Multiple “stable” states and the conservation of marine ecosystems

Nancy Knowlton *

Center for Marine Biodiversity and Conservation, Scripps Institution of Oceanography, University of California, San Diego, La Jolla, CA 92093-0202, USA

Abstract

Marine ecosystems were among the first to provide potential examples of multiple stable states. However, remarkably few of these have been explored in detail, and none have been rigorously confirmed. This may be because differences between alternative states are too subtle to document in the context of regular disturbance, because one state is naturally far more likely to occur than any other, or because most environments naturally support only one type of stable community. It is also possible that the temporal and spatial criteria required to document alternate stable states rigorously may be too difficult to meet in most circumstances. Nevertheless, the possibility of alternative stable states has recently received renewed attention in the context of marine conservation biology. People may be widening the range of habitats in which alternate stable states are possible, or they may be shifting communities to new domains of attraction that rarely occur in the absence of massive anthropogenic perturbations. The ability of people now to alter ocean ecosystems on global scales may eliminate “edge effects” that might otherwise rescue perturbed communities. Ecosystems with alternate stable states are characterized by positive feedback mechanisms that stabilize transitions; even if return to original conditions is predicted (that is, the alternate states are not stable), the same mechanisms will retard recovery. This may explain in part why return to original conditions following anthropogenic disturbance is slower than expected. Slow recovery times and transitions to new states are both potentially costly to human societies. Thus, from a conservation perspective the indefinite persistence of an alternate state may be less important than the presence of feedback loops that slow recovery. Both possibilities reinforce the arguments for application of the precautionary principle in managing marine ecosystems.

© 2004 Elsevier Ltd. All rights reserved.

Contents

1. History of the concept ................................................................. 388

2. Possible examples of multiple “stable” states in marine ecosystems ................. 389

* Corresponding author. Tel.: +1-858-822-2486; fax: +1-858-822-1267.
E-mail address: nknowlton@ucsd.edu (N. Knowlton).
Modern interest in the concept of multiple stable states began with Lewontin’s (1969) simple question: “Can there be more than one stable community in a given habitat?” This question was posed as part of the search for an “exact theory” of “the evolution of communities of organisms”. In that context, multiple stable states were of intrinsic importance because they implied that natural ecosystems were unpredictable in the absence of historical information. Subsequent theoretical work has confirmed Lewontin’s initial suggestion that multiple stable states are plausible and indeed might be common features of communities (e.g., Case & Casten, 1979; Gilpin & Case, 1976; Law & Morton, 1993; May, 1977; Noy-Meir, 1975).

Sutherland’s (1974) study of communities of fouling organisms was the first substantial application of the idea of multiple stable states in marine community ecology. He argued that adult organisms that occurred on settling plates could prevent the settlement of later-arriving larvae, and thus community composition at any point in time was dictated by history – by what larvae were present when space was available. With the widespread acceptance of the importance of disturbance in natural communities (and the implied irrelevance of equilibrium-based theory) in the 1980s, Sutherland’s study belatedly elicited a detailed critique from Connell and Sousa (1983), who argued that stability was rare and the existence of multiple stable states undocumented. They concluded that previously described apparent examples were flawed because: (1) they concerned spatial comparisons of communities existing under different physical conditions, (2) the communities had not been studied long enough to confirm stability, or (3) differences across space or time were not stable in the absence of continuing anthropogenic perturbations.

Connell and Sousa’s (1983) critique prompted more discussion in the literature (Peterson, 1984; Sousa & Connell, 1985; Sutherland, 1990) about what did and did not constitute satisfactory evidence for multiple stable states. From these exchanges a clearer set of criteria emerged. In particular, (1) persistence of alternate states over multiple generations (i.e., the existence of mechanisms that prevent return to the previous state) and (2) persistence of alternate states following the cessation of the perturbations (typically human-derived) that caused the initial changes were generally accepted as prerequisites for documenting their stability (but see Sutherland, 1990, and discussion below). In this context, several previously suggested examples fail to meet the criteria of multiple stable states. For example, the alternate fouling communities described by Sutherland (1974) lack mechanisms to maintain themselves after the death of the initial settlers, and thus are only stable for a single generation. The alternate stable states of kelps and sea urchin barrens described by Simenstad, Estes, and Kenyon (1978) probably also fail to qualify as a non-anthropogenically maintained example, as the barrens only persist with human hunting of sea otters. The validity
of Peterson’s (1984) point, that lack of identity of physical conditions between two areas with different communities would not rule out the existence of alternate stable states if the organisms themselves were responsible for the different conditions, has also been accepted. Finally, the existence of pervasive disturbance in all marine ecosystems suggests that the concept might be better termed multiple or alternate attractors or regimes rather than stable states (Scheffer & Carpenter, 2003); here I use quotations to indicate the same idea.

2. Possible examples of multiple “stable” states in marine ecosystems

The tighter criteria that have evolved for recognizing multiple “stable” states result in greater conceptual clarity, but pose serious challenges for rigorous experimental demonstration (particularly for ecosystems like coral reefs where the major players are too long lived to be studied over several generations). Below I outline the features of some marine ecosystems that are consistent with the idea of multiple “stable” states, although in every case questions remain.

2.1. Coral reefs

The failure of coral reefs to recover from disturbances over the last several decades has led a number of authors to suggest that tropical hard bottom environments can exist in two alternate states, one dominated by corals and the other dominated by seaweeds (Done, 1992; Hughes, 1994; Hatcher, 1984; Knowlton, 1992; Ostrander, Armstrong, Knobbe, Gerace, & Scully, 2000). In the case of the Caribbean, two different phenomena have been interpreted in this context – the replacement of corals by macro-algae following disturbances such as storm damage (Knowlton, 1992; Ostrander et al., 2000) and the failure of the sea urchin Diadema antillarum to recover from catastrophic mortality across most of its range (typically >95%) associated with disease in the early 1980s (Lessios, 1988, 1995).

Although Diadema may now be recovering (Edmunds & Carpenter, 2001), the length and pattern of the recovery process (over two decades, with little evidence of recovery until very recently) was unexpected for several reasons. Diadema releases one million eggs per month (Lessios, 1995), and size-frequency patterns prior to the die-off suggested that recruitment used to occur regularly (Lessios, 1988). Moreover, one of its competitors (the sea urchin Echinometra viridis) facilitates recruitment by Diadema, and the pathogen has by and large not returned (Lessios, 1995).

Together these data suggest that there is a threshold density of Diadema, below which population growth is negative – a classic example of an Allee effect (Courchamp, Clutton-Brock, & Grenfell, 1999). Allee effects can be caused in a variety of ways, but in free-spawning organisms like Diadema, they are most likely due to fertilization failure when males and females are very widely separated. This interpretation is supported in the case of Diadema by the fact that recruitment during the course of the epidemic only stopped once upstream populations were decimated (Lessios, 1988).

Single species population dynamics with Allee effects are the most basic type of dynamic leading to alternate stable states (Knowlton, 1992). However, simple Allee models would predict either collapse to extinction or steady recovery, neither of which occurred in the case of Diadema. Interestingly, however, more complex models do allow for persistence at both low and high densities with assumptions that are not unreasonable for this situation (Collie, Richardson, & Steele, 2004). In this context, prolonged persistence at low levels followed by apparent localized recovery (Edmunds & Carpenter, 2001) could reflect a recent chance occurrence of factors that was favorable to recruitment, leading to a transition between states [as Lessios (1995) predicted].

Alternatively, the dynamics observed over the last two decades may have nothing to do with alternative stable states and transitions between them, but rather reflect a steady recovery process that was so slow in
its initial stages that it was unmeasurable. Positive feedback loops [e.g., inhibition of *Diadema* settlement by the macroalgae that became established following the sea urchin’s collapse (Bak, 1985)] could play important roles in either case, so that simply documenting the existence of a positive feedback loop is not in and of itself adequate to demonstrate the existence of alternative stable states (Scheffer, Carpenter, Foley, Folke, & Walker, 2001). We may never have the data necessary to distinguish rigorously between these two possibilities.

Transitions from coral to algal domination are not limited to reefs affected by the die-off of *Diadema*, however. Indeed the first application of the concept of alternative stable states to reefs concerned the Great Barrier Reef (Hatcher, 1984), an area with relatively intact populations of herbivores (Pandolfi et al., 2003). Hatcher described a ship grounding that resulted in the long-term establishment of an unpalatable alga. Because herbivorous fish avoid some types of long-lived macroalgae once they are large enough to be recognized, a sufficiently massive disturbance could in principle outstrip the ability of herbivorous fish to keep unpalatable algae from becoming established, resulting in an alternate, algal-dominated state (Hatcher, 1984; Knowlton, 1992). As Hatcher (1984) noted, this positive feedback loop leading to the dominance of macroalgae could either retard recovery or result in indefinite persistence.

Interactions between corals and their predators have also been inferred as agents responsible for unexpected delays in coral recovery. Catastrophic loss of staghorn coral following Hurricane Allen in 1980 in Jamaica resulted in the concentration of corallivores on the surviving remnants of this highly preferred species, leading to further declines (Knowlton, 1992; Knowlton, Lang, & Keller, 1990). Alternate stable states of high and low staghorn coral abundance would be expected if corallivores were able persist on non-preferred prey when staghorn was rare but were limited by other factors when staghorn was abundant. Unfortunately, the subsequent die-off of *Diadema* swamped the effects of corallivores on coral growth, making further study of this predator–prey interaction impossible.

### 2.2. Rocky intertidal

In the rocky intertidal of the western north Atlantic, shorelines may be dominated by either the seaweed *Ascophyllum nodosum* or the mussel *Mytilus edulis*. This pattern has been interpreted as either an example of alternate stable states (Dudgeon & Petraitis, 2001; Petraitis & Dudgeon, 1999; Petraitis & Latham, 1999), or the result of differences in the extent of consumer pressure in physically distinct environments (Bertness, Trussell, Ewanchuk, & Silliman, 2002; Bertness, Trussell, Ewanchuk, Silliman, & Mullan, in press). Although all these papers concern Gulf of Maine communities, Petraitis and colleagues studied sheltered rocky bays, whereas Bertness and colleagues studied tidal rivers and the open coast. In the studies by Bertness and colleagues, mussels dominated high-flow regions of both habitat types and *Ascophyllum* dominated low-flow regions. This predictable association between community type and physical environment is contrary to the expectations for alternate stable states and hence experimental results indicating deterministic community composition are not surprising. In contrast, in the sheltered bays studied by Petraitis and Dudgeon (1999), there were no obvious habitat differences associated with *Ascophyllum* versus mussel dominance. This suggests one possible explanation for the conflicting results of the two sets of studies – that the sheltered rocky bays represent areas of intermediate water flow where either type of community can persist once established.

Petraitis and colleagues also emphasized the importance of scale in understanding the stability of alternate states and transitions between them (as did Knowlton, 1992). Edge effects might drive the community back to its original state when disturbances are small but not when they are large. For example, when large clearings are made in *Ascophyllum* canopies, recruitment by *Ascophyllum* could be reduced because of the limited dispersal ability of its germlings and their vulnerability to dessication and water movement in large clearings, and recruitment by mussels could be enhanced because of the absence of physical disturbance by *Ascophyllum* fronds and of predation by organisms that *Ascophyllum* typically
shelters. As discussed above for coral reefs, such positive feedback loops could either slow recovery of disturbed *Ascophyllum* beds or prevent recovery indefinitely.

A different set of alternative communities was described by Barry (1988) on the Pacific coast of southern California – a higher zone dominated by chitons and a lower zone by macro-algal turf. On steeply sloping coastlines the tidal elevation of the transition between zones was highly predictable, and his experiments showed that the zone where either chitons or turf could dominate was very narrow. On gently sloping coastlines, however, there was a mosaic of patches that he interpreted as evidence of alternate stable states. He suggested that the tidal zone where either chitons or turf could occur was much wider on gently sloping shores, because even at elevations where chitons could dominate by grazing the turf, the accumulation of sediment in the turf prevented them from doing so.

Finally, Ruesink (1998) studied the nature of the predator–prey interaction between hermit crabs and diatoms on low intertidal rocky shores in Washington. She documented a pattern of non-linear interaction strengths between them that could lead to two alternate states – coexistence of diatoms and crabs or the elimination of the diatoms. She argued that generalist consumers not linked to the abundance of specific resources might characterize many ecological systems (see, for example, the discussion of corallivores and staghorn coral above), implying the widespread potential for alternate stable states.

### 2.3. Rocky subtidal

The rocky subtidal habitats of Nova Scotia are either dominated by the kelp *Laminaria* or are heavily grazed by urchins, forming “barrens”. This pattern has been described as an example of alternate stable states by Johnson and Mann (1988). They argued that urchin densities drive a threshold response in the community: at low densities urchins hide in crevices and consume drift algae, whereas at high densities they live openly and destructively graze the kelp beds. Transitions from high to low urchin densities can be caused by disease, whereas the factors responsible for the reverse transition are less clear. Both increases in recruitment success and removal of urchin predators are possibilities; the latter would imply that urchin barrens would not generally occur in the absence of harvesting of urchin predators by people.

Subtidal urchin barrens are a global phenomenon, and the factors responsible for their establishment and maintenance are poorly understood in detail (Sala, Bourdouresque, & Harmelin-Vivien, 1998). However, the bulk of the evidence suggests that hunting or fishing of predators of urchins by people probably makes urchin barrens more likely. At intermediate levels of human impact, however, both barrens and kelp beds appear to be stable, with sharp boundaries suggestive of alternative stable states (Konar & Estes, 2003).

A very different set of alternate stable states was described by Barkai and colleagues (Barkai & Branch, 1987, 1988; Barkai & McQuaid, 1988) for two islands in South Africa. On one island, the subtidal benthic community is dominated by seaweeds, and rock-lobsters are very abundant, consuming a variety of food items including whelks. On another island only 4 km away, rock-lobsters and seaweeds are scarce, and the community is dominated by mussels, whelks and other marine invertebrates. In the latter community lobsters cannot reinvade, as they are immediately attacked and consumed by whelks. Until the 1960s, both islands were dominated by rock-lobsters; the original cause of their disappearance is unclear, although it may be related to pollution or overfishing. Regardless of the original cause, the alternative community state in which lobsters are prey rather than predators appears to be stable.

### 2.4. Soft sediments

Soft sediment communities may be especially likely candidates for alternate stable states, because infaunal organisms act as bioengineers, changing the sediment characteristics so as to inhibit hetero-specific competitors (Peterson, 1984). Possible examples include communities dominated by ghost shrimp vs.
bivalves or phoronids. Ghost shrimp prefer coarser grained sediments on average, but in addition increase the grain of the sediments they colonize by virtue of their active burrowing behavior. Thus in sediments of intermediate graininess, the dominant organism might be determined by the vagaries of the history of settlement.

A similar process on a smaller scale has been described for the presence or absence of diatoms on the surface of tidal flats (Van de Koppel, Herman, Thoolen, & Heip, 2001). Extensive diatom communities are disrupted by erosion of the sediment surface, and hence are absent in high-flow regimes, whereas they are abundant and form biofilms in low-flow regimes. At intermediate flow levels, the extracellular polymeric substances produced by diatoms connect the diatoms, surface grains and other particles, thereby reducing erosion and promoting their persistence.

3. Are multiple “stable” states rare in the ocean?

In freshwater and terrestrial ecosystems, evidence for multiple stable states continues to be cited (e.g., Drake, 1990; Drake et al., 1993; Scheffer, Hosper, Meijer, Moss, & Jeppesen, 1993; Scheffer et al., 2001).

![Diagram illustrating alternate stable states](image)

Fig. 1. Diagrammatic illustrations of why alternate stable points (indicated by asterisks) might be rarely documented. Each square defines possible relative abundances of two species (i.e., species 1 on horizontal axis, species 2 on vertical axis), dotted lines indicate boundaries of basins of attraction, and ovals represent isoclines of standard units of perturbation strength. (a) Few stable points exist, but they are very similar. (b) Many stable points exist so that transitions between them cannot be readily identified. (c) One stable point has a much large basin of attraction than the other. (d) Stronger perturbations are required to move system from one basin of attraction than from the other. (e) Perturbations are more frequent for one direction of transition than for the other. (f) Environmental conditions supporting alternate stable states (thicker portion of line in center of gradient) are relatively uncommon.
Nevertheless, it remains the case that relatively few studies have even attempted to document the importance of multiple stable states in marine ecosystems, despite heated early debates in the literature stimulated by Sutherland’s (1974) paper. Evidence for most of the proposed examples remains inferential (Bertness et al., 2002).

There are several possible biological reasons for the scarcity of clear examples of alternate stable states. These can be illustrated in the abstract by considering dynamics that might influence the abundance of two species in a model ecosystem. First, multiple stable states may be common but difficult to detect. This would be the case if multiple stable states did not differ to an enormous degree in the dominance of conspicuous organisms (Fig. 1(a)), or if there were so many multiple stable states that transitions among them resembled the effects of disturbance (Fig. 1(b)). Second, multiple stable states may be common, but for any general type of environment, one of the possible states may be much more common than the others. This could be the result of one basin of attraction being much larger than the others (Fig. 1(c)) or having much steeper sides (Fig. 1(d)) making transitions less likely. Alternatively, the basins themselves could be comparable but the probability of events pushing communities from one basin to the other could differ, such that most of the time communities were in only one of the multiple possible states (Fig. 1(e)). Third, multiple stable states may be intrinsically rare, because across physical gradients, only a small proportion of conditions support more than one state (Fig. 1(f)). Of course, natural ecosystems are far more complex than anything in Fig. 1. This, coupled with the stringent criteria (Connell & Sousa, 1983) necessary for documenting alternate stable states, may discourage scientists from even attempting to look for them.

Nevertheless, the examples summarized above are compatible with several of the scenarios outlined in Fig. 1. Studies across gradients in flow regimes, sediment characteristics and tidal exposure suggest that relatively narrow zones support alternate stable states, with much wider areas supporting a single community type (Fig. 1(f)). Studies of coral reefs and kelp beds suggest their basins of attractions are naturally broader (Fig. 1(c)) than are the basins for their alternatives, macro-algae and urchin barrens, respectively. It is also clear that certain types of transitions (e.g., geographically widespread, catastrophic mortality events) might be naturally less common than others (e.g., comparably widespread pulses in recruitment) (Fig. 1(e)). What is certainly true, however, is that people are increasingly changing many of the relevant parameters, suggesting that alternate stable states may be more important now than they were before.

4. Multiple “stable” states and marine conservation biology

The role of people was largely ignored by theoretical ecologists of the 1960s and 1970s, and Lewontin’s (1969) only discussion of people was the brief mention that “Man...stabilizes unstable species composition by making non-linear adjustments in the species abundances.” Today the impact of people on marine ecosystems is widely recognized (Jackson et al., 2001), and Lewontin’s concept, while technically correct, is expressed rather differently now.

Nevertheless, work on multiple stable states, particularly in marine habitats, has remained largely rooted in the traditions from which it emerged – that of trying to understand “natural” communities. Most of the “flawed” examples discussed by Connell and Sousa (1983) involved communities perturbed by people, the implication being that were such perturbations to be eliminated, the communities would relatively rapidly return to their original state (or rather states, with natural disturbance playing the dominant role).

There has also been a strong focus on documenting the potentially indefinite persistence of alternative states, whereas in the context of marine conservation, there may be little difference between very slow recovery (relative to human timescales) and indefinite persistence of an alternative stable state. In this context, feedback loops that inhibit recovery may be as important as the existence of alternative stable states in the strictest sense. Indeed, a major departure from prevailing conditions that has a very long return
time will often have greater ecological and economic implications than a transition between two alternative stable states that differ in relatively trivial ways.

There is certainly much evidence that humans can have enormous impacts on marine ecosystems, and that recovery from anthropogenic disturbance is often much slower than might be expected based on a simple consideration of life history characteristics of the affected organisms. Hutchings (2000, 2001), for example, documented that when populations of marine fish collapse due to the effects of overfishing, they do not typically recover rapidly. Moreover, this is true even when fishing mortality is strongly reduced and when the fish in question have high fecundities. Similarly, Caribbean queen conch have been slow to recover even when fisheries have been closed (Stoner & Ray-Culp, 2000). Allee effects have been suggested in the case of the conch, but for many marine fishes (e.g., cod), the reasons remain obscure. The possibility that entire food-webs have been altered, such that top predators have become prey (sensu Barkai & McQuaid, 1988), is of particular concern given the implications for recovery. Moreover, the restoration of three-dimensional spatial structure, which collapses later in the trajectory of ecosystem decline (Pandolfi et al., 2003), is likely to be even slower (e.g., Precht, Aronson, & Swanson, 2001).

Recent studies also suggest that the activities of people may make alternate stable states more likely (Scheffer et al., 2001). For example, overfishing appears to increase the probability of a transition from coral to algal dominated seascapes (Hughes, 1994; Knowlton, 1992), and hunting of otters allows transitions from kelp-dominated environments to urchin “barrens” (Konar & Estes, 2003). Similarly, one might expect that the severe kinds of catastrophic mortality events required to invert trophic relationships, such as those documented in the studies of Barkai and colleagues described above, might be extremely rare under natural conditions but not unexpected in the context of severe overfishing or environmental contamination.

Finally, people are now capable of modifying the structure of marine ecosystems at a global scale (Myers & Worm, 2003). In doing so, they eliminate the potential for edge effects emanating from unimpacted ecosystems to eventually restore damaged ecosystems. The clarity of evidence for alternate stable states in lakes (Scheffer et al., 1993) is in part due to the fact that lakes are small enough to be altered in their entirety. Humans are well on their way to impacting oceans as if they were merely a very large lake, with implications for alternate stable states that are very worrying.

In sum, multiple stable states, and especially the non-linearities and positive feedback loops that underpin them, have important implications for marine conservation biology. They provide strong support for the precautionary principle, because communities often respond to stress in unexpected ways (Paine, Tegner, & Johnson, 1998). Moreover, populations and ecosystems that collapse may be very slow to recover or require active (and expensive) intervention (Scheffer et al., 2001), regardless of whether the system has alternative stable states. For this reason, preserving the resilience of ecosystems (e.g., Nystrom & Folke, 2001) should be an essential component of all conservation strategies, since the disturbances that provoke departures from stable states or transitions between states cannot always be prevented.

5. A paleontological perspective on human impacts

Although direct observations of marine ecosystems are relatively recent, the history of many marine organisms and communities can be examined via the fossil record. Studies of coral reefs, for example, show that although small-scale ecological observations suggest a lack of stability (as emphasized by Connell & Sousa, 1983), much greater stability is seen when reefs are examined over large spatial and temporal scales (Pandolfi, 2002). The fact that reefs repeatedly reassembled in similar form following massive sea level fluctuations in the Pleistocene makes the scale of recent anthropogenic changes all the more remarkable.

In the proverbial global scheme of things, the changes we have witnessed in the last several decades to centuries, severe as they are, pale by comparison with major upheavals in the history of life. The five largest mass extinctions each removed more than 50% of the abundant marine invertebrate genera (Jablonski,
2001). In the case of the end-Permian extinction, over 90% of all forms of marine life went extinct. Moreover, recovery from such major upheavals was exceedingly slow, taking tens of millions of years, with many of the survivors continuing to decline (Jablonski, 2002). Whether the impact of people can measure up to the impact of an asteroid fortunately remains to be seen, but given the dominance of humans in most of the major biogeochemical cycles of the planet, it is not a possibility that can be dismissed out of hand.

References


