Journal of Aquatic Ecosystem Health **1**: 1–5, 1992.
M. Munawar (ed.), Assessing Aquatic Ecosystem Health: Rationale, Challenges, and Strategies.

Can ecosystems be healthy? Critical consideration of concepts

P. Calow
Department of Animal and Plant Sciences, University of Sheffield, Sheffield S10 2UQ, U.K.

Keywords: analogy, control, fitness, optimization, teleology

Abstract. Health, it is argued, implies that a system has an optimum state that can be defended. For organisms and populations this can be understood objectively and generally in terms of neo-Darwinian principles. Similar reasoning cannot be applied to ecosystems. The possible advantages and difficulties of applying this concept of health to ecosystems are critically considered.

1. Introduction

The concept of ecosystem health was probably first applied by James Hutton, a Scottish physician and geologist, who in 1788 delivered a paper to the Royal Society of Edinburgh on a theory of the earth as a superorganism capable of self-maintenance (Hutton, 1788; cited in Lovelock, 1988). More recently, the view that aquatic ecosystems can be unhealthy has been advocated notably by Rapport *et al.* (1985) and Schaeffer *et al.* (1988).

That a system has health implies an analogy with the human condition. The aim of drawing analogies is usually to further an understanding of the unfamiliar by reference to features of the familiar. But there are various strengths of analogy. In the present context, health might simply signal normality and hence the converse, ill-health, signals abnormality. I refer to this as the *weak form* of the analogy. Another interpretation though, and one consistent with what Hutton intended, is where 'health' defines a condition favourable (i.e. optimum) for the functioning of the whole organism that is actively defended by homeostatic processes. Ideally this optimum state should also be generalizable between individuals for only then can objective health criteria be defined. I refer to this as the *strong form* of the analogy.

Here, I begin by assessing the implications of the strong form of the analogy before going on to consider to what extent the strong and weak forms of the analogy are indeed applicable to ecosystems and then to their constituent populations.

2. Control systems

From the above, the notion of a controlled 'optimum state' is obviously central in applying the concept of health, in the strong sense of the analogy, to a system. Control is identified by systems remaining unchanged with perturbation (i.e. they resist perturbation), or by their ability to return to previous states after perturbation (i.e. their resilience). But there are two ways that control can be achieved:

(a) by active, feedback (usually negative) control according to a 'goal' state (or states) specified in a programme. This leads to an accepted form of teleology; the system moves towards a future 'goal state' that is programmed into it. This is different to the future controlling the present — a scientifically unacceptable form of teleology (Calow, 1976).
(b) the dynamics of interacting parts simply lead to an equilibrium state. All complex systems have one or more equilibria but these are achieved 'passively' by what systems' engineers often describe as 'fictitious feedback' (Calow, 1976).

Cybernetic machines such as thermostats and robots are programmed. Also *organisms* can be considered to be programmed because they contain molecular systems that code for phenotypes that are capable of replicating molecular programmes to a greater or lesser extent. This depends on survival probability (S), reproductive output (n), and time between reproductive events

(t) — i.e. their neo-Darwinian fitness. Programmes that lead to enhanced S and n and reduced t will become more common. Organisms can therefore be considered to be active control systems, programmed according to neo-Darwinian criteria, and can be understood in terms of the acceptable form of teleology described above (Calow, 1976).

Ecosystems, on the other hand, do not reproduce themselves in any sense equivalent to organisms. Therefore, they cannot be programmed (except in an unacceptable teleological way) and hence cannot be considered as active control systems. Of course, they are made up of programmed systems; and must in some sense be programmed themselves. However, whatever this means, it is unlikely that the way component parts are programmed will be such as to lead to active control for a 'balanced economy' in the ecosystem as a whole. This is because natural selection on individuals and populations will favour programmes that maximize command of resources even if this is at the expense of the rest of the system. The alternative is a group-selectionist interpretation (Dawkins, 1982) and though certain kinds of a group selection now seem possible, these can only occur under what some believe to be very specific conditions (Maynard Smith, 1984). Ecosystems, however, are complex systems, and those that persist will have dynamics that lead to stable equilibria in terms of resistance and resilience (Westman, 1978). This is more akin to passive control. The outcome, in terms of systems' dynamics, is similar to that for active control; but the way that it is achieved is different. In the strong sense of the analogy, the concept of ecosystem health is therefore at best misleading and at worst flawed by a group-selectionist approach, and/or the unacceptable form of teleology.

3. Ecosystem properties and processes

It follows from the last section that there cannot be actively controlled properties and processes that can be used to define healthy ecosystems. There may of course be properties and processes that we want to defend for our own human purposes (e.g. aesthetics, production, recreation). This is imposing active control from outside — something that is often done in terms of defining environmental quality objectives and from these specifying environmental quality standards. Nevertheless, there will be equilibrium states for ecosystems. Can these be used to define a health state in the weak sense of the analogy?

In structural terms, it is obvious that diversity, both with respect to the quality and quantity of taxa, varies with 'natural' environmental conditions (e.g. MacArthur, 1965). To use this effectively as a baseline, therefore, would require rigorous definition of the relationship between diversity and environmental factors. For example, this has been done for British rivers and incorporated into a model called RIVPACS (Armitage *et al.*, 1987; Wright *et al.*, 1989). From an initial survey of clean rivers a model was constructed that related diversity to a few key, environmental factors. This can be used to predict an expected fauna for a stretch of river. Deviations in the observed fauna from expectations give an indication of disturbance. There are three problems with this approach: (a) in defining the initial set of 'clean' sites; (b) deviations from expected are likely to be related to the intensity of stress in a complex way; (c) the models are empirically derived by correlation, so there is no guarantee that the relationship in the models are causal and will therefore apply generally.

On the last point, a firmer theoretical framework is needed. Here it is important to recognise that what theory is available suggests that diversity does not necessarily covary positively with stability (May, 1972, 1976). More diverse communities are more likely to be susceptible to stress, but not straightforwardly because stability also depends upon connectance; more diverse communities are less likely to be susceptible to stress if connectance is low. The product of species diversity and connectance is an important property that may tend towards some constant for stable ecosystems. However, the relationship between species diversity, connectance, and stability is further complicated. Pimm (1979) found in his models that as connectance was increased the proportion of stable communities reduced, but that in the subset that was stable, increasing connectance actually increased resilience. Clearly, more attention needs to be given to species diversity and connectance in attempting to define base states for

stable ecosystems. Moreover, the whole concept of stability may not be straightforwardly applicable to flowing-water ecosystems, where dynamism in physico-chemical conditions and flora and fauna seem to be the rule (Townsend, 1989).

Turning to functional states, it is clear that for any stable open system the input of energy must equal output; which in turn means that production (P) within biological systems must be balanced by respiratory heat loss (R). This might be thrown into imbalance by:

(i) enhanced input (e.g. attendant on organic pollution) that will not only upset the energy dynamics but will lead to changes in the physical state of the system. For example, organic enrichment of littoral habitats (and possibly stream-bed habitats) at first increases microhabitat diversity and hence species diversity, but ultimately progresses to a single microhabitat, a sediment cover, and reduced species diversity (Calow, 1984).

(ii) enhanced output — it is generally assumed that there are metabolic costs to resisting stress (see below). Hence, the respiration per total biomass (B) will increase under these conditions and relative production, i.e. P/R and P/B, will fall.

These process responses have been thoroughly discussed and reviewed by Odum (1985). The extent to which structural and process changes are coupled and hence their potential importance as stress indicators has also been considered in detail elsewhere (e.g. Cairns & Pratt, 1986). However, still much more needs to be done in rigorously defining the links between the two.

Normality, what properties a system should have in the absence of disturbance and interference (in the case of ecosystems, by man), might be defined in terms of lists of properties observed in putatively normal systems. This is certainly what happened in early medicine. Physicians sought to correlate body states with conditions of health and ill-health; but this only became credible and lead to sensible and sound treatment when it was embedded in an understanding of anatomy and physiology etc. Neither is there any guarantee that a compilation of observations on systems states in itself will lead to an understanding of the way the system works (the fallacy of induction). The published lists of ecosystem health criteria (Rapport, 1989a, 1989b) are empirically derived as with early medicine and suffer from similar problems. What we need here is a theory of ecosystem structure and function that goes beyond correlational techniques (cf. RIVPACS) and includes predictions about normal states from an experimentally substantiated understanding of ecosystems. This is as yet rudimentary and requires developments in general ecology.

4. Organisms, optima and active control

It has been suggested above that 'health' of organisms can be objectively defined, in a very general way, in terms of neo-Darwinian fitness. It is presumed that within most populations, fitness will be maximized, so anything that impairs fitness in the dominant genotype will lead to a reduction in the abundance of that population and possibly to extinction. In these simple terms, selection will maximize S (survival) and n (reproductive output) and minimize t (developmental time); and environmental stress can be recognized as impairment of these (Calow & Sibly, 1990). One consequence of this view, is that active control systems evolve within organisms according to neo-Darwinian criteria; 'goal states' will be those that maximize biomass production [hence growth rates (so t is minimized) and reproduction rates maximized] and minimize survival risks.

However, the situation is not as simple as this. There are, for example, good metabolic grounds for expecting negative correlations, trade-offs, between fitness components. So S, n, and somatic growth rate cannot be maximized simultaneously. They will be optimized according to environmental circumstances. These points are summarized in Fig. 1. Here it is presumed that reproductive output (n) and growth are positively related to the production term (P) of the energy budget — which as 'scope for growth' has often been used as a short-term, physiological measure of stress (e.g. Naylor *et al.*, 1989). It is further presumed that S is positively related to respiratory metabolism (R); this is because stress resistance is metabolically costly (Calow, 1991). Hence fitness increases with P and S but, for a given level of fitness, P and S are negatively correlated; i.e. there is a trade-off

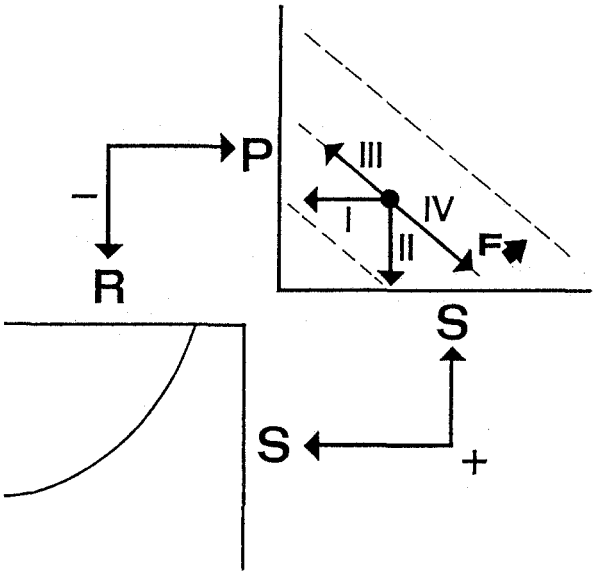

Fig. 1. A very simplified representation of the production/survival space. The point in this space can represent a genotype, an individual, or a population. It can move in the directions indicated by arrows and roman numerals and these are discussed further in the text. Broken lines are supposed to be combinations of P/S that give similar fitness (F) and these isoclines represent increasing fitness in the direction of the arrow. P is related negatively to respiratory metabolism (R). The functional relationship between R and S is hypothetical. For a more detailed argument see Calow & Sibly (1990).

between them. The precise form of these relationships depends upon the precise form of the functional relationships between P and n & t, and S and R (Calow & Sibly, 1990).

An individual, genotype, or population can be located as a point within the P/S space, and we can imagine how environmental factors will cause it to shift. Stress will cause the point to shift to a lower fitness isocline but this might occur by:

(I) predominantly reduced S;
(II) predominantly reduced P.

Type (I) shifts have been referred to as a mortality (μ) stress by Sibly & Calow (1989) — and disturbance by Grime (1979); whereas Type (II) shifts have been referred to as a growth (G) stress by Sibly & Calow — and simply stress by Grime. So stress, i.e. impairment of health, might be usefully classified in terms of these differing symptoms. Note, however, that neither reduced survival nor reduced production in themselves necessarily signal reduced fitness because reduced S might be compensated for by enhanced P (see III in Fig. 1); e.g. as might arise from density-dependent effects. Reduced P might arise due to increased investment in stress resistance, enhancing metabolic costs R (see IV in Fig. 1). The consequent increased S might compensate for the reduced R. Hence reduced scope for growth should be treated cautiously as an indicator of stress.

Not all genotypes within a population need respond in the same way to stress (see Calow & Sibly, 1990). Possible evolutionary outcomes in the context of the tradeoffs illustrated in Fig. 1 have been explored in detail by Sibly & Calow (1989). Tolerance, leading to enhanced survival under stress may, because of metabolic costs, lead to reduced production when the stress is removed (e.g. Cook *et al.*, 1972). Hence, the evolution of tolerance will, amongst other things, depend upon the frequency and duration of stressful conditions. The evolution of tolerance means, of course, that by definition the tolerant forms are not stressed by what were originally stressful conditions. Nevertheless, the presence of tolerant forms can be used as an indication that the community has been subjected to stress (Blanck *et al.*, 1988), and the syndrome of traits that evolve under stress can give insights into the response mechanism deployed by biological systems to environmental stress (Grime, 1979).

5. Conclusions

It is often useful to work by analogy in the development of science (Calow, 1976; Rapport, 1989a, 1989b); but care has to be taken of not by so doing importing fundamentally inappropriate principles from one class of system to another. This was one of the root problems of vitalism and mechanism (Calow, 1976.) I have argued, that the concept of 'health', as applied to organisms, involves different principles from the concept of 'health' as applied to ecosystems. Moreover, currently 'health states' are more easily and objectively definable for organisms, than they are for ecosystems. To make sound predictions about what properties would be expected of undisturbed ecosystems will require progress in fundamental as well as applied ecology.

References

Armitage, P. D., R. J. M. Gunn, M. T. Furse, J. F. Wright & D. Moss, 1987. The use of prediction to assess macro-invertebrate response to river regulation. Hydrobiologia 44: 25—33.

Blanck, H., S. Wangberg & S. Molander, 1988. Pollution-induced community tolerance — a new ecotoxicological tool. In: J. Cairns & J. R. Pratt (eds), Functional Testing of Aquatic Biota for Estimating Hazards of chemicals. Amer. Soc. Testing & Materials, STP 988: 219—230.

Cairns, J. & J. R. Pratt, 1986. On the relation between structural and functional analyses of ecosystems. Environ. Toxicol. Chem. 5: 785—786.

Calow, P., 1976. *Biological Machines. A Cybernetic Approach to Life.* Edward Arnold, London, 134 pp.

Calow, P., 1984. Factors affecting the diversity of littoral invertebrates: a hypothesis. Ing. Sanitaria 30: 41—46.

Calow, P., 1991. Physiological costs of combating chemical toxicants: ecological implications. Comp. Biochem. Physiol. 100C: 3—6.

Calow. P. & R. M. Sibly, 1990. A physiological basis of population processes: ecotoxicological implications. Func. Ecol. 4: 282—288.

Cook, S. C. A., C. Lefebvre & T. McNeilly, 1972. Competition between metal tolerant and normal plant populations on normal soil. Evolution 26: 366—372.

Dawkins, R., 1982. *The Extended Phenotype.* Freeman, London. 307 pp.

Grime, J. P, 1979. *Plant Strategies and Vegetation Processes.* John Wiley, Chichester. 222 pp.

Hutton, J., 1788. Theory of the earth; or an investigation of laws observable in the composition, dissolution and restoration of land upon the globe. Trans. Roy. Soc. Edinburgh 1: 209—304.

Lovelock, J., 1988. *The Ages of Gaia.* Oxford Univ. Press, Oxford. 252 pp.

MacArthur, R. H., 1965. Patterns of species diversity. Biol. Rev. 40: 510—533.

May, R. M., 1972. Will a large complex system be stable? Nature 238: 414.

May, R. M., 1976. Patterns in multi-species communities. In: R. M. May (ed.), *Theoretical Ecology: Principles and Applications*, pp. 142—162. Blackwell Sci. Pub., Oxford.

Maynard Smith, J., 1984. The population as a unit of selection. In: B. Shorrocks (ed.), *Evolutionary Ecology*, pp. 193—202. Blackwell Sci. Pub., Oxford.

Naylor, C., L. Maltby & P. Calow, 1989. Scope for growth in *Gammarus pulex*, a freshwater detritivore. Hydrobiologia 188—189: 517—523.

Odum, E. P., 1985. Trends expected in stressed ecosystems. BioScience 35: 419—422.

Pimm, S. L., 1979. Complexity and stability: another look at MacArthur's original hypothesis. Oikos 33: 351—357.

Rapport, D. J., 1989a. Symptoms of pathology in the Gulf of Bothnia (Baltic Sea): ecosystems response to stress from human activity. Biol. J. Linnean Soc. 37: 33—49.

Rapport, D. J., 1989b. What constitutes ecosystem health? Perspect. Biol. Med. 33: 120—132.

Rapport, D. J., H. A. Regier & T. C. Hutchinson, 1985. Ecosystem behaviour under stress. Amer. Natur. 125: 617—640.

Schaeffer, D. J., E. E. Herricks & H. W. Kerster, 1988. Ecosystem Health: 1. Measuring ecosystem health. Environ. Managem. 12: 445—455.

Sibly, R. M. & P. Calow, 1989. A life-cycle theory of response to stress. Biol. J. Linnean Soc. 37: 101—116.

Townsend, C. R., 1989. The patch dynamics concept of stream community ecology. J. North Amer. Benthol. Soc. 8: 36—50.

Westman, W. E., 1978. Measuring the inertia and resilience of ecosystems. BioScience 28: 705—710.

Wright, J. F., P. D. Armitage & M. T. Furse, 1989. Prediction of invertebrate communities using stream measurements. Regulated Rivers: Res. Managem. 4: 147—155.