Adaptive Inference for Distinguishing Credible from Incredible Patterns in Nature

C. S. Holling[1] and Craig R. Allen[2]*

[1]Department of Zoology, 110 Bartram, University of Florida, Gainesville, Florida 32611, USA; and [2]US Geological Survey, Biological Resources Division, South Carolina Cooperative Fish and Wildlife Research Unit, Clemson University, Clemson, South Carolina 29634, USA

ABSTRACT

Strong inference is a powerful and rapid tool that can be used to identify and explain patterns in molecular biology, cell biology, and physiology. It is effective where causes are single and separable and where discrimination between pairwise alternative hypotheses can be determined experimentally by a simple yes or no answer. But causes in ecological systems are multiple and overlapping and are not entirely separable. Frequently, competing hypotheses cannot be distinguished by a single unambiguous test, but only by a suite of tests of different kinds, that produce a body of evidence to support one line of argument and not others. We call this process "adaptive inference". Instead of pitting each member of a pair of hypotheses against each other, adaptive inference relies on the exuberant invention of multiple, competing hypotheses, after which carefully structured comparative data are used to explore the logical consequences of each. Herein we present an example that demonstrates the attributes of adaptive inference that have developed out of a 30-year study of the resilience of ecosystems.

Key words: adaptive inference; cross-scale dynamics; ecosystem structure; hypothesis testing; lumps; resilience; strong inference; textural discontinuity hypothesis.

INTRODUCTION

This article is a review of the approach and methods we have used in a continuing effort to develop and test a novel theory for a particular class of ecological systems—that is, ecosystems, or sets of ecosystems. In this class, interactions occur across scales from centimeters and days to 10s of kilometers and centuries, so that opportunities for manipulative experiments are limited and partial and causes are multiple and nonlinear. This is the class emphasized by Pickett and others (1994) in their critique of ecological theory, method, statistics, and practice. It is a class that requires pluralistic scientific methods, several competing hypotheses, and multiple lines of evidence that converge on a particular argument.

Understanding the complexity of structure and function in ecosystems requires a sequence of cycles of investigation where each cycle starts with the general, moves to the specific, and returns to the general again. At each phase in a given cycle, there are different goals for acceptable evidence. Hence, an investigation moves from general and suggestive goals to specific and credible ones. For major studies of complex systems, such as ecosystems, the sequence of cycles takes decades and a consistent and persistent pattern of research that searches for novelty, comparability, and credibility. We argue that the goal is not to arrive at certain conclusions but to develop credible lines of argument that include some conclusions that might be certain, others that are likely, and still others that are uncertain.

This article was initially motivated by two specific critiques (Manly 1996; Siemann and Brown 1999) of one recent cycle of research on the "lumpy" character of the body masses of organisms in ecosystems (Holling 1992). But it has grown from that inception and has been generalized to deal with the fuller range of issues entailed by such investigations. The specific examples we present here concern ecosystem dynamics and structure. However, the profile is one that applies generally to studies of global climate change, to analyses of the causes and functions of biodiversity, to studies of ecosystem trophic cascade dynamics, and to designs for restoring and maintaining resilience in regional systems where people and nature interact (Gunderson and Holling 2002). Our examples are drawn from our own work in the latter area.

The paper suggesting that ecosystem attributes are distributed discontinuously (that is, they are "lumpy") (Holling 1992) arose from 30 years of research aimed at understanding the dynamics and structure of ecosystems. It triggered a strong reaction from some of those active in statistical and community ecology, because it seemed to fall within those areas of inquiry. The critical responses dealt only with the question of whether the distributions considered were lumpy, arguing that either there were at most two lumps in the distribution of animal body masses (Manly 1996) or no lumps at all (Siemann and Brown 1999) rather than the five to 10 lumps originally proposed (Holling 1992). The significant point we want to address here concerns our belief that these two critiques are based on an approach and methods appropriate for a different tradition of science—one that does not place an emphasis on the issues we dealt with and that, taken alone, cannot readily identify properties of such complex systems. As an example of complex systems, ecosystems have causes that are nonlinear and multiscalar. Interactions and self-organization occur across scales; in particular, they involve self-organization among both biotic and abiotic elements.

Our purpose here is therefore not to establish how lumpy animal body mass distributions are, but rather to stand back and summarize the methods that we, along with others, have developed for the analysis of systems with complex dynamics. It happens that these approaches and methods were the ones that led us to the conclusion that the world is lumpy and that continue to guide our ongoing research on these and related subjects. We will illustrate the features of the approach and methods by reference to research on the resilience of ecosystems, including more recent work on lumpiness and cross-scale dynamics.

We will deal principally with the challenge such studies present in terms of approach and method. They pose the following problem: How can we maintain a degree of skepticism about the existence of patterns and causes while at the same time encouraging the innovation needed to stimulate creative discovery about those patterns and their causes? Meeting this challenge is essential to studies of complexity.

APPROACH AND METHODS

It is not useful to frame the approach as a battle that pits a philosophy that insists upon avoiding type I error and assumes the primacy of falsification against one that argues for avoiding type II error and sees value in confirmatory observational evidence. The first position runs the risk of over-restricting innovation for the sake of certainty; the second, of an endless accumulation of anecdotes for the sake of speculation. At their extremes, the first erects premature stopping rules for inquiry; whereas the second presents no clear stopping rule at all.

But at their intersection is a cycle of inquiry that we call "adaptive inference". The cycle of investigation is initiated by speculation. Speculation is based on untested but novel ideas that then can be explored using simple models, empirical observations, and hard data. The ideas are then roughly tested in as quick and effective a manner as possible, using easily obtained data from any source. For example, well before its publication, the original research on lumps (Holling 1992) was submitted to just such a process. This was done as an extension of earlier resilience research (Holling 1986) that suggested that there are a small number of structuring processes that control dynamics in ecosystems, each operating at its own scale. If so, one of the consequences would be a lumpy distribution in the attributes of ecosystems. One attribute that was easy to obtain and measure was the body mass of animals living in an extensive landscape. Analysis was first meant to be quick and to provide an approximate test of a large proposition. The first check of the full distribution of body masses of a group of animals did not reveal any obvious lumpiness. But a further analysis of departures from a unimodal representation did seem to show that body weights were grouped into a small number of size classes — at least five to eight, and perhaps more.

Thereafter, we made extensive efforts to devise several competing hypotheses of causation, to ac-

cumulate evidence from several systems and situations that would help to sift among hypotheses, to develop and test specific predictive models, to form generalizations, and finally to return to speculation. Ultimately, through this well-articulated process, the expanded study (Holling 1992) reconnected with the full program of resilience studies first described in 1973 (Holling 1973).

PERCEPTION OF ERROR

Adaptive inference is, in fact, not a completely novel concept or practice for large and multiscale ecosystem science. Although no single philosophical or conceptual approach to ecology has been employed by all ecologists, many have been forced to use strategies similar to adaptive inference when focusing on ecosystemic or landscape scales (Lakatos 1978; Pickett and others 1994; Allen and Starr 1982; Carpenter forthcoming; Carpenter and Leavitt 1991; Carpenter and others 1999; Walters 1986, 1997; Levin 1999). But the essential approach is rarely discussed. Here we will attempt to give adaptive inference some formal identity.

One of the intrinsic features of adaptive inference is a shifting concern for the two types of statistical error when progressing through the examination of new data. The desire to minimize type II error encourages the scientist to embrace a wide range of possible hypotheses. Thus, speculation dominates the initial phase of each cycle of inquiry. The goal then shifts to identifying potentially interesting patterns or causes. The need to minimize type I error and winnow out false hypotheses dominates toward the end of each cycle. At this point, the goal is to either negate or build support for the existence of specific and narrower patterns or causes. The initial phase uses data visualization and liberal tests so that a great deal of possible pattern is suggested in the data. This process helps to stimulate the search for novel attributes and novel explanations that otherwise might be overlooked. Many variables and scales are considered. Later phases use sequences of more conservative tests to remove as much noise as possible while retaining and testing what may be real pattern. Overall, a set of evidence—some confirming, some weakly negating, some strongly negating—accumulates support for a particular line or lines of argument.

This strategy stands in contrast to the almost exclusive concern for type I error that dominates many biological fields of ecology. As Kaplan (1964) notes, "The scientist usually attaches a greater loss to accepting a falsehood than to failing to acknowledge a truth. As a result, there is a certain conser-

vatism or inertia in the scientific enterprise, often rationalized as the healthy skepticism characteristic of the scientific temper." But according to, Churchman (1948), the best procedure is the one that minimizes *both* types of error. A strategy that shifts the emphasis between the two types of statistical error as the investigation proceeds achieves that balance.

But does the inherent complexity of ecosystems make this more general approach ineffective? What is the requisite simplicity needed to capture the relevant complexity of that aspect of nature? We argue that the answer for ecosystems lies neither in the elegant simplicity of classical physics nor in the fascination for detail of natural history. The search for an appropriate level of simplicity is a Goldilocks process of testing to see whether the porridge is too hot, too cold, or just right. It is an adaptive, sequential process that ideally starts with several competing caricatures of hypotheses and models that are rejected, expanded, or modified only when their predictions clearly do not resonate with reality. After 3 decades of investigation, we argue that there is a minimal level of complexity that is suited to the approach described here. It is also reviewed in a recent book that synthesizes ecological, economic, and social dimensions of renewable resource systems (Gunderson and Holling 2002).

THE PROCESS

We find it useful to distinguish among propositions, hypotheses, and models. All are similar in that they represent causal explanations, but each has a different level of abstraction, and each differs in its degree of generality and testability. That is, good propositions (both in courtship and science) are ones that are extremely interesting, are untestable, but invite ways of being tested. Concern for type II error dominates. The goal of propositions is to suggest interesting hypotheses that are derived from the propositions, are testable, but are narrower. Resilience, in its first expression (Holling 1973), was based on one such proposition: that ecosystems had more than one stable state. When hypotheses are then tested, some hypotheses begin to receive more support than others, and others are discouraged, so that more specific models can be formed to test the more likely explanations. The good models are transparent, are more quantitatively specific than hypotheses, and provide a more precise but narrower test of arguments. That is where avoidance of type I error tests becomes important, and a comparative analysis of a number of specific examples is then necessary. Together, propositions, sets of hy-

potheses, and suites of models lead to generalizations. Each serves its own function in a balanced approach to discovery.

To demonstrate, consider the sequence used over the past 30 years to assess the resilience and (now) the evolvability of sets of ecosystems. It started with a synthesis of separate observations that suggested that ecosystems could have more than one stable state (Holling 1973). That notion emerged from earlier work on population processes such as predation, parasitism, reproduction, and competition. Studies of population processes provided explanations that, contrary to traditional beliefs at the time, covered a wider range of densities, particularly at low levels, and that demonstrated empirically their essential nonlinear nature (Holling 1965, 1966). It was shown that these processes could interact among themselves within larger systems to produce multiple stable states. Moreover, published evidence of the behavior of organisms at large scales in space and time seemed to reflect just such a reality.

So the proposition was that ecosystems have multiple stable states and that those properties have profound, identifiable consequences for management. Resilience was defined not as a near equilibrium response but rather as the maximum magnitude of a large perturbation that could be tolerated without causing a flip into a new state—that is, one that would just allow the system to persist with the same controls. In short, resilience allows the system to move somewhere within one of the stability regions. This view therefore shifts the focus from an equilibrium state to a stability region and its neighbors.

This early work led to a decade and a half of research devoted to demonstrating causation within several different ecosystems at the large and multiple scales needed. The question necessarily required us to leave behind the lab and quadrat for a full region where experimental approaches are partial and incomplete. Each system had key processes and structures at spatial scales ranging from centimeters to multisystem regions and time scales from months to centuries. Our own work, with other collaborators, culminated in some 23 examples of ecosystem dynamics, structure, and management published separately and then synthesized and reviewed in Holling (1986). Each of these systems was used as a test of hypotheses about causation and structure in forests, grasslands, and ocean and lake systems. Others were doing similar work in other systems at scales of lakes and watersheds (Carpenter forthcoming). After 30 years, the review of both lines of inquiry has led to the conclusions that multiple stable states are common, that Bayes-

ian and probabilistic explanations are necessary, and that sustainability requires a strategy of adaptive management of ever-changing systems (Carpenter forthcoming).

All of the systems were found to share two properties. All of them demonstrated the possibility of more than one stability region, and each of them could be explained by a handful of relationships operating at very different scales. Specifically, the essential properties of the systems could be explained by the interactions among three to four sets of variables, each of which operated at a very different scale from the others. That observation evolved into a "rule of hand" (Gunderson and Holling forthcoming)—a designation of the minimum complexity needed to understand ecosystem structure and behavior. The observation was then tested by specific hypotheses suggesting the lumping of ecosystem attributes and using quantitative models of the consequences (Allen and others 1999) of lumpy cross-scale organization in ecosystems. These are discussed in more specific form in the last section.

Hence, in one of the cycles of investigation, adaptive inference moves from propositions, to hypotheses, and thence to models, with several competing causal explanations being proposed and tested at each step. The procedure is one that moves from the most general to the most specific and back again for another cycle (Figure 1). Progress is paced by exclusive predictions and empirical tests to sift among competing hypotheses, competing expectations, and competing predictions. It is therefore based on repeated efforts to first imagine alternative explanations and then deduce and test the logical expectations of each explanation. Some of them become less likely, others more so. Wherever possible, clear and apparently unambiguous evidence is sought to negate hypotheses; but where that is not possible, suggestive but consistently confirmatory evidence is retained that at the least provides clues for the direction to follow in search of understanding. Confirmation constructs potential theory as it builds support for one or more hypotheses; negation removes inappropriate theory by dismissing the hypotheses that are distinctly unable to describe reality (Pickett and others 1994). The two processes of confirmation and disproof, acting together, take each raw, naïve proposition through maturation toward theory.

The results at any stage can begin to explain a diversity of separate issues. For example, this process provided a way to assess, in eastern North American forests, the role of bird predation in producing surprising flips in conditions of insect out-
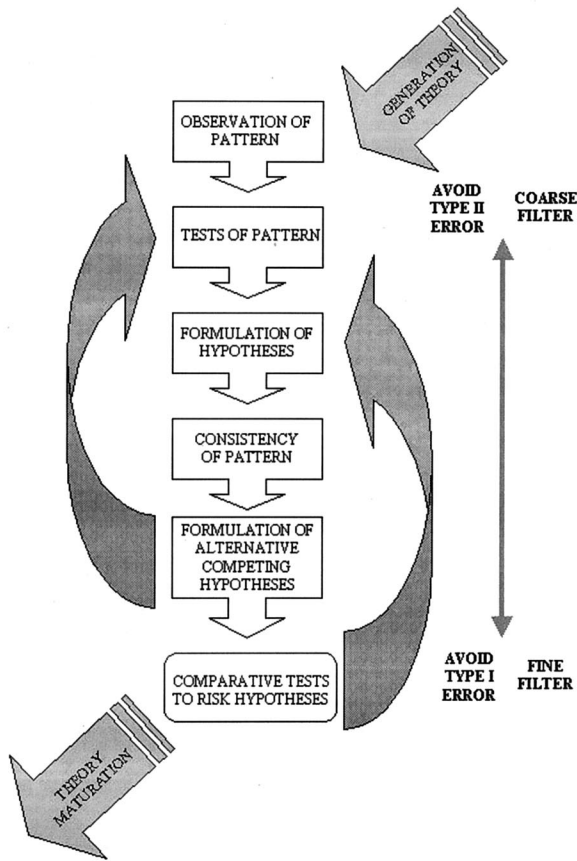
Figure 1. The adaptive cycle of adaptive inference. An intrinsic feature of adaptive inference is a shifting concern for the two types of statistical error as progress is made through the examination of new data. Avoidance of type II error dominates the initial phase of each cycle of inquiry; avoidance of type I error dominates toward the end of each cycle.

break in spruce and balsam tree stands. It provided an approach and models to assess the degree to which migratory birds (predators of an insect defoliator) would have to decline to cause such surprises. The crucial question was how great a change in the populations of migratory birds the system could tolerate before it changed from a state controlled by trees, on a time dynamic of 45–60 years, to a state controlled by foliage, on a dynamic of 12–15 years (Holling 1988). We concluded that bird populations would have to decline by about 65%, and there is in fact some evidence that this may be happening due to habitat impacts from Canada to South America (Gauthreaux 1992). In the course of this study, we established, for the first time, that the body masses of the 35 species of insectivorous birds were distributed among three size groups, thus making it possible to use empirical data from nature

to establish ranges of scales of effective budworm control by the 35 species of insectivorous birds. More importantly, this study offered a hint about lumpiness that was later explored extensively when theory suggested that it could be a more universal property (Holling 1992).

Adaptive inference involves the accumulation of multiple lines of mutually reinforcing evidence to build a line of credible argument. Falsification, confirmation, deduction, and induction are all employed to evaluate the empirical evidence, which is obtained by all manner of observations and tests for the purpose of sorting among competing hypotheses. This suite of methods provides the redundancy and cross-checks that Pickett and others (1994) prescribe for approaches to achieve ecological understanding and improve environmental policy and practice. It depends on cycles of investigation, each cycle starting with propositions and ending with tested hypotheses and models. But the result is a never-ending sequence in which credible arguments and explanations are accompanied by less credible ones. We identify likely, credible explanations and practical needs for management, assessment, and response, not many certainties and unambiguities.

Adaptive Inference and Strong Inference

Strong inference, as described by Platt (1964), shares some important features with adaptive inference. Both methods pose and test branch points on a tree of logically alternative sets of hypotheses. As the process proceeds, causal explanations are progressively refined, deepened, and generalized. But strong inference relies on situations where causes can be single and separable and where discrimination between pairwise alternative hypotheses can be determined experimentally by a simple yes or no answer. As Platt (1964) demonstrates, strong inference is a powerful and rapid way to deal with questions in molecular biology, cell biology, and physiology. But because ecosystems are more open than organisms, causes in ecological systems are multiple and overlapping and are not entirely separable (Hilborn and Stearns 1982; Pickett and others 1994). Frequently, competing hypotheses cannot be distinguished by a single unambiguous test or set of controlled experiments, but only by a suite of tests that accumulates a body of evidence that supports one line of argument and not others. Instead of pitting each member of a pair of hypotheses against each other, adaptive inference relies on the exuberant invention of multiple, competing hypotheses followed by the assessment of carefully

structured comparative data to explore the logical consequences of each.

We further demonstrate the process of adaptive inference in the following section by providing a concrete example, based upon the proposition that terrestrial ecosystems are controlled, or self-organized, by a small nested set of critical processes, each operating over a particular scale range. The world-is-lumpy proposition is centered on a hypothesis that encompasses patterns not only across ecosystems, but across scales as well. And this is the class where standard statistical methodologies fail. The type of system-level and cross-scale inquiry we are undertaking disallows much manipulation, replication, and—often—simple hypothesis testing.

One View of the Structure of Ecosystems

Growing evidence from nature, models, and theory suggests that ecosystem structure and dynamics are dominated by the influence of a small set of plant, animal, and abiotic processes (Holling 1986, 1992; Carpenter and Leavitt 1991; Iwasa and others 1991; Levin 1992). Each set of processes operates at characteristic periodicities and spatial scales (Allen and Starr 1982; O'Neill and others 1986; Holling 1992). Small and fast scales are dominated by biophysical processes that control plant morphology and function. At a larger and slower scale of patch dynamics, interspecific plant competition for nutrients, light, and water, interacting with climate, affects the composition and regeneration of local species. At a still larger scale of stands in a forest, mesoscale processes of fire, storm, pathogen outbreak, and large mammal herbivory determine structure and successional dynamics from 10s of meters to kilometers and years to decades. The largest landscape scales have geomorphological and evolutionary processes that affect structure and dynamics over 100s of kilometers and millennia.

Holling (1986, 1992) derived a version of the world-is-lumpy proposition by identifying the common properties found in the suite of 23 studies of the dynamics of regional scale ecosystems mentioned earlier. They included forest systems affected by insect outbreak and fire, semi-arid systems affected by grazing and fire, and marine systems structured by trophic relationships. In every instance, the essential behavior in space and time could be captured by three sets of interacting variables whose generation times or speeds varied by an order of magnitude from each other.

To use an example from one of the 23 ecosystems studied, the dynamics of the birch/balsam fir/spruce forest of east central North America can be ascribed to a fast set of variables involving the defoliating insect spruce budworm, its natural enemies, and spruce needles (a generation time of 1 year); an intermediate set involving the crown of foliage and energy storage (a generation time of 8–12 years); and a slow set involving the trees and their immediate competitors (a generation time of 70–100 years). The consequence of the interactions among these variables can be explored in full space/time simulation models built from detailed knowledge of processes, as described in Clark and others (1979). It was subsequently shown that the essential properties in this example and several others could also be represented with three coupled differential equations, each one reflecting the dynamics of one of the three sets of variables (see, for example, Walker and others 1969; Ludwig and others 1978; McNamee and Holling 1981).

The significant point is not that there were exactly three sets of critical variables. Presumably, the three identified were drawn from a larger hierarchical set that includes variables involved in smaller and faster processes, as well as larger and slower ones. Rather, the key point is that their temporal and geometric properties were distributed in a lumpy, or discontinuous, manner. That is, the frequency of occurrence of attributes in time and space were clustered into a small number of categories along an axis of increasing magnitude of scale in space and time.

The result therefore implies a nested set of cycles, each of a distinct scale in time and space and each controlled by one of the sets of controlling variables. Each set is characterized not only by its unique speed but also by its unique spatial scale. Fast variables are small; slow variables are large. This, therefore, is a hierarchical representation of a nested set of variables (Allen and Starr 1982; O'Neill and others 1986) where each set is controlled by processes sufficiently different in speed and size to introduce discontinuities or lumpiness in the distribution of ecosystem attributes. Each set controls or self-organizes a persistent pattern over a particular range of time and space scales.

Because the resulting ecosystem patterns at different scales on landscapes are persistent and repetitive between successional stages, they should have the tendency to entrain attributes of other variables that are not directly involved in producing this self-organized structure. And that realization identifies a new cycle of inquiry and a new proposition. If those landscape patterns are persistent enough, then biological processes unrelated to the original structuring processes will become entrained by and adapted to the pattern. They could well amplify the originating pattern to create an enhanced signature

of ecosystem and landscape structure. Hence, the life history and behavioral and morphological attributes of plants and animals could all become adapted to the lumpy landscape pattern and amplify its signal.

To test this hypothesis, two features are then sought—indications of lumpiness beyond that expected to be generated from random processes and evidence that such lumpiness is common. For example, time series data should show a small number of significant frequencies around which different data sets cluster. Spatial vegetation data should show a distribution of object sizes, distances, and fractal dimensions that all cluster around a few significant categories.

Time series analysis of charcoal in cores from stable lake sediments in the transition boreal forest zone in Minnesota shows four dominant periodicities (3–5 years, 10–15 years, 35–40 years, and over 80 years) (Clark 1988) that are the same periodicities predicted by models of the boreal forest (Clark and others 1979; Holling 1988). Analysis of remote imagery similarly shows spatial structures whose geometric attributes cluster around a few values across scales (O'Neill and others 1991). In addition, a variety of attributes of the animal communities living in a particular landscape (for example, body masses, home ranges, or geographic ranges) should cluster around a small number of size categories. The world-is-lumpy proposition proposes that body mass distributions of animals are discontinuous because they become entrained and reflect the discontinuous landscape structure that results from the signatures, in terms of structural patterns, of key processes. If this proposition is true, we should be able to detect discontinuous patterns in animal body mass distributions.

Detecting lumpy, or discontinuous, patterns in body mass distributions is far from elementary. Statistics for this type of pattern detection simply do not exist. So, early in the cycle of inquiry, the bias is to detect pattern, if it exists at all. That is, a premium is placed on avoiding type II error because clues to guide intuition are critically important early in the process. At this stage, the cost of discarding propositions that potentially reap great rewards is high, whereas the cost of pursuing new ideas that may ultimately fail is relatively low.

In comparison tests of the bird and mammal fauna of boreal forests and prairies, Holling (1992) originally used visualization to conclude that the lumpy proposition was the only one that seemed to be consistent with the evidence. Subsequent tests using different organisms and wet and dry systems led to the weakening of some explanatory hypoth-

eses (phenologogical, founding effects, locomotory, trophic) and the strengthening of others (entrainment by landscape patterns). Multiple causes might have a role, but the initiation seems to involve patterns established by the small number of critical processes, each operating at its own scale.

Visualization may be as valid as probabilistic inference (Cleveland 1993), and graphic display may well reveal significant relationships and patterns as well as formal statistics and provide deep insight into the structure of data (Chambers and others 1983). However, it is a weakness of the original lump paper (Holling 1992) that the visualization techniques used to identify gaps and lumps appear to be subjective, even when independent observers come to the same conclusions. In formal statistical tests, although the choice of a level of significance is equally subjective, it is a familiar type of subjectivity and therefore more comfortable to accept. For this reason, after the first visual discovery of interesting pattern, it is useful to add more formal statistical tests to the visualization procedures to develop the multiple methods useful in evaluating similarity and differences in pattern. However, there are currently few methods available that have the generality and levels of resolution appropriate to the question.

We agree with Wiens (1984) that confirmation of a predicted pattern may lend little credence to the arguments that generate the prediction. Similar patterns can easily be generated by a variety of causes. Therefore, alternative, competing hypotheses to explain lumpiness are needed, followed by efforts to negate them wherever possible. Where falsification of any of a few remaining hypotheses eludes us, relative degrees of corroboration are obtained through tests designed to sort among them. But before that step, some idea is needed whether it is worthwhile to proceed. That is where the first decision for relevant statistical tests is made.

Holling followed visualization tests by using an index of body mass difference (Holling 1992) to confirm what the eye readily perceives. An example is shown in Figure 2. Holling (1992) defined significant gaps using a body mass difference index (a type of split moving window) with the placement of a criterion line at one standard error above the mean and found that the data were lumpy.

In an attempt to test whether the lumpy pattern generated was exceptional compared to what one would expect from samples from a unimodal distribution, Holling (1992) simulated data sets from a fitted log normal, unimodal, null distribution. However, Manly (1996) clearly shows that unimodal kernel density estimates yield null distributions that
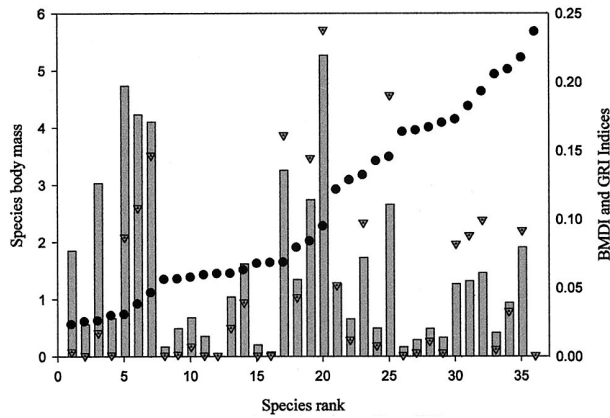
Figure 2. Are body mass patterns discontinuous? The body mass distribution of the mammals of the (Canadian) boreal forest (Holling 1992) in this graphic shows species rank versus body mass (left y-axis, circles) and split moving window (Mn/Mn+1; vertical bars, right y-axis) and gap rarity indexes (triangles, range 0 to 1, axis not shown). Where the body masses of adjacent animals are very similar, the line represented by the circles is nearly flat. Large differences in body masses are indicated by jumps. Similarly, high values for both indexes indicate "gaps," or unusually large size differences between adjacent species. More familiar statistical procedures, cluster analysis, and CART analysis provide the same results qualitatively. The divergent methodologies converge on the same interpretation of lumpy structure, and all of them confirm what the eye readily tells us.

are notably superior and more general. Manly (1996) went on to apply a conservative test of lumpiness using Silverman's "bump" hunting method (Silverman 1981, 1986). This process requires a test of each data set independently to determine the number of significant modes that are needed to describe each distribution. The test is designed to minimize the chance of accepting a false lump or mode (type I error). As a consequence, Manly found weak evidence for no more than two lumps and in fact identified only one mode for our example data (Figure 2). In contrast, Holling (1992) identified up to eight lumps in the same data sets, including eight identified for the data displayed in Figure 2. As Manly (1996) stated, however, using Silverman's method, "The probability of estimating the correct number of modes may be quite small."

And that is the point. The cycle of inquiry in adaptive inference is a process of screening possibilities—initially with a coarse screen, and then with progressively finer ones. Hence it is essential not to reject possibilities too soon by applying fine screens prematurely. Several years after the first substantial challenge to the world-is-lumpy propo-

sition (Manly 1996), we note with some satisfaction that a recent paper coauthored by Manly concluded: "This conservation gives considerable weight to Holling's (1992) proposition that body size distributions in nature may indeed reflect abrupt discontinuities in habitat texture" (Raffaelli and others 2000). Yet if we had used conservative tests early in our investigations, we would have concluded that the world is continuous.

Therefore, at early stages in the process of adaptive inference, it is better to obtain an approximate answer to the right question (for example, Is there significant lumpiness?) rather than a precise answer to the wrong question (How many significant lumps are there?). The issue at this stage is not how many lumps there are, nor where they are, but whether there is unexpected pattern that occurs often enough to deserve more rigorous subsequent exploration. A conservative test at this stage is inappropriate because it is likely to prematurely foreclose opportunities for discovery.

The qualitative pattern generated in these animal body mass data sets were not particular to the body mass difference index (BMDI), since they also result from application of the gap rarity index (Restrepo and others 1997), split moving window indexes (Cornelius and Reynolds 1991), classification and regression tree analyses (Clark and Pregibon 1992), or hierarchical clustering analysis (SAS Institute 1989; Allen and others 1999). In fact, for data in Figure 2, all of these methods indicate multiple aggregations, ranging from eight using Holling's BMDI (Holling 1992) to as few as five using cluster analysis (cluster analysis actually supports two interpretations—five lumps or eight lumps). However, only the gap rarity index was specifically designed to detect discontinuities in body mass distributions. Although all methods do seem to generally provide the same results, each particular methodology has its inherent biases.

Sendzimir (personal communication) has compiled data sets for mammals and birds from 75 ecosystem sites located in forests, savannas, and prairies from boreal, temperate, and tropical regions of the Western Hemisphere. He used kernel density estimates, as described by Silverman (1981, 1986) and Manly (1996), to obtain a unimodal null distribution for each data set. From the unimodal distribution were drawn 1000 simulated data sets to which the body mass difference index criterion was applied to define lumps and gaps. The frequency distribution of lumps for the real data set differed from, and was much greater than, that generated in the simulated data sets. We interpret those results as a strong indication that lumpiness in body mass

distributions in nature is overwhelmingly the rule, not the exception.

The landscape pattern hypothesis has been supported while other competing hypotheses have been inconsistent with the evidence, although they could describe causes that reinforce a pattern originated by landscape pattern. Assuming for the moment that those results are solid, a step of still finer screening of possibilities can be pursued by designing specific models of causation and testing them with independent data sets.

Each competing hypothesis generates different predictions for the degree of match and mismatch of patterns in these comparisons. They therefore provide a way to sift among the competing hypotheses in a process that can gradually build a case for one cause or a specific combination of causes. At that point, a specific model can be constructed whose parameter values define specific relationships. Modeling then leads to a third phase, where each model is tested using data sets chosen to be as independent as possible from the ones that defined the model.

Adaptive inference balances potential options against potential costs, in an iterative manner that first emphasizes maintaining future options by not falsely supporting null hypotheses, and then focuses upon shifting among competing hypotheses. In the example used, the textural discontinuity hypothesis, we are still in the fairly early stages of inference. However, there is a growing body of evidence that suggests the validity of the first proposition and continues to generate further hypotheses and some quantitative models. Some of this evidence supports the existence of lumpiness in numerous systems (see, for example, Restrepo and others 1997; Allen and others 1999; Raffaelli and others 2000; Havlicek and Carpenter 2001); it includes the discovery that independent attributes of species are associated with lumpy body mass patterns (Allen and others 1999; Allen and Saunders 2002).

Lumpiness in ecosystems is merely the specific issue that occupies us at present and that we have therefore chosen to use as an example of the process of adaptive inference. Although it is not yet formalized, adaptive inference has been the approach and guide throughout many journeys of ecological discovery. Other ecological examples that have, in our opinion, by necessity employed adaptive inference to uncover ecosystem properties include the description of trophic cascades and ecological resilience. In cases where the level of inquiry is the system, replicated field studies are not possible and the options for understanding are severely limited. It is not possible to identify single scales of investigation when the level of interest includes interactions both within and across scales. Without the adaptive inference approach, investigation into the lumpy nature of ecosystems would have been halted long ago, rather than continuing as a healthy, albeit controversial, research effort.

ACKNOWLEDGMENTS

REFERENCES

Allen CR, Forys EA, Holling CS. 1999. Body mass patterns predict invasions and extinctions in transforming landscapes. Ecosystems 2:114–121.

Allen CR, Saunders DA. 2002. Variability between scales: Predictors of nomadism in birds of an Australian Mediterranean-climate ecosystem. Ecosystems 5:348–59.

Allen TFH, Starr TB. 1982. Hierarchy: perspectives for ecological complexity. Chicago: University of Chicago Press.

Carpenter SR. Alternate states of ecosystems: evidence and its implications for environmental decisions. In: Huntley N, Levin S, editors. Ecology: achievement and challenge. Oxford (England): Blackwell. Forthcoming.

Carpenter SR, Leavitt PR. 1991. Temporal variation in a paleolimnological record arising from a tropic cascade. Ecology 72:277–285.

Carpenter SR, Ludwig D, Brock WA. 1999. Management of eutrophication for lakes subject to potentially irreversible change. Ecol Appl 9:751–771.

Chambers JM, Cleveland WS, Kleiner B, Tukey PA. 1983. Graphical methods for data analysis. Boston: Duxbury Press.

Churchman CW. 1948. Theory of experimental inference. New York: Macmillan.

Clark JS. 1988. Effect of climate change on fire regimes in northwestern Minnesota. Nature 334:233–235.

Clark LA, Pregibon D. 1992. Three-based models. In: Chambers JM, Hastie TJ, editors. Statistical models in S. Pacific Grove (CA): Wadsworth & Brooks/Cole. p 377–419.

Clark WC, Jones DD, Holling CS. 1979. Lessons for ecological policy design: a case study of ecosystem management. Ecol Model 7:1–53.

Cleveland WS. 1993. Visualizing data. Summit (NJ): Hobart Press.

Cornelius JM, Reynolds JF. 1991. On determining the statistical significance of discontinuities within ordered ecological data. Ecology 72:2057–2070.

Gauthreaux SA Jr. 1992. The use of weather radar to monitor long-term patterns of trans-Gulf migration in spring. In: Hagan JM, Johnston DW, editors. Ecology and conservation of neotropical migrant landbirds. Washington (DC): Smithsonian Institution Press. p 96–100.

Gunderson L, Holling CS, editors. 2002. Panarchy: understand-

ing transformations in human and natural systems. Washington (DC): Island Press.

Havlicek T, Carpenter SR. 2001. Pelagic size distributions in lakes: are they discontinuous? Limnol Oceanogr. 46:1021–1033.

Hilborn R, Stearns S. 1982. On inference in ecology and evolutionary biology: the problem of multiple causes. Acta Biotheoret 32:145–164.

Holling CS. 1992. Cross-scale morphology, geometry and dynamics of ecosystems. Ecol Monogr 62:447–502.

Holling CS. 1966. The functional response of invertebrate predators to prey density. Mem Entomol Soc Can 48:1–86.

Holling CS. 1965. The functional response of predators to prey density and its role in mimicry and population regulation. Mem Entomol Soc Can 45:1–60.

Holling CS. 1986. Resilience of ecosystems; local surprise and global change. In: Clark WC, Munn RE, editors. Sustainable development of the biosphere. Cambridge (England): Cambridge University Press. p 292–317.

Holling CS. 1973. Resilience and stability of ecological systems. Annu Rev Ecol Syst 4:1–23.

Holling CS. 1988. Temperate forest insect outbreaks, tropical deforestation and migratory birds. Mem Entomol Soc Can 146:21–32.

Iwasa Y, Sato K, Nakashima S. 1991. Dynamic modeling of wave regeneration (Shimagare) in subalpine Abies forests. J Theoret Biol 152:143–158.

Kaplan A. 1964. The conduct of inquiry. San Francisco: Chandler.

Lakatos I. 1978. The methodology of scientific research programmes. Cambridge (England): Cambridge University Press.

Levin S. 1999. Fragile dominion. Reading (MA): Perseus.

Levin SA. 1992. The problem of pattern and scale in ecology. Ecology 73:1943–1967.

Ludwig D, Jones DD, Holling CS. 1978. Qualitative analysis of insect outbreak systems: the spruce budworm and forest. Animal Ecol 44:315–332.

McNamee PJ, Holling CS. 1981. The structure and behavior of defoliating insect/forest systems. Res Pop Ecol 23:280–298.

Manly BF. 1996. Are there clumps in body-size distributions. Ecology 77:81–86.

O'Neill RV, DeAngelis DL, Waide JB, Allen TFH. 1986. A hierarchical concept of ecosystems. Princeton (NJ): Princeton University Press.

O'Neill RV, Turner SJ, Cullinam VI, Coffin DP, Cook T, Conley W, Brunt J, Thomas JM, Conley MR, Gosz J. 1991. Multiple landscape scales: an intersite comparison. Landscape Ecol 5:137–144.

Pickett, STA, Jones CG, Kolasa J. 1994. Ecological understanding: the nature of theory and the theory of nature. San Diego: Academic Press.

Platt JR. 1964. Strong inference. Science 146:347–353.

Raffaelli D, Hall S, Emes C, Manly BF. 2000. Constraints on body size distributions: an experimental approach using a small-scale system. Oecologia 122:389–398.

Restrepo C, Renjifo LM, Marples P. 1997. Frugivorous birds in fragmente neotropical montane forests: landscape pattern and body mass distribution. In: Laurance WF, Bierregaard RO, Moritz C, editors. Tropical forest remnants: ecology, management and conservation of fragmented communities. Chicago: University of Chicago Press. p 171–189.

SAS Institute. 1989. SAS/STAT user's guide. Version 6, 4th ed. Cary (NC): SAS Institute.

Siemann E, Brown JH. 1999. Gaps in mammalian body size distributions reexamined. Ecology 80:2788–2792.

Silverman BW. 1986. Density estimation for statistics and data analysis. London: Chapman & Hall.

Silverman BW. 1981. Using kernel density estimates to investigate multimodality. J R Stat Soc 43:97–99.

Walker BH, Ludwig D, Holling CS, Peterman RM. 1969. Stability of semi-arid savanna grazing systems. Ecology 69:473–498.

Walters CJ. 1986. Adaptive management of renewable resources. New York: Macmillan.

Walters CJ. 1997. Challenges in adaptive management of riparian and coastal ecosystems. Conserv Ecol: onlineURL: htp://www.consecol.org/voll/iss2/art1.

Wiens JA. 1984. On understanding a non-equilibrium world: myth and reality in community patterns and processes. In: Strong DR Jr, Simberloff D, Abele L. Thistle AR, editors. Ecological communities: conceptual issues and the evidence. Princeton (NJ): Princeton University Press. p 439–457.