EPA/600/R-06/038
ERASC-006
April 2006

ASSESSING RISKS TO POPULATIONS AT SUPERFUND AND RCRA SITES CHARACTERIZING EFFECTS ON POPULATIONS

by

Wayne R. Munns, Jr. and Matthew G. Mitro
U.S. Environmental Protection Agency
Office of Research and Development
National Health and Environmental Effects Research laboratory
Atlantic Ecology Division
Narragansett, RI 02882

Ecological Risk Assessment Support Center
Office of Research and Development
U.S. Environmental Protection Agency
Cincinnati, OH

AUTHORS, CONTRIBUTORS AND REVIEWERS

AUTHORS

Wayne R. Munns, Jr.
U.S. Environmental Protection Agency
Office of Research and Development
National Health and Environmental Effects Research Laboratory
Atlantic Ecology Division
Narragansett, RI 02882

Matthew G. Mitro
U.S. Environmental Protection Agency
Office of Research and Development
National Health and Environmental Effects Research Laboratory
Atlantic Ecology Division
Narragansett, RI 02882
(currently employed at Wisconsin Department of Natural Resources, Monona, Wisconsin)

REVIEWERS

Wayne G. Landis
Western Washington University
Institute of Environmental Toxicology and Chemistry
516 High Street
Bellingham, WA 98225-9180

Mary T. Sorensen
ENVIRON International Corporation
1600 Parkwood Circle, Suite 310
Atlanta, GA 30339

ACKNOWLEDGMENTS

Programmatic review of the document was conducted by Susan Roddy, EPA Region 6; Gina Ferreira, EPA Region 2; Dale Haroski (currently, Office of the Administrator); and Dale Hoff, EPA Region 8. Glenn Suter of EPA's National Center for Environmental Assessment, ORD, kindly provided text describing development of the assessment population concept. Technical review of the previous draft of this document was conducted by Rick Bennett, Gina Ferreira, Tim Gleason, Anne Kuhn, Joe LiVolsi, William “Skip” Nelson, Nathan Schumaker, Charlie Strobel, Glenn Suter, Sharon Thoms, and Steve Wharton.

TABLE OF CONTENTS

Page

TABLE OF CONTENTS (cont.)

Page

GLOSSARY OF KEY TERMS

assessment population	a group of conspecific organisms occupying a defined area which has been selected to serve as an assessment endpoint entity for an ecological risk assessment
bottleneck effect	a reduction in genetic heterogeneity within a population as a result of stressor-induced mortality
carrying capacity	the maximum abundance of a biological population that is sustainable by a habitat or environment
cohort	a group of similarly-aged members of a population
compensation	a feedback between the density of a population and some biological property of that population (typically demographic rates) (synonymous with *density dependence*)
compensatory mechanism	a biological mechanism, such as homeostatic acclimation of individuals, genetic adaptation, and density dependence in vital rates and migration, that can ameliorate adverse effects over the short or long term
critical abundance	the specific population density or abundance below which adverse population-level effects are known or suspected to occur
demographic rates	age- or stage-specific birth and death rates of individuals within the population (synonymous with *vital rates*)
density dependence	a feedback between the density of a population and some biological property of that population (typically demographic rates) (synonymous with *compensation*)
emergent property	a property of a system that cannot be predicted simply by understanding processes occurring at lower levels of organization in that system
equilibrium abundance	the abundance of a population at steady state
error	uncertainty resulting from the use of the wrong methods, models, and data in assessment activities

ignorance	uncertainty resulting from a lack of knowledge about the true value of a parameter that can result from inadequate or imperfect measurement
inbreeding depression	detrimental changes in birth and death rates resulting from reduced genetic diversity (usually a problem of small effective population sizes)
life-table	age- or stage-specific schedules of fecundity and survivorship
meta-analysis	an analysis that compares the outputs of two or more models constructed using different assumptions
metapopulation	a biological population consisting of two or more subpopulations separated in space
parameterization	quantification of the variables in a model
population	variously: a collection of individuals of a single species that occupy some defined geographical space; a subset of all individuals of a given species that share a common area and that interbreed
population attribute	a characteristic of the assessment population
population measure	a metric used to quantify a population attribute
prediction	a quantitative description of the future abundances or behavior of a population
projection	a qualitative description of the future abundance or behavior of a population assuming constant environmental conditions
sensitivity analysis	an evaluation of the influences of model variables on model outputs
state variable	a component or property of the system being modeled that, when aggregated with other state variables, determines what the system looks like
subpopulation	an internally coherent subdivision of the larger population

transition probability	a demographic model parameter describing the likelihood that individuals will move from one age or stage to another
transmutation	qualitative changes in the causes and mechanisms of response at different levels of biological organization
variability	uncertainty resulting from actual differences in the value of a parameter or attribute among units in a statistical population
vital rates	age- or stage-specific birth and death rates of individuals within the population (synonymous with *demographic rates*)

1. Statement of Problem and Scope of Response

In July 2000, ORD's Ecological Risk Assessment Support Center received a request from the Ecological Risk Assessment Forum (ERAF) to provide insight to issues associated with site-specific ecological risk assessments involving populations (Appendix). This request contained questions that have frequently been asked in the context of several regulatory programs, including Superfund, Resource Conservation and Recovery Act (RCRA), pesticide registration, water quality criteria and standards development, and National Pollutant Discharge Elimination System (NPDES) permitting, as well as in efforts undertaken by natural resource trustees. These questions include:

- How should populations be defined?
- What should we measure to characterize populations?
- How do we make the connection between measurements made on individual organisms and population effects?
- How can we measure and interpret effects on natural populations?
- What spatial and temporal scales are appropriate when assessing risks to populations?

Although relevant to almost all risk assessments involving populations, the problem reflected in these questions is particularly acute in the Superfund Remedial Investigation/Feasibility Study (RI/FS) and RCRA Facility Investigation/Corrective Measures Study (RFI/CMS) processes. While objectives like "population effects will be addressed" are something to be desired and are often required, Remedial Project Managers (RPMs) and site managers may lack practical knowledge of the concepts, approaches, and methods for assessing the effects of environmental stressors on populations. This problem creates difficulties during the problem formulation stage of the assessment and often leads to uncertainty in the interpretation and use of assessment results in support of management decisions.

This white paper attempts to support performance and interpretation of population-level risk assessments conducted during Superfund RI/FS and RCRA RFI/CMS processes by providing information relevant to the above questions. Its structure and content are determined by the ERASC request and subsequent discussions with ERASC and Regional Office staff. It does not attempt to provide specific, "how to" guidance for performing site-specific population-level risk assessments, and it specifically avoids explicit consideration of salient exposure analysis methods and approaches for risk characterization (i.e., for comparing and interpreting site-related population-level effects). Although much needed, development of such guidance is well beyond the scope of the ERASC support function. Rather, the focus of this white paper is strictly on issues relevant to characterizing effects on populations. And although structured in a way to support site-specific risk assessment for Superfund and RCRA, it is written primarily from the viewpoint of population ecology. (As this white paper was being reviewed, the Society of Environmental Toxicology and Chemistry (SETAC) organized an international workshop on population-level ecological risk assessment. This workshop addressed many of the issues salient to the ERASC request, and considered other issues that may be germane to assessing risk to populations at Superfund and RCRA sites. The proceedings of that workshop (Barnthouse et al., in review) were undergoing external peer review as the current document was being finalized. Agency staff are encouraged to refer to Barnthouse et al. (in review) for discussions of several scientific issues, management and population protection issues, empirical and modeling tools, and application of the Agency's Ecological Risk Assessment Guidelines to population-level ecological risk assessment.)

In addition to this context-setting material, the white paper consists of seven major sections: Section 2 describes ways in which populations can be defined relative to site-specific risk assessments, settling upon the concept of "assessment population" as described by the Risk Assessment Forum's recently completed Generic Endpoints for Ecological Risk Assessments project; Section 3 enumerates the attributes that describe populations, and offers some opinion about which might be relevant to Superfund and RCRA risk assessments; Section 4 surveys mathematical and other models that can be used to extrapolate effects on individuals to population response, and includes information about using toxicity test data in such extrapolations; Section 5 describes techniques and issues associated with estimating population attributes, and the demographic parameters used to calculate them, in populations at field sites; Section 6 identifies some of the major outstanding research needs relative to population-level risk assessments, and describes some of the work ORD is conducting to meet those needs; and Section 7 lists the extensive literature cited throughout the paper, identifying key references for additional description of relevant topics.

Although reasonably complete, no attempt was made in this white paper to provide an exhaustive review of available concepts and methods in population ecology. The information contained herein should be used as an entrée and primer to the state-of-the-science that supports evaluation of population-level effects. With time and significant additional effort, this work might form the basis of guidance for assessing risks to populations. Plans for accomplishing this have yet to be identified.

2. Definition of Population

What is a population?

This question is of central importance when establishing assessment endpoints and identifying analysis activities as part of a site-specific risk assessment. Several definitions have been used for the term *population*, including general definitions used in classical ecology and modern biology, and operational definitions developed for specific purposes. The usefulness of any of the definitions in a Superfund or RCRA risk assessment depends upon the context of the assessment and the risk it is attempting to characterize.

2.1 Ecological Definitions

In the most general sense, a population is simply a collection of individuals of a single species that occupy some defined geographical space. This definition works well for many ecological studies, as long as the rates of migration into and out of that collection of individuals are small relative to rates of internal replacement (i.e., reproduction). More specific definitions add requirements for interbreeding and the exchange of genetic material. Thus, a population can be defined on the basis of a shared gene pool. Such a collection of individuals can be called a *Mendelian population* (Pianka, 1974), reflecting classical concepts of Mendelian segregation and heredity. Mendelian populations might be defined on the basis of genotypes and allele frequencies, using modern techniques of population genetics and molecular biology. This definition, of course, applies to organisms that reproduce sexually, as true asexual organisms have no opportunity for genetic exchange.

Thus, the term *population* as used in ecological studies denotes a subset of all individuals of a given species that share a common area and that interbreed. In reality, it is unlikely that all individuals in a given geographical space truly interbreed. Population studies therefore generally rely on demarcations imposed by boundaries that constrain organism movement (or more specifically, genetic exchange) such that the dynamics of population size and structure are controlled by processes operating internally as opposed to by external processes such as migration. Such boundaries can be natural (as functions of geography and geology) or artificial and imposed. However, when boundaries are not obvious, or when they don't appear to coincide with those of the hazardous waste site or operable unit/waste management unit, the question of how to define the population for the purposes of the assessment likely remains.

There is a substantial literature describing the concepts and theory of population ecology that build from these general definitions. Because most of this is beyond the scope of this white paper, we refer those interested to any modern, general ecology text.

2.2 Operational Definitions - the Assessment Population Concept

Academic definitions notwithstanding, a population needs to be defined in a manner meaningful to site-specific risk assessments to be useful to the Superfund and RCRA processes. The U.S. EPA's Risk Assessment Forum (RAF) has undertaken a project to develop generic endpoints for ecological risk assessment. Recently finalized, this effort has drafted two operational definitions relevant to site-specific population-level risk assessments (U.S. EPA, 2003):

Assessment Population – A group of conspecific organisms occupying a defined area which has been selected to serve as an assessment endpoint entity for an ecological risk assessment.

Assessment Community or *Assemblage* – A group of organisms occupying a defined area which has been selected to serve as an assessment endpoint entity for an ecological risk assessment. The group may include all organisms in the area, in a taxon (a plant community or bird community), or in certain samples (macroinvertebrates in Hester-Dendy samples).

Reflecting the intent of the Generic Ecological Assessment Endpoints project, the remainder of this document relies on the *assessment population* concept (see Text Box 1 for further discussion on assessment populations and communities).

Factors to consider in problem formulation when defining the assessment population include various ecosystem and receptor characteristics, such as the biology and life history of the species, it's range relative to the hazardous waste site or operable unit, and the potential for movement between *subpopulations.* In addition, conservation management and policy goals and other factors that define the decision context of the risk management question are important considerations. Probably the most important consideration is driven by site management goals that the risk assessment is intended to support (see Munns et al., 2002). These goals establish the specific risk questions to be asked in the assessment, the analysis approaches to answer the questions, and ultimately, the definition of *assessment population*.

Text Box 1. Assessment Populations and Communities

(Prepared by Glenn Suter, ORD National Center for Environmental Assessment (NCEA),
for the RAF generic endpoints project; U.S. EPA, 2003)

Because the conventional ecological meaning of populations and communities presents problems in practice, this document introduces the terms “assessment population” and “assessment community” (defined in text). Although ecological assessment endpoints inevitably include population properties such as abundance and production, and community properties such as species richness, it is difficult to delineate populations and communities in the field. Classically defined populations are discrete and interbreeding. Classically defined communities are discrete and their constituent species are relatively consistent and interact in predictable ways. Although these classical definitions have been important to the development of genetics, evolution, and ecology (e.g., Hardy-Weinberg equilibrium and the competitive exclusion principle), they have always had manifest limitations in practice. More recently, ecology has become more focused on temporal dynamics, spatial patterns and processes, and stochasticity that belie the notion of static, independent populations. One example of this is metapopulation analysis which reveals that population dynamics are significantly determined by exchange of individuals among habitat patches or differential movement across a landscape that continuously varies in suitability (Hanski, 1999). Communities are subject to the same dynamics. For example, the species diversity of Pacific coral reefs is apparently determined by the availability of recruits from other reefs within 600 km (Bellwood and Hughes, 2001). If the composition of coral reefs, which would appear to be classic discrete communities, is in fact determined by regional dynamics, there is little chance of delimiting discrete communities in general.

Populations may be readily delimited if they are physically isolated within a broader species range (e.g., a sunfish population in a farm pond) or if the species consists of only one spatially discrete population (e.g., the endangered Florida panther, whose current range is restricted almost exclusively to southwest Florida). Otherwise, population boundaries are difficult to define because they are typically structured on multiple scales. Genetic analyses, which are needed to define discontinuities in interbreeding frequencies, are not a practical option for most ecological risk assessments.

The practical problems are even greater for communities. Although the members of a population consist of a single species, it is not always clear whether a particular group of organisms constitutes an instance of a particular community type. This is because the species composition of communities varies over space and time.

Text Box 1 (cont.)

To protect properties such as population production or community species richness, it is necessary to develop a pragmatic solution to these problems. An example of such a solution is the approach taken by the Nature Conservancy and NatureServe (formerly the Association for Biodiversity Information) to inventory and map biodiversity (Stein et al., 2000). Because it is not feasible to define discrete populations or communities, these organizations inventory and map occurrences of conservation elements, which may be defined at various scales, depending on the elements and circumstances. For example, a plant community occurrence may be "a stand or patch, or a cluster of stands or patches." However, an occurrence of a bird species would be defined quite differently.

We propose a similar solution for GEAEs [generic ecological assessment endpoints]. For individual assessments, the population or community entities to be protected must be defined during the problem formulation stage of risk assessment. These assessment populations and assessment communities should be defined in a way that is biologically reasonable, supportive of the decision, and pragmatic with respect to policy and legal considerations. For example, it would not be reasonable to define the belted kingfishers occurring in a 20 m stream reach as an assessment population if that reach cannot fully support one belted kingfisher pair. On the other hand, even though the kingfisher's range is effectively continuous, it would not be reasonable to define the entire species as the assessment population, given that it ranges across nearly all of North America. Rather, it may be reasonable to define the kingfishers on a watershed or a lake as an assessment population.

Assessment populations may be defined by nonbiological considerations as well. For example, for Superfund ecological risk assessments on the Department of Energy's Oak Ridge Reservation, populations of large terrestrial vertebrates were delimited by the borders of the reservation (Suter et al., 1994). This definition was reasonable not only because the Superfund site was defined as the entire reservation, but also because the reservation was large enough to sustain viable populations of deer, wild turkey, and bobcat, among others. Although the reservation is more forested than are the surrounding agricultural and residential lands, its borders are not impenetrable and are not ecologically distinct at all points. However, the pragmatic definition proved useful and acceptable to the parties. For similarly practical reasons, one might define an assessment community of benthic invertebrates in the first fully mixed reach of a stream receiving an effluent.

The selection of a scale to define an assessment population or community involves a tradeoff. If the area is large relative to the extent of the stressor, the effects of that stressor will be diluted. However, if the area is small, the assessment population or community may be significantly affected but may seem too insignificant to prompt stakeholder concern or action by the decisionmaker. Hence, appropriate spatial scales should be determined during the problem formulation stage for individual risk assessments, taking into consideration both the ecological and policy aspects of the problem; it should not be manipulated during the analysis to achieve a desired result.

Imprecisely framed or communicated management goals can lead to confusion about how to conduct the assessment and interpret its results. It probably is too vague to state the goal simply as "to control risks of the site to a naturally occurring population." Such a goal can open the door to a range of analysis approaches, and can result in answers that may not address the intent of site managers directly. Often, this problem boils down to one of incomplete specification of the goal relative to spatial and scalar relationships between the site and the natural range of the population. It may be useful to consider three general scalar situations or cases (Figure 1) as they relate to management goals, lead to different assessment questions, and therefore different analysis approaches:

Case 1 – The site boundaries encompass the majority of the natural range of the population.

Case 2 – The site is located within, and is substantially smaller than, the natural range of the population.

Case 3 – The site is located outside the seasonal range of the population, but some portion of individuals utilize the site during migration.

The simple management goal articulated above is applicable in each of these cases, and while it is possible (and reasonable) to assess the risks of the site to the entire population in each case, the relevant assessment questions and analysis approaches likely would differ. Certainly, the level of adverse effect on overall population dynamics expected from a given site would differ across cases, with Case 1 representing the greatest risk and Case 3 the smallest (simply as a function of the proportion of the population exposed). Therefore, differences in the analysis approaches to address the simply-stated goal, and their outcomes, can contribute to problems in interpretation of assessment results.

Clarification of the management goal by defining the assessment population should help to minimize such problems. In situations like Cases 2 and 3, if the goal is to manage site risks as if the population's range was limited to the site itself (making the analysis most conservative and analogous to that for Case 1), the assessment population should be defined as those organisms using the site. This might be the appropriate goal when the health and sustainability of local subpopulations are a concern for aesthetic or other societal reasons. In this situation, special consideration may need to be given in the analysis to migration of individuals from and to the surrounding area. Conversely, when the goal is to manage risks of the local site in its broadest context, the assessment population might be defined as the entire natural population. This situation might arise when particular advocacy groups, operating as stakeholders in the risk assessment/risk management process, are concerned more about overall cumulative risks to a resource species (for example, northern pintail ducks) than they are about local site effects. The analysis in this situation should include evaluation of how localized adverse effects on

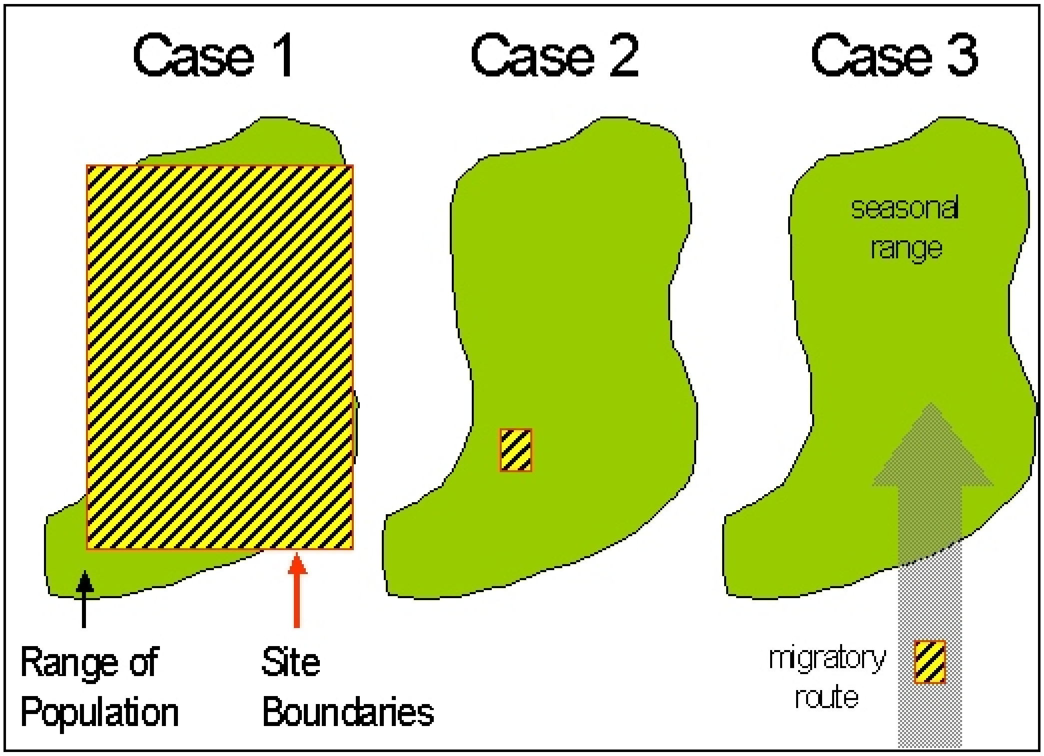

Figure 1. Spatial and scalar relationships between site and range of population. See text for explanation of cases.

individuals at the site impact the dynamics of the broader population as a whole; if severe effects occur within the site, the site might act as a “sink” that continually removes individuals from the assessment population in a manner that significantly reduces overall population abundance.

In final analysis, the definition of population used within a site-specific risk assessment is situationally-dependent and specific to the management goal and risk problem being addressed. Clear communication of how the population is defined and the assumptions made to support that definition will guide analysis approaches and facilitate understanding of assessment results and limitations by site managers and stakeholders.

3. Survey of Population-Level Attributes

Populations and their vitality can be characterized using a number of different attributes. However, not all population-level attributes are useful in site-specific risk assessments; use of some of the more esoteric attributes is perhaps best reserved for theoretical exercises of population biology. Some 16 attributes potentially useful in Superfund and RCRA assessments are described below, and are summarized in Table 1. These are organized into four broad, and potentially overlapping, categories: abundance, population growth, structure, and persistence. Much of the following has been abstracted from Maltby et al. (2001); additional considerations about specific attributes can be found there and in modern ecological texts.

3.1 Population Abundance

Population size - Population size, expressed as total number of individual organisms, total biomass, or any other summed quantity of mass or energy, is one of the most basic attributes of a population. It usually is denoted by N or N_t, where t equals time, although when expressed as biomass or some other quantity, other symbols may be used.

Population density - Density is simply the size of a population expressed on a per area or per volume basis. It is a convenient way to describe abundance in that it facilitates comparisons of abundance among populations or areas. Probably more importantly, however, the density of a population has implications to regulation of the population, in that density dependence (or *compensation*) can influence the demographic rates that determine population abundance. Although simple in concept, the idea that density can affect birth and death rates (say, through intraspecific competition for limiting resources) can be difficult to quantify and model.

Equilibrium abundance - The equilibrium abundance, or steady state abundance of a population is that population size at which inputs to the population (births, immigration) are exactly balanced by losses (deaths, emigration). It is related to the concept of carrying capacity, the number of individuals that the resources of the environment can support without increasing or decreasing. Maltby et al. (2001) suggest that changes in equilibrium abundance have greater significance than do changes in population size, since the former implies some long-term effect whereas the latter implies a temporary effect. Equilibrium abundance is probably best thought of as a theoretical concept in that most populations actually fluctuate around some average abundance as a result of environmental variability, demographic stochasticity, and so on.

Production and yield - These two attributes generally relate to the amount of a population (in terms of biomass or numbers) that can be harvested from a population. Generally applied in fishery management, they are appropriate for any situation in which the population is managed for take.

TABLE 1

Population-level Attributes for use in Ecological Risk Assessment

(see Maltby et al., 2001 for other descriptive information)

Attribute	Measure	Example Units of Measurement	Applicability to Site-Specific Risk Assessment*
Abundance			
population size	numbers, biomass, energy content, etc.	#, kg	general
population density	population size per unit area or volume	$\#/km^2$, g/m^3	general
equilibrium abundance	steady-state numbers, biomass, energy content, etc.	#, kg	specific
production/yield	harvestable quantity	kg, #	specific
spatial distribution	extent of spatial occupation, number of habitat patches used	hectares, # of patches	specific
Growth Rate			
population growth rate	change in population size per unit time	$\Delta\#/\Delta t$, $d\#/dt$	general
Structure			
age/size/stage structure	distribution of population among classes of age/size/developmental stage	#/class, kg/class	general
genetic diversity	heterogeneity in gene frequency	various indices	specific
tolerance distribution	variability in resistence to stressor effects	# affected in each treatment	specific
sex ratio	proportion of one sex relative to the other	♂:♀	specific
spatial distribution	extent of spatial occupation, number of habitat patches used	hectares, # of patches	specific

TABLE 1 (cont.)

Attribute	Measure	Example Units of Measurement	Applicability to Site-Specific Risk Assessment*
Persistence			
probability of extinction	likelihood that population abundance will go to zero	probability curves	specific
time to extinction	time between introduction of stressor and population extinction	years	specific
quasi-extinction	likelihood that population size will fall below some defined abundance	probability curves	specific
minimum viable population	smallest population abundance that will persist for a fixed period of time	#, kg	specific
stability	resistence or resilience to change		specific
recovery time	time between removal of stressor and return of population to equilibrium abundance	years	specific
spatial distribution	extent of spatial occupation, number of habitat patches abused	hectares, # of patches	specific

* Information in this column reflects value judgments by the authors

Spatial distribution - The spatial distribution of a population might be thought of as an attribute related to abundance, structure, and/or persistence. As an abundance attribute, it is related to population size - as the range of the population increases, its abundance also increases (assuming density to be constant). As a structure attribute, it describes the locations of the population and its uses of various habitats in the landscape. And as a persistence attribute, it can be related to the probability that a population will become extinct due to the local effects of anthropogenic stressors and environmental variability. However, few examples exist of the direct use of this attribute as an indicator of population level effect (but see Goldingay and Possingham, 1995).

3.2 Population Growth

Population growth rate - Changes in population abundance can be characterized by population growth rates that express those changes as a function of time. Population growth rate is generally denoted as r, the intrinsic rate of natural increase (also referred to as the Malthusian parameter), or as λ, the finite rate of population increase (also called the population multiplication rate). These two rates are related mathematically ($r = \ln(\lambda)$), and their use to express population growth rate generally is a matter of preference, although the nuances of how they are calculated and how they are used in population models (see below) can dictate which expression to use. Generally, r ranges from $-\infty$ to $+\infty$, and is symmetric around the value 0, which represents zero population growth. Thus, values less than 0 indicate a declining population abundance, whereas positive values of r indicate a growing population. λ, on the other hand, ranges from 0 to $+\infty$, with a value of unity representing zero population growth, values less than unity a declining population, and values greater than unity a growing population. Thus, at $r = 0$ and $\lambda = 1$, births and immigration into the population are exactly balanced by deaths and emigration out of the population.

In addition to characterizing changes in abundance through time, population growth rate can be thought of as an attribute related to population fitness (Leslie, 1945; Demetrius, 1975; reviewed in Caswell, 2001). That is, in a theoretical sense, populations with genotypes more suited to a particular environment should have greater rates of population growth than do those less well suited, and as a result, the better suited (i.e., more “fit”) population will out compete all others (all else being equal). Conversely, the degree to which population growth rate is adversely affected by environmental conditions (e.g., contaminant loadings) is a measure of risk to the population. In extreme cases, population growth rate may be reduced to such an extent that the population is projected to go (at least locally) extinct. Less radical reductions might increase the risk of extinction when such impacts occur in conjunction with other environmental insults or with stochastic fluctuations in controlling processes.

Population growth rate, and the extent to which it is impacted by stressors, is therefore of central importance in ecological and evolutionary theory. It is generally acknowledged to be the key variable linking individual effects to populations (Sibly, 1999; Calow et al., 1997), is a fundamental underpinning to population regulation (Sinclair, 1996), and underlies most fisheries and wildlife management concepts and approaches (Sutherland and Reynolds, 1998). However, the use of population growth rate in ecotoxicology has been limited (Sibly, 1996), despite its usefulness as a measure of stress having been demonstrated some 40 years ago (Marshall, 1962). As with changes in all population-level attributes, care should be taken in the interpretation of differences in population growth rate potentially attributable to a hazardous waste site, since factors other than toxic chemicals can and do influence this attribute.

3.3 Population Structure

Age/size/stage structure - Population structure can be characterized by the relative distributions of individuals (or biomass) among classes of age, individual size, and developmental stage. These classes can be delimited on a fairly arbitrary basis (such as years in calendar time), or can be tightly linked to the biology of the species (e.g., eggs, larvae, pupae, adults). In general, the distribution of individuals among classes is an indication of the status of the population. In populations that are growing rapidly, the distribution of individuals typically is skewed towards the younger classes. The opposite may be true for populations experiencing reductions in overall size. To illustrate this idea, human populations in developing countries that are experiencing relatively rapid population growth generally have lots of young and few old adults, whereas the distribution among ages in industrialized countries with more stable population sizes tends to be more even (Thomlinson, 1965).

Anthropogenic stressors can affect the age or size structure of a population by modifying the processes of births and deaths. Depending upon how these effects manifest, the age/size/stage structure can shift from pre-exposed conditions in ways that do not map neatly onto the generalization above. If, for example, the susceptibility of individuals to a particular chemical increases with age, the resulting distribution might still be skewed towards younger age classes, even though a population-level effect is occurring. Thus, caution is needed when interpreting age/size/stage distributions. However, a comparison of assessment population structure with that of a pre-exposed or reference population would still indicate an effect.

Genetic diversity - The genetic diversity of a population may be related to its fitness (i.e., its ability to persist through time). Population biology theory suggests that populations with greater genetic heterogeneity in fitness-related genes should be able to withstand a wider range of environmental conditions (including the presence of anthropogenic stressors) than can those with less diversity. Thus, genetic diversity can be an indicator of susceptibility to future environmental impact. It also can be an indicator of current and past stress, because such stress can reduce genetic heterogeneity when population size decreases rapidly in response to disturbance (the so-called *bottleneck effect*; Weins, 1977). Chronic or multiple episodic stressors may cause multiple bottlenecks, magnifying the reduction in genetic diversity, and there is mounting evidence that nonmigratory populations at sites that have experienced pollution or other forms of disturbance have reduced genetic diversity (Lavie and Nevo, 1982; Lavie et al., 1984; Nevo et al., 1986; Benton and Guttman, 1990). Reduced genetic diversity can expose the population to greater risks of adverse impact from other anthropogenic stressors, disease, and environmental stochasticity, and can cause detrimental changes in birth and death rates (inbreeding depression). The importance of genetic diversity in fitness-related genes to the persistence of wild and captive populations has been recognized by conservation biologists and is incorporated into management strategies for the preservation of threatened and endangered species. It should be noted that pollution stress can also increase genetic diversity through selection for tolerant genotypes in populations that still contain pollution-sensitive genotypes, but interesting, that mechanism would produce a pattern consistent with the general heterogeneity/persistence argument (i.e., increased heterogeneity and increased likelihood of persistence).

Tolerance distribution - Basically a concept of ecotoxicology, the distributions of tolerances (or conversely, susceptibilities) to chemical stressors is directly related to the previous attribute in that tolerance is a *phenotypic* measure of the underlying genetics. As with shifts in genetic diversity, temporal or spatial changes in tolerance distribution of the assessment population relative to pre-exposed or reference populations can provide an indication of population-level effects.

Sex ratio - The sex ratio of a population describes the relative abundances of the two sexes in sexually reproducing dioecious populations (and sometimes seasonally parthenogenetic species, such as daphnids, rotifers, and aphids). Substantial theory exists concerning the evolutionary consequences of skewed sex ratios in populations (Maynard Smith, 1978), although the use of this attribute as an indicator of population-level effects is limited primarily to quantifying differential effects among the sexes. However, it also can provide information regarding the mechanisms of such effect. For example, certain chemicals have been demonstrated to interfere with hormonal systems controlling sexual differentiation in certain birds (Fry and Toone, 1981), mammals (Jones and Hajek, 1995; Gray and Kelce, 1996), reptiles and amphibians (Bergeron et al., 1994; Guillette et al., 1994, 1995), and fish (Gimeno et al., 1996; Jobling et al., 1996; Gray and Metcalfe, 1997). In such cases, ecological theory (if not empirical evidence) suggests that a change in sex ratio affects population dynamics by influencing total reproductive output of the population and by altering frequencies of encounter between the sexes during the reproductive season (Kalmus and Smith, 1960; Hamilton, 1967). Stressor-induced imposex, such as that reported for certain invertebrates (Gibbs and Bryan, 1986; Moore and Stevenson, 1991), should have population-level ramifications similar to changes in sex ratio. Thus, sex ratio is a potentially useful attribute in descriptions of the effects of stressors, particularly in situations where such effects might be diagnostic.

3.4 Population Persistence

Probability of extinction and time to extinction - Probability of extinction and time to extinction are related measures of population persistence. Probability of extinction as an attribute describes the likelihood that population abundance will go to zero under defined demographic, genetic, and environmental conditions. Time to extinction can be derived directly from the probability of extinction (Foley, 1994; Gillman and Hails, 1997). Examples of the use of these attributes are described by Harrsion et al. (1991) and Foley (1994) for butterflies, and Gillman and Silvertown (1997) for plants. Snell and Serra (2000) recently suggested that probabilities of extinction can be used to interpret the ecological significance of toxicity test results. Because they are difficult to quantify in natural populations, and are only slightly less so in laboratory experiments (due to the replication required), estimates of these attributes are best made using modeling and simulation techniques.

Quasi-extinction - Related to probability of extinction is quasi-extinction, defined as the probability that the population will fall below some *critical abundance* (Ginzburg et al., 1982). Critical abundance can be thought of in terms of specific population densities below which adverse effects are known, or suspected to occur. For example, critical densities of puma may exist below which individuals are no longer able to find mates for reproduction. Similarly, critical population sizes of commercially-harvested shellfish might be identified (through, say, a cost-benefit analysis) below which harvest is no longer economically feasible. As with probability of extinction, quasi-extinction is best analyzed using modeling and simulation techniques. Maltby et al. (2001) describe a method by which the quasi-extinction probability curves that result from such analyses could be used to evaluate the ecological significance of estimated risks.

Minimum viable population - Also related to probability of extinction is the concept of minimum viable population (MVP), defined as the smallest population abundance that will persist for some specified length of time with a given probability. Based on the expectation of a negative relationship between absolute population abundance and the likelihood of extinction (due to the effects of stressors or environmental variability), it, like the previous two attributes of persistence, is best quantified using modeling and simulation techniques. Examples of MVP as an attribute are given by Samson et al. (1985), Shaffer and Samson (1985), and Goldingay and Possingham (1995).

Stability - Broadly defined, population stability is the tendency of a population to remain at or near its equilibrium abundance when disturbed (Haberman, 1977; see Lewontin (1969) and May (1973) for discussions of the various meanings of stability in an ecological context). A stable population is one that “resists” adverse impacts due to disturbance, or one that recovers to its equilibrium abundance once disturbed. Destabilization of population dynamics can lead to large changes in abundance, thereby increasing the probability of extirpation (as well as disruption of community and ecosystem structure and function). Stability (in addition to population abundance itself) also is inversely related to the probability that the population will be driven to local extinction as a result of stressor exposure. Because it reflects the resilience of a population following disturbance, stability may be more important to population sustainability than is population growth rate. Unfortunately, stability is very difficult to measure directly (Connell and Sousa, 1983), although it can be evaluated using models. Because of this, stability as an attribute for population-level ecological risk assessment likely is limited.

Recovery time - When a population is disturbed from its equilibrium or pre-exposed abundance, the time it takes to return to that abundance is called its recovery time. Recovery time is a highly attractive population-level attribute in that it is readily understood and appreciated by regulators, managers, and the public. In part because of this, recovery has been advocated in the consideration of adversity of effect by the Agency (U.S. EPA, 1998). However, several conceptual and methodology issues can confound its use in site-specific risk assessments. Because it may be unrealistic to expect a population to return *exactly* to its predisturbed abundance, criteria must be established to define when recovery has occurred (or is sufficient). These might include aspects of absolute population abundance (e.g., 90% of original population size; Sherratt et al., 1999), natural variability in that abundance (e.g., within two standard deviations of long-term mean abundance; Weins, 1996), its abundance relative to a reference or control population (e.g., 90% of reference population density; Thacker and Jepson, 1993), or perhaps its population growth rate relative to that of a reference or control population (e.g., 90% of the growth rate of an unaffected population; Kareiva et al., 1996).

3.5 Considering Attribute Selection

Which attributes should be evaluated to characterize population-level effects at Superfund and RCRA sites? Unfortunately, there likely is no one single best answer to this question, and, as is the case with defining the assessment population (the “ecological entity” element of the operational definition of the assessment endpoint; *sensu* U.S. EPA, 1998), selection of the population attribute (the “characteristic of the entity” in definition of the assessment endpoint) should depend on the goals of the assessment. However, consideration of several criteria should help guide attribute selection in any given risk assessment. Four of these would be: the relevance of the attribute to the assessment population, the susceptibility of the attribute to contaminants at the site, its relevance to stakeholders and the risk management and communication processes, and the tractability (ease) of its measurement. In most regards, these criteria mirror the general guidance provided by U.S. EPA, (1998) and program-specific documentation.

There are few, if any, broad ecological considerations to help evaluate the importance of population attributes to the assessment population generically. However, some attributes may have more meaning than others. For example, sex ratio has no meaning for species displaying obligatory parthenogenesis, whereas other attributes reflecting population structure (e.g., genetic diversity) may be critically important. From an evolutionary point of view (admittedly not a typical view of site-specific risk assessments), attributes reflecting population persistence likely have more meaning than those reflecting abundance, growth rate, or structure, but have less meaning in the context of shorter-term ecological phenomena (e.g., role in community dynamics, nutrient cycling). Part of the difficulty when evaluating this criterion is a lack of understanding and consensus as to the definition of a “healthy” population.

There also are no *a priori* reasons to suspect certain attributes to be more susceptible to chemical stressors than others. Plausible biological mechanisms exist to link effects on any particular attribute to chemical exposure, oft times in interrelated fashion. For example, a chemical linked to differential mortality or reproduction may affect population structure as measured by age structure or genetic diversity, which in turn may be reflected directly in abundance, directly in growth rate, or indirectly in persistence as a function of changing resistence to additional stressors. Despite this, understanding biological mechanisms may help to identify which attributes would respond more quickly or signal the effect more strongly.

From a stakeholder and risk communication perspective, changes in abundance and persistence arguably are more understandable and obvious effects than are attributes of growth rate or structure. And certain attributes are more easily measured (e.g., those reflecting abundance) than others (e.g., stability). But in the end, there is no one “best” attribute for evaluation in a population-level risk assessment. As with definition of the assessment population (Section 2), selection of the population attribute is situationally-dependent and specific to the management goal and risk problem being addressed. Clear communication of why any particular attribute is selected will facilitate understanding of assessment results and limitations by site managers and stakeholders.

3.6 Ecological Significance of Responses

The ability to measure or quantify population-level effects is only part of the problem in site-specific risk assessment – changes in population attributes due to exposure to stressors must be interpreted to understand the significance of estimated risks. Appreciation of this significance is crucial to developing, executing, and interpreting ecological risk assessments. A question central to this issue is: “What do the magnitude and direction of observed attribute response mean from an ecological standpoint?”. This question can be restated simply as: “So what?”. Answering this question is not straightforward, and presents complex challenges that have yet to be overcome by science. The following material is intended only to be an introduction to this issue; considerable work remains to advance the theory and practice of population ecology to the point where the ecological significance question can be answered satisfactorily.

Except in extreme situations (e.g., extinction), ecological significance is difficult to establish for at least four reasons. First, all ecological systems, including populations, display natural variability: population abundances fluctuate (perhaps around some steady-state abundance) due to environmental and demographic stochasticity, age structures change as reflections of year-class strengths, sex ratios vary randomly (or otherwise) around values of 1:1, and so on. Such variation has been called *process variation*. The challenge is to distinguish changes caused by anthropogenic stressors against this background of natural variability. Historically, inferential statistics has been used to make such distinctions. This has lead to application in a number of programs of seemingly arbitrary "bright lines" that form the basis of regulatory decisions (Chapman et al., 2002). Yet, statistical differences (the detection of which are functions of the test and sampling designs, natural variation, and the magnitude of "treatment" differences) do not necessarily equate with changes that are important ecologically. Establishing ecological significance is also hindered for a second reason: ecological systems typically vary in near-continuous fashion, such that in the absence of thresholds or obvious breakpoints, the importance of subtle changes is difficult to establish relative to less subtle changes. A third reason confounding interpretation of ecological significance is the existence of compensatory mechanisms, such as homeostatic acclimation of individuals, genetic adaptation, and density dependence in vital rates and migration, that can ameliorate adverse effects over the short or long term. Finally, there are few (if any) "values" that ecological systems place on themselves. In the absence of such values, ecological change is interpreted in the context of societal desires, preferences, needs, and policies.

In 1994, the U.S. EPA's RAF published an issue paper on the general topic of ecological significance as it pertains to risk assessment (Harwell et al., 1994). That paper offers a definition of ecological significance consisting of components that reflect both ecological structure and function, and societal values. Paraphrased from Harwell et al. (1994), an ecologically significant change is one that is important to the structure, function, or health of the system, exceeds natural variation, and is of sufficient type, intensity, extent, or duration to be important to society. Although their definition was offered in the context of the significance of estimated risk, it also can be applied to the significance of changes measured in population attributes themselves. Harwell et al. (1994) offer a tiered ecological significance framework or road map and general criteria for establishing ecological significance that include consideration of the nature and variability of the attribute, the temporal and spatial scales of effect, the magnitude and reversibility of the effect, and other considerations. They also suggest how issues of ecological significance can be used to support decision-making. Given our current state of understanding of population ecology, and in the absence of objective, quantifiable criteria against which the importance of observed or predicted change in population-level attributes can be judged, establishing ecological significance will continue to rely on sound professional judgement reflecting the considerations offered by Harwell et al. (1994) and the work cited therein. General guidance for evaluating ecological significance is given in U.S. EPA (1998).

Although discussed in terms of changes measured or predicted in population-level attributes themselves, questions of ecological significance are also cogent from the standpoint of predicting population-level effects from those measured in individuals. Methods for linking effects on individuals to population response are described in the next section.

4. Linking Effects on Individuals to Population Response

Most population-level attributes, including abundance attributes, are determined by the vital rates (births, deaths) of individuals within the population, as well as the rates of migration into and out of the population. These four rates - births, deaths, immigration, and emigration - are necessary and sufficient to describe changes in population numbers through time, although additional information (e.g., individual growth rates, weight-age relationships) may be needed for some expressions of abundance (e.g., biomass). Any factors that influence these processes, either directly or indirectly, can have population-level consequences. Linkages between effects on vital rates and those on population dynamics, or between effects on the biochemical and physiological processes that determine vital rates and effects on populations, can either be established empirically by correlating responses at different levels of biological organization, or by determining causal relationships and constructing mechanistic models (see Maltby et al., 2001 for a conceptualization and discussion of these linkages). Thus, with some degree of understanding of how changes in vital rates manifest into population consequences (i.e., mechanistic understanding), effects measured at the level of individuals can be extrapolated to expected population-level responses. Extrapolation of this latter type is usually accomplished using models that integrate the effects of stressors on survivorship and fecundity.

This section focuses primarily on modeling approaches that can be used to assess risks to populations. Its content is derived primarily from Munns (1988) and Maltby et al. (2001). Although their description is beyond the intended scope of this white paper, empirical approaches (including extrapolation among species as well as among attributes) have been developed that can be useful for site-specific risk assessments [see Munns (2002) for a general discussion of extrapolation issues in risk assessment].

4.1 About Models

In his broad discussion of ecological theory and models, Levins (1968) describes a triangular scheme for ordinating ecological models that has the attributes of *generality*, *realism*, and *accuracy* (originally *precision*[1]) as its apices (Figure 2). In this context, general models are those that tend to be simple and apply to a broad range of situations, and therefore are appropriate for exploring relationships among model parameters and outputs. Realistic models attempt to account for known relationships and processes in ecological systems, and as a result can be relatively complex. Accurate models are constructed with an objective to minimize numerical differences between model outputs and actual ecological dynamics. Their case-specific nature limits their use in broader applications.

---

[1] Use of the terms *accuracy* and *precision* in this white paper follows their connotations in the field of inferential statistics. Accuracy refers to how well an estimate matches the true value of a particular parameter or value being estimated (in this case, population abundance), and typically is quantified using some measure of bias. Precision refers to the amount of variation among multiple estimates made of the parameter, and usually is expressed using some measure of scatter.

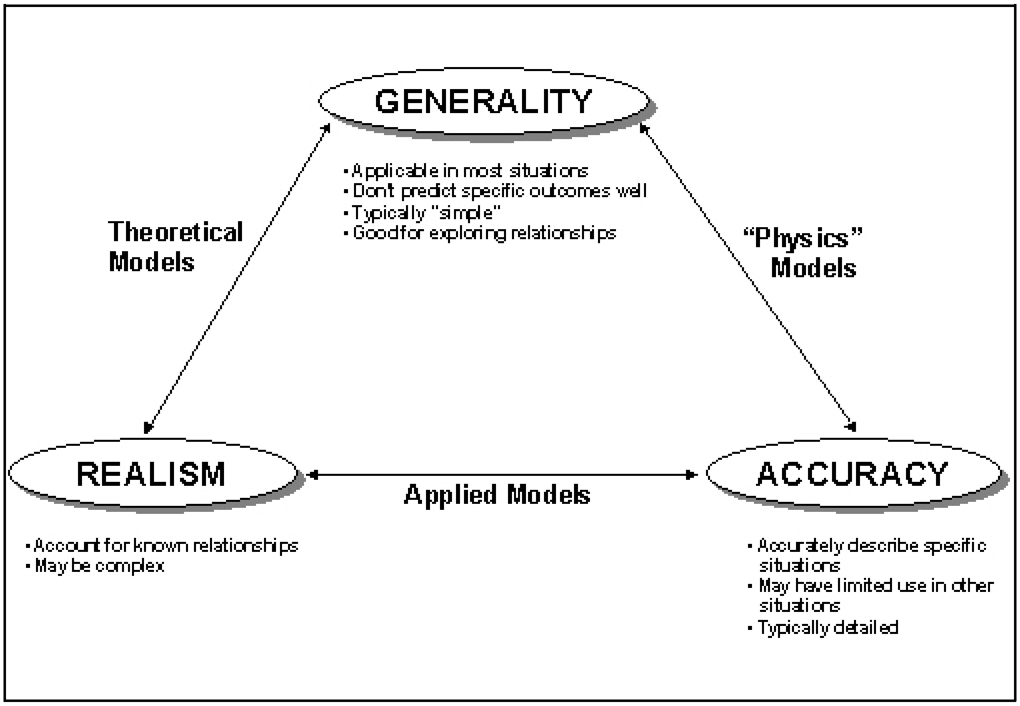

Figure 2. Ordination scheme for ecological models (modified from Levins, 1968).

Levins (1968) pointed out that any two of these attributes can be maximized at one time, but that it is not possible to maximize all three with a single model. Typically, models developed for use in applied situations (e.g., conservation biology) are intended to give realistic, accurate answers; model parameterization depends upon the actual conditions of the situation being modeled (e.g., the specific life history and demographic characteristics of the species of interest). Increased generality can be achieved, for example, by expanding the range of values assumed for particular model coefficients, or by assuming broad functional relationships among parameters, but such actions necessarily reduce the accuracy achievable in any particular application. Model accuracy can be enhanced by increasing the specificity of model parameterization relative to a particular species or environmental situation.

With this ordination scheme in mind, ecological models, and specifically population models, can be used in site-specific risk assessments for at least three, arguably different purposes. The first is to detect (and perhaps diagnose the causes of) previous or ongoing adverse effects on population dynamics. Such uses typically require sufficient high quality data to be able to detect changes in population abundance (methods for which are described in the next section), and to relate those changes to variation in chemical exposure, habitat quality, or other forms of disturbance. The second purpose is to project the consequences of a given set of environmental conditions (or changes in conditions) to the dynamics of a population. Here the intent may be to evaluate the ramifications of particular environmental management decisions as determined by trends in population numbers or changes in extinction probabilities. The final purpose is to forecast or *predict* the future behavior of the population based on a understanding of environmental variability and the dynamic interactions of density and biological processes (e.g., births and deaths). [The distinction between projection and forecasting used here follows that given by Caswell (1986, 2001).] This last use of population models may produce the most accurate results, although the generality of the analysis will suffer. It requires knowledge of how the major environmental and biological determinants of population dynamics (including such things as climate, prey availability, etc.) will themselves change through time (or, in a probabilistic sense, their future distributions), and sufficiently detailed understanding of the mechanisms through which these changes affect the population. Also required is understanding of how population density itself influences births, deaths, and migration through density dependence of these rates. Except in rare situations, such knowledge and understanding is difficult to obtain.

Prior to description of specific model formulations, it may be instructive to consider the assumptions and potential limitations of different modeling approaches. As a first cut, population models (indeed, almost all mathematical models) can be either *deterministic* or *stochastic*. Deterministic models treat all internal relationships (processes, mechanisms, transfer functions, etc.) as nonrandom; that is, they assume no random variation in how one *state variable* (in its simplest definition, a component or property of the system being modeled that, when aggregated with other state variables, determines what the system looks like) relates to another. This is not to say that one state variable cannot be a function of another; rather that even those functions contain no randomness. Because of this, the outputs of deterministic models are determined completely by their initial starting conditions - no matter how many times the model is run, the output will be the same as long as the values of model parameters are not changed by the user. Deterministic models sometimes can be solved analytically. Stochastic models, on the other hand, assume that at least some of the internal relationships have a random component. Their outputs are only partially dependent on initial starting conditions, and the exact values of output generally cannot be predicted in advance (although they might be bounded). Stochastic modeling approaches reflect the randomness inherent in biological and environmental processes, and usually assume specific probability distributions for the values that state variables can take. Stochastic models cannot be solved analytically except under certain conditions (such as when parameters are distributed normally), and *Monte Carlo* and other simulation techniques have been developed to support analysis of these models (see, for example, U.S. EPA, 1996).

Models (and their variables) can also differ in how they treat time. *Discrete time* models assume time to be “jumpy”, that is, time is handled as a series of blocks (seconds, days, years, etc.), with the values of model variables being updated only once during each block or time step. Discrete time models often are represented by difference equations. In contrast, *continuous models* assume time to flow continuously, such that model variables take on the value appropriate to any instant in time. Continuous time models can take the form of differential equations with respect to time.

Similar to how they treat time, models can also assume that their variables (and therefore the state of the model system itself) are either discrete or continuous. *Discrete state* models (and their variables) can take on values selected from a finite set of possible alternatives only. That is, parameter and model outputs are characterized by discontinuous changes in value; they can be X_1, X_2, or X_3, but not something between X_1 and X_2. *Continuous state* models include variables that can take on any value (perhaps within some bounded range), and outputs are characterized by smooth changes in value.

Any particular model can, of course, assume a combination of the three approaches above. For example, a modeling approach might assume stochasticity in the values that its discrete state variables take on through continuous time. The approach taken in model formulation in part determines it strengths and limitations for use in site-specific risk assessments. Together with other considerations, these issues are explored to support of model selection in this section.

4.2 Extrapolation Models

A plethora of models useful in describing population dynamics has evolved since the early demographic investigations of Malthus (1798) and his contemporaries. The following presentation is intended to provide a cursory overview of the types and classes of population models potentially useful in site-specific risk assessments. Both general constructs, and those developed specifically for analysis of fishery production, are given. More detailed reviews of the fishery models are given in Ricker (1975) and Sissenwine et al. (1979). Vaughan et al. (1984) discuss the utility of fishery models to assessing the effects of stress on fish populations. The discussion of fishery approaches given here is based in part upon the Vaughan et al.'s treatment of such models. Barnthouse et al. (1986) review population models for use in risk assessment, and Barnthouse (1992) provides additional information about individual-based models. Emlen (1989) provides a then-state-of-the-art review of use of population models in ecological risk assessment focusing on terrestrial species, and Emlen and Pikitch (1989) offer views on how to approach structuring population modeling exercises to support risk assessment. Barnthouse (1993) describes modeling approaches for evaluating population-level effects in the context of ecological risk assessment generally, and Barnthouse (1996) reviews population modeling approaches with particular reference to pesticide risk assessments. Pastorok et al. (2002) discuss the use of ecological models, including some population models, to assess risks of chemicals. Most recently, Munns et al. (in review) describe modeling approaches to population-level risk assessment, offering sidebar examples of model applications in a variety of regulatory contexts.

The general forms taken by models in each class, their equilibrium solutions (where appropriate), important assumptions, and data requirements necessary for parameterization, are briefly discussed. Most of the approaches described here do not include explicit terms for contaminant effects in the general form of the model. In these cases, impacts important to the dynamics of the population may be modeled through variation in rates and coefficients associated with fecundity, mortality, and in the case of yield-type models, with individual growth rate. When treated in this manner, contaminants are viewed no differently than other sources of variation in these vital rates.

General, heuristic models — Several models have been developed which are of heuristic value in the general description of population growth. Two such models are those describing simple exponential growth:

$$dN/dt = rN \qquad \text{Eq. 1}$$

and logistic growth (Verhulst, 1838):

$$dN/dt = rN\ (1 - N/K) \qquad \text{Eq. 2}$$

where N is population size at time t, r is the intrinsic population growth rate [= birth rate (b) - death rate (d)], and K is the so-called carrying capacity of the environment. Heuristic models typically are of limited value in field settings (except for short periods of time following new species introductions), and therefore will not be discussed further. However, several classes of models which may be useful in site-specific risk assessments have terms which take on the forms of Eqs. 1or 2. An investigation of the behavior of these models may be found in May (1976).

Demographic models - Many models used in the analysis of population dynamics incorporate age-specific demographic information. A relatively simple formulation which has been popular in the investigation of contaminant effects on population growth rate (e.g., Daniels and Allan, 1981; Allan and Daniels, 1982; Gentile et al., 1982, 1983) is the *stable age equation* (Lotka, 1925), or so-called *life table* approach. Age-specific schedules of fecundity and survivorship are the sole data requirements for this model, information that is easily obtained in laboratory experiments or field studies. The discrete version of this model (Leslie, 1948; Michod and Anderson, 1980; Caswell, 2001, p. 197) takes the form:

$$1 = \Sigma\lambda^{-(x+1)} \ell_x m_x \qquad \text{Eq. 3}$$

where λ is the geometric rate of population increase, ℓ_x is the probability of an individual surviving to age x, and m_x is the fecundity of an individual of age x. λ is solved for either in an iterative fashion or through algebraic manipulation of the sum of $\ell_x m_x$. λ is related to the more commonly utilized (incorrectly so in these applications; Michod and Anderson, 1980) population growth rate (r) by:

$$\lambda = e^r \qquad \text{Eq. 4}$$

Some important assumptions of this model are that: 1) age-specific schedules of fecundity and survivorship are independent of density (i.e., the population grows exponentially); 2) fecundity and survivorship do not change with time; and 3) fecundity and survivorship are constant within each age class. With the additional assumption of no net migration into or out of the population,

$$N_t = \lambda N_{t-1} \qquad \text{Eq. 5}$$

Thus, while zero population growth occurs when $r = 0$ in Eq. 1, no growth occurs in Eq. 3 when $\lambda = 1$. Projected impacts of environmental stressors could be examined using the stable age equation through modifications in values of ℓ_x and m_x.

A class of models frequently used in both fisheries and nonfisheries applications are the so-called *Leslie population projection matrix models* (Lewis, 1942; Leslie, 1945, 1948). As with the stable age equation, these models incorporate age-specific schedules of fecundity and survivorship to make projections of population dynamics given specified environmental conditions. Unlike that model, however, discrete time steps are incorporated explicitly into projection matrix models, allowing simulation of population dynamics through time. The sizes, n, of m individual age classes at time t are described by a series of difference equations as:

$$
\begin{aligned}
n_{0,t} &= \Sigma(n_{x,t-1}\, f_x) \\
n_{1,t} &= n_{0,t-1}\, P_1 \\
&\ \ \cdot \\
&\ \ \cdot \\
&\ \ \cdot \\
n_{m-1,t} &= n_{m-2,t-1}\, P_{m-1}
\end{aligned}
\qquad \text{Eq. 6}
$$

where f_x is the fecundity of age class x and P_x is the probability of survival of females from age class x-1 to class x. In matrix notation:

$$\mathbf{n}_t = \mathbf{M}\,\mathbf{n}_{t-1} \qquad \text{Eq. 7}$$

where $\mathbf{n}_t$ and $\mathbf{n}_{t-1}$ are vectors of age class sizes, and **M** is the Leslie projection matrix consisting of fecundities across the top row, survivorship probabilities down the first subdiagonal, and zeros elsewhere. In addition to describing total population size and the distribution of individual ages through time, Eq. 7 can be solved directly for its characteristic root, λ, the geometric rate of population increase:

$$|\mathbf{M} - \lambda\mathbf{I}| = 0 \qquad \text{Eq. 8}$$

where **I** is the identity matrix of **M**. Some important assumptions associated with the basic model are that: 1) schedules of fecundity and survivorship are independent of density and time; 2) fecundity and survivorship are constant within each age class; and 3) no net migration occurs. With these assumptions, population projection matrix models behave similarly to Eq. l: that is, growth is exponential (once a stability in age distribution has been reached). The impacts of environmental stressors can be projected in a manner similar to that used in the stable age equation: modifications can occur to fecundity and survivorship rates in the projection matrix. Projection matrix models have been generalized to stage-specific models, where stages can be defined as developmental stages, length classes, weight classes, or any other stage appropriate to a particular species (Lefkovitch, 1965). Changes from one stage to another are accomplished by non-zero transition or growth probabilities in **M** such that individuals can remain in the current class through time (i.e., no individual growth), or can move to other classes as dictated by individual growth and development rates instead of time alone [see Werner and Caswell (1977) and Caswell (1982, 2001) for further discussion of stage-classified matrix models]. With this modification, contaminant effects may also be modeled as changes in these transition probabilities.

The use of projection matrix models to characterize the adverse effects of chemical stressors on population dynamics has increased over the last decade or so, examples of which can be used to illustrate this approach. For instance, Pesch et al. (1991) evaluated the effects of contaminated dredged material on population growth rates of the marine infaunal polychaete, *Neanthes arenaceodentata*. They used an experimental design consisting of replicate exposures to four sediment treatments to obtain estimates of vital rates for each of nine 18-day ages classes. The resulting schedules of survivorship and fecundity were used as input to an age-classified deterministic matrix model to calculate treatment-specific estimates of λ, the finite rate of population increase. Population growth rates were found to decrease with increasing concentrations of the contaminated sediment, a relationship that compared favorably (albeit qualitatively) with observations made at an experimental disposal site (Pesch et al., 1991).

In an evaluation of potential adverse impacts associated with offshore disposal of municipal sewage sludge, Munns et al. (1996) conducted a prospective assessment of risk to populations of marine copepod that used an age-classified projection matrix model, the results of standard toxicity tests of the sewage sludge, and exposure models developed to describe sludge concentrations in the water column around and downstream of the disposal site. They simulated passive advection of the population with prevailing currents through the sludge exposure field, adjusting vital rates on a daily time step in accordance to the concentration of sludge encountered by the population on each day. Environmental stochasticity was simulated by allowing concentrations in the exposure field to vary lognormally. Munns et al. (1996) expressed risk estimates as a 3-dimensional response surface defined by axes of sludge loading rate, environmental stochasticity, and population growth rate with the intent of providing regulators with the information needed to select acceptable levels of disposal activity. Similar population-level assessments were conducted for relatively sedentary tilefish, *Lopholatilus chamaeleonticeps*, and the migratory Atlantic bluefin tuna, *Thunnus thynnus* (Munns, 1988).

Two examples of the use of projection matrix models relate to Superfund sites. First, Munns et al. (1997) developed a stage-classified model to describe the population-level effects of exposure to polychlorinated biphenyl (PCB)-contaminated sediment on the estuarine fish, *Fundulus heteroclitus*, residing at the New Bedford Harbor (MA) Superfund site. Using data obtained from field-collected fish, they developed exposure-response curves relating population growth rate to PCB liver burdens. In that same paper, Munns et al. developed a similar relationship for total dioxin body burden for fish fed contaminated diets. In support of an RI/FS at the Portsmouth Naval Shipyard (ME), Gleason et al., (2000) modeled the population responses of the purple sea urchin, *Arbacia punctulata*, resulting from exposure to lead, a primary contaminant of concern in the estuarine waters surrounding the Shipyard. They used standard bioassay data collected during site investigations to parameterize a stage-classified model designed explicitly to reflect the life history stages and vital rates represented by the bioassays. The population growth rate exposure-response relationship resulting from this effort was used subsequently in a weight-of-evidence characterization of ecological risks at the site (U.S. Navy, 2000; Johnston et al., 2002).

Three recent examples illustrate the development of projection matrix models for the expressed use of interpreting the ecological significance of toxicity test results as indicators of population-level effects. Reflecting the empirical approach mentioned earlier for extrapolating individual-level attributes to population response, Kuhn et al. (2000) used data from life cycle chronic tests to project the effects of some 20 chemicals on the mysid shrimp, *Americamysis* (formerly *Mysidopsis*) *bahia*, population dynamics. They then compared the concentration-based toxicity test statistics derived from standard 96-hr acute tests, 7-day rapid chronic tests, and the full life cycle tests for each of these chemicals to the chemical concentration projected by the model to represent the threshold of adverse population effects (denoted C*) using correlation analyses. Several strong correlations were observed, including (surprisingly) between C* and the acute LC_{50}. Kuhn et al. (2001) subsequently demonstrated that the age-classified model, parameterized for the chemical nonylphenol using standard toxicity test data, projected the dynamics of *A. bahia* reasonably well in a multigenerational experiment conducted in the laboratory. [As an aside, these data have been used by Maltby et al. (2001) to evaluate model formulations other than the original projection matrix model.] Most recently, Kuhn et al. (2002) have developed a projection matrix model to assist in understanding the ecological significance of standard toxicity test results for the estuarine amphipod, *Ampelisca abdita*.

Many other examples of the use of projection matrix models in the fields of ecotoxicology, fishery and game management, and conservation management can be found in the scientific literature. Given the ease with which the responses of individuals, measured in toxicological studies, can be linked to population-level attributes, and their relative flexibility to accommodate a wide variety of populations and environmental situations, projection matrix models appear to hold much promise for site-specific risk assessment. Further, analytical methods have been developed to evaluate systematically the relative sensitivity of λ to proportional changes in the transition probabilities (demographic rates) of the projection matrix. Called *elasticity analysis* (Caswell et al., 1984; de Kroon et al., 1986), this technique can be used to identify which demographic parameters influence population growth rate the most when changed, thereby focusing attention on those parameters (say, through targeted toxicity tests) in population-level risk assessments. Caswell (2001) describes matrix models in great detail, and discusses other forms of structured population models (discrete-state delay-differential equations in continuous time, continuous-state integrodifference equations in discrete time, and continuous-state partial difference equations in continuous time), the use of which in Superfund and RCRA site-specific risk assessments likely is limited at present.

Fishery models - This broad class of models, developed primarily for use in fisheries management, typically are framed in terms of biomass or numbers yield to the fishery. For example, *surplus production models* incorporate terms of biomass production and fishing effort to describe changes in population biomass. As opposed to the approaches described earlier, knowledge of the demographic characteristics of the population is not required to make this assessment; time series of catch biomass and fishing effort are the sole data requirements. The general form of such models is:

$$d\mathbf{B}/dt = \mathbf{B}\ \mathbf{R}[\mathbf{B}] - qf\mathbf{B} \qquad \text{Eq. 9}$$

where **B** is population biomass, q is a coefficient of catchability, f is fishing effort (thus qf represents fishing mortality) and **R** is some regulatory function which is itself a function of **B**. The term **B R**[**B**] describes changes in biomass in the absence of fishing mortality. **R**[**B**] can take any appropriate form, such as that proposed by Graham (1935) and used by Schaefer (1954, 1957):

$$\mathbf{R}[\mathbf{B}] = r\,(1 - \mathbf{B}/K) \qquad \text{Eq. 10}$$

where *r* and K are as described previously, but in units of biomass instead of numbers. In the absence of fishing, the population is self-regulating, and Eq. 9 describes a parabola symmetric about ½K. Integration of Schaefer's surplus production model yields a logistic trajectory of biomass through time. Pella and Tomlinson (1969) generalized **R**[**B**] to include nonsymmetric biomass yields. Some important assumptions associated with surplus production models are that: 1) catch per unit effort is proportional to total population biomass; 2) fishing effort is proportional to fishing mortality, but independent of total population biomass; 3) population size is always at equilibrium given constant fishing effort; 4) the age structure of the population is stable given constant fishing effort; and 5) no net migration occurs. The effects of environmental stressors could be incorporated into surplus production models by modifying the form or values assumed in the regulatory function (changes in *r* or K), in the coefficient of catchability (due to modification of individual growth rates or behavioral changes), or as a loss to production in addition to fishing mortality.

Another class of models developed for fishery applications are the *yield models*. These models incorporate age-specific schedules of weight and *cohort* (groups of similarly-aged members of the population) numbers, and are used to estimate the yield of individual cohorts following recruitment to the fishery. They take the general form:

$$dY/dx = F\ N(x)\ W(x) \qquad \text{Eq. 11}$$

where Y is the biomass yield to the fishery, F is the instantaneous fishing mortality rate, N(x) is an expression of population numbers of age x, and W(x) is an expression of the weight of an individual of age x. N(x) can be described by some appropriate function such as that used by Beverton and Holt (1957):

$$N(x) = N(x_c) \exp\{-(M + F)(x - x_c)\} \qquad \text{Eq. 12}$$

where x is the age of recruitment into the fishery ($x>x_c$), $N(x_c)$ is the number of recruits into the fishery, and M is the instantaneous natural mortality rate. Thus, the size of the cohort decreases exponentially with time. Weight at age x can be described by an appropriate function, such as the von Bertalanffy (1938) growth equation:

$$W(x) = W_\infty \{1- \exp[-k(x-x_0)]\}^3 \qquad \text{Eq. 13}$$

where W_∞ is the assumed asymptotic weight of an individual, k is some individual growth rate, and x_0 is the hypothetical age of an individual of weight zero. With the substitution of Eqs. 12 and 13 into Eq. 11, integration of Eq. 11 yields:

$$Y = F\, N(x_c)\, W_\infty\, \Sigma\{(U_n \exp[-nk(t_c-t_0)]) / (F + M + nk)\} \qquad \text{Eq. 14}$$

where U_n are integration constants taking on values of 1, -3, 3, and 1 for n = 0, 1, 2, and 3, respectively. Some important assumptions associated with yield models are that: 1) mortality is independent of density with rates remaining constant following recruitment into the fishery; 2) the growth function adequately describes individual growth following recruitment; 3) all individuals of a cohort recruit at the same instant; and 4) no net migration occurs. The effects of environmental stressors can be incorporated into such models through variation in growth rate (k) in Eq. 13, or through changes in mortality occurring beyond age x_c as characterized by N(x). Additionally, reproductive effects can be incorporated implicitly through reduction in the number of recruits entering the fishery at age x_c.

Stock recruitment models describe recruitment of numbers into the fishery, rather than biomass yield to the fishery subsequent to recruitment. The data requirements of such approaches include estimates of stock size and the subsequent number of recruits over time. The general form of such models is similar to the simple exponential growth model (Eq. 1):

$$dN/dt = -MN \qquad \text{Eq. 15}$$

where M is a parameter involving mortality that may incorporate both density-dependent (m_d) and density-independent (m_0) terms (Ricker, 1954):

$$M = m_0 + m_d N(t_s) \qquad \text{Eq. 16}$$

where $N(t_s)$ is the stock size at time of spawning, t_s. Integration of Eq. 15 with Eq. 16 substituted for M yields the number of recruits, R, Δt units of time after spawning:

$$R(t_s + \Delta t) = f \exp(-m_0 \Delta t) N(t_s) \exp[m_d \Delta t N(t_s)] \qquad \text{Eq. 17}$$

where f is the fecundity of a recruited fish. Some assumptions of this model include that: 1) effects of density occur prior to recruitment, with $N(t_s)$ being the important density; 2) fecundity is unaffected by population density or fish size; and 3) no net migration occurs. Contaminant effects could be incorporated into the mortality terms, or in the average fecundity of recruited individuals.

Other models - Classes of models other than the demographic and fisheries approaches already discussed may have potential utility in assessing site-specific risks; four are worth mentioning briefly here. The first of these is based upon statistical description of *time series* of population abundance or biomass parameters. These empirical models attempt to describe temporal fluctuations in population size based largely upon the past behavior of the population and associated environmental data. No explicit parameterization of demographic processes need be involved. Instead, regression analysis is used to estimate coefficients associated with terms for environmental processes deemed important to the temporal fluctuations, as well as with autocorrelative terms. Thus, population size, N, is described as some function, *f,* of environmental parameters, Ψ_i, and past population size with time as:

$$N = f(\Psi_i, N, t) \qquad \text{Eq. 18}$$

Typically, these models are built by the parsimonious inclusion of variables in an iterative fashion, until the fitted model meets certain criteria (for example, the model reasonably describes the behavior of N, and no autocorrelation of regression residuals exists). Impacts ascribable to changes in anthropogenic stressors could be detected through use of intervention terms (if relatively discrete changes in stressor exposure are known to occur at specific times), or through development of transfer functions relating time series of stressor loading or concentration to the time series of population size. Population size may also be forecast into the future once a reasonable model is developed (although not without a great deal of uncertainty). One potential drawback inherent in this approach is the requirement of long series of uniformly spaced data points; such data sets may not exist for the assessment population. Discussion of this and other problems, and of the techniques associated with time series analysis, is given in Box and Jenkins (1976). Examples of the application of this approach to investigation of near-shore pollution loadings and fishery stocks is given in Polgar et al. (1985) and Summers et al. (1985).

As a class, *individually based models* (IBMs) cover an enormous range of specific model formulations (de Roos et al., 1991). The theme common to this class of models is that the basic unit of model formulation is the individual organism, with descriptions of population dynamics simply being the summation of simulated stochastic outcomes for each individual. As a result, IBMs can accommodate a great deal of variation and complexity in the behavior of individuals and their interactions with biotic and abiotic components of the modeled environment, and therefore can reflect a great deal of realism. IBM simulations are computationally intensive, and their formulation can be non-trivial. Choices must be made between simulating in discrete time (in which case the order of biological and environmental "events" within each time step become important), or using event-driven accounting methods with time handled as a continuous variable and population state being updated with each event. Effects of environmental stressors can be incorporated in IBMs directly, using information derived from toxicity tests, time-to-death studies, and any number of other information sources that relate stressor exposure to physiological and individual-level attributes. As an example of this approach, Nisbet et al. (1989) describes an IBM developed for *Daphnia*. Grimm (1999) reviews some 50 IBMs developed for animals.

Vaughan et al. (1984) discuss one form of IBMs, introduced to fishery management by Ursin (1967), which is based upon the *bioenergetics* of individual organisms within the population. Weight-specific rates of important physiological processes are summed to describe changes in biomass (B) of individuals within a cohort as:

$$(1/B)\, dB/dt = C - (R + F + U + P) \qquad \text{Eq. 19}$$

where C is consumption rate, R is respiration rate, F is egestion rate, U is excretion rate, and P is reproductive loss. Typically, each of the terms in Eq. 19 is expanded to include appropriate influencing processes such as temperature, dissolved oxygen concentration, or activity level. Changes in biomass are summed over individuals within a cohort, and some mortality process is introduced. Although fairly data-intensive in that the environmental factors influencing each physiological rate must be quantified, Vaughan et al. (1984) indicate a major advantage of this approach over some others: contaminant stresses may be modeled at a basic, mechanistic level. However, because of the difficulties associated with quantifying the physiological processes of assessment populations, the utility of bioenergetics models, like other data-intensive modeling approaches, may be limited in site-specific assessments to "high priority" assessment endpoints (e.g., for endangered species).

All of the above models are, to a large degree, quantitative in nature. Qualitative population models have also been formulated which may be appropriately applied in this context. Among the more promising is *loop analysis* (Levins, 1975). This approach can be used to examine the behavior of a partially specified system based solely upon the signs of the interactions between various components of that system. In the context of population modeling, the components of the system would be identified as different ages or developmental stages in a fashion similar to that used in population projection matrix models. The methods employed in loop analysis depend upon the equivalence of differential equations near equilibrium, and matrices and life history diagrams (Levins, 1975). Although loop analysis appears to have limited value in exploring the dynamics of populations in isolation, it may be valuable in examining age/stage structure effects resulting from contaminant insult. An advantage of this approach in site-specific risk assessments lies in its minimal data requirements.

4.3 Extrapolating from Toxicity Test Data

Use in risk assessment of any of the extrapolation models described above typically requires quantitative information about the values that model parameters take under different exposure scenarios. Section 5 of this white paper describes some of the techniques that can be used to obtain this information for field populations occupying the site (and associated reference areas). Data to parameterize models can also be obtained from laboratory-based toxicity tests and experiments. In both approaches, the question is one of how to link responses measured in individuals to their population-level consequences. Although the material provided here is most relevant to the demographic modeling approaches described earlier, it applies to a majority of the other model formulations as well.

Assuming no net migration into or out of the population, the necessary and sufficient rates to describe population dynamics are births and deaths within the population (see above). Although the number of offspring produced by an individual female and the probability that an individual will die (or survive) at any particular instant are determined by any number of environmental, physiological, and historical factors, the effects of chemical stressors on these two vital rates can be estimated using toxicity tests and laboratory experiments. Such tests usually are performed in a manner that isolates the effects of the chemical stressor(s) from all other environmental influences through standardization – all non-chemical variables are held constant (within and across like-tests) such that the differences in responses observed across exposure conditions are attributable almost solely to the chemical(s). Such standardization supports understanding of differences among species in their sensitivities to single chemicals (or environmental media in the case of complex mixtures like sediments), and the “potency” of different chemicals to a single species, but results in effects information that likely underestimates effects expected to occur at the site. This is because only direct toxic effects are measured in standard toxicity tests; effects on vital rates that result from changes in species interactions (competition, predation, disease) and the interactions among biological and environmental factors (temperature, water availability, etc.) are purposely minimized. Thus, standard toxicity tests isolate the effects of the chemical (or environmental medium) from all others.

Of course, standard toxicity tests have been developed for a number of different species and exposure conditions (principally duration). Excluding for the moment bioassays designed primarily to measure physiological attributes (biomarkers and bioindicators; but see Maltby et al., 2001 for a description of extrapolation across multiple levels of biological organization), the measurements made in such tests usually include mortality (in short-duration or "acute" tests), but also can include reproduction (in longer-duration or "chronic" tests), and in some cases, changes in individual growth (as measured by size or weight). These first two attributes can be used to estimate the vital rates required by extrapolation models with appropriate consideration of the conditions of toxicity testing and the requirements of the model.

Standard toxicity tests provide information most relevant to the purpose and conditions of the test. Thus, if the test measures acute mortality of neonatal *Daphnia* over a 96-hr exposure period (and assuming all quality assurance requirements have been met), the resulting data adequately describe the rate at which newborn *Daphnia* will survive a 96-hr period of exposure to each concentration tested. It does not, however, directly provide survivorship data for older individuals, nor information about what reproductive effects might be elicited. Standard tests rarely measure both survival and reproductive effects over the full life span of the test species; this creates some practical difficulties with respect to extrapolating test results to population level effects. Yet, with appropriate acknowledgment of their concomitant uncertainties, a number of approaches can be employed that can utilize even minimal toxicity data sets to estimate population responses. The value of any of these approaches depends upon the goals of the risk assessment with respect to desired level of conservatism (safety) in the risk estimated, and the cost and likelihood of obtaining additional data for model parameterization.

- *Using data only as measured* - Probably the least conservative use of standard toxicity data is to model population response as if the chemical (or environmental medium) only affected the life stage tested. For example, the *Daphnia* test data described above might be used to estimate population-level risk if the assumption was that the chemical only affected neonates. Clearly, this analysis approach would underestimate risk if the survivorship of other age classes were also affected, or reproductive effects could occur.

- *Extrapolating effects to other ages or life stages* - Although differences in the sensitivities of age- and stage-classes may exist, the responses of a tested class can be extrapolated to those of untested classes by assuming functional relationships. The simplest relationship to assume is equality, that is, the response of class X is equal to that of class Y. With additional information, more complex relationships could be assumed. By extrapolating test data to the responses of all classes, (presumably) more realistic estimates of population risk can be made. However, two cautions are appropriate here. First, without appropriate knowledge of the relationships among the responses of different classes, errors can be introduced into the analysis which may be difficult to identify and interpret (class X may be more, equal, <u>or</u> less sensitive than class Y, and the resulting influence of risk estimates would be unrecognized). The second is related to the axiom that "an individual can't be killed twice", meaning that younger individuals that die may shift the distribution of individual sensitivities in older age classes to the right (less sensitive) in long-term exposures (typical of site-specific risk problems). Without recognition of this shift, estimates of risk would be relatively conservative.

- *Extrapolating effects to other vital rates* - Similarly, reproductive effects could be extrapolated assuming some relationship between survivorship and fecundity. The cautions offered above are relevant here as well. [As an aside, individual growth effects can also be extrapolated to changes in reproduction and survivorship when their functional relationships (e.g., smaller individuals produce fewer offspring) are known.]

- *Extrapolating effects across test exposure durations* - As indicated earlier, how toxicity test data are used in extrapolation models depends on the requirements of the model itself. A model that describes the population as a sequence of year-long age-classes requires data couched in terms of annual rates. The potential mismatch between model requirements and toxicity test durations and data is obvious, but it can be overcome mathematically (see details in Caswell, 2001, and an example in Munns et al., 1997). Uncertainties associated with using data obtained on time scales different from those required by the model include not only the propagation error inherent in mathematical manipulations, but also those due to differences in the mechanisms of toxicity that may be in operation under different combinations of concentration and exposure duration (i.e., acute versus chronic).

A laboratory-based approach for obtaining the data needed to parameterize extrapolation models that avoids some of the issues above involves performance of a life table response experiment (LTRE). In ecology, life tables are age- or stage-specific schedules of survivorship and reproduction that can be used in demographic extrapolation models to estimate population-level effects associated with some environmental situation or experimental treatment. LTREs can be used to generate these schedules for different chemical exposure conditions. Typically, the duration of LTREs correspond with the life span (or at least life cycle) of the species involved in the experiment, although partial life-cycle tests can, in some situations, provide sufficient information (e.g., Daniels and Allan, 1981). Use of LTREs to help parameterize population models is described in detail in Caswell (2001). Because LTREs evaluate effects on all vital rates for all classes, the need to extrapolate limited data is reduced or eliminated. Unfortunately, "standardized" (in the sense of toxicity tests) LTREs typically do not exist, and LTREs generally are more expensive to perform than are standard toxicity tests.

Expanded descriptions of two examples of extrapolation modeling introduced earlier illustrate the use of toxicity test data to parameterize population models for use in risk assessment. The first involves Kuhn et al.'s (2002) development of a projection matrix model to assist in extrapolating standard toxicity test results for the estuarine amphipod, *Ampelisca abdita*. Used commonly in a number of regulatory and monitoring programs, the standard solid-phase *Ampelisca* test involves exposing approximately 20-day old animals (based on size) for 10 days under static conditions (ASTM, 1993). Kuhn et al. (2002) constructed a model that divided *Ampelisca*'s life cycle into seven 10-day age classes, such that data from the standard test corresponded to survivorship between the second and third (juvenile) age classes. Mean survivorship over the course of a standard test (equivalent to the transition probability between classes 2 and 3 in the model) of Cd-spiked sediment was calculated as the number of individuals in a treatment replicate alive at the end of the test divided by the number at the start of the test, averaged across replicates. [As an aside, the variance observed among replicates can be used to quantify uncertainty in population effect estimates, or to describe probability distributions in probabilistic assessments.] Kuhn et al. (2002) also had conducted a full 70-day experiment to provide demographic information to parameterize a "base" model that reflected *Ampelisca* vital rates in the absence of Cd exposure (i.e., using reference sediment). Under various assumptions of how standard 10-day mortality related to *Ampelisca* vital rates, Kuhn et al. (2002) developed empirical extrapolation relationships between test results and population effects. These relationships can be used to interpret sediment toxicity test data in the context of population-level risks to amphipods at contaminated sediment sites.

The second example is Kuhn et al.'s (2000) use of data obtained from standard life cycle chronic tests to project the effects of some 20 chemicals on mysid shrimp, *Americamysis* (formerly *Mysidopsis*) *bahia*, population dynamics. The life cycle test for *A. bahia* is in most regards comparable to an LTRE, in that mortality and reproduction of test populations are followed on a regular (in this case, daily) basis over the course of a long-term exposure. Survivorship was calculated daily, using the same method as in the *Ampelisca* example, to parameterize a projection matrix model constructed with 24-hr age classes. The fecundity of each age class was calculated as the number of female offspring produced each 24-hr period, divided by the number of females alive during that period. Population models were uniquely parameterized for each treatment level of each chemical, and were used to develop chemical-specific exposure-(population growth rate) response relationships. These relationships could then be used to estimate the population-level effects of any given exposure concentration of each chemical. As mentioned earlier, Kuhn et al. (2001) subsequently demonstrated that the age-classified model, parameterized for the chemical nonylphenol using standard toxicity test data, projected the dynamics of *A. bahia* reasonably well in a multigenerational experiment conducted in the laboratory, adding confidence to the use of the models for predicting population-level risk. [As an aside, Maltby et al. (2001) describe how alternate formulations influence the accuracy of population models in predicting the results of the multigenerational experiment.]

It should be noted that descriptions of the use of toxicity test data in population extrapolations models to this point have been restricted to situations in which exposure is assumed to be constant. Except under certain conditions (e.g., benthic organisms in contaminated sediment), it is unlikely that the risk scenarios evaluated at Superfund and RCRA sites would involve static exposure. More often, organisms are thought to experience variable exposure as they move through the environment, or the environment (e.g., water) moves past them. With appropriate caveats and acknowledgment of uncertainties, toxicity test data can also be used in these situations. One approach for this is to use test data to construct treatment-wise exposure-response curves relating either vital rates or population-level attributes to exposure concentration. These relationships can be used with dynamic simulations of exposure through time or across space to provide more realistic assessments of population risk by dividing time into discrete units, each with a potentially unique exposure concentration which is translated into an effect through the response curve(s). This approach is illustrated in Munns et al. (1996) in an assessment of population risks associated with aquatic sewage sludge disposal. As with all of the uses described here, this approach comes with limiting assumptions that influence interpretation of the analysis.

The information presented above applies to situations in which the assessment population is of the same species as that used in the toxicity test or LTRE. Although their description is beyond the scope of this white paper, approaches exist for extrapolating effects across species when needed. Such extrapolations can be made either at the level of the test data themselves (e.g., mortality of species X extrapolated to mortality of species Y, followed by use of the extrapolated data to model a population of species Y), or at the population level (e.g., mortality of species X used to model a population of species X, followed by extrapolation to a population of species Y). A general description of cross-species extrapolation issues can be found in Munns (2002); more detailed accounts are given in Mayer et al. (1987), Suter (1993) and references cited therein.

4.4 Considering Spatial Scale

The extrapolation modeling approaches considered above typically assume the environment, and how individuals react to it, to be spatially homogeneous. Moreover, considerations of the geographical boundaries chosen when defining the assessment population are conceptually irrelevant to application of the models. However, natural populations exist in landscapes (or waterscapes) that consist of mosaics of habitat type and stressor concentration. The vital rates of individuals residing primarily in one patch (defined in terms of habitat quality and stressor concentration) likely differ from those living in other patches. The dynamics of these local groups, in aggregate, determine the dynamics of the overall population. Furthermore, refugia in the landscape (patches with low concentrations of anthropogenic stressors) can act as population sources for movement of individuals into areas impacted by stressors, potentially ameliorating the effects of those stressors on the local scale. Conversely, impacted areas can act as sinks into which individuals from unimpacted areas move, depressing overall population abundance (see Landis, 2002 for a demonstration of these effects, and Landis, in review for a discussion of the implications of patch dynamics to understanding population-level effects).

Historically, population models incorporating a spatial context have been applied to questions of conservation biology, pest dispersal and invasion, and other population management issues. Their use to address population-level effects of chemical stressors is relatively new, although examples exist in the literature (e.g., Thomas et al., 1990; Sherratt and Jepson, 1993; Pulliam, 1994; Jepson and Sherratt, 1996; see chapters in Barnthouse et al., in review for convenient summaries). Valuable discussions of modeling population dynamics in a spatial context are given by Okubo (1980), Hanski (1991, 1996, 1997, 1999), Hanski and Gilpin (1991), Dunning et al. (1995), and Tilman and Kareiva (1997). Specific model formulations accounting for spatial heterogeneity fall into two broad classes, as described below. Table 2 identifies particular models and their assumptions.

Metapopulation models divide the overall population into subpopulations that communicate with one another through immigration and emigration. Conceptually, subpopulations are treated almost as individuals, with extinction and recolonization driving local dynamics. Such models do not necessarily incorporate habitat quality and stressors explicitly; rather, the environment can be described mathematically as a series of more-or-less identical patches. Metapopulation models can be used to explore the influence of: 1) the size, spacing, and density of patches, 2) rates and forms of movement among patches, and 3) rates of extinction within patches on population persistence and spatial distribution. See Levins (1969) for early considerations of the metapopulation modeling approach.

Spatially-explicit models increase ecological realism by incorporating landscape structure and habitat quality explicitly. These models differ primarily from metapopulation models in that the spatial context of patches is important not only to migration among patches, but also to the internal dynamics of subpopulations within patches. Internal subpopulation dynamics can be modeled by relating aspects of habitat quality to average vital rates, or even by allowing environmental parameters to affect individual physiological and behavioural responses using an individual based modeling approach. Variation in habitat quality can be incorporated using spatially-referenced data sets of land use patterns, landscape characteristics, and stressor distributions. For example, Akçakaya and Atwood (1997) developed a spatially-explicit model of threatened California gnatcatcher (*Polioptila californica californica*) dynamics to support protection of this subspecies and its habitat from land development in southern California, USA. The habitat requirements of *P. c. californica* were determined by relating habitat characteristics to patterns of habitat use. These relationships were used to develop habitat suitability functions, which in turn were used to score the quality of grid cells modeled in the landscape based on geographic information system (GIS) data on land cover and topography. A two-stage stochastic projection matrix model with density dependence was used to model gnatcatcher dynamics within patches. Population-wide vital rates in this model were estimated from existing data sets, independent of habitat suitability, and were allowed to vary randomly to reflect environmental and demographic stochasticity. Akçakaya and Atwood (1997) used their model to relate probabilities of population decline (quasi-extinction) to a variety of environmental and demographic parameters, with an intent of supporting evaluation of various resource management options. More recently, Schumaker (1998) developed a spatially-explicit modeling platform that permits incorporation of GIS-referenced habitat data, specification of functional relationships between habitat quality and demographic response, and specification of rules to describe movement of individuals among patches. Schumaker's model, PATCH, is being updated to include the effects of chemical stressors on vital rates (N. Schumaker, ORD NHEERL, personal communication).

The ability to model populations in a spatially-explicit manner offers the opportunity to address some aspects of the question raised earlier about the scales appropriate for defining the assessment population. When clear boundaries that delimit the population cannot be established, if issues remain relative to the adverse effects of local hazardous waste sites on populations inhabiting large geographic ranges, or uncertainty exists about the extent to which local effects might be ameliorated by immigration from surrounding areas, a modeling analysis that varies the spatial boundaries in the definition of the population may provide some answers. One objective of this type of analysis might be to define the relationship between expected population-level effect and the spatial scale used to delimit the population. This relationship might take the form of a "scale-response" curve (conceptually analogous to a stressor-response curve) that could be used to help understand the ecological significance of site risks. The analysis could provide insights to the severity and extent of adverse effects, as well as the potential for population recovery as local effects are "diluted" by migration into the site from surrounding areas. When weighed against the regulatory requirements, policy considerations, and stakeholder interests driving problem formulation of the assessment, the results of this type of analysis could provide significant support to the decision-making process.

4.5 Considering Model Selection

Given the range of options for linking individual-effects to population response, which model should be used? That choice is situationally-dependent, and should reflect a number of considerations, including the assessment endpoint and how risk estimates will be used in the decision process, the model's intended use, the biology of the assessment population, the availability of information and data describing that biology, the resources available for performing the assessment, and so on. Additionally, the informational requirements of the selected model must be satisfied.

TABLE 2

Population Modeling Approaches Incorporating Spatial Context
(modified from Maltby et al., 2001)

Model	Assumptions	Comments	References/ Applications
Levins' model (and variants)	continuous time, discrete space	breaks habitat into an infinite number of distinct sites; assumes global random dispersal of recruits	Levins (1969) Hanksi (1991, 1996, 1997) Lamberson et al. (1992)
incidence function model	continuous time, discrete space	breaks habitat into a finite number of distinct sites of differing size; assumes localized dispersal of recruits	Hanski (1994)
reaction-diffusion models	continuous time, continuous space	local populations grow ("reaction") and spread through space ("diffusion")	reviewed in Okubo (1980)
cellular-automaton-like models	discrete time, discrete space	divides habitat into contiguous sites ("cells"); assumes dispersal among adjacent sites	Akçakaya and Atwood (1997) Schumaker (1998)

The data requirements of models deserve special consideration in selecting the appropriate model for assessment purposes. The degree to which these requirements can be satisfied determines, at least in part, which models may be successfully employed. In this regard, some of the modeling techniques discussed earlier (e.g. time series analysis) require only that the total abundances of the population or specific life stages be known. However, such approaches do not easily permit incorporation of mechanistic information regarding toxicity or population control, and therefore are of limited use in examining scenarios involving short-term temporal variation in exposure, incremental effects of multiple stressors, or the ramifications of various remediation options. At the other end of the spectrum are formulations such as IBMs which require detailed knowledge of physiological and metabolic processes and how these influence the vital rates of fecundity and survivorship. The realism of these approaches is further enhanced through incorporation of density influences on or stochastic variation in these processes. Such data are difficult to obtain, yet their inclusion into appropriate models permits the most detailed assessments. Recent useful discussions of density dependence relevant to ecological risk assessment can be found in Sabo et al. (2004) and Moe (in review).

Consideration of the ease with which assessment population-specific information can be obtained is important. Does valuable information exist regarding temporal and spatial patterns of abundance? Perhaps the most cost-effective source of such data is historic data bases obtained for other purposes. Can the life history characteristics of critical life stages be determined from field collections or from the literature? Can sensitivity and the modes of toxicity be examined in the laboratory? Such determinations, if they can be made, are usually performed at increased cost. Thus, the availability of information regarding the life history of the species and the dynamics of individual growth and development is crucial to determining model selection. Further, difficulties in obtaining information about populations at hazardous waste sites may affect the ability to evaluate the veracity of estimates of risk.

The model formulation used to examine effects must also permit isolation of the population attribute of interest (as defined by the assessment endpoint) for detailed analysis. Models which do not incorporate parameterization of specific vital rates offer little in the way of examining subtle effects on those rates. It also may be desirable to be able to parameterize aspects of the measured or expected exposure regime experienced as a result of disposal and site conditions, and further, to incorporate the functional form (e.g. linear, threshold) of the response displayed by target attributes. Dynamic computer simulations incorporating both exposure conditions and the resulting population-level effects can go a long way towards satisfying these requisites. Finally, the model selected for application should permit incorporation of the detail of species-specific information available, because such models arguably yield the most realistic assessments of risk. Additional discussion of model selection and the factors to be considered is given in Munns et al. (in review).

4.6 Considering Uncertainty In Extrapolation

Extrapolation and the use of models introduces uncertainty and potential error into risk assessment. Knowledge about the nature and sources of uncertainty associated with extrapolation can improve assessment planning and inform risk managers about how assessment results can be used in decision making. As summarized by Munns (2002), a useful way to think about uncertainty is to partition its sources among three distinct components (Suter, 1993; see Seiler and Alvarez, 1995 for other terminology and definitions): *variability*, also called heterogeneity or stochasticity, *ignorance* or lack of knowledge, and *error*. Variability is a component of all ecological systems, and represents actual differences in the value of a parameter or attribute among units in a (statistical) population. For example, organisms display variability in their susceptibility to chemical stressors, such that some individuals die at lower exposure concentrations than do others, a fact described by the classic exposure (dose)-response relationship. Variability cannot be reduced by taking additional measurements of a parameter, although it can be quantified more accurately using larger sample sizes. Ignorance represents a lack of knowledge about the true value of a parameter that can result from inadequate or imperfect measurement. It can, however, be reduced with the collection of additional data and information. Continuing the toxicity example, there is a true value at any point in time for the mean susceptibility of individuals in the population, and our estimate of that value improves as a function of the number of toxicity tests we run to estimate it.

The final source of uncertainty in risk assessment, error, results from the use of the wrong methods, models, and data in assessment activities. Error can be corrected or minimized by understanding the uses and limitations of various methods and models to answer specific risk questions, and by rigorous attention to assessment quality.

Extrapolation of population response from effects measured in individuals introduces uncertainty primarily through ignorance. This uncertainty arises from a lack of understanding of the processes and mechanisms that determine how populations operate, how they respond to environmental stressors, and how best to model them. For example, substantial evidence exists, and it is general knowledge that, population abundance is influenced by processes that are affected by density. However, the exact mechanisms by which density dependence operates, how best to model those mechanisms, and the values that should be assigned to model parameters involved in the density effect(s) generally are unknown (except in isolated instances). Incorporation of density-dependence into models without this knowledge introduces uncertainty that may not be compensated by supposed increase in realism achieved by that incorporation (see below). Additionally, inferences of population response based on observations made at lower levels of biological organization can err if emergent properties (Harré, 1972; Salt, 1979) control population dynamics. Qualitative changes in the causes and mechanisms of response at different levels of biological organization, also called transmutation (O'Neill, 1979), produce nonlinearities that can confound our ability to predict the nature and magnitude of risks. The existence of nonlinearities and emergent properties potentially limits the usefulness of extrapolation models to assess risk (see, for example, Bella, 2002).

Uncertainty arising from ignorance can be addressed in at least four ways. One is to perform a meta-analysis (Arnqvist and Wooster, 1995) of the type that compares the outputs of two or more models constructed using different assumptions about how population-controlling processes operate. For example, two or more fishery models might be used to make projections, and their resulting outputs used to bound the range of possible effects. Differences in output could also suggest closer examination of model assumptions and approaches, leading to the selection or rejection of specific formulations. Such meta-analyses clearly increase the costs of the assessment, but that added burden might be warranted if the costs associated with making a wrong management decision are high.

An arguably less costly approach to addressing ignorance is to perform sensitivity analyses of model performance and output in which parameter values are varied systematically. In this manner, the influences of particular assumptions and modeled relationships on model output can be evaluated. Attention to those parameters and relationships that most influence model output, say by improving data quality or by conducting additional research to enhance understanding of how biological processes operate, should reduce uncertainties with the model and thereby improve the estimates of risk.

A third approach for reducing ignorance is to strive for enhanced realism in the modeling effort. This can be accomplished in a number of ways. For example, the demographic and fishery approaches described above are deterministic in their general formulation. Introduction of stochastic variation in mechanistic processes or vital rates permits, for example, description of the distribution of population numbers under steady state (Costantino and Desharnais, 1981; Dennis and Costantino, 1988) or impacted conditions, as opposed to single point estimation. With such, different realizations of model outputs are assigned probabilities, offering a means to bound the range of possible outcomes and to identify those that are more likely. [As an aside, characterization of steady state distributions for unimpacted populations would permit estimation of the magnitude of population change needed to detect contaminant impact using inferential statistics.] Of course, introduction of stochasticity in model parameters also addresses the first source of uncertainty mentioned above, that is, variability.

Realism can also be enhanced by incorporating homeostatic and genetic compensatory mechanisms explicitly into the model formulation. An implicit assumption of the population models described above is that individuals are genetically identical to one another. That is, the dynamics of the population are treated as being independent of genetic variation and allele frequency. However, the impact that a given concentration of contaminants will have on a population is related to the fitness of individuals relative to such exposure. Populations displaying allele frequencies near those which accrue maximum fitness under conditions of contaminant insult will show minimum impact relative to native populations. Although long a goal in population biology, marriage of population dynamics theory with population genetics into single models has taken place only relatively recently (e.g., Desharnais and Costantino, 1983). Despite their increased realism, and except for selected applications developed in conservation biology that include effects of inbreeding as population size decreases (e.g., Lacy, 1993), such models are not sufficiently developed to be of much use in the context of site-specific risk assessments. However, genetic adaptation to contaminant insult is an important consideration in Superfund and RCRA assessments that should be explored more fully (Nacci et al., 2002).

With the exception of metapopulation and spatially-explicit formulations, all of the mechanistic models discussed here assume no net migration. Clearly, realistic and precise characterization of population dynamics requires quantification of the rates of both immigration and emigration. In closed populations, such as those bounded by geographical or environmental constraints, these rates are sufficiently low to be treated as zero. However, migration into and out of many populations associated with hazardous waste sites may not be so constrained. In these situations it may be appropriate to parameterize migration as a time or population density-dependent process.

As suggested earlier, model realism can be enhanced by incorporation of density-dependence in fecundity, survivorship, individual growth, and migration rates into those formulations where no such dependency exists. Instances in nature where demographic processes are unaffected by such influences probably are rare (Slobodkin, 1987). Further, introduction of time lags in the realization of density (Marchesseault et al., 1976) as well as other effects arguably result in more realistic models (May, 1973). Levin and Goodyear (1980) explore the behavior of a population projection matrix model with density dependence added. Parameterization of density-dependence (termed compensation in fishery science) is problematic in the formulation of population dynamics models. An overview of the difficulties associated with measurement and parameterization of such affects in fisheries management are given by Goodyear (1980).

One last modification affecting model realism is incorporation of species interactions (e.g., competition, predation). Populations don't exist in isolation; they interact with other populations in the community in ways that can have as profound impacts on their dynamics as do anthropogenic stressors. Yet, many of the formulations described above treat populations as if they were isolated from such interactions. Exceptions to this include those that include fishing or harvesting pressure (where humans are a predator), and IBM formulations that include interspecific competition and predation. Numerous approaches for modeling species interactions and community dynamics exist, description of which is beyond the scope of this white paper. However, the extent to which such approaches improve risk assessment of populations remains to be determined.

The most obvious way to reduce uncertainty arising from both ignorance and error, is to test model outputs against known results. Generally, questions of uncertainty relative to the use of models in risk assessment often devolve to ones of model *verification* and *validation*. These terms have been used inconsistently (and sometimes interchangeably) in the ecological literature, leading to considerable confusion and, we suggest, obfuscation of the important underlying question: "Does the model work as intended?" (Mitro, 2001). Rather than adding to assessment uncertainties by questioning model verification/validation, we recommend that RPMs and site managers seek to understand how models were tested and evaluated in a manner that facilitates understanding of their uses and potential limitations in site-specific risk assessments. The goal of model evaluation is to characterize the relationships between predictions or projections generated from available data and modeling constructs, and the actual population-level responses observed at the site. Field evaluation of the modeling and assessment procedures, and monitoring in general, are complex subjects which cannot be addressed adequately here. However, three points mentioned below are worth considering, as they potentially affect the feasibility of population-level assessments. More detailed discussions of problems associated with monitoring and model evaluation are provided in Cairns et al. (1984) and others. [One of us (MM) currently is conducting an evaluation of the degree to which population models have been tested, and how they performed against their intended use, based on a review of existing literature.]

Biological systems are often characterized by considerable temporal and spatial variation, the degree of which changes across scales of measurement. For instance, local densities of zooplankton vary in response to variation in water circulation processes, such as Langmuir cell formation, but also change in absolute abundance as a result of intrinsic and external regulatory processes. This variation should be adequately addressed through timely application of appropriately designed sampling programs so that changes in population attributes can be ascribed to either contaminant impact or other causes. Inadequate replication in sampling designs may hinder detection of change, let alone proper quantification of that change.

Successful evaluation methods should also incorporate examination of exposure conditions to avoid drawing incorrect conclusions concerning the correspondence between model projections or predictions and responses observed in the field. Potentially confounding effects of multiple stressors should be identified and factored into a reevaluation of expected population-level change. Problems of scale are also cogent here. Evaluation of models used to predict or describe exposure regimes is therefore extremely important to successful evaluation of population modeling assessments.

As a final consideration relative to model testing, such evaluations may only be relevant for those cases in which actual forecasts (as defined earlier) of population responses are made. Risk assessments based upon projections of response, at least at some level, require constancy in environmental conditions. Constancy in natural biological systems is rare indeed. This is not to say that models that project population change cannot be used in dynamic simulation. However, only qualitative correspondence between simulation results and the actual behavior of the population in the field might be expected in those cases.

5. Measuring Population Attributes and Parameters in the Field

Data collected through field measurements can be used to parameterize population models described in the previous section, to evaluate the outputs of those models, or, independent of models, to assess the effects of toxic chemicals at a site directly. In all these cases, there are inherent difficulties in measuring characteristics of natural populations in the field as compared to laboratory settings. The opportunity exists in laboratory studies to have almost complete knowledge of the population: exactly how many individuals are in the population, its age structure, the rate at which individuals are added to or removed from the population (through the processes of birth, death, and migration), and the rate at which population abundance changes. Such complete knowledge is nearly impossible of natural populations, especially animal populations. Statistical sampling is almost always necessary to estimate model parameters, evaluate model outputs, or characterize the status and dynamics of populations in response to chemical exposure. This section describes some techniques for field sampling and data analysis to support estimation of population-level attributes and the parameters used to calculate them, and is organized around the major categories of attributes described in Section 3 (additional information can be found in Suter et al., 2000). It does not, however, provide a step-by-step procedure for conducting field-based assessments of population risk from chemicals at Superfund and RCRA sites.

5.1 Measuring Population Abundance and Density

Estimation of population abundance and population density are two separate yet interrelated problems. If population abundance and the total area or range of distribution are known, then density can be estimated simply by their ratio. However, techniques exist to estimate population abundance without precise knowledge of the area over which the population is distributed. Alternatively, area can be fixed and the density of the animal population in that area can be estimated.

If every individual in a population (or in a defined area) can be detected or captured (i.e., the probability of detection or capture, p, equals 1.0), then a census of the population can be taken to determine its abundance or density. For many plants, larger birds (common loon, trumpeter swan), and mammals (elk, bison) that are highly visible, p often can be assumed to be 1, and a census performed. However, if p is actually less than 1, the census will be biased negatively (i.e., abundance or density will be underestimated). Aquatic animals such as fish, and many birds and mammals are not readily detectable or amenable to capture. For these populations, information from individuals that are seen can be used to make inferences about population size. Two techniques that can be used in this regard are *capture-recapture* and *distance sampling*; capture-recapture studies focus on the estimation of abundance (among other parameters), while distance sampling studies focus on the estimation of density. Because a large and well-developed literature exists on capture-recapture and distance sampling techniques for animals (Otis et al., 1978; White et al., 1982; Seber, 1982, 1986, 1992; Buckland et al., 1993; Schwartz and Seber, 1999), these methods are described here only generally.

Capture-recapture - Capture-recapture techniques are used to estimate abundance. In capture-recapture sampling, individuals are captured and are in some way marked, tagged, or banded (Neilsen, 1992; Murray and Fuller, 2000), such that these individuals can be identified on subsequent recapture occasions. Marks may be unique to the animal or unique to the capture occasion such that capture histories can be identified. This information is used to construct a capture history matrix that comprises a coded list of unique capture histories for two or more sampling occasions and identifies how many individuals have each capture history. For example, if a population of fish were captured and marked on five days and 57 fish were captured on days 2 and 5 (but not on days 1, 3, and 4), then 57 fish would have the unique capture history “01001". A set of capture histories can be analyzed using a suite of models to estimate abundance.

The assumption of population *closure* is important to consider when using capture-recapture to estimate abundance. A population is considered closed if no births, deaths, or migration occur during the sampling period. Different estimators of abundance are used for closed versus open populations. Other important assumptions for capture-recapture studies include: 1) every animal in the population on the j^{th} sampling occasion has the same probability, p_j, of capture (i.e., equal catchability); 2) marks, tags, or bands are neither lost nor overlooked over the course of the study; and 3) all samples are effectively instantaneous relative to the period between samples, and captured individuals are released immediately after sampling. Violations of these assumptions can result in biased or imprecise estimates of abundance.

Violations of the closure assumption can be minimized by sampling over a restricted time period during which births, deaths, and migration do not occur or are minimal. Closure with regards to migration can be ensured for some populations by physically isolating them. For example, blocking nets can be used in small streams to prevent fish from moving into or out of a sample reach. Physically isolating a population is often not feasible, but for short sampling periods, physically open populations can sometimes be treated as closed (Pollock, 1982). For example, sampling for a bird population might be restricted to a time period just prior to nesting, when individuals are not migrating and natural mortality is minimal. If evidence of biological closure is desired for a physically open population, then a multistrata sampling and modeling approach can be used to quantify any losses from the population attributable to death or migration (Hestbeck et al., 1991; Brownie et al., 1993; Mitro and Zale, 2002). Multistrata model can be implemented in the computer program MARK (White and Burnham, 1997).

The goal of capture-recapture techniques is to develop *estimators of abundance*. If the closure assumption is satisfied, a set of closed-population abundance estimators can be developed from the capture histories. Estimators based on multiple capture occasions can be calculated using the computer program CAPTURE (Otis et al., 1978; White et al., 1982; Rexstad and Burnham, 1991). Included are estimators that are robust to departures from the equal catchability assumption: behavioral response to capture, temporal variability in capture probabilities, and heterogeneity among individuals to capture. These estimators typically require three or more capture occasions; the *Lincoln-Petersen estimator* (Ricker, 1975) can be used for two capture occasions and is robust to temporal variability in capture probabilities.

An estimator for abundance used when behavioral response to capture is of concern is the *removal estimator* developed by Zippin (1956). In removal sampling, animals are numerically depleted from the closed population by physical removal or marking on successive capture occasions, such that catch per unit effort decreases proportionately to the number of animals remaining. Animals are not subject to recapture (or marked animals are ignored on subsequent capture occasions), thus avoiding behavioral responses to first capture which may increase or decrease the probability of recapture. Three or more removal occasions are typically required to estimate abundance. A single removal occasion can be used in conjunction with a calibrated model to estimate abundance in a large area by sampling many smaller areas (Mitro and Zale, 2000a). This approach supports a more rapid assessment of population abundance compared to the multiple removal occasions techniques.

If the closure assumption cannot be satisfied, the *Jolly-Seber estimator* for open populations can be used to estimate population abundance from capture-recapture data (Jolly, 1965; Seber, 1965). The Jolly-Seber estimator parameterizes the processes of birth and death, thereby allowing additions to or losses from the population during the sampling period. The Jolly-Seber abundance estimator can be calculated using the computer program POPAN (Arnason et al., 1998).

Sampling effort, probability of capture, and population size all influence the precision of abundance estimates. Estimator precision increases with increasing sampling effort and capture probability. Greater sampling effort translates into higher capture probabilities; as more individuals in a population are captured, more marked individuals are recaptured, providing additional information upon which to base an estimate of abundance. Sampling effort can be increased by sampling an area more intensively or by sampling on more occasions. However, as population size increases, precision decreases because a relatively smaller portion of a population is marked and recaptured for a given amount of effort.

The number of parameters in an estimator also influences abundance estimate precision, as well as bias. The goal is to select an estimator that minimizes both bias and precision, and the most appropriate estimator will depend on the nature of the capture-recapture data available. If the data indicate a violation of the equal catchability assumption, such as temporal variability in capture probabilities, then an estimator parameterized for temporally varying capture probabilities will minimize both bias and precision. Selecting an estimator with fewer parameters will improve precision but introduce bias; an estimator with more parameters will be unbiased but imprecise. Estimator selection can be aided by using information theoretic criteria, likelihood ratio tests, goodness of fit tests, and simulation (Burnham and Anderson, 1998; Mitro and Zale, 2002)

Distance sampling - Distance sampling is used to estimate density, and has been used for populations of birds, mammals (terrestrial and marine), and fish (Buckland et al., 1993). Distance sampling is equally applicable to estimating the density of inanimate objects related to animals, such as bird nests or fish spawning redds (Buckland et al., 1993; Mitro and Zale, 2000b). This technique involves only the sighting of objects (animals, plants, other) along a line transect or around a point transect. In line transect distance sampling, an observer moves along a line transect of a study area and records the location, relative to the line, of each object detected within some distance either side of the line. In point transect sampling, an observer stands at a fixed point in the study area and records the location, relative to the point, of each object detected within some distance around the point.

There are three assumptions necessary for reliable density estimation from line or point transect sampling (Buckland et al., 1993): 1) objects on a transect are detected with certainty; 2) objects are detected at their initial location before any movement in response to the observer; and 3) distances between objects and the transect are measured accurately. Minimizing deviations from these assumptions will minimize bias in density estimates. That said, objects can remain undetected without undermining the validity of the density estimate. Distance sampling theory accommodates decreases in the detectability of objects as their distance from the line or point transect increases (Buckland et al., 1993). Perpendicular distances from a line transect (or distances from a point transect) are “sampled”, and the distances are modeled so that detectability can be estimated. As distance from the line or point transect increases, detectability decreases, allowing estimation of the effective area sampled and hence density.

In addition to the assumptions above, the position of individual objects in the study area ideally should be random relative to the line or point transect. If this condition is met, there is no requirement about how transects are positioned – that is, transects may follow a random or systematic pattern. If individuals are not randomly distributed, the transects should be selected randomly (or systematically from a random starting point). Transect surveys should be designed to avoid sampling effects that are systematic, such as when transects are placed along streams, roads, or some other physical feature that may influence the distribution of objects.

Models for estimating density from line or point transect data are provided in the computer program DISTANCE (Laake et al., 1994). This program uses sample data to parameterize a *detection function*, which is used to estimate density. The detection function describes the distribution of animal locations relative to the line or point transect, and is selected from a series of models comprising a key function and a series expansion, the latter which is used to optimize the key function to fit the sample data

(Buckland et al., 1993). The key function and series expansion that fit the data best can be selected using a combination of *Akaike's information criterion* (AIC) (Akaike, 1973), likelihood ratio tests, and goodness of fit tests. Density estimate precision depends on sample size: the larger the sample size, the more information available to describe the detection function. Buckland et al. (1993) suggest a minimum sample size of 60-80 detected individuals; smaller sample sizes will yield less precise but valid density estimates.

5.2 Measuring Population Growth

Population growth rate is a quantification of the rate at which population abundance changes and can be expressed as an intrinsic rate, r, or a finite rate, λ (see Section 3). If the rate of change is constant over some time period t, then (as the solution to Eq. 1):

$$N_t = N_0 e^{rt} \qquad \text{Eq. 20}$$

and solving for r:

$$r = \frac{1}{t} \cdot \ln\left(\frac{N_t}{N_0}\right) \qquad \text{Eq. 21}$$

Similarly,

$$N_t = \lambda N_{t-1} \qquad \text{Eq. 5}$$

and solving for λ:

$$\lambda = \frac{N_t}{N_{t-1}} \qquad \text{Eq. 22}$$

Because both r and λ can be calculated from ratios of population abundance (Eqs. 22 and 23), direct estimates of the growth rate of field populations can be made using time series of population abundances estimated using the methods described above. Additionally, because population growth rate integrates reproduction and survivorship into a scalar multiplier that defines how a population changes over some time period, it can be calculated for natural populations using estimates of fecundity and survivorship from field data and the demographic extrapolation models described in Section 4. Techniques for both approaches are described below, together with an approach that uses capture-recapture data.

Estimates from time series analysis - A time series of abundance is simply a sequence of population sizes through (preferably equally spaced) time. And over a finite length of time, population growth rate can be estimated simply as the ratio of the final abundance to the initial abundance. As with most field-oriented endeavors, however, estimation of population growth from field data is not as simple as it sounds. Time series analysis for estimating population growth rate is a method dealing with complexities of field data.

Time series analysis uses the complete sequence of population abundances. In terms of bias and precision of the resulting estimate of growth rate, it performs best if abundances are obtained by census or complete enumeration of the population at each time point in the series; observation error is introduced if abundances are estimated. Additionally, process error is introduced if stochastic demographic or environmental forces are operating and reflected in the time series. When observation error predominates, r or λ can be estimated from the slope of a linear regression of the natural log of population size, ln(N), versus time (McCallum, 2000). When process error predominates, abundances may be autocorrelated such that random variations in time periods influence population size in successive time periods (i.e., errors are associated with successive data points) and the estimate of standard error will be an underestimate of the true standard error. If autocorrelation is detected, r or λ can be estimated from the mean in the change in ln(N) between sequential censuses (McCallum, 2000). The Durbin-Watson test can be used to detect autocorrelation (Sen and Srivastava, 1990).

The approach for addressing autocorrelation in time series is to use differences between successive observations as the variable for analysis (McCallum, 2000). For population abundance observations N_0, N_1, . . ., N_n, mean population growth rate can be estimated as:

$$\bar{r} = \left[\sum_{i=1}^{n} \ln\left(\frac{N_i}{N_{i-1}}\right)\right] \Big/ n \qquad \text{Eq. 23}$$

which can be rearranged as:

$$\bar{r} = \left[\ln(N_n) - \ln(N_0)\right] \Big/ n \qquad \text{Eq. 24}$$

Note that only the first and last values of abundance in a time series are used to estimate $\bar{r}$, and the intermediate values do not contribute, which is not a problem if N_0 and N_n are true counts. However, intermediate values will contribute to the estimate of the standard error of mean growth rate. If only N_0 and N_n are used to estimate r, then any error in their estimation can greatly affect the estimate of r.

The average growth rate of the population, $\bar{r}$, is converted to the growth rate of the average population $\hat{r}$ by the equation:

$$\hat{r} = \bar{r} + \frac{\hat{\sigma}^2}{2} \qquad \text{Eq. 25}$$

where:

$$\hat{\sigma}^2 = \frac{\sum (r_i - \bar{r})}{(n - 1)} \qquad \text{Eq. 26}$$

and the estimated standard error of $\hat{r}$ is:

$$se_{\hat{r}} = \sqrt{\hat{\sigma}^2 \left[\frac{1}{t_n} + \frac{\hat{\sigma}^2}{2(n - 1)} \right]} \qquad \text{Eq. 27}$$

where t_n is the number of observations in the time series.

Estimates from demographic matrix population models - The population growth rate λ is equal to the maximum eigenvalue of a projection or transition matrix (Section 4), which can be parameterized with age- or stage-specific estimates of fertility and survival derived from field data. The challenge here is to estimate age- or stage-specific fecundity and survival from field data. If fecundity and survival are constant over time, then the matrix population model is deterministic. If temporal variation in fecundity or survival is observed in the field and estimates of that variation can be derived, then fecundity and survival parameters in the matrix population model can be functions of time and the model is termed stochastic. Population growth rate of a stochastic matrix model is calculated as the average of the maximum eigenvalues for each realization of the projection matrix.

Fertility is a quantification of the per-capita reproductive output by age or stage, or the number of surviving offspring produced per individual per projection interval. The value of fertility used in the projection matrix is a product of: 1) survival of the breeding adult from the beginning of the time interval to reproduction; 2) the reproductive output of an individual; and 3) survival of the offspring from reproduction to the end of the time interval:

$$F_i = S_i^{1-p} \cdot m_i \cdot S_0^p \qquad \text{Eq. 28}$$

where F is fertility, S is survival, and m is reproductive output at age i, and p is the fraction of the time interval after reproduction. If a population is censussed or sampled immediately after reproduction (postbreeding) then $p \rightarrow 0$. If a population is censussed immediately before reproduction (prebreeding) then $p \rightarrow 1$. Field studies are commonly conducted such that p is effectively 0 or 1 and one of the survival terms drops out of the equation for fertility.

Information on reproductive output is a relatively easy demographic parameter to measure and can be obtained by direct observation of a random sample of individuals in each age- or stage-class. Estimates of clutch or litter size from observations of captive-bred animals can also be applied to wild populations (McCallum, 2000), although care must be taken to ensure that holding or culture conditions do not unintentionally influence the estimates. For example, female fish in a gravid state can be sampled to determine the number of eggs produced by age or size class. Many birds have clutch sizes that rarely vary in size. In a prebreeding census, these numbers are then multiplied by the survival rate through the remainder of the first year to obtain an estimate of the number of surviving offspring produced. For example, the clutch size for a bird can be multiplied by rates of hatching success, survival from hatch to fledging, and survival from fledging through the remainder of the first year. In wild populations, an estimate of breeding propensity, or the percentage of females of reproductive age that are breeding, is also required and should be included in the fertility equation, unless it can be assumed that all individuals of reproductive age breed in a given year.

Estimation of *survival* of individuals in a field population is a more difficult problem than the estimation of fertility. Whereas the fate of every individual can be followed in a laboratory population, such information may only be available for some fraction of a wild population, and this information must be used to make inferences about the population as a whole. There are a number of methods of estimating survival that depend on various assumptions of detectability of previously identified individuals. If individuals can be marked and followed with certainty, methods of analysis for *known fate* data can be used. Data for individuals that are marked but that may or may not be observed again can be analyzed using *live-recapture* or *dead-recovery* type models. If animals cannot be marked and later identified, age structure and population size can be used under restrictive assumptions to estimate survival (McCallum, 2000). The three methods involving data from marked animals (known fate, live recaptures and dead recovery) are further discussed below.

Known fate - Animals may be individually tagged with radio transmitters such that the fate of individuals can be known with certainty through the study period. Survival can be estimated without the confounding problem of nuisance parameters such as capture probability, because radio-tagged animals are detected with certainty over time. The data for each time interval of a known fate study are simply the number of tagged animals alive and the number dead. For age-specific survival estimates it is necessary to follow animals tagged at a known age. Known fate data can be analyzed using models in the computer program MARK (White and Burnham, 1997).

Live recaptures - The survival of animals that are marked and released and only reencountered when recaptured alive on subsequent sampling occasions can be estimated using *open population models* of the Cormack-Jolly-Seber type (Cormack, 1964; Jolly, 1965; Seber, 1965; Lebreton et al., 1992). Open population models parameterize the process of survival. Survival rate estimates from live recapture data are actually estimates of apparent survival: the probability that an animal is alive and has remained in the study area such that it is available for recapture (i.e., survival times site fidelity). The assumptions of no tags lost or overlooked, and effectively instantaneous samples relative to periods between samples (see abundance estimation procedures above), still apply to open population models for survival estimation. The assumption of equal catchability of individuals in the population is not as important for survival estimation as it is for abundance estimation because survival is estimated only for marked individuals. It is necessary, however, to assume that the survival of marked and unmarked individuals is the same.

The data for live recaptures is similar to the capture histories described for estimating abundance. These data can be analyzed using models in the computer program MARK (White and Burnham, 1997). Open population models for survival analysis require estimation of survival probabilities and recapture probabilities. These models are flexible in allowing estimation of time and group effects on survival and recapture probability. In other words, these parameters can be allowed to vary or be held constant over any or all time periods or groups. [However, for time-specific models, the estimates of survival and recapture probability for the final time period are confounded.] Groups may be identified, for example, by age, sex, geographic location, or treatment. Two-parameter models can also be used to identify linear trends in a parameter over time, such as a declining survival rate over the time period of the mark-recapture study. Modeling of individual covariate effects is also permitted (e.g., individuals identified by size, condition, body burden).

The approach for analyzing live recapture data should begin with the identification of biologically plausible models, and may include a global model in which all parameters are allowed to vary (Burnham and Anderson, 1998). *Akaike's information criterion* (AIC) can be used to identify the model or models most strongly supported by the data. The AIC is a combination of a likelihood value, which measures the discrepancy between the model and the data, and a penalty term based on the number of parameters in the model (Burnham and Anderson, 1998; McCallum, 2000). As parameters are added to a model, the likelihood value will not increase and may decrease, but the penalty term increases. The model with the lowest AIC has the strongest support from the data and best describes the survival and recapture rates during the study period given the data. Model averaging can be used to obtain unconditional parameter estimates; that is, parameter estimates that are not conditional on one chosen model. In model averaging, parameters are obtained from all tested models weighted by their AIC.

Further statistical tests can be performed to obtain a deeper understanding of the data (Anderson et al., 1995). For example, say AIC is used to select a model and the model incorporates a linear trend in survival over time. Likelihood ratio tests can be used to determine whether or not there is significant time-specificity in survival rates, that the linear trend is significant, and how many parameters should be used to let survival vary by time. A test can also be conducted to determine the significance of the slope in the linear trend. Goodness-of-fit tests can be used to test how well a particular model fits the data.

Dead recoveries - The survival of animals that are marked and released and only reencountered when recovered dead can be estimated using dead recovery models. Dead recovery models estimate survival and either reporting probability (the probability an animal is reported given it has died; White and Burnham, 1997) or recovery probability (a function of both survival and reporting probability; Brownie et al., 1985). In the Brownie et al. (1985) model, which uses recovery probabilities, there is no confounding of parameters for time-specific models. However, modeling a covariate for recovery probabilities (e.g., recovery as a function of animal size) is not always clear given that recovery is a function of both survival and reporting probabilities.

In dead recovery studies (also referred to as band-recovery studies), individuals are banded and recovered at different times, individuals are recovered only one time, and data may be collected by different people. Band-recovery studies are typically used to quantify survival for populations subject to harvest, but can be used in any situation in which individuals are found dead and reported. There can be many more recovery periods than banding periods in band-recovery studies, as compared to capture-recapture studies which require an equal number of marking and recapture periods. The assumptions for band-recovery studies are similar to those for capture-recapture studies: no band or tag loss, no band effect on survival or recovery, and all banded individuals in the same cohort have the same survival and recovery rates.

Dead recovery models can be parameterized for study time or group effects and individual covariate effects as described above for live recapture models. The analysis procedure of identifying biologically plausible models a priori and testing models using AIC, likelihood ratio tests, and goodness of fit tests also applies for dead recovery models.

Direct estimation of λ from capture-recapture data - The population growth rate parameter λ can also be estimated directly from capture-recapture data using a model by Pradel (1996). This model is available in the computer program MARK (White and Burnham, 1997). Capture histories of the live recaptures are required for this model. Population growth rate λ is estimated from a likelihood equation that simultaneously uses recruitment parameters and apparent survival parameters. As in the models described earlier for live recaptures and dead recoveries, the Pradel model for population growth rate can be parameterized for study time, group, and individual covariate effects.

5.3 Measuring Population Structure

The structure of a population may be defined in terms of the distribution by size or age of individuals in a population. A temporally consistent age distribution is a characteristic of a stable population, or a population at equilibrium. Such populations tend to have a larger number of smaller or younger individuals and progressively fewer larger or older individuals. Changes in age distributions over time can indicate instances of excessive mortality (although they can also indicate episodic recruitment events or particularly strong recruitment of year classes). If a population is stressed, say by exposure to a toxic contaminant, the structure of the population may be disrupted. For example, a contaminant that adversely affects reproduction and recruitment may result in a population dominated by larger or older individuals. A contaminant that adversely affects longevity may result in a population dominated by smaller or younger individuals, to a greater extent than in a stable population.

Knowledge of the age or size structure of a population may show how a population has been stressed in the past, but may also show how a population can respond in the short-term if released from the stressor. The short-term or transient dynamics of a population are heavily influenced by the structure of the population (Caswell, 2001). Transient dynamics show what changes can be expected in the short run, such as during a recovery period, whereas asymptotic dynamics of a population show how the population can be expected to change in the long run.

A representative sample of individuals is necessary for characterizing the age or size structure of field population. Many species of fish can be accurately aged by counting annual increments formed in scales, otoliths, or other hard body parts. Such increments are discernable because of seasonal differences in growth rates. Alternatively to direct aging procedures, a length-frequency analysis can be performed to assign ages to length modes. Determining the age of bird or mammal species is more problematic, and usually requires marking individuals at birth (i.e., a known age) and following them over time. However, most species can be easily grouped into stages such as juvenile, sub-adult, and adult stages based on size or other physical characteristics.

5.4 Measuring Population Persistence

Population persistence attributes provide information on the viability of a population, or the likelihood that a population will persist through some period of time. There are a number of different persistence attributes (see Section 3), not all of which are easily estimated for wild populations. However, *population viability analysis* (PVA) is the method for estimating the probability that a population of a certain size will persist for a certain length of time (see Lacy, 1993, 1994). White (2000) describes the data requirements and techniques for PVAs. These include construction of a realistic population model (such as the matrix population model described in Section 4) that includes fertility and survival components that contribute to population dynamics over time. Techniques for estimating these parameters are described above. The model should also include a regulation mechanism, such as density dependence, because persistence will be overestimated for populations allowed to grow indefinitely. Realistic descriptions of regulating mechanisms, and accurate field data to parameterize them, are usually difficult to obtain.

A PVA should also include information about demographic and temporal variation. Demographic variation is important to consider for small populations. When survival and births between time steps are treated as random events, then the number of individuals moving from one stage to the next or entering the population via birth is a random number. Consider a population with an individual survival rate of 0.5. About 50% of the individuals in the population will survive to the next time step on average, but randomness inherent in applying a probability to a finite number of individuals in a population may preclude exactly 50% surviving. Variation in the number surviving can be quite large for small populations, and a series of time steps in which less than 50% actually survive can lead to extinction in small populations versus large populations.

Temporal variation occurs when fertility and survival are random variables that can change over time in response to environmental condition (White, 2000). For example, years with harsh weather conditions can result in lower fertility and survival rates compared to years with mild weather conditions. It can be difficult to obtain field estimates of temporal variance in demographic parameters; long-term studies are typically necessary. However, it is acceptable to borrow information from other studies as they relate to the assessment population, or even from a related species for which long-term data are available. Data from other studies may be used to construct rules relating temporal variance in a population parameter to an environmental variable. Note that it is acceptable to borrow information on the temporal and spatial variance of demographic parameters but not the parameters themselves. Fertility and survival rate estimates from stable or increasing populations are inappropriate for use in models for declining populations.

If a population has a wide geographic distribution (as in Cases 2 and 3 described in Section 2), then the inclusion of spatial variation in the PVA may be necessary (White, 2000). Spatial variation in fertility and survival rates can occur if there is spatial variation in the quality of habitat across a landscape. If movement rates are high, then local populations can become depleted as a result of emigration or replenished as a result of immigration (Stacy and Taper, 1992). High movement rates may therefore result in high persistence. The failure to include such variation in a PVA may result in an underestimate of population persistence.

Individual heterogeneity or variation in fertility or survival rates may also be included in PVAs. This type of variation is typically the most difficult to estimate, especially individual heterogeneity in survival rates. Individual heterogeneity can be quantified using live recapture data that includes individual covariates, such as some measure of body condition associated with each marked individual. Rules can then be constructed relating individual variation in a population parameter to some characteristic of the individual. With appropriate caveats, estimates of individual heterogeneity may be borrowed from other studies or related species. If individual heterogeneity is not included in a PVA then viability may be underestimated.

The major obstacle to PVA is obtaining the data to drive the model (Ruggiero et al., 1994; White, 2000). Population persistence is a stochastic phenomenon (White, 2000), requiring complex models that include demographic, temporal, spatial, or individual variation. A lack of data may suggest that a simpler model should be used, but White (2000) notes that a lack of data means a lack of information – no valid estimate of population persistence is possible without adequate information and there is no reason to believe that an unstudied field population (i.e., lack of information) is any simpler compared to well-studied populations. The predictive accuracy of several PVA approaches as used in conservation biology is evaluated by Brook et al. (2000).

5.5 Considering Scalar Issues

As acknowledged earlier, an issue central to assessment of site-specific risks to populations is the scale or scales at which a population should be evaluated. There is no single natural scale for ecological studies (Levin, 1992). We may investigate the characteristics of what can be termed an assessment population (Text Box 1), but an assessment population is a dynamic concept that needs to be defined in each application. The question then is what spatial and temporal scales define an assessment population. Although relevant both to modeling and field-based assessments, the material below considers scalar issues from the field-based perspective.

The spatial scale at which an assessment population is investigated can be defined based on biological characteristics of the species and the physical characteristics of the landscape in which the assessment population occurs. Some populations may be physically isolated, such as fish in a pond. Some populations may be connected to other populations in a homogenous or fragmented landscape, such as song birds using forest habitat. Quantifying the connectedness of populations over a broad spatial scale is particularly important for estimating population persistence. If immigration and emigration are negligible, then spatial variation in the landscape or habitat may divide a population into smaller subpopulations, which may be more vulnerable to local demographic extinction.

The temporal scale at which an assessment population is investigated is important because different species have different life histories. Sampling design with regard to temporal scale concerns the time increment between samples and the length of the study (or number of time increments sampled). The time increment between samples or observations of a population should not be so great such that important dynamics are unobserved. Time increments should certainly not be longer than the reproductive cycle of a species but may be shorter if survival is related to ontogenetic changes in an individual. The length of an assessment may be limited because of decision-making requirements. However, longevities and the reproductive strategy of the assessment population should be taken into consideration when interpreting study results and in making a decision to borrow information from other long-term investigates. For a short-lived species, multiple generations may be observed in an assessment; for long-lived species, only a fraction of a generation may be observed. Population-level effects on long-lived species may therefore be more difficult to detect in assessments of insufficient duration. Population models may be particularly useful for projecting long-term responses of long-lived species to stressors.

The field methods described earlier for estimating population attributes and demographic parameters, such as abundance, fertility, and survival, can be used at most any scale provided sufficient sampling effort is expended. These methods may also be used at smaller scales and the results extrapolated over larger scales. For example, abundance may be estimated in a random selection of small patches of habitat and extrapolated to areas not sampled (Mitro and Zale, 2000a). Relations between fertility and size or age may be used to estimate fertility for a population.

5.6 Considering Uncertainty in Field Measurements

The identification and quantification of uncertainty is critical to risk assessment. As described throughout this section, uncertainty in the estimation of population attributes and demographic parameters can arise from a number of sources. Variability is also a characteristic of populations themselves, which respond stochastically to demographics and environmental conditions. Variability in population attributes cannot be reduced, but can be estimated. Ignorance in parameter estimates often can be estimated and reduced by increasing sampling effort and therefore the information available to describe those parameters.

The methods described earlier for quantifying population attributes and demographic parameters such as abundance and survival are firmly grounded in statistical theory; estimation of uncertainty in parameter estimates (also called sampling variance) is therefore a straight-forward task. The computer programs CAPTURE, MARK, POPAN, and DISTANCE include estimation of parameter uncertainty. Program MARK also includes a model averaging capability that quantifies a measure of uncertainty in model selection, which results in unconditional parameter estimates (i.e., parameter estimates that are not conditional on one particular model).

If parameter estimates are extrapolated over larger scales, then additional sources of uncertainty have to be considered in addition to the uncertainty in the parameter estimates. A total estimate of uncertainty should include the variation within each parameter estimate, the variation among parameter estimates, and an extrapolation error term for making predictions for areas not sampled (Mitro and Zale, 2000a, 2002).

Estimating the uncertainty or stochasticity of the dynamics of a population requires a time series of estimates for the parameter of interest. If such data are not available for the assessment population, an estimate of variance can be borrowed from another study of the species or from investigations of related species. Estimation of temporal or spatial variation (also called process variation) requires removing the sampling variance from the series of parameter estimates. Burnham et al. (1987) and White (2000) describe a procedure for removing sampling variance from a series of estimates to estimate the underlying process variation.

The uncertainty of parameter estimates can also be incorporated into estimates of population attributes such as persistence (White, 2000). Models to estimate persistence have multiple parameters, each with an associated variance; nonzero covariances may also be present for parameters estimated from the same data (e.g., age-specific survival rates). The variance of a function of parameters can be estimated using the delta method or the parametric bootstrap (Efron and Tibshirani, 1993; White, 2000).

6. Outstanding Research Issues for Superfund and RCRA

In a general sense, the concepts, theories, and methods of population ecology are sufficiently developed to conduct many kinds of site-specific risk assessment. Despite this, key uncertainties remain with respect to population-level risk assessment at Superfund and RCRA sites. Many of these uncertainties have been captured as research needs identified and communicated by the Superfund program in December 2000. Similar research needs have been identified through interactions with other Agency offices and programs. These can be summarized as:

- Develop predictive population models to support site-specific risk assessment for particular species
- Develop methods that allow extrapolation of effects on individuals, as measured in toxicity tests, to populations
- Develop methods that support cross-species extrapolations of effect
- Develop methods and models to estimate effects on populations from exposure to multiple stressors
- Develop methods and models that support diagnosis of causes of adverse effects on populations, and that can be used to understand the relative risks of multiple stressors
- Develop approaches to incorporate spatial and temporal variability in stressor concentrations and distribution in assessment of risks to populations, and to delimit the spatial scales at which such assessments should be performed
- Develop and describe approaches for defining ecological significance, both of effects on populations and of indicators of those effects measured at lower levels of biological organization.

In response to many of these research needs, ORD has developed a draft strategic plan to guide its population-level risk assessment research. The goal of NHEERL's Wildlife Research Strategy (U.S. EPA, 2004) is to develop scientifically valid approaches for assessing risks to wildlife and other populations from multiple stressors. Through a series of EPA's Science Advisory Board reviews and consultations, other EPA peer-reviews, and discussions with Program Offices that reflect the needs listed above, four key areas of research have been identified where advances in the science would be instrumental in improving population-level risk assessment techniques and criteria methodology. These areas include:

1. Extrapolation research that improves the basis for predicting toxicological responses among wildlife and other species and exposure scenarios of concern

2. Coordinated population biology and toxicology research to improve predictions of population dynamics in spatially-explicit habitats

3. Research to advance techniques for assessing the relative risk of chemical and non-chemical stressors on wildlife and other populations

4. Research to define appropriate geographical regions/spatial scales for population-level risk assessments.

Reflecting a conceptual approach to population-level risk assessment that combines aspects of toxicology, population biology, and landscape ecology, the NHEERL strategy is focused on three major research objectives:

1. Develop mechanistically-based approaches for extrapolating toxicological data across species, media, and individual-level response endpoints

2. Develop approaches for predicting population-level responses to stressors. Identify the responses at the individual level that have the greatest influence on population-level responses

3. Develop approaches for evaluating the relative risks from chemical and non-chemical stressors on spatially structured populations across large areas or regions.

The activities communicated in the NHEERL draft research strategy to meet these objectives are designed to produce methods, models, and findings that will provide scientifically-credible approaches to EPA Program and Regional Offices, and States and Tribes for conducting population-level risk assessments and to develop associated criteria.

In addition to ORD's efforts, the need for improved methods for population-level risk assessment has been acknowledged by the broader scientific community. SETAC has taken a lead role in facilitating advancement of the underlying science by organizing an international workshop (Barnthouse et al., in review) to explore the management context and goals for protecting populations, key scientific issues relevant to describing population response to environmental stressors, and the empirical and modeling approaches that can be used to characterize population-level effects. It also developed broad guidance for conducting population-level risk assessment in the context of the Agency's ecological risk assessment guidelines. Working through its Ecological Risk

Assessment Advisory Group, SETAC continues to pursue related projects, including collation of examples of population-level risks assessments conducted at hazardous waste sites. Updates on these projects can be found at http://www.setac.org/eraag/era_pop_index.htm.

7. References

References identified in **bold print** are considered important or key.

Akaike, H. 1973. Information theory as an extension of the maximum likelihood principle. In: Second International Symposium on Information Theory, B. Petrov and F. Csaki, eds. Akadermiai Kiado, Budapest, pp. 267-281.

Akçakaya, H.R. and J.L. Atwood. 1997. A habitat-based metapopulation model of the California gnatcatcher. Conservation Biology 11: 422-434.

Allan, J.D. and R.E. Daniels. 1982. Life table evaluation of chronic exposure of *Eurytemora affinis* (Copepoda) to kepone. Marine Biology 66: 176-184.

Anderson, D.R., G.C. White and K.P. Burnham. 1995. Some specialized risk assessment methodologies for vertebrate populations. Environmental and Ecological Statistics 2: 91-115.

Arnason, A.N., C.J. Schwarz and G. Boyer. 1998. POPAN-5: a data maintenance and analysis system for mark-recapture data. Release 5.0. http://www.cs.umanitoba.ca/~popan/

Arnqvist, G. and D. Wooster. 1995. Meta-analysis: synthesizing research findings in ecology and evolution. Trends in Ecology and Evolution 10: 236-240.

ASTM (American Society for Testing and Materials). 1993. Proposed new standard guide for conducting solid phase 10-day static sediment test with marine and estuarine amphipods. E 1367-92, American Society for Testing and Materials, Philadelphia, PA.

Barnthouse, L.W. 1992. The role of models in ecological risk assessment: a 1990's perspective. Environmental Toxicology and Chemistry 11: 1751-1760.

Barnthouse, L.W. 1993. Population-level effects. In: *Ecological Risk Assessment* (G.W. Suter II, ed.), Lewis Publishers, Boca Raton, FL, pp. 247-274.

Barnthouse, L.W. 1996. Modeling ecological risks of pesticide application: A review of available approaches. Prepared for Office of Pesticide Programs, U.S. Environmental Protection Agency under Interagency Agreement No. 1824-D073-A1 with the U.S. Department of Energy.

Barnthouse, L.W., R.V. O'Neill, S.M. Bartell, and G.W. Suter II. 1986. Population and ecosystem theory in ecological risk assessment. In: *Aquatic Toxicology and Environmental Fate*, ASTM STP 921 (T.M. Poston and R. Purdy, eds.), American Society for Testing and Materials, Philadelphia, PA, pp. 82-96.

Barnthouse, L.W., W.R. Munns, Jr., and M.T. Sorensen (eds.). (in review). *Population-Level Ecological Risk Assessment*, SETAC Press, Pensacola, FL.

Bella, D.A. 2002. Salmon and complexity: challenges to assessment. Journal of Human and Ecological Risk Assessment 8(1): 55-74.

Bellwood, D.R and T.P. Hughes. 2001. Regional-scale assembly rules and biodiversity of coral reefs. Science 292: 1532-1534.

Benton, M.J. and S.I. Guttman. 1990. Effects of allozyme genotype on the survivorship of mayflies (*Stenonem femoratum*) exposed to copper. Journal of the North American Bethological Society 9: 271-276.

Bergeron, J.M., D. Crews and J.A. McLachlan. 1994. PCBs as environmental estrogens: turtle sex determination as a biomarker for environmental contamination. Environmental Health Perspectives 102: 780-781.

Beverton, R.J.H. and S.J. Holt. 1957. On the dynamics of exploited fish populations. United Kingdom Ministry of Agriculture, Fisheries and Food. (Ser. 2) 19: 1-533.

Box, G.E.P. and G. Jenkins. 1976. *Time Series Analysis: Forecasting and Control*. Holden-Day, San Francisco, CA.

Brook, B.W., J.J. O'Grady, A.P. Chapman, M.A. Burgman, H.R. Akçakaya and R. Frankham. 2000. Predictive accuracy of population viability analysis in conservation biology. Nature 404: 385-387.

Brownie, C., D.R. Anderson, K.P. Burnham and D.S. Robson. 1985. Statistical inference from band recovery data: a handbook. Second edition. U. S. Fish and Wildlife Service, Resource Publication 156, Washington, DC.

Brownie, C., J.E. Hines, J.D. Nichols, K.H. Pollock and J.B. Hestbeck. 1993. Capture-recapture studies for multiple strata including non-Markovian transitions. Biometrics 49: 1173-1187.

Buckland, S.T., D.R. Anderson, K.P. Burnham and J.L. Laake. 1993. *Distance Sampling: Estimating Abundance of Biological Populations*. Chapman and Hall, London.

Burnham, K.P. and D.R. Anderson. 1998. *Model selection and inference: a practical information-theoretic approach*. Springer-Verlag, New York, NY.

Burnham, K.P., D.R. Anderson, G.C. White, C. Brownie and K.H. Pollock. 1987. Design and analysis experiments for fish survival experiments based on capture-recapture. American Fisheries Society Monograph 5, Bethesda, MD.

Cairns, V.W., P.V. Hodson and J.O. Nriagu (eds.). 1984. *Contaminant Effects on Fisheries*. Wiley, New York, NY.

Calow, P., R.M. Sibly and V.E. Forbes. 1997. Risk assessment on the basis of simplified life-history scenarios. Environmental Toxicology and Chemistry 16: 1983-1989.

Caswell, H. 1982. Stable population structure and reproductive value for populations with complex life cycles. Ecology 63: 1223-1231.

Caswell, H. 1986. Life cycle models for plants. Lectures on Mathematics in the Life Sciences 18: 171-233.

Caswell, H. 2001. *Matrix Population Models*. 2nd Edition. Sinauer Associates, Inc., Sunderland, MA.

Caswell, H., R.J. Naiman and R. Morin. 1984. Evaluating the consequences of reproduction in complex salmonid life cycles. Aquaculture 43: 123-134.

Chapman, P.M., K. Ho, W.R. Munns, Jr., K. Solomon and M.P. Weinstein. 2002. Issues in sediment toxicity and ecological risk assessment. Marine Pollution Bulletin 44: 271-278.

Connell, J.H. and W.P. Sousa. 1983. On the evidence needed to judge ecological stability or persistence. American Naturalist 121: 789-824.

Cormack, R.M. 1964. Estimates of survival from the resighting of marked animals. Biometrika 51: 429-438.

Costantino, R.F. and R.A. Desharnais. 1981. Gamma distributions of adult numbers for *Tribolium* populations in the region of their steady states. Journal of American Ecology 50: 667-681.

Daniels, R.E. and J.D. Allan. 1981. Life table evaluation of chronic exposure to a pesticide. Canadian Journal of Fisheries and Aquatic Sciences 38: 485-494.

de Kroon, H., A Plaisier, J. von Groenendael and H. Caswell. 1986. Elasticity: the relative contribution of demographic parameters to population growth rate. Ecology 67: 1427-1431.

Demetrius, L. 1975. Natural selection in age-structured populations. Genetics 79: 535-544.

Dennis, B. and R.F. Costantino. 1988. Analysis of steady-state populations with the gamma abundance model: application to *Tribolium.* Ecology 69: 1200-1213.

de Roos, A.M., E. McCauley and W.G. Wilson. 1991. Mobility versus density limited predator-prey dynamics on different spatial scales. Proceedings of the Royal Academy of London (B) 256: 117-122.

Desharnais, R.A. and R.F. Costantino. 1983. Natural selection and density-dependent population growth. Genetics 105: 1029-1040.

Dunning, J.B., D.J. Stewart, B.J. Danielson, B.R. Noon, T.L. Root, R.H. Lamberson and E.E. Stevens. 1995. Spatially explicit population models: current forms and future uses. Ecological Applications 5: 3-11.

Efron, B. and R.J. Tibshirani. 1993. *An Introduction to the Bootstrap*. Chapman and Hall, New York, NY.

Emlen, J.M. 1989. Terrestrial population models for ecological risk assessment: A state-of-the-art review. Environmental Toxicology and Chemistry 8: 831-842.

Emlen, J.M. and E.K. Pikitch. 1989. Animal population dynamics: Identification of critical components. Ecological Modelling 44: 253-273.

Foley, P. 1994. Predicting extinction times from environmental stochasticity and carrying capacity. Conservation Biology 8: 124-137.

Fry, D.M. and C.K. Toone. 1981. DDT-induced feminization of gull embryos. Science 213: 222-224.

Gentile, J.H., S.M. Gentile, N.G. Hairston, Jr. and B.K. Sullivan. 1982. The use of life-tables for evaluating the chronic toxicity of pollutants to *Mysidopsis bahia*. Hydrobiologia 93: 179-187.

Gentile, J.H., S.M. Gentile, G. Hoffman, J.F. Heltshe and N. Hairston, Jr. 1983. The effects of a chronic mercury exposure on survival, reproduction and population dynamics of *Mysidopsis bahia*. Environmental Toxicology and Chemistry 2: 61-68.

Gibbs, P.E. and G.W. Bryan. 1986. Reproductive failure in populations of the dog-whelk, *Nucella lapillus*, caused by imposex induced by tributyltin from anti-fouling paints. Journal of Marine Biological Association of the UK 66: 767-777.

Gillman, M.P. and R. Hails. 1997. *An introduction to ecological modelling: Putting practice into theory*. Blackwell Science, Oxford, UK.

Gillman, M.P. and J. Silvertown. 1997. Population extinction and IUCN categories: the uncertainty of ecological measurement. In: *The Role of Genetics in Conserving Small*

Populations (T.E. Tew, T.J. Crawford, J.W. Spencer, D.P. Stevens, M.B. Usher and J. Warren, eds.) Joint Nature Conservation Committee, Peterborough. p. 155-162.

Gimeno, S., T. Gerritsen, T. Bowmer and H. Komen. 1996. Feminization of male carp. Nature 384: 221-222.

Ginzburg, L.R., L.B. Slobodkin, K. Johnson and A.G. Bindman. 1982. Quasi-extinction probabilities as a measure of impact on population growth. Risk Analysis 2: 171-182.

Gleason, T.R., W.R. Munns, Jr. and D.E. Nacci. 2000. Projecting population-level response of purple sea urchins to lead contamination for an estuarine ecological risk assessment. Journal of Aquatic Ecosystem Stress and Recovery 7: 177-185.

Goldingay, R. and H. Possingham. 1995. Area requirements for viable populations of the gliding marsupial *Petaurus australis*. Biological Conservation 73: 161-167.

Goodyear, C.P. 1980. Compensation in fish populations. In: *Biological Monitoring of Fish* (C.H. Hocutt and J.R. Stauffer, Jr., eds.), D.C. Heath and Company, Lexington, MA, pp. 253-280.

Graham, M. 1935. Modern theory of exploiting a fishery and application to North Sea trawling. Journal du Conseil. Conseil Permanent International pour l'eloration de la mer 10: 264-274.

Gray, L.E., Jr. and W.R. Kelce. 1996. Latent effects of pesticides and toxic substances on sexual differentiation of rodents. Toxicology and Industrial Health 12: 515-531.

Gray, M.A. and C.D. Metcalfe. 1997. Induction of testis-ova in Japanese medaka *Oryzias latipes* exposed to *p*-nonylphenol. Environmental Toxicology and Chemistry 16: 1082-1086.

Grimm, V. 1999. Ten years of individual-based modelling in ecology: what have we learned and what could we learn in the future? Ecological Modelling 115: 129-148.

Guillette, L.J., Jr., T.S. Gross, T.R. Masson, J.M. Matter, H.F. Percival and A.R.Woodward. 1994. Developmental anomalies of the gonad and abnormal sex hormone concentrations in juvenile alligators from contaminated and control lakes in Florida. Environmental Health Perspectives 102: 680-688.

Guillette, L.J., Jr., D.A. Crain, A.A. Rooney and D.B. Pickford. 1995. Organization versus activation: The role of endocrine-disrupting contaminants (EDCs) during embryonic development in wildlife. Environmental Health Perspectives 103(Suppl 7): 157-164.

Haberman, R. 1977. *Mathematical Models*. Prentice Hall, Englewood Cliffs, NJ.

Hamilton, W.D. 1967. Extraordinary sex ratios. Science 156: 477-488.

Hanski, I. 1991. Single-species metapopulation dynamics: Concepts, models and observation. Biological Journal of the Linnean Society 42: 17-38.

Hanski, I. 1994. A practical model of metapopulation dynamics. Journal of Animal Ecology 63: 151-162.

Hanski, I. 1996. Metapopulation ecology. In: *Population Dynamics in Ecological Time and Space* (O.E. Rhodes, R.K. Chesser and M.K. Smith, eds.), University of Chicago Press, Chicago, IL, pp. 13-43.

Hanski, I. 1997. Metapopulation dynamics: From concepts and observations to predictive models. In: *Metapopulation Biology: Ecology, Genetics and Evolution* (I. Hanski and M. Gilpin, eds.), Academic Press, San Diego, CA, pp. 69-72.

Hanski, I. 1999. *Metapopulation Ecology*. Oxford University Press, London, UK.

Hanski, I. and M. Gilpin. 1991. Metapopulation dynamics: A brief history and conceptual domain. Biological Journal of the Linnean Society 42: 3-16.

Harré, R. 1972. *The Philosophies of Science*. Oxford University Press, London, UK.

Harrsion, S. J.F. Quinn, J.F. Baughman, D.D. Murphy and P.R. Ehlich. 1991. Estimating the effects of scientific study on 2 butterfly populations. American Naturalist 137: 222-243.

Harwell, M., J. Gentile, B. Norton and W. Cooper. 1994. Issue paper on ecological significance. In: Ecological Risk Assessment Issue Papers, EPA/630/R-94/009, Risk Assessment Forum, Office of Research and Development, Washington, DC, pp. 2-2 -- 2-49.

Hestbeck, J.B., J.D. Nichols and R.A. Malecki.1991. Estimates of movement and site fidelity using mark-resight data of wintering Canada geese. Ecology 72: 523-533.

Jepson, P.C. and T.N. Sherratt. 1996. The dimensions of space and time in the assessment of ecotoxicological risks. In: *ECOtoxicology: Ecological Dimensions* (D.J. Baird, L. Maltby, P.W. Greig-Smith and P.E.T. Douben, eds.), Chapman & Hall, London, pp. 43-54.

Jobling, S., D. Sheahan, J.A. Osborne, P. Matthiessen and J.P. Sumpter. 1996. Inhibition of testicular growth in rainbow trout *Oncorhynchus mykiss* exposed to estrogenic alkylphenolic chemicals. Environmental Toxicology and Chemistry 15: 194-202.

Johnston, R.K., W.R. Munns, Jr., P.L. Tyler, P. Marajh-Whittemore, K. Finkelstein, K. Munney, F.T. Short, A. Melville and S. Hahn. 2002. Weighing the evidence of ecological risk of chemical contamination in the estuarine environment adjacent to the Portsmouth Naval Shipyard, Kittery, Maine, USA. Environmental Toxicology and Chemistry 21: 182-194.

Jolly, G.M. 1965. Explicit estimates from capture-recapture data with both death and immigration - stochastic model. Biometrika 52: 225-247.

Jones, L.A. and R.A. Hajek. 1995. Effects of estrogenic chemicals on development. Environmental Health Perspectives 103(Suppl 7): 63-67.

Kalmus, H. and C.A.B. Smith. 1960. Evolutionary origin of sexual differentiation and the sex-ratio. Nature 186: 1004-1006.

Kareiva, P., J. Stark and U. Wennergren. 1996. Using demography theory, community ecology and spatial models to illuminate ecotoxicology. In: *ECOtoxicology: Ecological Dimensions* (D.J. Baird, L. Maltby, P.W. Greig-Smith and P.E.T. Douben, eds.), Chapman & Hall, London, pp. 13-23.

Kuhn, A., W.R. Munns, Jr., S. Poucher, D. Champlin and S. Lussier. 2000. Prediction of population-level response from mysid toxicity test data using population modeling techniques. Environmental Toxicology and Chemistry 19: 2364-2371.

Kuhn, A., W.R. Munns, Jr., D. Champlin, R. McKinney, M. Tagliabue, J. Serbst and T. Gleason. 2001. Evaluation of the efficacy of extrapolation population modeling to predict the dynamics of *Americamysis bahia* populations in the laboratory. Environmental Toxicology and Chemistry 20: 213-221.

Kuhn, A., W.R. Munns, Jr., J.R. Serbst, P. Edwards, M.G. Cantwell, T. Gleason, M. Pelletier, and W. Berry. 2002. Evaluating the ecological significance of laboratory response data to predict population-level effects for the estuarine amphipod *Ampelisca abdita*. Environmental Toxicology and Chemistry 21: 865-874.

Laake, J.L., S.T. Buckland, D.R. Anderson and K.P. Burnham. 1994. DISTANCE User's Guide V2.1. Colorado Fish and Wildlife Research Unit, Colorado State University, Fort Collins, CO.

Lacy, R.C. 1993. VORTEX: A computer simulation model for Population Viability Analysis. Wildl. Res. 20: 45-65.

Lacy, R.C. 1993/1994. What is Population Viability Analysis? Primate Conservation 14/15: 27-33.

Lamberson, R.H., R. McKelvey, B.R. Noon and C. Voss. 1992. A dynamic analysis of northern spotted owl variability in a fragmented forest landscape. Conservation Biology 6: 505-512.

Landis, W.G. 2002. Uncertainty in the extrapolation from individual effects to impacts upon landscapes. Human and Ecological Risk Assessment (in press).

Landis, W.G. (in review). The spatial structure of populations and ecological risk assessment. In: *Population-Level Ecological Risk Assessment* (L.W. Barnthouse, W.R. Munns, Jr. and M.T. Sorensen, eds.), SETAC Press, Pensacola, FL.

Lavie, B. and E. Nevo. 1982. Heavy metal selection of phosphoglucose isomerase allozymes in marine gastropods. Marine Biology 71: 17-22.

Lavie, B., E. Nevo and U. Zoller. 1984. Differential ability of phosphoglucose isomerase allozyme genotypes of marine snails in nonionic detergent and crude oil-surfactant mixtures. Environmental Research 35: 270-276.

Lebreton, J.-D., K.P. Burnham, J. Clobert and D.R. Anderson. 1992. Modeling survival and testing biological hypotheses using marked animals: a unified approach using case studies. Ecological Monographs 62: 67-118.

Lefkovitch, L.P. 1965. The study of population growth in organisms grouped by stages. Biometrics 21: 1-18.

Leslie, P.H. 1945. On the use of matrices in certain population mathematics. Biometrika 33: 183-212.

Leslie, P.H. 1948. Some further notes on the use of matrices in population dynamics. Biometrika 35: 213-245.

Levin, S.A. 1992. The problem of pattern and scale in ecology. Ecology 73: 1943-1967.

Levin, S.A. and C.P. Goodyear. 1980. Analysis of an age-structured fishery model. Journal of Mathematical Biology 9: 245-274.

Levins, R. 1968. *Evolution in Changing Environments*. Monographs in Population Biology 2, Princeton University Press, Princeton, NJ.

Levins, R. 1969. *Some demographic and genetic consequences of environmental heterogeneity for biological control.* Princeton University Press, Princeton, NJ.

Levins, R. 1975. Evolution in communities near equilibrium. In: *Ecology and Evolution of Communities* (C.L. Martin and J.M. Diamond, eds.), Belknap Press, Cambridge, MA, pp.16-50.

Lewis, E.G. 1942. On the generation and growth of a population. Sankhya 6: 93-96.

Lewontin, R.C. 1969. The meaning of stability. In: *Diversity and Stability in Ecological Systems*, Brookhaven Symposium in Biology No. 22, pp. 13-24.

Lotka, A.J. 1925. *Elements of Physical Biology*. Williams and Wilkins Co., Baltimore, MD.

Maltby, L, T.J. Kedwards, V.E. Forbes, K. Grasman, J.E. Kammenga, W.R. Munns, Jr., A.H. Ringwood, J.S. Weis and S.N. Wood. 2001. Linking individual-level responses and population-level consequences. In: *Ecological Variability: Separating Natural from Anthropogenic Causes of Ecosystem Impairment* (D.J. Baird and G.A. Burton, Jr., eds.), SETAC Special Publication Series, Society of Environmental Toxicology and Chemistry, Pensacola, FL, pp. 27-82.

Malthus, T.R. 1798. *An Essay on the Principle of Population as It Affects the Future Improvement of Society*. Johnson, London.

Marchesseault, G.D., S.B. Saila and W.J. Palm. 1976. Delayed recruitment models and their application to the American lobster (*Homarus americanus*) fishery. Journal of the Fishery Research Board of Canada 34: 1779-1787.

Marshall, J.S. 1962. The effects of continuous gamma radiation on the intrinsic rate of natural increase of *Daphnia pulex*. Ecology 43: 598-607.

May, R.M. 1973. *Stability and Complexity in Model Ecosystems*. Princeton University Press, Princeton, NJ.

May, R.M. 1976. Simple mathematical models with very complicated dynamics. Nature 261: 459-467.

Mayer, F.L., C.H. Deans and A.G. Smith. 1987. Inter-taxa correlations for toxicity to aquatic organisms. EPA/600/X-87/332, US Environmental Protection Agency, Office of Research and Development, Gulf Breeze, FL.

Maynard Smith, J. 1978. *The Evolution of Sex*. Cambridge University Press, New York.

McCallum, H. 2000. *Population parameters: estimation for ecological models*. Blackwell Science, Malden, MA.

Michod, R.E. and W.W. Anderson. 1980. On calculating demographic parameters from age frequency data. Ecology 61: 265-269.

Mitro, M.G. 2001. Ecological model testing: verification, validation, or neither? Bulletin of the Ecological Society of America 82: 235-237.

Mitro, M.G. and A.V. Zale. 2000a. Predicting fish abundance using single-pass removal sampling. Canadian Journal of Fisheries and Aquatic Sciences 57: 951-961.

Mitro, M.G. and A.V. Zale. 2000b. Use of distance sampling to estimate rainbow trout redd abundances in the Henry's Fork of the Snake River, Idaho. Intermountain Journal of Sciences 6: 223-231.

Mitro, M.G. and A.V. Zale. 2002. Estimating abundance of age-0 rainbow trout by mark-recapture in a medium-sized river. North American Journal of Fisheries Management 22: 188-203.

Moe, S.J. (in review). Density dependence in ecological risk assessment. In: *Population-Level Ecological Risk Assessment* (L.W. Barnthouse, W.R. Munns, Jr. and M.T. Sorensen, eds.), SETAC Press, Pensacola, FL.

Moore, C.G. and J.M. Stevenson. 1991. The occurrence of intersexuality in harpacticoid copepods and its relationship to pollution. Marine Pollution Bulletin 22: 72-74.

Munns, W.R., Jr. 1988. Feasibility of applying population models to impact assessment at offshore disposal sites. EPA/600/11-88/381, Office of Research and Development Environmental Research Laboratory, Narragansett, RI.

Munns, W.R., Jr. 2002. Axes of extrapolation in risk assessment. Human and Ecological Risk Assessment 8: 19-29.

Munns, W.R., Jr., H.A. Walker, J.F. Paul and J.H. Gentile. 1996. Ocean disposal at the 106-Mile Deepwater Dumpsite as an ecological risk assessment case study. Journal of Marine Environmental Engineering 3: 279-297.

Munns, W.R., Jr., D.E. Black, T.R. Gleason, K. Salomon, D. Bengtson and R. Gutjahr-Gobell. 1997. Evaluation of the effects of dioxin and PCBs on *Fundulus heteroclitus* populations using a modeling approach. Environmental Toxicology and Chemistry 16: 1074-1081.

Munns, W.R., Jr., W.N. Beyer, W.G. Landis and C. Menzie. 2002. What is a population? SETAC Globe 3: 29-31.

Munns, W.R., Jr., J. Gervais, A.A. Hoffman, U. Hommen, D.E. Nacci, M. Nakamaru, R. Sibly, and C.J. Topping. (in review). Modeling approaches to population-level ecological risk assessment. In: *Population-Level Ecological Risk Assessment* (L.W. Barnthouse, W.R. Munns, Jr. and M.T. Sorensen, eds.), SETAC Press, Pensacola, FL.

Murray, D.L. and M.R. Fuller. 2000. A critical review of the effects of marking on the biology of vertebrates. In: *Research Techniques in Animal Ecology: Controversies and Consequences* (L. Boitani and T.K. Fuller, eds.), Columbia University Press, New York, NY, pp. 15-64.

Nacci, D.E., T.R. Gleason, R. Gutjahr-Gobell, M. Huber and W.R. Munns, Jr. 2002. Effects of chronic stress on wildlife populations: a modeling approach and case study. In: *Coastal and Estuarine Risk Assessment: Risk on the Edge* (M.C. Newman, M.H. Roberts, Jr. and R.C. Hale, eds.), CRC Press/Lewis Publishers, New York, NY, pp. 247-272.

Neilsen, L.A. 1992. Methods of marking fish and shellfish. American Fisheries Society, Special Publication 23, Bethesda, MD.

Nevo, E., R. Noy, B. Lavie, A. Bieles and S. Muchtar. 1986. Genetic diversity and resistance to marine pollution. Biological Journal of the Linnean Society 29: 139-144.

Nisbet, R.M., W.S.C. Gurney, W.W. Murdoch and E. McCauley. 1989. Structured population models: a tool for linking effects at individual and population level. Biological Journal of the Linnean Society 37: 79-99.

Okubo, A. 1980. *Diffusion and Ecological Problems: Mathematical Models*. Springer-Verlag, Berlin.

O'Neill, R.V. 1979. Transmutation across hierarchical levels. In: *System Analysis of Ecosystems* (G.S. Innis and R.V. O'Neill, eds.), International Cooperative Publishing House, Fairland, MD, pp. 59-78.

Otis, D.L., K.P. Burnham, G.C. White and D.R. Anderson. 1978. Statistical inference from capture data on closed animal populations. Wildlife Monographs 62: 1-135.

Pastorok, R.A., S.M. Bartell, S. Ferson and L.R. Ginzburg. 2002. *Ecological modeling in risk assessment: chemical effects on populations, ecosystems, and landscapes*. CRC Press, Lewis Publishers, Boca Raton, FL.

Pella, J.J. and P.K. Thomlinson. 1969. A generalized stock production model. Bulletin of the Inter-American Tropical Tuna Committee 14: 420-496.

Pesch, C.E., W.R. Munns, Jr. and R. Gutjahr-Gobell. 1991. Effects of contaminated sediment on life history traits and population growth rate of *Neanthes arenaceodentata* (Polychaeta: Nereidae) in the laboratory. Environmental Toxicology and Chemistry 10: 805-815.

Pianka, E.R. 1974. *Evolutionary Ecology*. Harper & Row, NY.

Polgar, T.T., J.K. Summers, R.A. Cummins, K.A. Rose and D.G. Heimbuch. 1985. Investigation of relationships among pollutant loadings and fish stock levels in northeastern estuaries. Estuaries 8: 125-135.

Pollock, K.H. 1982. A capture-recapture design robust to unequal probability of capture. Journal of Wildlife Management 46: 752-757.

Pradel, R. 1996. Utilization of capture-mark-recapture for the study of recruitment and population growth rate. Biometrics 52: 703-709.

Pulliam, H.R. 1994. Incorporating concepts from population and behavioral ecology into models of exposure to toxins and risk assessment. In: *Wildlife Toxicology and Population Modeling* (R.J. Kendall and T.E. Lacher, Jr., eds.), Lewis Publishers, Boca Raton, FL, pp. 13-26.

Rexstad, E. and K. Burnham. 1991. User's guide for interactive program CAPTURE. Colorado Cooperative Fish and Wildlife Research Unit, Colorado State University, Fort Collins, CO.

Ricker, W.E. 1954. Stock and recruitment. Canadian Journal of Fisheries and Aquatic Sciences 11: 559-623.

Ricker, W.E. 1975. Computation and interpretation of biological statistics of fish populations. Bulletin of the Fisheries Research Board of Canada 191.

Ruggiero, L.F., G.D. Hayward and J.R. Squires. 1994. Viability analysis in biological evaluations - concepts of population viability analysis, biological population, and ecological scale. Conservation Biology 8: 364-372.

Sabo, J.L, E.E. Holmes and P. Kareiva. 2004. Efficacy of simple viability models in ecological risk assessment: Does density dependence matter? Ecology 85: 328-341.

Salt, G.W. 1979. A comment on the use of the term *emergent properties*. American Naturalist 113: 145-148.

Samson, F.B., F. Perez-Trejo, H. Salwasser, L.F. Ruggiero and M.L. Shaffer. 1985. On determining and managing minimum population size. Wildlife Society Bulletin 14: 425-433.

Schaefer, M.B. 1954. Some aspects of the dynamics of populations important to the management of the commercial marine fisheries. Bulletin of the Inter-American Tropical Tuna Committee 1: 25-56.

Schaefer, M.B. 1957. A study of the dynamics of the fishery for yellowfin tuna in the eastern tropical Pacific Ocean. Bulletin of the Inter-American Tropical Tuna Committee 2: 245-268.

Schumaker, N.H. 1998. A users guide to the PATCH model. EPA/600/R-98/135, Office of Research and Development, Corvallis, OR. Model available at: http://www.epa.gov/wed/pages/models.htm.

Schwartz, C.J. and G.A.F. Seber. 1999. Estimating animal abundance: review III. Statistical Science 14: 427-456.

Seber, G.A.F. 1965. A note on the multiple-recapture census. Biometrika 52: 249-259.

Seber, G.A.F. 1982. *The Estimation of Animal Abundance and Related Parameters*. 2nd edition. Arnold, London.

Seber, G.A.F. 1986. A review of estimating animal abundance. Biometrics 42: 267-292.

Seber, G.A.F. 1992. A review of estimating animal abundance II. International Statistical Review 60: 129-166.

Seiler, F.A. and Alvarez J.L. 1995. Toward a new risk assessment paradigm: variability, uncertainty, and errors. Technology Journal of the Franklin Institute 332A: 221-235.

Sen, A. and M. Srivastava. 1990. *Regression analysis: theory, methods, and applications*. Springer-Verlag, New York, NY.

Shaffer, M.L. and F.B. Samson. 1985. Population size and extinction: A note on determining critical population sizes. American Naturalist 125: 144-152.

Sherratt, T.N. and P.C. Jepson. 1993. A metapopulation approach to modelling the long-term impacts of pesticides on invertebrates. Journal of Applied Ecology 30: 696-705.

Sherratt, T.N., G. Roberts, P. Williams, M. Whitfield, J. Briggs, N. Shillabeer and S.J. Maund. 1999. A life-history approach to predicting the recouvery of aquatic invertebrate populations after exposure to xenobiotic chemicals. Environmental Toxicology and Chemistry 18: 2512-2518.

Sibly, R.M. 1996. Effects of pollutants on individual life histories and population growth rates. In: *Ecotoxicology. A Hierarchical Treatment* (M.C. Newman and C.H. Jagoe, eds.), Lewis Publishers, Boca Raton, FL, pp. 197-223.

Sibly, R.M. 1999. Efficient experimental designs for studying stress and population density in animal populations. Ecological Applications 9: 496-503.

Sinclair, A.R.E. 1996. Mammalian populations: Fluctuation, regulation, life history theory and their implications for conservation. In: *Frontiers of Population Ecology* (R.B. Floyd, A.W. Sheppard and P.J. De Barro, eds.), Commonwealth Scientific and Industrial Research Organisation (CSIRO), Australia, pp. 127-154.

Sissenwine, M.P., B.E. Brown and J. Brennan-Hoskins. 1979. Brief history and state of the art of fish production models and some applications to fisheries off the northeastern United States. In: *Climate and Fisheries: Proceedings from a Workshop*. University of Rhode Island Center for Ocean Management Studies, Kingston, RI.

Slobodkin, L.B. 1987. Background to a discussion of compensation. In: *Mechanisms of Compensatory Response of Fish Populations: Workshop Proceedings*, Electric Power Research Institute, pp. 2-1 -- 2-17.

Snell, T.W. and M. Serra. 2000. Using probability of extinction to evaluate the ecological significance of toxicant effects. Environmental Toxicology and Chemistry 19: 2357-2363.

Stacy, P.B. and M. Taper. 1992. Environmental variation and the persistence of small populations. Ecological Applications 2: 18-29.

Stein, B.A., L.S. Kutner and J.S. Adams. 2000. *Precious heritage: The status of biodiversity in the United States*. Oxford University Press, Oxford, UK.

Summers, J.K., T.T. Polgar, J.A. Tarr, K.A. Rose, D.G. Heimbuch, J. McCurley, R.A. Cummins, G.F. Johnson, K.T. Yetman and G.T. DiNardo. 1985. Reconstruction of long-term time series for fisheries abundance and estuarine pollution loadings. Estuaries 8: 114-124.

Suter, G.W., II (ed.). 1993. *Ecological Risk Assessment*. Lewis Publishers, Boca Raton, FL.

Suter, G.W., II, B.E. Sample, D.S. Jones and T.L. Ashwood. 1994. Approach and strategy for performing ecological risk assessments for the U.S. Department of Energy's Oak Ridge Reservation: 1994 revision. ES/ER/TM-33/R1, Oak Ridge National Laboratory, Oak Ridge, TN.

Suter, G.W., II, R.A. Efroymson, B.E. Sample and D.S. Jones. 2000. *Ecological risk assessment for contaminated sites*. Lewis Publishers, Boca Raton, FL.

Sutherland, W.J. and J.D. Reynolds. 1998. Sustainable and unsustainable exploitation. In: *Conservation Science and Action* (W.J. Sutherland, ed.), Blackwell Science, Oxford, UK, pp. 90-115.

Thacker, J.R.M. and P.C. Jepson. 1993. Pesticide risk assessment and non-target invertebrates: Integrating population depletion, population recovery and experimental design. Bulletin of Environmental Contamination and Toxicology 51: 523-531.

Thomas, C.F.G., E.H.A. Hol and J.W. Evert. 1990. Modelling the diffusion component of dispersal during the recovery of a population of linyphiid spiders from exposure to an insecticide. Functional Ecology 4: 357-368.

Thomlinson, R. 1965. *Population Dynamics - Causes and Consequences of World Demographic Change*. Random House, New York, NY.

Tilman, D. and P. Kareiva. 1997. *Spatial Ecology. The Role of Space in Population Dynamics and Interspecific Interactions*. Monographs in Population Biology 30, Princeton University Press, Princeton, NJ.

Ursin, E. 1967. A mathematical model of some aspects of fish growth, respiration, and mortality. Canadian Journal of Fisheries and Aquatic Sciences 24: 2355-2443.

U.S. EPA. 1996. Summary report for the workshop on Monte Carlo analysis. EPA/630/R-96/010, Risk Assessment Forum, Washington, DC.

U.S. EPA. 1998. Guidelines for ecological risk assessment. EPA/630/R-95/002F, Risk Assessment Forum, Washington, DC.

U.S. EPA. 2003. Generic ecological assessment endpoints (GEAEs) for ecological risk assessment. EPA/630/P-02/004F, Risk Assessment Forum, Washington, DC.

U.S. EPA. 2004. NHEERL Wildlife Research Strategy. EPA 600/R-04/050, Office of Research and Development, RTP, NC.

U.S. Navy. 2000. Estuarine ecological risk assessment for Portsmouth Naval Shipyard, Kittery, Maine. Volume I: Technical report. Marine Environmental Support Office, Space and Naval Warfare Systems Command, San Diego, CA.

Vaughan, D.S., R.M. Yoshiyama, J.E. Breck and D.L. DeAngelis. 1984. Modeling approaches for assessing the effects of stress on fish populations. In: *Contaminant Effects on Fisheries* (V.W. Cairns, P.V. Hodson and J.O. Nriagu, eds.), Wiley, New York, NY, pp. 259-278.

Verhulst, P.F. 1838. Notice sur la loi que la population pursuit dans son accroissement. Correspondences in Mathematics and Physics 10: 113-121.

von Bertalanffy, L. 1938. A quantitative theory of organic growth. Human Biology 10: 181-223.

Weins, J.A. 1977. On competition and variable environments. American Scientist 65: 590-597.

Weins, J. 1996. Coping with variability in environmental impact assessment. In: *ECOtoxicology: Ecological Dimensions* (D.J. Baird, L. Maltby, P.W. Greig-Smith and P.E.T. Douben, eds.), Chapman & Hall, London, pp. 55-70.

Werner, P.A. and H. Caswell. 1977. Population growth rates and age versus stage distribution models for teasel (*Dipsacus sylvestris* Huds.). Ecology 58: 1103-1111.

White, G.C. 2000. Population viability analysis: data requirements and essential analyses. In: *Research Techniques in Animal Ecology: Controversies and Consequences* (L. Boitani and T.K. Fuller, eds.), Columbia University Press, New York, NY, pp. 288-331.

White, G.C. and K.P. Burnham. 1997. Program MARK—survival estimation from populations of marked animals. http://www.cnr.colostate.edu/~gwhite/mark/mark.htm

White, G.C., D.R. Anderson, K.P. Burnham and D.L. Otis. 1982. Capture-recapture and removal methods for sampling closed populations. Los Alamos National Laboratory, LA 8787-NERP, Los Alamos, NM.

Zippin, C. 1956. An evaluation of the removal method of estimating animal populations. Biometrics 12: 163-189.

APPENDIX

ECOLOGICAL RISK ASSESSMENT SUPPORT CENTER REQUEST FORM

Request #0006: How do we make the connection between individual measurements and population effects; what do we measure?

Requestor: Clarence A. Callahan, Region 9

Problem Statement: How do we make the connection between individual measurements and population effects; what do we measure? For example, if we measure mortality, what can we say about the ultimate effect on the population? What can we say about the effect on a population of amphipods if we perform a laboratory test using amphipods and measure mortality? How do we make the connection between mortality on the individual and an impact on the population? Can this relationship be shown in both theory and practice? Of course, the question relates to invertebrates (tending to R-strategy), other animals (tending to *K*-strategy) as well as plants.

Background: For several years the general assumption was that toxicity endpoints, even acute effects, e.g. mortality, were representative of population effects. This is especially true for invertebrate receptors and assumed for receptors in general. Over the last 10 to 15 years, there seems to have been a leap of faith that what is true about invertebrates is also true with larger animals. The relationship between measurements at the individual level and effects at the population level is not generally provided in any ERAs that I have read. There is generally a statement that "population effects will be addressed" and then a presentation of acute and sometime chronic testing in single species laboratory tests. In other words, toxicity measures using single species tests are assumed to be indicative of population effects although the measurement was mortality of individuals in a laboratory test. This was assumed to be true even when the laboratory test species was not the same or sometimes even close to the receptor population assumed to be "represented" in the assessment.

If we truly are assessing population level effects, how do we make the connection between individual measurements and population effects? If we measure mortality, what can we say about the ultimate effect on the population? What exactly will be the effect on a population of amphipods if we perform a laboratory test using amphipods and measure mortality? How do we make the connection between mortality on the individual and the impact on the population? This, I believe has to be shown in both theory and practice if we continue to maintain this position.

For larger animals, I don't believe that we are capable of doing justice to anything called an ecological risk assessment at a population level. Then what do we do for sites where larger animals are potentially impacted? What kinds of studies do we perform? Do we measure individual effects and make that work somehow, or do we continue to assume that these individual measures relate to population effects?

Expected Outcome: Some text to explain the connection between laboratory testing and field results. Further explanations about the relationship between the measurement of mortality and the impact to populations. A list of measurements for laboratory tests that can be related to population effects.