Assessing habitat quality and disturbance in Coastal Sage Scrub using an index of biological integrity

Prepared for:
California Department of Fish and Game

James E. Diffendorfer and Rosalie del Rosario

Department of Biology
San Diego State University
5500 Campanile Drive
San Diego, CA 92182

February 28, 2002

TABLE OF CONTENTS

Executive Summary

The report is divided into two parts. The first reviews current methodology for assessing habitat quality and the second describes the methods we will use to develop an Index of Biological Integrity (IBI) for Coastal Sage Scrub (CSS) in Southern California (S. CA). We reviewed only common and modern methods of assessing habitat quality and do not give an historical account of the concept and its applications. Currently, habitat quality is assessed using single species approaches such as indicator or umbrella species, GIS-based methods, or multi-metric approaches such as IBIs. Assessing habitat quality over large spatial scales for use in reserve design is typically done by a variety of GIS-based approaches. However, at smaller spatial scales, the reliability of the GIS-based approaches declines. At these smaller spatial scales, multi-metric indices, such as the Index of Biological Integrity have been successfully developed.

IBI's are developed using a four-step process, which includes: establishing biological dose-response curves, developing scoring systems, selecting metrics and statistical analysis of the IBI, and IBI validation or verification. In the report, we explain each of these steps in detail giving examples from successful IBIs.

Measuring disturbance independently of the biological data is a critical aspect of IBI development. Measurements of disturbance are often specific to the site of interest and to the type of disturbance. Professional judgment often categorizes sites between low, moderate and severely impacted. Only three terrestrial IBI's have been developed and these used professional opinion, GIS-data, historical records of disturbance, and abiotic measures of the environment to place sites into disturbance categories.

In the second part of the report we describe our sampling protocols and IBI development. Sampling will occur across 38 sites, four times a year. The timing of the sampling is optimized to gain the maximum amount of information across all taxa sampled. At each site, a 50 x 50 m grid will sample vegetation, small mammals, birds, arthropods, and herpetofauna (in conjunction with Dr. Robert Fisher, United States Geological Service). Disturbance will be estimated using the level of invasion by non-native annual grasses, GIS-derived landscape variables, and historical information on past land use (for some sites). A host of statistical methods, from exploratory data analysis, to ANOVA's, to boot strapping methods of power analysis, and clustering algorithms will be used to develop and test the IBI.

We end the report by discussing the ancillary information we will generate relating to the ecology of CSS, the development of robust sampling protocols, and integrating data collection and management procedures with USGS.

Part I. Literature review of assessing habitat quality

Introduction

Increased human population size in Southern California (S.CA) alters the diversity and function of natural habitats either directly through displacement by urban and agricultural land uses, or indirectly by a host of processes such as edge effects, road impacts, or altered fire regimes. Land use changes and resulting impacts on habitats engage land managers who must comply with federal and state regulations such as the Endangered Species Act, Clean Water Act, and National Environmental Policy Act mandating some level of protection of natural systems. Government, non-profit, private, and military land stewards face the challenge of balancing economic uses and ecological function of lands they govern. Assessing the balance requires evaluating the quality of specific habitats.

Historically, most land stewards managed one or a few species at a time and typically targeted those of particular concern to the goals of their organization (e.g. sport fish, game animals, livestock, endangered species). In more recent history, land stewards are attempting to manage multiple-species, entire ecological communities of organisms simultaneously, or simply put, focal habitats (such as Coastal Sage Scrub – "CSS").

Managers and scientists, struck by the immense complexity inherent in ecological systems, now attempt to select and develop reliable methods of representing habitat quality for such a diverse array of species. However, standard protocols for assessing habitat quality or the impacts of disturbance on various types of ecosystems do not exist. This lack of uniformity is not surprising given the diversity of ecosystems, complexity of species interactions, and multiplicity of goals among land management organizations. However, without an accepted protocol for characterizing habitat quality, land managers face significant disadvantages when legislatures, courts, or the general public ask them to justify their management decisions. Public sentiment, as well as most federal and state environmental statutes, demands that land managers base their actions on the best-available science. Objective and accurate methods for measuring habitat quality would help land managers communicate the logic and evidence behind their decisions.

Scope of work

For this report, we were tasked with describing "1) a method for estimating human disturbance in CSS habitat including a 2) review of the IBI literature." In addition, the report was to "summarize how estimates of disturbance are used in the development of IBI's and how disturbance has been quantified in other systems."

We review the approaches land stewards have commonly used to assess habitat quality, and we describe how managers assess the degree to which changing land uses (as a form of ecological disturbance) may affect habitat quality of natural lands. Because the literature related to habitat quality is enormous and system specific, we limited our focus to assessment methods currently being used in management, understanding the historical development of habitat quality as a concept is complex and not yet synthesized by science historians. The methods discussed cover a variety of approaches for assessing habitat quality. These approaches vary considerable primarily because the estimates of habitat quality they generate are used for different reasons. In particular, this report focuses on the index of biological integrity (IBI; (Karr 1981, Karr et al. 1986), an empirically tested approach for measuring habitat quality. We end by describing our plans to measure habitat quality and disturbance in CSS communities in S. CA.

Existing approaches for measuring habitat quality

Habitats of any type, found in deserts to rainforests, consist of numerous, interacting species each with unique life cycles and adaptations shaped by long evolutionary histories. As such, measuring habitat quality is extremely difficult because the complexity of species interactions makes defining or gauging "quality" seemingly intractable. Nevertheless, several attempts to develop standard measures of habitat quality for biologically diverse ecosystems exist.

In conservation planning, managers sometimes seek effective shortcuts to conserve biodiversity. Commonly, this involves managing for a few species and assuming protecting these surrogates will confer protection on other species in a region. There are several categories of surrogate species. According to a review[1] of the use of surrogate species in conservation management, (Andelman and Fagan 2000) consider the three most prevalent categories: flagships (i.e. charismatic species that attract public support), umbrellas (i.e. species requiring large areas of habitat, whose protection serves to protect many co-occurring species), and biodiversity indicators (i.e. sets of species or taxa whose presence may indicate areas of high species richness). Other categories of surrogate species commonly used in habitat quality management include: big carnivore, habitat generalists, habitat specialists, high age at first reproduction, long-lived, health indicators, population indicators, population turnover, residency status, etc. (Landres et al. 1988, Caro 1999, Andelman and Fagan 2000).

One of the simpler approaches to conservation management is to focus on a single keystone species, as a proxy for habitat quality. The rationale behind keystone-based measures of habitat quality is straightforward: habitats lacking keystone species must be significantly altered and of lower quality. Keystone species are typically considered those species which, when removed from an ecosystem, cause a cascade of changes in abundance of

[1] This review included biological database from Southern California coastal sage scrub communities (Natural Diversity Database, California Department of Fish and Game).

other species (Paine 1966, 1969). A classic example of a keystone species is the predaceous sea star, *Pisaster ochraceus*, whose presence enhances species diversity by allowing other invertebrates to colonize rocky substrates that would otherwise be dominated by barnacles and mussels.

Given the binary nature of the metric (i.e. keystone present or absent), this concept has limited utility in management. For example, many management units might lack their historical keystone species, yet still support large amounts of diversity. Furthermore, the concept of keystone species has been thoroughly criticized by Hurlbert (1997), who considers it operationally impossible to define (i.e. pine trees might be considered a keystone species in pine forests as much as Grizzly bears) and thus meaningless. Mills et al. (1993) critiqued the use of keystone species in part because although it has applications for conservation and food web theory, it remains largely undemonstrated in nature. In fact, Mills et al. (1993) argue the formalization of the term in laws and policy guidelines would do more harm than good, and the use of this broadly applied and poorly defined term does not allow for the practical use of keystone species in management.

Another common approach is to use a single species as an indicator of biological diversity, species richness or composition. An indicator species is "an organism whose characteristics (e.g. presence or absence, population density, dispersion, reproductive success) are used as an index of attributes too difficult, inconvenient, or expensive to measure for other species or environmental conditions of interest" (Landres et al. 1988). This approach to habitat conservation has had mixed success. The probability a single species could serve as a surrogate measure of habitat quality given the complexity of natural systems is small. In CSS, indicator species of conservation concern could not be assumed to be indicators of hotspots for either bird or small-mammal richness (Chase et al. 2000). Furthermore, in their examination of 40 species of birds and mammals the presence of bird and mammal species were poorly correlated, suggesting managing for a single species would not be effective conservation planning.

A related single-species approach, the umbrella species, attempts to manage for multiple species by conserving a species "with large area requirements, which if given sufficient protected habitat area, will bring many other species under their protection" (Noss 1990). Examples of umbrella species that are proposed to protect other species within their ecosystems include: spotted owls (Franklin 1994), desert tortoises (Tracy et al. 1995), black-tailed deer (Hanley 1993) and butterflies (Launer and Murphy 1994). However, landscapes managed for a single species may fail to meet the needs of other species in a complex ecosystem (Franklin 1994). For example, management plans for large-scale forest reserves to protect the umbrella species, Northern Spotted Owl, did not protect aquatic ecosystems, Marbled Murrelets, and failed to include a large portion of the late-successional forests (Franklin 1994).

Conservation based on surrogate species is a common approach to habitat and ecosystem management because managing for a single species is easier and more practical than managing for complex ecosystems, which require monitoring several biotic and abiotic factors. However, the ability of surrogate species to protect other species in the region are considered inadequate and cost-ineffective by some (e.g., (Franklin 1993, Lambeck 1997, Andelman and Fagan 2000). Debates on how to manage habitats and ecosystems have led to the continued use of both single and multiple species approaches.

Geographic Information Systems (GIS) based approaches to measuring habitat quality

A number of GIS-based methods exist for modeling or predicting habitat quality. All of these methods rely on either known, or assumed relationships between focal species (typically an animal) and habitat (typically vegetation, but may also include other features such as snags, rock outcroppings, urban edge, etc). These methods model species distributions using the habitat relationships and maps of existing land-types, and are typically used to predict both the range of a single species or group of species as well as a map of habitat quality for the species of interest. These types of models are typically called Habitat Suitability Indices (HSI) (Terrell and Carpenter 1997). However, once the predicted species maps are in place, they are often over-laid to predict areas of high species richness and guide reserve design (Habitat Evaluation Procedures - HEP), or discover gaps in the extent of reserve systems (GAP analyses). More computationally elaborate methods allow the weighting of various map layers and rules for inclusion in a potential reserve. Mathematical algorithms then predict a potential reserve design across a landscape that optimizes the reserve selection given the constraints originally set (Chikumbo et al. 2001, McDonnell et al. In Press). The Bureau of Land Management, US Forest Service, and US Fish and Wildlife Service among others have developed several HSI models for different species (e.g. salmon, red tailed hawk), as technical notes to serve as the basis of management decisions. The California Department of Fish and Game (CDF&G) maintains the California Wildlife Habitat Relationship System a system of HSI models for 675 vertebrate species (CDF&G 1999).

HSI models are best used as hypotheses of species-habitat relationships as opposed to causal functions (Morrison et al. 1998). HSI is defined as a linear index representing the capacity of a particular habitat to support a focal species (US Fish and Wildlife Service 1981). HSI's combine a suite of variables thought to correlate with the population size of a species or group of species. Variables might include the number of downed logs, old snags, percent cover of a particular vegetation type, etc. The index typically scales from 0-1 and is the ratio of actual habitat conditions compared to optimal habitat conditions for the species in a specified unit of measure (a km^2 for example). Optimal habitat is defined as that combination of variables resulting in the maximum carrying capacity. The HSI model produces an index assuming a linear relationship between HSI value and carrying capacity (i.e. units of biomass/unit area or units of biomass production/unit area; (US Fish and Wildlife Service 1981).

The Habitat Evaluation Procedure (HEP) is another hybrid single-species / multiple-species approach. It combines the Habitat Suitability Index (quality of the habitat) with the total area of available habitat (quantity). The HEP is a collection of procedural and habitat suitability index models for fish and wildlife species (US Fish and Wildlife Service 1980). The models predict changes to carrying capacity of habitats of the particular species of concern. Some examples of applications of HEP include: assessment of timber-sale impacts on wildlife habitat in the Sierra Nevada, California (e.g. (Doering and Armijo 1986), as well as developing a model for rocky mountain bighorn sheep (Smith et al. 1991). This latter model "combined (1) a quantitative assessment of bighorn range to determine if there are adequate quantities of resources to support a minimum viable population of bighorn sheep, and (2) a qualitative assessment of a range to predict the probable density of bighorns the range can support." HEP guidelines suggest selection of indicators "can be arbitrary or according to some ranking scheme," where the "availability of habitat data" is used as a component of the ranking scheme (US Fish and Wildlife Service 1980).

HEP and HSI rely on evaluation species as indicators of habitat quality as well as predictors of future impacts to habitat quality for other species. (Morrison et al. 1998) argue the strength of HSI and HEP "lies in documenting a repeatable assessment procedure and providing an index to particular environmental characteristics that can be compared with alternative management plans." However, there are several critiques on the use of these models in habitat and ecosystem management. (Landres et al. 1988), for example, suggest the arbitrary nature of selecting indicator species based on availability of habitat data compromises the ability of the indicator species to reflect habitat quality. In addition, several reviews of HSI and other habitat-relationship models of birds and mammals have shown large deviations from species habitat requirements and model assumptions (e.g. (Dedon et al. 1986, Raphael and Marcot 1986, Stauffer and Best 1986). Malanson and Westman (1985) argue HSI models developed from single-species experimental data assume optimum habitat for a species in isolation (i.e. absence of competition) is equal to optimum in the field, without taking into account differences in habitat optima for a species in a community versus in isolation.

GIS based methods of assessing habitat quality can effectively aid decision-making in situations covering large spatial scales at relatively low resolutions. Implementation of these approaches is common in reserve design, or landscape-scale habitat conservation planning. For example, HEP’s are used in a number of the NCCP plans to predict total loss of habitat for target species. Even more recently, optimization-based GIS approaches are being used in the North County MSCHP (W. Spencer, Pers. Comm.). Unfortunately, these techniques are not adequate to assist in management decisions requiring finer spatial resolution, because the estimates of habitat quality they produce are unreliable at these smaller scales. Furthermore, because GIS-based models of habitat quality use wildlife-habitat relationships, some of which are based on assumptions and best professional estimates, they can give little indication of the detailed processes causing changes in habitat quality. Indeed, the habitat quality maps derived from GIS-based approaches are, in reality, spatially-explicit working hypotheses requiring additional testing and study to verify both the patterns predicted as well as the processes generating those patterns (Beutel et al. 1999).

Biological indicators of disturbance

At smaller spatial scales, detailed sampling allows the accumulation of fine-scale data and has lead to numerous efforts to develop indices or scoring systems containing information about habitat quality. The vast majority of work describing habitat quality does so for a single species, typically measuring demographic variables such as survival or reproduction to delineate high from low quality sites. The goal of these studies is to find habitat where the average fitness of individuals is greater than 1, and populations are growing or stable. In this report, we ignore these single species efforts and instead focus on efforts to capture the overall quality of a habitat for multiple species, a primary goal of our research effort.

An increasingly popular multi-species approach is the Index of Biological Integrity (IBI), which evaluates habitat quality through combining a series of empirical tested species response curves. These indices incorporate many attributes of the biological community, which can encompass multiple functional groups (i.e. ground nesting birds, terrestrial insects), trophic levels (producers, herbivores, meso and top predators), or unique species to evaluate human disturbance effects on habitats (Karr 1991). The IBI was first developed for fish communities in mid-western streams of the US (Karr 1981, Fausch et al. 1984), and is predominately used to evaluate aquatic ecosystems[2]. However, this approach has increasingly been adopted for use in other ecosystems to evaluate habitats of aquatic invertebrates (Lenat 1988, Lang et al. 1989, Plafkin et al. 1989), birds (O'Connell et al. 2000), terrestrial insects (Kimberling et al. 2001) and coral reefs (Jameson et al. 2001).

The IBI is developed by sampling across a gradient in anthropogenic disturbance and quantifying the systems response to human impact. Thus, IBI's are based on empirical relationships showing the response of taxa in a system to varying levels of disturbance. Scientists who developed the IBI realized ecosystems vary naturally across space and through time yet managers and decision makers were more interested in the added variation caused by human disturbance than they were in natural levels of variation. IBI's attempt to extract and measure this additional, human caused variation

[2] Over 91% of biological assessments using IBI are in aquatic systems, according to a Biosis search spanning 1985-2001.

Ecologists have long known the quality of habitats can be significantly altered by disturbance events. These disturbance events or processes are likely correlated with one another and rarely, if every, impact CSS in isolation. Disturbance events can be natural events such as fires or they can be anthropogenic, such as livestock grazing, residential and commercial development, and road construction. Disturbance adds to the levels of variability we see in ecosystems and may play a critical part in maintaining biodiversity. For example, some unique plants in prairie systems are only found on prairie dog mounds or along the edges of the Buffalo "dirt bath" depressions. However, anthropogenic disturbances are typically much larger in magnitude than most natural disturbances, making them far more devastating than the kinds and types of disturbance in which a system has evolved. Disturbance events occurring too frequently or fundamentally altering basic life-support systems (such as soil quality, hydrology, and light/shade) can have severe impacts on biological diversity. For example, abnormally high fire frequency in S. CA often results in a substantial change in species composition dominated by fast growing non-native grasses (Zedler et al. 1983, Haidinger and Keeley 1993). Because the response of natural systems to disturbance is often specific to the ecosystem, habitat type and/or region, managers must select biological indicators (metrics) that respond to human disturbance in detectable and consistent ways. In our work, we consider the following disturbance: Fire, Grazing, Edge, Roads, Mechanical disturbance, Agriculture, Air pollution, Light pollution, Habitat Fragmentation, and Recreation.

Introduction/overview of IBI's.

The general components of an IBI model involve an iterative process of establishing a biological response to gradients of human disturbance, developing scoring criteria, selecting a subset of biological metrics to include, and having independent data sets to verify and validate the IBI model[3]. Here we describe each of these steps in detail.

Step 1. Establishing biological dose-response curves. The first step in developing an IBI as a model to evaluate ecological condition is to establish empirical relationships between biological metrics and human disturbance. This step consists of two parts. First, a gradient of human disturbance must be developed. This entails sampling areas, which serve as reference sites. The reference sites are assigned a ranking along the disturbance gradient based on a combination of abiotic factors, depending on the scope of the study. Some examples of abiotic factors in stream systems include measures of water quality, level of urbanization or

[3] For detailed descriptions describing the method used to develop an IBI, refer to (Karr et al. 1986, Kerans and Karr 1994, Fore et al. 1996, Brooks et al. 1998, O'Connell et al. 2000)

agriculture along a stream, and distance downstream from a point source of pollution. The ranking is often qualitative and includes a three-category scale of low, moderate, and severely disturbed sites. Professional judgment of field biologists familiar with the study sites is typically used to assign site ranks. We discuss our methods for ranking sites in more detail below. Having established a gradient of human disturbance, the next step is to sample the biota. When combined across sites, a biological dose-response curve is created and indicates if and how particular species, or taxa respond to disturbance.

Step 2. Scoring systems. Once empirical relationships between disturbance and biological metrics of the system are established, a scoring system is developed to allow the ranking of the sites and a method of comparison. In general, each site gets a score for each metric and then the sum of the scores across all metrics is used to rank sites.

To date, scoring systems are quite arbitrary and vary across investigators, the ecosystems where IBI's are being developed, and across the taxa being used in the IBI. For example, in many aquatic IBI's, and in the methods described by Karr and Chu (1999), scoring entails trisecting the range of values of the metric and assigning an arbitrary value of either 1, 3, or 5 to each section, with 5 representing the least impacted site. However, O'Connell et. al (2000) used a ranking scheme of 1, 2, and 3 in their bird-based IBI. They ranked sites with highest occurrence of specialist guilds, reflecting highest biological integrity, with a "3", next highest a "2", and the lowest a "1", and used the reverse order to assign sites with highest occurrence of generalist guilds a "1", and lowest a "3".

As a site's ranking is dependent on undisturbed sites, selecting reference sites plays an important role in developing an IBI model. Biological communities at a disturbed site are compared to communities at a relatively undisturbed "reference site". However, because unimpaired ecosystems may no longer exist, an estimate of expected biological integrity in ecosystems is often based on "least-impacted" conditions (Davis and Simon 1995). It is these least-impacted sites that represent one end of the spectrum in the gradient of disturbance. Suggested criteria for reference sites are that they be: 1) accessible for monitoring over multiple years (e.g. usually public lands), 2) representative of land types and landscape settings commonly impacted during the permitting process, and 3) selected at random (Brooks et al. 1996).

Step 3. Metric selection and statistical analyses

A study conducted across sites varying in the degree of disturbance will generate massive amounts of data and a large number of biological dose response curves. A successful IBI does not need to include all possible metrics collected in a study. Unfortunately, predicting (*apriori*) those biological attributes showing detectable responses to disturbance is difficult if not impossible.

Thus, biologists often start with a list of all pertinent biological attributes, knowing it will be reduced. Karr and Chu (1999) recommend an IBI includes three categories of metrics: species richness and composition, trophic composition, and taxa abundance and condition. Many IBI's we reviewed follow these general guidelines. The justification for the categories is not arbitrary as each of the categories makes up a unique aspect of a biological system. Species richness metrics include information about the make-up or composition of a system. Typically metrics include overall species richness of tolerant taxa, or the number of invasive species. Trophic composition includes information about food web complexity, which can often decline in degraded systems because species interactions are altered. Thus, trophic composition metrics indirectly measure the integrity of trophic interactions (i.e. predation) and include metrics such as the number of herbivores and/or carnivores, and the presence/absence of top-predators. Finally, taxa abundance and condition metrics reflect information about individual and population level processes. Metrics might include the proportion of individuals with deformities, average body condition of key species, or the relative abundance of particular taxa.

In selecting metrics for an IBI model, several criteria must be met. First, each candidate metric's sensitivity to human disturbance should be evaluated, such that explicit hypotheses on how each metric responds to a particular disturbance should be tested. For example, a confirmed hypothesis that intensive cattle grazing in coastal sage scrub should decrease species richness of plants would demonstrate that plant species richness is a suitable biological metric for evaluating grazing disturbance effects on coastal sage scrub systems. This relationship between biological response (e.g. species richness) to different levels of disturbance (e.g. light, moderate, heavy grazing intensity) is an underlying principle in generating a "dose-response curve," (Karr et al. 1986) where the biological attribute is plotted against a gradient of disturbance. For a successful biological metric, a clear relationship between the IBI and disturbance gradient should be detectable or obvious (Figure 1).

Second, the metric should adequately distinguish sites with different levels of disturbance (Kerans and Karr 1994). Not only should plant species richness decrease in grazed areas, but species richness should have distinguishable responses to low, moderate, and heavy grazing intensity.

Third, the successful metrics should not be redundant. Because the IBI model is a multi-metric index, redundant metrics should be avoided. For example, an IBI model that contains all metrics evaluating species richness in coastal sage scrub vegetation would be redundant.

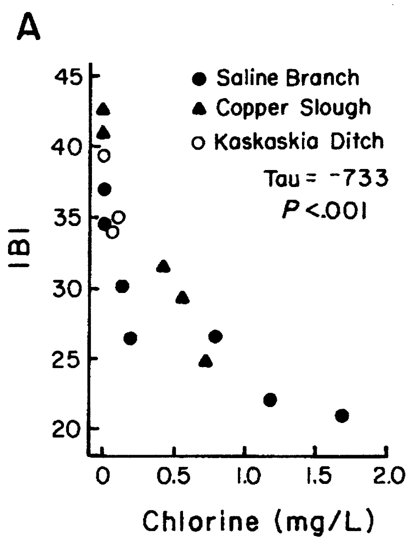


Figure 1. Index of biotic integrity (IBI) as a function of total residual chlorine content in three streams in east-central Illinois with wastewater inflow from standard secondary treatment with chlorination (from Karr, JR., R.C. Heidinger, and E.H. Helmer. 1985. Sensitivity of the index of biotic integrity to changes in chlorine and ammonia levels from wastewater treatment facilities. Journal of the Water Pollution Control Federation 57:912-915).

Fourth, a robust metric should correlate with more than one measure of disturbance (e.g. land use, soil condition, etc.) An IBI's ability to assess biological responses to human disturbance is strengthened by including in the model several measures of disturbance that can be independently quantified. If plant species richness were a robust metric, the metric would respond along the gradients of independent measures of livestock grazing, such as grazing intensity, soil compaction, and time since grazing.

Lastly, the metric should respond to disturbance despite natural temporal or spatial variability. One challenge to biological assessments is the inherent natural variations of populations, which may undergo statistically significant fluctuations even in undisturbed systems (McBride et al. 1993). The difference between statistically significant and ecologically significant results is often clarified by graphing the biological attribute against the established gradient of disturbance to detect potential trends (Figure 2).

Once a relationship is established, the biological attribute is a candidate component of an IBI. Furthermore, the selected metrics should have little or no overlap between poor and good habitat conditions. Fore et al. (1996) argued it is more biologically useful to select metrics from plots of metric vs. a specific or cumulative measure of disturbance, than to rely on statistical tests that typically focus on organisms' abundance rather than their biology.

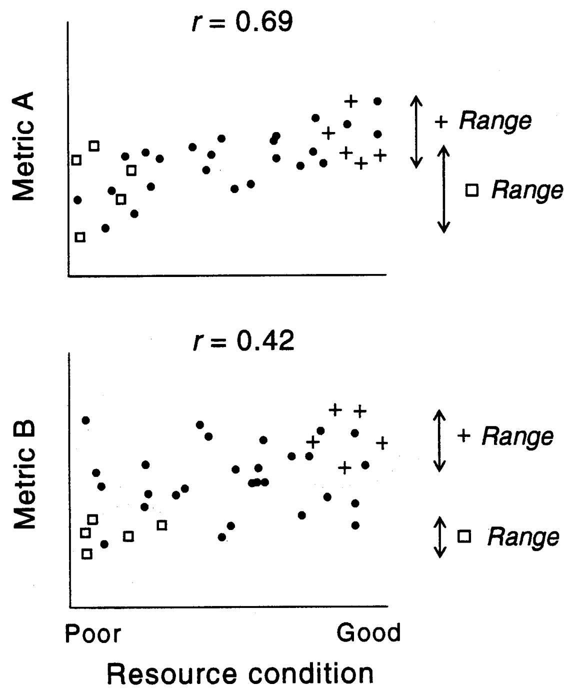


Figure 2. Hypothetical relationships between resource condition and candidate biological metrics. Metric A is more strongly correlated with resource condition (or r2 is higher if using regression) than metric B, suggesting it is a better metric. A more reliable test compares the metric's ability to distinguish between minimally disturbed (+) and severely degraded (square) sites (ranges noted by arrows). Note that moderately degraded sites (shaded circle) span the range between the extremes. Metric B is the more effective metric in spite of its smaller statistical correlation. (From Fore et al. 1996).

The process of selecting metrics and testing IBIs. A general methodology exists for selecting metrics given the framework described above. First, exploratory data analysis is performed with the goal of identifying potential metrics, then measuring their response to disturbance and their correlation with other potential metrics. Second, as metrics are chosen and a scoring system devised, statistical methods are used to determine the ability of the IBI to discriminate between levels of disturbance. These two steps may occur iteratively as initial metrics are placed in a scoring system then tested and metrics dropped or added to improve the IBI. The specific statistical tests or procedures used to select metrics and test an IBI vary across researcher and the system they study.

For example, Kimberling et al. (2001) used nonparametric Mann-Whitney U tests to determine if metrics could distinguish undisturbed from disturbed sites and Spearman rank correlations to test for correlations between a metric and levels of disturbance. The authors also performed separate tests in each of the 2 years and only used metrics producing similar patterns in both years to increase the chance of choosing consistent metrics. Finally, they performed exploratory Discriminate Function Analysis (DFA) to determine if a multivariate statistical approach would rank the sites similarly to the multi-metric IBI. In this study, two metrics were redundant so one was discarded. The DFA ranked sites similarly to the IBI.

O'Connell et al (2000) used ANOVA's, and Spearman Rank Correlations to determine if metrics generated by their bird surveys varied in value across three categories of wetland rankings (high, medium, low quality - the ANOVA's) or if the rank of wetlands was correlated with the value of the metric. Furthermore, the authors used Cluster Analysis to determine the maximum number of categories of sites with statistically distinguishable bird communities. Metrics generated from the bird point count data varied across wetland categories, indicating they could be reliable indicators of disturbance. Furthermore, a number of large-scale variables measured using GIS explained significant amounts of variation in the bird community, suggesting both localized disturbances as well as larger-scale changes in landscape impact bird communities. The cluster analysis indicated the IBI could distinguish between five categories of disturbance with statistical confidence.

Finally, Fore et. al (1994) provide an excellent example and discussion of distributional considerations and methods for use when evaluating the statistical properties of an IBI. They demonstrate the effectiveness of bootstrapping and power analysis to determine the number of categories of disturbance an IBI can distinguish.

Step 4. Validation and verification.

After demonstrating an IBI can discriminate sites with varying levels of disturbance, the final model should be verified. Verification is a process where the models predictions are tested using a new set of data, independent of those data used to construct the model. In the case of an IBI, one can apply the IBI to an independent set of study sites that are representative of the study area. If possible, data on biotic and abiotic condition is collected on the new sites using identical techniques to those used during IBI development. The biotic data from the new sites are then used to generate a ranking of the sites based on the IBI model. This ranking is then compared to the actual level of disturbance at the site derived from the abiotic variables. If the predicted ranking and the actual ranking are similar, the IBI has been successfully validated.

O'Connell et al. (2000) could not collect identical types of abiotic data for new sites, so instead of validating the model, they performed an ingenious analysis they called model "verification". First, they showed the original abiotic data used to rank sites by disturbance and generate the IBI was highly correlated with a new ranking system generated from a GIS using landscape variables such as land use and vegetation type near the bird transects. With this correlation in hand, they then ranked 126 new sites using the GIS methods. Thus, they now had a new set of 126 sites ranked using an algorithm highly correlated to the original ranking method used to generate the IBI. They then sampled the bird community at the 126 sites and used the IBI to rank the sites. The IBI model was verified by a strong positive correlation between the ranking of the 126 sites using the GIS approach and the bird-based IBI.

We note that very few of the papers we reviewed on IBI development included the critical phase of model validation. Model validation is expensive, basically requiring a repeat of the same experiment but in a new set of locations. Given that after the IBI is developed, statistical tests show the metrics vary across levels of disturbance and the IBI can distinguish sites, it is not difficult to see why the motivation to continue work on the model would wane. However, data collected during one or two year IBI development period, may not be representative of the true state of a system, thus model validation through time is critical to a well supported IBI.

How disturbance is measured in IBI's

Measuring the effects of disturbance on habitats and ecosystems in IBI studies requires 1) characterizing disturbance in terms of type (e.g. logging, fire, grazing, etc.), and 2) establishing the scales at which disturbance will be measured (e.g. landscape, habitat, local). These steps are crucial in the development of an IBI because the disturbance gradient is the foundation on which biological responses are measured and sites are assessed. Measuring several levels of disturbance along a gradient allows for more sensitive detection of biological response to particular levels of disturbance, though in practice, this is rare.

When assessing levels of disturbance, professional judgment is commonly involved. Professional judgment typically draws the line between minimally, moderately, and severely impacted sites. Measurements of disturbance are often specific to the site of interest and to the type of disturbance. The majority of disturbance evaluations are entirely qualitative in their assessments (e.g. low, moderate, highly disturbed sites; (Brooks et al. 1996) and only in rare cases do IBIs include a quantifiable scale of disturbance (Kimberling et al. 2001). Brooks et al. (1996) classified wetland sites under three categories (vegetation, water quality, and surrounding landscape condition), they considered a site to be pristine if it was considered pristine for two of three categories, while being mildly disturbed for the third category.

In aquatic systems, disturbance levels are directly measured using both local and larger-scale variables. For example, disturbance is readily estimated at local scales using characteristics of the stream environment such as the presence/absence of channelization, impoundments, and stream bank vegetation, as well as water quality variables such as turbidity, dissolved oxygen, or specific contaminant levels. Larger scale variables include amount of impervious area (i.e. concrete) within a set distance to the stream, the distance from an upstream point source, or estimates of land-use (urban/agricultural vs. natural) within the watershed. Thus in aquatic stream systems, ranking sites based on their level of disturbance is relatively straightforward. For example, Fore et al. 1996 measured human influences at two scales: 1) at a watershed scale based on USGS data on watershed area, % logged, road length, and 2) at the riparian scale based on a resident hydrologist's professional assessments of the conditions of the riparian corridor, stream bed, bank stability, and influences of road building and logging on stream channels. They then plotted each measure of human influence against the other (e.g. riparian vs. watershed assessments) to confirm data consistency and identify unexpected outliers.

Measuring disturbance in terrestrial IBIs. Unlike aquatic systems, few IBI's exist for terrestrial systems. A number of authors have discussed the use of biological data as an indicator of either disturbance or biodiversity in terrestrial systems (Kremen 1992, Kremen et al. 1993, Oliver 1993, Weaver 1995, Oliver 1996, McGeoch 1998), yet few terrestrial IBI's exist. As such, there is not yet a prescribed method for measuring disturbance in terrestrial systems.

In terrestrial systems, local disturbances at fairly small spatial scales (1-10 m) can impact local vegetation and hence potentially alter the value of metrics used in an IBI. Thus, developing IBI's for use at spatial scales where many smaller scale management decisions take place; controlled burns, restored sites, invaded patches, road cuts or fire breaks (i.e. 1-100ha) would seemingly require detailed information about current and past disturbance at a fine spatial scale. Thus, ranking sites *apriori* based on non-biological data to develop the biological-dose response curves is potentially problematic because gaining such fine-scale information may be impossible for some sites.

Kimberling et al. (2001) were able to develop disturbance estimates by using past land-use histories for 25 sites at the Hanford Nuclear Reservation. They categorized each site coarsely into undisturbed or disturbed and further subdivided disturbed sites into those with mechanical disturbance, those with past agriculture, sites where buildings once stood, or sites used to dump toxic chemicals. Given the known history of Hanford, each site was also scored based on the extent, time, frequency, and impact to soil of the disturbance. Given this ranking, they successfully developed an IBI based on arthropods.

Work by Brooks and O'Connell focused on using bird communities in wetland systems in the Mid-Atlantic region (Brooks et al. 1996, O'Connell et al. 2000). In these studies, the wetlands chosen to create the biological dose response curves were previously ranked during a large effort to assess and protect wetlands in Pennsylvania. In these studies, wetlands were evaluated and ranked in a three-category scale based on soil properties, sediment deposition, vegetation characteristics and amphibian surveys. In addition to these rankings, this successful IBI collected bird data at relatively large spatial scales (up to 2 km transects), allowing the researchers to measure disturbance using aerial photography and GIS. They characterized the amount of different land uses (urbanization, agriculture) or vegetation types within a circle (1 km in diameter) surrounding the site where biological metrics were measured, in addition to collecting local vegetation data. Given the large-scale sampling of birds, the metrics responded well to changes in landscape structure caused by urbanization or agriculture.

Finally, Bradford et al. (1998) when developing a bird-based IBI for Great Basin rangelands, used the professional opinion from "local range scientists" to categorize sites into low, medium and high levels of impacts from cattle grazing. In addition, low impact sites were protected from grazing for "many decades", while high impacts sites had known heavy grazing.

Part II. IBI methods in Southern California

Introduction

We describe a procedure for assessing the biological integrity of CSS, and describe how we plan to quantify levels of disturbance. Because we are interested in assessing the biological integrity of the entire CSS community, we plan to develop several IBI's specific to the taxa reviewed in Diffendorfer et al. 2002: vegetation, mammals, birds, reptiles and amphibians, and arthropods. After developing these separate IBI's, we will devise a new scoring system for combining the data across metrics. To date, no IBI's have simultaneously used data from so many different taxa. The goal is to develop a hierarchical IBI scoring system allowing a single summary score across all taxa but in addition, a series of sub-scores for each taxa. In addition, within a taxa, information about individual guilds will also be present. Thus, a user could first attain an overall score for a site then "drill down" into the IBI to begin determining what aspects of the community show evidence of disturbance relatively to intact sites. We feel this hierarchical IBI will create a robust framework for interpreting monitoring data across reserves. Furthermore, because we will have separate IBI's for each taxa, a manager can still use an IBI for a specific taxa even if they have not collected data on all taxa.

Methods

We will begin sampling vegetation, mammals, birds, herpetofauna, and arthropods in late March, 2002. We have selected 38 sites located in San Diego, Orange, and Riverside counties for the work (Table 1). These sites were selected using three criteria. First, in order to obtain herpetofauna data, the sites had to be located adjacent to sampling arrays used by Dr. Robert Fisher of the United States Geological Service. Second, we used vegetation data from Dr. Fisher to determine the % relative cover of exotic grasses at a site and chose sites to span a gradient of invasion by exotic grasses (Table 1 and Figure 3). Third, we stratified a subset of sites near urban edges. We justify the use of the vegetation data below in the "measuring disturbance" section.

Table 1. Summary of the 38 sampling sites spread across four reserve systems. The USGS site array code refers to the label used by Dr. Robert Fisher for the herpetofauna sampling sites. Each reserve has a minimum of 8 sampling sites. Relative % exotic refers to the % of exotic vegetation hits along a 100 point-intercepts at the herpetofauna array. Seven of the sites are located within 200m of an urban edge. AWC = Aliso Woods Canyon, Chino= Chino Hills State Park, SJHW = San Joaquin Hills West, Rancho Jamul = Rancho Jamul Ecological Reserve. AWC and SJHW are part of the Nature Reserves of Orange County.

Sample size within a reserve area	USGS Site-Array code	Location	Relative % Exotic
1	Awc14	Urban Edge	0
2	Awc1		7.09
3	Awc15	Urban Edge	10.87
4	Awc16		17.42
5	Awc2		23.21
6	Awc17		26.72
7	Awc13		34.17
8	Awc3		100.00
1	Chino10		0
2	Chino8		5.88
3	Chino17		11.48
4	Chino18		27.46
5	Chino6		29.84
6	Chino11		50.00
7	Chino4		66.04
8	Chino5		72.97
1	Sjhw7	Urban Edge	2.01
2	Sjhw14		5.21
3	Sjhw19		7.21
4	Sjhw18		12.36
5	Sjhw5		14.29
6	Sjhw16		17.61
7	Sjhw17		18.75
8	Sjhw11	Urban Edge	21.02
9	Sjhw6		22.88
10	Sjhw12	Urban Edge	37.01
11	Sjhw20	Urban Edge	47.25
12	Sjhw9		52.38
13	Sjhw10		58.00
14	Sjhw21	Urban Edge	64.03
15	Sjhw13		100.00
1	Rjer18		5.95
2	Rjer7		11.97
3	Rjer11		43.15
4	Rjer16		52.86
5	Rjer9		61.82
6	Rjer17		89.29
7	Rjer6		97.97
8	Rjer4		99.39

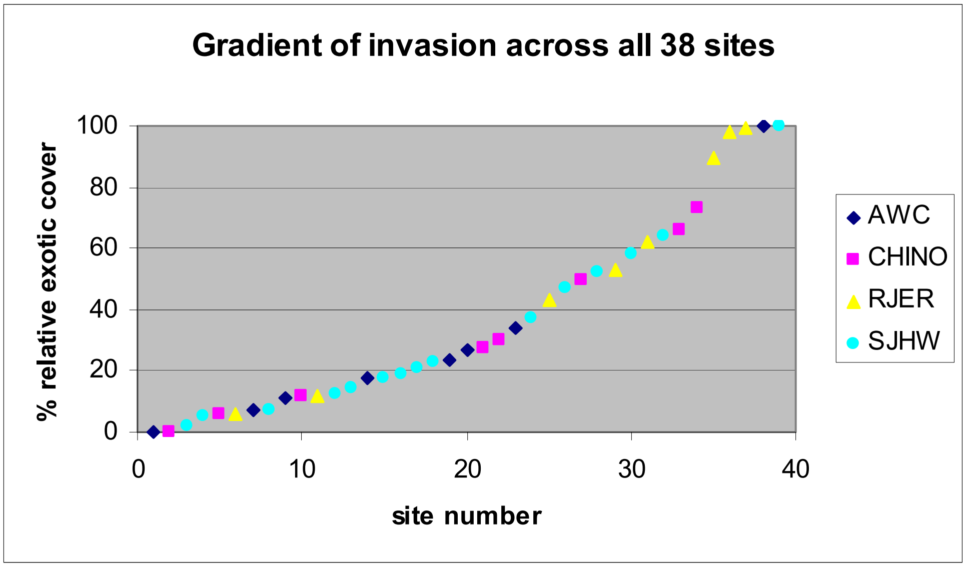


Figure 3. Gradient of invasion by exotic plants we will sample across to develop an IBI for CSS. Data are subdivided by reserve area. AWC = Aliso Woods Canyon, CHINO = Chino Hills State Park, RJER = Rancho Jamul Ecological Reserve, SJHW = San Joaquin Hills West

Grid design. At each site we have established 50 x 50 m grids (Figure 4). These grids contain 49 Sherman traps for sampling small mammals, 6 small bowls for sampling terrestrial arthropods, 4 sticky traps (surrounded by nylon mesh to keep out birds) for sampling flying arthropods, and 4-100 m vegetation transects (Figure X). In addition, we will conduct bird point counts at each site. We are currently discussing sampling methodologies with plant and ecosystem ecologists, as well as a soil microbial ecologist at San Diego State University and will likely employ a number of methods to sample ecological processes such as biomass production, seed production, and CO_2 flux, as well as below-ground microbial diversity, and mychrorrizal fungi.

Timing. We will sample each site 4 times a year. We will not sample in 4, evenly spaced sessions across the year. Instead, we have chosen sampling periods in an attempt to gather as much information on species diversity across all taxa, given a logistical constraint of 4 sampling periods. We discussed the sample timing with Dr. Ted Case (University of California, San Diego) and Dr. Gerald Braden (San Bernadino County Museum), who both have extensive experience sampling a wide variety of species and taxa in CSS. We have attempted to sample close to the periods described below while including constraints revolving around the academic calendar (final exams, spring break, etc) and Dr. Fisher's sampling schedule:

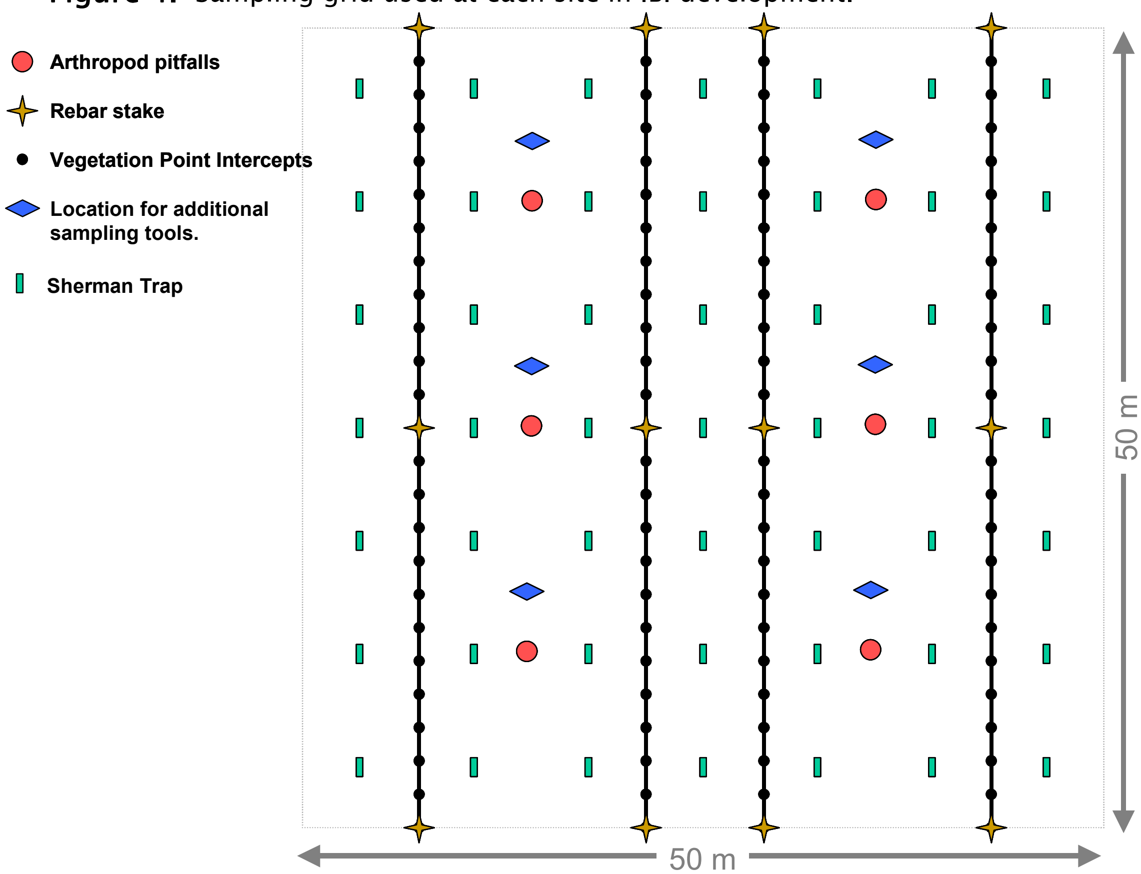
- Late-January to mid-February. Given enough rainfall, amphibians become active. Small mammal captures for some species increase, with peak abundances typically from January through May. Resident bird community is readily sampled.

- Mid-April to early May. Peak bird migration and peak herpetofauna activity. Largest insect biomass and peak in diversity.

- June. Allows insight into patterns of decline as CSS vegetation begins to dry. Final sample before the summer. Sampling during hotter summer months is not productive.

- Mid-September to early October. Return of neotropical migrant birds. A second peak in activity for some herpetofauna and young of the year appear for some species.

Figure 4. Sampling grid used at each site in IBI development.



Sampling sessions last 4 days. Given the size of the field crew, we can employ two teams, each checking 8 sites in a morning. In total, it will take 5, 4-day trapping periods to sample all of the sites (8 sites x 5 trapping periods = 40 sites). Given our schedules, we will sample all sites in a 2-3 week period, reducing differences across sites caused by temporal changes.

On day one, small mammal and arthropod traps are opened. Small mammal traps are checked each morning, closed, then reopened every afternoon, employing standard mark and recapture techniques. After three mornings of trapping, the traps are closed, and the arthropod pitfall traps emptied and turned over. Sticky traps are also collected. We have coordinated with Dr. Fisher and our sampling efforts will coincide within one week of herpetofauna sampling at any given site. Thus, the temporal overlap in our sampling efforts is within the 2-3 week period to sample all of the sites.

Vegetation sampling will occur during the first two years of the study, to verify the sampling protocol and determine levels of sampling error. Over a longer time period, vegetation will not need annual sampling. Vegetation sampling will occur only during the growing season, independent of the 4 sampling sessions described above. At each grid, vegetation foliar cover will be estimated utilizing point-intercept methods. Measurements will occur every 2 m on the 4, 50 m transects, resulting in 100 point-intercepts. In addition, shrub structure and recruitment will be assessed along each transect by sampling size, density, and status (ie. seedling, live, dead) of each shrub species in 1m^2 quadrats located every 2 m along the transect lines (100 quadrats total). Species richness data will also be collected by recording all species noted to occur within the 50 x 50m grid.

Measuring disturbance

We chose sites to span a gradient in invasion by non-native grasses. Karr argues against using biotic data to estimate disturbance levels (Karr and Chu 1999) although the practice has occurred in other terrestrial IBI's (Brooks et al. 1996, Bradford et al. 1998, Brooks et al. 1998, O'Connell et al. 2000). In general, we must be careful to avoid circularity when developing our IBI's. For example, if we developed a vegetation-based IBI and used the ratio of exotic to native cover as a metric in the IBI, we would be using the same criteria to rank the sites as we used in the IBI. This logical fault would fundamentally flaw the IBI. However, if we use data on exotic annuals to rank the sites, and metrics from the various fauna to develop the IBI, this would not cause problems. In this case, the IBI would gauge the impacts of invasion by non-native grasses on CSS fauna.

Our review of the vegetation literature leads us to conclude the levels of non-native annuals in a CSS stand are highly correlated with past or present levels of disturbance (Lozon and Macisaac 1997, Diffendorfer et al. 2002). Thus, though we may not know exactly how many head of cattle grazed our sampling site and the fire return interval, we can still gauge the level of disturbance at a site based on invasion. Indeed, we asked a number of local plant experts how they tell if a sight has been disturbed. In all cases, they invariable pointed to the level of invasion as the "tell-tale" metric of disturbance. Furthermore, invasion by non-natives does, itself, constitute a disturbance and is perhaps the most critical threat to the long-term maintenance of native biodiversity in the NCCP reserves. Thus, we focus on invasion for good reason.

However, in addition to levels of invasion, we are developing other independent measures of disturbance for each site. As O'Connell et al. (2000), we are currently developing a GIS to allow the quantification of landscape variables at each site. We will test the impacts of metrics such as distance to edge, patch size in which the site is embedded, road or trail density, reserve size, etc., on various components of the communities we sample. We have already selected 7 sites because they occur near urban edges. In addition, we are discussing possible measurements of disturbance derived from soil samples with the Dr. Lipson, the new soil microbial ecologist at San Diego State University. For example, soil density and organic matter content are considered good indicators of soil conditions (Rapport et al. 1998, Zaady et al. 2001).

Finally, between sampling periods, we will begin gathering additional information on each site from historical aerial photographs, historical records and maps of burns to further assess levels of disturbance. Our goal is to determine a reasonably accurate history of disturbance on a subset of our sites to allow us to categorize them into low, medium, and high levels of disturbance. Once in place, we can use statistical techniques to compare the rankings of the sites based on historical information to those based on GIS methods and relative cover of invasive species. In addition, we will attempt to identify new locations across S.CA with known levels of disturbance. These sites can be noted and considered for future field validation studies.

IBI development

Given our current sampling protocols, we will begin exploratory data analysis immediately following our first census. However, we will not develop a full IBI until after an entire year of sampling. Given the multiple samples through a year, we will be able to test the effects of census timing and effort on IBI development and robustness. In a perfect world, we hope to develop a cost effective IBI that can be estimated using only one sample per year. However, given the large seasonal (rainfall and temperature driven) changes in community structure in CSS, it may be necessary to sample multiple times within a year to gain an adequate sample of species for specific taxa. We envision employing all of the statistical techniques described above to both find candidate metrics as well as test the final IBI.

Ancillary products from the IBI project.

In addition to creating an IBI, the current field effort will produce a wealth of additional information about the biology of CSS, sampling techniques, and data management methods. The project is currently supporting 2 MS theses, and chapters of 2 dissertations. In addition, the project represents a significant collaboration between Dr. Diffendorfer and Dr. Fisher from USGS. Combined, we are sampling the vast majority of biodiversity in CSS and view our field methods as one of the few examples of an "in progress" test of long-term monitoring protocols for the NCCP reserves.

Biological information. To date, no single study in CSS has simultaneously studied so many taxa across gradients of disturbance. As such, our study will advance our understanding of CSS ecology on a number of fronts. First, the patterns of species occurrence and relative abundance across gradients of invasion will lead to a number of hypotheses regarding the mechanisms shaping particular taxa's responses to invasion. We anticipate some species will drop out of CSS at various thresholds of disturbance while other species will only occur at the most highly disturbed sites. Once these patterns are in place, we can begin additional field studies to elucidate the mechanisms. For example, if a suite of bird species drops out of CSS after certain levels of disturbance, we can begin experiments to determine if the cause is predator-induced, a lack of resources (i.e. food or nesting locations), or other factors.

Second, our field protocols will allow us to compare responses of taxa to the same disturbance regime and come away with an understanding of how CSS food webs react to anthropogenic impacts. Three studies have elucidated how anthropogenic effects to CSS may cascade through food webs. Crooks and Soule' (1999) produced results suggesting meso-predator release is potentially caused by habitat fragmentation in S. CA. Longcore (1999) showed a similar form of meso-predator release in the arthropod food webs of CSS. His data suggest in restored CSS sites, scorpions are rare and do not regulate ground-dwelling spiders, which in turn prey heavily on terrestrial beetles, reducing their abundance and perhaps diversity. Finally, work by Suarez and Case (2002) shows how argentine ants influence horned lizard abundance by altering native ant communities. Taken together, these studies strongly suggest trophic interactions may play critical roles in determining how biodiversity in CSS responds to disturbance.

Third, the GIS metrics will allow us to begin comparing the impacts of local variables (i.e. vegetation composition at the site) to larger-scale variables (i.e. patch size, road density) on community structure in CSS. Understanding the spatial scales at which CSS taxa respond to disturbance is critical for long-term management. For example, if our data indicate large-scale variables explain a greater amount of variation in community structure than local variables, then the overall ability of a reserve to sustain biodiversity may be influenced more by the matrix of land-use surrounding it and less by within reserve management practices. We hope this is not the case. However, it is possible we can discover particular size or shape thresholds that explain significant amounts of variation in the diversity patterns at reserves.

Methodological issues related to sampling protocols. Our sampling protocols are designed to allow us to test the efficacy of different sampling regimes for estimating various biological metrics. For example, a number of recent developments in statistical estimation of species richness lend themselves directly to NCCP monitoring protocols but have not been tested in CSS (Boulinier et al. 1998, Nichols 1998, Nichols et al. 1998). Furthermore, we can also perform tests of our ability to accurately estimate demographic rates in rodents and some commonly capture lizards given our current sampling efforts. Overall, these analyses will help establish robust, yet cost-effective methods of performing long-term monitoring in CSS.

Data collection and management protocols. We will implement data collection and management methods spearheaded by Dr. Fisher. We are currently developing data entry forms for our handheld computers and Access database tables. These tables will be integrated into the Sequel database system being developed by Dr. Fisher and a system of automatic hot-synching and data uploading from SDSU to the USGS server will be implemented. We envision this model of a central server and remote data collection and entry a successful tool for overall NCCP monitoring plans.

Literature Cited

Andelman, S. J., and W. F. Fagan. 2000. Umbrellas and flagships: Efficient conservation surrogates or expensive mistakes? Proceedings of the National Academy of Sciences of the United States of America **97**:5954-5959.
Beutel, T. S., R. J. S. Beeton, and G. S. Baxter. 1999. Building better wildlife-habitat models. Ecography **22**:219-223.
Boulinier, T., J. D. Nichols, J. R. Sauer, J. E. Hines, and K. H. Pollock. 1998. Estimating species richness: The importance of heterogeneity in species detectability. Ecology (Washington D C) **79**:1018-1028.
Bradford, D. F., S. E. Franson, A. C. Neale, D. T. Heggem, G. R. Miller, and G. E. Canterburg. 1998. Bird species assemblages as indicators of biological integrity in Great Basin rangeland. Environmental Monitoring and Assessment **49**:1-22.
Brooks, R. P., C. A. Cole, D. H. Wardrop, L. Bishel-Machung, D. J. Prosser, D. A. Campbell, and M. T. Gaudette. 1996. Wetlands, Wildlife, and watershed assessment techniques for evaluation and restoration: final report for the project: evaluating and implementing watershed approaches for protecting Pennsylvania's wetlands. Report Number 96-2, Volumes I and II. Penn State Cooperative Wetlands Center, University Park, Pennsylvania, USA.
Brooks, R. P., T. J. O'Connell, D. H. Wardrop, and L. E. Jackson. 1998. Towards a regional index of biological integrity: The example of forested riparian ecosystems. Environmental Monitoring and Assessment **51**:131-143.
Caro, T. M., and G. O'Doherty. 1999. On the use of surrogate species in conservation biology. Conservation Biology **13**:805-814.
CDF&G. 1999. CWHR version 7.0 personal computer program. *in*. California Department of Fish and Game, Sacramento, CA.

Chase, M. K., W. B. Kristan, A. J. Lynam, M. V. Price, and J. T. Rotenberry. 2000. Single species as indicators of species richness and composition in California coastal sage scrub birds and small mammals. Conservation Biology **14**:474-487.

Chikumbo, O., R. Bradbury, and S. Davey. 2001. Large-scale ecosystem management as a complex systems problem: multi-objective optimization with spatial constraints. Complexity International **8**:1-31.

Crooks, K. R., and M. E. Soule. 1999. Mesopredator release and avifaunal extinctions in a fragmented system. Nature (London) **400**:563-566.

Davis, W. S., and T. P. Simon. 1995. Introduction. Pages 3-6 *in* T. P. Simon and W. S. Davis, editors. Biological assessment and criteria: tools for water resource planning and decision making. Lewis Publishers, Boca Raton, Fla.

Dedon, M. F., S. A. Laymon, and R. H. Barrett. 1986. Evaluating models of wildlife-habitat relationships of birds in black oak and mixed-conifer habitats. Pages 115-119 *in* J. Verner, M. L. Morrison, and C. J. Ralph, editors. Wildlife 2000: Modeling Habitat Relationships of Terrestrial Vertebrates. Univ. Wisconsin Press, Madison.

Diffendorfer, J. E., G. Fleming, R. Chapman, J. Duggan, M. Mitrovitch, and M. E. Rahn. 2002. Coastal Sage Scrub response to disturbance. A literature review and annotated bibliography. California Department of Fish and Game, San Diego.

Doering, J. P. I., and M. B. Armijo. 1986. Habitat evaluation procedures as a method for assessing timber-sale impacts. Pages 407-410 *in* J. Verner, M. L. Morrison, and C. J. Ralph, editors. Wildlife 2000: Modeling Habitat Relationships of Terrestrial Vertebrates. University of Wisconsin Press, Madison.

Fausch, K. D., J. R. Karr, and P. R. Yant. 1984. Regional application of an index of biotic integrity based on stream fish communities. Transactions of the American Fisheries Society **113**:39-55.

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.

Franklin, J. F. 1993. Preserving biodiversity: Species, ecosystems, or landscapes. Ecological Applications **3**:202-205.

Franklin, J. F. 1994. Developing information essential to policy, planning, and management decision-making: The promise of GIS. Pages 18-24 *in*.

Haidinger, T. L., and J. E. Keeley. 1993. Role of high fire frequency in destruction of mixed chaparral. Madrono **40**:141-147.

Hanley, T. A. 1993. Balancing economic development, biological conservation, and human culture: The sitka black-tailed deer Odocoileus hemionus sitkensis as an ecological indicator. Biological Conservation **66**:61-67.

Hurlbert, S. H. 1997. Functional importance vs keystoneness: Reformulating some questions in theoretical biocenology. Australian Journal of Ecology **22**:369-382.

Jameson, S. C., M. V. Erdmann, J. R. Karr, and K. W. Potts. 2001. Charting a course toward diagnostic monitoring: A continuing review of coral reef attributes and a research strategy for creating coral reef indexes of biotic integrity. Bulletin of Marine Science **69**:701-744.

Karr, J. R. 1981. Assessment of biotic integrity using fish communities. Fisheries **6**:21-27.

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

Karr, J. R., and E. W. Chu. 1999. Restoring life in running waters : better biological monitoring. Island Press, Washington, D.C.

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. Illinois Natural History Survey Special Publication #5.

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. Responses of terrestrial invetebrates to human disturbance in shrub-steppe in eastern Washington. Ecological Indicators 1:63-81.

Kremen, C. 1992. Assessing the indicator properties of species assemblages for natural areas monitoring. Ecological Applications 2:203-216.

Kremen, C., R. K. Colwell, T. L. Erwin, D. D. Murphy, R. F. Noss, and M. A. Sanjayan. 1993. Terrestrial arthropod assemblages: their use in conservation planning. Conservation Biology 7:796-808.

Lambeck, R. J. 1997. Focal species: A multi-species umbrella for nature conservation. Conservation Biology 11:849-856.

Landres, P. B., J. Verner, and J. W. Thomas. 1988. Ecological Uses of Vertebrate Indicator Species A Critique. Conservation Biology 2:316-328.

Lang, C., G. L'Eplattenier, and O. Reymond. 1989. Water quality in rivers of western Switzerland: Application of an adaptable index based on benthic invertebrates. Aquatic Sciences 51:224-234.

Launer, A. E., and D. D. Murphy. 1994. Umbrella species and the conservation of habitat fragments: A case of a threatened butterfly and a vanishing grassland ecosystem. Biological Conservation 69:145-153.

Lenat, D. R. 1988. Water quality assessment of streams using a qualitative collection method for benthic macroinvertebrates. Journal of the North American Benthological Society 7:222-233.

Longcore, T. R. 1999. Terrestrial anthropods as indicators of restoration success in coastal sage scrub. Doctoral in Philosophy of Geology. University of California, Los Angeles, Los Angeles.

Lozon, J. D., and H. J. Macisaac. 1997. Biological invasions: are they dependent on disturbance? Environmental Reviews 5:131-144.

Malanson, G. P., and W. E. Westman. 1985. Postfire succession in Californian coastal sage scrub: the role of continual basal sprouting. American Midland Naturalist 113:309-318.

McBride, G. B., J. C. Loftis, and N. C. Adkins. 1993. What do significance tests really tell us about the environment? Environmental management 17:423-432.

McDonnell, M., H. P. Possingham, I. R. Ball, and E. Cousins. In Press. Mathematical methods for spatially cohesive reserve design. Environmental modelling and assessment.

McGeoch, M. A. 1998. The selection, testing and application of terrestrial insects as bioindicators. Biological Reviews (Cambridge) 73:181-201.

Mills, L. S., M. E. Soule, and D. F. Doak. 1993. The keystone-species concept in ecology and conservation. BioScience 43:219-224.

Morrison, M. L., B. G. Marcot, and R. W. Mannan. 1998. Wildlife-Habitat Relationships, Concepts and Applications, 2nd edition. The University of Wisconsin Press, Madison, Wisconsin.

Nichols, J. D., T. Boulinier, J. E. Hines, K. H. Pollock, and J. R. Sauer. 1998. Inference methods for spatial variation in species richness and community composition when not all speices are detected. Conservation Biology **12**:1390-1398.
Nichols, J. D., Thierry Boulinier, James E. Hines, Kenneth H. Pollack, and John R. Sauer. 1998. Estimating rates of local species extinction, colonization, and turnover in animal communities. Ecological Applications **8**:1213-1225.
Noss, R. F. 1990. Indicators for Monitoring Biodiversity A Hierarchical Approach. Conservation Biology **4**:355-364.
O'Connell, T. J., L. E. Jackson, and R. P. Brooks. 2000. Bird guilds as indicators of ecological condition in the central Appalachians. Ecological Applications **10**:1706-1721.
Oliver, I., and A. J. Beattie. 1993. A possible method for the rapid assessment of biodiversity. Conservation Biology **7**:562-568.
Oliver, I., and A. J. Beattie. 1996. Designing a cost effective invertebrate survey: A test of methods for rapid assessment of biodiversity. Ecological Applications **6**:594-607.
Paine, R. T. 1966. Food web complexity and species diversity. American Naturalist **100**:65-75.
Paine, R. T. 1969. The Pisaster-Tegula interaction: prey patches, predator food preference, and intertidal community structure. Ecology **50**:950-961.
Plafkin, J. L., M. T. Barbour, K. D. Porter, S. K. Gross, and R. M. Hughes. 1989. Rapid bioassessment protocols for use in streams and rivers: benthic macroinvertebrates and fish. United States Environmental Protection Agency, Office of Water, Washington, D.C.
Raphael, M. G., and B. G. Marcot. 1986. Validation of a wildlife-habitat relationships model: Vertebrates in a Douglas-fir sere. Pages 129-138 *in* L. Verner, M. L. Morrison, and C. J. Ralph, editors. Wildlife 2000: Modeling habitat relationships of terrestrial vertebrates. University of Wisconsin Press, Madison.
Rapport, D. J., C. Gaudet, J. McCullum, and M. Miller. 1998. Ecosystem health and its relationship to the health of the soil subsystem: a conceptual and management perspective. Pages 341-359 *in* P. M. Huang, editor. Soil chemistry and ecosystem health. The Society, Madison, Wis., USA.
Smith, T. S., J. T. Flinders, and D. S. Winn. 1991. A habitat evaluation procedure for rocky mountain bighorn sheep in the intermountain west. The Great Basin Naturalist **51**:205-225.
Stauffer, D. F., and L. B. Best. 1986. Effects of habitat type and sample size on habitat suitability index models. Pages 71-77 *in* J. Verner, M. L. Morrison, and C. J. Ralph, editors. Wildlife 2000: modeling habitat relationships of terrestrial vertebrates. The University of Wisconsin Press, Madison.
Suarez, A. V., and T. J. Case. 2002. Bottom-up effects on persistence of a specialist predator : Ant invasions and horned lizards. Ecological Applications **12**:291-298.
Terrell, J. W., and J. Carpenter, editors. 1997. Selecting habitat suitability index model evaluations. Information and technology report. USGS/BRD/ITR.
Tracy, C. R., P. F. Brussard, T. Esque, L. Defalco, K. Dean-Bradley, K. T. Castle, C. C. Peterson, and B. Henen. 1995. Requirements of the threatened desert tortoise: Competition with domestic cattle. Bulletin of the ecological society of america **76**:395.
US Fish and Wildlife Service. 1980. Habitat evaluation procedures (HEP). Division of Ecological Services, U.S. Fish and Wildlife Service, Dept. of the Interior, Washington, D.C.
US Fish and Wildlife Service. 1981. Standards for the development of habitat suitability index models. Division of Ecological Services, U.S. Fish and Wildlife Service, Dept. of the Interior, Washington, D.C.

Weaver, J. C. 1995. Indicator species and scale of observation. Conservation Biology **9**:939-942.

Zaady, E., R. Yonatan, M. Shachak, and A. Perevolotsky. 2001. The effects of grazing on abiotic and biotic parameters in a semiarid ecosystem: A case study from the Northern Negev Desert, Israel. Arid Land Research and Management **15**:245-261.

Zedler, P. H., C. R. Gautier, and G. S. McMaster. 1983. Vegetation change in response to extreme events: the effect of a short interval between fires in California chaparral and coastal scrub. Ecology **64**:809-818.