James R. Karr[1] and **Diana N. Kimberling,**[2] University of Washington, Box 355020, Seattle, Washington 98195-5020

A Terrestrial Arthropod Index of Biological Integrity for Shrub-Steppe Landscapes

Abstract

We studied arthropods in shrub-steppe at the Idaho National Engineering and Environmental Laboratory to identify scientifically sound indicators of ecological condition and to develop a terrestrial index of biological integrity (T-IBI), an analog of the multimetric indexes used to manage water resources. We sampled terrestrial arthropods with pitfall traps in late spring at nine sites selected to reflect little or no human disturbance and diverse histories of human activity (livestock grazing, chemical contamination, physical disturbance, restoration). Our evaluation of 56 assemblage attributes—measured as taxa richness or the relative abundance of key taxonomic, trophic, or ecological groups—sought measures that varied systematically with human influence. From 21 attributes that did vary significantly with disturbance, we selected 8 for a T-IBI. This Idaho T-IBI correlated significantly with human influence across the nine sites. We compared this index with an independently developed T-IBI, also comprising eight metrics, from three years' study at the Hanford Nuclear Reservation in Washington State. Six metrics were identical in both indexes. Combining our Idaho and Hanford results, we propose a nine-metric T-IBI for shrub-steppe lands. Concordance of metrics and the index in two widely separated shrub-steppe environments suggests that T-IBI offers an effective means of measuring biological condition. Parallels in this terrestrial system with the extensive aquatic work applying the same principles further suggest that this T-IBI could be effective in guiding conservation and restoration decisions in shrub-steppe landscapes.

Introduction

Human activity changes environments, sometimes in big, obvious ways, sometimes in small, subtle ways. Converting shrub-steppe to agriculture or clearcutting a forest inevitably and visibly alters the mix of plants, insects, birds, and mammals at a site. Damming, channeling, or polluting a river inevitably, but perhaps less visibly, alters the river's biota. Such biotic changes are most often documented as counts of threatened and endangered species or as declining populations or production of species with commercial or recreational value. Unfortunately, the preoccupation with imperiled and commodity species obscures wider consequences of human activity for living systems. Without a report card for those wider consequences, society cannot effectively identify and protect ecologically intact places, restore those that are degraded, or make informed decisions about permits for development.

Understanding human impacts on the living world requires well-selected indicators. Good indicators provide ecologically sound, quantitative information relevant to public policy. They represent diverse dimensions of the system under study, simplify information about complex phenomena to improve communication with interested parties, and are cost-effective (National Academy of Sciences 2000). Identifying indicators that meet these standards requires a rigorous scientific process beginning with proper study design, including appropriate sampling and analytical methods, and ending with synthesis and communication of study results (Karr and Chu 1999).

The use of multimetric biological indexes in aquatic systems—specifically the index of biological integrity (IBI)—transformed water resource programs (Davis 1999). IBI was a direct response to the Clean Water Act's 1972 mandate to restore and maintain the biological integrity of the nation's waters. The broad measurement categories included in the original IBI—taxonomic richness, species composition, trophic composition, abundance, and condition—offer robust indicators at multiple biological levels (e.g., individual, population, assemblage, and landscape) (Karr 1991, Karr and Chu 1999). Federal, state, and local agencies now apply these indexes to establish conservation priorities, determine the effects of land use practices, define restoration needs, evaluate the success of restoration activities, and guide regulatory actions throughout the United States (Davis and Simon

[1] Author to whom correspondence should be addressed. E-mail: jrkarr@u.washington.edu

[2] Current address: Department of Forest Science, 321 Richardson, Oregon State University, Corvallis, Oregon 97331-5752

© 2003 by the Northwest Scientific Association. All rights reserved.

1995, Davis et al. 1996, Cuffney et al. 1997, McCarron and Frydenborg 1997, Yoder and Rankin 1998, Barbour et al. 1999, Black and MacCoy 2000, Fore and Grafe 2002, U.S. Environmental Protection Agency 2002, Fore 2003).

Research in aquatic systems has demonstrated the utility of, and parallel conclusions from, studies of fish, benthic macroinvertebrates, and algae (Davis et al. 1996, Karr 1998, Fore 2003). Any major taxonomic group can be examined in an effective monitoring program; more important than which taxon is studied is the wisdom of the process of identifying and selecting metrics (Karr 1991).

The work reported here provides the first terrestrial analog of the aquatic IBI using the IBI development protocol, from study design and data collection to analysis and synthesis (Karr and Chu 1997, 1999, 2000). Key features of this approach include study of multiple study sites with varying histories of human influence, careful examination of data to identify biological measures (metrics) that provide reliable biological signal, and validation that patterns are consistent in multiple data sets (e.g., sample sites, regions, and years). Our studies of shrub-steppe landscapes included data collection at Department of Energy facilities in Washington (Hanford Nuclear Reservation) and in southeastern Idaho (Idaho National Engineering and Environmental Laboratory; INEEL). We sampled arthropods, vascular plants, and the macrobiotic soil crust; here we report only on our arthropod studies.

Arthropods are widely regarded as useful in monitoring and assessing biological systems (Moldenke and Lattin 1990, Kremen 1992, Didham et al. 1998). Although many assessments focus on one arthropod order or family as indicators of disturbance (Refseth 1980, Klein 1989, Neumann 1992, Niemala et al. 1993, Spitzer et al. 1997, Fabricius et al. 2003, Jones et al. 2003), the strength of multimetric indexes derives from the deliberate combination of diverse taxonomic, relative abundance, and ecological or functional measures into a single index. Using multiple metrics strengthens our ability to evaluate multiple dimensions of complex living systems. In addition, multimetric indexes dampen natural variability, or noise, at the metric level, because measures with high variance are excluded. When strong metrics are combined, the resulting terrestrial index of biological integrity (T-IBI), like IBIs for aquatic systems, provides a richer signal than, say, counts of endangered species: first, because it encompasses a number of biological contexts and, second, because it operates much like mathematical averages (Fore et al. 1994, Kimberling et al. 2001).

At Hanford we collected data for three years to identify metrics indicating the effects of human actions (Kimberling et al 2001). Our 25 Hanford study sites included undisturbed sites and disturbed sites (construction, dump, and agricultural). We identified attributes that were significantly correlated with history of human disturbance at our sites in two consecutive years (1997-98). (The probability that an attribute would have been chosen twice because of chance alone was 0.05^2, or 0.0025.) Through this process we constructed a preliminary eight-metric Hanford T-IBI and validated those results with an independent data set from additional Hanford sites in a third year (1999) (Kimberling et al. 2001).

This paper describes the results of a similar study at INEEL and our effort to provide a more robust T-IBI that can be applied to the broader geographic area of northwest shrub-steppe. INEEL has both a large expanse of protected shrub-steppe and a recent history of diverse human activities. For example, approximately 40% of INEEL has been free from grazing for the past 45 yr (Anderson et al. 1996). The site was designated as a National Environmental Research Park in 1975 to protect its high ecological and biological value (Burger et al. 2003).

We sampled sites with various histories of human influence and posed the following questions: (1) How different are terrestrial arthropod assemblages at INEEL sites with different histories of human activity? (2) Can we select robust metrics and construct a multimetric index that is a reliable indicator of site condition? (3) Can the index be used to recognize biological differences among sites with different types of disturbance? (4) Are biological differences similar to those found at Hanford Nuclear Reservation in Washington State (Kimberling et al. 2001), where we applied the same methods, thereby allowing us to propose an index applicable to both regions?

Methods

Study Region and Sites

The 2300 km^2 INEEL features volcanic buttes, rhyolite domes, and unevenly surfaced basalt flows.

The region's relatively high elevation (average 1500 m) and its proximity to northern and western mountain ranges (Lost River, Lemhi, and Bitterroot) result in cold winters and hot, dry summers that constrain plant growth. Native vegetation consists of a shrub overstory and grass and forb understory. Sagebrush (*Artemesia* spp.) inhabits about 80% of the land; perennial grasses (*Poa* spp., *Agropyron* spp., *Festuca* spp., *Stipa* spp.) constitute the most abundant understory plants. Saltbush (*Atriplex* spp.) and winterfat (*Eurotia* spp.) are also widespread.

Cattle and sheep overgrazed most of the Intermountain West in the late 1800s (Wilkinson 1992, Belsky and Blumenthal 1997), but the extent to which grazing affected native plant communities at INEEL is unknown. Historical trails and roads indicate at least some grazing (Anderson et al. 1996). The construction of hundreds of kilometers of canals designed to carry water also influenced the region, although most canals were abandoned early in the 1900s. Before 1949 the military used the area as a testing range for naval guns from the U. S. Naval Ordnance Station in Pocatello, Idaho. Protected from public access for security reasons, the site was designated as a National Reactor Testing Station in 1949, renamed the Idaho National Engineering Laboratory in 1974, and given its present name in 1996.

We selected and sampled nine study sites in 1999 (Table 1, Figure 1) to match major environmental features (slope, soil type, elevation); all naturally supported shrub-steppe vegetation. Although we cannot know if all of our sites had equivalent natural biotas before human activity expanded in these regions in the past 200 yr, we match these natural attributes in an effort to minimize the probability of major differences in initial conditions among our sites. Our experimental design required sampling both minimally disturbed sites (natural areas) and sites with varying histories and types of human activities.

Quantifying Disturbance

To select biological indicators, one must test and calibrate them against a gradient of human disturbance. Ecological principles and theory (e.g., from island biogeography, landscape ecology, and disturbance ecology) suggest that several key landscape features may serve as surrogates of human influence: area disturbed, frequency of disturbance, and time since last disturbance (Table 1). These measures seem both general and largely independent of type of human disturbance (physical, chemical, or biological). A fourth disturbance measure, extent of soil profile disruption, reflects the importance of soil disturbance and the role of the macrobiotic soil crust in regulating soil moisture, plant establishment, and plant and animal persistence in shrub-steppe environments (Belnap et al. 2001). For this study, we rated each criterion as high (2), medium-high (1.5), medium (1), medium-

TABLE 1. Principal human activity, and corresponding disturbance ratings, of the INEEL study sites shown in Figure 1. Disturbance ratings were based on areal extent of disturbance, soil profile disruption, time since disturbance, and frequency of disturbance. A site's overall rating equals the sum of its extent, soil, time, and frequency ratings.

Site code[1]	Activity	Disturbance ratings[2]				
		Extent	Soil	Time	Frequency	Overall
UN-1	Undisturbed	0	0	0	0	0
UN-2	Undisturbed	0	0	0	0	0
CH-1	Aerial chemical	1.5	0	2	2	5.5
RE-1	Restoration	1	2	2	1	6
RE-2	Restoration	1	2	2	1	6
GR-1	Cattle grazing	2	1	2	2	7
GR-2	Cattle grazing	2	1	2	2	7
PH-1	Borrow pit	1	2	2	1	6
PH-2	Gravel pit	1.5	2	2	2	7.5

[1] First two letters of site code designate type of disturbance (UN = no known disturbance, CH = chemical disturbance, RE = physical disturbance followed by restoration, GR = cattle grazing, and PH = physical disturbance).

[2] Extent: 2 = > 200 ha, 1.5 = 21–199 ha, 1 = 6–20 ha, 0.5 = < 0.5 ha, 0 = undisturbed. Soil: 2 = deeply disturbed, 1 = minor surface disturbance, 0 = undisturbed. Time: 2 = < 10 yr, 1 = 10–20 yr, 0 = > 20 yr. Frequency: 2 = chronically disturbed, 1 = infrequently disturbed, 0 = undisturbed

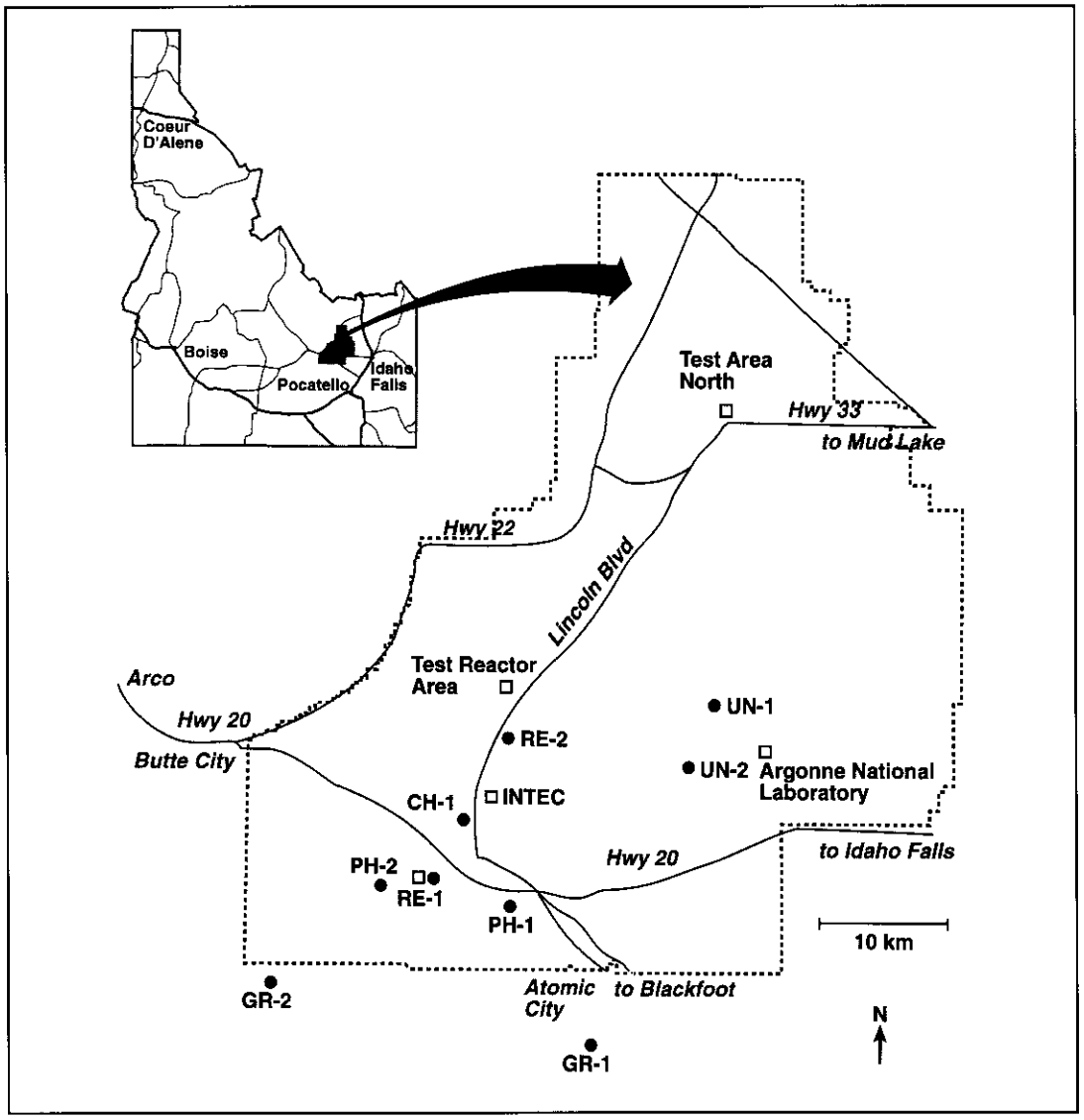

Figure 1. Idaho National Engineering and Environmental Laboratory study site locations (see Table 1 for human activities and degrees of disturbance at each site).

low (0.5), or low (0) to indicate the degree of each type of disturbance. The sum of scores for all disturbance classes is assumed to reflect relative disturbance levels at each site.

Collecting Methods

We collected insects and other arthropods in simple pitfall traps (3 grids of 15 traps at each of 9 sites, see below). Each trap consisted of a 0.5 l plastic cup with a 0.2 l disposable cone cup—its bottom removed to create a funnel—inserted below the larger cup's lip. Each trap held a mixture of 50% water and 50% propylene glycol ~1.5 cm deep in the bottom. Traps were in place for seven nights and all sites were sampled within a 2-wk period (25 May to 7 June 1999). This period coincides with spring emergence and activity of herbaceous plants, grasses, and invertebrates but comes before the severe drought of summer, when many organisms become dormant.

We set three grids of 15 pitfall traps within each study site in three rows of five traps each at 5-m intervals; all study sites and the placement of grids within each site were defined to minimize border effects from surrounding areas with different disturbance histories. At Hanford, 15 traps per grid were sufficient to detect differences among sites reliably (Kimberling et al. 2001); the percentage of total variance due to measurement error (representing variability associated with replicate samples from the same site) declined from 58% with only one trap to 11% with 15 traps per replicate. Measurement error was similar at INEEL with 15 traps per replicate (Figure 2).

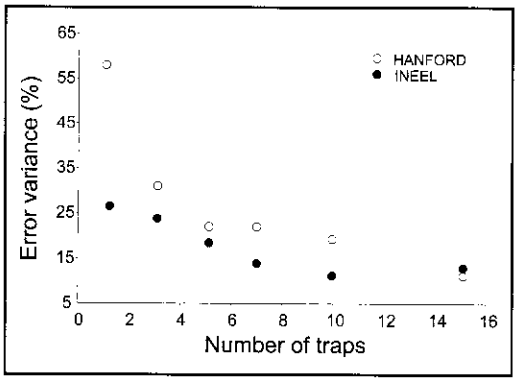

Figure 2. Error variance relative to sampling effort at Hanford (Kimberling et al. 2001) and INEEL. For overall species richness, the percentage of variance due to error declined as the number of traps per composite sample rose. The percentage error at INEEL (13%) came close to that at Hanford (11%) at 15 traps per sample.

Traps were collected at each site in the order they were set. We poured the contents of each trap onto a fine mesh screen (placed over a plastic jug for proper disposal); specimens remained on the screen. Samples were sealed into plastic bags with a 70% ethanol solution, and each grid of samples was put into a jar with 70% ethanol.

We sorted and identified the contents of each sample at the University of Washington in Seattle. Insects and spiders were sorted to order, family, genus (when possible), and species (when possible). Specimens that could not be identified to genus or species were assigned unique numbers as morphospecies. (For choosing metrics, either species or morphospecies would give equivalent taxa richness, or biodiversity.) We also evaluated attributes that applied to broader taxonomic categories than species (e.g., family: all Tachinidae are parasitic flies). We consulted taxonomic experts for further identification within orders and to confirm morphospecies.

Data Analysis

First, we developed a list of 56 attributes of arthropod assemblages for evaluation. These were tested for significant differences (Spearman's r, $P < 0.05$) along a gradient of human disturbance, a winnowing process that yielded 21 attributes for further evaluation. Second, we selected the best eight for inclusion in a multimetric IBI (T-IBI) by removing those that were biological redundant, behaved erratically relative to our hypotheses, or were not significant at $P < 0.01$. Finally, we calculated a T-IBI for each study site and tested the T-IBI values for their correlation with disturbance using Spearman's r.

Whenever an IBI is developed, disparate quantitative scales of the original attributes require converting the attributes' units of measure to a common scoring base (Karr et al. 1986, Kerans and Karr 1994, Fore et al. 1996). We assigned scores of 5, 3, or 1 to each of the metrics with clear dose-response curves, where 5 represented the quantitative values found at minimally disturbed sites (natural or reference condition), 3 represented moderate deviation from reference condition, and 1 represented a strong deviation from reference condition. We assigned the range of metric values associated with undisturbed sites a score of 5 and divided half the range of metric values for disturbed sites and assigned a score of 3 or 1 to each half (Kimberling et al. 2001). Metric scores (5, 3, or 1) were then summed to yield a site T-IBI.

Results

Quantifying Disturbance

Measured by our four disturbance criteria—area of disturbance, soil profile disruption, time since disturbance, and disturbance frequency—disturbance ratings ranged from 0 to 7.5 (Table 1). Sites with rating 0 were the least disturbed (UN-1, UN-2), and sites scoring 7 or more were most disturbed (GR-1, GR-2, PH-2).

Field Samples

Our sample of 405 pitfall traps across the 9 study sites contained 23,165 insects (excluding

Collembola), spiders, and other arthropods representing 19 orders, 140 families, and 531 species. The number of specimens per grid varied widely among the sites, although within a site, sample size varied much less. Undisturbed sites UN-1 and UN-2 averaged 2171 and 773 specimens per grid, whereas the most disturbed site, PH-2, averaged 100 specimens per grid. We captured the fewest arthropods at PH-2, the relatively barren gravel pit often disrupted by heavy equipment. Large numbers of ants (69% of captured individuals) increased sample sizes at UN-1.

Biological Attributes and Human Disturbance

Of the 56 attributes tested, 21 showed significant differences along a gradient of disturbance ($P < 0.05$) (Table 2). For three attributes (relative abundance of the harvester ant *Pogonomyrmex*, relative abundance of parasitoids, and relative abundance of parasitic Hymenoptera), quantitative changes along the gradient, although significant, contradicted our predictions; we predicted that the relative abundance of the ant *Pogonomyrmex* would increase and the relative abundance of parasitoids and parasitic Hymenoptera would decrease along our disturbance gradient. Those attributes were not considered further.

Eight attributes were significant at $P < 0.05$ and 10 at $P < 0.01$. Because our data were limited to only one year, we used a high standard of association and focused on the 10 attributes with $P < 0.01$. Two pairs of biologically redundant measures were included among the 10 (total species

TABLE 2. Predicted responses of attributes of the arthropod assemblage at INEEL sites, and the statistical correlation of these attributes to human disturbance. A *T* preceding a name indicates taxa richness of that group; a *P*, its relative abundance expressed as a percentage of total individual organisms. Correlation was determined by Spearman's *r*; X indicates that the actual response was the opposite of the predicted response. INEEL T-IBI metrics appear in bold.

Attribute	Predicted response	Spearman's r ($n = 9$)	Attribute	Predicted response	Spearman's r ($n = 9$)
Tspecies	Decrease	$P < 0.01$	PCollembola	Increase	
Tfamily	Decrease	$P < 0.01$	**TAcarina**	Decrease	$P < 0.01$
TColeoptera	Decrease	$P < 0.05$	PAcarina	Decrease	$P < 0.01$
PColeoptera	Decrease		TLepidoptera	Decrease	
TCarabidae	Decrease		THemiptera	Decrease	$P < 0.05$
PCarabidae	Increase		PHemiptera	Increase	
Tpolyphagous Carabidae	Increase		Tparasitoids	Decrease	
Ppolyphagous Carabidae	Increase		Pparasitoids	Decrease	X
TElateridae	Decrease		Tparasitic Hymenoptera	Decrease	
PElateridae	Decrease		Pparasitic Hymenoptera	Decrease	X
TTenebrionidae	Decrease	$P < 0.01$	**Tdetritivores**	Decrease	$P < 0.01$
PTenebrionidae	Decrease		Pdetritivores	Decrease	
TEleodes	Decrease	$P < 0.05$	Pdetritivores (minus Collembola)	Decrease	
PEleodes	Decrease		**Tpredators**	Decrease	$P < 0.01$
TAraneae	Decrease	$P < 0.05$	Ppredators	Increase	
PAraneae	Increase		Therbivores	Decrease	$P < 0.05$
Tweb/Araneae	Decrease		Pherbivores	Increase	
Tweb/Araneae	Decrease		Tpollinators	Decrease	
Tno web/Araneae	Decrease	$P < 0.05$	Ppollinators	Increase	
Pno web/Araneae	Increase		**Tground**	Decrease	$P < 0.01$
THomoptera	Decrease	$P < 0.05$	Pground	Decrease	
PHomoptera	Increase		Pdominance (3 families)	Increase	
TDiptera families	Decrease	$P < 0.01$	Pdominance (minus Collembola)	Increase	
TTachinidae	Decrease		Pdominance (minus Formicidae)	Increase	$P < 0.05$
PTachinidae	Decrease		Trare	Decrease	
THymenoptera	Decrease	$P < 0.01$	No. individuals (minus Collembola)	Increase	
PHymenoptera	Increase				
TFormicidae	Decrease				
PFormicidae	Increase				
P*Pogonomyrmex*	Increase	X			

TABLE 3. Metric and index values for all INEEL sites (average of three samples per site; see Table 1 for site descriptions). Metrics beginning with a T measure taxa richness. Numbers in parentheses represent metric scores (5, reference or undisturbed conditions; 3, moderate deviation from reference; 1, strong deviation from reference), assigned on the basis of research at Hanford (Kimberling et al. 2001). Minimum T-IBI value equals 8, maximum equals 40.

Site code	Metrics[1]								
	Tfamily	TTenebrionidae	TDiptera	THymenoptera	TAcarina	Tpredators	Tdetritivores	Tground	T-IBI
UN-1	67 (5)	4.3 (3)	14.7 (3)	22.3 (5)	6.3 (5)	28.7 (5)	17.7 (5)	37.3 (5)	36
UN-2	63 (5)	5.7 (5)	13.7 (3)	25.3 (5)	6.3 (5)	28 (5)	16.3 (5)	36.3 (5)	38
CH-1	55.3 (5)	3.7 (3)	12.3 (3)	21.0 (3)	5.7 (5)	17.7 (3)	12 (3)	29.7 (3)	28
RE-1	55.3 (5)	2.7 (1)	13.7 (3)	21.7 (3)	1.3 (3)	22.7 (3)	14 (5)	33.7 (5)	28
RE-2	54.3 (5)	3.3 (3)	12.3 (3)	22.7 (5)	0.3 (1)	21 (3)	16.3 (5)	31.7 (5)	30
GR-1	49 (5)	2.7 (1)	10.3 (1)	22.0 (3)	3 (5)	14.7 (1)	12 (3)	27 (3)	22
GR-2	48 (5)	2.7 (1)	12 (3)	14.3 (1)	1.7 (3)	15.3 (3)	11.7 (3)	26.3 (3)	22
PH-1	43.3 (3)	1.7 (1)	9.7 (1)	17.0 (3)	2.3 (5)	14.3 (1)	9.3 (1)	25.7 (3)	18
PH-2	26.7 (1)	0.7 (1)	5.3 (1)	9.0 (1)	0 (1)	8.3 (1)	4.7 (1)	7.7 (1)	8

[1] Metrics: Tfamily = overall family richness; TTenebrionidae = Tenebrionid species richness; TDiptera = Diptera family richness; THymenoptera = bees, wasps, and allies species richness; TAcarina = Acarina species richness, Tpredators = predator species richness, Tdetritivores = detritivore species richness, and Tground = ground-dwelling species richness.

richness and total family richness; taxa richness of mites and relative abundance of mites). We therefore dropped one measure from each pair, total species richness and relative abundance of mites, from further consideration. Eight attributes with significant ($P < 0.01$) dose-response curves remained (Table 3): total family taxa richness; fly (Diptera) family richness; species richness of darkling beetles (Tenebrionidae), mites, detritivores, predators, adult ground dwellers, and bees, wasps, and allies (Hymenoptera).

We chose those eight metrics to define a T-IBI specific to our INEEL data set that is significantly correlated ($P = 0.011$) with disturbance rating (Figure 3). The INEEL T-IBI yields site assessments that range from 36 and 38 (Table 3) for the two undisturbed sites (maximum = 40) to 8 and 18 for physically disturbed sites (minimum = 8).

Many dimensions of INEEL's biota—individual attributes and the resulting index—varied among our study sites. Physical disturbance (PH-2) was most devastating to biological condition: family richness at physically disturbed sites differed by 60% from that at undisturbed sites (UN-1, UN-2); richness of fly families by 64%; species richness of ground beetles by 85%, predators by 70%, and detritivores by 73%. Continual operation of heavy equipment in the gravel pit (PH-2), for example, where the soil was compacted, apparently prevented most organisms from establishing resident populations there. Limited recent activity at the borrow pit (PH-1) permitted some

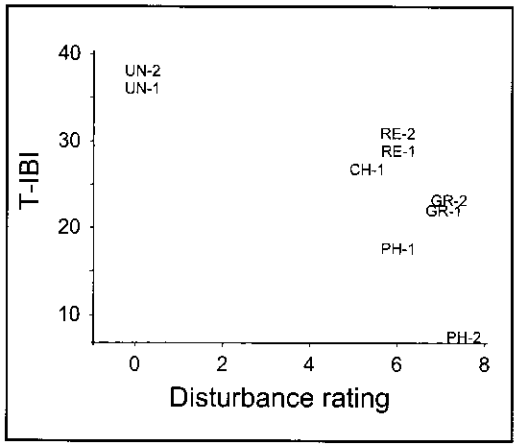

Figure 3. Relationship between disturbance rating at nine INEEL shrub-steppe sites and T-IBI (Spearman's $r = -0.814$).

reestablishment of vegetation and a more diverse arthropod fauna. Sites with some restoration had improved biological condition compared with physically disturbed but unrestored sites. Grazed (22, 22), restored (28, 30) and chemically contaminated (28) sites had intermediate T-IBIs. Severe soil compaction, either from the hooves of livestock or from past construction, influenced these sites.

Discussion

Our research at INEEL demonstrates that we can formulate a measure of human activity reflecting multiple human influences. Moreover, our

exploration of the differences in terrestrial arthropods at sites with different histories of human influence shows that eight biological measures correlate significantly with degree of human disturbance. Those eight measures can be combined to produce a multimetric biological index (T-IBI)—an integrative evaluation of the biological effects of varying histories of human activities at each of the nine INEEL study sites.

T-IBI captures the often extreme biological changes that have occurred, changes that are not normally reflected by more narrow measures such as counts of threatened and endangered species. The level of significance for the correlations between metrics and disturbance rating ($P < 0.01$), combined with divergent index values for sites with different disturbance histories, reinforces our confidence in this approach to biological assessment.

Still, our limited data set (nine sites and one year) make it inappropriate to overinterpret (or extrapolate) our results, even if we can find detailed, statistically significant patterns in the data. Very detailed statistical analyses of small data sets lack credibility in validating a measurement tool for broad applicability. We therefore compared our INEEL results with a larger data set from the Hanford Nuclear Reservation in Washington State (Kimberling et al. 2001).

The Hanford T-IBI was strongly associated with human disturbance in all 3 yr (Kimberling et al. 2001). Index values varied widely, depending on land-use history and current human activity. We found, for example, that the Arid Lands Ecology (ALE) Reserve had a T-IBI of 34 (maximum 40), but the old Hanford town site, abandoned in the 1940s, had a T-IBI of only 10 (minimum 8). Differences in specific metrics from minimally disturbed sites to severely altered areas included species richness, or biodiversity, that was lower in ground dwellers (23%), predators (34%), detritivores (51%), and mites (87%). Family richness was lower by 40%. Our Hanford T-IBI captured variation in arthropods associated with our study sites' different histories of human influence.

Although the terrestrial arthropod faunas at INEEL and Hanford are not identical, they are similar; 35% (187) of the 531 species found at INEEL also occurred in our Hanford samples. We evaluated 56 biological measures from one year of INEEL data; 21 (38%) showed significant dose-response relationships to human influence, a number that is strikingly similar to the 19 (33%) and 22 (39%) of 57 significant attributes in the Hanford data for 1997 and 1998 (Kimberling et al 2001). We used these similarities to develop a generally applicable shrub-steppe T-IBI, in much the same way as a benthic invertebrate IBI (B-IBI) was developed for streams in Tennessee, the Pacific Northwest, and Japan (Karr 1998).

A total of 12 biological attributes met our standards of association between disturbance rating and biological condition at Hanford (8 metrics out of the 12, 2 yr, $P < 0.05$) and INEEL (10 metrics out of 12, 1 yr, $P < 0.01$) (Table 4). Six metrics with clear dose-response relationships to disturbance—total family richness; Diptera family richness; Acarina species richness; predator species richness; detritivore species richness; and adult ground dweller species richness—overlapped at Hanford and INEEL. In other words, the same dominant biological characters changed in response to diverse human influences at shrub-steppe sites of Washington and Idaho. Thus it is reasonable to include at least those core 6 metrics in a preliminary shrub-steppe T-IBI. Indeed, a T-IBI composed of those six metrics correlated very strongly with the disturbance gradient for our nine INEEL sites ($P < 0.001$) (Table 5).

Two metrics in the Hanford index—polyphagous Carabid species richness and relative abundance of Collembola—did not correlate significantly with disturbance rating at INEEL. Conversely, two metrics in the Idaho index—species richness of darkling beetles (Tenebrionidae) and species richness of bees, wasps, and allies (Hymenoptera)—did not go into the Hanford index. When the two Hanford metrics are added to the overlapping core six metrics—the Hanford T-IBI (WA 8)—the resulting index correlates with disturbance at INEEL to the same extent as the original INEEL T-IBI (ID-8, $P = 0.011$; WA-8; $P = 0.004$) (Table 5, Figure 4). In effect, the correlation between the disturbance gradient and biological condition at INEEL is significant whether measured by metrics defined by analysis of Hanford or INEEL data.

Combining analyses at both INEEL and Hanford yields nine metrics correlated significantly with disturbance rating in three out of four years, regardless of study region (Best 9, $P < 0.002$) (Table 5). These 9 constitute the most inclusive list of metrics; biological condition as measured by those 9 metrics is correlated with human disturbance rating at INEEL (Figure 4). We conclude that a

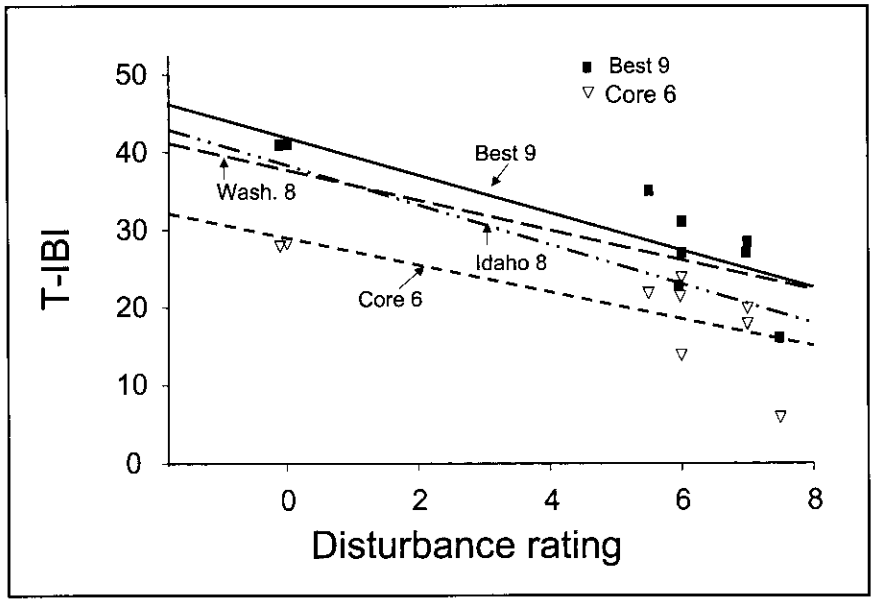

Figure 4. Relationship between disturbance rating at nine INEEL sites and T-IBIs based on four alternative combinations of metrics. Data and regression line shown for Core 6 and Best 9 T-IBIs; regression line only for ID 8 and WA 8 T-IBIs. See Tables 4 and 5 for description of alternative combinations of metrics.

this metric, then, adjusting the boundary levels needed to score a 5 (e.g., from 47 to 60) would likely provide a more accurate assessment of sites at INEEL.

We do not know whether the differences in taxa richness between the two areas reflect differences in the biota (higher natural biodiversity at INEEL) or differences in the history of human influence (more human-caused degradation at Hanford). It will take more study to resolve such scoring-threshold questions. For present purposes, we did not adjust scoring thresholds. The result is that the absolute values of T-IBI for the best Hanford sites fall below those for the best INEEL sites. Without further research we cannot know if the discrepancy reflects real discrepancies in biological condition (histories of human influence) or imprecise scoring because we do not recognize real differences in biological diversity in our two study regions.

Observations

We offer other observations that may help with interpretation of future monitoring and assessment studies. First, the only two metrics that increased significantly with increasing human influence at Hanford—the percentage relative abundance of Collembola and polyphagous carabid taxa richness—were not significant at INEEL. None of the metrics in our recommended nine-metric T-IBI increased with human influence. When faced with severe or incessant disturbance (for example, like that in the gravel pit at INEEL, where little survives), even those groups favored by disturbance may not be able to persist, a situation also encountered in aquatic systems and for which appropriate scoring adjustments are available (Rankin and Yoder 1999).

Second, when choosing metrics for a multimetric index, one should avoid those that are biologically redundant (Karr and Chu 1999). Our decision to exclude two significant INEEL metrics (species richness and mite relative abundance, which were redundant with family richness and mite richness) from the final nine-metric T-IBI did just that. Nevertheless, three of the nine T-IBI metrics were more redundant biologically at INEEL than at Hanford. Nearly two-thirds of detritivore species were also adult ground dwellers, and more than one-third of predator taxa were also ground dwellers. Ground-dwelling taxa are

likely sensitive to soil and understory disruptions that interfere with their life cycles; detritivores may also be more susceptible than other taxa to changes in the soil or litter composition because of their often sedentary nature (Moldenke and Lattin 1990). Predators may be more affected by disturbance as their prey populations decline or become fragmented (Didham et al. 1998). Although excessive biological redundancy among metrics may justify eliminating metrics in some circumstances, we believe that the overlap among our nine metrics is not large enough to outweigh the increased resolution that comes from including more highly significant metrics in the index.

Our combined Hanford and INEEL studies show that T-IBI metrics are sensitive to human influences over more than one isolated geographic area. Our results reinforce the view that the principles embodied in multimetric indexes meet the National Research Council's criteria for good ecological indicators: they reflect diverse dimensions of the system under study, simplify information about complex phenomena to improve communication with interested parties, and are cost-effective (National Academy of Sciences 2000). Measures of biological condition, such as T-IBI, incorporate a depth and breadth that goes beyond counts of threatened, endangered, or commodity species. They also highlight the important links among taxa, composition, and function in ecological systems—the parts and processes that constitute a living system's integrity.

Ecological, especially biological, indicators are central to any effort to measure and understand the environmental consequences of human actions. Because biological indicators track the status of, and trends in, living systems, they provide the single most relevant means of evaluating the effectiveness of conservation, cleanup, remediation, and restoration. They also serve as the most appropriate report card of compliance with federal and state laws, such as the Clean Water Act or state and local growth management acts. Such indicators can and should guide policy and direct research (National Academy of Sciences 2000).

Acknowledgements

Department of Energy Cooperative Agreements #DE-FCO1-95EW55084 and #DE-FG26-00NT40938 to the Consortium for Risk Evaluation with Stakeholder Participation (CRESP) supported this research. Our sincerest thanks to the many individuals and institutions without whose assistance this work would not have been completed: Kris Rein contributed substantially to field and laboratory work. The Environmental Science and Research Foundation (ESRF) in Idaho Falls helped us obtain permits and training required to work on site as well as with solving other logistical problems; R. Morris, S. Majors, and C. Horman of ESRF went beyond their normal work requirements to give exceptional assistance. T. Gooch (BLM) supplied information about grazing allotments, and D. Lowry and B. Pace (Lockheed) cleared our work of archaeological concerns. J. LaBonte, J. C. Miller, R. Zack, J. Longino, T. Griswold, and G. Krantz identified many organisms in our collections; L. Fore provided expert statistical assistance; and E. W. Chu edited the final version of the manuscript.

Literature Cited

Anderson, J. E., K. T. Ruppel, J. M. Glennon, K. E. Holte, and R. C. Rope. 1996. Plant communities, ethnoecology, and flora of the Idaho National Engineering Laboratory. Environmental Science and Research Foundation Report Series, Number 005, ESRF, Idaho Falls, Idaho.

Barbour, M. T., J. Gerritsen, B. D. Snyder, and J. B. Stribling. 1999. Rapid bioassessment protocols for use in streams and wadeable rivers: Periphyton, benthic macroinvertebrates and fish, second edition. Publication EPA 841-B-99-002. U.S. Environmental Protection Agency, Office of Water, Washington, D.C.

Belnap, J., R. Rosentreter, S. Leonard, J. H. Kaltenecker, J. Williams, and D. Eldridge. 2001. Biological Soil Crusts: Ecology and Management. Publication Technical Reference 1730-2. United States Department of Interior, Bureau of Land Management, United States Geological Survey Forest and Rangeland Ecosystem Science Center, National Science and Technology Center, Denver, Colorado.

Belsky, A. J., and D. M. Blumenthal. 1997. Effects of livestock grazing on stand dynamics and soils in upland forests of the interior west. Conservation Biology 11:315-327.

Black, R. W., and D. E. MacCoy. 2000. The development and evaluation of a benthic index of biological integrity for the Cedar River Watershed, Washington. Water Resources Investigations Report 99-4203. U.S. Geological Survey, Tacoma, Washington.

Burger, J., T. M. Leschine, M. Greenberg, J. R. Karr, M. Gochfeld, and C. W. Powers. 2003. Shifting priorities at the Department of Energy's bomb factories: Protecting human and ecological health. Environmental Management 31:157-167.

Cuffney, T. F., M. R. Meador, S. D. Porter, and M. E. Gurtz. 1997. Distribution of fish, benthic invertebrate, and algal communities in relation to physical and chemical

conditions, Yakima River Basin, Washington, 1990. Water-Resources Investigations Report 96-4280. U.S. Geological Survey, Raleigh, North Carolina.

Davis, W. S. 1999. Foreword. Pages iii-iv *In* T. P. Simon (editor). Assessing the Sustainability and Biological Integrity of Water Resources Using Fish Communities. CRC Press, Boca Raton, Florida.

Davis, W. S., and T. P. Simon (editors). 1995. Biological Assessment and Criteria: Tools for Water Resource Planning and Decision Making. Lewis Publishers, Boca Raton, Florida.

Davis, W. S., B. C. Snyder, J. B. Stribling, and C. Stoughton. 1996. Summary of state biological assessment programs for streams and rivers. Publication EPA 230-R-96-007. Environmental Protection Agency, Office of Policy, Planning, and Evaluation, Washington, D.C.

Didham, R. K., J. H. Lawton, P. M. Hammond, and P. Eggleton. 1998. Trophic structure stability and extinction dynamics of beetles (Coleoptera) in tropical forest fragments. Philosophical Transactions of the Royal Society of London. Series B 353:437-451.

Fabricius, C., M. Burger, and P. A. R. Hockey. 2003. Comparing biodiversity between protected areas and adjacent rangeland in xeric succulent thichet, South Africa: arthropods and reptiles. Journal of Applied Ecology 40:392-403.

Fore, L. S. 2003. Responses of diatom assemblages to human disturbance: Development and testing of a multimetric index for the Mid-Atlantic region (USA). Pages 445-480 *In* T. P. Simon, editor. Biological Response Signatures: Indicator Patterns Using Aquatic Communities. CRC Press, Boca Raton, Florida.

Fore, L. S., and C. Grafe. 2002. Using diatoms to assess the biological condition of large rivers in Idaho (USA). Freshwater Biology 47:2015-2037.

Fore, L. S., J. R. Karr, and L. L. Conquest. 1994. Statistical properties of an index of biological integrity used to evaluate water resources. Canadian Journal of Fisheries and Aquatic Sciences 51:1077-1087.

Fore, L. S., J. R. Karr, and R. W. Wisseman. 1996. Assessing invertebrate responses to human activities: evaluating alternative approaches. Journal of the North American Benthological Society 15:212-231.

Jones, D. T., F. X. Susilo, D. E. Bignell, S. Hardiwinoto, A. N. Gillison, and P. Eggleton. 2003. Termite assemblage collapse along a land-use intensification gradient in lowland central Sumatra, Indonesia. Journal of Applied Ecology 40:380-391.

Karr, J. R. 1991. Biological integrity: A long-neglected aspect of water resource management. Ecological Applications 1:66-84.

Karr, J. R. 1998. Rivers as sentinels: Using the biology of rivers to guide landscape management. Pages 502-528 *In* R. J. Naiman and R. E. Bilby (editors). River Ecology and Management: Lessons from the Pacific Coastal Ecosystems. Springer, New York.

Karr, J. R., K. D. Fausch, P. L. Angermeier, P. R. Yant, and I. J. Schlosser. 1986. Assessing biological integrity in running waters: A method and its rationale. Special Publication 5, Illinois Natural History Survey, Urbana, Illinois.

Received 12 July 2002
Accepted for publication 22 April 2003

Karr, J. R., and E. W. Chu. 1997. Biological monitoring: Essential foundation for ecological risk assessment. Human and Ecological Risk Assessment 3:993-1004.

Karr, J. R., and E. W. Chu. 1999. Restoring Life in Running Waters: Better Biological Monitoring. Island Press, Washington, D.C.

Karr, J. R., and E. W. Chu. 2000. Sustaining living rivers. Hydrobiologia 422/423:1-14.

Kerans, B. L., and J. R. Karr. 1994. A benthic index of biotic integrity (B-IBI) for rivers of the Tennessee Valley. Ecological Applications 4:768-785.

Kimberling, D. N., J. R. Karr, and L. S. Fore. 2001. Measuring human disturbance using terrestrial invertebrates in the shrub-steppe of eastern Washington (USA). Ecological Indicators 1:63-81.

Klein, B. C. 1989. Effects of forest fragmentation on dung and carrion beetle communities in central Amazonia. Ecology 70:1715-1725.

Kremen, C. 1992. Assessing the indicator properties of species assemblages for natural areas monitoring. Ecological Applications 2:203-217.

McCarron, E., and R. Frydenborg. 1997. The Florida bioassessment program: An agent for change. Human and Ecological Risk Assessment 3:967-977.

Moldenke, A. R., and J. D. Lattin. 1990. Dispersal characteristics of old-growth soil arthropods: The potential for loss of diversity and biological function. Northwest Environmental Journal 6:408-409.

National Academy of Sciences. 2000. Ecological Indicators for the Nation. National Academy Press, Washington, D.C.

Neumann, F. G. 1992. Responses of foraging ant populations to high intensity wildfire, salvage logging, and natural regeneration processes in *Eucalyptus regnans* regrowth forest of the Victorian central highlands. Australian Forestry 55:29-38.

Niemela, J., D. Langor, and J. R. Spence. 1993. Effects of clear-cut harvesting on boreal ground-beetle assemblages (Coleoptera: Carabidae) in western Canada. Conservation Biology 7:551-561.

Rankin, E. T., and C. O. Yoder. 1999. Methods for deriving maximum species richness lines and other threshold relationships in biological field data. Pages 611-624 *In* T. P. Simon (editor). Assessing the Sustainability and Biological Integrity of Water Resources Using Fish Communities. CRC Press, Boca Raton, Florida.

Refseth, D. 1980. Ecological analyses of carabid communities: Potential use in biological classification for nature conservation. Biological Conservation 17:131-141.

Spitzer, K., J. Jaros, J. Havelka, and J. Leps. 1997. Effects of small-scale disturbance on butterfly communities of an Indochinese montane rainforest. Biological Conservation 80:9-15.

U.S. Environmental Protection Agency. 2002. Summary of biological assessment programs and biocriteria development for states, tribes, territories, and interstate commissions: streams and wadeable rivers. EPA-822-R-02-048. U.S. Environmental Protection Agency, Washington, DC.

Wilkinson, C. F. 1992. Crossing the Next Meridian: Land, Water, and the Future of the West. Island Press, Washington, D.C.

Yoder, C. O., and E. T. Rankin. 1998. The role of biological indicators in a state water quality management process. Environmental Monitoring and Assessment 51:61-88.