Recommendations for Improving Recovery Criteria under the US Endangered Species Act

DANIEL F. DOAK, GINA K. HIMES BOOR, VICTORIA J. BAKKER, WILLIAM F. MORRIS, ALLISON LOUTHAN, SCOTT A. MORRISON, AMANDA STANLEY, AND LARRY B. CROWDER

Recovery criteria, the thresholds mandated by the Endangered Species Act that define when species may be considered for downlisting or removal from the endangered species list, are a key component of conservation planning in the United States. We recommend improvements in the definition and scientific justification of recovery criteria, addressing both data-rich and data-poor situations. We emphasize the distinction between recovery actions and recovery criteria and recommend the use of quantitative population analyses to measure the impacts of threats and to explicitly tie recovery criteria to population status. To this end, we provide a brief tutorial on the legal and practical requirements and constraints of recovery criteria development. We conclude by contrasting our recommendations with other alternatives and by describing ways in which academic scientists can contribute productively to the planning process and to endangered species recovery.

Keywords: conservation, ecology, population biology, extinction, endangered species

Over the past 20 years, ecologists and conservation biologists have conducted multiple reviews of the US Endangered Species Act (ESA) focused on legal, policy, and especially scientific elements of the act's implementation (e.g., Tear et al. 1993, 1995, Foin et al. 1998, Boersma et al. 2001, Gerber and Hatch 2002, Hoekstra et al. 2002, Lawler et al. 2002, Morris et al. 2002, Moyle et al. 2003, Scott et al. 2005, Gibbs and Currie 2012). These reviews have shown numerous shortcomings in the effectiveness and scientific basis of recovery plans and recovery criteria and have suggested just as many remedies. In response to these academic reviews and to court decisions interpreting the ESA, the two government agencies that implement the act (the US Fish and Wildlife Service [USFWS] and the National Marine Fisheries Service [NMFS]; henceforth, *the services*) have continued to update their procedures for recovery planning (NMFS and USFWS 2010).

Despite these efforts, recent reviews of the ESA's implementation have still shown little improvement in key metrics of scientific rigor, including the clear articulation and biological justification of recovery criteria (Neel et al. 2012). This situation prompted us to convene a workshop to find pragmatic ways to improve this central part of ESA recovery planning. To increase the odds that our recommendations would have traction, we sought to understand the viewpoints of representatives from many parts of the conservation community and to focus on one key element of the ESA—recovery criteria and their use—rather than conducting a general critique of the act or its implementation.

We focus on recovery criteria for three reasons. First, they specify the conditions under which a species may be considered for downlisting (i.e., being moved from *endangered* to *threatened* status) or delisting (i.e., being removed from ESA protection), thereby defining what characteristics the services expect a population to exhibit once it reaches a state of recovery. Criteria therefore serve as a structuring element for a recovery plan as a whole and guide the actions of government agencies and other entities. Second, the ESA stipulates that recovery criteria be "measurable and objective" and that delisting decisions be based on "the best scientific and commercial data available" (16 U.S.C. § 1533); both requirements inject a primary role for science, although exactly how recovery standards are to be defined or assessed is left unclear. Finally, a vast amount has been written about assessing extinction risk, establishing targets for healthy populations in the face of harvest and habitat loss, analyzing the consequences of population size and connectivity for inbreeding, and other topics directly relevant to setting recovery thresholds. Therefore, recovery criteria appeared to be a relatively tractable target for improving the scientific implementation of the ESA.

BioScience 65: 189–199.
doi:10.1093/biosci/biu215
Advance Access publication 21 January 2015

Box 1. Sociopolitical factors influencing recovery criteria.

Multiple analyses have shown that sociopolitical factors have strong influences on many aspects of Endangered Species Act (ESA) implementation, including recovery criteria (Vucetich et al. 2006, Goble 2009). Two crucial components of recovery criteria that are particularly influenced by social and policy considerations are the following:

The portion of the range to which a species should be restored

The ESA calls for a species to be listed if it is endangered or threatened in all or a significant portion of its range (SPR), and, therefore, delisting should specify the geographic area to which healthy populations must be restored. Despite ongoing debate about the meaning of SPR (Vucetich et al. 2006, Carroll et al. 2010), the issue of where endangered species must or should be restored is clearly influenced by the sociopolitical setting and constraints imposed by feasibility and societal desirability. Within existing recovery plans, the extent of occupied range for recovered populations is typically addressed through viability needs. Similarly, the US Fish and Wildlife Service recently issued guidance on SPR, clarifying that a portion of the range is considered significant if "its contribution to the viability of the species is so important that, without that portion, the species would be in danger of extinction" (76 Fed. Reg. 237 (December 2011), pp. 76987–77006). The viability-based approach to recovery criteria that we advocate neither requires nor precludes broader definitions of SPR arising from the policy arena.

Acceptable risk of extinction

Under the ESA, *recovery* implicitly means that a species is not in danger of extinction (box 2), but any population has some possibility of extinction, and the ESA does not quantitatively define acceptable or unacceptable risk. Several authors have advocated for normative standards for acceptable extinction risk (e.g., Gilpin 1987, Mace and Lande 1991, Gerber and Demaster 1999) and National Marine Fisheries Service documents have proposed some guidelines (Demaster et al. 2004, McElhany et al. 2000, Regan et al. 2009). Similarly, the IUCN has established extinction risk levels for its categories of endangerment (IUCN 2012). Nonetheless, the acceptable risk of extinction for a recovered species has, so far, been determined on a case-by-case basis. We surveyed plans from 2009 to the present and show in figure 1 the combinations of extinction risk and time horizons for species for which both risk and horizon were defined in recovery criteria. We also indicate IUCN viability standards. Across plans, there is high variation but also a negative association between time horizon and extinction risk (Spearman rank correlations –0.59 and –0.83, $p < 0.02$, for delisting and downlisting, respectively), further exacerbating the high variance in acceptable extinction risk across plans. Society is willing to accept a higher extinction risk for some species (upper left) than for others (lower right). One striking trend was how few of these plans (only 6 of 23) employed quasiextinction thresholds, with the majority using complete extinction in defining risk. Although we do not propose or advocate for any universal standards for risk here, viability-based recovery criteria are compatible with the establishment of either universal or taxon-specific standards arising from the policy arena.

Although we see a crucial role for science in setting recovery criteria, defining what *recovery* should mean for a population or species involves more than scientific analysis. In particular, the risk of partial or complete failure (i.e., extinction) that we as a society are willing to accept and the degree to which we try to restore species to former numbers, distributions, and ecological functions blend into matters legal and ethical. These decisions are often made in part by biologists, but we emphasize that they are not objective biological decisions and that they require careful attention (box 1).

We begin with a brief tutorial on recovery planning, emphasizing the development of criteria. Even though all of us have read or reviewed numerous plans, served on recovery teams, or both, we nonetheless did not appreciate the practical constraints that several key legal and administrative rulings impose on how recovery plans must be written. Given our advocacy of increased involvement of academics in recovery planning, this description of *everything you (should have) always wanted to know about recovery planning, but were too ignorant to ask* is especially germane.

Legal and policy context

Recovery plans describe the biology of the species and its threats, develop a strategy for attaining recovery, outline actions needed to carry out the strategy, and detail the criteria by which attainment of recovery (box 2) can be assessed. Although a bevy of requirements and recommendations shape how recovery criteria are developed (NMFS and USFWS 2010), a handful of rules and legal decisions are also of key importance. The only explicit guidelines in the ESA regarding recovery criteria and actions are that recovery plans must "to the maximum extent practicable," contain "objective, measurable criteria, which, when met, would result in a determination, in accordance with the provisions [of the ESA], that the species be removed from the list" and "a description of such site-specific management actions as may be necessary to achieve the plan's goal for the conservation and survival of the species" (16 U.S.C. § 1533(f)(1)(B)). The ESA definition of *endangered* ("in danger of extinction throughout all or a significant portion of its range") highlights the role of extinction risk and spatial distribution in defining recovery but otherwise provides little guidance for recovery criteria and, in fact, injects additional need for policy clarification for undefined terms such as *in danger of* and *significant portion of its range* (Vucetich et al. 2006, Carroll et al. 2010). The services' *Recovery Planning Guidance* (NMFS and USFWS 2010), intended to provide more explicit guidelines for recovery planning and to outline

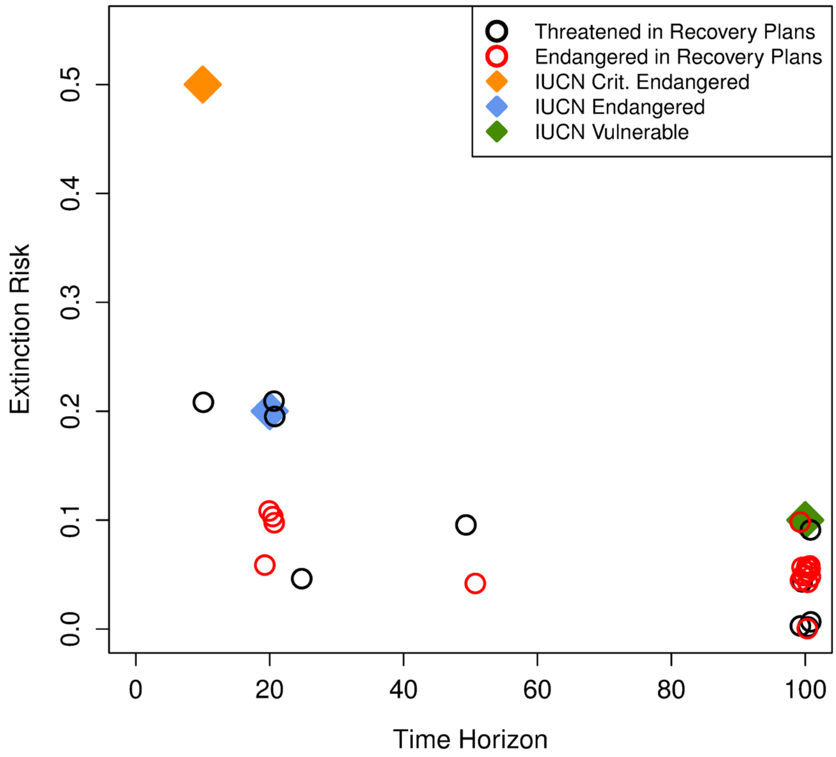

Figure 1. Viability criteria used to assess recovery are highly variable across recovery plans. A common way to assess population viability is the estimated probability (risk) that a population will become extinct or will fall below a specified quasiextinction threshold, over some time horizon. Endangered Species Act (ESA) recovery plans published between 2009 and 2013 used a wide range of acceptable risks of extinction or quasiextinction for delisting or downlisting, and there was a similarly broad set of time horizons employed. Viability standards defined for different International Union for Conservation of Nature (IUCN) categories of risk are also shown; for a given time horizon, the ESA criteria are generally more restrictive (i.e., require a lower risk of extinction) than are those used by the IUCN.

policy directives, indicates that they do not consider the measureable and objective requirement to mean that criteria must be quantitative (section 5.1.8.3). The guidance document defines recovery *actions* to be all activities "necessary to achieve full recovery of the species," as well as "the monitoring actions necessary to track the effectiveness of these actions and the status of the species" (NMFS and USFWS 2010).

One aspect of the services' approach to recovery criteria stems from the ESA requirement that prior to listing, the services must conduct a formal review to assess the extent to which the species is affected by five specific "threat factors": (1) destruction, modification, or curtailment of habitat or range; (2) overuse; (3) disease or predation; (4) inadequacy of existing regulation; and (5) any other natural or manmade factors. A species can only be removed from the list when none of the five factors threatens or endangers it. The courts have ruled that recovery criteria must address all five threat factors and must measure whether they have been ameliorated (*Fund for Animals v. Babbitt*: 903 F. Supp. 96 (D.D.C 1995)). The services interpret this ruling literally and recommend that plan writers formulate separate recovery criteria targeted at each threat factor (GAO 2006, NMFS and USFWS 2010). The services also suggest that demographic criteria (which we use in the sense of any estimates of population status; e.g., population size, trends through time, demographic rates, genetic factors, spatial distribution, or population viability indices) be listed separately from "threat-based" criteria (NMFS and USFWS 2010).

A final aspect of real-world recovery planning worth highlighting is that only about half of recovery plans are written by recovery teams, which include nonagency experts and may include agency staff as well. The rest are written by only one or a few agency personnel or contractors (Debby Crouse, US Fish and Wildlife Service, personal communication, 27 September 2014). This limited authorship demonstrates that resources (expertise, time, and money) for writing recovery plans are even more restricted than is widely recognized.

Current approaches to defining recovery criteria

How do these requirements and constraints affect the formulation of recovery criteria? Even very recent plans differ greatly in the number, range, format, quantity, and degree of specificity of their recovery criteria (see supplemental appendix S1 for examples of criteria from different plans, including many of those referred to in this section). For example, some plans contain only demographic criteria, such as the short-tailed albatross (*Phoebastria albatrus*) plan, whose sole delisting criterion stipulates requirements for population size, growth rate, and spatial distribution of the population.

However, most recent plans also—or primarily—use threat-based criteria that specify control or reduction of threats. The level of threat reduction required can vary in specificity and may or may not be linked explicitly to demography or viability. For example, one delisting criterion for the vermilion darter (*Etheostoma chermocki*) requires the attainment of specific water quality standards for turbidity over 10 consecutive years under a specified sampling regime. In contrast, the sei whale (*Balaenoptera borealis*) threat-based recovery criteria are more general, requiring that each

Box 2. Key definitions under the ESA.

The Endangered Species Act (ESA) protects species listed under the act as *endangered* or *threatened*:

Endangered

"In danger of extinction throughout all or a significant portion of its range" (16 USC § 1532).

Threatened

"Likely to become an endangered species within the foreseeable future throughout all or a significant portion of its range" (16 USC § 1532).

The ESA requires the development of recovery plans whose purpose is "to restore a species to ecological health" (USFWS 2013a). Several closely related concepts form the foundation of a recovery plan:

Recovery or recovery goal

The ESA's "ultimate goal is to 'recover' species so they no longer need protection under the ESA" (USFWS 2013). Therefore, at a minimum, *recovery* means that the species is not in danger of extinction in the foreseeable future. Translating this to the terms of quantitative conservation biology, recovery is the attainment of the conditions by which the species is viable over a long time frame. According to the services, "some recovery planning efforts may attempt to set goals higher than those needed to achieve delisting of the species" (NMFS and USFWS 2010). An example of such a goal might be reaching densities and distributions that allow the species to fulfill key ecological roles.

Recovery objective

The services use recovery objectives to link the recovery goal and criteria, stating "recovery objectives are the parameters of the goal, and criteria are the values for those parameters" (NMFS and USFWS 2010).

Recovery criteria

Recovery criteria are the conditions that signify recovery has been attained. As was stated by the services, "recovery criteria are the values by which it is determined that [a recovery] objective has been reached" (NMFS and USFWS 2010). Therefore, a clearly stated concept of recovery might be a 95% probability of persistence over 100 years.

Recovery actions

Recovery actions are the steps the services or other managers take to manage the species to achieve the goal of recovery. As was stated by the services, recovery actions are the steps "that will alleviate known threats and restore the species to long term sustainability. These actions might include (but are not limited to) habitat protection, limitations on take, outreach, research, control of disease, control of invasive species, controlled (including captive) propagation, reintroduction or augmentation of populations, and monitoring actions" (NMFS and USFWS 2010).

threat identified in the plan, such as reduced prey abundance due to climate change, anthropogenic noise, ship collisions, and gear entanglement, continue "to be investigated and any necessary actions being taken to address the issue are shown to be effective or this is no longer believed to be a threat."

Some threat-based criteria essentially consist of actions, including administrative or monitoring directives, focused on specific threats. For example, the downlisting criteria for the smalltooth sawfish (*Pristis pectinata*) stipulate that public education programs about the species and the prohibitions against harming it be in place. Similarly, delisting criteria for Kemp's ridley sea turtle (*Lepidochelys kempii*) include the establishment of a network of monitoring sites.

Occasionally, threats are accounted for by weighing their impacts on demographic processes. For example, the delisting criteria for the Gila trout (*Oncorhynchus gilae*) are focused on the number of populations and length of occupied streams, that were determined by quantitative analysis to best demonstrate resilience to the effects of catastrophic fires, the primary proximal threat to the species. More generally, the Gulf Coast jaguarundi (*Puma yagouaroundi cacomitli*) plan calls for habitat loss, degradation, and fragmentation to be reduced to the point that the species is no longer in danger of extinction. Similarly, the Wyoming toad (*Anaxyrus baxteri*) plan calls for chytridiomycosis infection rates to be maintained at levels that ensure the long-term sustainability of the population.

Other demographic criteria take the form of "viability criteria" that are either direct measures of a population's risk of extinction or quasiextinction (e.g., 5% risk of extinction within 100 years) or demographic measures (e.g., population size or trend) that have been shown to directly relate to a target recovery threshold—commonly, extinction risk. For example, one delisting criterion for the island fox (*Urocyon littoralis)* is based on extinction risk, calculated from population size and mortality rates. This criterion also details the time period, the quasiextinction threshold, and the number of years of consistently meeting the risk threshold required before recovery is declared. This plan also explicitly states that the analyses of risk can and should be updated as more data become available. Many more variations on demographic- and threat-based criteria exist among recent plans (appendix S1).

Regardless of their content, the ESA mandates that recovery criteria be measurable, but there is no history of this requirement being interpreted in the narrowest, most literal sense. Rather, a wide variety of measures, most of which are indirect and imprecise in the sense that they require statistical extrapolation from partial information (e.g., population sizes estimated from mark–recapture analyses, indirectly assayed threat abatement standards, estimated genetically effective population sizes, and probabilities of future extinction) have all been included in plans.

Some plans specify that additional evaluation, such as monitoring, population viability analyses (PVA), and threat assessment will be needed to develop or clarify criteria that are not immediately measureable. For example, some criteria (e.g., for the Mariana fruit bat, *Pteropus mariannus mariannus*; Bexar County karst invertebrates; the dwarf lake iris, *Iris lacustris*) state specific viability targets for a PVA that is yet to be developed. Others (e.g., for gentian pinkroot, *Spigelia gentianoides*; the scaleshell mussel, *Leptodea leptodon*; Guthrie's ground plum, *Astragalus bibullatus*; the Puerto Rican parrot, *Amazona vittata*) merely state criteria stipulating that future analyses must show that populations are "viable," without defining *viability*. Many threat-based criteria also call for additional analyses to specify target levels. For example, the criteria may state that habitat adequate in extent, quality, and quantity will be identified and protected (e.g., the plan for the Caribbean manatee, *Trichechus manatus*) or that a threat will continue to be investigated and ameliorated until it no longer limits recovery (e.g., entanglement for sei whales, water flows for the Caribbean manatee).

Common problems with current recovery criteria

We see two problems with the way criteria are often framed and justified. First, many plans fail to link the recovery criteria, either demographic or threat based, to some clearly specified definition of population recovery. In other words, many plans do not articulate how meeting recovery criteria will result in a population that is at reduced risk of extinction or otherwise deemed to be "recovered." This issue has a considerable history in critiques of recovery plans (Schemske et al. 1994, Gerber and Hatch 2002) and continues to be a problem in even the most recent plans (Neel et al. 2012, Himes Boor 2014).

A second but related problem is the conflation of recovery criteria and recovery actions. Although these two aspects of a plan are described as distinct elements in the ESA (box 2), in practice, many plans include what would commonly be considered *actions* (Salafsky et al. 2008) among their recovery "criteria." For example, many plans include criteria requiring the establishment of monitoring programs or other biological studies (appendix S1). We heard from both service personnel and personnel from conservation nongovernmental organizations that recovery plan writers may seek to highlight the importance of actions by listing them as criteria and that funding may be more available for actions that are listed as criteria. Still, we view this mixing of actions and criteria as problematic. Recovery *criteria* should reflect something about the status of the species itself (e.g., population size or distribution, rate of population growth, rate of mortality from some threat) that indicates that it has reached a state of recovery, whereas recovery *actions* are what managers do to achieve and evaluate recovery (box 2).

Recommendations for improved recovery criteria

Regardless of the exact degree of risk that a plan's recovery criteria embrace—part of the societal decisions that underlie any plan—a scientifically defensible plan should include quantitative recovery criteria establishing that the species is safe from extinction or extreme declines for the moderate-term future or that the species is likely to maintain an even higher number or wider geographical distribution deemed necessary for it to play its proper ecological role. Such criteria must account for existing and anticipated or potential future threats (Salafsky et al. 2008), including climate change effects and the shifting regulatory and threat landscapes faced by delisted species (Soulé et al. 2005, Goble et al. 2012). The broad set of analytical methods used to judge whether a population or set of populations meets such a standard is usually called *population viability analysis*, or PVA. Although we use this acronym, we emphasize that we are not advocating solely for matrix population models or similar approaches (e.g., integral projection models). Many simpler and also more sophisticated modeling tools exist in which quantitative approaches are used to synthesize a wide range of information (e.g., demographics, spatial distribution, threats) to do more than simply assess extinction risk (e.g., Bayesian, hierarchical, state-space, or integrated population models; also see supplemental appendix S3). These tools can and should be used to judge the likelihood of sustaining a wide range of desired attributes of a recovered species, including the number and density of individuals, the number and geographic distribution of populations, fulfillment of ecological functioning, and other population traits that are associated with the conservation biology principles of representativeness, redundancy, and resiliency.

Within this broad suggestion, we offer three more specific recommendations:

Recommendation 1: The central recovery criteria should be quantitative, biologically-based, and clearly justified. To the greatest extent possible, criteria should be quantitative; focused on traits of the species, itself, rather than external factors; and based on clear scientific reasoning. To ensure this direct link between criteria and species biology, plans should have a distinct section that outlines the biological justification for each criterion, with evidence of how the quantitative standards are objectively linked to a clearly stated definition of recovery (box 3). Given the ambiguity in the ESA regarding what recovery is, this recommendation serves to facilitate both an unambiguous statement of how recovery is defined

Box 3. Illustrative wording for recovery criteria.

We present the following templates for demographic, threat-based, and combined criteria that follow the recommendations outlined in the text as well as recommendations made in Himes Boor (2014). They are presented as illustrative examples; many other criteria could be formulated that meet the standards set out in our recommendations.

Demographic criterion with adequate data

Estimated intrinsic growth rate for the entire population must meet or exceed _ with _% probability of certainty for more than _ years based on our analyses that such growth will result in a population with less than _% probability of quasiextinction within _ years (see appendix _ for analysis details, including assumptions and definitions).

Demographic criterion with inadequate data

The species as a whole should have less than an _% probability of extinction within _ years and each individual population should maintain a probability of extinction less than _% within _ years. The PVA models must be peer reviewed and must take into account uncertainty in parameter estimates and future scenarios, including potential impacts of climate change and threat factors _ and _. The data to complete such an assessment should meet the standards outlined in section _ of this recovery plan.

Threat-based criterion with adequate data

Threat _ must be reduced to an estimated (based on the upper _% confidence interval) _ units per year across the entire species current range and must remain at or below that level for _ years. We estimate that this reduction will result in a _% increase in vital rate _, therefore allowing a population growth rate consistent with less than an _% probability of extinction within _ years (see appendix _ for the detailed analysis and model assumptions).

Threat-based criterion with inadequate data

Threats _ and _ should be reduced such that their cumulative impact on the species is no longer threatening its viability and the population has greater than _% probability of persistence for more than _ years. The model developed to estimate viability must be peer reviewed and must take into account uncertainty in parameter estimates, future management scenarios, and threat impacts.

Combined demographic and threat-based criterion

Estimates of total population size must meet or exceed _ breeding individuals with _% probability of certainty for more than _ years, based on our analyses demonstrating that a population maintaining that number of breeding individuals has less than _% probability of quasiextinction within _ years and has overcome threats _ and _.

Each criterion for which a model (e.g., PVA or threats analysis) is used should also specify the section of the recovery plan containing the detailed model description, including all model assumptions and justifications. Criteria should also be accompanied by references to the section of the recovery plan describing explicit methodologies for collecting data and estimating parameters, including acceptable levels of uncertainty surrounding estimated parameters. This will ensure the appropriate data are collected for the desired analysis.

for a species and how the specified criteria demonstrate that the species has a high probability of remaining in this “recovered” state. Both the definition and the rationale are essential to ensure that the connections between available information about the species and the plan’s recovery criteria are transparent to the public and to plan reviewers. We recognize that many other ancillary criteria will often be included in plans that address less direct aspects of recovery and population management, but without the inclusion of criteria that are directly related to biological recovery, a plan is not scientifically defendable.

Recommendation 2: All plans should include demographic criteria. Plans should include one or more demographic criteria (criteria focused on population number, dynamics, or demography) and should state how analyses have been (or will be) done to tie these criteria to the probability of populations meeting specific quasi-extinction risk thresholds or other indices of population health (box 3). If adequate data are available at the time a plan is written, plan developers should conduct analyses of population viability and identify quantitative population metrics, such as population size, population trends over a specified time period, and geographical distribution, that indicate that the population has an acceptably low risk of falling below recovery thresholds. If the data are not in hand to support such analyses when a plan is written, the criteria can state the thresholds and risks that are deemed acceptable, and recovery actions can specify collection of the data that will be needed to assess when that criterion has been met (figure 2). Both of these approaches are preferable to setting arbitrary demographic thresholds that have no clear link to a species’ ecosystem role or its future viability (Schemske et al. 1994, Tear et al. 1995). As was noted above, these approaches have already been taken in some approved plans (e.g., the sei whale, the Mariana fruit bat) and have been advocated by NMFS scientists (Demaster et al. 2004) and others (Himes Boor 2014), so they are not untested nor too uncertain to pass muster under the ESA. In practice, many of the best plans take a combined approach, defining demographic standards that predict a certain safety

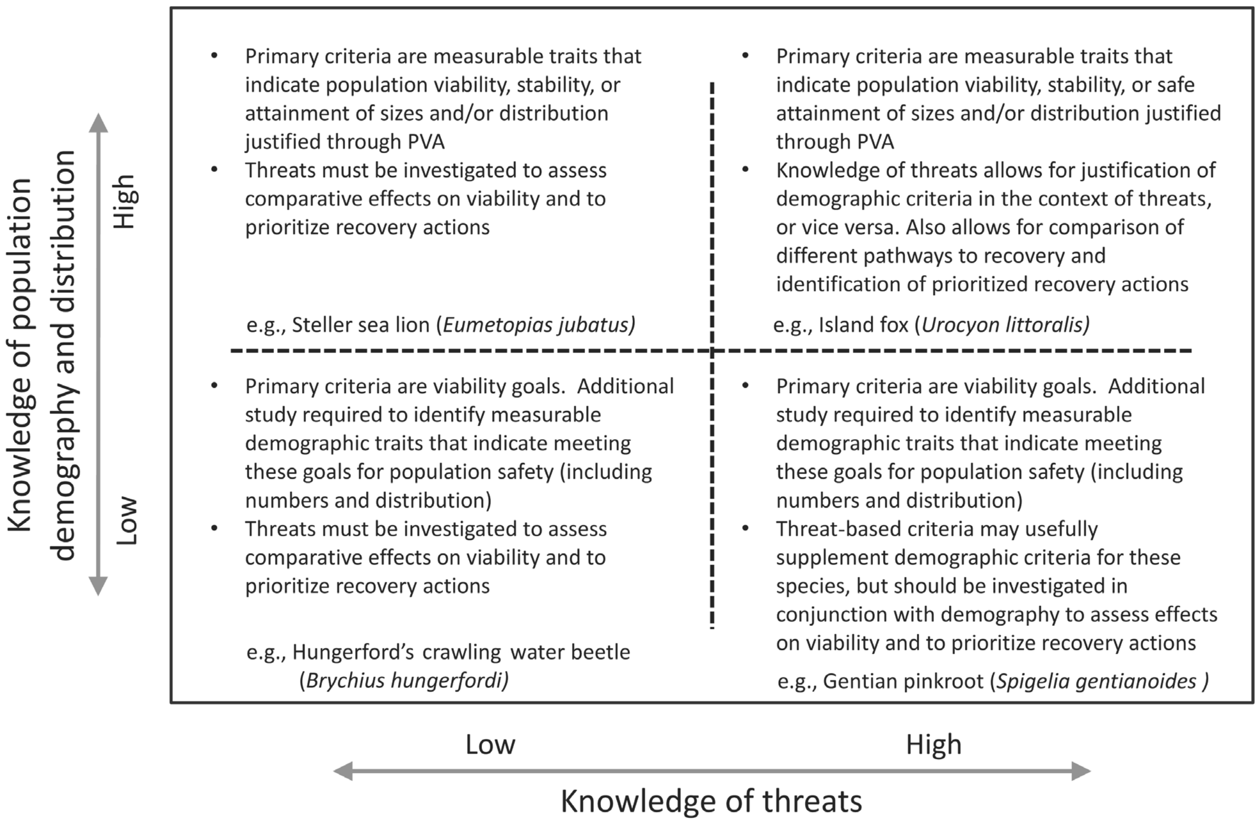

***Figure 2. Formulating the path to recovery for threatened and endangered species is influenced by the degree of knowledge of threats and of population demography and distribution. We present general guidelines for developing demographic and threat-based recovery criteria for listed species based on the initial levels of knowledge about the species and its threats. All completed recovery plans, including those listed here as examples, are available at* www.fws.gov/endangered/species/recovery-plans.html.**

from falling below desired thresholds but also stipulating further data collection to refine the link between numbers and safety, which will, in general, involve use of some type of PVA (appendix S3).

Recommendation 3: Threat-based criteria should derive from the population consequences of threats. A plan that has only threat-based criteria, unlinked to population trends or demographic measurements, is difficult or impossible to defend scientifically. When quantitative estimates of the impacts of threats on demographic processes or population growth rates are available, the level of threat reduction stipulated as a goal for recovery should be based on their population-level effects, in the context of other threats and the species' life history. As the classic case of the loggerhead sea turtle (*Caretta caretta*) shows, such analyses are necessary to correctly prioritize among different threats (Crouse et al. 1987, Crowder et al. 1994) and therefore to gauge the threat reduction needed to achieve self-sustaining populations, in part because threat factors, themselves—let alone specific levels for their abatement—are inherently difficult to precisely and defendably define. We recommend that threat-based recovery criteria—that is, criteria that specify the threat abatement necessary for the removal of a species from ESA protection—be expressed in terms of the level of threat reduction needed for population viability. Specifically, the impacts of current and anticipated future threats (including the loss of ESA protections) should be included in population models so that the interactive effects of multiple threats or threat reductions are folded into an overall assessment of viability (see supplemental appendix S2). An alternative option, already taken in some plans (e.g., for the black-footed ferret, *Mustela nigripes*), is to justify the demographic thresholds in the context of threats and describe how achievement of the criteria would indicate that the threats have been adequately addressed. All criteria should specify that any new information about the demographic impacts of threats and the expected impact of regulatory changes after delisting be incorporated when assessing whether the population is recovered. Although accurately anticipating novel or changing threats is not trivial, our approach incorporates this uncertainty into a framework that is flexible and that requires any new threats to be controlled to the levels necessary to achieve population safety before delisting can be considered.

If the demographic impacts of a threat cannot be adequately quantified when a plan is written, one alternative is to define criteria addressing this threat in terms of viability (box 3). In these data-poor situations (figure 2), this would involve a two-pronged approach that takes advantage of the requirement for plans to define actions, as well as criteria. First, recovery *criteria* would specify that the threat must be low enough to allow the population to meet a specific viability standard. Second, recovery *actions* would include activities that lower threat levels and also that collect data to quantify the demographic or population-level responses to these threat reductions.

This approach to threat reduction can also effectively address conservation-reliant species. Managers are increasingly aware that many endangered species will require conservation measures in perpetuity (Goble et al. 2012). Well-executed PVA analyses can take into account future threat management scenarios, including the effects of delisting on the regulatory mechanisms needed to ensure that essential management continues. In our view, assessing whether even the seemingly nonbiological threat factor 4 (the inadequacy of existing regulation) has been sufficiently ameliorated requires a population perspective (e.g., will laws limiting future harvest allow the species to sustain numbers above the desired population thresholds?). In some cases, a realistic consideration of a species' biology and future threat scenarios (e.g., climate change, regulatory changes) may preclude recovery criteria that are attainable in the foreseeable future; nevertheless, such a determination would be a successful outcome of quantitative analyses and of the ESA, rather than a failure (Doremus and Pagel 2001).

Implications of these recommendations

Our recommendations contrast with the services' current guidelines on viability-based criteria, which state that such criteria should be ancillary to "traditional population and listing factor-based recovery criteria" because, they state, PVAs rely on estimates of vital rates and on assumptions about threat conditions and their effects on demographic rates (NMFS and USFWS 2010; as has been noted elsewhere, PVAs can be based on many other kinds of data). However, "traditional" criteria not linked to PVA are also based on guesses or assumptions about population processes, including demographic rates, as well as assumptions about threat conditions and their effects on demography, with the important difference that these assumptions and estimates are often unclear, implicit, and indirect. This lack of transparency in the estimates and assumptions linking traditional criteria and population health is their key weakness. In viability-based criteria, assumptions about the effects of threats on recovery are explicitly stated, which allows for updating of criteria (and models) as assumptions are tested and additional data are collected.

We acknowledge that we are setting a high bar for the services, especially in suggesting that threats be tied to population responses and extinction risk, but we believe that it is crucially important to push for these high standards so as to ensure the long-term survival of imperiled species and efficient and effective management. Following our recommendations will make criteria more scientifically and legally defensible, more transparent, and more aligned with the already-developed conservation planning literature (e.g., Salafsky et al. 2002, 2008). In particular, our recommendations are intended to create a scientifically justifiable approach that can accommodate the diverse situations of different listed species (figure 2). For some species, large, long-term data sets are available, the effects of threat factors have been experimentally estimated, and adequate financial resources to support management are in hand. For most species, none of these advantages exist, and a recovery plan can count on only modest monitoring and analysis efforts, which make rigid numerical recovery criteria set at the time the plan is written impractical and indefensible. The approach that we suggest can accommodate both these extremes, without resorting to weak generalizations, guesswork, or untested expert opinion. Furthermore, they are designed to be flexible enough to allow recovery criteria to stay relevant in the face of shifting threat conditions such as climate change, exotic species, and land-use change.

Just as importantly, an emphasis on recovery criteria that are tied to a population's status, rather than to the amelioration of specific threats, can give the services flexibility to change management tactics if new threats arise after the recovery plan is written. Using demographic criteria, the degree of threat abatement needed can be directly tied to the ultimate goal of recovery, and when new information indicates that more or less attention to a given threat is needed, the criteria can accommodate this updated information.

Finally, having to show that recovery criteria actually mean that a population is relatively safe from extinction or from dropping to a low level that impedes its functional role in an ecosystem may mean that some species are not removed from the list as quickly. We underscore, however, that this is not a valid objection to these recommendations. If we are slower to remove species from ESA protections because we cannot say with an acceptable degree of certainty that they are indeed recovered, that is the scientifically justifiable, legally required, and precautionary outcome. That said, making clearer statements of how recovery is defined and justified should also mean faster delisting of some species, as well as making recovery actions more targeted and delisting decisions less contentious.

In considering our first and most fundamental recommendation, it is important to address several aspects of PVA and related population analysis tools. First, this is not a recommendation to adopt hopelessly complex approaches to viability assessment. Population analyses can be quite simple, even when applied to spatially complex situations (see appendix S3 for examples); this recommendation does not require mountains of data or cutting-edge analysis, nor is it designed to be a job creation program for population modelers. What it does require is a clear statement of what risk of population deterioration is deemed acceptable and why the recovery criteria proposed would indicate that a species has likely met this goal. The need to define such clear standards is the most fundamental advantage of taking this approach to recovery criteria development.

Second, implementing these recommendations does not require that PVA and other population analysis methods be flawless. The strengths and weaknesses of predicting population fates have been thoroughly dissected in the conservation literature (Beissinger and Westphal 1998, Ludwig 1999, Coulson et al. 2001, Ellner et al. 2002). However, the core shortcomings of PVA as a predictive tool are shared with all other predictive methods. Some may argue that, because they are based on analyses more complex than simple statistics, viability-based criteria may be less palatable to policymakers and managers. But this objection applies to many types of scientific evidence already used in legal and social contexts, such as genetic analyses used in criminal cases or the formulation of ecotoxicological standards in pollution control and, in this case, can be addressed by clear explanation of the details of the data and assumptions used to estimate population viability and its uncertainty.

Finally, with regard to the use of population analysis methods to assess recovery, the limitations of PVAs must be judged against the shortcomings of alternative methods for determining recovery. We do not see a good argument for the use of criteria justified mostly or solely by expert opinion, as opposed to standards based on actual analysis of population status and dynamics. Another potential option would be to adopt the International Union for Conservation of Nature's (IUCN) listing criteria (IUCN 2012). However, we believe that this would be a poor way to improve recovery planning. Although their adoption would standardize recovery criteria, IUCN benchmarks were designed as a one-size-fits-all system for global priority setting across all taxa and multiple conservation situations and, as such, do not take into account species-specific biology and threat conditions, which we believe is a crucial component of effective management and precautionary delisting. With that said, our recommendations are not incompatible with the IUCN approach, because one of the requirements for moving a species to a lower IUCN threat level is the completion of a quantitative analysis to evaluate its risk of extinction.

Implementing the recommendations

Criticism of ESA implementation is easy, but practical improvements likely to be adopted given the services' legal, political, and budgetary constraints are hard. Based on our conversations with service personnel, we offer these suggestions for how to implement our recommendations.

First, we suggest that the recovery planning guidelines be revised to provide clear guidance to recovery plan authors on why and how to set quantitative, scientifically defensible criteria. We have tried to describe as lucidly as possible how such criteria could be formulated (box 3; appendix S3).

Second, we suggest that the services develop mechanisms to encourage both natural and social scientists from outside the agencies to contribute their expertise and time to the process of developing recovery criteria. Writing a well-articulated, objective, and defensible plan would seem nearly impossible without input from individuals with multiple perspectives and expertise, including those with an understanding of the legal and regulatory sideboards of recovery planning; knowledge of the species and its ecosystem, as well as the threats the species faces and their biological impacts; knowledge of the political, social, and land-use settings where the species occurs; and expertise in analytical and modeling methods necessary to assess a species' recovery status in a scientifically defensible way. For high-profile species, it is easier for the services to assemble recovery teams that include members with each of these types of expertise. But the many species for which plans are written by individuals or small teams will often not have the benefit of this complete set of knowledge and skills. This is not a trivial obstacle to improving recovery planning.

One possibility to redress this limitation is for university biologists to incorporate recovering planning into their teaching. For example, graduate students in a population ecology course could construct, parameterize, and use population models to craft demography-based threat reduction actions and recovery criteria. If adequate data are not available, students and faculty members could work with plan writers to design effective recovery actions to collect the data needed to assess recovery. Close coordination with the services in such efforts is essential so that the contributions of academic partners are useful to the planning process. A different approach to achieve the same end would be to find funding for postdoctoral researchers or other individuals outside the services to contribute expertise that could allow the services to more rapidly produce defensible plans. An added benefit of either scenario is that a cohort of young scientists will gain real-world experience at the intersection of conservation science, practice, and policy, and thereby foster their careers in conservation. Experts on planning, policy, social science, and environmental law could likewise be tapped to work on other elements of recovery planning.

Finally, the services are required to review the status of each listed species every five years, including the evaluation of new information and threats that can trigger a revision of an outdated recovery plan (NMFS and USFWS 2010). We encourage nonagency scientists with species-specific expertise to communicate their work to the services, and we urge the services to create openings for nonagency experts to participate in these reviews, including updating population assessments in light of new data. This phase of the recovery process presents another opportunity for early-career scientists to make substantive contributions to conservation practice.

Conclusions

We believe that we have presented practical and important ways to enhance the scientific integrity of the recovery planning process. Similarly, we think that creating ways to better tap the expertise, time, and enthusiasm of scientists outside of the services can be a means to implement these recommendations and overcome the very real constraints faced by the services in writing strong recovery plans. For that external involvement to be efficient and effective, however, the services must be open to working with outsiders, and scientists must understand the needs and constraints inherent in ESA implementation.

Although we have focused here on recovery planning under the US ESA, many other nations have similar legislation with provisions for endangered species recovery. Although there are parallel sets of proposed approaches to endangered species assessment and recovery planning in other jurisdictions, these proposals and critiques are similar to those of the US ESA; there are many suggestions but little evidence of on-the-ground improvement (Salafsky et al. 2008, Mooers et al. 2010, but see Salafsky and Margoluis 1999). The general approaches that we suggest here can help improve the management of threatened species elsewhere and may also be applicable to other aspects of ESA planning, such as critical habitat designation and section 7 consultation. With our emphasis on defining clear standards by which to judge recovery and requiring that recovery criteria and threat reductions be explicitly linked to these measures of population safety, our recommended approach will help ensure that recovery plans more effectively and efficiently guide recovery of imperiled species.

Acknowledgments

We gratefully acknowledge the contributions of Tyler Abbott, Charlotte Ambrose, Carlos Carroll, Deborah Crouse, Maile Neel, Larry Rabin, Jay Tutchton, and Shaye Wolf, all of whom participated in our October 2012 workshop but who could not or chose not to be authors of this article. Nonetheless, they provided key perspectives and information and deserve more than a standard acknowledgement. The Wilburforce Foundation and The Nature Conservancy provided funding, and the Gordon and Betty Moore Foundation hosted the workshop. WFM was supported by the Swedish Science Council. GKHB was supported in part by the US Marine Mammal Commission. Authors are listed in order of their contributions to the paper.

Supplemental material

The supplemental material is available online at *http://bioscience.oxfordjournals.org/lookup/suppl/doi:10.1093/biosci/biu215/-/DC1*.

References cited

Beissinger SR, Westphal MI. 1998. On the use of demographic models of population viability in endangered species management. Journal of Wildlife Management 62: 821–841.

Boersma PD, Kareiva P, Fagan WF, Clark JA, Hoekstra JM. 2001. How good are endangered species recovery plans? Bioscience 51: 643–649.

Carroll C, Vucetich JA, Nelson MP, Rohlf DJ, Phillips MK. 2010. Geography and recovery under the US Endangered Species Act. Conservation Biology 24: 395–403.

Coulson T, Mace GM, Hudson E, Possingham H. 2001. The use and abuse of population viability analysis. Trends in Ecology and Evolution 16: 219–221.

Crouse DT, Crowder LB, Caswell H. 1987. A stage–based population model for loggerhead sea turtles and implications for conservation. Ecology 68: 1412–1423.

Crowder LB, Crouse DT, Heppell SS, Martin TH. 1994. Predicting the impact of turtle excluder devices on loggerhead sea turtle populations. Ecological Applications 4: 437–445.

Demaster DP, Angliss R, Cochrane J, Mace P, Merrick R, Miller M, Rumsey S, Taylor B, Thompson G, Waple R. 2004. Recommendations to NOAA Fisheries: ESA Listing Criteria by the Quantitative Working Group. National Oceanic and Atmospheric Administration. Technical Memorandum no. NMFS-F/SPO-67.

Doremus H, Pagel JE. 2001. Why listing may be forever: Perspectives on delisting under the US Endangered Species Act. Conservation Biology 15: 1258–1268.

Ellner SP, Fieberg J, Ludwig D, Wilcox C. 2002. Precision of population viability analysis. Conservation Biology 16: 258–261.

Foin TC, Riley SPD, Pawley AL, Ayres DR, Carlsen TM, Hodum PJ, Switzer PV. 1998. Improving recovery planning for threatened and endangered species. Bioscience 48: 177–184.

[GAO] Government Accountability Office. 2006. Endangered Species: Time and Costs Required to Recover Species Are Largely Unknown. GAO. Report no. GAO-06-463R.

Gerber LR, Demaster DP. 1999. A quantitative approach to Endangered Species Act classification of long-lived vertebrates: Application to the north Pacific humpback whale. Conservation Biology 13: 1203–1214.

Gerber LR, Hatch LT. 2002. Are we recovering? An evaluation of recovery criteria under the U.S. Endangered Species Act. Ecological Applications 12: 668–673.

Gibbs KE, Currie DJ. 2012. Protecting endangered species: Do the main legislative tools work? PLOS ONE 7 (art. e35730).

Gilpin ME. 1987. Spatial structure and population vulnerability. Pages 125–139 in Soule ME, ed. Viable Populations for Conservation. Cambridge University Press.

Goble DD. 2009. The Endangered Species Act: What we talk about when we talk about recovery. Natural Resources Journal 49: 1–44.

Goble DD, Wiens JA, Scott JM, Male TD, Hall JA. 2012. Conservation-reliant species. BioScience 62: 869–873.

Himes Boor GK. 2014. A framework for developing objective and measurable recovery criteria for threatened and endangered species. Conservation Biology 28: 33–43.

Hoekstra JM, Clark JA, Fagan WF, Boersma PD. 2002. A comprehensive review of Endangered Species Act recovery plans. Ecological Applications 12: 630–640.

[IUCN] International Union for Conservation of Nature. 2012. IUCN red list categories and criteria, version 3.1, 2nd ed.. IUCN.

Lawler JJ, Campbell SP, Guerry AD, Kolozsvary MB, O'Connor RJ, Seward LCN. 2002. The scope and treatment of threats in endangered species recovery plans. Ecological Applications 12: 663–667.

Ludwig D. 1999. Is it meaningful to estimate a probability of extinction? Ecology 80: 298–310.

Mace GM, Lande R. 1991. Assessing extinction threats: Toward a reevaluation of IUCN threatened species categories. Conservation Biology 5: 148–157.

McElhany P, Ruckleshaus MH, Ford MJ, Wainwright TC, Bjorkstedt. EP. 2000. Viable salmonid populations and the recovery of evolutionarily significant units. National Oceanic and Atmospheric Administration. Technical Memorandum no. NMFS-NWFSC-42.

Mooers AO, Doak DF, Findlay CS, Green DM, Grouios C, Manne LL, Rashvand A, Rudd MA, Whitton J. 2010. Science, policy, and species at risk in Canada. Bioscience 60: 843–849.

Morris WF, Bloch PL, Hudgens BR, Moyle LC, Stinchcombe JR. 2002. Population viability analysis in endangered species recovery plans: Past use and future improvements. Ecological Applications 12: 708–712.

Moyle LC, Stinchcombe JR, Hudgens BR, Morris WF. 2003. Conservation genetics in the recovery of endangered animal species: A review of US endangered species recovery plans (1977–1998). Animal Biodiversity and Conservation 26: 85–95.

Neel MC, Leidner AK, Haines A, Goble DD, Scott JM. 2012. By the numbers: How is recovery defined by the US Endangered Species Act? Bioscience 62: 646–657.

[NMFS and USFWS] National Marine Fisheries Service, US Fish and Wildlife Service. 2010. Interim endangered and threatened species recovery planning guidance, version 1.3. NMFS and USFWS.

Regan T, Taylor B, Thompson G, Chochrane J, Merrick R, Nammack M, Rumsey S, Ralls K, Runge M. 2009. Developing a structure for quantitative listing criteria for the U.S. Endangered Species Act using performance testing, Phase 1 Report. National Oceanic and Atmospheric Administration. Technical Memorandum no. NMFS-TM-NMFS-SWFSC-437.

Rohlf DJ. 1991. Six biological reasons why the Endangered Species Act doesn't work—and what to do about it. Conservation Biology 5: 273–282.

Salafsky N, Margoluis R. 1999. Threat reduction assessment: A practical and cost-effective approach to evaluating conservation and development projects. Conservation Biology 13: 830–841.

Salafsky N, et al. 2008. A standard lexicon for biodiversity conservation: Unified classifications of threats and actions. Conservation Biology 22: 897–911.

Schemske DW, Husband BC, Ruckelshaus MH, Goodwillie C, Parker IM, Bishop JG. 1994. Evaluating approaches to the conservation of rare and endangered plants. Ecology 75: 584–606.

Scott JM, Goble DD, Wiens JA, Wilcove DS, Bean M, Male T. 2005. Recovery of imperiled species under the Endangered Species Act: The need for a new approach. Frontiers in Ecology and the Environment 3: 383–389.

Soulé ME, Estes JA, Miller B, Honnold DL. 2005. Strongly interacting species: Conservation policy, management, and ethics. BioScience 55: 168–176.

Tear TH, Scott JM, Hayward PH, Griffith B. 1993. Status and prospects for success of the Endangered Species Act: A look at recovery plans. Science 262: 976–977.

———. 1995. Recovery plans and the Endangered Species Act: Are criticisms supported by data? Conservation Biology 9: 182–195.

[USFWS] US Fish and Wildlife Service. 2013. ESA basics: 40 years of conserving endangered species. USFWS. (9 December 2014; *www.fws.gov/endangered/esa-library/pdf/ESA_basics.pdf*).

Vucetich JA, Nelson MP, Phillips MK. 2006. The normative dimension and legal meaning of endangered and recovery in the US Endangered Species Act. Conservation Biology 20: 1383–1390.

Dan F. Doak (daniel.doak@colorado.edu) and Allison Louthan are affiliated with the Environmental Studies Program at the University of Colorado, Boulder. Gina K. Himes Boor and Victoria J. Bakker are affiliated with the Ecology Department at Montana State University, in Bozeman. William F. Morris is affiliated with the Department of Ecology and Genetics at Uppsala University, in Uppsala, Sweden and with the Biology Department at Duke University, in Durham, North Carolina. Scott A. Morrison is affiliated with the The Nature Conservancy, in San Francisco, California. Amanda Stanley is affiliated with the Wilburforce Foundation in Seattle, Washington. Larry B. Crowder is affiliated with the Center for Ocean Solutions in Monterey, California.