NATURAL RESOURCES JOURNAL

Volume 44
Issue 4 *Fall 2004*

Fall 2004

Beyond Definitions: Maintaining Biological Integrity, Diversity, and Enivironmental Health in National Wildlife Refuges

James R. Karr

JAMES R. KARR*

Beyond Definitions: Maintaining Biological Integrity, Diversity, and Environmental Health in National Wildlife Refuges

ABSTRACT

Throughout its century-long history, the National Wildlife Refuge System has been dedicated to the protection of living systems. For many refuges, management emphasis involved a subset of nature, such as migratory waterfowl. Passage of the 1997 National Wildlife Refuge System Improvement Act defined a new mission but framed it with elusive terms such as biological integrity and environmental health. Although the context and meaning of these words have been explored for several decades, as in implementation of the Clean Water Act, they remain the focus of debates. Because the controversy is unlikely to be resolved entirely, the Fish and Wildlife Service should place emphasis on moving beyond debates about definitions to actually understand status and trends in refuge living systems. That understanding can come only with a rigorous sampling and analytical framework focused on practical and technically sound measures to track refuge condition. Defining precisely what parameters are to be measured and documenting how they behave in the face of human-induced and natural disturbances must form the centerpiece of these efforts. Equally important is the task of communicating the status and trends of living resources within refuge boundaries to Fish and Wildlife Service administrators, political leaders, and the public.

INTRODUCTION

As knowledge expands and perspectives and philosophies shift, societal goals evolve. So too do the approaches used to accomplish those goals. Nowhere is that evolution more striking than in environmental

* Professor, Department of Biology and School of Aquatic and Fishery Sciences, University of Washington, Seattle. Ph.D. 1970, M.S., 1967, University of Illinois at Urbana-Champaign; B.S. 1965, Iowa State University.

protection, where advances in environmental science,[1] shifting legislative mandates,[2] and altered societal philosophies[3] challenge historical approaches to land and water management.

That evolution is vividly displayed in the management goals defined in the 1997 National Wildlife Refuge System Improvement Act (Improvement Act), especially the provision "to ensure that the biological integrity, diversity, and environmental health of the System are maintained."[4] This mandate creates challenges at three levels. First, all key parties must converge on a framework of definitions for nettlesome concepts such as biological integrity, biodiversity, and environmental health. Definitions must go beyond the general and academic to the practical; they must be sensible guides for refuge management. Second, program success depends on a rigorous sampling and analytical framework, including practical and technically sound measures for tracking these elusive concepts. Defining precisely what parameters are to be measured and documenting how they behave in the face of human-induced and natural disturbances are key components of this second issue. Third, patterns and trends must be communicated so that citizens and policy makers can understand both trends and their consequences for the values refuges are established to protect.

DEFINING AND APPLYING KEY CONCEPTS

Aldo Leopold pioneered use of the concepts of health and integrity in the environmental arena. He defined "land health" as "the capacity of the land for self-renewal." He also noted, "A thing is right when it tends to preserve the integrity, stability, and beauty of the biotic community. It is wrong when it tends otherwise."[5]

By the late twentieth century, the terms *health* and *integrity* became lightning rods, especially among scientists.[6] Some argue that such value-laden words should not be applied to multispecies

---

1. James R. Karr & Chris O. Yoder, *Biological Assessment and Criteria Improve Total Maximum Daily Load Decision Making*, 130 J. ENVTL. ENG'G 594, 599 (2004).
2. WILLIAM H. RODGERS, JR., ENVIRONMENTAL LAW 53 (2d ed. 1994); ROBERT W. ADLER ET AL., THE CLEAN WATER ACT 20 YEARS LATER 6 (1993).
3. LESLIE PAUL THIELE, ENVIRONMENTALISM FOR A NEW MILLENNIUM: THE CHALLENGE OF COEVOLUTION 3 (1999).
4. 16 U.S.C. § 668dd(a)(4)(B) (2000).
5. ALDO LEOPOLD, A SAND COUNTY ALMANAC: WITH ESSAYS ON CONSERVATION FROM ROUND RIVER 258, 262 (1966).
6. James R. Karr, *Vignette 11.1: Biological Integrity and Ecological Health*, *in* FUNDAMENTALS OF ECOTOXICOLOGY 245 (Michael C. Newman & Michael A. Unger eds., 2d ed. 2003).

assemblages, such as ecosystems or landscapes;[7] others hold that talking about ecological health or biological integrity is beyond the purview of science (*e.g.*, not an observable ecological property).[8] Yet the words are particularly useful in policy-making arenas precisely because they are familiar and imply values worth protecting. It seems a natural intuitive leap from "my health" or the nation's "economic health" to "ecological health" or "land health." And, as a legal or policy goal, protecting a place's health and integrity has greater direct appeal than abstractions such as "system dynamics" or "ecosystem functions."[9]

Like people, ecosystems or wildlife refuges can be more or less "ill." An unhealthy person may be suffering from a cold or dying of cancer; an unhealthy refuge may be degraded by loss of a top predator, a few sensitive species, or all of its vegetation. An unhealthy river may have game fish populations depleted by overfishing or no fish at all. Perhaps only a few of the river's most tolerant invertebrates may remain after severe chemical pollution. The healthiest places, those with integrity, have undergone little or no disturbance at human hands. These places support a balanced, integrated, adaptive biota having the full range of elements or parts (genes, species, assemblages; plants, animals, microbes) and processes (mutation, demography, biotic interactions, nutrient cycling, energy flow, metapopulation dynamics) characteristic of the region and expected in areas with minimal human influence.[10] Because such places support a thriving living system, they retain the capacity to regenerate, reproduce, sustain, adapt, develop, and evolve; they retain the full legacy of wild nature, or, in Leopold's words, they still have "all the parts."[11] Complete, unimpaired living systems, then, possess biological integrity; they support a biota that is the product of evolutionary and biogeographic processes with little or no influence from industrial society.

Both parts and processes are important to human and nonhuman living systems. Places with this biological integrity provide a

---

7. David Policansky, *Application of Ecological Knowledge to Environmental Problems: Ecological Risk Assessments*, *in* COMPARATIVE ENVIRONMENTAL RISK ASSESSMENT 44-46 (C. Richard Cothern ed., 1993).

8. Peter Calow, *Can Ecosystems Be Healthy? Critical Consideration of Concepts*, 1 ECOSYSTEM HEALTH 1, 1 (1992); Glenn W. Suter, *A Critique of Ecosystem Health Concepts and Indexes*, 12 ENVTL. TOXICOLOGY & CHEMISTRY 1533 (1993).

9. Karr, *supra* note 6, at 245.

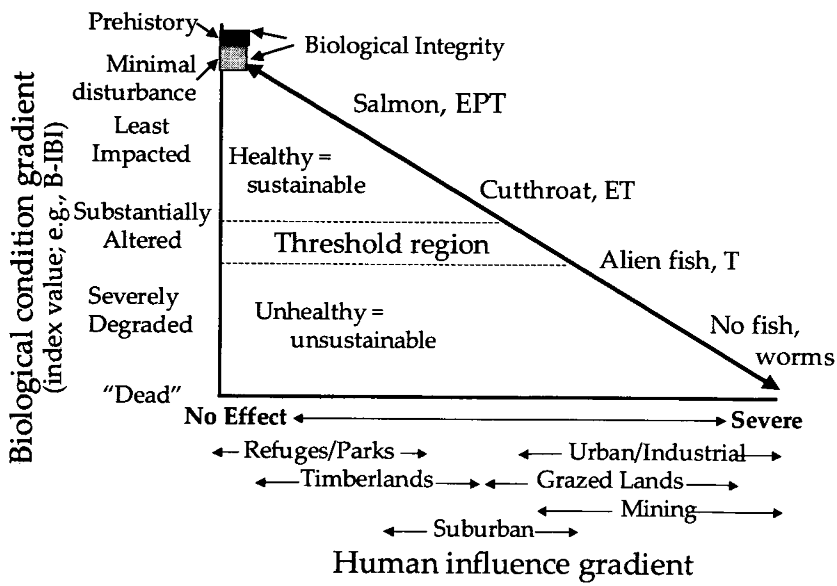
10. *See generally* James R. Karr, *Biological Integrity: A Long-Neglected Aspect of Water Resource Management*, 1 ECOLOGICAL APPLICATIONS 66-67 (1991); Paul L. Angermeier & James R. Karr, *Biological Integrity Versus Biological Diversity as Policy Directives*, 44 BIOSCIENCE 690, 692-95 (1994).

11. LEOPOLD, *supra* note 5, at 190.

benchmark (reference condition) against which other sites can be evaluated. For example, a "normal," or benchmark, body temperature of 37° C (98.6° F) provides a similar standard for individual humans.

Biological integrity refers to the biological character of places with little or no recent human influence, a condition that defines one end of a gradient of biological condition (Figure 1, *y* axis).[12] As human activity alters living systems, they—and we along with them—move along a gradient of measurable biological condition, ultimately to a state where there is little or nothing left alive. The most remote sites in Mt. Rainier National Park approximate biological integrity, and a channelized stream in downtown Seattle, Washington, approaches the little or nothing left alive at the end of the gradient.

Figure 1: Relationship between biological condition and a hypothetical, synthetic measure of human activity, with examples. Different human activities result in biological changes such as different dominant organisms along a descending slope of biological condition. EPT stands for mayflies (Ephemeroptera), stoneflies (Plecoptera), and caddisflies (Trichoptera); see text for details.



12. James R. Karr, *Health, Integrity, and Biological Assessment: The Importance of Whole Things*, *in* ECOLOGICAL INTEGRITY: INTEGRATING ENVIRONMENT, CONSERVATION, AND HEALTH 209, 213 (David Pimentel et al. eds., 2000).

Health is not the same as integrity.[13] Ecological health should be the goal for sites that are "cultivated for crops, managed for tree harvest, stocked for fish, urbanized, or otherwise intensively used."[14] Because integrity in an evolutionary sense is not possible in such areas, healthy land use can be defined as land use that will not degrade the site for future use or degrade areas beyond the site. That is, a socially defined "preferred use" might then be the goal. Because the "preferred use" of a refuge depends on one's point of view, not everyone defines a "healthy" refuge by the same criteria. For a bird watcher, a healthy refuge supports all the bird species expected in the refuge. Duck hunters might prefer a refuge that maximizes local populations of breeding waterfowl. Or they might prefer the maximum number of ducks using the refuge as a stopover site on the way south. For a carp fisherman, a healthy refuge supports an abundance of carp, but someone wishing to go swimming might find the turbid, carp-supporting river unhealthy. Smallmouth bass fishers might prefer a river without carp or swimmers. Defining health for a particular refuge, then, requires scientific, cultural, and social consensus.[15]

GRADIENTS OF BIOLOGICAL CONDITION

Many people would agree that a minimally disturbed site with condition approximating biological integrity is healthy (Figure 1, upper left). Many would also agree that a severely degraded site is unhealthy (Figure 1, lower right). But more than the natural sciences have input in such decisions; social and cultural perspectives also play an important role in judging such sites. If a site's condition is culturally acceptable, it can perhaps be deemed healthy.

Once human actions alter a place so that it no longer possesses biological integrity, the question arises as to whether the place is ecologically healthy. That decision is largely a judgment tied to social values. Does society value the new biological condition? Are there shifts in the parts and processes of living systems (*e.g.*, loss of salmonids as stream temperature increases or of area sensitive birds as forests are fragmented)? These decisions cannot be solely value based because value-based land uses that are acceptable may not be sustainable. When

---

13. James R. Karr, *Ecological Integrity and Ecological Health Are Not the Same, in* ENGINEERING WITHIN ECOLOGICAL CONSTRAINTS 97, 100–02 (Peter Schulze ed., 1996).

14. Karr, *supra* note 13, at 102.

15. James R. Karr & Eriko M. Rossano, *Applying Public Health Lessons to Protect River Health,* 4 ECOLOGY & CIV. ENG'G 3, 14 (2001).

that threshold is crossed, the situation shifts from healthy to unhealthy, from sustainable to unsustainable (Figure 1[16]).

This model explicitly connects diverse concepts such as biological integrity, biological diversity, gradients of human influence and biological condition, and societal values that define the threshold between healthy and unhealthy.[17] It also connects these to the concept of sustainability. These concepts and connections are essential to initiate an operational and integrative management program.

Using this model to understand the integrity and health goals of the Improvement Act, then, requires a biologically defined benchmark (integrity), a measurement of condition scaled to reflect a divergence from integrity, a societal judgment on whether that divergence is acceptable, and, finally, an evaluation of whether that divergence is sustainable in the long term. If it is not, the conclusion must be that, whether valued or not, the situation is not healthy because long-term maintenance of the system and the values that derive from it are threatened.

Defining these concepts, however, is only the first step toward using them in science, policy making, or law. For credibility, practitioners need tools for translating the subjective concept into something objective; they need tools both to quantify and to describe. Fortunately, the toolbox has been expanded in recent decades, enabling practitioners to evaluate sites and rank them with respect to divergence from integrity along a gradient of biological condition.[18]

MEASURING BIOLOGICAL CONDITION

A century or more of land management in refuges emphasized a small subset of species, usually "hook and bullet" species pursued by sportfishers and hunters, although a diversity of refuge goals is obvious from the very early days of the system.[19] Even nonindigenous species were introduced by wildlife managers throughout the nation, an activity that is now generally avoided because it arguably reflects a divergence from biological integrity.

When I was an undergraduate in a fish and wildlife biology program, many students and faculty marginalized species not harvested

16. Karr, *supra* note 12, at 209, 213.

17. *Id.* at 212-13.

18. *See generally* JAMES R. KARR & ELLEN W. CHU, RESTORING LIFE IN RUNNING WATERS: BETTER BIOLOGICAL MONITORING (1999).

19. ROBERT L. FISCHMAN, THE NATIONAL WILDLIFE REFUGES: COORDINATING A CONSERVATION SYSTEM THROUGH LAW 35 (2003).

by sportsmen, also marginalizing fellow students with interest in such species. Small non-game, passerine birds, for example, were referred to with the implicitly derogatory phrase "spatzies."

Classroom and management discussions, as well as the dominant textbooks of the time, commonly focused on tracking population size for sport species or made inferences about those populations on the basis of "habitat" conditions assumed to support those "commodity" species. Management of habitat was the mantra of the time; it was assumed that high populations of priority species were the inevitable result of management for those "preferred" habitats. Many refuges also generated revenue to be transferred to local economies by harvest of commodities such as timber, hay, grass, or row crops.[20]

Slowly, often reluctantly, state and federal agencies expanded their vision to include a diverse array of nongame species and, more recently, a rapidly expanding list of threatened and endangered species. The new goals defined in the Improvement Act complete the transition to more comprehensive biological goals in management of the refuge system. But the transition also means that state and federal agencies need new measurement approaches to track and evaluate refuge condition and management success.

Refuge managers face numerous challenges, as do all professionals evaluating system condition.

> A doctor evaluating a patient depends on the patient to communicate what is amiss....[The doctor may ask the patient to describe her symptoms, request relevant laboratory tests, and]...gather information on relevant environmental factors (home and work; recent travel; tobacco, alcohol, and drug use). But not all patients—infants, someone with severe dementia, or a pet dog or cat, for example—can volunteer such information. Neither can wildlife refuges.
>
> Like competent medical practitioners, refuge managers can deduce refuge condition through standardized "evaluation" procedures. Sampling a refuge's biota with a standard protocol, followed by appropriate data analysis, is a robust way to both measure condition and identify...causes of...degradation. The process combines biological monitoring (sampling the biota of a place) and biological

20. *Id.* at 39.

> assessment (using the samples to evaluate the condition of the sampled place).[21]

Such monitoring programs may include efforts to track the conventional sport species, but they must also track threatened and endangered species, as well as the broader biological context implicit in biological integrity and diversity goals. By combining biological monitoring and assessment with knowledge of relevant environmental hazards (*e.g.*, the kind and extent of human actions in the vicinity of the refuge), managers can improve their ability to protect and restore the biota of refuges. They may even be able to restore integrity in degraded refuges.

Living organisms not only give clear signals about refuge condition, they also attract popular attention, often reaching diverse groups emotionally. Arguably, the primary factors of interest to refuge visitors involve the biota of those places, especially the vertebrate macrofauna and the dominant plants. Migratory waterfowl, for example, are central to the lives of people near many refuges. Citizens identify with carefully expressed measures of biological condition because signals from the biota are more easily grasped intuitively than is, for example, chemical water quality. Photos of a die-off of migratory birds have far greater impact than water chemistry data indicating contamination.

THE HIERARCHY OF INDICATORS

Over the past century, many indicators have been used to understand, regulate, and manage environmental quality. Thirty years of steps and missteps by the U.S. Environmental Protection Agency (EPA) yield important insights for managers of wildlife refuges.[22] Indicators can be arranged in a hierarchy (Figure 2[23]) from counts of management actions (bureaucratic "bean counts") to measures of biological response.[24] Unfortunately, this framework, initially developed in the late 1980s, is not widely employed in the United States, a fact that has been

---

21. Karr & Rossano, *supra* note 15, at 9.

22. ADLER ET AL., *supra* note 2, at 23; ENVIRONMENTAL ANALYSIS: THE NEPA EXPERIENCE 3-5 (Stephen G. Hildebrand & Johnnie B. Cannon eds., 1993); MARC K. LANDY ET AL., THE ENVIRONMENTAL PROTECTION AGENCY: ASKING THE WRONG QUESTIONS 6-9 (1990); Karr & Yoder, *supra* note 1, at 594.

23. *See* Karr & Yoder, *supra* note 1, at 595.

24. ENVTL. STATISTICS & INFO. DIV., OFF. OF POL'Y, PLAN., & EVALUATION, EPA 239-R-95-012, A CONCEPTUAL FRAMEWORK TO SUPPORT DEVELOPMENT AND USE OF ENVIRONMENTAL INFORMATION IN DECISION-MAKING (1995).

criticized recently by both the National Research Council (NRC)[25] and the General Accounting Office (GAO).[26]

Figure 2: Hierarchy of monitoring and assessment indicators. All can be used to measure and manage environmental progress, but only biological responses focus on end outcomes.

1. Management actions 2. Response to management	Administrative indicators [permits, enforcement, plans, grants, enforcement] [technologies used, BMPs installed]
3. Stressor abatement	Stressor indicators [effluent reduction, changes in land-use practices]
4. Ambient conditions 5. Direct exposure to effects of pollution	Exposure indicators [pollutant conc., flow or physical habitat alteration] [assimilation and uptake of pollutants, reduced spawning success, nutrient dynamics changed, sedimentation effects]
6. Biological response	Response indicators [biological metrics, multimetric indexes, target species, other biological measures]

Endpoint: "Ecological Health" or Biological Condition

Key to the selection of indicators is an understanding of the relevance of indicators at different levels of the hierarchy. Although sometimes easier to develop and measure, indicators more removed from direct biological measurements (*e.g.*, administrative, chemical, or habitat measures) are not reliable surrogates of biology. Criteria positioned closer to biological condition are more integrative and accurate indicators to evaluate if defined environmental goals have been attained. Attainment of goals is more dependable, accurate, and scientifically rigorous with biological indicators.[27]

---

25. NAT'L RESEARCH COUNCIL, ASSESSING THE TMDL APPROACH TO WATER QUALITY MANAGEMENT 37 (2001).

26. U.S. GEN. ACCOUNTING OFF., GAO-03-112, MAJOR MANAGEMENT CHALLENGES AND PROGRAM RISKS: ENVIRONMENTAL PROTECTION AGENCY, PERFORMANCE AND ACCOUNTABILITY SERIES 13 (2003).

27. NAT'L RESEARCH COUNCIL, *supra* note 25, at 24-26; Karr & Yoder, *supra* note 1, at 596.

GAO notes that dependence on administrative performance measures (*e.g.*, number of environmental standards established, permits issued, and enforcement actions taken, all referred to as *outputs*) still limits EPA program effectiveness, including the agency's ability to assess risk. In 1999, for example, 86 percent of 278 EPA performance measures were outputs rather than *end outcomes* (direct measures of environmental conditions); the percentage of environmental performance measures increased from 7 percent in 1999 to 27 percent in 2003. Even so, most end outcomes in GAO's analysis were chemical, not biological (the gold standard endpoint). The national refuge system may not fare well if managers do not use end outcomes, especially biological measures, of environmental condition.[28]

Criticisms of EPA performance measures for not focusing on end outcomes have parallels in management of wildlife refuges where, as already noted, a few commodity species and their presumed habitats are the primary focus. Habitat dimensions most commonly mentioned involve measures of vegetation structure, often supplemented with information on plant species composition. But just as water chemistry is not an adequate surrogate of river biology, wildlife habitat measures need to be more rigorously and directly connected to important dimensions of biological condition. Managers should not assume species or more broadly defined biological contexts are present when indicators in the middle of the hierarchy are selected.

Disconnects between human conceptions of habitat requirements and wildlife's actual habitat needs are often large. Woody debris was removed from stream channels in the Pacific Northwest, for example, for years under the assumption that fish passage was enhanced. We know now how important woody debris is to the maintenance of healthy rivers and healthy salmon populations;[29] fishery biologists' early conceptions of optimal habitat were flawed.

BIOLOGICAL MONITORING

Biological monitoring involves sampling the biota of a place (*e.g.*, a stream, a woodlot, or a wetland)[30] to "use...a biological entity as a detector and its response as a measure to determine environmental conditions. Ambient biological surveys and toxicity tests are common

28. U.S. GEN. ACCOUNTING OFF., *supra* note 26, at 13-14.
29. Robert E. Bilby & Peter A. Bisson, *Function and Distribution of Large Woody Debris*, *in* RIVER ECOLOGY AND MANAGEMENT: LESSONS FROM THE PACIFIC COASTAL ECOREGION 324, 324 (Robert J. Naiman & Robert E. Bilby eds., 1998).
30. KARR & CHU, *supra* note 18, at 2, 47.

biological monitoring methods."[31] Data collected in biomonitoring are then used to assess the condition or health of those places (biological assessment).[32] Biological monitoring to track the condition of water bodies goes back to the early twentieth century, when concerns were first raised about organic pollution and associated oxygen depletion in water bodies.[33] Within a few decades, the proliferation of toxic chemicals shifted the focus from biological to chemical monitoring; most efforts to protect water quality, at least in North America, relied on chemical standards.[34]

Direct biological monitoring in North America has come back since the development, widespread testing, and application of multimetric biological approaches to water resource assessment.[35] Recent applications of this approach to wetland and terrestrial environments demonstrate the generality of this approach.[36] Multimetric

---

31. OFF. OF WATER, U.S. ENVTL. PROT. AGENCY, USE OF BIOLOGICAL INFORMATION TO TIER DESIGNATED AQUATIC LIFE USES IN STATE AND TRIBAL WATER QUALITY STANDARDS (forthcoming 2005) (review draft at 148, on file with author).

32. KARR & CHU, *supra* note 18, at 47.

33. Richard Kolkwitz & M. Marsson, *Okologie der Pflanzlichen Saprobien [Ecology of Plant Saprobia]*, 26a BERICHT DER DEUTCSHEN BOTANISCHEN GESELLSCHAFT [REPORTS OF THE GERMAN BOTANICAL SOCIETY] 505 (1908), *translated in* BIOLOGY OF WATER POLLUTION 47 (U.S. Joint Publications Research Service trans., Lowell E. Keup et al. eds., 1967).

34. James R. Karr & Daniel R. Dudley, *Ecological Perspective on Water Quality Goals*, 5 ENVTL. MGMT. 55 (1981); ADLER ET AL., *supra* note 2, at 14.

35. *See generally* James R. Karr, *Assessment of Biotic Integrity Using Fish Communities*, FISHERIES, Jan. 1981, at 21; JAMES R. KARR ET AL., ASSESSMENT OF BIOLOGICAL INTEGRITY IN RUNNING WATERS: A METHOD AND ITS RATIONALE (Ill. Natural History Survey, Special Publication No. 5, 1986); ECOLOGICAL ASSESSMENT SECTION, OHIO ENVTL. PROT. AGENCY, BIOLOGICAL CRITERIA FOR THE PROTECTION OF AQUATIC LIFE (1988), *available at* http://www.epa.state.oh.us/dsw/bioassess/BioCriteriaProtAqLife.html (last visited Dec. 7, 2004); BIOLOGICAL ASSESSMENT AND CRITERIA: TOOLS FOR WATER RESOURCE PLANNING AND DECISION MAKING (Wayne S. Davis & Thomas P. Simon eds., 1995); MICHAEL T. BARBOUR ET AL., EPA 841-B-99-002, RAPID BIOASSESSMENT PROTOCOLS FOR USE IN STREAMS AND WADEABLE RIVERS: PERIPHYTON, BENTHIC MACROINVERTEBRATES AND FISH (2d ed. 1999), *available at* http://www.epa.gov/owowwtr1/monitoring/rbp/index.html (last visited Nov. 21, 2004); OFF. OF WATER, *supra* note 31.

36. Among many examples, see Robert B. Blair, *Land Use and Avian Species Diversity Along an Urban Gradient*, 6 ECOLOGICAL APPLICATIONS 506 (1996), *available at* http://links.jstor.org/sici?sici=1051-0761%28199605%296%3A2%3C506%3ALUAASD%3E2.0.CO%3B2-S (last visited Nov. 21, 2004); THOMAS J. DANIELSON, WETLANDS BIOASSESSMENT FACT SHEETS (EPA 843-F-98-001, 1998), *available at* http://www.epa.gov/owow/wetlands/wqual/bio_fact/index.html (last visited Nov. 21, 2004); Orie L. Loucks, *Pattern of Forest Integrity in the Eastern United States and Canada: Measuring Loss and Recovery*, *in* ECOLOGICAL INTEGRITY: INTEGRATING ENVIRONMENT, CONSERVATION, AND HEALTH 177 (David Pimentel et al. eds., 2000); Sandra A. Bryce, et al., *Development of a Bird Integrity Index: Using Bird Assemblages as Indicators of Riparian Condition*, 30 ENVTL. MGMT. 294 (2002), *available at* http://springerlink.metapress.com/media/1A48DDN1WMCYUMB1XXWW/Contributions/R/G/X/Q/RGXQYVHC81Y83F44.pdf (last visited Nov. 21, 2004); James R. Karr & Diana N.

indexes, modeled after econometric indexes such as the Dow Jones Industrial Average or the index of leading economic indicators, provide more comprehensive and robust assessments than narrowly focused chemical standards, measures of habitat, or population size for target species. They provide a view through multiple biological lenses, a view that was ignored through most of the twentieth century.

Biological assessments as recently developed and applied in many regions of the United States do more than indicate the condition or health of local and regional landscapes. Biological assessments can also aid diagnosis of the cause(s) of degradation, suggest treatments to halt or reverse damage, and evaluate the effectiveness of management actions.[37] Conventional monitoring and evaluation studies, such as tracking chemical pollution or population size of target species, are inadequate to protect overall ecological condition (such as in a wildlife refuge) because they are conceptually narrow. Moreover, they also fail because they are not well suited for distinguishing variation caused by natural events from variation caused by human actions.[38] Neither are they very useful in diagnosis of the specific human actions causing degradation.

SELECTING BIOLOGICAL METRICS FOR ASSESSMENTS

Multimetric biological assessment evaluates multiple dimensions of complex living systems. Success depends on a rigorous process to identify measurable biological attributes that provide reliable, relevant, and easily interpreted signals about the biological effects of human actions. Choosing from the profusion of biological attributes that could be measured is a winnowing process, where each attribute is essentially a hypothesis to be tested for its merit as a metric.[39] One accepts or rejects each hypothesis by asking whether an attribute varies systematically through a range of human influence. Through orderly

---

Kimberling, *A Terrestrial Arthropod Index of Biological Integrity for Shrub-Steppe Landscapes*, 77 NORTHWEST SCI. 202 (2003).

37. For more detailed discussion and examples, see Karr, *supra* note 10, at 66; BIOLOGICAL ASSESSMENT AND CRITERIA: TOOLS FOR WATER RESOURCE PLANNING AND DECISION MAKING, *supra* note 35; KARR & CHU, *supra* note 18; ASSESSING THE SUSTAINABILITY AND BIOLOGICAL INTEGRITY OF WATER RESOURCES USING FISH COMMUNITIES (Thomas P. Simon ed., 1999); Symposium, *Assessing the Ecological Integrity of Running Waters*, 422/423 HYDROBIOLOGIA 1 (2000), *available at* http://ipsapp007.kluweronline.com/IPS/frames/toc.aspx?J=4758&I=79 (last visited Nov. 21, 2004).

38. James R. Karr, *Rivers as Sentinels: Using the Biology of Rivers to Guide Landscape Management*, *in* RIVER ECOLOGY AND MANAGEMENT: LESSONS FROM THE PACIFIC COASTAL ECOSYSTEMS 502 (Robert J. Naiman & Robert E. Bilby eds., 1998).

39. Kurt D. Fausch et al., *Fish Communities as Indicators of Environmental Degradation*, 8 AM. FISHERIES SOC'Y SYMPOSIUM 123, 135 (1990).

selection and organization of metrics, an effective multimetric index can emerge from the chaos of potential biological measures.[40]

Knowledge of natural history and familiarity with ecological principles and theory guide the selection of attributes and the prediction of their behavior under varying human influences. Successful biological monitoring depends most on demonstrating that an attribute has a reliable empirical relationship—a consistent quantitative change—across a range, or gradient, of human influence. Unfortunately, this crucial step is often omitted in many local, regional, and national efforts to define measures of biological condition or to develop multimetric indexes.[41]

Throughout the twentieth century, scientists and managers have used the size of a population (expressed as abundance or density) as the primary measure of population health and thus management success. This use may seem concordant with statutory language from the 1997 Refuge Improvement Act, which establishes a mission for the system of maintaining and, where appropriate, restoring healthy populations of animals and plants.[42] But the situation is not so simple. One might understand healthy "to describe only the quantitative threshold where population levels are sustainable."[43] Other interpretations might focus on quantitative (population size or density) or qualitative (presence of disease or other measures of the vitality of the population such as sex and age distribution) measures.[44] A more inclusive approach would aid a more ecologically informed implementation of the Act.

In my experience, most researchers and refuge managers still assume that population size provides a reliable signal about refuge condition. Yet because species abundances vary so much as a result of natural environmental variation, even in pristine areas, population size is rarely a reliable indicator of human influence except for extreme population densities. Other attributes—such as taxa richness (number of unique taxa in a sample, including rare ones) and percentages of individuals belonging to tolerant taxa—vary consistently and systematically with human influence in many kinds of situations. Such attributes, when graphed, give rise to analogues of the toxicological dose-response curve—called ecological dose-response curves—where

---

40. KARR & CHU, *supra* note 18, at 46.

41. *Id.*

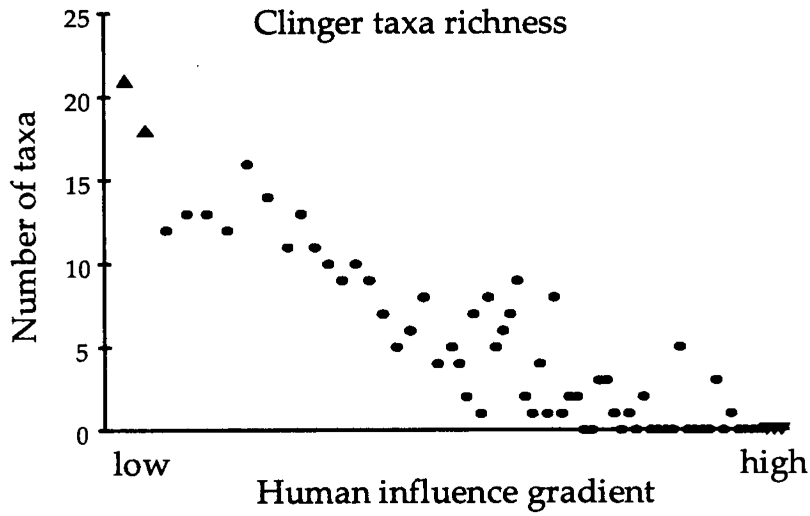
42. National Wildlife Refuge System Improvement Act of 1997, 16 U.S.C. § 668ee(5)(4) (2000).

43. Fischman, *supra* note 19, at 81.

44. *Id.*

the y-axis represents the measured attribute and the x-axis the measures of human influence (Figure 3[45]).[46]

Figure 3: Dose-response curve for taxa richness of clingers—benthic invertebrates that cling to rocks enabling them to live in the interstitial spaces between rocks—in standard samples from 65 Japanese streams ranked according to intensity of human influence.



Ecological dose-response curves differ in one critical aspect from toxicological dose-response curves. Toxicological dose-response curves often illustrate changing survival rates or disease incidence as a function of the amount of a single toxin;[47] as the concentration of DDT in a sequence of aquariums increases, the proportion of fish dying in the aquariums increases. Ecological dose-response curves document biological responses to the cumulative ecological exposure, or "dose," of all events and human activities that influence a place, such as a national wildlife refuge. Measures of human activity may be expressed in diverse ways: percentage of area logged, riparian condition, or percentage of impervious area in a basin.[48]

Unlike streams, where the definition of metrics for use with fish, benthic invertebrates, and diatoms to detect river condition is advanced, the documentation of reliable metrics for application to wetlands or

45. KARR & CHU, *supra* note 18, at 98.
46. *Id.* at 70.
47. MICHAEL C. NEWMAN & MICHAEL A. UNGER, FUNDAMENTALS OF ECOTOXICOLOGY 175 (2nd ed. 2003).
48. KARR & CHU, *supra* note 18, at 48.

terrestrial environments, although less advanced, has still produced important insights in recent years.[49] The approach used to identify the best metrics in streams and rivers, for example, yields ecologically sound and mathematically understandable results in diverse environment types.

Three other issues should be kept in mind during the metric selection process. First, the array of selected metrics should incorporate diverse dimensions of living systems. Robust metrics typically include taxa richness (biodiversity) and composition, tolerance or intolerance of specific environmental stressors, trophic organization (measured as relative abundance of selected trophic groups), health or condition of individuals, and richness or relative abundance of selected ecological groups.[50] The latter may be organized in a variety of ways: by autecology, morphology, reproductive biology, and others. Second, primary measures should capture diverse components of biology, ranging from biomarkers and individual health to population, community, ecosystem, and landscape attributes. Third, measures should be selected that are sensitive to a range of types and levels of human influence (pollutants; agriculture; urbanization; logging; water withdrawal; alteration of physical environments; environmental fragmentation; overharvest by sport, commercial, and subsistence harvesters; introduction of nonindigenous taxa; and so on).

ASSESSING BIOLOGICAL DEGRADATION CAUSED BY HUMANS

Links between biology and human actions came to the forefront more than a century ago. Pollution, particularly from raw sewage, harmed waterways, resulting in passage of the nation's first water quality legislation, the Rivers and Hárbors Appropriations Act of 1899.[51] Overharvesting of waterfowl and other migratory birds combined with destruction of wetlands to decimate populations of many species; a 1913 survey of experts in California concluded that the decline in populations

49. *See* sources cited *supra* note 36.

50. KARR & CHU, *supra* note 18, at 62, 68; LESKA S. FORE, U.S. ENVTL. PROT. AGENCY, DEVELOPING BIOLOGICAL INDICATORS: LESSONS LEARNED FROM MID-ATLANTIC STREAMS, EPA/903/R-03/003 (2003), *available at* http://www.epa.gov/bioindicators/pdf/MAIA_lessons_learned_biology.pdf (last visited Nov. 21, 2004).

51. 33 U.S.C. §§ 403, 407, 411 (2000); *see generally* William H. Rodgers, Jr., *Industrial Water Pollution and the Refuse Act: A Second Chance for Water Quality*, 119 U. PA. L. REV. 761 (1971).

of ducks averaged 50 percent, and geese were down 75 percent.[52] Hunting regulations and wetland protection improved the situation for waterfowl. Efforts to track the health of water bodies were less successful because they focused on the presence of chemical contaminants, under the assumption that chemically clean water was sufficient to protect river health. This assumption proved wrong.

We now know that human influences on terrestrial and aquatic living systems fall into five major classes: physical habitat alteration, modification of seasonal flows, addition of pollutants, changes in energy sources, and shifts in biotic interactions[53] (Figure 4[54]). Given the choice of measuring all such influences or of measuring the condition of the biota—which includes the prime witnesses, and victims, of environmental change—many agencies and institutions are shifting to direct measurement of biological condition. As noted earlier, both NRC and GAO outline the importance of this shift.[55]

Biological monitoring and assessment detects and evaluates human-caused biotic changes apart from those occurring naturally; the techniques are gaining widespread acceptance as part of water and land managers' toolkits. Effective monitoring and assessment programs incorporate four key activities: classify natural environment types, sample the biota, employ rigorous analytical procedures, and communicate results to citizens and decision makers.[56]

52. DAVID S. WILCOVE, THE CONDOR'S SHADOW: THE LOSS AND RECOVERY OF WILDLIFE IN AMERICA 148 (1999).

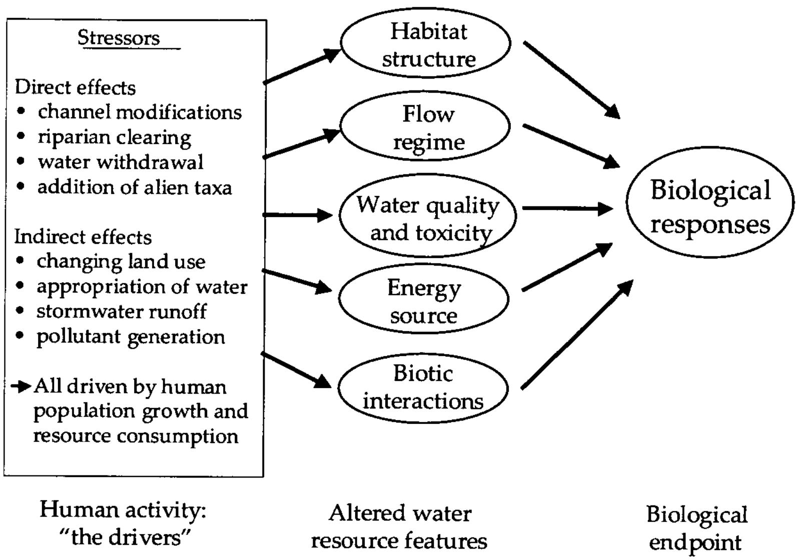
53. *See* James R. Karr, *Biological Integrity: A Long-Neglected Aspect of Water Resource Management*, 1 ECOLOGICAL APPLICATIONS 66, 73 (1991).

54. Karr & Yoder, *supra* note 1, at 599.

55. *See* Nat'l Research Council, *supra* note 25; U.S. GEN. ACCT. OFF., *supra* note 26.

56. James R. Karr & Ellen W. Chu, *Biological Monitoring: Essential Foundation for Ecological Risk Assessment* 3 HUM. & ECOLOGICAL RISK ASSESSMENT 993, 995-96 (1997).

Figure 4: Linkages from human activity (the stressors or drivers of system change) through the five major features of water resources altered by human activity to the biological responses that produce ambient condition, the biological endpoints of primary interest in biological assessment programs. This model illustrates the multiple causes and pathways of water resource change associated with human activities.



Classify Environments to Define Homogeneous Sets Within or Across Regions

Successful biological monitoring and assessment depends on judicious classification to organize sites into relatively homogeneous groups. Excessive emphasis on classification, or inappropriate classification, can impede development of cost-effective and sensible assessment programs. Using too few classes fails to recognize important distinctions among places; using too many unnecessarily complicates development of effective monitoring approaches. Inappropriate levels of classification can lead to poorly informed decisions. An ideal classification system has only as many classes as are needed to represent the range of relevant biological variation in a refuge and the level appropriate for detecting and defining the biological consequences of human activity.[57]

57. KARR & CHU, *supra* note 18, at 118.

Like a taxonomy of places, classification attempts to distinguish and group distinct environments, communities, or ecosystem types; the proper approach to classification may vary, however, according to specific goals.[58] Both abiotic and biotic contexts can be important to an effective classification system (Table 1).

The characteristics that make places similar or different biologically—and thus make classification important for biological monitoring—are determined first by the geophysical setting (including climate, elevation, basin shape, and source and volume of water), and second by natural biogeographic processes. Together they are responsible for local and regional biotas. Prairie pothole wetlands in North Dakota, for example, are likely to be biologically similar, as are bogs in northern Michigan and coastal wetlands in Washington, yet those three wetland types differ from each other in numerous attributes.[59] Within each of those wetland classes, one might expect biological variation as a function of wetland size; the diversity of birds nesting in a wetland increases as wetland size increases.

Classification systems should be used to *guide* monitoring and assessment programs, not define them. The point of classification is to group places where the biology is similar in the absence of human disturbance and where the responses are similar after human disturbance. In some cases, these groupings may coincide with ecoregion boundaries; in others, they may cross those boundaries. To evaluate sites over time and place, we need groupings that will give reliable metrics and accurate thresholds for scoring metrics to represent biological condition. Thus, classification based on rigid application of ecological theory, on strictly chemical or physical criteria, or even on the logical biogeographical factors used to define biological or ecological regions is not necessarily sufficient for biological monitoring. The good biologist uses the best natural history, biogeographic, and analytical information and resources available to choose a classification system.[60]

Finally, classification is only part of the management picture. Scientists and managers too often focus on classifying sites almost to the exclusion of systematically evaluating the effects of human activity. Simply recognizing differences among wetlands or groups of wetlands, among forests or groups of forests, does not mean that all those different classes are relevant to all scientific or management goals. Classification

---

58. *Id.*

59. Mark A. Brinson, *A Hydrogeomorphic Classification for Wetlands*, U.S. ARMY WATERWAYS EXPERIMENT STATION TECHNICAL REPORT WRP-DE-4 (1993).

60. KARR & CHU, *supra* note 18, at 121.

systems—whether based on abiotic factors, vegetation types, ecoregions, or other factors—must match specific management needs.

Table 1. Matrix of classification factors at three spatial scales, from geographical regions to microhabitat zones, for three major environment types (rivers, reefs, and wetlands)

Classification Scale	Rivers	Reefs	Wetlands
Geographic regions	**Comparison among continents** Tropical Asia, Africa, South America, North America **Comparison within continents (N. Amer.)** New England coastal rivers, Tennessee River basin, Puget Sound streams	**Comparison among oceans and seas** Indo-Pacific, Western Pacific, Red Sea, Caribbean **Comparison within oceans and seas (Caribbean)** Northeast Florida coast, Florida Keys, Bahamas, Jamaica, Atlantic Panama	**Comparison among continents** South America, Europe, Africa, North America, **Comparison within continents (N. Amer.)** Florida Everglades, prairie potholes, Puget Sound coast
Stream, reef, or wetland type	Stream size: stream order, watershed area, stream volume Gradient and elevation Water temperature: warm, cool, or cold Water chemistry: hardness	Reef type (pancake reef, fringing reef, atoll) and degree of isolation Temperature (NE Florida or Keys) and salinity (near river mouth or not)	Three hydro-geomorphic variables: 1. Geomorphic: depositional, riverine, fringing 2. Water source: precipitation, surface flow, groundwater discharge 3. Hydrodynamics: vertical fluctuation, unidirectional flow (riparian), or bidirectional flow (e.g., tidal)
Microhabitat zones	Dominant microhabitat features in streams: pools, riffles, raceways or runs, and cascades, defined by the interaction of current velocity, depth, and substrate type.	Dominant reef zones, such as fore reef slope, rubble zone, sediment zone (named for kind of physical environment) or for dominant organisms and primarily determined by light intensity (depth) and wave energy, modified by other factors such as outer reef crests.[61]	Dominant vegetation zones [emergent (smartweed, cattail, bulrush), floating, and submerged plants] or open water, determined largely by depth or duration of flooding (depressional wetlands) or water movement (riparian or coastal wetland).

61. R.R. Graus & Ian G. Macintyre, *The Zonation Patterns of Caribbean Coral Reefs as Controlled by Wave and Light Energy Input, Bathymetric Setting and Reef Morphology: Computer Simulation Experiments*, 8 CORAL REEFS 9 (1989).

Develop Rigorous Sampling Designs and Effective Sampling Protocols

Successful biological monitoring and assessment also depends on the development of rigorous sampling designs and sampling protocols. Inferences about biological condition at a site depend on accurate measures of a site's fauna or flora. These measures should derive from standardized sampling programs, avoid statistical "bias," and focus on components of living systems that are influenced most by human disturbance. Biological monitoring programs need not amass information on all dimensions of natural variation, a point that scientists and managers have often lost sight of. Rather, the goal is to track, evaluate, and communicate the condition of biological systems and the consequences to those systems of human activities. Thus, the spatial and temporal scale of sampling should detect and foster understanding of human influences, not document the magnitude and sources of natural seasonal or successional variation in the same system.[62]

The appropriate sampling protocol varies with environment type and the group of organisms (*e.g.*, fish, birds, plants, insects) selected as the focus of an assessment program. Lessons from streams provide important insights, but the bottom line is that sampling issues will have to be worked out for each environment.[63] For example, sampling of stream fish emphasizes sampling that covers all major microhabitats whereas invertebrate sampling is most convenient and effective if single microhabitats are selected (*e.g.*, riffles).[64]

Define Analytical Procedures to Extract and Understand Relevant Patterns

A critical component of an effective biological monitoring program is the selection of analytical procedures to display and interpret biological data. Graphical methods are "often the most effective way to describe, explore, and summarize a set of numbers (even a very large set)."[65] Graphs reveal the biological responses important for evaluating metrics more clearly than do strictly statistical tools because they exploit

62. Karr & Chu, *supra* note 56, at 999.
63. Karr & Chu, *supra* note 18, at 94-95.
64. *Id.* at 95-96.
65. EDWARD R. TUFTE, THE VISUAL DISPLAY OF QUANTITATIVE INFORMATION 9 (1983).

unexpected patterns that can be discovered in graphs.[66] Researchers and managers must then confront and explain the patterns in those graphs.

For samples where the relationship between human influence and biological response is strong, statistical and graphical analyses lead to the same inferences about resource condition. In other cases, meaningful biological patterns can be lost by excessive dependence on the outcome of menu-driven statistical tests. First, dependence on statistical correlation can miss important relationships if the x-variable (the "dose" of human activity) is measured with low precision. In addition, other factors beyond those plotted on the x-axis may influence metric values. These situations are common.[67] Our ability to measure the condition of biological systems with precision is advanced, but our ability to measure accurately the many dimensions of human activities and their effects on living systems is, by comparison, primitive. Until recently, for example, associations between human influence and stream biological condition were explored by focusing on a narrow range of chemical pollutants.

Second, not all aspects of human influence can be easily captured in a single graph or statistical test. When human actions exist in discrete classes (*e.g.*, road near refuge or not), a single plot against one dimension of human influence will not tell the whole story; neither will a single statistical test.

Third, weak statistical correlation can also miss important biological patterns when the distribution of the data does not lend itself to tests based on standard correlation techniques that detect only linear relationships. Nonlinear patterns are common in field data, as are "factor ceiling distributions"[68] where extensive data fall into a wedge-shaped distribution whose scatter shows little or no statistical significance but can be interpreted biologically.[69] Such situations are likely, for example, where number of species increases across a gradient of forest or wetland patch sizes. But for each refuge size, disturbed refuges have depressed

---

66. *See generally* FREDERICK MOSTELLER & JOHN W. TUKEY, DATA ANALYSIS AND REGRESSION: A SECOND COURSE IN STATISTICS (1977).

67. James R. Karr & Ellen W. Chu, *Sustaining Living Rivers*, 422/423 HYDROBIOLOGIA 14, 10 (2000).

68. James D. Thomson et al., *Untangling Multiple Factors in Spatial Distributions: Lilies, Gophers, and Rocks*, 77 ECOLOGY 1698, 1700 (1996).

69. Kurt D. Fausch et al., *Regional Application of an Index of Biotic Integrity Based on Stream Fish Communities*, 113 TRANSACTIONS AM. FISHERIES SOC'Y 39, 44 (1984). *See also* Tim M. Blackburn et al., *A Method for Estimating the Slope of Upper Bounds of Plots of Body Size and Abundance in Natural Animal Assemblages*, 65 OIKOS 107, 108 (1992); Thomson, *supra* note 68; Frederick S. Scharf et al., *Inferring Ecological Relationships from the Edges of Scatter Diagrams: Comparison of Regression Techniques*, 79 ECOLOGY 448, 448 (1998).

taxa richness, where some human activity in or adjacent to the refuge has reduced the number of species present, "dragging" species richness below the line. Quantile regression, a statistical approach recently imported to ecology from economics, has considerable potential to aid the analysis of situations where only a subset of limiting factors is measured.[70] These new statistical procedures discern more complicated underlying patterns, especially changes near the maxima as in factor ceiling distributions, rather than at the center of response distributions.[71] Graphs force one to search for insights that rote application of statistical tests cannot discover.

Third, graphs highlight idiosyncrasies in data distributions that, when examined closely, may provide insight into the causes of a particular biological pattern. Outlying points on a graph may offer key insights about the complex influence of human activities in or near a wildlife refuge; one can, for example, explore what unique situations exist near the site to cause them to appear as outliers.

Fourth, graphs can be a superior approach to methods that focus on maximum variance extracted because graphs, when used correctly, emphasize ecological rather than mathematical associations, a more appropriate criterion for organizing and understanding complex information.[72]

In short, graphs can be among the scientist's or manager's most useful tools, permitting the exploration of ecological data "before, after, and beyond the application of 'standard' analyses."[73] Rather than choose an inappropriately linear statistical model before plotting their data, ecologists should exploit the power of graphs for "reasoning about quantitative information,"[74] and then choose and apply appropriate statistics. Both extremes—being a slave of standard statistical rules and procedures or judicious avoidance of all statistics—are inappropriate, even myopic.[75]

Another advantage of multimetric indexes is that they are designed to aid scientists, citizens, and policy makers faced with decisions about complex systems—economies, a family member's health, an ecological system. Such decisions require multiple levels of

---

70. Brian S. Cade & Barry R. Noon, *A Gentle Introduction to Quantile Regression for Ecologists*, 1 FRONTIERS IN ECOLOGY & ENV'T 412, 412–19 (2003).

71. Brian S. Cade et al., *Estimating Effects of Limiting Factors with Regression Quantiles*, 80 ECOLOGY 311, 311 (1999).

72. Edward W. Beals, *Ordination: Mathematical Elegance and Ecological Naïveté*, 61 J. ECOLOGY 23, 24 (1973).

73. Carol Augspurger, *Editor's Note*, 77 ECOLOGY 1698 (1996).

74. TUFTE, *supra* note 65, at 9.

75. KARR & CHU, *supra* note 18, at 144.

information, as evidenced by the indexes used to track the health of the national economy: the index of leading economic indicators, the producer price index, the consumer price index, the cost-of-living index, or the Dow Jones industrial average. Although some may prefer maximizing the growth of one econometric—professors' salaries—more rational heads would advocate the integration of multiple economic factors.

The index of leading economic indicators[76] tracks the U.S. economy with 12 metrics: length of work week; unemployment claims; new manufacturing orders; vendor performance; net business formation; equipment orders; building permits; change in inventories, sensitive materials, and borrowing; stock prices; and money supply. These metrics are combined to form the overall index, which takes as its reference point a standardized year (*e.g.*, 1967); the value of the current year's index is expressed in terms of its value in the reference year. Composite economic indexes like these have survived six decades of discussion and criticism and remain widely used by economists, policy makers, and the media to interpret economic trends.[77]

Similarly, physicians and veterinarians rely on multiple measures and multiple tests to assess the health of individual patients. On a single visit to the doctor, a patient might be "sampled" for urine chemistry, blood-cell counts, blood chemistry, body temperature, a throat culture, weight, or a chest X-ray. Clearly, these measurements are not independent of one another, for they come from a single individual whose health is affected by many interacting factors. A doctor would be irresponsible to depend on only one specialized blood test to diagnose overall health; rather, multiple measures yield more-accurate diagnoses. Patterns emerging from multiple measurements enable the doctor to recognize the "signature" of a particular ailment and to suggest more-targeted measurements if she suspects a certain disease. Only then could she prescribe treatment.

The APGAR method for scoring newborns even combines multiple measures into a single index. Newborns are scored for five factors [Activity (muscle tone), Pulse, Grimace (reflex irritability), Appearance (skin color), and Respiration] one minute and five minutes after birth.[78] Scores for each measure are summed with three classes of

---

76. WESLEY C. MITCHELL & ARTHUR F. BURNS, STATISTICAL INDICATORS OF CYCLICAL REVIVALS 162 (1938).

77. Alan J. Auerbach, *The Index of Leading Indicators: "Measurement Without Theory," Thirty-five Years Later*, 64 REV. ECON. & STAT. 589, 589 (1982).

78. *See APGAR Scoring for Newborns, at* http://www.childbirth.org/articles/apgar.html (last visited Nov. 20, 2004).

infant condition recognized on the basis of those scores: normal, requires some resuscitative measures, requires immediate resuscitation.

A core aspect of all these measures of condition, or health, is the concept of reference condition, the expected "normal" in human health. In an ecological context, "reference" is a standard defined as the biological condition and character at minimally disturbed sites.

Rigorous application of concepts such as multimetric indexes and reference condition has not yet developed in refuge management, but it is now a standard for biological assessment in streams.[79] Multimetric biological indexes calculated from ambient biological monitoring data provide an integrative approach for "diagnosing" the condition of complex living systems. The same logical sequence applies in compiling multimetric economic, health, or biological indexes. First, identify reliable and meaningful response variables through testing; then, measure and evaluate the system against expectations; and finally, interpret the measured values in terms of an overall assessment of system condition. The resulting index (for economic or biological resources) or diagnosis (for patients) allows people without specialized expertise to understand overall condition and to make informed decisions that will then affect the health of those economies, living systems, or patients.

Most multimetric biological indexes developed to date, such as the index of biological integrity (IBI), contain 8 to 12 metrics. A few, such as those for species-poor cold-water fish assemblages,[80] have only 4 to 6 metrics, as does the plant IBI recently developed for sagebrush steppe in Washington and Idaho.[81]

The proper number of metrics for a specific refuge and taxonomic group must be developed through study of the system. One should guard against the application of a simple criterion like redundancy as a determinant of what metrics to include. Metrics are not independent because they are calculated from a single collection of organisms, just as multiple personal health tests are done on a single individual. But even if metrics are statistically correlated, they are not

---

79. KARR & CHU, *supra* note 18, *passim*; Robert M. Hughes, *Defining Acceptable Biological Status by Comparing with Reference Conditions*, *in* BIOLOGICAL ASSESSMENT AND CRITERIA: TOOLS FOR WATER RESOURCE PLANNING AND DECISION MAKING 31, 31 (Wayne S. Davis & Thomas P. Simon eds., 1995).

80. John L. Lyons et al., *Development and Validation of an Index of Biotic Integrity for Coldwater Streams in Wisconsin*, 16 N. AM. J. FISHERIES MGMT. 241, 247 (1996); *but see* Robert M. Hughes et al., *A Biointegrity Index (IBI) for Coldwater Streams of Western Oregon and Washington*, 133 TRANSACTIONS AM. FISHERIES SOC'Y 1497, 1500–02 (2004) (describing an 8-metic coldwater IBI).

81. J.R. Karr et al. (unpublished manuscript, on file with author).

necessarily biologically redundant. Rather, just as fever and high white-blood-cell count reinforce a diagnosis of bacterial infection, multiple metrics all contribute to a diagnosis of biological degradation (ecological disease). Extending IBI to new taxa, environment types, and geographic areas is like learning to practice medicine in humans, pets, livestock, and so on: the expectation of what constitutes "health" depends on the species or assemblage of species, but the same fundamental diagnostic strategy applies in all cases.[82]

Critical elements of a rigorous diagnostic strategy are proper classification, metric definition, and the definition of protocols for field sampling and data analysis. Halfhearted or otherwise inadequate treatment of any of these aspects of biological monitoring and assessment is likely to lead to errors that will result in potentially unavoidable refuge degradation.

Communicate Results to Citizens and Policy Makers

Communicating the condition of biological systems and the consequences of human activities to those systems should be the ultimate purpose of refuge assessment. Effective communication can transform refuge assessment from a scientific exercise into an effective tool to influence environmental decision making. Because politics plays such an enormous role in environmental policy decisions, the clarity of scientific information on refuge condition is crucial. Thus, assessment should not be conceived, developed, and implemented as directed by today's flawed regulatory approaches or management frameworks. People need, want, and deserve to understand these issues; refuge assessment focused on biological condition will facilitate such understanding.

Multimetric indexes make it possible to compare sites objectively across geographic regions. Using these explicit cross-region comparisons, citizens and decision makers can better see and understand the consequences of present and planned land-use activities and thereby set priorities for use, protection, restoration, or mitigation. Biological monitoring provides an effective tool for ecological risk assessment. Indeed, biological monitoring is the essential foundation of ecological risk assessment because it measures present *biological* conditions. It provides the means to compare current condition with conditions expected in the absence of humans, or after varied development or restoration scenarios.[83] Biological monitoring is crucial to efforts to

---

82. Karr & Rossano, *supra* note 15, at 14; KARR & CHU, *supra* note 18, *passim*.

83. Karr & Chu, *supra* note 56, at 993.

determine if there is divergence from integrity (natural condition) and why and to avoid activities that degrade refuges and thus erode their ability to support healthy living systems, including priceless ecological goods and services. It is also key to understanding if restoration might be possible and what approaches to restoration might be most likely to succeed.

To protect society's interests in refuges as living systems, we must measure and interpret biological signals. If we do not understand how biological systems respond, and the consequences of those responses for humans, we cannot understand what is at risk from human actions. When biological monitoring and assessment is integrated with knowledge of regional human activities, managers, policy makers, and citizens can use this information to decide if measured alterations in biological condition are acceptable and set policies accordingly.

We cannot halt degradation of the nation's ecological resources if we continue to act as if our activities carried no ecological risks.[84] By enabling us to identify the biological consequences of human actions, biological monitoring and assessment provides an essential foundation to judge ecological risks, as well as to determine if managers are ensuring "that the biological integrity, diversity and environmental health" of the system are being maintained.[85]

84. James R. Karr, *Risk Assessment: We Need More Than an Ecological Veneer*, 1 HUMAN & ECOLOGICAL RISK ASSESSMENT 436, 437 (1995).

85. 16 U.S.C. § 668dd(a)(4)(B) (2000).