WQC in America

Section 304(a) (1) of the CWA directs the USEPA to develop chemical-specific WQC for the protection of both human and aquatic life (USEPA 1994). These WQC are applied to all waters of the United States, which includes rivers, streams, seasonally flowing streams, lakes, natural ponds, wetlands, and marine waters, but does not include groundwater. These criteria serve as non-regulatory guidelines and can be adopted directly or modified by States and Tribes. It is the responsibility of the States and Tribes to set a water quality standard (WQS) for each water body based on its designated uses (ex. recreation, water supply, aquatic life, agriculture) and apply the relevant, legally-binding WQC. If a water body does not meet the applicable criteria for its designated use, it is placed on the Section 303(d) list of impaired water bodies that necessitates the development of a Total Maximum Daily Load (TMDL). The States and Tribes must review and solicit public comment on these WQSs and WQCs every three years.

WQC were initially developed for 65 compounds and families of compounds that were considered priority pollutants. Further chemical-specific numeric criteria for toxicants that were not priority toxic pollutants have been developed based on biological monitoring and whole effluent testing. Currently, approximately 150 chemical criteria have been developed, of which approximately 60 are ALC (USEPA 2012a). In addition to numeric criteria, narrative ALC such

as “no algal blooms” may also be issued; however, narrative criteria are not within the scope of this project and will not be discussed further.

Methodology of Stephen et al. 1985

Numeric ALC are currently developed using the methodology put forth in the 1985 document *Guidelines for deriving numerical national water quality criteria for the protection of aquatic organisms and their uses*, henceforth known as the *Guidelines* (Stephen et al. 1985). The *Guidelines* are a generic methodology that can then be applied to specific chemicals. While there some of the recent CDs have deviated from the *Guidelines* methodology (discussed below), this is the only official guidance document.

As mentioned above, the goals of this document are to develop criteria that protect 1) economically and recreationally important species and 2) fish and benthic invertebrate assemblages. There are typically separate criteria for marine and freshwater assemblages, although one may be used to inform the other (see Leung et al. 2001 for further information). In order to protect important species and assemblages from both acute and chronic exposure, a criterion maximum concentration (CMC) and a criterion chronic concentration (CCC) are developed. The CMC is a one-hour average maximum concentration that is derived from a sensitivity distribution of genus mean acute values (GMAV) of the geometric mean of the LC50s (concentration resulting in 50% mortality) and EC50s (50% effect level concentration) of different species. Based on this GMAV, a final acute value (FAV) is calculated by approximating the 95^{th} percentile. The CMC is then set as half of the FAV. The CCC is the maximum four-day average that can be calculated similarly to the CMC, or it can be based on an acute-chronic ratio (ACR). The CCC is then determined by dividing the FAV by the ACR.

These values can be checked alongside a Final Residue Value (FRV), meant to approximate bioconcentration, and a Final Plant Value (see Stephen et al. 1985 for more information). The CMC and CCC are assumed to be protective as long as they are not exceeded more than once every three years.

Besides providing a methodology, one of the most important functions of the *Guidelines* is promoting data quality guidelines. The *Guidelines* specify what species must be included in the sensitivity distribution and the appropriate testing conditions. This rigidity aims to eliminate uncertainty in the ALC so that it can be applied in a binary pass/fail situation. However, this inflexibility leads to ambiguity as to how to appropriately evaluate non-traditional tests that may still contain valuable information. Also, it can be argued that in attempting to erase uncertainty, the *Guidelines*, by avoiding even using the word "risk", ignore an important component of the CD methodology.

Ecological Risk Assessment

As described above, ERAs are used to quantify the potential for harmful effects to ecological endpoints following exposure to a stressor. ERAs can be retrospective or prospective and are stressor-initiated, source-initiated, or ecological value-initiated (USEPA 1998). The framework for an ERA is set forth in the USEPA's *Guidelines for Ecological Risk Assessment*, and can be broken down into three main steps (please see Figure 2).

1) The first step is ecological problem formulation- the process of generating the hypothesis as to why ecological effects have occurred or may occur. During this step, risk assessors first identify assessment endpoints that adequately reflect management goals and the

ecosystems they represent. Appropriate assessment endpoints are characterized by ecological relevance, susceptibility to known or potential stressors, and relevance to management goals. An example of an ecologically relevant assessment endpoint would be changes in community structure that could be quantified by a measurement endpoint such as some form of biotic index (USEPA 1998). The risk assessor then develops a conceptual model relating the stressor to the assessment endpoint. The conceptual model includes both exposure and effects and is important to evaluate primary, secondary, and tertiary endpoints based on factors such as species distribution and predatory-prey interactions. The conceptual model is also important as a reference for identifying holes in the available data. The final step during the problem formulation stage is to develop an analysis plan for going forward with the risk assessment.

2) The second step is an analysis phase in which exposure to and effects of the stressor are quantified. This is the step in which the exposure-response profile is characterized. This phase shares the most similarity to a CD.

3) The final step is risk characterization, in which the relationship between exposure and effects are examined along with ecosystem characteristics. Weight of evidence (WoE), the process of evaluating the strengths of multiple pieces of evidence for or against a hypothesis, is then used reach a conclusion regarding the risk of a stressor to the ecosystem being evaluated.

Throughout this whole process, uncertainty is being evaluated and emphasized. The process can also be iterative as missing information is identified (USEPA 1998).

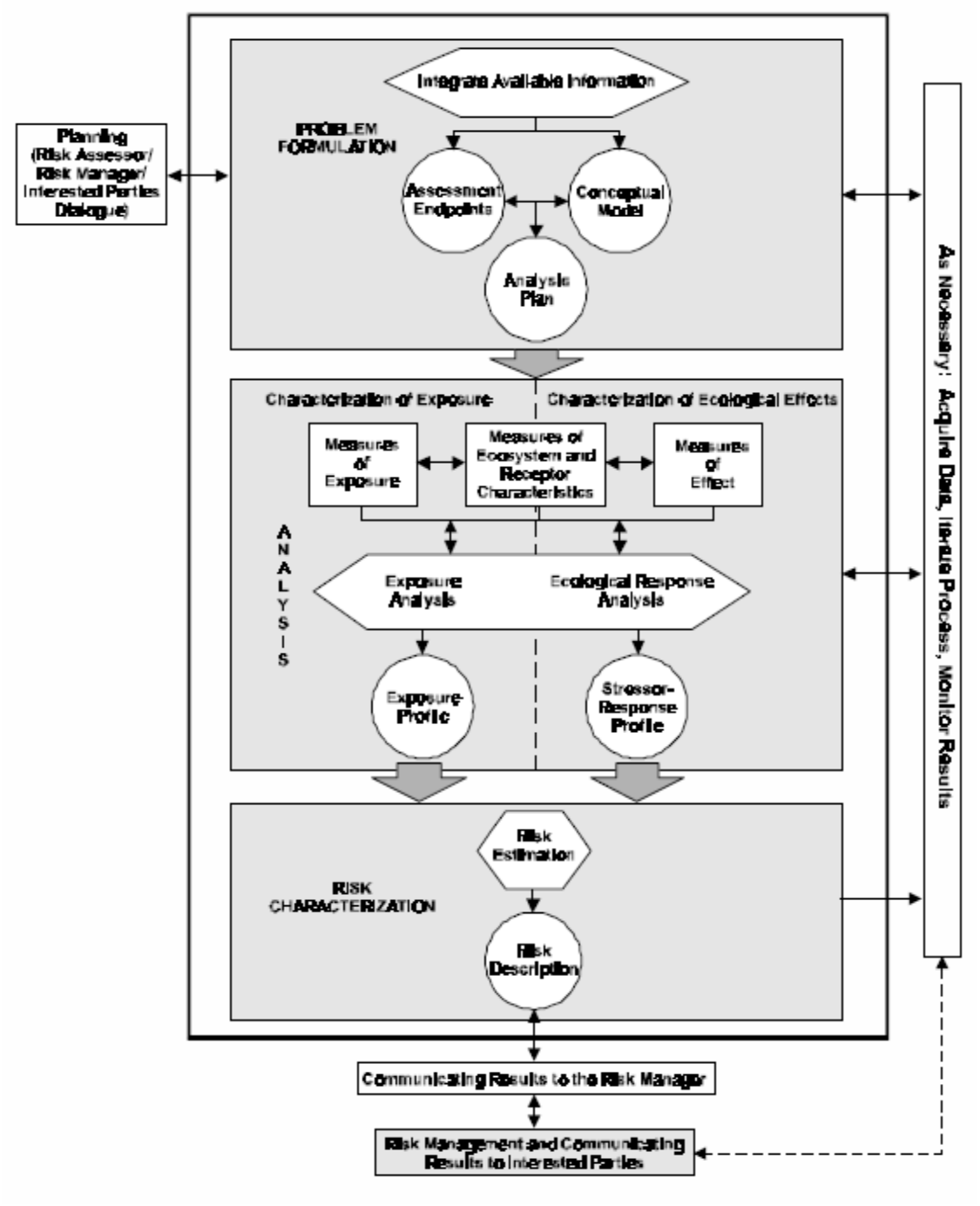

Figure 2. A graphical representation of an ERA framework based on the USEPA's Guidelines for Ecological Risk Assessment (USEPA 1998.)

Ecotoxicological Risk Assessment

The four-core problems from an ecological standpoint of an ERA methodology hinged on an exposure-response model were discussed in the introduction. While the *Guidelines for Ecological Risk Assessment* may have attempted to acknowledge these issues, the requisite ecotoxicological data was lacking because of a focus on testing a limited variety of species in unrealistic situations with an inability to extrapolate spatially, temporally, or from the individual level. ERA was reduced to a "book-keeping exercise," with criticism that it relied more on producing a number rather than evaluating ecological risk (van den Brink 2008).

Developments in ecotoxicology have allowed EtRA to respond to some of the most important paradigms of ecology as related to chemical disturbance: 1) it is important to preserve species composition in order to achieve the ultimate goal of protecting ecological function and 2) communities are resilient due to recovery potential and functional redundancy (Calow and Forbes 2003; Chapman 2002).

EtRA methodologies accomplish this with the use of more diverse and ecologically relevant species that account for site-specific variation in eco-region and keystone species. Furthermore, the increased diversity in data collection methods such as modeling, micro/mesocosms, and field studies has produced more complex data that will more likely reflect what is found in reality (Brix et al 2011). Now, it is possible to study the variability within aquatic systems due to abiotic and biotic factors and to analyze changes at the population and community level. For example, while exposure profiles are a component of ERA, EtRA-based modeling translates laboratory derived exposure profiles into different scenarios based on varying temporal and spatial dynamics (De Laender et al. 2005, Topping and Lagisz 2012). Models also allow for the

translation of individual data to higher level-effects based on extrapolation. Population-level modeling can be used to assess the impact of a stressor on species' spatial distribution and genetic diversity; community-level modeling can be used to assess impacts due to altered inter-species interactions such as changes in predatory-prey dynamics or competition (Grimm et al. 2009, Hanson and Stark 2011). These extrapolations can be performed quickly to evaluate different, varying scenarios in order to compile a range of risk scenarios under uncertainty.

Microcosms and mesocosms allow for the study of interactions of multiple species under more realistic abiotic conditions (Giddings et al. 2001). The fate of different stressors can be better evaluated than under constant laboratory conditions. Indirect effects such as behavioral changes and altered trophic dynamics can also be more easily studied. Field studies, quantified by measurement endpoints such as the index of biotic integrity (IBI), can look at abiotic influences and interactions even more realistically and encompass higher-order species (Yeom et al. 2009). In these more complex studies, resilience can also be more easily studied.

Advances in laboratory techniques should also be mentioned, including testing with more diverse species and evaluating non-traditional exposure routes and endpoints such as dietary exposure and endocrine disruption (Kefford et al. 2005, Tierney et al. 2005).

While it should be emphasized that there is no formal cohesive methodology or framework for EtRA, there is extensive literature documenting these developments. It should also be acknowledged that there are areas of ecology where EtRA is still lacking, especially in regards to quantifying community resilience to stressors and the impacts of mixtures of stressors. However, promising developments are being made to better characterize these processes (De Lange et al. 2010).

METHODS

In this project, available criteria and draft criteria guidelines for chemicals produced in the past 10 years and new draft guidelines for non-chemical stressors will be compared to the USEPA's original *Guidelines for Deriving Numerical National Water Quality Criteria for the Protection of Aquatic Organisms and Their Uses* by Stephen et al. 1985 and previously accepted chemical specific guidelines. Criteria produced in the last 10 years include tributyltin (2003), nonylphenol (2005), diazinon (2005), copper (2007), and acrolein (2009); draft criteria produced in the last 10 years include atrazine (2003), selenium (2004), ammonia (2009), and carbaryl (2011); new draft guidelines for non-chemical stressors that will be discussed include a white paper on chemicals of emerging concern (CECs), suspended and bedded sediments (SABS), and nutrients (USEPA 2012a). Guidelines produced in the last 10 years were chosen as they represent the most up to date methods used by the USEPA.

The comparison will focus on the increased use of specific components of developments in ecotoxicology (see Table 1). Due to the similarities between ERA and CD, application of the ERA and EtRA framework will also be evaluated. Areas in which the methods of EtRA could be further incorporated will be identified. Implications of these changes for implications to environmental managers and further suggestions will then be discussed.

Table 1. Components of ERA Framework and EtRA Methodologies that will be identified in draft AQL CDs

Component
Year
ERA Framework
Problem Formulation
Weight of Evidence
Uncertainty
EtRA Methodologies
Ecologically Relevant/Keystone Species
Range of Abiotic Factors
Micro/mesocosm data
Field Studies
Population Level Effects
Community Level/Indirect Effects
Non-traditional Exposure
Non-traditional Effects
Recovery/Long-term Effects

RESULTS AND DISCUSSION

ERA and EtRA in Guidelines for Aquatic Life Water Quality Criteria

The *Guidelines* were evaluated to determine how the risk framework had already been implemented and where the demands for greater consideration for ecology had already been met. Two of the main components were identified: consideration of site-specific abiotic impacts and use of data from a variety of species.

Since states and tribes may alter the criteria under the approval of the USEPA as part of section 303(c) of the CWA, there was already an accommodation for customizing the criteria to be more site-specific or protective (2002). Stephen et al. 1985 provided direction for modifying the ALC as a function of the site's water chemistry, such as for differing levels of hardness, pH, or salinity. This is important for approximating realistic exposure instead of solely relying on laboratory conditions.

The *Guidelines* also required that the FAV be compiled from species representing at least 8 different genera, and specified that the most sensitive life stage should be represented, although methods for identifying the most sensitive life stage were not provided. These requirements are important from an ecological standpoint in representing a range of sensitivities to the stressor and are a marked improvement over using a standard base set of algae, daphnia and brown trout data, as happened in the early stages of ecotoxicology (Forbes et al. 2008). The guidelines also specify that criteria can be modified to give greater consideration to commercially, recreationally, or other important species. While not explicitly stated, this could include modification for keystone or other ecologically important species.

However, there are still identifiable problems with this methodology from both a risk and ecological perspective. From a risk standpoint, there is no discussion of problem formulation, uncertainty, or WoE as the guidelines focus on a narrow set of "objective, internally consistent" tests. Ignoring uncertainties leads to gaps in this methodology, as there is a need for criteria for chemicals in which not much testing has been done, such as pesticides and personal care products (TenBrook 2010). This is especially important as a lack of chronic laboratory data often leads to the CCC being derived from an ACR. However, the species most sensitive to an acute exposure may not be the most sensitive to a chronic exposure (Forbes et al. 2008). Furthermore the ACR is inappropriate for many CECs that have very high acute toxicities (USEPA 2008). Uncertainty regarding this data is not discussed or quantified.

Ecological deficiencies arise due to the paradox of using lab data based on individual species to achieve the higher-level goals. Sensitivity distributions attempt to address this concern by including data from a range of species that approximate different parts of an ecological community. By representing the structure of a community, a sensitivity distribution is assumed to protect a community's function (Forbes and Calow 2002). Therefore, the sensitivity distribution approach appears suitable for the CWA goal of protecting assemblages. However, sensitivity distributions do not take into effect species interactions or the potential for trophic cascades, an important component of higher-order ecology (De Lange et al. 2010). Ignoring ecological interactions can result in sensitivity distributions that are not protective of an ecosystem, particularly for toxicants that directly target phytoplankton (De Laender et al. 2008). As an ecosystem is more than just a sum of its parts, a sensitivity distribution that evaluates each species in isolation may not be representative of reality. Furthermore, sensitivity distribution are

often not comprised of site-relevant species or realistic assemblages, which may influence their usefulness (Brix et al. 2005).

Furthermore, this methodology encourages a focus on ecotoxicity data from studies that only evaluate the endpoints of survival, immobility, growth, or reproduction (TenBrook et al. 2010). This renders the methodology ineffective in deriving criteria for stressors with non-traditional endpoints such as immunological suppression or endocrine disruption (Cormier et al. 2008). It is also not relevant to non-chemical stressors or that have multiple modes of action such as sediment. Indirect effects, such as behavioral changes and alterations of predator-prey relationships are also ignored. While dietary exposure routes can be evaluated in the *Guidelines* as a secondary exposure route in the final residue value, the methodology primarily excludes non-traditional exposure routes for ambient exposure.

The shared lack of ecology in the *Guidelines* and early ERA emphasizes the importance of applying developments of ecology in ecotoxicology to the field of CD.

ERA and EtRA in Chemical Specific Aquatic Life Water Quality Criteria

Since the 1985 guidelines, no official general update has been produced although a need for revision was acknowledged formally in 2003 (USEPA 2003d). Additional methods for modifying criteria to be more site-specific were provided in the *Water Quality Standards Handbook* by the USEPA in 1994. Procedures for criteria modification included in this handbook are the Recalculation Procedure, the Water-Effect Ratio procedure, and the Resident Species Procedure.

The Recalculation Procedure takes into account relevant differences between the sensitivities of the aquatic organism in the national dataset and the sensitivity of the organisms that occur at the site. The Water-Effect Ratio Procedure (previously called the Indicator Species Procedure) takes into account relevant differences between the toxicities of the chemical in laboratory dilution water and in the site water. Although, it should be noted that environmental managers must give additional consideration to confounding variables such as alkalinity, end-point variability, temporal and spatial variability, interspecies variability, and ensuring fish acclimation to ensure adequate protection when using this procedure (Welsh et al. 2009). Finally, the Resident Species Procedure intended to take into account both kinds of differences simultaneously (USEPA 1994).

Although no official updates to the 1985 guidelines have been offered, it is still possible to evaluate the chemical specific guidelines offered by the USEPA under section 304 of the CWA for incorporations of risk and the developments of ecology in eco-toxicology. In comparing the newest chemical-specific guidelines versus those of older chemical specific guidelines, there is a marked difference due to the greater inclusion of both ERA and EtRA. In this section, the new guidelines will be analyzed for components of ERA and EtRA principles to identify ways in which these developments can be incorporated into a generic guideline and other chemical-specific ALCs.

Criteria- final

Tributyltin

Tributyltin is a biocide that was commonly used on the hulls of ships (USEPA 2003c). The criteria was developed by integrating multiple lines of evidence including laboratory tests of

traditional endpoints of adverse effects as well as endocrine disruption and bioaccumulation capabilities observed in the field. The criteria also took the fate of Tributyltin into consideration. Microcosm and long term data were included in the "Other Data" section.

Nonylphenol

Nonylphenol is produced from cyclic intermediates in the refinement of petroleum and coal tar crudes, and is metabolized by the CYP450 enzyme (USEPA 2005a). While nonylphenol has been associated with estrogenic effects, data on estrogenicity were not used, because they did not meet data quality requirements. Instead, the criteria development focused on reproductive endpoints. A saltwater plant discussed in the "Other Data" section was more sensitive than the saltwater FAV; however no plants were included in the criteria calculation. The "Other Data" section included a mesocosm study, longer term tests, and the population growth of *Americamysis bahia.*

Diazinon

The ALC for diazinon, a broad spectrum organophosphate insecticide, was released in 2005 (USEPA 2005b). Certain components of developments in ecotoxicity were identified in the "Other Data" section including pulsed exposure, microcosm data, multi-generational tests, recovery, and non-traditional endpoints.

Copper

The copper ALC was updated in 2007 to incorporate new toxicity data and the biotic ligand model (USEPA 2007). Using the biotic ligand model increased the accuracy of the criteria, but resulted in increased uncertainty. However, the ability of the CD process to shift and

accommodate this diversity demonstrates that other forms of models could be incorporated into the CD process. While this revision focused on an increase in knowledge about the bioavailability, rather than the ecology, of copper, long-term exposure, microcosm and dosed stream data, and community and population level endpoints were included in the "Other Data" section.

Acrolein

The ALC for acrolein, a biocide used for aquatic weed control, was released in 2009 and followed the guidelines very closely (USEPA 2009a). While there was a discussion on the fate of acrolein and its mechanisms of toxicity, this information was not used to inform the CD. Furthermore, while acrolein targets aquatic plants, plant toxicity data was only included in the "Other Data" section and a final plant value was not calculated. This could leave unintended plant endpoints vulnerable and lead to ecosystem level effects.

Criteria- Draft

Atrazine

Atrazine, a commonly used herbicide, has been the subject of intense ecological concern and a wealth of research (Solomon et al. 2008). In 2003, the USEPA published a revised draft of their 2001 atrazine ALC (USEPA 2003b). The revised draft incorporated a large amount of the risk framework and EtRA methodologies, especially in regards to spatial distribution, recovery, and species interactions. Before introducing toxicity data, the guideline discussed the distribution of atrazine in the environment and within a system. This sort of background information was not present in early ALC CDs and is a necessary component of determining an exposure profile. The

exposure profile can then be used to direct the attention of the criteria developer. For example, by determining that atrazine is distributed at the micro-surface level, managers would want to focus on ensuring that the ALC was protective of top-dwelling plants rather than benthic organisms. The exposure profile for atrazine is also complemented by the information about atrazine's fate, persistence, and pulsed release. These characteristics are useful for analyzing recovery and could be used in the spatial and temporal models of EtRA.

In a strong contrast to earlier ALC CDs, the atrazine CD incorporated a large amount of ecologically relevant data to complement laboratory derived acute and chronic toxicity data. This includes identification of species of concern, discussion of atrazine impacts on habitat, community and population level endpoints, and recovery potential.

In a discussion of available acute and chronic toxicity data, studies with *Cladophora glomerata* were given particular attention as it was a fairly sensitive keystone species. Studies with amphibians, which were typically ignored in earlier ALC CDs, were also sought, although they were determined to be fairly insensitive to acute toxicity. The atrazine CD also included an entire section on ecosystem level effects. Community based assessment endpoints were evaluated based on the results from 77 micro and mesocosm studies. These assessment endpoints included changes in structure and density. From these studies, a range of exposure levels was determined that resulted in indirect effects such as feeding habitat loss and decreases in dissolved oxygen. These studies also occurred on a much longer timetable than recommended by Stephen et al. 1985, some examining application over a period of up to 323 days. These longer studies allowed for the determination longer-term effects as well as of recovery potential.

In order to protect aquatic life from chronic exposure to atrazine, the USEPA developed an entirely new set of methods to estimate community response than the *Guidelines.* This new

methodology determined concentrations of realistic and complex exposure profiles that resulted aquatic community structural changes based on modeled relationships between laboratory and micro/mesocosm data. The assessment endpoints of aquatic community structure change and function of primary producers were chosen based on the results of the 77 mesocosm and microcosm studies discussed within the ecological effects section. Plant growth rate and biomass was used as the measurement endpoint for the laboratory studies; reduction in primary producers and changes in the structure of primary producer communities was used as the measurement endpoint for the micro/mesocosm studies. These endpoints were the most ecologically relevant, because plants were the most sensitive receptor, and by protecting plants, indirect effects to higher order species such as loss of feeding habitat would also be avoided. The methods used in the atrazine CD to relate these measurement endpoints to the exposure profile will be briefly described here to demonstrate how they include some of the developments in ecotoxicology and EtRA methodologies.

From the micro/mesocosm studies, 77 results for effects/no effects for the identified measurement endpoints were reported including 24 from ponds or lakes, 20 from artificial streams, and 33 from microcosms. These 77 results were then quantified for severity of effect according to the methods of Brock et al. 2000. These effect scores were represented on a plot of the studies' exposure profile based on concentration and duration (see Figure 3).

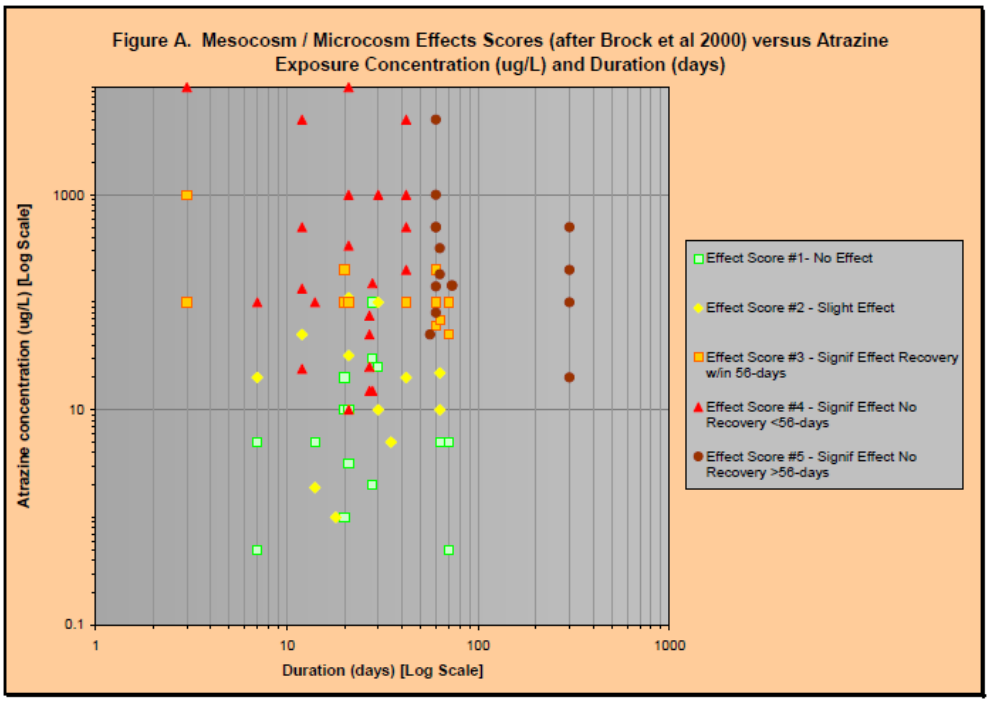

Figure 3. Process of determining concentration of community level effects in atrazine CD (USEPA 2003b).

To define an exposure profile that related to a specific level of concern, the ecological food chain model- The Comprehensive Aquatic Systems Model (CASM) was then used to estimate the exposure concentration for specific effect levels of concern based on the species sensitivity distribution (SSD) single species laboratory toxicity tests. The CASM model, which incorporates direct and indirect changes to community structure, incorporated the pulsed exposure conditions discussed in the CDs introduction to evaluate the % change in the Steinhaus similarity index of the aquatic community structure of primary producers. A similarity index value based on % deviation was then estimated between the micro/mesocosm test results and the model data that had the most similar exposure profiles.

The similarity index was then plotted against the Brock et al. 2000 scores to confirm that a 5% change in community structure based on the model was equal to no effect or slight effect in the

mesocosm study. The chronic freshwater criteria was set as the Average Primary Producer Steinhaus Similarity Deviation for a site less than 5% as determined using CASM or other appropriate models or indices. The report also included a considerable discussion of uncertainty throughout all stages of the CD.

This CD is exciting in its strong reliance on the risk framework and EtRA methodologies. By using laboratory and micro/mesocosm data to support each other, this CD produced a criterion that is much more ecologically defensible. Through the CAMS model, more realistic exposure profiles and recovery potential could be estimated. While not a straightforward number, this criterion is easily customized to site-specific scenarios and goals. Finally, discussion of uncertainty lets environmental managers identify where this process is vulnerable to being under or over protective. However, it should be noted that the method was difficult to follow and has not been formalized. Also, while this method included indirect effects, non-traditional effects that have been linked to atrazine, such as endocrine disruption, were not included. Based on the sensitivity of species to the potential endocrine disrupting properties of atrazine, this could lead to population level impacts.

Selenium

In 2004, the USEPA released a revised draft CD to update the selenium ALC of 1987 (USEPA 2004). Since 1987, considerable data has been put forth proving that diet is the primary route of selenium exposure. Therefore, this draft was modified to focus on the non-traditional exposure of diet for determination of chronic selenium toxicity. In addition to being a non-traditional exposure route, dietary exposure requires consideration of food chain transfer, a community-level process. Studies used to determine the final chronic value, many of which

occurred over a long time scale, had to include some form of dietary exposure. Among the studies was an evaluation of embryo larval deformities and embryo cranio-facial deformities, also caused by a non-traditional exposure route, from eggs collected in a field study. Studies in which exposure occurred for multiple generations were also considered, but not used due to their low sensitivity.

The chronic value put forth from this CD is unique in that it is expressed as a whole body tissue concentration of selenium on a dry weight basis. This tissue specific criterion helps to nullify the impact of site-specific factors that affect bioaccumulation such as chemical speciation and rate of transformation. This criterion can be translated back to a water concentration using a site-specific bioaccumulation factor or Bayesian Monte Carlo analysis (Toll et al. 2005).

In addition to including the dietary route of exposure, this CD used the EtRA principle of incorporating abiotic influences. Since temperature has such a strong effect on selenium concentration, the chronic value was lowered because of a result from a laboratory test simulating winter low temperatures. Also, while not included in the final CCC, recovery and long-term effects were mentioned within the suggestions for making site-specific modifications.

This CD is notable in the reliance on dietary exposure, use of diverse types of data, and production of a tissue specific guideline. However, there were still many places in which the EtRA methods could have been better used. Almost all of the data tested was on bluegill sunfish at the individual level, although population growth of algae was discussed in the bioconcentration section. While fish are some of the most sensitive species to selenium, this ignores food chain dynamics and is irrelevant to those sites without fish where birds are exposed to selenium via an invertebrate-based diet (Ohlendorf et al. 2011). This also ignores any sort of indirect effect. Also, while the introduction discussed how selenium was distributed in the

environment, this distribution or resulting exposure profiles were not included during criteria determination.

Ammonia

In 2009, the USEPA released a revised draft for ammonia to update the WQC put forth in 1999 (USEPA 2009b). The update occurred in order to include data for mussels because they are wide spread, commercially important, and sometimes threatened or endangered. Specifically, consideration was given to protecting the glochidia life stage in which the mussels are free-living and most sensitive. Protecting the most sensitive life stage is important from a population level. However, this CD was lacking in risk principles and EtRA methodologies. Outside of mussels, no consideration was given to species of concern. Furthermore, there was no discussion of the potential for community level effects, or use of micro or mesocosm data.

Carbaryl

A draft CD for carbaryl, a commonly used carbamate insecticide and molluscide was put forth for the first time in 2011 (USEPA 2011a). The CD drew upon an ecological risk assessment conducted by the USEPA's Office of Pesticide program in 2004 and was reviewed by the ecological risk assessment team in the Office of Water. Therefore, this CD incorporated multiple aspects of the risk assessment framework such as problem formulation and discussion of uncertainty. The CD used the problem formulation to provide a guiding strategic framework for focusing on the most relevant effects and properties. This problem formulation was then translated into a conceptual model that visualized different exposure routes and effects (see Figure 4).

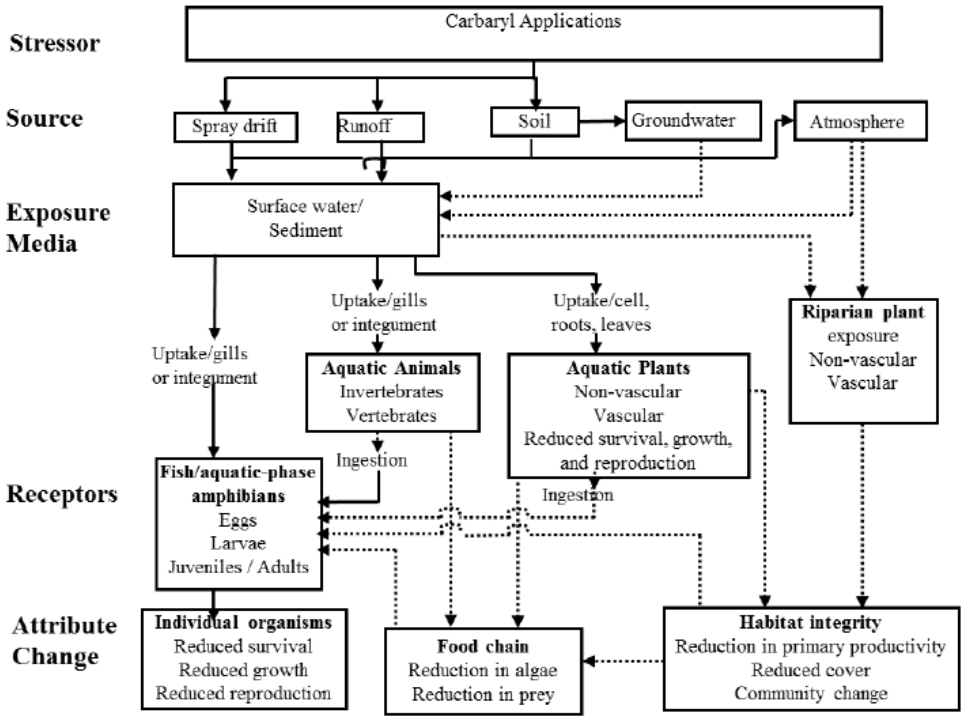

Figure 4. Conceptual model of carbaryl risk (USEPA 2011a)

The conceptual model is incredibly useful to a CD, because it identifies different routes of exposure and the attribute changes. This is important for determination of non-traditional routes of exposure and indirect effects. For example, reduction in primary productivity may lead to reduced feeding habitat. This alerts the criteria developer to look for tests measuring these endpoints. The conceptual model was then verified against field and mesocosm results. Field studies also identified the indirect effect of drift (the release, dispersal, and downstream displacement of insects) following carbaryl exposure. A mesocosm study found that high concentrations of carbaryl also led to delayed metamorphosis of *Rana clamitans* tadpoles. While the assessment endpoints used were still survival, growth, and reproduction from laboratory tests,

due to their lower sensitivity, at least there was confirmation that the WQC would be protective of other, non-traditional endpoints.

In addition to developing a conceptual model, this CD also outlined the use of carbaryl, the chemical's fate and transport, and its degradation products. This information is useful for determining a realistic exposure profile and for beginning to approximate recovery potential.

Criteria- Frameworks

Chemicals of Emerging Concern

The USEPA released a white paper for a draft guideline for determining AQLs for CECs, those contaminants, often personally care products and pharmaceuticals, which are not routinely monitored (USEPA 2008). These guidelines were developed to adapt the principles set forth in the *Guidelines* with respect to the common toxicological mechanism of endocrine disruption shared by many CECs. The white paper responded to particular deficiencies with the *Guidelines*, especially as related to calculating the CCC, rather than completely modifying the methodology (although problem formulation was recommended). These deficiencies include the use of the ACRs for EDCs, incorporation of sub-organismal effect data, and the length of life cycle being tested.

Many synthetic estrogens have high ACRs (up to greater than 30,000) that render calculation of acute toxicity irrelevant and introduce uncertainty when the ACR is used to calculate the CCC. As an alternative, the white paper suggests allowing biomarkers and other non-traditional endpoints as acceptable surrogate for traditional chronic endpoints. Biomarkers, a sub-organismal indicator effects such as an enzyme change, are easier to develop for EDCs than for other stressors since many interact with the hypothalamic /pituitary/gonadal axis. Currently,

tests that are not conducted on a whole organism are not acceptable for use in AQL CDs. Other non-traditional endpoints that could be included would be histopathological changes, biochemical measures such as vitellogenin production, gross morphology, or behavioral change. Since these endpoints do not often result in individual death, they are often ignored in aquatic life CDs. For example, histopathological changes (tissue lesions) were reported in the 1995 draft for the ammonia ALC, but were not considered in the CCC calculation. Even though these effects don't result in individual deaths, they could represent damage to ecological health or population level effects. For example, alteration of estrogen and androgen levels by CECs can result in changes in breeding behavior, such as nest-holding, which leads to reproductive and ultimately population level impacts (Weis et al. 2001). These non-traditional endpoints may also be more sensitive, as the concentration of tributyltin reported to cause imposex in bivalves was lower than the final chronic value calculated using the standard derivation procedure (USEPA 2003c).

The white paper also suggested that the guidelines for determining chronic toxicity be updated to include life-cycle tests to evaluate multi-generational effects and to remove early-life stage tests (ELS). They argued that CECs have effects on life period outside of the exposure period of the early-life stage tests currently allowed under the 1985 guidelines.

While this white paper was put forth to modify the guidelines for CECs, many of the suggested updates incorporated additional ecological information, especially the inclusion of non-traditional endpoints that may have population level effects and the suggestion for studying

Suspended and Bedded Sediments

In 2006, the USEPA produced the Framework for Developing Suspended and Bedded Sediments Water Quality Criteria (USEPA 2006). While many states already have a SABS ALC, this framework is one of the first applications of an ALC CD to a non-chemical stressor by the USEPA. The Framework draws upon many of the principles of ERA, likely because risk assessors were involved in composing the document. The Framework acknowledges that it is difficult to derive a SABS ALC because of both the difficulty in testing with SABS and the differences in SABS composition. Rather than recommending a single method of calculation, the Framework is a flexible, consistent step-wise process that allows for the integration of various types of information. In the Framework, alternative candidate criterion values are derived from multiple data sets, and then evaluated using WoE to select a final criterion (Cormier et al. 2008). Allowed data explicitly includes models, field observations, field manipulations, as well as laboratory tests (USEPA 2006). Direction for use of field data is provided. For example, field data should be grouped based on ecological similarity during the analysis phase to reduce the influence of natural variables. The ability to use field studies is important due to the potential for SABS to have cascading trophic level effects (Crowl 1989).

While the entire seven-step framework will not be described in detail here, notable uses of risk principles and EtRA methods will be acknowledged. To begin the CD process, the Framework recommends describing SABS effects on a specific site using a diagram or conceptual model, paying attention to both direct and indirect pathways. The reliance on conceptual models and multiple types of evidence is important for non-chemical stressors such as SABS that have complicated exposure and effect relationships. The model should then be used to set specific exposure and effect indicators, in which indicators are relevant to the biology

of the species. For example, smothering potential would be an appropriate indicator for benthic but not surface dwelling creatures. These exposure/response indicators should then be characterized based on the strength of data. Response indicators should be comparable across a region, for example, cold-water species should be grouped together. It may also be possible to tier the data. Finally, a transparent and justifiable method should be used to weight the data and select the ALC. It may also be possible to tier the ALC, depending on the management goals.

This framework emphasized substantial background gathering and consideration of site ecology, resulting in a promising, risk-based, ecologically relevant methodology. Although the methodology must be hashed out and tested, the SABS CD incorporates many of the developments in ecotoxicology.

Nutrients

There is currently no USEPA guideline for nutrient WQC even though nutrients are one of the most common ecological stressors to aquatic life (USEPA 2010). States have requested additional technical guidance from the USEPA in determining criteria for nutrients, which may result in a comprehensive guideline (USEPA 2003d). Currently, partial guidelines exist such as how to derive numerical nutrient criteria from exposure-response relationships (USEPA 2010). The stressor-response guideline is very similar to the analysis phase of ERA and includes conceptual modeling and WoE. In compiling data to analyze, the guideline advocates for models and field studies as nutrients often exert negative effects through intermediate steps that are hard to parse out in lab situations. Therefore, while not a complete guideline, it is possible to identify some of the EtRA methodologies that could be used to set a nutrient WQC.

Summary

There have been a number of improvements within the chemical specific ALC produced in the past 10 years as compared to older ALCs in regards to the incorporation of the ERA framework and improvements in ecotoxicology/EtRA methodologies (see Tables 2 and 3). This is especially true for chemicals supported by an ERA or data collected under the Federal Insecticide, Fungicide, and Rodenticide Act.

Table 2. Inclusion of ERA framework and developments in ecotoxicology in final chemical specific ALC.

Component	Tributyltin	Nonylphenol	Diazinon	Copper	Acrolein
Year	2003	2005	2005	2007	2009
ERA Framework					
Problem Formulation					
Weight of Evidence	X				
Uncertainty				X	
Developments in Ecotoxicology					
Ecologically Relevant/Keystone Species	X		X		
Range of Abiotic Factors			X	X	
Micro/mesocosm data	X	X	X	X	
Field Studies	X			X	
Population Level Effects		X	X	X	
Community Level/Indirect Effects				X	
Non-traditional Exposure			X		
Non-traditional Effects	X	X	X		X
Recovery/Long-term Effects	X	X	X	X	

Table 3. Inclusion of ERA framework and developments in ecotoxicology in draft chemical specific ALC.

Component	Atrazine	Selenium	SABS	Ammonia	Carbaryl
Year	2003	2004	2006	2009	2011
ERA Framework					
Problem Formulation					X
Weight of Evidence	X		X		
Uncertainty	X	X	X		X
Developments in Ecotoxicology					
Ecologically Relevant/Keystone Species	X	X	X	X	
Range of Abiotic Factors	X	X	X	X	
Micro/mesocosm data	X	X	X		X
Field Studies	X	X	X		X
Population Level Effects	X	X	X		
Community Level/Indirect Effects	X	X	X		X
Non-traditional Exposure					
Non-traditional Effects		X	X		X
Recovery/Long-term Effects	X	X	X		X

These improvements have enhanced the CD process and allowed for greater protection of ecological health. However, the irregularity and informality in which these improvements have been applied has significantly reduced the transparency of the ALC CD process. This makes it more difficult for state and tribes to adapt the ALC to a specific site or to apply the process to additional criteria. Furthermore, it means that ecological protection is not consistent, as shown in the extreme differences between the methodologies used in the atrazine and ammonia ALC CDs. This can lead to holes in ecological protection.

Since these improvements have not been formalized, there are no quality standards for the newer forms of data. This is especially true for field studies, which were explicitly rejected in the 1985 *Guidelines* as well as mesocosms and models. Therefore, this data, while informative, is relegated to the "Other Data" section, rather than being included in the CD. As a result, data produced under more up to date methods is ignored or unverified. Finally, with additional forms

of data, there is an increased reliance on WoE such as with the SABS framework and tributyltin CD (Cormier et al. 2008, USEPA 2003c). However, there has not been a consensus on an appropriate WoE methodology. Therefore, it is imperative the USEPA develop data quality standards for non-traditional testing including mesocosms, models, field tests, and evaluation of indirect effects as well as a methodology for using WoE. This is crucial to maintaining the technical rigor and transparency of the CD process.

Suggestions for Further Incorporation of EtRA and Developments in Ecotoxicology

Based on chemical-specific ALCs described above, all ALC CDs should include, at a minimum: a section of background information where exposure profiles are discussed; a conceptual model relating all possible stressor-effect relationships; a variety of data types in which community level endpoints and field data are either incorporated or verified against each other; and a transparent process for weighting evidence, if necessary.

Before recommending suggestions for further incorporation of EtRA methodology, it is important to consider what is desired from an ALC. The ALC must balance the need to evolve beyond a prescriptive derivation methodology, as agreed upon by the Office of Science and Technology and Office of Research and Development in 2003, while remaining easily tested as desired by decision makers (USEPA 2003d, Suter and Cormier 2008b). The ALC CD methodology must also be accomodating of the rise and fall of the presence of different stressors including non-traditional ones such as SABS and nutrients. To achieve the CWA goal of protection of different assemblages, the CD should consider the mechanism of action of the chemical, the natural variability in exposure and biota, and drivers of community change that

may impact how effects are distributed amongst and between assemblages. Therefore, I propose that the following proceedures be implemented to further incorporate developments in ecotoxicology and EtRA methodologies:

Regarding Species Diversity and Sensitivity Distributions

- Separate the sensitivity distributions to best approximate realistic communities such as lentic versus lotic or cold versus warm water dwelling
- Ensure that the sensitivity distribution is comprehensive of different trophic orders.

In a meeting on developing WQC for the protection of wildlife, it was noted that the genera requirements ignore data from amphibians and reptiles (USEPA 1994). This could be an important omission as tadpoles can be particularly sensitive and the commonly tested African clawed frog (though non-native) is more sensitive toward some compounds than daphnids, rainbow trout, fathead minnows or midges (Boyer and Grue 1995). If the compound is persistent or bioaccumulative, the sensitivity distribution should also include higher- order species such as mink.

- Include and highlight keystone species. Keystone species were only discussed in the atrazine CD, but from an ecological perspective should be given the same consideration as economically or recreationally important species.
- Guidelines should also be developed to determine if local species are needed to supplement the sensitivity distribution (Breitholtz et al. 2006). Trait-based ecology could be used to identify particularly sensitive species at a site (Baird and van den Brink 2007).

Regarding Development of Endpoints

- Identify all endpoints during the problem formulation stage; include non-traditional endpoints and community level effects.
- Study population level endpoints such as *r*, the intrinsic rate of population growth, in micro/mesocosm settings as a function of exposure.
- Develop multimetric indices such as the Index of Biotic Integrity or the Invertebrate Community Index that serve as measurement endpoints during field or mesocosm studies.

For priority toxic pollutants for which EPA has not issued section 304 criteria guidance, CWA 303 (c)(2)(b) requires that the States and Tribes adopt criteria based biological monitoring and assessment methods (USEPA 1994). These biocriteria describe the expected reference biological integrity of an aquatic community and are established as part of Congress' intent for State programs to expand beyond chemical-specific approaches. These biocriteria are used to determine if the site is meeting its intended uses and can provide conclusive evidence of nonattainment under the Independent Applicability Policy (Davies 1991).

However, it is not enough to only use biocriteria to look retrospectively at a site's ecological status when ALC are supposed to protect from damage occurring in the first place. Therefore, there is a need to:

- Create a diagnostic method of bioassessment that uses endpoints of macroinvertebrate structure and function as assessment endpoints and assemblage attributes as measurement attributes under field conditions. This data could then be used to calculate the ALC or to ensure that the ALC is protective of communities under realistic conditions.

King and Richardson 2003 used changepoint analysis to detect a threshold response for assemblage attributes to stressor exposure. This technique could also be used when stressor response relationships are nonlinear or heteroscedastic

Regarding Incorporation of multiple stressors

- Deploy approaches to incorporate effects of multiple stressors.

Rainbow trout and daphnids experienced nearly 100% mortality following exposure to combinations of the CMC of arsenic, cadmium, chromium, copper, mercury and lead (Spehar and Fiandt 2009). While EtRA methodologies cannot fully solve this problem, multiple factor microcosm and mesocosm, as well as field based studies can be a starting point for evaluation. For example, a multiple factor microcosm study was used to determine synergistic effects between low oxygen and carbaryl that led to smaller sized daphnids that were more susceptible to predation by *Chaoborus* (Schulz and Dabrowski 2001). While it may not be realistic to incorporate this data into a CD unless the presence of multiple stressors is known, mixture combinations can be tested on a site-specific basis to ensure protection. Mixture models are also being developed that may facilitate this process (Rider and LeBlanc 2005)

Regarding Recovery

- Incorporate recovery potential into CDs as well as on a site-specific basis.

Recovery potentials are useful for addressing the ecological principle that ecosystems are resilient. Recovery was considered in the atrazine CD, but could be used for other chemicals in which chemical fate data is known. Care would have to be given, however, to ensure protection before assuming biota-mediated recovery factors such as acclimation. Incorporating recovery into ALC would parallel one of the newest developments in EtRA, which is a shift to thinking of risk in terms of vulnerability: the function of exposure, effect, and recovery at a site (De Lange et al. 2010). By incorporating recovery potentials, the ALC is less likely to be overprotective.

Regarding Modeling

- Use predictive modeling to approximate different exposure profiles. Examples include: exposure to non-human fecal matter in the 2012 draft guidelines for recreational water criteria (USEPA 2012b). Models used in ALC CD can examine effects of abiotic conditions such as temperature, i.e. for selenium.
- Use models to better relate micro and mesocosm data to various community endpoints simultaneously (van den Brink 2008).
- Utilizing models in geospatial analysis software achieve greater insight on chemical-spatial dynamics in a specific exposure area. (Lahr et al. 2010).

Regarding Uncertainty

There was no discussion of uncertainty in the *Guidelines*. However, as risk principles are increasingly incorporated into CD and multiple forms of data are used, evaluation of uncertainty becomes more important. The 1998 *Guidelines for Ecological Risk Assessment* provide some discussion of uncertainty that may be useful to CD until the different processes become formalized. The more explicitly uncertainty is articulated, the easier it will be for environmental managers to apply the criteria to a specific site.

Implications for Environmental Managers

The current shift in the ALC CD process and the potential for further incorporation of EtRA methodologies is important for environmental managers, including the state and tribal leaders setting WQC as well as those in the private sector who wish to apply WQC to specific water bodies. The current trend indicates that until new guidelines are produced, the CD process may become less transparent. This could have negative consequences should an environmental manager decide to modify the USEPA's criteria.

The 1994 *Water Quality Standards Handbook* provides some guidance for determining if environmental managers should modify the USEPA's ALCs. Reasons for modifying the ALCs include sensitive species or a different goal for protection. For example, the Central Valley Regional Water Quality Control Board's objective is to maintain its waters free of "toxic substances in concentrations that produce detrimental physiological responses in plant, animal, or aquatic life", and the state of North Carolina has a trout water standard to protect the propagation and survival of stocked trout on a year round basis (TenBrook et al. 2010, North Carolina Department of Environment and Natural Resources 2011). If an environmental manager

is concerned that the ALC is not protective of their sensitive species, they could perform a risk assessment to evaluate the original ALC. The Risk Assessment Technical Forum promotes USEPA-wide consensus on controversial risk assessment issues and could be an asset during this process (USEPA 2011c).

In addition to modifying the criteria, the environmental manager could also alter the period of allowed exceedences or the averaging time of the CCC and CMC. An environmental manager may choose to change this depending on vulnerability and other site specific factors.

Regardless of the environmental managers decision to alter the criteria, it is important to implement a monitoring program and potentially look for biomarkers (Calow and Forbes 2003). This follow-up is crucial for determining the efficacy of the ALC and to produce an outcome assessment.

CONCLUSION

The changing landscape of ecotoxicology has made for a more complex and scientifically rich field. However, to translate these developments into actual policy, great effort is needed. In this project, I analyzed the current status of ecotoxicology in the aquatic life water quality criteria put forth by the USEPA. I found that while there have been improvements over the original *Guidelines,* they have been applied inconsistently, leading to decreased transparency. Many of the same limitations exist for integrating developments in ecotoxicology into ALCs as ERA; however, I believe that the formalization and standardization of the CD process makes it the best route to concentrate efforts. Therefore, it is imperative that the USEPA release new guidelines and data quality standards to encompass developments in ecotoxicology so that all relevant data and methodologies can be utilized in the CD process. The atrazine ALC CD and the SABS framework could serve as partial examples.

In the meantime, environmental managers should ensure that ALC are protective of their site by 1) developing a conceptual model to identify all potential routes of exposure and effects, 2) identify gaps in the ALC and complement with additional species that represent the site, 3) verify the ALC with a field study from a similar site, and 4) implement a monitoring program.

Advances in WQC for the protection of aquatic life will go beyond ensuring the best possible risk-based criteria for our waterways. What is learned here can also be used to improve other types of EQS, such as soil screening levels, and to better inform processes such as EtRA. By redesigning the CD process to be a collaborative effort of ecology and toxicology, ecotoxicology can best realize its intended goals.

WORKS CITED

[Anonymous]. 2002. Federal Water Pollution Control Act. In: 33 USC 1251 et seq. United States of America: http://www.epa.gov/lawsregs/laws/cwa.html, 234 pp.

Baird DJ and van den Brink P. 2007. Using biological traits to predict species sensitivity to toxic substances. Ecotoxicology and Environmental Safety. 67(2): 296-301.

Boyer R, Grue CE. 1995. The need for water quality for frogs. Environmental Health Perspectives 103(4): 352-357.

Breitholtz M, Ruden C, Hansson SO, Bent-Erik B. 2006. Ten challenges for improved ecotoxicological testing in environmental risk assessment. Ecotoxicology and Environmental Safety 63: 324-336.

Brix KV, DeForest DK, Burger M, Adams WJ. 2005. Assessing the relative sensitivity of aquatic organisms to divalent metals and their level of representation in toxicity data sets compared to natural aquatic communities. Human and Ecological Risk Assessment.11: 1139-1156.

Brix KV, De Forest DK, Adams WJ. 2011. The sensitivity of aquatic insects to divalent metals: a comparative analysis of laboratory and field data. Science of the Total Environment 409:4187-4197.

Brock TCM, Lahr J, van den Brink PJ. 2000. Ecological risks of pesticides in freshwater ecosystems. Part 1: Herbicides. Wageningen, Alterra, Green World Research. Alterra-Rapport 088. 124 pp.

Brock TCM, Maltby L, Hickey CW, Chapman J, Solomon KR. 2008. Spatial extrapolation in ecological effect assessment of chemical. In: Extrapolation practice for ecotoxicological effect characterization of chemicals, (Solomon KR, Brock TCM, De Zwart D, Dyer SD, Posthuma L, Richards S, et al., eds):SETAC and CRC Press, 224-256.

Cairns Jr. JH. 1988. Putting the eco in ecotoxicology. Regulatory Toxicology and Pharmacology 8: 226-238.

Calow P, Forbes VE. 2003. Does ecotoxicology inform ecological risk assessment? Environmental Science and Technology 37(7): 146A-151A.

Chapman PM. 2002. Integrating toxicology and ecology: putting the "eco" into ecotoxicology. Marine Pollution Bulletin 44(1): 7-15.

Cormier SM, Paul JF, Spehar RL, Saw-Allen P, Berry WJ, Suter GW. 2008. Using field data and weight of evidence to develop water quality criteria. Integrated Environmental Assessment and Management 4(4): 490-504.

Crowl TA. 1989. Effect of crayfish size, orientation, and movement on the reactive distance of largemouth bass foraging in clear and turbid water. Hydrobiologia 183: 133-140.

Dalkvist T, Topping CJ, Forbes VE. 2009. Population level impacts of pesticide induced chronic effects on individuals depend more on ecology than toxicology. Ecotoxicology and Environmental Safety 72: 1663-1672.

Davies TT. 1991. Transmittal of final policy on biological assessment and criteria. (I-X WMDDR, ed). Washington D.C., 18 pp.

De Laender, FD, De Schamphelaere KAC, Verdonck FAM, Heijerick DG, Van Sprang PA, Vanrolleghem PA, Janssen CR. 2005. Simulation of spatial and temporal variability of chronic copper toxicity to *Daphnia magna* and *Pseudokirchneriella subcapitata* in Swedish and British surface waters. Human and Ecological Risk Assessment 11: 1177-1191.

De Laender FD, De Schamphelaere AC, Vanrolleghem PA, Janssen CR. 2008. Do we have to incorporate ecological interaction in the sensitivity assessment of ecosystems? An examination of a theoretical assumption underlying species sensitivity distribution models. Environnent International 34:390-396.

De Lange HJ, Sala S, M. V, Faber JH. 2010. Ecological vulnerability in risk assessment- a review and perspectives. Science of the Total Environment 408: 3871-3879.

European Commission. 2012. Common implementation strategy for the water framework directive. Available: http://ec.europa.eu/environment/water/water-framework/objectives/status_en.htm [accessed February 3 2012].

Forbes VE, Calow P. 2002. Species sensitivity distributions revisited: a critical appraisal. Human and Ecological Risk Assessment 8(3):473-492.

Forbes VE, Calow P, Sibley RM. 2008. The extrapolation problem and how population modeling can help. Environmental Toxicology and Chemistry 27(10): 987-994.

Giddings JM, Solomon KR, Maund SJ. 2001. Probabilistic risk assessment of cotton pyrethroids: II aquatic mesocosm and field studies. Environmental Toxicology and Chemistry 20(3): 660-668.

Grimm V, Ashauer R, Forbes VE, Hommen U, Preuss TG, Schmidt A, van den Brink P, Wogram J,Thorbek P. 2009. CREAM: A European project on mechanistic effect models for ecological risk assessment of chemicals". Valery Forbes Publications. Paper 16. http://digitalcommons.unl.edu/biosciforbes/16.

Hanson N, Stark JK. 2011. Extrapolation from individual-level responses to population growth rate using population modeling 17(6):1332-1347.

Hickey GL, Kefford BJ, Dunlop JE, Craig PS. 2008. Making species salinity sensitivity distributions reflective of naturally occurring communities: using rapid testing and Bayesian statistics. Environmental Toxicology and Chemistry. 27:2403-2411.

Kefford BJ, Palmer CG, Jooste, S, Warne, MS, Nugegoda D. 2005. What is meant by"95% of species"? An argument for the inclusion of rapid tolerance testing. Human and Ecological Risk Assessment 11: 1025-1046.

Kefford BJ, Nugegoda D, Metzaling L, Fields EJ. 2006. Validating species sensitivity distributions using salinity tolerance of riverine macroinvertebrates in the southern Murray Darling Basin (Victoria, Australia). Canadian Journal of Fisheries and Aquatic Science. 63:1865-1877.

King RS, and Richardson CJ. 2003. Integrating bioassessment and ecological risk assessment: An approach to developing numerical water quality criteria. Environmental Management. 31 (6):795-809.

Lahr J, Munier B, De Lange HJ, Faber JF, Sorensen PB. 2010. Wildlife vulnerability and risk maps for combined pollutants. Science of the Total Environment. 408(18):3891-3898.

Leung KMY, Morritt D, Wheeler J, Whitehouse P, Sorokin N, Toy R., Holt M, Crane M. 2001. Can saltwater toxicity be predicted from freshwater data? Marine Pollution Bulletin 42: 1007-1013.

Leung KMY. 2011. Preface to programme. In: International Conference on Deriving Environmental Quality. International conference on Deriving Environmental Quality Standards for the Protection of Aquatic Ecosystems. IV-VII.

North Carolina Department of Environment and Natural Resources. 2011. Surface water classifications. Available: http://portal.ncdenr.org/web/wq/ps/csu/classifications#classes. [Accessed January 19 2012].

Ohlendorf HM, Covington SM, Byron ER, Arenal CA. 2011. Conducting site-specific assessments of selenium bioaccumulation in aquatic systems. Integrated Environmental Assessment and Management. 7(3):314-324.

Preston BL. 2002. Indirect effects in aquatic ecotoxicology: implications for ecological risk assessment. Environmental Management 29(3): 311-323.

Relyea R, Hoverman J. 2006. Assessing the ecology in ecotoxicology: a review and synthesis in freshwater systems. Ecology Letters 9(10): 1157-1171.

Rider C, LeBlanc G. 2005. An integrated addition and interaction model for assessing toxicity of chemical mixtures. Toxicological Science 87(520:528): 520-528.

Rohr JR, Kerby JL, Sih A. 2006. Community ecology as a framework for predicting contaminant effects. Trends in Ecology and Evolution 21(11): 606-613.

San Joaquin River Restoration Program. 2010. Exhibit B: Water quality criteria. Fisheries Management Plan: A Framework for Adaptive Management in the San Joaquin River Restoration Program. EPA/822/B-00-002.USEPA.

Schulz R and Dabrowski JM. 2001. Combined effects of predatory fish and sublethal pesticide contamination on the behavior and mortality of mayfly nymphs. Environmental Toxicology and Chemistry. 20(11): 2537-2543.

Solomon KR, Carr JA, Du Preez LH, Giesy JP, Kendall RJ, Smith EE, Van Der Draak, GJ. 2008. Effects of atrazine on fish, amphibians, and aquatic reptiles: A critical review. Critical Reviews in Toxicology. 38(9):721-772.

Spehar RL, Fiandt JT. 2009. Acute and chronic effect of water quality criteria-based metal mixtures of three aquatic species. Environmental Toxicology and Chemistry 5(10): 917-931.

Stephen CE, Mount DI, Hansen DJ, Gentile JR, Chapman GA, Brungs WA. 1985. Guidelines for deriving numerical national water quality criteria for the protection of aquatic organisms and their uses. PB85-227049. 59 pp.

Suter GW, Cormier SM. 2008a. What is meant by risk-based environmental quality criteria? Integrated Environmental Assessment and Management 4(4): 486-489.

Suter GW, Cormier SM. 2008b. Revitalizing environmental assessment. Integrated Environmental Assessment and Management 4(4): 385.

TenBrook PL, Palumbo AJ, Fojust TL, Hann P, Karkoski J, Tjeerdema RS. 2010. The University of California-Davis Methodology for deriving aquatic life pesticide water quality criteria. In: Reviews of Environmental Contamination and Toxicology, Vol. 209, (Whitacre DM, ed). New York: Springer, 1-155.

Tierney KB, Baldwin DH, Hara TJ, Ross PS, Scholz NL, Kennedy CJ. 2010. Olfactory toxicity in fishes. Aquatic Toxicology. 96(1):2-26.

Toll JE, Tear Lucinda LM, DeForest DK, Brix KV, Adams WJ. 2005. Setting site-specific water-quality standards by using tissue residue criteria and bioaccumulation data. Part 1. Methodology. Environmental Toxicology and Chemistry. 24(1): 224-230.

Topping CJ, Lagisz M. 2012. Spatial dynamic factors affecting population-level risk assessment for a terrestrial arthropod: An agent-based modeling approach. Human and Ecological Risk Assessment: An International Journal 18(1):168-180.

USEPA. 1994. Water Quality Standards Handbook. Second Edition. EPA 823-B-94-005a. http://water.epa.gov/scitech/swguidance/standards/handbook/index.cfm. [Accessed January 20 2012].

USEPA. 1998. Guidelines for Ecological Risk Assessment. EPA/630/R-95/002. 188 pp.

USEPA. 2003a. Draft Strategy: Proposed revisions to the “Guidelines for deriving numerical and national water quality criteria for the protection of aquatic organisms and their uses". 5 pp.

USEPA. 2003b. Ambient aquatic life water quality criteria for atrazine- revised draft. EPA-822-R-03-023. 178 pp.

USEPA. 2003c. Ambient aquatic life water quality criteria for tributyltin (TBT)-final. EPA 822-R-03-031. 175 pp.

USEPA. 2003d. Strategy for water quality standards and criteria. EPA-823-R-03-010.48 pp.

USEPA. 2004. Draft aquatic life water quality criteria for selenium- 2004. EPA-822-D-04-001. 90 pp.

USEPA. 2005a. Aquatic life ambient water quality criteria- nonylphenol. EPA-822-R-05-005. 96 pp.

USEPA. 2005b. Aquatic life ambient water quality criteria- diazinon. EPA-822-R-05-006. 85 pp.

USEPA. 2006. Framework for developing suspended and bedded sediments (SABS) water quality criteria EPA-822-4-06-001. 169 pp.

USEPA. 2007. Aquatic life copper criteria. EPA-822-R-07-001. 204 pp.

USEPA. 2008. White Paper: Aquatic life criteria for contaminants of emerging concern (part 1). 86 pp.

USEPA. 2009a. Ambient aquatic life water quality criteria for acrolein. http://water.epa.gov/scitech/swguidance/standards/criteria/aqlife/pollutants/acrolein/upload/aq-final.pdf. Accessed 2/10/12.

USEPA. 2009b. Draft 2009 update aquatic life ambient water quality criteria for ammonia-freshwater. USEPA-822-D-09-001. 192 pp.

USEPA. 2010. Using Stressor-response Relationships to Derive Numeric Nutrient Criteria. EPA-820-S-10-001. 93 pp.

USEPA. 2011a. Draft aquatic life ambient water quality criteria for carbaryl. EPA-820-D-11-0001.157 pp.

USEPA. 2011b. EPA risk assessment: basic information. Available: http://epa.gov/riskassessment/basicinformation.htm#risk [accessed January 18 2011].

USEPA. 2011c. Risk Assessment Forum. Available: http://www.epa.gov/raf/index.htm [accessed February 12 2012].

USEPA. 2012a. Aquatic Life Criteria. Available: http://water.epa.gov/scitech/swguidance/standards/criteria/aqlife/index.cfm [accessed January 20 2012].

USEPA. 2012b. Recreational water quality criteria- Draft. 820-D-11-002. 73 pp.

van den Brink PJ. 2008. Ecological risk assessment: From bookkeeping to chemical stress ecology. Environmental Science and Technology 42: 8999-9004.

Wang J-Y, Mu J-L, Wang Y. Method development for scientific derivation of marine water quality criteria in China. In: Proceedings of the International Conference on Deriving Environmental Quality. 42.

Weis JS, Smith G, Zhou T, Santiago-Bass C, Weis P. 2001. Effects of contaminants on behavior: biochemical mechanisms and ecological consequences. Bioscience 51(3): 209-217.

Welsh PG, Lipton J, Chapman GA. 2009. Evaluation of water-effect ratio methodology for establishing site-specific water quality criteria. Environmental Toxicology and Chemistry 19(6): 1616-1623.

Wiegers JK, Feder HM, Mortensen LS, Shaw DG, Wilson VJ, Landis WG. 1998. A regional multiple-stressor rank-based ecological risk assessment for the fjord of Port Valdez, Alaska. Human and Ecological Risk Assessment: An International Journal 4(5): 1125-1173.

Yamakazi K. 2011. Regulatory Standards for Conservation of Aquatic Life in Japan. In: International conference on Deriving Environmental Quality Standards for the Protection of Aquatic Ecosystems. Hong Kong. 30.

Yeom DH, Chung KY, Kim YH, Adams SM. 2009. Ecological health and causal assessment of fish communities experiencing multiple stressors in Gap Stream, South Korea. Toxicology and Environmental Health Sciences. 1(2):97-108.

Yillia PT, Kreuzinger N, Zessner M, Krois H. 2011. Blueprint for developing international water quality guidelines for aquatic ecosystems: output of a UN-Water scoping study. In: International conference on Deriving Environmental Quality Standards for the Protection of Aquatic Ecosystems. Hong Kong. 74.