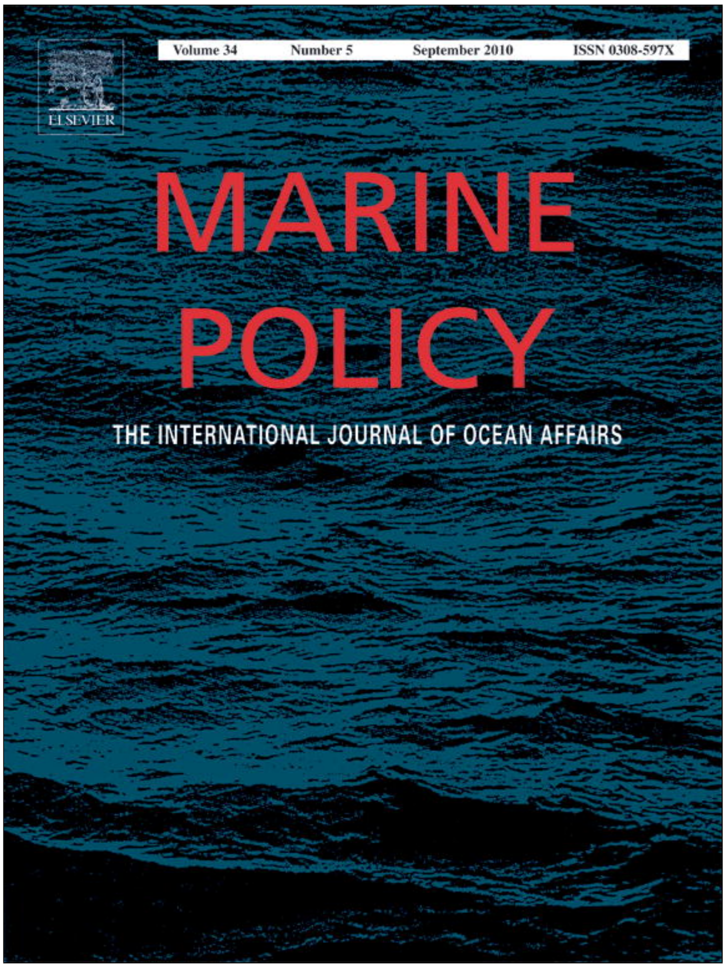

Volume 34 Number 5 September 2010 ISSN 0308-597X

ELSEVIER

MARINE POLICY

THE INTERNATIONAL JOURNAL OF OCEAN AFFAIRS

Marine Policy 34 (2010) 836–843

Contents lists available at ScienceDirect

Marine Policy

journal homepage: www.elsevier.com/locate/marpol

Scientific inference and experiment in Ecosystem Based Fishery Management, with application to Steller sea lions in the Bering Sea and Western Gulf of Alaska

Marc Mangel *

Center for Stock Assessment Research and Department of Applied Mathematics and Statistics, MS E-2, University of California, Santa Cruz, CA 95064, USA

ARTICLE INFO

Article history:
Received 24 November 2009
Received in revised form
11 January 2010
Accepted 11 January 2010

Keywords:
Adaptive management
Ecosystem Based Fishery Management
Steller sea lions
Experiment

ABSTRACT

Learning about ecosystem processes and patterns is an essential component of Ecosystem Based Fishery Management and the sustainable use of natural resources.

Currently, such learning is usually done through adaptive management (passive or active) or Management Strategy Evaluation, which are explained. An example of adaptive management in northwestern Australia shows the strengths and limitations of management experiments and raises the question of how to learn if an experiment is not practicable. Both adaptive management and Management Strategy Evaluation are examples of scientific inference, an idea introduced by Sir Harold Jeffreys nearly 80 years ago. With sufficient variation, even if it is not through controlled experiments, scientific inference is possible by combining mechanistic models with statistical methods; the recently proposed paradigm of 'adaptive monitoring' is another case of scientific inference. The decline of Steller sea lions in the Bering Sea and Aleutian Islands is reviewed, including the only work in which 10 hypotheses concerning the decline were simultaneously compared. It is concluded that scientific inference using mechanistic models and fine scale data at the level of the rookery can provide useful information about the interactions of fisheries, fish populations, and Steller sea lions.

1. Introduction

1.1. *Ecosystem Based Fishery Management is a reality*

An ecosystem based approach to fishery management (EBFM) is now recognized as requisite for the 21st century [1]. The overall objective of EBFM is to sustain healthy marine ecosystems and the fisheries they support. There are a variety of components to achieving this goal: (i) avoiding degradation of ecosystems, (ii) minimizing the risk of irreversible change to assemblages of species and ecosystem processes; (iii) finding ways to obtain and maintain long-term socioeconomic benefits without compromising the ecosystem; and (iv) generating knowledge of ecosystem processes sufficient to understand the likely consequences of human actions, which is the focus of this essay. In almost every situation knowledge is limited [2] and one must learn about the ecosystem.

The limitation of knowledge implies that robust and precautionary fishery management measures that favor the ecosystem should be adopted. For example, it is clear that consistently removing more biomass than the system productivity, even after accounting for the requirements of other ecosystem components, is a non-sustainable practice. Similarly, maintaining system characteristics within certain bounds may protect ecosystem resilience and avoid irreversible changes [2]. Single-species management has been successful at reducing incidental killing of non-target species or undersized individuals of the target species [1]. Indirect effects are more complicated [3] but can also be addressed in the context of single-species management. Thus, EBFM can be characterized by effective single-species management with the addition of precautionary set-asides for unknown ecosystem components [1].

Furthermore, EBFM can be implemented in systems that differ in levels of information and uncertainty through the judicious use of a precautionary approach [1]. This means erring on the side of caution in setting management targets and limits when information is sparse or uncertain, and learning as one proceeds. Greater uncertainty should be associated with more stringent management measures. Indeed, EBFM shifts the burden of proof so that fishing would not take place unless it could be shown not to harm key components of the ecosystem [4].

In the last two decades, there have been rapid developments in the statistical tools for learning about marine ecosystems and their components [5–7] and stochastic dynamic programming has become popular as a management tool [8–11]. There are many advantages in model formulation that accrue because of the process of formalizing our thinking about dynamics and costs and benefits [9,12].

---

* Tel.: +1 831 234 2970; fax: +1 831 459 4482.
E-mail address: msmangel@ucsc.edu

0308-597X/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.marpol.2010.01.005

1.2. *Steller sea lions in the Western Bering Sea and Gulf of Alaska*

The population of Steller sea lions (*Eumetopias jubatus*) in western Alaska fell by more than 80% between the late 1970s and the early 2000s. The initial decline was characterized by a loss of about 15% of the population per year, accompanied by reduced size-at-age and other symptoms of nutritional stress. In the 1990s, the rate of decline slowed to about 5% per year (reviewed in [13]). The available evidence suggested that animals in the declining western population were in better condition than those in the growing eastern population at this time, in line with the general expectation for a population below carrying capacity. The western stock was listed as Endangered in 1997. Despite the efforts of many scientists to identify a ''smoking gun'' responsible for the decline, there is still no definitive answer.

In [13], one also finds the first comprehensive analysis of the numerous hypotheses that have been proposed to explain the decline. They fall into five main categories (Table 1): I. Insufficient prey availability. II. Unsuitable prey species composition. III. Direct mortality related to fishing activities. IV. Enhanced depredation by killer whales or sharks (also see [14,15]). V. Disease or pathogen outbreaks. There is some evidence in support of each of these, but none is conclusive ([13,16] for review).

The lack of consensus results in part because the different hypotheses make non-exclusive predictions regarding the population trends to which they are fitted. A second issue is that the historical data (rookery counts based on aerial photos taken during the breeding season) are both noisy and sparse, leading most researchers to sacrifice some useful variation by pooling data across rookeries or across years [13]. This exacerbates the first problem, because the spatial and temporal pattern of the decline is crucial for differentiating between alternative effects. Thus, different researchers have found support for different hypotheses, leading to confusion and dissension.

Table 1
The 10 hypotheses concerning the decline of Steller sea lions and their respective parameters ([13]).

Hypothesis	Parameter and mechanism of decline
Category I: Insufficient prey availability	
H1: Lower prey density → Lower fecundity rate	c_1, Prey density increment per unit encounter rate; foraging shortfalls terminate pregnancy
H2: Lower prey density → Less pup recruitment	c_2, Prey density increment per unit encounter rate; foraging shortfalls cause pup starvation
H3: Lower prey density → Lower non-pup survival	c_3, Prey density increment per unit encounter rate; shortfalls cause non-pup starvation
Category II: Unsuitable prey species composition	
H4: Higher pollock fraction → Lower fecundity rate	c_4, Exponent of non-pollock prey fraction in fecundity multiplier
H5: Higher pollock fraction → Less pup recruitment	c_5, Exponent of non-pollock prey fraction in pup recruitment multiplier
H6: Higher pollock fraction → Lower non-pup survival	c_6, Exponent of non-pollock prey fraction in non-pup survival multiplier
Category III: Direct mortality due to fishing activities	
H7: More fishing activity → Less pup recruitment	c_7, Pup mortality rate per fishery gear deployment within 20 km of rookery
H8: More fishing activity → Lower non-pup survival	c_8, Non-pup mortality rate per fishery gear deployment within 20 km of rookery
Category IV: Enhanced depredation by killer whales or sharks	
H9: Fewer harbor seals (more predation) → Less pup recruitment	c_9, Fraction of potential pup recruitment lost when harbor seal density $< h_{crit}$
H10: Fewer harbor seals (more predation) → Lower non-pup survival	c_{10}, Fraction of potential non-pup survival lost when harbor seal density $< h_{crit}$
H9, H10	h_{crit}, Harbor seal density below which sea lions become prey to killer whales

The report of the National Research Council [16] recommended the use of finer scale data (as in [13]) and ''Use [of] replicated closed and open rookeries to experimentally evaluate localized fishery impacts'' (p. 157) and ''[to] Implement a 'titration experiment' that progressively increases restrictions on fisheries (such as area closures) until it is clear whether or not a positive response can be achieved via fishery management'' (p. 158). About the former, the authors of [16] write ''Although option #3 [use of replicated closed and open rookeries to experimentally evaluate localized fishery impacts] provides the best chance of providing new information about the role of fisheries in the Steller sea lion decline, there is no guarantee that the outcome of this adaptive management experiment will provide an unequivocal result. A small, positive response to the fishery closures may be masked by fluctuations in other factors that also contribute to the decline'' (p. 160; also the Australian example described below).

In discussing the potential of experiment in the pollock fishery, the At-sea Processors Association concluded [17]:

> **APA's Plan for Condition #10:** The Final Report on BS/AI Alaska Pollock recognizes the legal and practical impediments identified by fishery management authorities and scientists to conducting the controlled area experiments proposed by the National Research Council (NRC) in 2002. In addition, NOAA Fisheries' scientists have provided fishery management authorities with a detailed analysis of the substantial cost of such experiments, the decades-long commitment required for such a program and the likely prospect that the findings would be inconclusive. Notwithstanding the issues identified above, APA is aware that AFSC is in its fourth year of research testing the localized depletion hypothesis and will continue with its program if FY 2005 funding is available through Congressional appropriation. (See discussion under Condition #5 above.) NOAA Fisheries' previous work on possible fishing effects on SSLs has examined fisheries for Alaska pollock, Pacific cod and Atka mackerel. APA will request a meeting with AFSC and the certifier within six months to review research results to date and to discuss ongoing research. APA will consult with the certifier and AFSC prior to the meeting to ensure all issues relevant to both groups are addressed at the meeting. In addition, APA will propose that the meeting include a thorough discussion on the current state of research on hypotheses relating to possible effects of pollock fishing on foraging sea lions, including agency-sponsored research and research projects conducted under the auspices of the Alaska SeaLife Center, the Pollock Conservation Cooperative Research Center, the North Pacific Research Consortium, and other noted authorities. By the first annual audit, APA will prepare and provide a report to the certification body detailing actions and timelines for meeting the objectives of this condition should the results of the meeting between APA, NMFS and the certification body identify continuing research needs to meet the condition.

One is thus lead to ask the question: if a controlled experiment is not practicable are there alternatives to experiment that are consistent with the goal of learning about the ecosystem and its components with sufficient detail to aid EBFM? Answering this question requires understanding adaptive management, management strategy evaluation, and scientific inference.

2. Adaptive management

2.1. Overview

The nature of management in ecology is often misunderstood, and adaptive management is most misunderstood. To begin, we do not manage, and will probably never manage, populations, communities, or ecosystems. Rather, we manage human intervention with them [2]. "Adaptive management" is not trying something, seeing it not work, and trying something else. Rather, it is learning as we are doing; adaptive management forces us into evidence-based policy, rather than policy-based evidence [18].

Adaptive management of renewable resources has about a 35-year history [19–21] and has been popularized very much by Possingham and colleagues (e.g. [10,11,22]). The essence of adaptive management is to learn. This can be done passively or 'off-line' from the management actions or actively, so that part of the management activity is to sacrifice physical gain to gain information. For example, in early papers [23,24] on the adaptive management of marine resources the value of information is computed and it is demonstrated that a little bit of contrast can go a long way in resolving uncertainties. Regardless of the form of adaptive management, monitoring and learning essential but experiment is not [25,26]. Adaptive management can be viewed as a form of evidence based policy (*sensu* [27]) in which evidence is used to answer questions such as 'what works' and 'what happens if we change these settings?'.

2.2. Active adaptive management in Northwestern Australia

In the final quarter of the 20th century, an adaptive management experiment was conducted on the fishery on the northwest shelf of Australia [28,29]. The fisheries involved a variety of species, but mainly from the families of fish that included emperor bream (*Lethrinus*), threadfin bream (*Nemipterus*), Bombay ducks (*Saurida*), and snappers (*Lutjanus*). *Lethrinus* and *Lutjanus* were high valued species and *Nemipterus* and *Saurida* were lower-valued. A Japanese trawl fishery took mainly emperor bream, a Taiwanese trawl fishery took all species, and in 1984 an Australian trap fishery began, targeting emperor bream, snapper, and sea bass (*Epinephelus*). *Lethrinus* and *Lutjanus* had a significantly higher probability of occurrence in areas where there were large benthic organisms (e.g. sponges) than in areas where there were not.

Survey vessel data were available for some years since 1960 and total catch data, separated by species and fishery, were available. During the period of interest (about 1958–1985) the demersal environmental was verified to have changed. Overall, the fish community maintained total biomass but changed in composition from a community that was 40–60% by mass emperor bream and snapper to one that was 10% emperor bream and snapper and about 25% threadfin bream and Bombay duck.

A decision to expand the Australian trap fishery would be most effective with a return to a community structure more similar to the one in 1960 than the one in 1980. Two key management questions emerged [28]. First, would the long-term rewards to Australia justify the effort to alter the structure of the fish community? Second, if so, what combination of trawling (or trawl closures) and trap would be best? Answering these questions required understanding the factors that structure the fish community.

At least four hypotheses could be identified [28]: H1: All species groups are controlled by intraspecific competition, in which case the dynamics of biomass can be characterized by a series of unrelated population dynamic equations (discrete time logistic equations with constant carrying capacity). H2: The population sizes of *Lethrinus* and *Lutjanus* are controlled by intraspecific processes while those of *Nemipterus* and *Saurida* experience competition determined by the total biomass of the first two species. H3: The converse of H2, i.e. that the population sizes of *Nemipterus* and *Saurida* are controlled by intraspecific processes while those of *Lethrinus* and *Lutjanus* experience competition determined by the total biomass of the first two species. H4: The carrying capacity of all species is determined by the amount of suitable habitat, which is negatively affected by trawling.

These hypotheses have different implications concerning the yield available to a trap fishery. Hypothesis 1 and 2 imply a low yield because of the decline of *Lethrinus* and *Lutjanus* under fishing pressure is interpreted as being due to low productivity of the stock (small maximum per capita growth rate r in a discrete logistic equation). On the other hand, Hypotheses 3 and 4 predict high potential yield from a trap fishery since the trap fishery does little damage to habitat (which is assumed to recover in Hypothesis 4) and a trap fishery removes few *Nemipterus* and *Saurida*, which were predicted to return to their unfished levels with a concomitant reduction in the negative effects that they have on *Lethrinus* and *Lutjanus* populations.

Three long-term fishing policies were considered [28]: (1) a trap fishery for *Lethrinus* and *Lutjanus* with fishing mortality close to the level providing Maximum Sustainable Yield (MSY) for Hypotheses 1 and 2; (2) a trap fishery with fishing mortality close to the MSY level for Hypotheses 3 and 4; or (3) continuation of trawling as *status quo* with no development of the trap fishery into areas then trawled. With the assumption of a learning period and prior probabilities for the different hypotheses, it was possible to compute the expected economic value of each of the fishing policies. In part because of this analysis, the Australian management agencies instituted a zonal management scheme with rotating closures of areas to trawling and the development of a trap fishery.

The full analysis for the adaptive management experiment consisted of a 5-year management regime that corresponded to a period of learning, during which both revenue from fish and information about fish were obtained. At the end of this 5-year period, Bayesian methods were used to update the probabilities of the hypotheses and a risk neutral manager selected one of the strategies for that time onwards. This period ended roughly 1985 when the management was set in place, although information continued to be collected, which allowed for passive adaptive management.

Updates of the status of the adaptive management experiment are given in [29,30] and are summarized in Table 2: 10 years into the adaptive management experiment one could give 2:1 odds that habitat destruction by trawling was the limiting factor in the recovery of the fish community to its pre-1960s level—these odds are far from 'statistically significant' in the cult of hypothesis testing [31]. Data collected in the 1990s provided further support for Hypothesis H4 [30], with the conclusion that trap fishing should replace trawling to achieve a recovery in the fish community structure and improve the fishery for the more valuable resources.

2.3. Passive adaptive management in the Bering Sea

The Fishery Management Plan (FMP) for the Bering Sea includes a detailed plan for passive adaptive management [32]. Management is divided into 6 Tiers, based on the information that is available. The ranking, from least informative to most

Table 2
Summary of the Adaptive Management Experiment on the Northwestern Shelf of Australia (modified from [29,30]).

Hypothesis	Probability		
	Prior	1985	1990
H1: No species interactions	0.25	0.01	0.02
H2: The population sizes of *Lethrinus* and *Lutjanus* negatively impact those of *Nemipterus* and *Saurida*	0.25	0.52	0.33
H3: The population sizes of *Nemipterus* and *Saurida* negatively impact those of *Lethrinus* and *Lutjanus*	0.25	0.01	0.03
H4: Carrying capacity determined by suitable habitat, which is negatively affected by trawling	0.25	0.46	0.62

informative is

> *Tier 6*: A reliable history of catch is available.
> *Tier 5*: There are reliable estimates of current biomass and natural mortality.
> *Tier 4*: There are reliable estimates of biomass, and the fishing mortalities that reduce spawners per recruit to 35% and 40% of the unfished levels.
> *Tier 3*: As in Tier 4 but there is also a reliable estimate of the biomass associated with the fishing mortality rates described above.
> *Tier 2*: There are reliable estimates of biomass, the biomass producing maximum sustainable yield and the fishing mortality rates in Tier 4.
> *Tier 1*: There are reliable estimates of biomass, the biomass producing maximum sustainable yield, and a probability distribution for the fishing mortality producing maximum sustainable yield.

A thorough review [33] of the harvest strategies used by the North Pacific Fishery Management Council (henceforth simply the Council) lead to conclusions that include the following:

(1) Of all the stocks, only Bering Sea Aleutian Islands (BSAI) pollock were well-enough quantified to qualify for Tier 1 management (in the Gulf of Alaska pollock are classified as Tier 3 and in the Aleutian Islands and Bogoslof as Tier 5). In Tier 1 the allowable biological catch is determined using the harmonic mean of the distribution for the fishing mortality producing maximum sustainable yield. The harmonic mean has the mathematical property that it is less than the simple average by an amount that increases with the spread of the distribution, so this establishes a margin that increases with the uncertainty in the estimate. However, the harmonic mean does not have the statistical property of ensuring a constant specified confidence that the associate fishing mortality does not exceed the true fishing mortality giving maximum sustainable yield, as would be ensured by using a lower confidence limit of the estimate of that distribution.
(2) In a single-species/target-stock context, the process for setting Total Allowable Catch (TAC) employed by the Council is a very conservative one, at least for Tiers 1 through 5, and the in-season monitoring and management system seems adequate for implementing the TACs with little risk of exceeding them.
(3) In practice, this management system seems to have worked well, judged simply by the continuing productivity of the target stocks, for the bulk of the BSAI/GOA stocks in recent decades. However, there are at least 21 species that consume Pollock [2], so it is possible to create a meaningful frequency distribution of proportion of standing stock consumed. To move human consumption from the extreme right hand tail of the distribution to the center would require approximately a 100-fold reduction in harvest rates; to move it roughly to one standard deviation above the mean would require about a 10-fold reduction in harvest rates [2].

The definite exceptions to this empirical record of success are the rockfish, which were overfished early on, and in general have not recovered; this is a common experience with the *Sebastes*, which have remarkably low productivity for their size [34].

3. From adaptive management to management strategy evaluation

In [28] it was noted “The extent to which short-term revenue should be compromised to gain information about resource dynamics will depend on both social acceptability and the value of the information in leading to a higher revenue in the long term” (p. 377). This Australian experience was instrumental in the development of Management Strategy Evaluation (MSE; [35,36]).

MSE involves using a model (or models) to assess the consequences of a range of management strategies and presenting the results in a way that makes the trade-offs clear. Unlike adaptive management, in MSE one does not try to prescribe an optimal strategy, but rather one tries to provide decision makers with information about the states of nature and the consequences of various actions, as in classical decision theory. Also unlike adaptive management, MSE does not necessarily focus on feedback harvest strategies or attempt to develop an consensus management procedure, but rather to provide a quantitative basis for short term and long term decision-making by framing differences and the consequences of those differences.

The components of MSE are [35,36]: (1) specifying clear management objectives; (2) developing quantifiable performance measures for each objective; (3) identifying alternative management strategies or decision options; (4) evaluating (using quantitative performance measures) the performance of each strategy or option against the range of objectives, taking suitable account of uncertainty; and (5) communicating the results to decision-makers.

Experience implementing MSE [35] showed that managers had a variety of concerns about MSE including (1) the requirement that objectives are explicit and measurable, (2) decision rules are made explicit, (3) the origin of the approach was in the scientific community, (4) single answers are not given, and (5) the method is complicated. However, managers ultimately warmed to MSE because it allows a formal approach to management by objectives, fits well with partnerships, and forces one to deal honestly with uncertainties.

Furthermore, industry were willing partners in MSE [35], particularly through industry-based vessel surveys of relative abundance, by industry making contributions to the major sources of uncertainty and contributions of environmental variability. As with managers, members of industry had to understand new kinds of operating and organizational procedures; they also recognized that the technical details of Bayesian analysis are very difficult to follow. Involving an industry-funded scientist in the MSE team lead to industry acceptance of the MSE process. The industry-funded scientist helped design and implement the survey program, could articulate the views of industry, and interpreted scientific debate to industry colleagues. Often (although increasingly less so) environmental agencies and non-governmental organizations (NGOs) have been unfamiliar with methods of fishery assessment and tend to distrust them. However, by involving members of NGOs and environmental agencies in the MSE team, the overall process maintained wider public support for assessment and management. Fishery biologists who were not MSE analysts had to cede some power and autonomy but "most feel that the more open process has many benefits" ([35], p. 975). These include dealing directly with industry and gaining useful biological information through industry–scientist interactions, having better industry support for stock assessment outcomes, having members of NGOs take pressure off scientists for having to be 'advocates for the stock', being able to discuss and evaluate alternative management approaches, and better targeted research.

In [35], the authors emphasize that the technical details of the MSE process are much less important than thorough examination of the data, discussion of multiple working hypotheses, and discussion of methods for weighing alternative scenarios and for selection of the base case scenario. They also emphasize that MSE scientists must be open to non-standard hypotheses and non-traditional ways of translating these into the formal process of quantitative assessment. It is especially important for them to be critically aware that something is always missed and that no matter how thorough one is, fundamental uncertainty remains.

Thus MSE is the prospective evaluation of the management procedures, which includes fishing controls, monitoring, and decision rules for altering fishing controls or monitoring, with the goals of determining those that satisfy the performance criteria [36–39] (Fig. 1).

Most importantly, MSE relies on wide-ranging simulation testing of the management process using performance measures derived from operational objectives rather than ecological experiments. MSE involves selecting (operational) management objectives, specifying performance measures, specifying alternative management strategies, and evaluating these using simulation methods, which become more and more powerful each year. The MSE framework emphasizes the identification and modeling of uncertainties, and propagates these through to their effects on the performance measures.

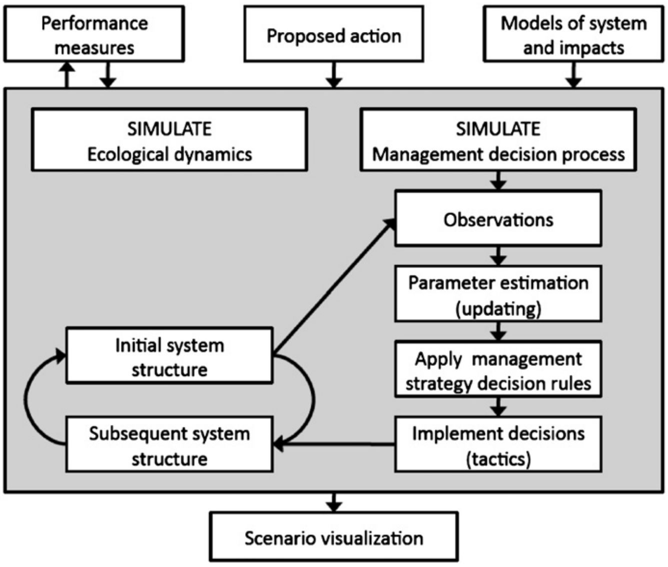

Fig. 1. The general framework for Management Strategy Evaluation (MSE), based on [36].

Indeed, it is often the lack of any management framework that causes problems, rather than a particular management framework. In MSE, the evaluation of management procedures before their implementation provides the opportunity to eliminate management options that would fail to meet the objectives, thus potentially preventing a trial-and-error approach. Such evaluation is more likely to lead to the implementation of management procedures that achieve the objectives despite uncertainties in the various parts of the system, including the limitations of a monitoring program, such as incomplete data and low power in assessments. It can also be used to ensure that the costs of management are commensurate with the value of the fishery. Such forward projection is essential if we are to be able to evaluate short-term costs and long-term benefits. Furthermore, prospective management allows the exploration consequences of possible management procedures, including understanding the range of dynamics of the stock, and providing a framework for dealing with discarding, incidental take, habitat damage, and food-web effects.

In the analysis of the connection between science and policy, the ultimate technical result is that the problem is not a technical one but rather a social one [40]. Furthermore, regardless of the technical or scientific considerations one must be pragmatic:

> Being pragmatic does not mean the rejection of rules or principles in favor of ad hoc decision making or raw intuition. Rather, it means a rejection of the view that rules, in and of themselves, dictate outcomes…. Hard policy decisions can't be programmed into a spreadsheet …. But we also need an analytic framework to help structure the process of making environmental decisions …. Rather than rigid rules or mechanical techniques, we need a framework that leaves us open to the unique attributes of each case, without losing track of our more general normative commitments [41], p. 17

MSE allows decision makers to be pragmatic and for a variety of different groups to be involved in the consideration of assessment advice to have the same access to information [35]. This is essential because effective management must be transparent and have clear and multiple metrics [42]. Management must be accountable (explicit about decision criteria), legitimate (explicit about policy strategies), and flexible (explicit about uncertainty) and must deal carefully with issues concerning burden of proof [43]. MSE allows this to be done.

Three applications of MSE include [36] benthic habitats and community composition (the adaptive management experiment described above), incidental take of species of high conservation value, and food-web interactions. It is concluded that MSE forces the clarification of objectives, the articulation of trade-offs, and the balancing of different views about the dynamics of resources and ecological dependencies and interactions. Doing so helps lead to agreement on management and monitoring measures in the face of uncertainties and clarifies the meaning of a precautionary approach and of adaptive management.

4. Reclaiming scientific inference through ecological detection

Quantitative methods are the grammar of science [44] that allows understanding to be articulated. Nearly 80 years ago, Sir Harold Jeffreys introduced the notion of scientific inference [45] by recognizing that all 'laws of nature' are statements about the probabilities of outcomes given inputs. Jeffreys used the notion of 'inverse probability' as a means of assessing these outcomes. Today inverse probability is more commonly called Bayesian analysis. The development of the cult of statistical significance

[31] deflected the attention of scientists and statisticians from understanding the natural world to focus on hypothesis testing. MSE is a first step towards reclaiming scientific inference, but not the whole way.

Scientific inference possible when experiment is not possible [46] because scientific inference can rely on observation as well as experiment. The key is that there must be a range of variation in inputs and outputs. In seeking 'laws' that are probabilistic statements about outcomes being more or less likely we often provide approximate answers that can vastly improve the quality of public discourse [47].

Ecological detection [48] is a form of scientific inference. In the case of Steller sea lions, the purpose of ecological detection was to confront each putative mechanism of the decline with the data and allow the data to arbitrate between the different models on a level playing field [13]. Ecological detection recognizes that understanding of the world will always be incomplete and that the goal should be to achieve the best understanding possible given the available data. In many ecological situations, it is essential to understand the role of multiple mechanisms, rather than trying to explain the entire phenomenon with any single hypothesis.

There are two reasons why it is so important to consider multiple hypotheses simultaneously. First, if ecological science teaches anything it is that multiple mechanisms are almost always at work: it is very common for more than one mechanism to be important in a system, and the nature of any interactions between different effects will not be captured if they are considered individually. Second, one can often find spurious support for any single hypothesis when it is considered in isolation. The only solution is to jointly solve for the strengths of all hypotheses at the same time, so that the dominant ones may emerge and leave little for the spurious ones to explain. Underlying this, and key to success, is the ability to find the means to express all of the hypotheses within a common ecological currency [13].

The concept of 'adaptive monitoring' has recently been proposed as a new paradigm [49]. This paradigm is based on the notion of conducting work that (1) addresses a well-defined and tractable set of questions specified before the commencement of the monitoring program; (2) is underpinned by a rigorous statistical design; (3) is based on a conceptual model of how the ecosystem might work or how the components of the ecosystem that are targets for monitoring might function, and (4) is driven by human need to know about the ecosystem. Adaptive monitoring is another form of scientific inference.,

5. Steller sea lions in the Bering Sea Redux

A comparison of the population growth rates at individual Steller sea lion rookeries with the amount of commercial fishing activity that occurred within 20 km of each rookery found an association between the decline and commercial fishing activity [50]. Because there were no controls "a conclusive cause and effect relationship between fishing and SSL population trend cannot be determined" (p. 715), but there was strong support for a negative effect of fishing activities on sea lions before 1991. However, using the same data, when the hypotheses describing the effects of fishing activities on pup recruitment and non-pup survival were allowed to compete against the other eight hypotheses (Table 1 here), there was little statistical support for the fishery effects [13]. The complete disappearance of support between the single-hypothesis study and the multiple hypothesis study suggests that the other factors fit the observed pattern more closely. Thus, one can surmise that an experiment with fishing mortality that did not include other factors might give a conclusion similar to the single hypothesis study, but could miss the true factors.

For situations such as this, what is needed are a model with mechanism and use of fine scale rather than aggregated data. In the case of the Steller sea lion model, when any of the 10 hypotheses are considered in isolation (with the other nine effects set to zero), the Maximum Likelihood Estimate (MLE) of the relevant parameter is non-zero. In other words, all 10 of the hypotheses appear to fit the data when they are tested individually. But when they are all tested simultaneously, with the model searching for the joint maximum likelihood solution of all 10 parameters, all but the four listed in Table 1 have their MLE's at 0, indicating that they have no effect. This observation serves as a reminder of the fact that likelihood is relative, and the apparent fit of any single hypothesis to the data does not rule out the possibility that other hypotheses may exist that have an even better fit. For example, in [13] the authors were not able to test hypotheses related to diseases, other pathogens, or negative interactions between sea lions and small boats (those lacking observers).

Other recent empirical and modeling studies complement the conclusions of the multi-hypothesis study. For example, it is proposed that a reasonable annual strategy for Steller sea lions is to forage on herring aggregations in the winter, spawning aggregations of forage fish in spring, salmon in summer and autumn, and pollock and hake throughout the year [51]. The frequency of occurrence of herring and Pollock in the diet of Steller sea lions did not differ significantly between seasons, although that of other prey species did [52]. In [53] it is concluded that Steller sea lions in the Northern Gulf of Alaska (the eastern population) are not food limited during summer months. A bioenergetics model used to investigate the hypothesis that competition between the Western population of Steller sea lions and the commercial ground-fishery for walleye pollock gave rise to the decline of the population [54] predicted population trends of Steller sea lions under a variety of scenarios of continued pollock harvest. Annual energy budgets were enumerated for Steller sea lions in the Gulf of Alaska, and compared with predicted available energy from pollock under various harvest scenarios. In particular, none of the fisheries management regimes that were simulated produced an energy deficit sufficient to account for the decline in the Western stock of Steller sea lions. However, the authors caution that if fishery–sea lion interactions are manifested at the level of the individual, then the data do not exist to allow extrapolation to the population. It is difficult to imagine an experiment that could resolve this question.

Three models for instantaneous rate of change of sea lion population at breeding rookeries with respect to fish abundance and longline fishing effort showed a negative association between pollock density and trends in abundance of Steller sea lions and that an order of magnitude (base e) decline in pollock is estimated to increase Steller sea lions abundance by about 1–2% [55], consistent with the conclusion [13] that it is too much pollock, in a relative sense, that is responsible for part of the decline. There were also statistically significant positive relationships between Steller sea lion abundance and commercial trawl effort but a negative relationship with longline effort [55]. The authors note "However, regression models are incapable of discerning whether walleye pollock density is a driving variable or merely correlated with some other environmental variable that is the driving force behind the Steller sea lion decline" ([55] p. 117). Regarding the possibility of an experiment, they conclude "If our results are correct, a manipulative experiment, for instance, using the longline fishery would need to compare a sixfold difference in fishing effort to see less than a 0.02 change in λ [per capita growth rate of the population]. Hence, any effective, manipulative experiment would need to test treatment extremes in order to have a chance of detecting impacts on Steller sea lion population trends.'' ([55], p. 117–118).

The key to successful learning is having contrast with which hypotheses can be assessed through scientific inference. In a large scale system such as the Bering Sea natural variation provides opportunities for such contrast. For example, to assess hypothesis H10 in Table 1, one would identify rookeries with high and low numbers of harbor seals (regardless of the number of sea lions). The prediction of H10 is that the per-capita attack rate of killer whales on sea lions will be higher around rookeries where harbor seal densities are low. Careful monitoring of killer whale attack rates would provide a natural test of the killer whale hypothesis. If the prediction is not confirmed (that is, if low harbor seal numbers are associated with declining sea lions, but not with elevated killer whale attack rates), then there must be some other factor to explain the observed correspondence between low harbor seals and declining sea lions. It is already well known (www.cf.adfg.state.ak.us/geninfo/finfish/grndfish/pollock/pollockhome.php) that there is wide variation of pollock in the Bering Sea. For example, in the eastern Bering Sea, the biomass of pollock is at a moderate level relative to those 30 years ago, and appears to be decreasing. In the Gulf of Alaska, the biomass of pollock is relatively low but is considered to be increasing. Localized depletion may occur due to commercial fishing – allowing variation on which scientific inference can be built – and pollock have the ability to repopulate areas that are locally depleted. This natural variation allows scientific inference in the absence of experiment since the finer spatial scale data are likely to provide sufficient variation in conditions, which is crucial for learning [56]. Sea lion vital rates could be monitored in the same areas and the model in [13] adapted to simulate forward and suggest a time scale over which results might be expected to appear.

6. Conclusions

This is an era of limited resources. To begin, there is limited political will, in which policy decisions emerge from politics, judgement and debate, rather than from empirical analysis or modeling [27] and policy is a tortuous process [57]. Fiscal resources are also limited and one challenge is to achieve the greatest gain for the conservation dollar [58,59]. The adaptive management experiment in northwestern Australia (a much simpler system than the Bering Sea) showed that even after a considerable period of time the odds of one hypothesis rather than the other was 2:1. As suggested by the Australian experiment and the difficulty of interpreting various controlled experiments on Steller sea lions in aquaria (e.g. [60]) the likelihood of an adaptive management experiment in the Bering Sea producing a conclusive result – at great political, social, and economic cost – is slight. What could be done, then?

First, the situation of Steller sea lions is an example of the more general principle that nature is always conducting ''Monte Carlo experiments'' in sustainability through natural variation leading to the dynamics of selective extinction and speciation [2]. As Darwin so aptly demonstrated, careful observation allows scientific inference in situations for which experiments are virtually impossible. Second, in Alaska, management regulations already include EBFM measures such as control of directed and incidental catches and a prohibition on fishing of forage species [61]. Third, it is essential to address the evidence concerning multiple hypotheses while using data at the finest resolution possible. Advances in understanding occur by learning through experience. Combining mechanistic models and fine scale data, rather than just statistical tests, will allow scientific inference about Steller sea lions in the Bering Sea and Aleutian Islands in cases for which experiment is not practicable.

Acknowledgments

This essay was inspired by discussions that I had at the Fenner Conference on the Environment at the Australian Academy of Sciences in March 2009; I thank Hugh Possingham and Ted Lefroy for the invitation to open the Conference. I thank Keith Sainsbury for general discussions on the topic and Nick Wolf for discussions over the years on the nature of adaptive management in the Bering Sea. Special thanks to Gia Kobara for help with manuscript preparation and submission.

References

[1] Pikitch EK, Santora C, Babcock EA, Bakun A, Bonfil R, Conover DO, et al. Ecosystem-based fishery management. Science 2004;305:346–7.

[2] Fowler CW. Systematic management. Sustainable human interactions with ecosystems and the biosphere. New York: Oxford University Press; 2009.

[3] Richerson K, Levin PS, Mangel M. Accounting for indirect effects and non-commensurate values in ecosystem based fishery management. Marine Policy 2009;34:114–9.

[4] Dayton PK. Reversal of the burden of proof in fisheries management. Science 1998;279:821–3.

[5] Hilborn R. Managing fisheries is managing people: what has been learned? Fish and Fisheries 2007;8:285–96.

[6] Clark JS. Models for ecological data. Princeton: Princeton University Press; 2007.

[7] McCarthy MA. Bayesian methods for ecology. Cambridge: Cambridge University Press; 2007.

[8] Mangel M, Clark CW. Dynamic state variable modeling in ecology. Princeton: Princeton University Press; 1988.

[9] Clark CW, Mangel M. Dynamic state variable models in ecology. Methods and applications. New York: Oxford University Press; 2000.

[10] McCarthy MA, Possingham HP. Adaptive management for conservation. Conservation Biology 2007;21:956–63.

[11] Rout TM, Hauser CE, Possingham HP. Optimal adaptive management for the translocation of a threatened species. Ecological Applications 2009;19:515–26.

[12] Hajkowicz S, Higgins A, Miller C, Marinoni O. Is getting a conservation model used more important than getting it accurate? Biological Conservation 2009;142:699–700.

[13] Wolf N, Mangel M. Multiple hypothesis testing and the declining-population paradigm in Steller sea lions. Ecological Applications 2008;18:1932–55.

[14] Springer AM, Estes JA, van Vliet GB, Williams TM, Doak DF, Danner E, et al. Sequential megafaunal collapse in the North Pacific Ocean: an ongoing legacy of industrial whaling? Proceedings of the National Academy of Sciences 2003;100:12223–8.

[15] Springer AM, Estes JA, van Vliet GB, Williams TM, Doak DF, Danner EM, et al. Mammal-eating killer whales, industrial whaling, and the sequential megafaunal collapse in the North Pacific Ocean: a reply to critics of Springer et al. 2003. Marine Mammal Science 2008;24:414–42.

[16] NRC. Decline of the Steller sea lion in Alaskan waters. Untangling food webs and fishing nets. Washington: National Academies Press; 2003.

[17] APA (At-sea Processors Association). Action Plan for Meeting the Conditions for Continued Certification of the Bering Sea/Aleutian Islands (BS/AI) Pollock Fishery, 2005. Available at 〈http://www.msc.org/track-a-fishery/certified/pacific/bsai-pollock/assessment-downloads-1/BSAI_-Pollock_Action_Plan_Final_Feb05.pdf〉.

[18] Pullin AS, Knight TM. Doing more good than harm – building and evidence – base for conservation and environmental management. Biological Conservation 2009;142:931–4.

[19] Walters CJ, Hilborn R. Adaptive control of fishing systems. Journal of the Fisheries Research Board of Canada 1976;33:145–59.

[20] Walters CJ, Hilborn R. Ecological optimization and adaptive management. Annual Review of Ecology and Systematics 1978;9:157–88.

[21] Walters CJ. Adaptive management of renewable resources. Caldwell, NJ: Blackburn Press; 1986.

[22] Richards SE, Possingham HP, Tizard J. Optimal fire management for maintaining community diversity. Ecological Applications 1999;9:880–92.

[23] Mangel M, Clark CW. Uncertainty, search, and information in fisheries. ICES Journal of Marine Science 1983;41:93–103.

[24] Mangel M, Beder JH. Search and stock depletion: theory and applications. Canadian Journal of Fisheries and Aquatic Sciences 1985;42:150–63.

[25] Shea K, Amarasekare P, Mangel M, Moore J, Murdoch WW, Noonburg E, et al. Management of populations in conservation, harvesting and control. Trends in Ecology and Evolution 1998;13:337–71.

[26] Levin PS, Fogarty MJ, Murawski SA, Fluharty D. Integrated ecosystem assessments: developing the scientific basis for ecosystem-based management of the ocean. PLoS Biology 2009;7:23–8.

[27] Head BW. Three lenses of evidence-based policy. The Australian Journal of Public Administration 2008;67:1–11.

[28] Sainsbury KJ. The ecological basis of multispecies fisheries, and management of a demersal fishery in tropical Australia. In: Gulland JA, editor. Fish population dynamics, 2nd ed. London: John Wiley and Sons; 1988. p. 349–82.

[29] Sainsbury KJ, Campbell RA, Lindholm R, Whitelaw AW. Experimental management of an Australian multispecies fishery: examining the possibility of trawl induced habitat modification. In: Pikitch EL, Huppert DD, Sissenwine MP, editors. Global trends: fisheries management, vol. 20. Bethesda: American Fisheries Society Symposium; 1997. p. 107–12.

[30] Hill SL, Watters GM, Punt AE, McAllister MK, Le Quéré C, Turner J. Model uncertainty in the ecosystem approach to fisheries. Fish and Fisheries 2007;8:315–36.

[31] Ziliak ST, McCloskey DN. The cult of statistical significance. How the standard error costs us jobs, justice, and lives. Ann Arbor: The University of Michigan Press; 2008.

[32] North Pacific Fishery Management Council. Fishery Management Plan for Groundfish of the Bering Sea and Aleutian Islands Management Area, 2009. Available at 〈http://www.fakr.noaa.gov/npfmc/fmp/bsai/BSAI.pdf〉.

[33] Goodman D, Mangel M, Parkes G, Quinn T, Restrepo V, Smith T, et al.. Scientific review of the harvest strategy currently used in the BSAI and GOA Groundfish Fishery Management Plans, 2002. Report to the North Pacific Fishery Management Council. Available at 〈http://users.soe.ucsc.edu/~msmangel/art_bookchap1.html〉.

[34] Ralston S. West Coast groundfish harvest policy. North American Journal of Fisheries Management 2002;22:249–50.

[35] Smith ADM, Sainsbury K, Stevens R. Implementing effective fisheries-management systems-management strategy evaluation and the Australian partnership approach. ICES Journal of Marine Science 1999;56:967–79.

[36] Sainsbury KJ, Punt A, Smith ADM. Design of operational management strategies for achieving fishery ecosystem objectives. ICES Journal of Marine Science 2000;57:731–41.

[37] de la Mare WK. Tidier fisheries management requires a new MOP (management-oriented paradigm). Reviews in Fish Biology and Fisheries 1998;8:349–56.

[38] Cooke JG. Improvement of fishery-management advice through simulation testing of harvest algorithms. ICES Journal of Marine Science 1999;56:797–810.

[39] Mangel M, Marinovic B, Pomeroy C, Croll D. Requiem for Ricker: unpacking MSY. Bulletin of Marine Science 2002;70:763–81.

[40] Ludwig D. A quantitative precautionary approach. Bulletin of Marine Science 2002;70:485–97.

[41] Farber DA. Eco-pragmatism: making sensible environmental decisions in an uncertain world. Chicago: The University of Chicago Press; 1999.

[42] Brodziak J, Link J. Ecosystem-based fishery management: what is it and how can we do it? Bulletin of Marine Science 2002;70:589–611.

[43] Charles AT. The precautionary approach and 'burden of Proof' challenges in fishery management. Bulletin of Marine Science 2002;70:683–94.

[44] Pearson K. The grammar of science. London: Walter Scott; 1892.

[45] Jeffreys H. Scientific inference. Cambridge: Cambridge University Press; 1931.

[46] Mangel M, Brodziak JKT, DiNardo G. Reproductive ecology and scientific inference of steepness, a fundamental metric of population dynamics and strategic fisheries management. Fish and Fisheries 2009, in press, doi:10.1111/j.1467-2979.2009.00345.x.

[47] Socolow RH, Lam SH. Good enough tools for global warming policy making. Philosophical Transactions of the Royal Society A 2007;365:897–934.

[48] Hilborn R, Mangel M. The ecological detective. Confronting models with data. Princeton: Princeton University Press; 1997.

[49] Lindenmayer DB, Likens GE. Adaptive monitoring: a new paradigm for long-term research and monitoring. Trends in Ecology and Evolutionary Biology 2009;24:482–6.

[50] Hennen D. Association between the Alaska Steller sea lion decline and commercial fisheries. Ecological Applications 2006;16:704–17.

[51] Womble JN, Sigler MF, Willson MF. Linking seasonal distribution patterns with prey availability in a central-place forager, the Steller sea lion. Journal of Biogeography 2009;36:439–51.

[52] Womble JN, Sigler MF. Seasonal availability of abundant, energy-rich prey influences the abundance and diet of a marine predator, the Steller sea lion *Eumetopias juhatus*. Marine Ecology Progress Series 2006;325:281–93.

[53] Adams CF, Foy RJ, Johnson DS, Coyle KO. Seasonal changes in pelagic fish biomass around the Chiswell Island Steller sea lion rookery in 2003. Fisheries Research 2008;93:179–85.

[54] Cornick LA, Neill W, Grant WE. Assessing competition between Steller sea lions and the commercial groundfishery in Alaska: a bioenergetics modelling approach. Ecological Modelling 2006;199:107–14.

[55] Dillingham PW, Skalski JR, Ryding KE. Fine-scale geographic interactions between Steller sea lion (*Eumetopias jubatus*) trends and local fisheries. Canadian Journal of Fisheries and Aquatic Sciences 2006;63:107–19.

[56] Parma AM, Amarasekare P, Mangel M, Moore J, Murdoch WW, Noonburg E, et al. What can adaptive management do for our fish, forests, food, and biodiversity? Integrative Biology 1998;1:16–26.

[57] Rosenberg AA. The precautionary approach in application from a manager's perspective. Bulletin of Marine Science 2002;70:577–88.

[58] Bottrill MC, Joseph LN, Carwardine J, Bode M, Cook C, Game ET, et al. Is conservation triage just smart decision making? Trends in Ecology and Evolution 2008;23:649–54.

[59] MacKenzie DI. Getting the biggest bang for our conservation buck. Trends in Ecology and Evolution 2007;24:1177–177–>1745.

[60] Fritz LW, Hinckley S. A critical review of the regime shift – 'Junk Food' – nutritional stress hypothesis for the decline of the western stock of Steller sea lion. Marine Mammal Science 2005;21:476–518.

[61] Witherell D, Pautzke C, Fluharty D. An ecosystem-based approach for Alaska groundfish fisheries. ICES Journal of Marine Science 2000;57:771–7.