PHILOSOPHICAL TRANSACTIONS B

rstb.royalsocietypublishing.org

Research

CrossMark click for updates

Cite this article: Kelly RP, Erickson AL, Mease LA, Battista W, Kittinger JN, Fujita R. 2015 Embracing thresholds for better environmental management. *Phil. Trans. R. Soc. B* **370**: 20130276.
http://dx.doi.org/10.1098/rstb.2013.0276

One contribution of 16 to a Theme Issue 'Marine regime shifts around the globe: theory, drivers and impacts'.

Subject Areas:
environmental science

Keywords:
tipping points, environmental state change, ecosystem management, nonlinear

Author for correspondence:
Ryan P. Kelly
e-mail: rpkelly@uw.edu

Electronic supplementary material is available at http://dx.doi.org/10.1098/rstb.2013.0276 or via http://rstb.royalsocietypublishing.org.

THE ROYAL SOCIETY PUBLISHING

Embracing thresholds for better environmental management

Ryan P. Kelly[1,2], Ashley L. Erickson[2], Lindley A. Mease[2], Willow Battista[1], John N. Kittinger[2] and Rod Fujita[3]

[1]School of Marine and Environmental Affairs, University of Washington, Seattle, WA 98105, USA
[2]Center for Ocean Solutions, Stanford University, Stanford, CA 93940, USA
[3]Environmental Defense Fund, San Francisco, CA 94105, USA

Three decades of study have revealed dozens of examples in which natural systems have crossed biophysical thresholds ('tipping points')—nonlinear changes in ecosystem structure and function—as a result of human-induced stressors, dramatically altering ecosystem function and services. Environmental management that avoids such thresholds could prevent severe social, economic and environmental impacts. Here, we review management measures implemented in ecological systems that have thresholds. Using Ostrom's social–ecological systems framework, we analysed key biophysical and institutional factors associated with 51 social–ecological systems and associated management regimes, and related these to management success defined by ecological outcomes. We categorized cases as instances of prospective or retrospective management, based upon whether management aimed to avoid a threshold or to restore systems that have crossed a threshold. We find that smaller systems are more amenable to threshold-based management, that routine monitoring is associated with successful avoidance of thresholds and recovery after thresholds have been crossed, and that success is associated with the explicit threshold-based management. These findings are powerful evidence for the policy relevance of information on ecological thresholds across a wide range of ecosystems.

1. Introduction

Nonlinear relationships between environmental variables occur in a wide range of ecosystems and result in biophysical thresholds ('tipping points') that can have dramatic social and ecological effects [1–4]. Threshold responses to anthropogenic stressors are thus good bases for environmental decision-making because they provide opportunities for managers to set non-arbitrary targets and to maximize ecosystem return on management investment [5–7].

In the past decade, thresholds have become increasingly relevant in the context of environmental management as stressor–response relationships have become better understood [8–11]. As a consequence, there is increased interest in identifying thresholds prospectively (i.e. before they are crossed) to inform decision-making [3,4,12,13]. Prospective management of thresholds requires managers to both recognize that the system is approaching a threshold and act in time to avert that threshold. Such action can be more or less precautionary, depending on the amount of information available and the degree of risk tolerance, with potentially significant social and economic benefits of avoiding the threshold.

Nevertheless, until recently, the primary way in which management has incorporated thresholds into permitting, planning or other regulatory decisions has been retrospective, i.e. only after the threshold is crossed does its existence become relevant to policy [14–16]. This limitation may be attributed to a number of factors, including (i) a legal mandate that is triggered only after a threshold is crossed (e.g. overfishing of managed stocks); (ii) managers not knowing that a threshold exists until it is crossed (e.g. eutrophication of many lakes); or (iii) social, economic and policy influences—including overcapitalization inhibiting management actions that would strand capital—that prevent management action in time to avoid the threshold (e.g. open-access conditions incentivize overharvesting) [17–19]. Retrospective management of thresholds—trying to recover an ecosystem state or set of services—can be challenging and costly, but it nevertheless remains an important management option in cases where ecosystem change is reversible.

Here, we evaluate the success of management measures implemented in systems with well-characterized biophysical thresholds. For the purposes of this study, we define a threshold as a nonlinear change in ecosystem state, property or phenomenon, where changes in an environmental driver produce disproportionately large responses in the ecosystem (figure 1) [5,8,20]. While thresholds are sometimes conceived of as time-dependent, wherein systems approach a threshold and cross over it as stressors increase with time, our usage is time-independent, depending only upon the relationship between input and outcome variables. We surveyed a diversity of systems including instances of lake and estuary eutrophication, single-species management using logistic growth curves (e.g. maximum sustainable yield), critical loads of airborne pollutants, dynamics between fire and forest cover, and others (electronic supplementary material, table S1). Our objectives in this paper are to (i) describe how ecological systems with threshold dynamics have been managed; (ii) determine whether management that is explicitly aimed at avoiding or reversing biophysical thresholds performs better than management that does not; (iii) identify social, institutional and biophysical factors related to successful management of ecosystems that have thresholds; and (iv) based on these findings, develop recommendations for making management more effective with respect to avoiding thresholds or recovering systems that have crossed thresholds.

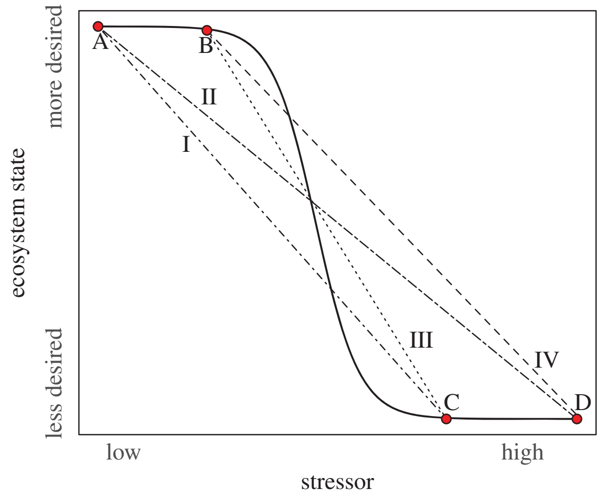

Figure 1. Nonlinear (solid) and linear (broken) curves describing a schematic view of a set of relationships between ecosystem state and the intensity of an anthropogenic stressor. In threshold cases, a small change in stressor intensity can drive a dramatic change in ecosystem state (e.g. from point 'B' to 'C' along the solid curve). Where a management or policy process assumes a linear relationship between stressor and state, the difference between the expected (linear) and observed (nonlinear) slopes creates different policy incentives depending upon starting and ending points (A, B and C, D), and upon the slope of the expected relationship (I, II, III, IV) versus the actual relationship (solid nonlinear curve). (Online version in colour.)

2. Methods

We employ a broad definition of 'threshold' that embraces not just the complex shifts among suites of ecosystem interactions [3,5], but also simpler nonlinear relationships between a driver and some metric of ecosystem state. Consequently, we include examples of single-species management (whose viability or reproductive capacity decline precipitously past a density threshold) in our dataset, because such management is common and represents an important source of data on the outcomes associated with management of threshold-based systems.

We use the term 'management' to mean a conscious effort to change the trajectory of some environmental state or outcome variable (i.e. dependent variable; such as net primary productivity in a eutrophying estuary) by altering one or more input variables (i.e. independent variables such as anthropogenic nitrogen input suspected of driving the eutrophication) that influence the ecosystem of interest. Management thus often aims to change the input variable in order to influence environmental state, but may also aim to change a third variable (e.g. phosphate input into the estuary) in order to reach the desired environmental state. Using indicators for drivers and states of the ecosystems in each of the cases we analysed, we were able to compare management attributes and environmental outcomes across a wide variety of ecosystems.

(a) Case-study selection

We limited our review to cases for which sufficient information was available to describe both an environmental threshold and the relevant management history. We collected ecological, social and economic information primarily from peer-reviewed literature, supplemented with grey literature, books and first-person accounts (electronic supplementary material, table S1). We excluded potential cases for which a management outcome was not identifiable, for which the drivers of the threshold change were substantially out of the control of the managers (e.g. El Niño Southern Oscillation effects on marine systems), or where spatial and temporal scales of management and environmental data did not align. These filtering criteria resulted in a dataset of 51 cases.

We corrected for several potential sources of non-independence among cases. For example, where a variety of fish stocks were managed under a common strategy such as the United States' Magnuson-Stevens Act, we assessed the controlling law or policy itself, rather than management of any individual fish stock. Analysing the highest hierarchical level of management provided a rational means of selecting from among potentially relevant cases, although it required that we average the management attribute scores across available instances of a common management practice. Species simultaneously managed under multiple regimes, such as whales, were another source of non-independence. We dealt with this by including whale species management only under the International Whaling Commission (IWC) case study, excluding whales from species averages in the US Endangered Species Act and Marine Mammal Protection Act case studies. Finally, many parallel cases were available to us—for example, eutrophication thresholds have been documented in lakes and estuarine systems worldwide—and in order to ensure adequate sampling of the different types of thresholds with management implications, we attempted to include cases from a variety of ecosystems rather than incorporating many duplicative cases. We selected cases from duplicative system types using the best-documented examples.

(b) Case-study categorization

We categorized each instance of environmental decision-making as either prospective or retrospective mitigation. We defined prospective cases as those in which a known or suspected threshold had not been crossed (the threshold being known through small-scale experiments, comparison with analogous systems, modelling, past observations or estimates), thereby giving managers the opportunity to take action to prevent threshold crossing. A retrospective case study was one in which parties managed a system that has already crossed the threshold. In general, both retrospective and prospective cases represent management efforts to preserve or improve the existing ecosystem state.

(c) Input variables

We used Ostrom's social–ecological systems governance framework [21,22] to collect and analyse important social, institutional and ecological attributes of our case studies. This framework is frequently used in similar comparative governance analyses [23–25], and identifies common system characteristics across different social–ecological contexts [25,26].

We collected information on 18 system attributes derived from Ostrom's second-tier variables, including information on resource and governance systems, users and social, economic and political settings (electronic supplementary material, table S4). Consistent with other applications of this framework [24,25], we selected the variables based on their relevance to our research objectives, and developed a coding scheme for each variable set. Electronic supplementary material, table S4 shows the complete set of input variables, and notes the attribute modifications that were necessary to apply Ostrom's framework to our dataset.

Where possible, we framed input variables to be explicitly threshold-relevant, intending to examine the attributes most important to threshold-based systems specifically, as opposed to environmental management generally (table 1; 18 input attributes, with seven (in grey) examined through an explicit threshold lens).

(d) Outcome score

Every ecological system has a different composition, structure and function, and our sample of cases varied by several orders of magnitude in spatial and temporal extent. We therefore created a common ordinal outcome metric in order to compare such divergent ecosystems and management schemes by defining the outcome of each management action relative to its own threshold as follows: 1 = net negative progress with respect to the desired ecological outcome relative to the onset of the management action of interest (e.g. in a retrospective case study, if water quality in a eutrophied estuary continues to degrade after management), 2 = no change in the state of the ecological outcome variable (e.g. an estuary exhibits constant degraded water quality on a timescale relevant to the threshold), 3 = a degree of progress or improvement in the ecological outcome variable (e.g. an estuary exhibits improved water quality in some measurable variable) and 4 = 'desired outcome' achieved relative to the outcome variable threshold (e.g. an estuary shifted back to a pre-existing mesotrophic state).

We scored all outcomes relative to a business-as-usual baseline; consequently, an action that has ameliorated a precipitous decline in ecosystem state received a positive score, even if the ecosystem state is nevertheless still declining. Management outcomes in our analysis therefore depend only on current system state relative to system state at the onset of the focal management action. We note that environmental management is complex: not all change in ecosystem state can be fairly attributed to management, and some ecosystems are more resilient or resistant to change than others [3,27]. Our analysis assumes, however, that the variability in these unmeasured variables is distributed randomly across our focal case studies.

We adopted a simple scoring scheme at the risk of missing important outcome complexities. For example, our scoring system weighed only environmental outcomes and not social outcomes (e.g. improved livelihoods). The metrics we employ here have the advantage of being comparable across a diverse suite of case studies, and in each case, the score assigned is supported by analysis of the available management history. As an internal control on assigning outcome scores, we reviewed the outcome scoring as a group after we independently scored subsets of case studies. Case-study outcomes were uncorrelated with the identity of the initial author scoring them (ANOVA $p = 0.44$ for prospective cases; $p = 0.65$ for retrospective cases). Where insufficient or ambiguous data were available to assess environmental outcomes, we informally contacted experts involved in management or research in the focal ecosystem to verify our understanding of particular outcomes (electronic supplementary material, table S1).

Finally, we used the Intergovernmental Panel on Climate Change rubric for assigning confidence descriptors for each of our outcome scores [28]. We accordingly classified each outcome score as having very high, high, medium, low or very low confidence, based upon the evidence available (multiple, consistent, transparent, independent lines of high-quality evidence) and upon the scientific agreement among sources of evidence (electronic supplementary material, tables S2 and S3). For example, a score determined using multiple sources, robust data on the ecological status of the system with regards to the known ecological threshold and evidence consistent across all cited sources would result in a confidence score of 'very high'. We then analysed the dataset using outcome scores weighted by their confidence descriptor. The weighting scheme was as follows: weighted outcome score = (confidence) × (raw outcome score), where numeric expressions of confidence were as follows: 2 (very high), 1.5 (high), 1 (medium), 0.75 (low), 0.5 (very low), resulting in weighted success scores that varied between 0.5 and 8 (electronic supplementary material, table S1).

(e) Statistical analysis

We performed all statistics using the R software package (v. 3.0.1; [29]), treating prospective and retrospective cases separately. We evaluated ordinal logistic regressions of weighted success scores for each of the case-study attributes (R function 'polr' in the MASS package; [30]), judging each likelihood-based model on its parameter estimates and its maximum-likelihood pseudo-r^2 value calculated using the package 'pscl' [31]. We carried out a bootstrap analysis of single-attribute regression results by randomly sampling 25 (of 31) retrospective cases and 15 (of 20) prospective cases, creating 300 subset replicates to evaluate the distribution of pseudo-r^2 values for each logistic regression (electronic supplementary material, figures S1 and S2), to provide an estimate of confidence in our single-attribute regression results. We assessed the independence of case-study attributes using Spearman's rank correlations.

An analysis of statistical power indicated that we could confidently detect an effect that changed the weighted success score by a relatively large margin: 1.55 (retrospective dataset, $n = 31$) or 2.24 (prospective, $n = 20$), representing 19.3% or 30% of the possible range of weighted outcome scores (0.5–8).

We used the R package 'mvpart' (v. 1.6–1; [32]) to carry out recursive multivariate partitioning, generating bifurcating regression trees for each dataset using 1000 iterations for cross-validation. This non-parametric method can identify combinations of attributes conditionally associated with successful outcomes [32], whereas the single-attribute analysis described above reveals information only about one attribute at a time. The two sets of analyses need not agree; for example, the single-attribute logistic regression may identify one attribute as being most strongly associated with greater log-odds of an outcome given the whole dataset, whereas multivariate partitioning may identify a different attribute as most strongly influencing the difference in mean outcome score between two subsets of the data.

3. Results

Attributes were not strongly correlated with one another (median Spearman's $\rho = 0.045$ (prospective), 0.033 (retrospective)), and so were considered to be independent with the following exceptions: within retrospective cases, the size

Table 1. Case-study attributes forming the dataset. Shaded attributes relate explicitly to threshold dynamics. 'Degree of threshold-based management' includes the term 'use of the threshold,' in its definition for categories 1 and 2. 'Use of threshold' can consist of setting limits or targets, incorporating buffers, monitoring leading indicators, or other similar control rules. Electronic supplementary material table S1 contains the full dataset and scoring scheme.

biophysical or management attribute	our definition
discrete/bounded system	the resource system is discrete and bounded (e.g. an estuary versus the population range of the bluefin tuna)
size	the size of the resource system, reported in square kilometres (log transformed for analysis)
quantitative threshold defined	a quantitative threshold has been enumerated/identified that defines the point at which nonlinear change in the resource system occurs
managing primary stressor	the manager of the resource system controls the primary driver of threshold change, an indirect variable that influences the threshold relationship, or both
cost–benefit analysis	a cost–benefit analysis that evaluates the consequences of management action versus inaction has been used. Any documented quantification of monetary costs and benefits of management action qualified as a cost–benefit analysis
leading indicators available	an environmental indicator exists which could be used to identify ahead of time when a management action should be taken in order to avoid or recover from a threshold change. Leading indicators are surrogates for some larger ecological state or process that is undergoing the nonlinear shift. Often indicators are easier or more reliably monitored than the larger ecological state or process. Primary system variables involved in the threshold relationship (i.e. the driver of the change or the variable changing) do not qualify as leading indicators
model type	if a model exists that describes the system and identifies where the threshold occurs, is it quantitative (capturing relationships mathematically) or qualitative (conceptually relating system features)?
degree of threshold-based management: 'threshold-based management'	1. managers have identified a quantitative ecological threshold, which they use to set targets and manage threshold variables; 2. managers have a qualitative idea of the ecological threshold, which they use to set narrative objectives without the use of numerical targets; 3. managers are unaware of, or ignore, the ecological threshold, but their management actions do impact the threshold relationship; 4. managers are unaware of, or ignore, the threshold, and their actions do not impact the threshold relationship
management duration	the time, in years, between the first set of management actions and the most recent report on the state of the ecosystem (i.e. the point at which the authors determined the outcome score)
number of managing entities	the number of entities that are involved in the management of the resource system
jurisdiction over key stressors	the managers of the resource system have jurisdiction over all of the sectors/drivers that influence the threshold relationship
adaptive management	the operational management rules require iterative updating of plans and targets to incorporate new information and changing environmental variables
hierarchical level of governance	the governance level of the primary managing entity: 1. local (e.g. municipality); 2. state; 3. national; 4. international (e.g. international commissions)
notable NGO presence	a non-governmental organization was/is involved in motivating, influencing, or creating accountability for management of the resource system
routine monitoring requirement	the management of the resource system includes routine monitoring of an environmental variable on a timescale relevant to the ecological threshold
larger legal framework present	a written, binding agreement that focuses on the relevant threshold exists. This agreement is separate from any management-level plans or strategy documents, and may be a treaty, regulation, or statute
human development index (HDI)	the human development status of the nation in which the resource system is located. If the system spans multiple nations, the average of these values
binding requirements	the legal framework (written, binding agreement) includes consequences if managers violate the agreement

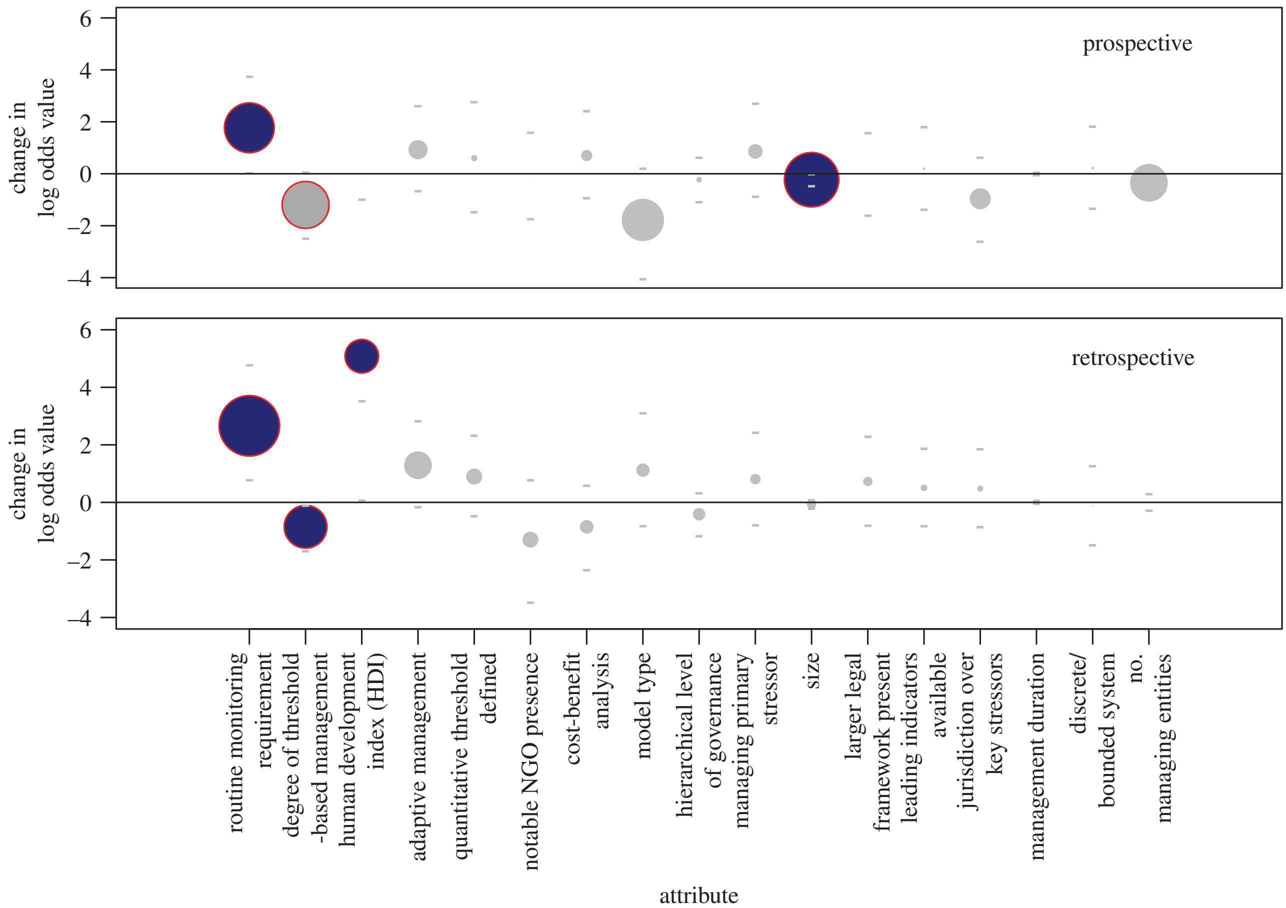

Figure 2. Ordinal logistic regression results for prospective and retrospective case studies. The change in log odds ratio is a measure of effect size and represents the difference between the log probabilities of a high outcome score over a low outcome score, given a 1-unit change in the independent variable. The odds ratio accounts for a marginal increase in the outcome score relative to the number of cases with a given attribute. Coloured circles are significant correlations (i.e. those with a logit parameter estimate whose 95% maximum-likelihood confidence interval does not include zero); 95% CIs are indicated on the chart with grey hash marks. Outlined circles indicate attributes most heavily discussed in the main text. The size of the circles is proportional to the maximum-likelihood estimate of the correlate's pseudo-r^2 value; the largest circle has a pseudo-$r^2 = 0.38$. For a full table of attribute logistic regression results, see electronic supplementary material, table S5. (Online version in colour.)

of the resource system was correlated with hierarchy of management ($\rho = -0.676$) and human development index (HDI) was correlated with hierarchy of management (-0.63). Each of these associations was significant after correction for multiple comparisons.

(a) Prospective cases ($n = 20$)

The most important correlates of outcome scores in prospective management were the use of routine monitoring (pseudo-$r^2 = 0.178$) and the size of the resource system (pseudo-$r^2 = 0.193$; figure 2). Each of these attributes demonstrated a statistically significant effect on outcome score as estimated by 95% CI for the logistic regression parameter (i.e. log-odds). Degree of explicit threshold-based management also accounted for a substantial portion of the variance in outcome scores (pseudo-$r^2 = 0.167$), although this attribute was marginally non-significant (95% CI -2.5 to 0.03; for complete results table, see electronic supplementary material, table S5).

(b) Multivariate partitioning and regression trees

Iterative multivariate partitioning of the prospective dataset into sets with greater and lesser outcome scores resulted in a regression tree having only a single, most-important node corresponding to the geographical size of managed area (explaining 27% of the variance in outcome scores; minimum number of cases = 4; figure 3). Cases with managed areas greater than or equal to 327 750 km^2 (approximately the size of the Great Barrier Reef; ln(size) = 12.7 km^2) had significantly lower outcome scores (mean = 2.4; s.d. = 0.78; $n = 6$) than those with smaller areas (mean = 5.2; s.d. = 1.2; $n = 14$; Wilcoxon $p = 0.019$). Other attributes had significantly weaker effects, excluding them from the single best tree.

(c) Retrospective cases ($n = 31$)

For retrospective management, the strongest correlates of outcome scores were routine monitoring (pseudo-$r^2 = 0.215$), threshold-based management (pseudo-$r^2 = 0.152$) and HDI (pseudo-$r^2 = 0.119$); all were statistically significant effects (i.e. 95% CI did not include zero; figure 2). Most notably, greater degrees of threshold-based management yielded higher outcome scores, highlighting the importance of incorporating explicit, quantitative threshold targets into management decision-making (figure 4).

(d) Multivariate partitioning and regression trees

The best-fit retrospective regression tree identifies the geographical size of the resource system as most strongly

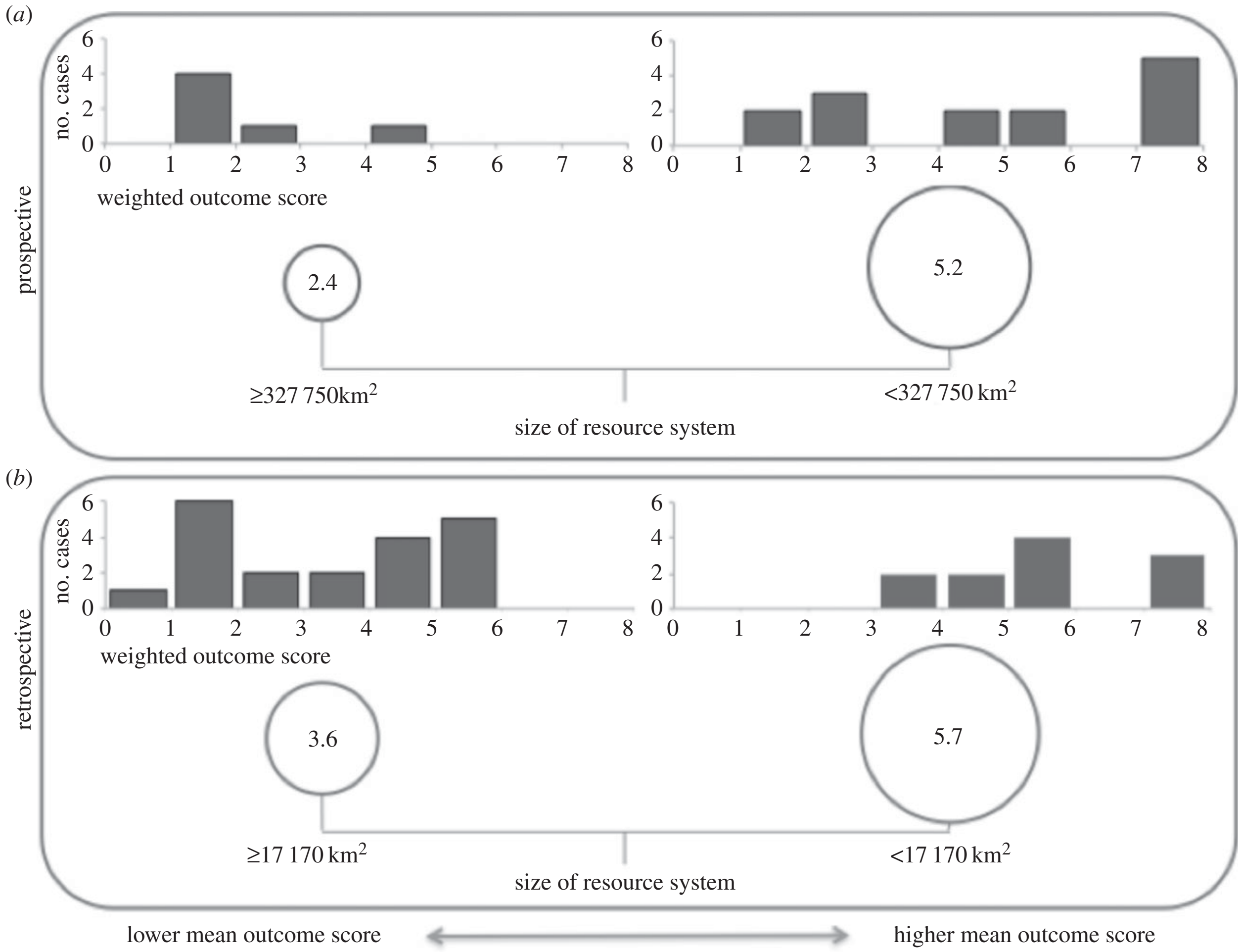

Figure 3. Best-fit regression tree from multivariate partitioning of the retrospective and prospective datasets demonstrates that the geographical size of the resource system is the most influential attribute for case-study outcome score. (*a*) The mean outcome score (bubble) and distribution of outcome scores for prospective cases with areas larger than or less than 327 750 km^2. (*b*) The mean outcome score (bubble) and distribution of outcome scores for retrospective cases with areas larger than or less than 17 170 km^2. The prospective tree explains 27% of the variance in weighted outcome score (error $=$ 0.73; CV error $=$ 1.4; s.e. $=$ 0.32); the retrospective tree explains 24% of the variance in weighted outcome score (error $=$ 0.76; CV error $=$ 1.4; s.e. $=$ 0.25).

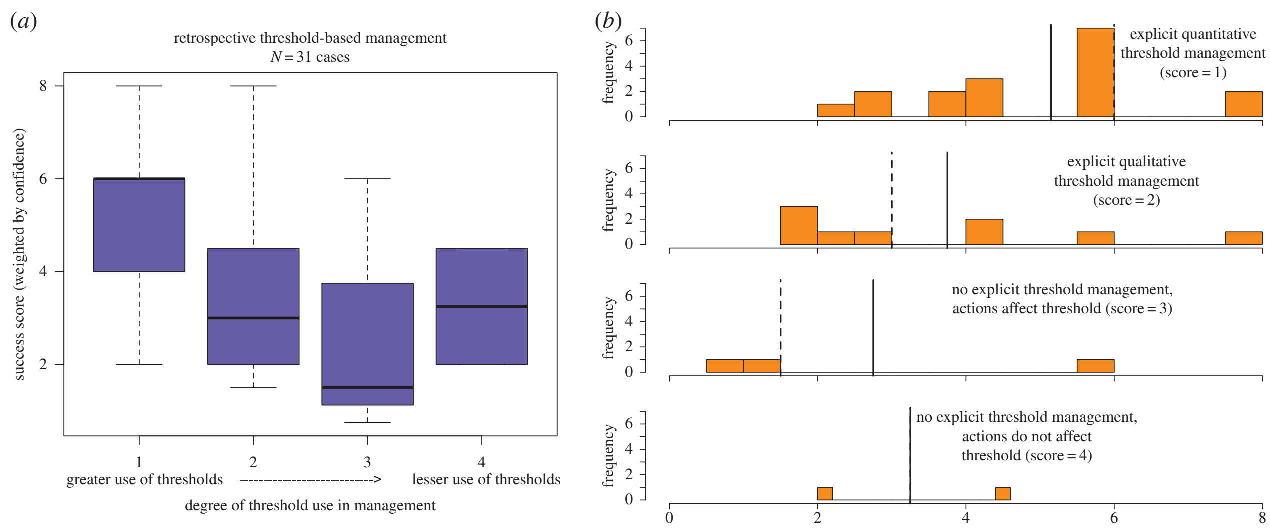

Figure 4. (*a*,*b*) The correlation between degree of threshold-based management and environmental outcomes. Categories of incorporating knowledge of biophysical thresholds into management decisions are as follows, ordered from most explicit to least: 1, managers use a known threshold quantitatively to set targets and manage threshold variables; 2, managers use a known/suspected threshold to manage threshold variables without quantitative targets; 3, managers do not recognize the threshold, but their actions affect threshold variables; 4, managers do not recognize the threshold, and their actions do not affect the threshold variables. Across all 31 retrospective data points, degree of threshold-based management is significantly correlated with outcome score (Spearman's rank correlation, $\rho = -0.42$, $p = 0.016$). (Online version in colour.)

influential (explaining 24% of the variance in outcome scores; minimum number of cases = 3), an identical result to the prospective regression tree. The regression tree contrasts with the retrospective single-variate analysis, in which size is only a minor logistic correlate. Cases with managed areas greater than or equal to 17 170 km^2 (approximately the size of the Neuse River Estuary; ln(size) = 9.75 km^2) had significantly lower outcome scores (mean 3.6; s.d. = 1.12; $n = 20$) than systems with smaller areas (mean = 5.7; s.d. = 0.83; $n = 11$; Wilcoxon $p = 0.009$; figure 3).

4. Discussion

Three key conclusions arise from our data on threshold-based management systems. First, routine monitoring is strongly associated with positive environmental outcomes in both prospective and retrospective cases. Such agreement across independent datasets underscores the importance of monitoring in threshold-based systems, where frequently updated data are necessary to gauge ecosystem state relative to a tipping point. Second, management actions that explicitly use known environmental thresholds have better outcomes, particularly in retrospective cases. This finding—that explicit threshold-based management performs better than threshold-blind management—demonstrates the utility and policy relevance of information about ecosystem thresholds across a wide range of ecosystems. Finally, the regression tree results revealed the influence of resource system size on ecological outcomes in both prospective and retrospective cases, suggesting that smaller systems may provide managers with greater ability to understand, track and control drivers of ecological thresholds.

A core aim of this paper was to evaluate whether management decisions based expressly on known or suspected ecological thresholds perform better than threshold-blind management. We found that this kind of threshold-based management is associated with better outcomes overall, and despite a relatively small sample size, we observed that greater degrees of threshold management yielded increasingly higher outcome scores.

(a) Important attributes in both prospective and retrospective management

It may be both possible and politically feasible to avoid a threshold or to reverse an ecosystem change once a biophysical threshold has been crossed. However, doing so depends upon the availability of routine monitoring data. This allows managers to avoid reactionary practices that often result from a lack of necessary quantitative information to make accurate management decisions until after ecological degradation occurs. A steady flow of environmental data—coupled with science-based decision-making—can allow managers to (i) identify whether a nonlinear change has occurred, is occurring, or will occur in their managed system; (ii) determine whether their degraded system can be returned to its previous state (retroactive management); (iii) identify and directly manage the stressors driving the ecological change (either prospectively or retrospectively); and (iv) measure and set targets around a quantitative threshold. Accurate, up-to-date scientific information is vital for setting and achieving ecosystem targets [33], and for identifying and managing the direct driver of the environmental threshold. Routine monitoring on a scale relevant to the threshold therefore increases the resolution with which managers may guide decisions, and our results provide empirical support for monitoring as a key investment for effective management [34]. Of course, successful management also requires policies that trigger action in response to data indicating proximity to thresholds.

Longleaf pine management at Eglin Air Force Base and salt marsh management in the southeastern US offer concrete illustrations of the critical management attributes identified above. On Eglin, the presence of the endangered red-cockaded woodpecker, which prefers to nest in endemic longleaf pines, compelled managers to develop an adaptive management programme with regular monitoring and prescribed burning at intervals designed to limit sand pine invasion. If intervals between fires are longer than approximately 5 years, the invasive sand pines are able to establish, after which they have the effect of suppressing future fires, making it more difficult to restore the system [35] (Wildland Fire Center personnel and Forest Management team, Eglin Air Force Base: B Hagedorn, J Furman, J Hiers, S Hassel, A Sutsko, T Chavers 2013, personal communication). Eglin managers' interventions succeeded owing to their rigorous monitoring regime, robust scientific analysis to identify proper management strategies relative to the threshold, and aggressive management of the environmental stressor.

By contrast, salt marshes in the southeastern US have been declining for more than a decade despite continued—and often successful—restoration efforts to reduce nutrient inputs and lower soil salinity. However, mass die-offs of salt marsh vegetation are actually driven in large part by grazing snails [36,37], which cause precipitous salt marsh declines when their densities exceed a threshold of 800 snails per m^2 under normal conditions or around 500 snails per m^2 during drought conditions [37] (BR Silliman 2013, personal communication). Managers have nonetheless failed to act on these known snail density thresholds, and mass die-off events continue across the southeast as current management has failed to take threshold-relevant actions such as reducing the over-harvest of the snail's main predator, the blue crab [38].

(b) Attributes of secondary importance

In both prospective and retrospective cases, management of geographically smaller systems had better outcomes. In Washington State, for example, Taylor Shellfish Farms, Inc. is a commercial aquaculture operation that maintains the Pacific oyster (*Crassostrea gigas*) below its growth-and-viability threshold by managing the acid–base chemistry of its intake water prospectively [39] at a less-than-10 km^2 facility. Taylor manages a relatively small system and thus has the ability to control and monitor the important drivers of state change—in this case, the stressor is pH, and the outcome is juvenile oyster growth and survival. At larger scales, the magnitude and diverse nature of stressors increases and may overwhelm institutional capacity. At the other end of the size spectrum, large, nonbinding management systems such as the IWC tend to have less positive outcomes, even when success is judged with a single-species focus. Such global governance structures present challenges associated with large spatial extent (including spatial complexities, e.g. multiple stressors, cross-jurisdictional boundaries), varied management capacities and authorities, and cultural differences [40]. Smaller systems may be more nimble in managing around thresholds because of reduced biophysical complexity, fewer stressors and less institutional and jurisdictional complexity. Although institutional jurisdiction and the size of the resource system are not highly correlated in the dataset, higher congruence would be expected to lead to increased control over the drivers of ecosystem state change by the primary management entity [41]. Because managing on a smaller scale may by definition eliminate comprehensive ecosystem-level management opportunities, smaller-scale management systems should increase coordination and collaboration with neighbouring management systems to tackle larger, ecosystem-level drivers.

In retrospective cases, high HDI of the managing countries was an important correlate of a high outcome score. As an indicator for development, HDI may be associated with management capacity for implementing threshold-based management. As data are often scarce and scientific resources are often limited in low HDI countries, identifying ecosystem thresholds, indicators of those thresholds, relevant drivers and effective control measures can be challenging, requiring specialized expertise and resources. These limitations result in a lack of management action and poor ecological outcomes (e.g. fisheries that are not scientifically assessed perform poorly; [42]). In such cases, simple risk-assessment methods that can characterize ecosystem and resource status without statistical or deterministic models [43,44] are perhaps attractive means of incorporating better scientific information into management without requiring extensive financial or institutional resources. Moreover, in cases with attributes outside the control of resource managers, such as HDI and the size of the ecological system, our findings highlight the importance of other management attributes such as routine monitoring and explicit use of quantitative, ecological thresholds to set management targets.

Given the influence of spatial scale in both prospective and retrospective cases, we note that in large countries smaller-scale management may require the national government to set goals and provide funding for smaller jurisdictions (e.g. local or state government) to develop smaller, discrete management units within larger-scale planning regions. The benefits of devolving management responsibility to multiple, smaller-scale entities include localizing the costs and benefits of management, thereby increasing control and investment of resource stewardship ([22]; but see [45]). However, multi-scale management is more likely to be successful in areas with higher levels of leadership and social capital as well as learning networks and bridging organizations [46–48]—putting large countries with low HDI and large management units (such as developing countries with large marine protected areas) at a potential disadvantage.

Moreover, many systems are managed based on the geographical range of the environmental variable of concern and, thus, the designated jurisdiction cannot be easily subdivided. It is difficult to strike a balance between the large scales of ecosystems and the small scale of effective governance. One way of balancing spatial scales is to create a common legal framework with nested spatial authority—examples include many US environmental regulatory schemes such as those implementing the Clean Water Act, Clean Air Act and other statutes—but here again, the resources and bureaucratic machinery to carry out such nested regulation are likely to be scarce in many areas. Explicit threshold-based management and decisions tied to routine monitoring therefore remain the greatest leverage points in resource systems that cannot be subdivided.

(c) Barriers

It is important to note that even with strict adherence to the principles listed above management efforts may sometimes fail to achieve conservation objectives [27]. It may be difficult to achieve desired ecological outcomes in some systems because of biophysical barriers to threshold-based management [49]. For example, prospective management actions may not be identifiable through analogous systems, modelling, past observations or small-scale experiments, and retrospective management will be difficult or impossible in systems that have undergone an irreversible or hysteretic change. Natural stochasticity and environmental variability may also reduce management effectiveness [27]. Additionally, there may be social barriers (such as political or economic) that limit a management system's ability to develop an appropriate management response, and many of the attributes of linked social–ecological systems are inherent to the system and therefore not malleable. In such cases, management system attributes that *are* malleable may take on outsized importance in efforts to increase management effectiveness.

(d) Elements of effective management in environments with known threshold dynamics

We suggest the following as preliminary recommendations for threshold-based environmental management: (i) the managed system includes routine monitoring of relevant environmental variables on a timescale relevant to the threshold; (ii) management decisions (such as setting targets for various stressors) are explicitly threshold-based; and (iii) where appropriate, the managed system is composed of geographically small management units with frequent interaction among units. These are consistent with prior recommendations for environmental management more generally [5,11,34,50]. Our results provide empirical support for their application to the management of systems that have thresholds.

These attributes may seem obvious; however, other 'obvious' attributes included in the analysis—such as adaptive management, binding legal agreements and cost–benefit analyses—were not strong correlates of outcome score. As such, our empirical analysis helps to develop a set of priorities among otherwise equally attractive elements of environmental management. The specifics of how effective management systems are designed will vary based on available information, extent of governance authority, accessible funding and a host of other factors. We hope that the priorities we identify can be useful to effectively allocate resources in the context of the inevitable trade-offs that environmental management entails.

The results of our study are particularly compelling because they are drawn from a compilation of real-world cases. In recent years, there has been a call to design studies that embrace the complex nature of *in situ* environmental management and systematically test hypotheses over a wide variety of situations, so that results can be externally validated [50,51]. Our study does just this, and our results may therefore be especially useful to managers wishing to incorporate scientific guidance into on-the-ground management efforts.

(e) Benefits of, and incentives for, threshold-based management

Threshold-based management may result in a variety of social and institutional benefits, including (i) increased management efficiency, (ii) improved assessment of social and ecological trade-offs, (iii) more accurate prioritization of management goals based on possible social and ecological outcomes, and (iv) increased opportunities to develop and achieve clear management targets. South Africa's Kruger National Park provides an example of improved efficiency, where managers monitor resources on a timescale relevant to the estimated ecosystem threshold, reducing the expenditure of valuable social and economic capital on unnecessary monitoring. They also use the knowledge of biophysical thresholds to assess management trade-offs, implementing a framework that systematizes trade-off decision-making to provide an efficient management scheme within a resource and data-poor environment and an effective feedback mechanism for adaptive management [52,53]. These trade-off analyses also allow managers to prioritize management goals based on social and ecological outcomes. For example, Kruger reached a critical threshold for woodland degradation owing to expanding elephant populations, but prioritized tourist revenue from elephant viewing over woodland health.

(f) Limitations of threshold-based management

Threshold-based management may not be relevant or desirable for all managed ecosystems. First, it can be difficult to distinguish human-caused shifts from background ecosystem variability, particularly in the absence of sufficient baseline data. Non-anthropogenic examples of dramatic, nonlinear change (such as the Pacific decadal oscillation) may require managers to adapt resource use to the new conditions (e.g. by reducing allowable fish harvest when ocean productivity is low), but may not require managers to address the fundamental state-change drivers as these are outside of their control. Second, understanding where managers can act to mitigate human-caused shifts in stable state is critical for employing threshold-based strategies to management. Such an understanding can be elusive, however, and may require greater research and monitoring capacity than is feasible in many systems. Third, prospective management entails uncertainty about the proximity of an ecological threshold, and therefore may be politically costly relative to benefits, at least in the short-term. Finally, there may be some cases in which the costs of prospective or retrospective management exceed the benefits associated with conserving or recovering ecosystem services associated with a desired ecosystem state. Conceivably, prospective avoidance of a threshold could result in the conservation of an ecosystem service of low value relative to alternative ecosystem states, resulting in an unfavourable cost–benefit ratio if management is narrowly focused on a single ecosystem service. These limitations reinforce the need for appropriate monitoring and study of threshold-prone systems prior to—and throughout—management engagement.

5. Conclusion

Many ecosystems respond to anthropogenic stressors in nonlinear ways; that is, small changes in a stressor result in large changes in ecosystem state and delivery of ecosystem services. Management stands to benefit from understanding these relationships to proactively avoid tipping points or restore system dynamics that have crossed thresholds. Notably, we found that management is most effective when it is explicitly using science to avoid thresholds or to reverse ecosystem change after a threshold has been crossed. Moreover, effective environmental management is associated with routine monitoring of the system on a temporal and spatial scale relevant to the ecological threshold, and with local- and regional-scale management rather than decision-making at larger spatial scales. These findings support increased integration of threshold-based science tools into ecosystem management and the continued investigation of thresholds as leverage points for efficiently maintaining ecosystem structure and functioning.

Acknowledgements. Many members of the larger project provided valuable feedback and logistical help, including (but not limited to): Meg Caldwell, Larry Crowder, Melissa Foley, Megan Mach, Becca Martone, Carrie Kappel, Ben Halpern, Kim Selkoe, Crow White, Mary Hunsicker, Courtney Scarborough, Phil Levin, Adrian Stier and Jameal Samhouri. Sincere thanks also to those who provided background information and conversations about the individual case studies.

Funding statement. The authors gratefully acknowledge support of the Gordon and Betty Moore Foundation, who supported this work as part of the Ocean Tipping Points project (www.oceantippingpoints.org).

References

1. Holling CS. 1973 Resilience and stability of ecological systems. *Annu. Rev. Ecol. Syst.* **4**, 1–23. (doi:10.1146/annurev.es.04.110173.000245)
2. Scheffer M, Carpenter SR. 2003 Catastrophic regime shifts in ecosystems: linking theory to observation. *Trends Ecol. Evol.* **18**, 648–656. (doi:10.1016/j.tree.2003.09.002)
3. Folke C, Carpenter S, Walker B, Scheffer M, Elmqvist T, Gunderson L, Holling CS. 2004 Regime shifts, resilience, and biodiversity in ecosystem management. *Annu. Rev. Ecol. Evol. Syst.* **35**, 557–581. (doi:10.1146/annurev.ecolsys.35.021103.105711)
4. Carpenter SR, Brock WA. 2006 Rising variance: a leading indicator of ecological transition. *Ecol. Lett.* **9**, 311–318. (doi:10.1111/j.1461-0248.2005.00877.x)
5. Suding KN, Hobbs RJ. 2009 Threshold models in restoration and conservation: a developing framework. *Trends Ecol. Evol.* **24**, 271–279. (doi:10.1016/j.tree.2008.11.012)
6. Samhouri JF, Levin PS, James CA, Kershner J, Williams G. 2011 Using existing scientific capacity to set targets for ecosystem-based management: a Puget sound case study. *Mar. Policy* **35**, 508–518. (doi:10.1016/j.marpol.2010.12.002)
7. Levin PS, Möllmann C. 2015 Marine ecosystem regime shifts: challenges and opportunities for ecosystem-based management. *Phil. Trans. R. Soc. B* **370**, 20130275. (doi:10.1098/rstb.2013.0275)
8. Groffman PM *et al.* 2006 Ecological thresholds: the key to successful environmental management or an important concept with no practical application? *Ecosystems* **9**, 1–13. (doi:10.1007/s10021-003-0142-z)
9. Osman RW, Munguia P, Zajac RN. 2010 Ecological thresholds in marine communities: theory, experiments and management. *Mar. Ecol. Prog. Ser.* **413**, 185–187. (doi:10.3354/meps08765)
10. Samhouri JF, Levin PS, Ainsworth CH. 2010 Identifying thresholds for ecosystem-based management. *PLoS ONE* **5**, e8907. (doi:10.1371/journal.pone.0008907)

11. Scheffer M *et al*. 2012 Anticipating critical transitions. *Science* **338**, 344–348. (doi:10.1126/science.1225244)
12. McClanahan TR, Graham NAJ, MacNeil MA, Muthiga NA, Cinner JE, Bruggemann JH, Wilson SK. 2011 Critical thresholds and tangible targets for ecosystem-based management of coral reef fisheries. *Proc. Natl Acad. Sci. USA* **108**, 17 230–17 233. (doi:10.1073/pnas.1106861108)
13. Lindegren M, Dakos V, Gröger JP, Gårdmark A, Kornilovs G, Otto SA, Möllmann C. 2012 Early detection of ecosystem regime shifts: a multiple method evaluation for management application. *PLoS ONE* **7**, e38410. (doi:10.1371/journal.pone.0038410)
14. Gunderson LH. 2001 South Florida: the reality of change and the prospects for sustainability: managing surprising ecosystems in Southern Florida. *Ecol. Econ.* **37**, 371–378. (doi:10.1016/S0921-8009(01)00179-3)
15. Roe E, van Eeten M. 2001 Threshold-based resource management: a framework for comprehensive ecosystem management. *Environ. Manage.* **27**, 195–214. (doi:10.1007/s002670010143)
16. Hunter ML, Bean MJ, Lindenmayer DB, Wilcove DS. 2009 Thresholds and the mismatch between environmental laws and ecosystems. *Conserv. Biol.* **23**, 1053–1055. (doi:10.1111/j.1523-1739.2009.01205.x)
17. Quinn II TJ, Fagen R, Zheng J. 1990 Threshold management policies for exploited populations. *Can. J. Fish. Aquat. Sci.* **47**, 2016–2029. (doi:10.1139/f90-226)
18. Kemp WM *et al*. 2005 Eutrophication of Chesapeake Bay: historical trends and ecological interactions. *Mar. Ecol. Prog. Ser.* **303**, 1–29. (doi:10.3354/meps303001)
19. Farley J. 2012 Ecosystem services: the economics debate. *Ecosyst. Serv.* **1**, 40–49. (doi:10.1016/j.ecoser.2012.07.002)
20. Huggett AJ. 2005 The concept and utility of 'ecological thresholds' in biodiversity conservation. *Biol. Conserv.* **124**, 301–310. (doi:10.1016/j.biocon.2005.01.037)
21. Ostrom E. 2007 A diagnostic approach for going beyond panaceas. *Proc. Natl Acad. Sci. USA* **104**, 15 181–15 187. (doi:10.1073/pnas.0702288104)
22. Ostrom E. 2009 A general framework for analyzing sustainability of social-ecological systems. *Science* **325**, 419–422. (doi:10.1126/science.1172133)
23. Cox M. 2014 Understanding large social-ecological systems: introducing the SESMAD project. *Intl J. Commons* **8**, 265–276.
24. Mills M, Pressey RL, Ban NC, Foale S, Aswani S, Knight AT. 2013 Understanding characteristics that define the feasibility of conservation actions in a common pool marine resource governance system. *Conserv. Lett.* **6**, 418–429. (doi:10.1111/conl.12025)
25. Basurto X, Ostrom E. 2009 Beyond the tragedy of the commons. *Econ. Delle Fonti Energ. Dell'ambiente* **52**, 55–60.
26. McGinnis MD. 2011 An introduction to IAD and the language of the Ostrom workshop: a simple guide to a complex framework. *Policy Stud. J.* **39**, 169–183. (doi:10.1111/j.1541-0072.2010.00401.x)
27. Game ET, Kareiva P, Possingham HP. 2013 Six common mistakes in conservation priority setting. *Conserv. Biol.* **27**, 480–485. (doi:10.1111/cobi.12051)
28. Mastrandrea MD *et al*. 2010 *Guidance note for lead authors of the IPCC Fifth Assessment Report on Consistent Treatment of Uncertainties*. Intergovernmental Panel on Climate Change (IPCC). Available at http://www.ipcc.ch.
29. The R Development Core Team. 2010 *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
30. Venables WN, Ripley BD. 2002 *Modern applied statistics with S*, 4th Edn. New York, NY: Springer.
31. Jackman S. 2011 *pscl: Classes and methods for R developed in the political science computational laboratory*. Stanford, CA: Stanford University.
32. De'ath G. 2007 *mvpart: Multivariate partitioning, R package ver. 1.2-6*. Available at http://cran.r-project.org/web/packages/mvpart/index.html.
33. Pressey RL, Cabeza M, Watts ME, Cowling RM, Wilson KA. 2007 Conservation planning in a changing world. *Trends Ecol. Evol.* **22**, 583–592. (doi:10.1016/j.tree.2007.10.001)
34. Biggs R, Carpenter SR, Brock WA. 2009 Turning back from the brink: detecting an impending regime shift in time to avert it. *Proc. Natl Acad. Sci. USA* **106**, 826–831. (doi:10.1073/pnas.0811729106)
35. Hardesty J, Adams J, Gordon D, Provencher L. 2006 Simulating management with models. *Conserv. Pract.* **1**, 26–32. (doi:10.1111/j.1526-4629.2000.tb00157.x)
36. Silliman BR, Bertness MD. 2002 A trophic cascade regulates salt marsh primary production. *Proc. Natl Acad. Sci. USA* **99**, 10 500–10 505. (doi:10.1073/pnas.162366599)
37. Silliman BR, van de Koppel J, Bertness MD, Stanton LE, Mendelssohn IA. 2005 Drought, snails, and large-scale die-off of southern U.S. salt marshes. *Science* **310**, 1803–1806. (doi:10.1126/science.1118229)
38. Bertness MD, Silliman BR. 2008 Consumer control of salt marshes driven by human disturbance. *Conserv. Biol.* **22**, 618–623. (doi:10.1111/j.1523-1739.2008.00962.x)
39. Dewey B. 2012 The canary in the coal mine: a cautionary tale of ocean acidification and shellfish. Presentation given at 1st Meeting of Washington State Blue Ribbon Panel on Ocean Acidification, 30 March 2012, Seattle, WA. Available at http://tvw.org/index.php?option=com_tvwplayer&eventID=2012030125A and http://www.ecy.wa.gov/water/marine/oceanacidification.html.
40. Kittinger JN, Dowling A, Purves AR, Milne NA, Olsson P. 2011 Marine protected areas, multiple-agency management, and monumental surprise in the Northwestern Hawaiian Islands. *J. Mar. Biol.* **2011**, 1–17. (doi:10.1155/2011/241374)
41. Cash DW, Adger WN, Berkes F, Garden P, Lebel L, Olsson P, Pritchard L, Young O. 2006 Scale and cross-scale dynamics: governance and information in a multilevel world. *Ecol. Soc.* **11**, 8.
42. Costello C, Gaines SD, Lynham J. 2008 Can catch shares prevent fisheries collapse? *Science* **321**, 1678–1681. (doi:10.1126/science.1159478)
43. Hobday AJ *et al*. 2011 Ecological risk assessment for the effects of fishing. *Fish. Res.* **108**, 372–384. (doi:10.1016/j.fishres.2011.01.013)
44. Samhouri JF, Levin PS. 2012 Linking land- and sea-based activities to risk in coastal ecosystems. *Biol. Conserv.* **145**, 118–129. (doi:10.1016/j.biocon.2011.10.021)
45. Gelcich S, Edwards-Jones G, Kaiser MJ, Castilla JC. 2006 Co-management policy can reduce resilience in traditionally managed marine ecosystems. *Ecosystems* **9**, 951–966. (doi:10.1007/s10021-005-0007-8)
46. Gutiérrez NL, Hilborn R, Defeo O. 2011 Leadership, social capital and incentives promote successful fisheries. *Nature* **470**, 386–389. (doi:10.1038/nature09689)
47. Rotarangi S, Russell D. 2009 Social-ecological resilience thinking: can indigenous culture guide environmental management? *J. R. Soc. NZ* **39**, 209–213. (doi:10.1080/03014220909510582)
48. Berkes F. 2009 Evolution of co-management: role of knowledge generation, bridging organizations and social learning. *J. Environ. Manage.* **90**, 1692–1702. (doi:10.1016/j.jenvman.2008.12.001)
49. Lyytimäki J, Hildén M. 2007 Thresholds of sustainability: policy challenges of regime shifts in coastal areas. *Sustain. Sci. Pract. Policy* **3**, 61–69.
50. Schultz L, Duit A, Folke C. 2011 Participation, adaptive co-management, and management performance in the world network of biosphere reserves. *World Dev.* **39**, 662–671. (doi:10.1016/j.worlddev.2010.09.014)
51. Naeem S, Duffy JE, Zavaleta E. 2012 The functions of biological diversity in an age of extinction. *Science* **336**, 1401–1406. (doi:10.1126/science.1215855)
52. Rogers K, Biggs H. 1999 Integrating indicators, endpoints and value systems in strategic management of the rivers of the Kruger National Park. *Freshw. Biol.* **41**, 439–451. (doi:10.1046/j.1365-2427.1999.00441.x)
53. Van Wilgen BW, Biggs HC. 2011 A critical assessment of adaptive ecosystem management in a large savanna protected area in South Africa. *Biol. Conserv.* **144**, 1179–1187. (doi:10.1016/j.biocon.2010.05.006)