

- Holling, C. S. 1988. Temperate forest insect outbreaks, tropical deforestation and migratory birds. *Memoirs of the Entomological Society of Canada* 146:21-32.
- Holling, C. S. 1992. Cross-scale morphology, geometry and dynamics of ecosystems. *Ecological Monographs* 62(4):447-502.
- Holling, C. S., D. D. Jones, and W. C. Clark. 1977. Ecological policy design: A case study of forest and pest management. IIASA CP-77-6:13-90 in *Proceedings of a Conference on Pest Management*, October 1976, G. A. Norton and C. S. Holling, eds. Laxenburg, Austria.
- Holling, C. S., D. W. Schindler, B. Walker, and J. Roughgarden. 1995. Biodiversity in the functioning of ecosystems: An ecological primer and synthesis. In *Biodiversity Loss: Ecological and Economic Issues*, C. Perrings, K. G. Måler, C. Folke, C. S. Holling, and B. O. Jansson, eds. Cambridge, England: Cambridge University Press.
- Morris, R. F. 1963. The dynamics of epidemic spruce budworm populations. *Memoirs of the Entomological Society of Canada* 21:332.
- O'Neill, R. V., D. L. DeAngelis, J. B. Waide, and T. F. H. Allen. 1986. *A Hierarchical Concept of Ecosystems*. Princeton, N.J.: Princeton University Press.
- Pimm, S. L. 1984. The complexity and stability of ecosystems. *Nature* 307:321-326.
- Schindler, D. W. 1990. Experimental perturbations of whole lakes as tests of hypotheses concerning ecosystem structure and function. *Proceedings of 1987 Crafoord Symposium*. *Oikos* 57:25-41.
- Schindler, D. W. 1993. Linking species and communities to ecosystem management. *Proceedings of the 5th Cary Conference*, May 1993.
- Schindler, D. W., T. M. Frost, K. H. Mills, P. S. S. Chang, I. J. Davis, F. L. Findlay, D. F. Malley, J. A. Shearer, M. A. Turner, P. J. Garrison, C. J. Watras, K. Webster, J. M. Gunn, P. L. Brezonik, and W. A. Swenson. 1991. Freshwater acidification, reversibility and recovery: Comparisons of experimental and atmospherically-acidified lakes. Volume 97B: 193-226 in *Acidic Deposition: Its Nature and Impacts*, F. T. Last and R. Watling, eds. *Proceedings of the Royal Society of Edinburgh*.
- Sinclair, A. R. E., P. D. Olsen, and T. D. Redhead. 1990. Can predators regulate small mammal populations? Evidence from house mouse outbreaks in Australia. *Oikos* 59:382-392.
- Tilman, D., and J. A. Downing. 1994. Biodiversity and stability in grasslands. *Nature* 367:363-365.
- Waide, J. B., and J. R. Webster. 1976. Engineering systems analysis: Applicability to ecosystems. Volume IV, pp. 329-371 in *Systems Analysis and Simulation in Ecology*, B. C. Patten, ed. New York: Academic Press.
- Walker, B. H. 1981. Is succession a viable concept in African savanna ecosystems? Pp. 431-447 in *Forest Succession: Concepts and Application*, D. C. West, H. H. Shugart, and D. B. Botkin, eds. New York: Springer-Verlag.
- Walker, B. H., D. Ludwig, C. S. Holling, and R. M. Peterman. 1969. Stability of semi-arid savanna grazing systems. *Ecology* 69:473-498.
- Walters, C. J. 1986. *Adaptive Management of Renewable Resources*. New York: McGraw Hill.

## ARTICLE 3

## The Resilience of Terrestrial Ecosystems

*Local Surprise and Global Change*

C. S. HOLLING

Pages 292-320 in *Sustainable development of the biosphere*:  
By W. C. Clark and R. E. Munn (eds). Cambridge University Press,  
Cambridge, UK.

Source: Clark WC, Munn RE (eds) (1986) *Sustainable Development of the Biosphere*. Cambridge University Press, Cambridge, UK.

ADEQUATE EXPLANATIONS of long-term global changes in the biosphere often require an understanding of how ecological systems function and of how they respond to human activities at local levels.

Outlined in this chapter is one possible approach to the essential task of linking physical, biological, and social phenomena across a wide range of spatial and temporal scales. It focuses on the dynamics of ecological systems, including processes responsible for both increasing organization and for occasional disruption. Special attention is given to the prevalence of discontinuous change in ecological systems, and to its origins in specific nonlinear processes interacting on multiple time and space scales. This ecological scale of analysis is linked "upward" to the global scale of biogeochemical relationships and the "Gaia" hypothesis (see Chapters 7,

8, and 9 in *Clark and Munn* volume), and "downward" to the local scale of human activities and institutions (see Chapters 3, 11, and 14 in *Clark and Munn* volume).

### Introduction

Considerable understanding has been accumulated during the last decade of the way the world "ticks" in its various parts, and progress has been made in recognizing what those parts are and the need to inter-relate them. Only after such developments, and therefore only recently, could we begin to address global ecological questions effectively. There are four key questions. How do the Earth's land, sea, and atmosphere interact through biological, chemical, and physical processes? How do ecosystems function and behave to absorb, buffer, or generate change? How does the development of man's economic activities, particularly in industry and agriculture, perturb the global system? How do people—as individuals, institutions, and societies—adapt to change at different scales? In this chapter I respond to the second question by exploring the way ecosystems function and behave, but with the other three problems in mind.

During much of this century global change has been slow, although some important cumulative effects have occurred. The gradual expansion, on a global scale, of economic and agricultural development is well represented by regular increases in atmospheric carbon dioxide ( $\text{CO}_2$ ) of approximately one part per million per year [1]. This increase is attributable to a 2–4% annual increase in burning fossil fuels and, in part, to deforestation. The climate during this century has been benign relative to other periods. Marine fisheries stocks, although typically variable, were largely steady during the period 1920–1970, at least in relation to the apparent sharp shifts that occur among such species as Pacific sardine or Atlantic herring every 50 to 100 years [2]. Problems of environmental pollution have increased in geographical scale from the highly local to the size of air basins or watersheds, but slowly enough that the effects have been largely ameliorated [3]. The deterioration of Lake Baikal in the USSR has been slowed and that of Lake Erie in North America has been stopped. Fish have returned to the River Thames and the extreme smogs of London are now only memories. Atmosphere, oceans, and land, coupled through biological, chemical, and physical forces have apparently been able to absorb the global changes of this century.

But now qualitative change [4], as distinct from gradual quantitative change, seems possible. Man's industrial and agricultural activities have speeded up many terrestrial and atmospheric processes, expanded them globally, and homogenized them. Four qualitative changes are suggested. First, such changes are being considered as ecological, not simply environmental. For example, pollution can no longer be viewed as inertly burdening the atmosphere. Rather, its impacts on vegetation can accelerate the consequences by impairing the regulatory processes that are mediated by vegetation. Second, the intensity of the impact of man's activities and their acceleration of the time dynamics of natural processes can influence the coupling of long-term regulatory phenomena that link atmosphere, oceans, and land. Third, some of these qualitative changes are likely to be irreversible in principle. So long as the change is local, it can be reversed because there are alternative sources both of genetic variability and species and for the renewal of air and water. But this becomes less and less an option as the change becomes more homogeneous with increasing scale, from local to continental to hemispheric, and then to global. Finally, with the option of reversibility reduced, increasing emphasis will be placed on adapting to the inevitable. But individual, institutional, and social adaptation each have their own time dynamics and histories. There has been little experience in translating the remarkable adaptive responses of individuals to local changes [5] into responses to international and global ones.

In order to analyze ecosystem function and behavior in such a way that global changes can be related to local events and action, I consider four topics. The first is a conceptual framework that can help focus treatment of the contrasts between global and local behavior and between continuous and discontinuous behavior. Since the framework describes different perceptions of regulation and stability, it provides the necessary background for the second topic: the particular causal relations and processes within ecosystems, the influence of external variation on them, and their behavior in time and space. The third topic synthesizes our present understanding of the structure and behavior of ecosystems in a way that has considerable generality, and organizational power. The fourth connects that understanding to our knowledge of global phenomena and of local perception and action.

### The Conceptual Framework: Gaia and Surprise

In this chapter I discuss ecosystems, but first the relationships between ecosystems and two other key aspects of the global puzzle must be established; namely, with global biogeophysical events, and with societal perception and management. In the former case, some image is needed of the way the global systems in the atmosphere, oceans, and land interact. That image is provided by the Gaia hypothesis [6]. And to relate our understanding of the behavior of local ecosystems to the way societies perceive and manage those systems the concept of surprise is needed: Gaia and surprise are dealt with in turn.

#### Gaia

Gaia is the "global biochemical homeostasis" hypothesis, proposed by Lovelock and Margulis [6,7], that life on Earth controls atmospheric conditions optimal for the contemporary biosphere. The Gaia hypothesis presumes homeostatic regulation at a global level. An example is the maintenance of 21% oxygen in the air, a composition representing the highest possible level to maximize aerobic metabolism, but just short of the level that would make Earth's vegetation inflammable. The residence time of atmospheric oxygen is of the order of thousands of years, a time scale that renders methane ( $\text{CH}_4$ ) production by anaerobic organisms an important regulator of oxygen concentration. The mechanism proposed includes the burying of a small amount of the carbonaceous material of living matter each year and the production of  $\text{CH}_4$ , which reacts with oxygen, thereby providing a negative feedback loop in the system of oxygen control. Similarly, linked biological and geological feedback mechanisms have been proposed for the regulation of global temperature. The regulation is mediated by control of  $\text{CO}_2$  in the atmosphere at concentrations that have compensated for increasing solar radiation over geological time [8].

Even though the Gaia hypothesis is speculative, at least there is more and more evidence for the dynamic role of living systems in determining the composition of many chemicals in the air, soil, and water [9]. And at a smaller geographical scale, as discussed later, there are many ecosystem processes that cybernetically regulate conditions for life.

There are three reasons why I use the Gaia hypothesis as one of my two organizing themes. First, by being rooted in questions of regula-

tion and stability through identifiable biological, chemical, and physical processes, it gives a direction for relevant scientific research—for disproof of the hypothesis if nothing else. Second, this is the only concept I know that can, in principle, provide a global rationale for giving priority to rehabilitation, protection of ecosystems, and land use management. If  $\text{CH}_4$  production, for example, provides an essential negative feedback control for ozone ( $\text{O}_3$ ) concentration, then the recent 1–2% annual increase in  $\text{CH}_4$  content is important and priority should be given to considering major changes in its primary sources—i.e., wetlands, biomass burning, and ecosystems containing ruminants and termites [10]. Finally, an examination of global change concerns not only science but also policy and politics. In a polarized society where certitude is lacking, Gaia has some potential for bridging extremes by providing a framework for understanding and action.

#### Surprise

Just as Gaia is global, the second organizing theme of surprise is, necessarily, local. Surprise concerns both the natural system and the people who seek to understand causes, to expect behaviors, and to achieve some defined purpose by action. Surprises occur when causes turn out to be sharply different than was conceived, when behaviors are profoundly unexpected, and when action produces a result opposite to that intended—in short, when perceived reality departs *qualitatively* from expectation.

Expectations develop from two interacting sources: from the metaphors and concepts we evolve to provide order and understanding and from the events we perceive and remember. Experience shapes concepts; concepts, being incomplete, eventually produce surprise; and surprise accumulates to force the development of those concepts. This sequence is qualitative and discontinuous. The longer one view is held beyond its time, the greater the surprise and the resultant adjustment. Just such a sequence of three distinct viewpoints, metaphors, or myths has dominated perceptions of ecological causation, behavior and management [11].

*Equilibrium-Centered View: Nature Constant:* This viewpoint emphasizes not only constancy in time, but also spatial homogeneity and linear causation. A familiar image is that of a landscape with a bowl-shaped

valley within which a ball moves in a way determined by its own acceleration and direction and by the forces exerted by the bowl and gravity. If the bowl was infinitely large, or events beyond its rim meaningless, this would be an example of global stability. Such a viewpoint directs attention to the equilibrium and near-equilibrium conditions. It leads to equilibrium theories and to empirical measures of constancy that emphasize averaging variability in time and "graininess" in space. It represents the policy world of a benign nature where trials and mistakes of any scale can be made with recovery assured once the disturbance is removed. Since there are no penalties or size, only benefits to increasing scale, this viewpoint leads to notions of large and homogeneous economic developments that affect other biophysical systems, but are not affected by them.

*Multiple Equilibria States: Nature Engineered and Nature Resilient:* This second viewpoint is a dynamic one that emphasizes the existence of more than one stable state. In one variant the instability is seen as maintaining the resilience of ecological systems [12]. It emphasizes variability, spatial heterogeneity, and nonlinear causation. A useful image is that of a landscape of hills and valleys with the ball journeying among them, in part because of internal processes and in part because exogenous events can flip the ball from one stability domain to another. This viewpoint emphasizes the qualitative properties of important ecological processes that determine the existence of stable regions and of boundaries separating them. Continuous behavior is expected over defined periods that end with sharp changes induced by internal dynamics or by exogenous events, at times large, at times small.

The length of the period of continuous behavior often determines the magnitude of the subsequent change and affects policy recommendations. For example, one would argue from an equilibrium-centered viewpoint that climate warming due to the accumulation of "greenhouse" gases will proceed slowly enough for ecological and social processes to adapt of their own accord. Efforts to facilitate adjustment are unnecessary because existing crop types, for example, are likely to develop and be well adapted to prevailing conditions. However, the second viewpoint of dynamic, nonlinear nature suggests just the opposite: that slow changes

of the type expected might be so successfully absorbed and ignored that a sharp, discontinuous change becomes inevitable.

Similarly, spatial graininess, which is small relative to the range of movement of an organism, is presumed to be averaged out in the equilibrium-centered approach [18]. The nonlinear viewpoint, however, presents the possibility that small-scale events cascade upward, as has been described for climatic behavior [14]. But for ecological systems, Steele [15,16] notes that widely ranging animals feed on small-scale spatial variability. For example, if fish could not discover and remain in plankton patches they could not exist.

This second viewpoint can produce two variants of policy. One assumes that the landscape is fixed or that sufficient knowledge is available to keep it fixed. It is a view of nature engineered to keep variables (the ball) away from dangerous neighboring domains. It occurs in the responsible tradition of engineering for safety, of fixed environmental and health standards, and of nuclear safeguards.

The alternative variant sees that key features of the landscape are maintained by the journeys of the ball, by variability itself testing and maintaining the configuration. This is resilient nature in which the experience of instability is used to maintain the structure and general patterns of behavior. It is assumed in the design that there is insufficient knowledge to control the landscape and hence one attempts to retain variability while producing economic and social benefit [12, 17]. In such cases variables are allowed to exceed flexible limits so long as natural and designed recovery mechanisms are encouraged. Designs have been proposed for example, for dealing with pollution [18], environmental hazards [5], water resources [19], and pest management [20].

*Organizational Change: Nature Evolving:* The final viewpoint is one of evolutionary change. Later a number of examples are presented to demonstrate that successful efforts to constrain natural variability lead to self-simplification and so to fragility of the ecosystem. A variety of genetic, competitive, and behavioral processes maintain the values of parameters that define the system. If the natural variability changes, the values shift: the landscape of hills and valleys begins to alter. Stability domains shrink, key variables become more homogeneous (e.g., species



composition, age structure, spatial distribution), and perturbations that previously could be absorbed no longer can be.

The resultant surprises can be pathological if continuing control requires ever-increasing vigilance and cost. But if control is internal and self-regulated, i.e., homeostatic, then the possibility opens for organizational change because the benefits of being embedded in a larger ecological or social system significantly exceed the costs of local control.

An example from biological evolution is the remarkably constant internal temperature maintained by endothermic (warm-blooded) animals in the presence of large changes in external temperature. A large metabolic load is required to maintain a constant temperature. As expected, the range of internal temperatures that sustains life becomes narrower than for {cold-blooded} ectotherms. Moreover, the typical endotherm body temperature of around 37° C is close to the upper lethal temperature for most living protoplasm. It does not represent a "policy" of keeping well away from a dangerous threshold.

The evolutionary significance of this internal temperature regulation is that maintenance of the highest body temperature, short of death, allows the greatest range of external activity for an animal [21]. Speed and stamina increase and activity can be maintained at high and low external temperatures, rather than forcing aestivation or hibernation. There is hence an enhanced capability to explore environments and conditions that otherwise would preclude life. The evolutionary consequence of such temperature regulation was the suddenly available opportunity for dramatic organizational change and explosive radiation of adaptive life forms. Hence the reduction of internal resilience as a consequence of effective self-regulation was more than offset by the opportunities offered by other external settings.

Hence the study of evolution requires not only concepts of function but also concepts of organization—of the way elements are connected within subsystems and the way subsystems are embedded in larger systems. Food webs and the trophic relations that represent them are an example and have long been a part of ecology. Recently some revealing empirical analyses have demonstrated remarkable regularities in such ecosystem structures [22], with food webs of communities in fluctuating

environments having a more constrained trophic structure than those in constant environments [23].

These and related developments, connected in turn to hierarchy theory [24] on the one hand, and the stability and resilience concepts described earlier, on the other, are starting to provide the framework required for comprehending organizational evolution [25]. Although not as well developed as equilibrium, engineering, and resilience concepts, such developments are an essential part of any effort to understand, guide, or adapt to global change.

These views of nature represent the different concepts people have of the way natural systems behave, are regulated, and should be managed. Surprise can occur when the real world is found to behave in a sharply different way from that conceived. The perception can be ignored, resisted or acknowledged depending on how extreme the departure is and depending on how flexible and adaptable the observer is. Although observer and system are interlinked, I do not explore the psychology and dynamics of individual, institutional, and social adaptation in this chapter, though this is ultimately necessary if we want to understand and design sustainable systems. But in the next section I examine a number of ecological systems to determine which of the views of nature most closely matches reality.

## Dynamics of Ecosystems

### *Resilience and Stability*

This chapter relies heavily on the distinction between resilience and stability. Since that distinction was first emphasized [12] a significant literature has developed to test its reality in nature, to expand the theory, and to apply this to management and design. Much of what follows is drawn from the literature. The distinction relies on definitions that recognize the existence of different stability structures of the kind described in the previous section. There are four main points. First, there can be more than one stability region or domain, i.e., multiequilibrium structures are possible. Second, the behavior is discontinuous when variables (i.e., elements of an ecosystem) move from one domain to another because they become attracted to a different equilibrium condition. Third, the precise

kind of equilibrium—steady state or stable oscillation—is less important than the fact of equilibrium. Finally, parameters of the system that define the existence, shape, and size of stability domains depend on a balance of forces that may shift if variability patterns in space and time change. In particular, reduced variability through management or other activities is likely to lead to smaller stability regions whose contraction can lead to sharp changes because the stability boundary crosses the variables, rather than the reverse.

This leads to the following definitions. Stability (*sensu stricto*) is the propensity of a system to attain or retain an equilibrium condition of steady state or stable oscillation. Systems of high stability resist any departure from that condition and, if perturbed, return rapidly to it with the least fluctuation. It is a classic equilibrium-centered definition.

Resilience, on the other hand, is the ability of a system to maintain its structure and patterns of behavior in the face of disturbance. The size of the stability domain of residence, the strength of the repulsive forces at the boundary, and the resistance of the domain to contraction are all distinct measures of resilience.

Stability, as here defined, emphasizes equilibrium, low variability, and resistance to and absorption of change. In sharp contrast, resilience emphasizes the boundary of a stability domain and events far from equilibrium, high variability, and adaptation to change.

However, one school of ecology so strongly emphasizes linear interactions and steady state properties [26, 27, 28] that resilience is treated in the opposite way to that described above. It is defined as “how fast the variables return towards their equilibrium following a perturbation [28]” and is measured by the characteristic return times. In terms of the definitions used in this chapter, this concerns only one facet of stability and has nothing to do with the qualitative distinctions that I believe are important.

In addition to the growing number of tests and demonstrations of the key features of resilience, there have been two major expansions of theory and example. One is Levin's excellent analysis and review of patterns in ecological communities [29]. Levin first placed experimental, functional, and behavioral descriptions within a formal mathematical frame. More important, he made two qualitative additions. One was to explore spatial patterns of multistable systems by analyzing the consequences of

diffusion. The second was to make a sharp distinction between variables associated with different speeds or rates of activity, partly to facilitate analysis but more to stress the consequences of coupling subsystems whose cycles are of different lengths. The second major expansion was that of Allen and Starr who extended the analyses of ecosystem patterns for a wide range of examples [25]. Most significantly, they embedded theory, measurement, and modeling relevant to resilience and stability into a hierarchical framework. More than any recent development, this framework provides a means of studying community structure and of treating evolution or organizational change.

#### *Ecosystem Scale*

These three developments in analysis—of multistable systems [12], of spatial diffusion [29], and of hierarchies [25]—concern the coupling of nonlinear subsystems of different scales in time and space. They are fundamental to understanding how predictable change is, whether or not historical accidents are important, and how to achieve a balance between anticipatory design and adaptive design. Clark [30] has provided a useful classification of the relevant scales for a wide range of geophysical, ecological and social phenomena. The scales range from square centimeters to global and from minutes to thousands of years. In the present analysis I concentrate on ecological systems covering scales from a few square meters to a few thousand square kilometers and from a few years to a few hundred years.

These scales represent ecosystems, which are defined here as communities of organisms in which internal interactions between the organisms determine behavior more than do external biological events. External abiotic events do have a major impact on ecosystems, but are mediated through strong biological interactions within the ecosystems. It is through such external links that ecosystems become part of the global system. Hence, the spatial scale is determined by the dispersal distance of the most mobile of the key biological variables. The structure of eastern North American spruce–fir forests, for example, is profoundly affected by the spruce budworm, which periodically kills large areas of balsam fir. The modal distance of dispersal of adult budworms is of the order of 50 km [20], but movements are known to extend up to 200 km. The relevant spatial area over which internal events dominate can therefore cover a

good part of east-central North America. And the minimum area for analysis has to be of the order of 70 000 km<sup>2</sup>.

Similarly, the time span of up to a few hundred years is set by the longest-lived (slowest-acting) key biological variables. In the case of the spruce-fir forests the trees are the slowest variables with a rotation age of about 70 years. Any effective analysis therefore must consider a time span that is a small multiple of that—of the order of 200 years.

Eugene Odum [31], more than anyone else, has emphasized that such ecosystems are legitimate units of investigation, having properties of production, respiration, and exchange that are regulated by biological, chemical, and physical processes. Hence they represent distinct subsystems of the biogeochemical cycles of the Earth. They are open, since they receive energy from the sun and material and energy from larger cycles. In regulating and cycling this material through biotic and abiotic processes, outputs are discharged to larger cycles. Ecosystems hence are Gaia writ small.

#### *Succession*

One dominant theme of ecosystem study has been succession—the way complexes of plants develop after a disturbance. Clements' scheme of succession has played an important role in guiding study and theory [32]. He emphasized that succession leads to a climax community of a self-replicating assemblage of plants. The species comprising the assemblage are determined by precipitation and temperature. Plant colonization and growth are seen as proceeding to the stable climax. Initial colonization is by pioneer species that can grow rapidly and withstand physical extremes. They so ameliorate these conditions as to allow entry of less robust but more competitive species. These species in turn inhibit the pioneers but set the stage for their own replacement by still more effective competitors. Throughout this process, biomass accumulates, regulation of biological, chemical, and physical processes becomes tighter, and variability is reduced until the stable climax condition is reached and maintained. This scheme represents a powerful equilibrium-centered view in which disturbances by fire, storm, or pest are treated as exogenous (and somehow inappropriate) intrusions into a natural order. Clements gave an analogy to an organism and its ability to repair damage [32].

During the past 15 years this view has been significantly modified by a wide range of studies of ecosystems—some dominated by disturbance, some not—and by experimental manipulation of defined ecosystem units such as the classic Hubbard Brook Watershed Study [33]. Before describing those developments, however, it is useful to relate this view of succession to another powerful equilibrium-centered notion: that of *r* and *K* strategies.

MacArthur and Wilson [34] proposed this classification to distinguish between organisms selected for efficiency of food harvest in crowded environments (*K*-selected) and those selected simply to maximize returns without constraint (*r*-selected). The designations come from the terms of the logistic equation, where *K* defines the saturation density (stable equilibrium population) and *r* the instantaneous rate of increase. MacArthur [35] pointed out the contrast between “opportunistic” species in unpredictable environments (*r*-strategists) and “equilibrium” species in predictable ones (*K*-strategists). Pianka [36] and Southwood et al. [37] have emphasized that these represent extremes of a continuum, but that a variety of life histories and biological and behavioral features correlate with the two strategies. Briefly, *r*-strategists have a high reproductive potential, short life, high dispersal properties, small size, and resistance to extremes. They are the pioneers of newly disturbed areas or the fugitive species that ever occupy transient habitats. *K*-strategists have lower reproductive potential, longer life, lower dispersal rates, large size, and effective competitive abilities. They represent, therefore, the climax species of Clements or those that occupy stable, long-lasting habitats.

There clearly are communities that have developed a climax maintained through plant-by-plant replacement in the manner proposed by Clements. Lorimer [38], for example, examined the history of pre-settlement forests in northeastern Maine, USA, and found that the time interval between severe disturbances was much longer than that needed to obtain a climax, all-age structure. Other examples are presented in an extensive review of forest succession edited by West et al. [39]. But the Clementsian view of succession as analogous to the recovery of an organism from injury, with an ordered and obligatory sequence of replacements of one species by another, is oversimplified and limiting for several reasons.

First, many communities are subjected to regular or irregular disturbances severe enough to kill established plants over areas of a few square meters (the size of a tree) to several thousand square kilometers. Traditionally these disturbances—fires, landslides, storms, floods, disease, insect pests, and herbivore grazing—have usually been viewed as external to the system. But when they occur at a frequency related to the life span of the longest-lived species, the plants themselves can become increasingly adapted to the disturbance and so make the event an internally triggered and maintained phenomenon. This is particularly well recognized for fire. Mutch [40], for example, demonstrated that vegetation of fire-adapted species was significantly more combustible than that of related species in communities not subject to fires. Similarly, Biswell [41] describes the twig development and proliferation of chaparral species that significantly increase the inflammability of plants that are 15 years and older. This coincides with a typical burn cycle of similar duration. Such “accidents designed to happen” are more common than is usually recognized and further examples involving agents other than fire are described later.

As a consequence, there are many examples of what I imagine consistent Clementsian ecologists would be forced to see as self-inflicted wounds to the ecosystem “organism”. Such disturbances have a wide range of periodicities set by the dynamics of the slowest variable [42]. Fire frequency in the Pacific Northwest of North America, for example, occurs every 400–500 years, and this period is related to the potentially 100-year life span of Douglas fir [43]. Eastern white pine forests experience a fire periodicity of 100–300 years in presettlement times [44, 45], while cyclic changes of 200-year periodicity are proposed for elephant populations; this period is due to the recovery time required for the tasty (to elephants) and long-lived baobab trees of East Africa [46]. The fire controlled period throughout much of the boreal forest of northern North America was between 51 and 120 years [47, 48]. Chaparral in California [49] is adapted to a more frequent cycle of 10–50 years.

The second significant departure from Clements' notions is even more fundamental. Some disturbances can carry the ecosystem into a different stability configuration or domain. At times this happens after a long period of exploitation has apparently reduced the resilience of the ecosystem. For example, fishing in the Great Lakes has been argued to

have set the stage for a radical change in fish communities from a system dominated by a large species to one of smaller species [50], the overall biomass remaining constant. The resilience was reduced to a level where small stresses from the physical environment, from man, or from biological invasion triggered the new configuration. In a similar way, shifts of savannas from mixed grass-shrub systems to ones dominated by shrubs are often triggered by a modest drought after being conditioned by an extended period of cattle grazing [51]. In other instances, the magnitude of the triggering event is so great or the resilience of the ecosystem so naturally low that new configurations emerge quite independent of previous management. Bormann and Likens [33, p. 189] present one such example of a burn in a spruce-hardwood forest on thin soils that transformed part of the ecosystem into a bare-rock-shrub system. Hence there is not just one climax state; there can be more than one.

Third, species that are important late in the sequence can be present together with pioneer species at the initiation of old field succession [52] or forest succession [43,48]. The resultant successional sequence is hence much more in the form of a competitive hierarchy as described by Horn [42]. Early in a sequence, the opportunist species grow rapidly, dominate for a short time, but ultimately cannot withstand crowding and competition from other more persistent species. Marks [53] presents a particularly clear demonstration of this sequence and of the opportunist role of pin cherry in reestablishing disturbed hardwood systems in New Hampshire, USA. Late in the sequence pin cherry trees are almost totally absent, having been squeezed out by more competitive trees like beech and red maple. After disturbance, however, seeds long dormant in the soil germinate, pin cherry trees flourish, and begin to be eliminated again after about 20 years.

Finally, invasion of species after disturbance as well as during succession is highly probable, particularly in tropical lowlands [54]. This, combined with the competitive hierarchical relations mentioned above, makes the tropical forest highly individual in character and very diverse. So many species are capable of filling a particular niche that succession is better described by life history traits and tree geometry. There is, moreover, considerable advantage in dealing with succession in terms of such properties, since they determine successional status [55]—whether in tropical, temperate, or arctic regions.



That is why the idea of *r*-selected and *K*-selected species was introduced earlier. Each strategy is associated with distinctive traits that, in exaggerated form, contrast the exploitative and opportunistic species that dominate early in the succession with the consolidating or conservative species that dominate later through competition. Moreover, the terms can be used to refer to two principal ecosystem functions: exploitation and conservation. Early in succession biotic and abiotic exploitative processes dominate. These lead to the organization and binding of nutrients, rapid accumulation of biomass, and modification of the environment. Eventually, conservative forces begin to dominate, with competition being the most important aspect. This leads to increased organization through trophic and competitive connections, to reduced variability, and, if not interrupted, to reduced diversity.

Ecosystems, however, are also systems of discontinuous change. In addition to the successional processes leading to increasing order there are periods of disorganization. The examples mentioned earlier were of large-scale disruptions that affect extensive areas. But change of this kind also occurs in the ecosystems that most closely achieve a climax condition. Individual trees senesce, creating local gaps. The only difference is that the spatial scale is small and the disruptions are not necessarily synchronous. A complete picture of the dynamics of ecosystems therefore requires additional functions to those of exploitation and conservation. Such functions relate more to the generation of change and the introduction of disorder.

#### *Forces of Change*

To identify these functions and their effects, I initially analyze a small number of examples of ecosystems that demonstrate pronounced change and that have been examined in detail. They can be classified as follows: forest insect pests, forest fires, grazing of semiarid savannas, fisheries, and human disease. Many of the examples also represent systems that have been subjected to management. In a sense, the management activities can be viewed as diagnostic, for they introduced external changes that helped expose some of the internal workings of the natural, unmanaged system. In addition, a number of the management approaches were very much dominated by the goal of achieving constancy through externally imposed regulations. Hence the implicit hypothesis was an

equilibrium-centered one and the experiences in managing forests, fish, and other organisms can be viewed as weak tests of that hypothesis.

To give an impression of the consequences I consider the following examples:

- (1) Successful suppression of spruce budworm populations in eastern Canada using insecticides certainly preserved the pulp and paper industry and employment in the short term by partially protecting the forest. But this policy has left the forest and the economy more vulnerable to an outbreak over a larger area and of an intensity not experienced before [20].
- (2) Suppression of forest fire has been remarkably successful in reducing the probability of fire in the national parks of the USA. But the consequence has been the accumulation of fuel to produce fires of an extent and cost never experienced before [56].
- (3) Semiarid savanna ecosystems have been turned into productive cattle-grazing systems in the Sahel zone of Africa, southern and east Africa, the southern USA, northern India, and Australia. But because of changes in grass composition, an irreversible switch to woody vegetation is common and the systems become highly susceptible to collapse, often triggered by drought [51].
- (4) Effective protection and enhancement of salmon spawning on the west coast of North America are leading to a more predictable success. But because this triggers increased fishing and investment pressure, less productive stocks become extinct, leaving the fishing industry precariously dependent on a few enhanced stocks that are vulnerable to collapse [57].
- (5) Malaria eradication programs in Brazil, Egypt, Italy, and Greece have been brilliant examples of sophisticated understanding combined with a style of implementation that has all the character of a military campaign. But in other areas of the world, where malaria was neither marginal nor at low endemic levels, transient success has led to human populations with little immunity, and mosquito vectors resistant to DDT. As a consequence, during the past five years some

countries have reported a 30- to 40-fold increase in malaria cases compared with 1969-1970, signaling a danger not only to the health of the population, but also to overall socio-economic development.

In each of these examples, the policy successfully reduced the probability of an event that was perceived as socially or economically undesirable. Each was successful in its immediate objective. Each produced a system with qualitatively different properties. All of these examples, and others that fall into the five classes, represent "natural" systems that are coupled to management institutions and to the society that experiences the success or endures the failure of management. Here I focus principally on the natural system.

Despite the large number of variables in each example, the essential causal structure and behavior can be represented by interaction among three sets of variables. These represent three qualitatively different speeds, or rates of activity, corresponding to rates of growth, generation times, and life spans (Table 1). It becomes possible, as a consequence, to proceed in two directions: to achieve a qualitative understanding of the natural system and to achieve a detailed policy design. The first objective draws upon the theory of differential equations [58]. The second draws upon more recent developments in simulation modeling and optimization [20]. Both are very much connected to hierarchy theory [25].

In many of the examples, both objectives have been pursued. Here I concentrate more on the efforts to achieve a qualitative understanding in order to define research agendas. The other objective, concentrating on detail, is also useful but more in terms of defining operational management agendas. Since distinctively different speeds can be identified for the variables, four steps of analysis are possible [58]:

- (1) Analyze the long-term behavior of the fast variable, while holding the slow variables constant.
- (2) Define the response of the slow variables when the fast ones are held fixed.
- (3) Analyze the long-term behavior of the slow variables, with the fast variables held at their corresponding equilibria.

TABLE 1: Key variables and speeds in five classes of ecosystems.

The system	The variables			Key references
	Fast	Intermediate	Slow	
Forest insect pest	Insect	Foliage	Trees	[58, 59]
Forest fire	Intensity	Fuel	Trees	[60]
Savanna	Grasses	Shrubs	Herbivores	[51]
Fishery	Phytoplankton	Zooplankton	Fish	[16]
Human disease	Disease organism	Vector and susceptibles	Human population	[61, 62]

- (4) Combine the preceding steps to identify needs for extra coupling, so that, when added, the behavior of the full system is described.

I now summarize the procedure, emphasizing the main conclusions. The "fast" dynamics are determined by the way key processes affect change in the fast variable at different fixed values of the slow variables. An example is shown in Figure 1. The important point is that a long history of experimental analysis of ecological processes has led to generalization of the qualitative form of system response, the condition for each distinct form, and the features that determine where impact is greatest (see Holling and Buckingham [63] for predation and competition and Peters [64] for a variety of ecological processes related to body size). Hence, this knowledge can be applied to understanding behavior at a more aggregate level of the hierarchy. Equally important, it frees the analysis from the need for detailed quantification, setting the stage for research designs that are both more economical and more appropriate.

In continuing this qualitative emphasis, attention is then focused on conditions for increase and conditions for decrease of the fast variable. The boundary between the two represents either transient or potentially stable equilibria. The conditions for these equilibria can be organized to

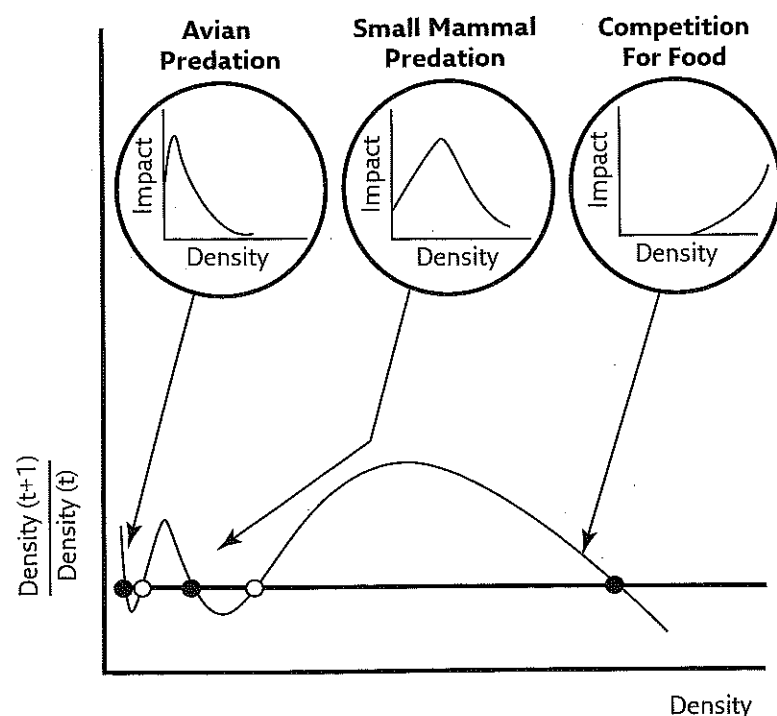


FIGURE 1: A stylized recruitment curve for jack-pine sawfly at a fixed level of slow and driving variables, showing the contributions of three of the key processes. The horizontal line represents the conditions where the population density of the next generation is the same as that of the present generation. Intersections of the recruitment curve with this line indicate potential equilibria, some potentially stable (closed circles), some unstable (open circles).

show the set of all equilibria, i.e., the zero isoclines for the fast variable. Four examples are shown in Figure 2: for spruce budworm (a), jack-pine sawfly (b), forest fire (c), and savanna grazing (d).

These equilibrium structures show that there are a number of stability states controlled by the slow variable. There are many other examples given in the literature: for 15 other forest insect pests [59], for other grazing systems [62,65], for fish [17,66], and for human host-parasite systems [62].

Two main points emerge at this stage of the analysis. First, discontinuous change occurs because there are multiple stable states. As the

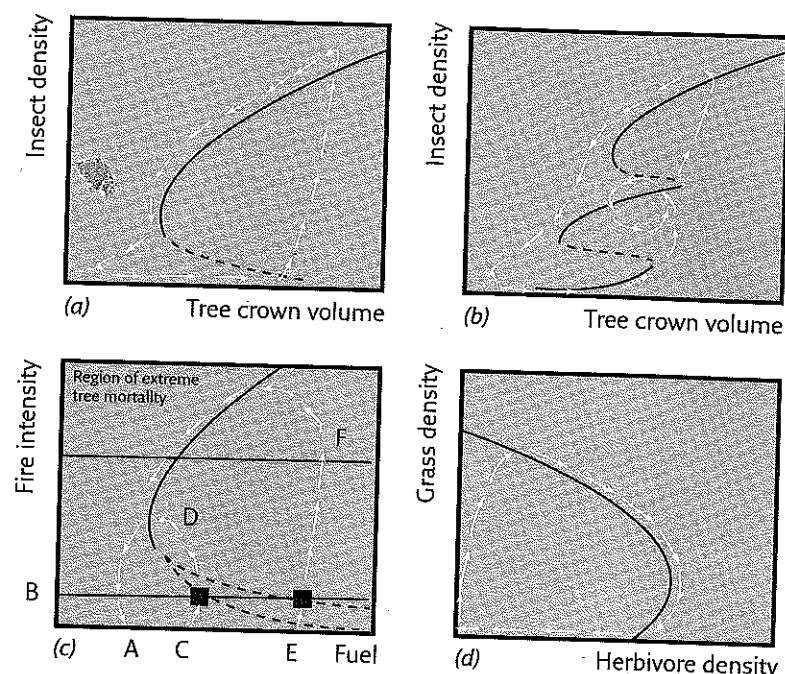


FIGURE 2: Zero-isocline surfaces showing the equilibrium values of the fast variable at different fixed levels of the slow variables. The full lines represent stable surfaces and the broken lines unstable ones. Typical trajectories are shown by the arrows. (a) Spruce budworm and balsam fir; (b) jack-pine and sawfly; (c) forest fire and fuel; (d) savanna grass and herbivore grazing.

slow variables change (tree growth, fuel accumulation, herbivore population increase), different equilibria suddenly appear, and when other equilibria disappear, the system is suddenly impelled into rapid change after periods of gradual change. The basic timing of these events is set by the dynamics of the slow variable.

Second, external stochastic events can lead to highly repetitive consequences. All of the surfaces shown in Figure 2 would be better represented as fuzzy probability bands to reflect "white noise" variability in weather conditions. But this modification typically changes the precise timing of events by a trivial amount. For long periods the systems are in a refractory state and the triggering event is totally or strongly inhibited.



In the insect pest cases [Figures 2(a) and (b)], for example, a variety of predators, chiefly birds, introduce such a strong "predator pit" that insect populations are either becoming extinct or being kept at very low densities. Similarly, the reflexively folded set of unstable equilibria for fire [Figure 2(c)] can turn stochastic ignition events, such as lightning strikes, into highly predictable outbreaks of fire. If the surface is low [dotted line in Figure 2(c)], then the average ignition intensity of B triggers a fire at C which consumes the fuel and hence extinguishes itself at A. This is similar to the cycles of ground fires experienced prior to fire management in the mixed-conifer forests of the Sierra Nevada in western USA [56,60]. In several areas they occurred with a remarkably consistent interval of seven to eight years, and helped maintain conditions for tree regeneration and nutrient cycling. In addition, these light fires killed only some of the young white fir, thereby introducing and maintaining gaps in the forest canopy and, in essence, producing natural fire breaks. However, if the undersurface is raised because of increased moisture or effective fire control practices [broken line in Figure 2(c)], more fuel must accumulate before an average ignition event triggers a fire [point E, Figure 2(c)]. This results in a longer period before a fire, but also in a more intense fire, corresponding to a natural long-period fire cycle of the kind mentioned earlier or to the unexpected failure of a fire control policy.

Because of these properties, pulses of disturbance should not be seen as exogenous events. Insect outbreaks, forest fires, overgrazing, sudden changes in fish populations, and outbreaks of disease are determined by identifiable processes affecting the fast variable, whose impacts are modified by the magnitudes of the slow variables. As a consequence, changes in the slow variables eventually result in a condition where a sharp disturbance is inevitable.

A fuller definition of those properties requires two more steps. First, the equilibrium structure set by the fast variable is affected by both intermediate-speed and slow variables. A three-dimensional representation of the zero isocline can then be shown. An example for spruce budworm is shown in Figure 3. Finally, zero-isocline surfaces are constructed for both the intermediate and slow variables in order to explore the intersections between them. A more formal and rigorous treatment can be found in Ludwig et al. [58]. Figure 4 shows, as an example, the iso-

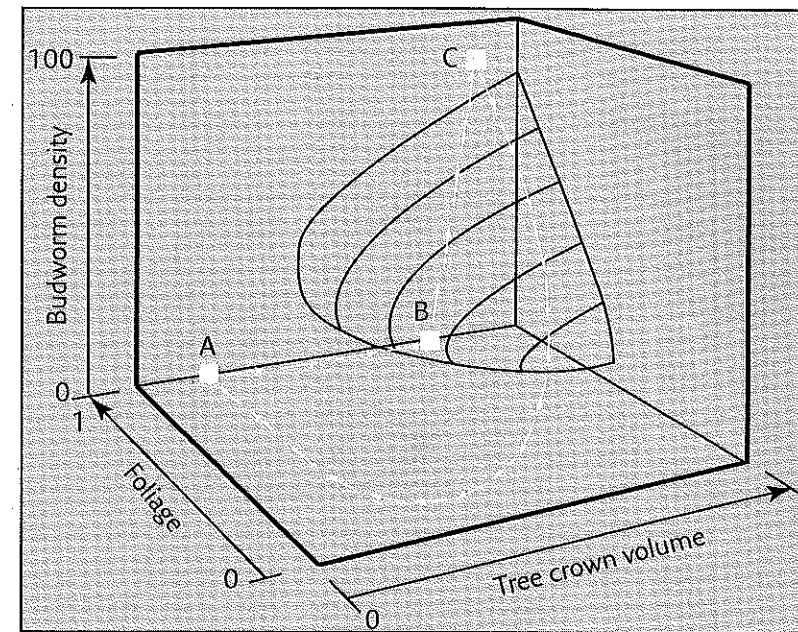


FIGURE 3: Zero-isocline surface for budworm as a function of foliage and tree crown volume. The trajectory shows a typical unmanaged outbreak sequence.

cline surface for tree crown volume laid over that for budworm. Where these two surfaces intersect (line AB) represent the only places where a stable equilibrium for both budworm and trees might be possible. But this can only be realized if the stable portion of the zero isocline for foliage, the intermediate-speed variable, also intersected the line AB. The surface for foliage is folded something like that of the budworm, with a stable surface and an unstable reflexed one. For values of foliage area below this unstable surface, foliage production cannot match natural foliage depletion, so that the foliage eventually disappears. Although it cannot be clearly shown in Figure 4, it happens that the foliage zero-isocline surface lies under the major portion of the budworm surface. As a consequence, there is no stable intersection.

Thus the unmanaged budworm system is in a state of continuous and fundamental disequilibrium. If one variable is on a stable zero isocline, the others are usually not on theirs. If two of the variables happen to be simultaneously on their stable isoclines, the third one is never on its



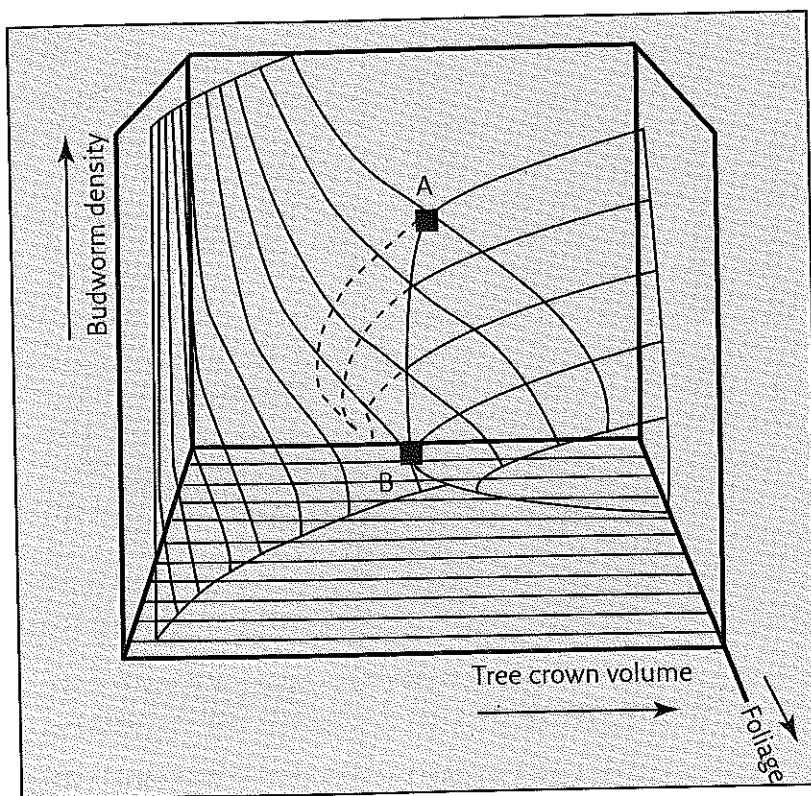


FIGURE 4: Overlay of the zero-isocline surface for tree crown volume (a measure of forest age) against the budworm surface of Figure 2. The line AB is where the two surfaces intersect.

stable isocline. It is a system under continual dynamic change, always chasing ever-receding equilibria. But control is never completely lost because of the existence of the independent single- or two-variable equilibrium states.

Only an analysis involving three variables would expose this behavior. In such cases, equilibrium-centered concepts, definitions, and measurements that require the existence of at least one non-zero equilibrium (termed here a system equilibrium) are simply irrelevant. There is no system equilibrium.

With these examples I can now add to the review of the main conclusions by summarizing the role of the three distinct speeds of variables in

producing cyclic variations of different periods. Under different conditions each of these speeds can dominate the dynamics.

In the spruce-budworm system the period is set by the slowest variable, the tree, and the other variables interact with it. The same is true of unmanaged savanna-grazing systems, where intensive periods of overgrazing lead to depletion of above-ground vegetation. This is followed by emigration or high mortality of ungulates allowing early recovery of perennial grasses with underground storage. Sometime later, annual grasses begin to dominate through competition with the former. Many forest fire systems are similarly controlled, in that the slowest variable sets the cycle and the fast and intermediate variables follow.

In all these examples the variability produces diversity as a consequence of a cyclic shifting of the competitive advantage between species. Balsam fir can outgrow spruce, and would do so except that budworm preferentially attacks balsam and suddenly shifts the balance [67]. The high photosynthetic efficiency of annual grasses places them at an advantage over those perennial grasses that invest a considerable part of their biomass in underground storage. But intense overgrazing tips the balance the other way, so that both types are refused [51]. And Loucks' analysis of long-period forest fire cycles makes a similar point regarding maintenance of species diversity [44]. In all these cases the cycle length is set by the slowest variables and other variables are driven according to that cycle. In every example, the high variability encourages species diversity and spatial heterogeneity.

In a second class of cases the basic timing is set by the intermediate-speed variable. The slowest variable is largely disengaged because a stable oscillation develops around a single system equilibrium in which this variable persists. An example is the ground fire cycle described earlier for the mixed-conifer forests of the Sierra Nevada [Figure 2(c), cycle C-A]. A number of forest insect systems show this pattern [59]. For example, the European larch-budmoth system in Alpine regions of Switzerland [68] shows a remarkably persistent cycle with an 8- to 10-year period that has persisted for centuries. Both insect and foliage follow this cycle but there is little tree mortality.

Finally, there are patterns in which the fast variable dominates and the intermediate and slow variables are little affected. A particularly interesting example is the jack-pine-sawfly system, [69] and Figure 2(b).

At intermediate tree ages, predation by small mammals establishes a "predator pit" at moderate population densities of sawfly. At these densities, the small-mammal predation causes enough mortality to allow the sawfly and its parasites to establish a high-speed stable limit cycle of 3-4 years. Foliage is little affected and, as a consequence, neither is tree mortality. Such oscillations can persist for some years, but eventually the system is shifted to a different pattern by a change in climate or forest stand. Other forest insect systems show this pattern, as does endemic malaria where vectoral capacity is high [61].

In summary, the key features of this analysis of the forces of change lead to the following observations:

- (1) There can be a number of locally stable equilibria and stability domains around these equilibria.
- (2) Jumps between the stability domains can be triggered by exogenous events, and the size of these domains is a measure of the sensitivity to such events.
- (3) The stability domains themselves expand, contract, and disappear in response to changes in slow variables. These changes are internally determined by processes that link variables and, quite independently of exogenous events, force the system to move between domains.
- (4) Besides exogenous stochastic events, different classes of variability and of temporal and spatial behavior emerge from the form of equilibrium surfaces and the manner in which they interact. There can be conditions of low equilibrium with little variability. There can be stable-limit cyclic oscillations of various amplitudes and periods. And there can be dynamic disequilibrium in which there is no global equilibrium condition and the system moves in a catastrophic manner between stability domains, occasionally residing in extinction regions. There also exists the possibility of "chaotic" behavior.

The one overall conclusion is that discontinuous change is an internal property of each system. For long periods change is gradual and discontinuous behavior is inhibited. Conditions are eventually reached,

however, when a jump event becomes increasingly likely and ultimately inevitable.

### Synthesis of Ecosystem Dynamics

#### *Ecosystem Functions*

It was mentioned in an earlier section that there are two aggregate functions that determine ecosystem succession: an exploitation function (related to the notion of *r*-strategies) that dominates early and a conservation function (related to *K*-strategies) that dominates late in the succession. The conclusion of the preceding analysis of forces of change is that there is a third major ecosystem function. The increasing strength of connection between variables in the maturing ecosystem eventually leads to an abrupt change. In a sense, key structural parts of the system become "accidents waiting to happen."

When the timing is set by the slowest variable, the forces of change can lead to intense, widespread mortality. When the timing is set by the faster variables the changes are less intense and the spatial impact, while synchronous over large areas, is more patchy. But even in those instances, individuals constituting the slow variables eventually senesce and die. The difference is that the impact is local and is not synchronous over space.

There is both a destructive feature to such changes and a creative one. Organisms are destroyed, but this is because of their very success in competing with other organisms and in appropriating and accumulating the prime resources of energy, space, and nutrients. The accumulated resources, normally bound tightly and unavailable, are suddenly released by the forces of change. Such forces therefore permit creative renewal of the system. I call this third ecosystem function "creative destruction," a term borrowed from Schumpeter's economic theory [70].

Although the change is triggered by such a function, the bound energy, nutrients, and biomass that accumulated during the succession are not immediately available. There is therefore a fourth and final ecosystem function. One facet of that function is the mobilization of this stored capital through processes of decomposition that lead to mineralization of nutrients and release of energy into the soil. The other facet includes

biological, chemical, and physical processes that retain these released nutrients, minimizing losses from leaching. This fourth function is one of ecosystem renewal.

These processes result in a pulse of available nutrients after disturbance. In many instances surprisingly little is lost from the ecosystem through leaching. In other instances so much is lost that algal blooms may be triggered in receiving waters [33, 71]. The kinds of retention mechanisms are not well understood because of the difficulty of studying soil dynamics at an ecosystem scale. But experimental manipulation of whole watersheds through harvesting, removal of structural organic material from the soil surface, and herbicidal inhibition of vegetative regrowth has begun to allow some of the mechanisms to be identified [33, 72]. They include colloidal behavior of soil, rapid uptake by the remaining vegetation whose growth is accelerated by the disturbance, and low rates of nitrification that keep inorganic nitrogen in ammonia pools rather than as the more soluble nitrates [73]. In addition a recent experiment demonstrated that rapid uptake of nutrients by microbes during decomposition is a major process preventing nitrogen losses from areas of harvested forests [72].

Such processes of release and retention after disturbance define the renewal function. Hill [74] emphasizes their importance in reestablishing the cycle of change and hence in determining the resilience of ecosystems. Of particular importance are the processes of retention. When savannas become dominated by woody shrubs, it is because of the loss of water retention capacity of perennial plants and soil. Similarly, intensified burning of upland vegetation in Great Britain has caused the vegetation to shift irreversibly from forest cover to extensive blanket bogs [75]. On sites where soils are poor, rainfall is high, and temperatures are low, the result has been loss of nutrients and reduced transpiration and rainfall interception, leading to waterlogged soils, reduced microbial decomposition, and the development of peat. The original tree species, such as oak, cannot regenerate because of wetness, acidity, and nutrient deficiency. In a similar vein, tropical rain forests may have a low resilience to large-scale disturbance. Many of the tree species have large seeds with short dormancy periods. These features facilitate rapid germination and regrowth of vegetation in small disturbed areas, but make it impossible to recolonize extensive areas of cleared land [76]. Partly as

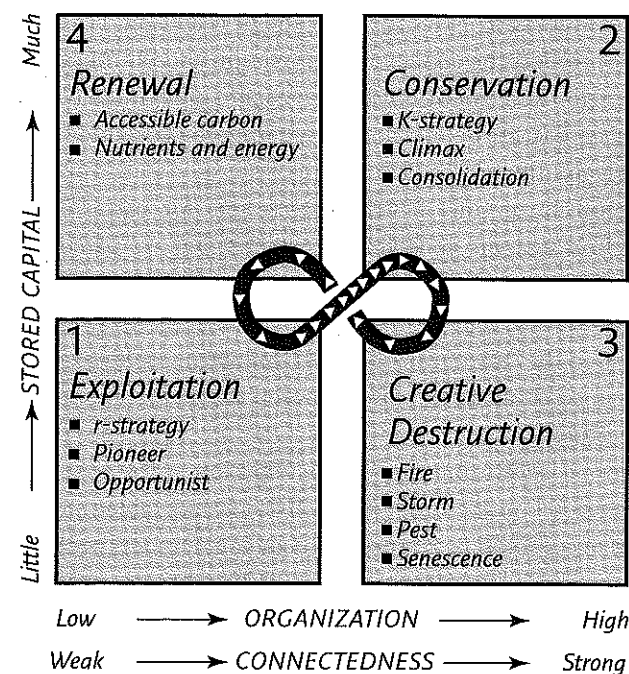


FIGURE 5: The four ecosystem functions and their relationship to the amount of stored capital and the degree of connectedness. The arrowheads show an ecosystem cycle. The interval between arrowheads indicates speed, i.e., a short interval means slow change, a long interval rapid change.

a result, extensive land clearance in the Amazon basin has led to permanent transformation of tropical forest areas into scrub savanna [77].

The full dynamic behavior of ecosystems at an aggregate level can therefore be represented by the sequential interaction of four ecosystem functions: exploitation, conservation, creative destruction, and renewal (Figure 5). The progression of events is such that these functions dominate at different times: from exploitation, 1, slowly to conservation, 2, rapidly to creative destruction, 3, rapidly to renewal, 4, and rapidly back to exploitation. Moreover, this is a process of slowly increasing organization or connectedness (1 to 2) accompanied by gradual accumulation of capital. Stability initially increases, but the system becomes so over-connected that rapid change is triggered (3 to 4). The stored capital is then released and the degree of resilience is determined by the balance



between the processes of mobilization and of retention. Two properties are being controlled: the degree of organization and the amount of capital accumulation and retention. The speed and amplitude of this cycle, as indicated earlier, are determined by whether the fast, intermediate, or slow variable dominates the timing.

These patterns in time have consequences for patterns in space. Rapidly cycling systems generate ecosystems that are patchy. Tropical ecosystems are an example. Slowly cycling systems produce higher amplitude, discontinuous change that tends to occur as a wave moving across space. In the case of uncontrolled budworm outbreaks, for example, the wave takes about 10 years to sweep across the province of New Brunswick, Canada.

The factors determining the size distribution of the areas of disturbance, however, are not well understood. If the considerable understanding of time dynamics could be connected to an equal understanding of spatial patch dynamics, then questions of global change could be better anticipated and better dealt with by local policies.

Levin [129, 78] has developed an effective framework for analysis and description of patches. There are two parts:

- (1) Patch size and age distributions as related to birth and death rates of patches.
- (2) The response of species to the regeneration opportunities existing in patches of different sizes and ages.

The "fast" and "slow" designations are part of the analysis, as well as diffusion and extinction rates. It is, therefore, completely compatible with the analysis presented here, and has begun to be applied to forest systems [79] together with Mandelbrot's theory of fractals to relate extinction laws and relative patchiness [80]. As a consequence Mandelbrot proposes a descriptive measure of patchiness and succession that is scale independent and has considerable value for any effort to measure patch dynamics and disturbance.

#### *Complexity, Resilience, and Stability*

This synthesis helps clarify the relationship between complexity and stability. It was long argued that more species and more interactions

in communities conferred more stability, the intuitive notion being that the more pathways that were available for movement of energy and nutrients, the less would be the effect of removal of one. However, May's analysis of randomly connected networks showed that increased diversity, in general, lowered stability [81]. This means that ecosystems are not randomly connected. The issue has been significantly clarified by Allen and Starr [25] and the treatment here provides further support of May's observation.

First, measures of stability referred to typically did not distinguish between stability and resilience—systems with low stability can often demonstrate high resilience. Second, ecosystems have a hierarchical structure, and for this reason it has been possible to capture the essential discontinuous behaviors with three sets of variables operating at different speeds. Other species and variables are dramatically affected by that structure and the resultant behavior, but do not directly contribute to it. Hence the relevant measures of species diversity, which is one measure of complexity, should not involve all species, but only those contributing to the physical structure and dynamics.

The significant measure of complexity, therefore, concerns the degree of connectedness within ecosystems. As Allen and Starr demonstrate, the higher the connectedness, i.e., the complexity, the lower the probability of stability. They present examples both from theory and from the empirical literature to demonstrate the point. A system can also become so underconnected that critical parts go their own way independent of each other—resilience disappears. Tropical farming based upon monocultures and extensive land clearance is highly overconnected, particularly through pest loads [82]. Hence stability and ultimately resilience are lost. Introduction of patches through traditional shifting of agriculture or through the breaking of connections by interplanting different cultivars produces a farming pattern more akin to the highly patchy, less connected natural system.

Third, the pattern of connectedness and the resultant balance between stability and resilience are a consequence of the pattern of external variability that the system has experienced. Systems such as those in the tropics, which have developed in conditions of constant temperature and precipitation, therefore demonstrate high stability but low resilience. They are very sensitive to disturbances induced by man. On the



other hand, temperate systems exposed to high climatic variability typically show low stability and high resilience, and are robust to disturbance by man.

The present analysis adds an important fourth ingredient. Hierarchies are not static in the kinds or strengths of connections. The degree of connectedness changes as the ecosystem is driven by the four ecosystem functions. Succession introduces more connectedness, and hence increasing likelihood of instability. An overconnected condition develops, triggering a discontinuous change. The connectedness is sharply reduced thereby, to be followed by reorganization and renewal. The destabilizing effect produced by overconnectedness generates variability, which in turn encourages the development and maintenance of processes conferring resilience, particularly during the period of low connectedness and recovery. Collapse of resilience, or escape to a different stability domain, can occur, however, if the system becomes too underconnected during the destabilized phase of this cycle. This can happen if processes of mobilization are not balanced by processes of retention. Since those processes occur dominantly in soil, any exploration of global change must place a high priority on developing a better and more extensive understanding of soil dynamics in relation to the cycles driven by the four ecosystem functions.

#### *Connections*

The previous section addressed the question posed in the introduction concerning the capacity of ecosystems to absorb, buffer, or generate change. It concentrated on the processes and functions that lead to cycles of ecosystem growth, disruption, and renewal. The periods and amplitudes of those cycles are defined by qualitatively distinct speeds of a small number of key variables. Their ability to maintain structure and patterns of behavior in the face of disturbance, i.e., their resilience, is determined by the renewal function whose properties are, in part, maintained by pulses of disturbance.

The timing and spatial extent of the pulses emerge from the interaction between external events and an internally generated rhythm of stability/instability. Industrial societies are changing the spatial and temporal patterns of those external events. Spatial impacts are more homogeneous; temporal patterns are accelerated. An understanding of

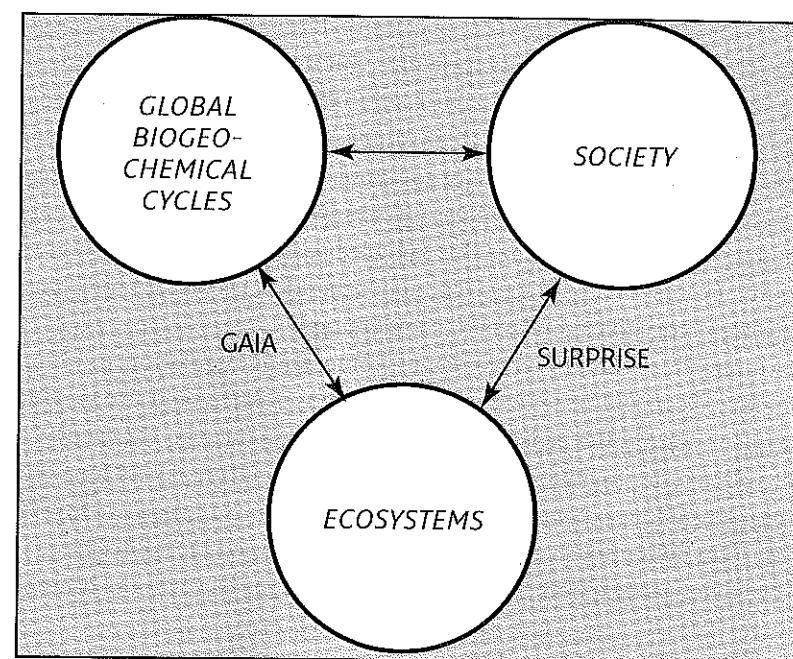


FIGURE 6: *Connections*

impacts of global change therefore requires a framework to connect the understanding developed here for ecosystem dynamics to that developed for global biogeochemical changes on the one hand and societal developments on the other hand. There are transfers of energy, material, and information among all three, as suggested in Figure 6. Biosphere studies now concentrate on changes in the amount, speed, and scale of those transfers and what should or should not be done about them. The Gaia hypothesis, as indicated earlier, provides a focus for discussing the interaction between ecosystems and global biogeochemical cycles. Surprise provides a focus for discussing the interaction between ecosystems and society.

#### *Gaia*

The spatial and temporal patterns generated by the four ecosystem functions form the qualitative structure of ecosystems. A small number of variables and species are fundamental to determining that structure. And the resultant architecture of an ecosystem offers a variety of niches

which are occupied by different species that are affected by the ecosystem structure, but contribute little to it. But while contributing little to the structure, they could contribute significantly to exogenous biogeochemical cycles. This could be determined by drawing on the extensive literature that identifies the major components of geochemical exchange and regulation in plants, animals, and soils. The scheme presented here can provide a way to organize this knowledge so that the interactions of specific ecosystems with external biogeochemical cycles and their possible regulatory roles can be better understood.

Moreover, if the key processes of homeostatic regulation in atmospheric cycles could be demonstrated, such an approach could also provide a way to identify the ecosystems that contribute most to the feedback control. In order to do so, the qualitative analysis outlined here for a few systems could be expanded into a comparative study of the structure of ecosystem dynamics in each of the life zones defined by Holdridge [83] or by Soviet geographers (as reviewed in Grigor'yev [84]) on the basis of climate data.

Such a study would provide a descriptive classification for determining ecosystem responses to global environmental changes. The responses that are most critical are the qualitative patterns of behavior. These patterns are determined by the fast, intermediate, and slow variables during ecosystem growth and disruption and by the mobilization/retention processes during renewal. They also emerge from the way the resultant internal dynamics modify climatic variability. The latter is determined by global atmospheric and oceanic processes, which in turn set the variability in the physical environment.

Steele has reviewed the temporal behavior of physical variables in the ocean and atmosphere [85]. If the well defined periods of days and seasons are removed, the underlying trend for physical conditions in the oceans is for variance to increase as a function of period. The increase is close to the square of the period and occurs at all time (and space) scales. This "red noise" is in contrast to "white" noise where variance is independent of scale (see Dickinson, Chapter 9, pp. 257-260, this volume).

For periods up to about 50 years, physical variation in the atmosphere, unlike in the oceans, is close to being white noise. Thereafter variation seems to follow a red spectrum, suggesting a coupling of atmospheric and oceanic processes. As described earlier, many terrestrial ecosystem

cycles have a period from a few decades to one or two hundred years, driven by the slowest variables. Even if this similarity in cycle periods is a coincidence rather than due to adaptation, changes in the external forcing frequency induced by man's activities could be transmitted and transformed by the existing response times of ecosystems in unexpected ways. Now that the qualitative dynamics of a number of ecosystems are beginning to be better understood, a fruitful area of research can be developed to demonstrate, by example, how changes in the frequency pattern of external forcing can affect ecosystem stability and resilience.

### *Surprise*

Man's efforts to manage ecosystems can be viewed as weak experiments testing a general hypothesis of stability/resilience. In many of the examples discussed earlier, the management goal was to reduce the *variability* of a target variable by applying external controls. Crudely, it represented an equilibrium-centered view of constant nature. All the cases examined were successful in achieving their short-term objectives, but as a consequence of that success, each system evolved into a qualitatively different one.

The evolution took place in three areas. First, the social and economic environment changed. More pulp mills were built to exploit the protected spruce-balsam forests; more recreational demand was developed in the parks protected from fire; more efficient and extensive fisheries were developed to exploit salmon; more land was used for cattle ranches on the savannas; and more development was possible in those areas protected from malaria.

Second, the management agencies began to evolve. Effective agencies were formed to spray insects, fight fires, operate fish hatcheries, encourage cattle ranching, and reduce mosquito populations. And the objective of these agencies naturally shifted from the original socioeconomic objective to one that emphasized operational efficiency: better and better aircraft, navigation, and delivery systems to distribute insecticide; better and better ways to detect fires and control them promptly.

These changes in the socioeconomic environments and in the management institutions were generally perceived and were rightly applauded. But evolution occurred in a third area—the biophysical—whose consequences were not generally perceived.

Because of the initial success in reducing the variability of the target variable, features of the biophysical environment which were implicitly viewed as constants began to change to produce a system that was structurally different and more fragile. Reduction of budworm populations to sustained moderate levels led to accumulation and persistence of foliage over larger and larger areas. Any relaxation of vigilance could then lead to an outbreak in a place where it could spread over enormous areas. Reduction of fire frequency led to accumulation of fuel and the closing of forest crowns so that what were once modest ground fires affecting limited areas and causing minor tree mortality became catastrophic fires covering large areas and causing massive tree mortality. And similarly, increased numbers of salmon led to increases in size and efficiency of fishing fleets and extinction of many native stocks; maintenance of moderate numbers of cattle led to changes in grass composition toward species more vulnerable to drought and errors of management; persistent reduction in mosquitoes led to gradual increases in the number of people susceptible to malaria, and to mosquitoes resistant to insecticide.

In short, the biophysical environment became more fragile and more dependent on vigilance and error-free management at a time when greater dependencies had developed in the socioeconomic and institutional environment. The ecosystems simplified into less resilient ones as a consequence of man's success in reducing variability. In these cases, connectedness increased because of spatial homogenization of key variables: foliage for budworm, fuel and canopy structure for fires, efficient but vulnerable grasses for savannas, numbers of stocks and ages of fish, and the number of people susceptible to malaria.

The hypothesis of constant nature encountered the surprising reality of resilient nature. If control falters, the magnitude and extent of the resultant disruptive phase can be great enough to overwhelm the renewal process.

Just as ecosystems have their own inherent response times, so do societal, economic and institutional systems. How long an inappropriate policy is successful depends on how slowly the ecosystem evolves to the point when the increasing fragility is perceived as a surprise and potential crisis.

The response to such surprises is alarm, denial, or adaptation and is similarly related to the response times of different groups in society and of the management institutions. For example, forest fire policy in the national parks of the western USA has recently changed radically to reinstate fire as the natural "manager" of the forests. This rather dramatic adaptation was not made easily and rested upon the existence of an alternative policy and of technologies to implement it, on a climate of understanding, and on costs that were relatively modest compared with my other examples. But it might be equally important that the critical variables of fuel and forest tree composition changed at the slowest rate of all the examples. It was some sixty years before the change became critical. I argue that the relevant time unit of change for a management institution is of the order of 20 to 30 years, the turnover rate of employees. As a consequence, by the time the problem became critical there was a new generation of experts and policy advisors who would be more willing to recognize failures of their predecessors than of their own. In addition, the slowness of change allowed the accumulation of knowledge of the processes involved and the communication of that growing understanding to a wide range of actors.

In contrast, the changes in the budworm/forest systems proceeded faster. Insecticide spraying began on a large scale in the mid-1950s with conditions of vulnerability building to a critical point by the early 1970s. In this case, the 15–20 year-period was insufficient to accumulate and, most important, disseminate an understanding of the problem. Alternative policies or technologies were not developed and the parents of the original policies were still central actors and defenders of the past. Adaptive change has been an agonizing process and is only now showing signs of occurring [86].

There are insufficient examples to make these remarks anything more than speculation: but they do identify a research priority to determine the time dynamics that lead to increasing dependencies of societies on policies that have succeeded in the past, to examine increasing rigidities of management institutions, and to increase sensitivity to surprise. The research effort should be based on case studies that cover as wide a spectrum of man's activities as possible—economic, technological, and behavioral.



TABLE 2: Possible analogies between ecosystem function and functions or typologies proposed for other systems.

Subject	Function or typology			
	Exploitation (r)	Conservation (K)	Creative destruction	Renewal
Ecosystem				
Economics [e.g., 70]	Innovation, Market, Entrepreneur	Monopolism, Hierarchy, Saturation, Social rigidity	Creative destruction	Invention
Technology [e.g., Brooks 87]	Innovation	Technological monoculture, Technological stalemate	Participatory paralysis	Expert knowledge
Institutions [e.g., 88, 89]	Entrepreneurial market	Caste, Bureaucracy	Sect	Ineffectual
Psychology [e.g., Jung as in 90]	Sensation	Thinking	Intuition	Feeling

Such a comparative study requires collaboration among a number of disciplines. But it is essential to involve practical experience in business, government, and international organizations as well. It is only possible now because so many place priority on understanding change. Equally important, frameworks for understanding change can be found in economics, technology, institutional behavior, and psychology that provide some possible connections to the framework presented here for ecosystems. Examples are suggested in Table 2.

The analogies suggested by Table 2 might simply represent common ways for people to order their ignorance. But there are strong hints, at least from analysis of institutional organizations from the perspectives of cultural anthropology [88, 89] and of technological developments [87], that functions similar to the four ecosystem functions operate in societal settings, although the results can be very different. Some comparative studies already exist that have both predictive and descriptive power. An example is Thompson's analysis of the very different deci-

sions that were made in the UK and California concerning the siting of liquefied gas plants [89]. And regarding technological development, consider Brooks' argument (Chapter 11, this volume) and this quote [87, p. 253]:

One reason for this situation is that, as a particular technology matures, it tends to become more homogeneous and less innovative and adaptive. Its very success tends to freeze it into a mould dictated by the fear of departing from a successful formula, and by massive commitment to capital investments, marketing structures and supporting bureaucracies. During the early stages of a new technology many options and choices are possible, and there are typically many small competing units, each supporting a different variation of the basic technology, and each striving to dominate the field. Gradually one variation begins to win, and the economies of scale in marketing and production then begin to give it a greater and greater competitive edge over rival options. The technical options worth considering become narrower and narrower; research tends to be directed increasingly at marginal product improvements or product differentiation, and the broader consequences of application tend to be taken more and more for granted. Elsewhere I have spoken of this as a "technological monoculture". What happens is that through its very success a new technology and its supporting systems constitute a more and more self-contained social system, unable to adapt to the changes necessitated by its success.

He later adds [87, p 256]:

The paralysis of the decision process by excessive participation will eventually result in a movement to hand the process back to elites with only broad accountability for results according to then current social values. Eventually effects on certain social expectations will become sufficiently serious so that distrust of the experts will revive and there will be a new wave of demands for participation until the frustration of more diffuse social interests will again result in reversion to experts.



This analysis and speculation is completely in harmony with the cyclic processes described here for ecosystems.

Such analogues at the least suggest that a formal comparative study of different cases could help provide an empirical basis to classify the timing of key phases of societal response to the unexpected: in detecting surprise, in understanding the source and cause of surprise, in communicating that understanding, and in responding to surprise. Such a classification can help introduce a better balance between prediction, anticipation, and adaptation to the known, the uncertain, and the unknown features of our changing world.

### Recommendations

Ecosystems have a natural rhythm of change the amplitude and frequency of which is determined by the development of internal processes and structures in a response to past external variabilities. These rhythms alternate periods of increasing organization and stasis with periods of reorganization and renewal. They determine the degree of productivity and resilience of ecosystems.

Modern technological man affects these patterns and their causes in two ways. First, traditional resource-management institutions constrain the rhythms by restricting them temporally and homogenizing them spatially. Internal biophysical relationships then change, leading to systems of increasing fragility, i.e., to a reduced resilience. Moreover, modern man and his institutions operate with a different historical rhythm that can mask indications of slowly increasing fragility and can inhibit effective adaptive responses, resulting in the increased likelihood of internally generated surprises, i.e., crises. Second, the increasing extent and intensity of modern industrial and agricultural activities have modified and accelerated many global atmospheric processes, thereby changing the external variability experienced by ecosystems. This imposes another set of adaptive pressures on ecosystems when they are already subject to local ones. As a consequence, locally generated surprises can be more frequently affected by global phenomena, and in turn can affect these global phenomena in a web of global ecological interdependencies.

We now have detailed examples and analyses of ecological patterns, largely from northern temperate regions, that demonstrate the role of variables of different rates of action and reveal the importance of functions that trigger change and renewal in maintaining resilience. The

resultant synthesis indicates that there is now less of a priority to develop predictive tools than to design systems with enough flexibility to allow recovery and renewal in the face of unexpected events—in short, there needs to be a better balance established between anticipation, monitoring, and adaptation [91].

The design effort would be facilitated by research to test and expand the conclusions in three ways. First, the ecosystem synthesis should be extended to further examples of four critical ecosystems: arctic, arid, humid tropical, and marine, since each has patterns and structures different from northern temperate ecosystems. Second, the analyses of time responses and rhythms of change described here should be extended more explicitly to the links between natural/societal systems, particularly regarding the history of economic, technological, and resource development. Third, there is a need and opportunity to develop a set of well replicated mesoscale experiments in order to reduce the ambiguity of problems occurring because of local surprise and global interconnection. These are given more specifically in the following sections.

#### *A Comparative Study of Resilience and Ecosystem Recovery*

*Purpose:* to define early warning signals of pathologically destructive change and to design self-renewing resource systems.

Data are required to extend the description of time patterns to allow comparison between northern temperate, arctic, arid, humid tropical, and marine systems. Processes that trigger change and facilitate renewal should be identified and classified in terms of their effects on stability, productivity, and resilience. The former requires information as to the role of slow variables in triggering pulses of disturbance. The latter particularly requires a study of soil processes, the balance between nutrient mobilization and retention, their sensitivity to disturbance, and their rates of recovery after small- to large-scale disturbances (e.g., from natural patch formations to man-made land clearance and drainage).

#### *A Comparative Study of Sources and Responses to Surprise in Natural-Social Systems, Particularly Economic Technological, and Resource Development*

*Purpose:* to define conditions that determine how much to invest in action (decide policy and act now), anticipation (delay and find out more), or adaptation (forget the immediate problem and invest in innovation).

The generation of sharp change, its detection, and adaptation of policy responses depend on the interaction between the response times of the managed (natural) system, of the institutions managing the systems, and of the economic and social dependencies that develop.

It now seems possible to classify resource, ecological, and environmental problems not only in terms of uncertainty of their consequences, but also in terms of uncertainty of societal response. Those requiring priority attention are not necessarily those that have the greatest impact, but those likely to generate a pathological policy response. The analysis presented here for ecosystems could be usefully applied to interactions between three components.

One of these components concerns the organization and time dynamics of management institutions. Focus and direction can be given by combining the analysis of surprise with the experience and orientation that has matured in hazards research studies [5] and in institutional analyses from the perspective of cultural anthropology [88, 89]. The second component concerns the geophysicochemistry of the atmosphere and oceans that increasingly connects regional economic development with global ecological interdependency through the ecosystems. Focus and direction can be given by the Gaia hypothesis of Lovelock [6] and the system dynamic studies of Steele [16] which view the atmosphere, oceans, and living systems as an interacting, self-regulated whole. The third and final component is society itself, particularly the historical patterns of economic and technological development that reveal how attitudes are formed, technological monocultures developed, and innovations either inhibited or enhanced. Focus and direction can be given by combining an understanding of ecosystem surprise with historical analyses of change, such as those of McNeill [92].

#### *International Mesoscale Experiments*

*Purpose:* to develop a set of internationally replicated experiments involving areas from a few square kilometers to a few thousand that can test alternative hypotheses developed to explain particular impacts of man's activities and to determine remedial policies.

Our understanding of the structure and behavior of ecosystems, and of how exploitation and pollution affect them, comes from a synthesis of knowledge of ecological, behavioral, physiological, and genetic

processes. Much of that knowledge has been developed from the solid tradition of experimental, quantitative, and reductionist science which can now be generalized and synthesized to propose quantitative structures, qualitative behaviors, and qualitative consequences of impacts. Although synthetic, they essentially represent hypotheses because the arguments are based on studies that could be accommodated in the laboratory or in a few hectares. Ecosystems (as well as people's responses to them) operate on scales of a few square kilometers to several thousand square kilometers. That is where our knowledge and experience is the weakest.

The experiments would be designed to clarify alternative explanations of and policies for problems that emerge from the extension and intensification of industrial and agricultural development. Rather than discussing or investigating these endlessly, it should now be possible to design experiments that distinguish between alternatives. It is essential to concentrate on experiments in which the tests are qualitative in nature, the duration short (less than 5 years by drawing on fast/slow definitions of variables), the spatial scale in the "meso" range, and the policy consequences international. International replication and collaboration then becomes part of the design, which could ultimately contribute to institutional solutions as well as to scientific and policy understanding [93].

#### **Notes and References**

- [1] See, for example, McElroy (Chapter 7) and Dickinson (Chapter 9).
- [2] Steele, J. H. and Henderson, E. W. (1984), Modeling long-term fluctuations in fish stocks, *Science*, 224, 985-987.
- [3] Clark, W. C. and Holling, C. S. (1985), Sustainable development of the biosphere: human activities and global change, in T. Malone and J. Roederer (Eds), *Global Change*, pp. 474-490, Proceedings of a symposium sponsored by the ICSU in Ottawa, Canada (Cambridge University Press, Cambridge).
- [4] Qualitative change, in the sense used here, is structural change; that is, changes in the character of relationships between variables and in the stability of parameters. Such changes challenge traditional approaches to development, as well as to control, which is in itself a qualitative change in the combined ecological-social system.