Synthesis

The Challenge of Collecting and Using Environmental Monitoring Data

Eric Biber

ABSTRACT. The monitoring of ambient environmental conditions is essential to environmental management and regulation. However, effective monitoring is subject to a range of institutional, political, and legal constraints, constraints that are a product of the need for monitoring to be continuous, long lived, and well matched to the resources being studied. Political pressure or myopia, conflicting agency goals, the need for institutional autonomy, or a reluctance of agency scientists to pursue monitoring all may make it difficult for ambient monitoring to be effectively undertaken. Even if effective monitoring data is gathered, it may not be used in decision making. The inevitable residual uncertainty in monitoring data allows stakeholders to contest the use of monitoring in decision making. Structural solutions, e.g., the creation of agencies to conduct monitoring separate from management or regulation and prompt use of that data in decision making, may be the most promising solutions.

Key Words: environmental law; monitoring; uncertainty

INTRODUCTION: WHY AMBIENT MONITORING MATTERS

Monitoring of environmental conditions, i.e., ‘ambient monitoring,’ is central to environmental decision making. Monitoring allows the evaluation of the success or failure of management projects. This is particularly true for ‘adaptive management’ in which the effectiveness of various management choices is monitored over time to reduce uncertainty (Holling and Meffe 1996). Monitoring can pinpoint new environmental problems that need regulatory attention, and therefore provide political or legal pressure for the implementation of new standards. It can help set the level of standards for an existing program (Hellawell 1991, Messer 2004). Ambient monitoring is part of regulatory programs under the Clean Air Act and the Clean Water Act.

Can we reliably and effectively obtain the ambient data needed for regulation and management? There are major gaps in our understanding of current environmental conditions (Christensen et al. 1996, O’Malley et al. 2009). These gaps exist because ambient monitoring is difficult to conduct effectively, and that difficulty creates political, legal, and organizational obstacles to the collection and use of monitoring data. Solutions require understanding why and how those obstacles deter the collection and use of ambient monitoring data.

Ambient monitoring includes a wide range of monitoring activities, such as effects and effectiveness monitoring, compliance monitoring, implementation monitoring, and status and trends monitoring. I focus on the concept of overall ambient monitoring, to provide an assessment of the baseline characteristics, obstacles, and possible solutions. Of course in the context of a particular monitoring program, specific political, economic, institutional, and ecological factors will produce certain kinds of incentives for and obstacles to monitoring. Examples of significant patterns associated with important kinds of monitoring are discussed. For instance, geographic and temporal scale will often limit the incentives that private parties have to develop effective monitoring data.

WHY AMBIENT MONITORING IS DIFFICULT

Continuity and longevity

Ambient monitoring requires measuring variables over extended periods of time; it must be both continuous and long-lived to be successful. Continuity matters because data gaps impair assessments of environmental variability that is often shaped by infrequent but significant events (Franklin 1990). The causes of missing data may also bias the results of the monitoring program (Rubin 1976). For instance, flood gauge data might be missing because a particularly large flood event destroyed the gauges (Gunderson 2003a).

Because there is always something special or unique about any individual time period, missing data can be used by critics to challenge the monitoring data and management recommendations based on that data.

Longevity matters because many environmental resources change slowly; subtle trends require significant time for identification (Franklin 1990). Declines in wildlife populations are often measured in a few percentage points annually and may take many years to identify.

Long-term monitoring can distinguish secular trends from temporary fluctuations, which is a difficult task given most environmental resources have high and uncertain variability (Kent 2003). Several years of measuring the atmospheric levels of carbon dioxide were needed to differentiate between seasonal changes and an overall increase in carbon dioxide levels over time (Keeling 1998). Satellite measurements of global environmental systems can require decades to identify long-term cycles and separate them from human impacts (NRC 2007).
Establishing causal linkages between a management or regulatory decision and changes in environmental resources may also require long periods of data collection, especially when an environmental resource is subject to a range of important, but infrequent impacts from other human or natural sources. For instance, decision makers may question whether changes in dam and reservoir management will benefit endangered salmon runs. There may be a causal relationship between management changes and initial improvements in a salmon run. Or the improvements might be the result of factors such as ocean conditions, changes in fisheries management, etc. Multiple years of data collection may be required to separate these factors. Just as with continuity, longevity may be essential to respond to political arguments that data sets exclude important events.

Finally, longevity may be necessary because data takes extended periods of time to collect. For instance, salmon runs occur a few times a year; collecting adequate data will often require many years of observations (Lee 1999).

**Effectiveness**

Effective monitoring requires collecting enough of the right kind of data needed to answer the questions the monitoring program was established to answer (Biber 2011). Reid (2001:815) found that 30% of flawed monitoring programs studied could not provide the kind of information that “was needed to meet the project objectives.” Programs must measure the right variables, at the right scale, and at sufficient levels. Choosing the wrong or misleading variables or indicators can lead agencies to ignore important environmental problems. For instance, wetlands managers and regulators have been criticized for overemphasizing total acreage of wetlands available instead of considering the quality of the wetlands being protected or restored (Dale and Gerlak 2007). This mismatch may occur as a result of the lack of good knowledge about how different variables interact, and therefore whether one variable is a good proxy for another. It may occur because goals have not been well defined, and that lack of clarity may lead to the adoption of inappropriate variables for measurement. Or, it may occur because of political pressure to demonstrate progress in addressing an environmental problem by using variables that are easy to measure or that make it easier to demonstrate progress. This last factor appears to have played a role in the problems of measuring wetland protection and restoration.

A mismatch in scale can drown any signal with large variability in the monitoring data, or alternatively make it impossible to detect important fine-grained variations (Hellawell 1991, Usher 1991). The appropriate scale will depend on the resource being monitored and the questions being asked.

The restoration program for the Chesapeake Bay monitored water quality by measuring the total amount of pollution entering the bay from major watersheds. However, these measurements could not assess the effectiveness of various management or regulatory efforts within a watershed, the key questions for decision makers considering the success or failure of various management or regulatory strategies (Ernst 2003).

The appropriate scale for a monitoring program is difficult to know as a general matter. Inexpert outsiders will therefore rarely be able to judge whether a monitoring program is operating at the correct scale.

Monitoring programs often require a statistically significant difference or correlation to justify management changes (Elzinga et al. 2001, Nichols and Williams 2006). But monitoring programs often collect inadequate data to detect management-relevant statistically significant differences (Usher 1991, Noon 2003). For instance, the National Oceanic and Atmospheric Administration (NOAA) was required by the U.S. Congress to determine whether a certain form of tuna fishing was harming dolphin populations. However, NOAA did not examine enough dolphins to answer the congressionally mandated questions (Earth Island Institute v. Hogarth, 494 F.3d 757 (9th Cir. 2007)). Determining how much to monitor is challenging because we often lack information about the full range of variability for many resources. Questions about statistical significance again require detailed understanding that makes them difficult for nonexperts to assess.

Longevity, continuity, scale, and statistical significance mean that ambient monitoring can be quite costly. Monitoring of the restoration of riparian habitat in one small creek cost “as much as actual construction of the habitat improvements and [required] three years of data for statistically reliable confirmation” (Lee and Lawrence 1986:447). The dynamism and variable scales of environmental resources also mean that even effective ambient environmental monitoring will have residual uncertainty.

**INSTITUTIONAL, POLITICAL, AND LEGAL OBSTACLES TO EFFECTIVE AMBIENT MONITORING AND ITS USE**

Ambient monitoring requires continuity, consistency, adequate scale, and significant expertise and effort. These characteristics in turn create three main challenges to the pursuit of effective ambient monitoring by public agencies: the need for institutional continuity, the relative opacity of assessing whether monitoring is effective or not, and the requirement of matching the jurisdictional scale of the agencies conducting monitoring with the optimal scale for monitoring important resources. Even if effective data is collected, agencies may not use it to make management or regulatory decisions, in large part because of residual uncertainty.
Scale
The geographic scale of the agency conducting the monitoring program may not match the optimal scale for monitoring, leaving the agency unable or unwilling to conduct effective monitoring. Even an agency with authority to engage in extraterritorial monitoring may have little incentive to monitor, or to do so well (Buzbee 2008). If monitoring data is not directly relevant for the agency’s own decision making, then data collection provides little benefit for the agency. A water quality agency may have little desire to monitor downstream water quality that it has no responsibility to protect. Even extraterritorial monitoring information that is useful for agency decision making may still not receive adequate investment: this information may be useful for other decision makers as well, but the benefits to those decision makers will be ignored in the agency’s own cost-benefit calculus.

For similar reasons, private entities will often have weak incentives to conduct effective ambient monitoring. Companies may have economic or publicity reasons to demonstrate that their actions are not harming the environment and that they are operating within the existing regulatory permits, i.e., what is often called ‘compliance monitoring.’ However, monitoring of ambient conditions usually will have limited benefits for the company, because the ambient condition of the environment will often be influenced by other human and natural factors outside of the company’s control (Biber 2011).

An example of how private parties will generally only focus on their impacts on the environment, rather than on the overall quality of the environment, involves Plum Creek, a Pacific Northwest timber company, which owned a substantial amount of forest containing bull trout populations, a federally listed endangered species. An agreement with the U.S. Fish and Wildlife Service allowed timber activities to proceed on the condition that efforts were made to mitigate the impact of logging on bull trout habitat. Plum Creek was required to monitor the status and trends of the bull trout. But even though the bull trout’s habitat crosses property lines, Plum Creek’s monitoring efforts were restricted to bull trout populations and habitat within its property borders. This monitoring would be effective in demonstrating that Plum Creek’s mitigation efforts would benefit the species. In contrast, monitoring outside its property borders would be much less effective in showing Plum Creek’s success, because there the status of bull trout would more likely be affected by the activities of other landowners. However, the status of the species as a whole is best determined by monitoring across the entire range of the species.

Problems can also occur when agency jurisdictional borders are substantive, rather than geographic. Effective monitoring of one resource, e.g., an endangered species, might require understanding and monitoring another resource, e.g., water quality. But even if the management agency tasked with one resource has the authority to monitor the second resource, the benefits of monitoring water quality, for example, will spill over and benefit another decision maker, e.g., a water quality agency. Thus, the management agency may not put enough resources into monitoring water quality (Sanchirico et al., in press).

Myopia and political pressure
Budgets limit monitoring programs (Spellerberg 2005) because monitoring is frequently portrayed as unnecessary information collection (Doremus 2008). One could increase monitoring’s relevance for policymaking (Lee 1993), but that creates other risks. The politics of environmental decision making are shaped by the differential organizational capacities of those who benefit from and those who pay for environmental protection: the costs of most environmental regulation tend to fall heaviest on a relatively small group of individuals or corporations; the benefits of environmental regulation, usually a public good, tend to be spread widely among a large number of individuals. The supporters of stronger environmental regulation accordingly face significantly greater organizational challenges compared to the opponents (Zinn 2002), particularly for the ongoing implementation of environmental statutes by agencies. Implementation is essential but low profile compared to the enactment of new laws (Farber 1999).

Monitoring is an exemplar of this dynamic. It may be relatively simple to diagnose whether any monitoring is occurring, but often quite difficult for nonexperts to diagnose the effectiveness of a monitoring program. Because effective monitoring requires continuity, outside parties must conduct ongoing supervision of agency monitoring programs. And because the legal or political burden of demonstrating the need for additional regulation is often on the agency, the lack of effective monitoring will usually benefit regulated parties. Politically meaningful monitoring is therefore vulnerable to asymmetric political pressure (Biber 2011).

The Fish Passage Center (FPC) was established to accurately assess the number of salmon that successfully migrate through the various hydroelectric facilities on the Columbia and Snake rivers in the Pacific Northwest. Those facilities are a major factor in the collapse of salmon runs. The FPC reported that increasing water flow through the Snake River facilities might improve the survival of juvenile salmon, albeit at the cost of reduced hydroelectric power and navigation. In response, a U.S. senator from Idaho pushed legislation to eliminate FPC (Bhattacharjee 2005). The FPC was later saved through litigation.

Political pressure and the power of the status quo
Political pressure might not just prevent the collection of monitoring data. It may also prevent an agency from using
effective monitoring data to implement management or regulatory changes that threaten important interest groups, in large part because of residual uncertainty in monitoring data. The Canadian cod fishery shows the importance of political pressure in shaping how monitoring data is used. Throughout the 1980s and 1990s critics raised serious questions about the adequacy of the monitoring program for the cod fishery. These critics questioned whether the fishery was truly rebounding, as the government claimed, and pointed to contrary data (Neis 1992, Hutchings et al. 1997). Eventually a sudden and dramatic population collapse forced the closure of the fishery. Most fisheries scientists concluded that overfishing was primarily responsible for the collapse, and that government scientists and policymakers failed to collect, downplayed, or ignored important monitoring data because of political pressure (Hutchings and Myers 1994).

Some monitoring data cannot be ignored. The cod fishery was closed when there was no longer any doubt that the cod was commercially extinct. But usually monitoring data has inherent uncertainty (Lee 1993, Moir and Block 2001), because of environmental complexity and dynamism (Moir and Block 2001) and the large spatial and temporal scales of many ecological questions (Hilborn and Ludwig 1993). Indeed, because there can be multiple methods for analyzing monitoring data, and there is no consensus on which methods are superior in many situations (Thomas 1996), even disputes over how best to analyze monitoring data may not be resolvable. Monitoring may not be able to answer key environmental policy questions (Roe 1996, Bormann and Keister 2004).

California’s sardine fishery collapsed in the mid-20th century. California’s fisheries agency had produced data that it concluded showed that overfishing was a serious risk. Using the exact same data, federal fisheries scientists concluded that the fishery was safe from overfishing and at no risk of collapse. Indeed, one state and one federal scientist coauthored an article on the status of the sardine fishery, relying on the same data, and each reached diametrically opposite conclusions (Radovich 1982).

Powerful political actors opposed to management changes can use residual uncertainty to argue that monitoring data does not require management changes, even when monitoring data might be considered effective by outside observers (Halbert 1993, Gunderson 2003b). These actors have at least four advantages. First, inaction can be made more appealing by arguing that additional research is required to narrow uncertainty (Walters 1997, Nichols and Williams 2006, Allen and Gunderson 2011). Second, statistical analysis in monitoring programs often gives a significant advantage to parties supporting the status quo. Because of the importance of uncertainty, decisions about the statistical burden of proof, how to resolve questions about whether monitoring data indicates that management should be changed, are often outcome determinative (Martin et al. 2009). Many resource management agencies have emphasized the risk of type I errors, i.e., “approving an action, taken to be sound (true) based on the best science, that later proves to be unsound (false),” even though type II errors, the opposite dynamic, may be just as important (Stankey et al. 2005:27). This risk aversion makes it hard to change management practices. Third, when the status quo involves significant economic exploitation of a resource (often), decision makers tend to minimize information that supports reducing exploitation. In fisheries management, when scientists give decision makers a range of possible quotas for fishing, decision makers regularly and consistently set fishing quotas as high as possible within that band of uncertainty (Eagle and Thompson 2003). This pressure to resolve uncertainty in favor of economic development or exploitation derives from a number of sources, such as the advantage that regulated interests often have in the political arena (Zinn 2002). Another important factor is the need among economic interests for regulatory certainty (Ruhl 2008). Certainty will create strong political resistance to changes in how much an environmental resource can be exploited. Fourth, political and legal inertia often prevent changes to existing management and regulatory systems (Lazarus 2009). An agency may simply not have enough legal authority, money, or time to adequately review or analyze effective monitoring data, or to implement management changes based on that data (Gregory et al. 2006, Doremus 2011). Inertia may make it difficult to change the statutory or regulatory structures that inhibit agency action. Overcoming that inertia requires political capital, time, and energy.

The adaptive management program for the Glen Canyon Dam sought to address significant downstream impacts on Grand Canyon National Park. A series of experimental floods were designed to test whether changes to dam operations might improve downstream conditions. The scientific consensus was that the experimental floods were successful and supported significant changes in dam operations. However, only 15 years after the adaptive management program began were any changes proposed by the management agency. Critics argued that this delay occurred because power and water interests hurt most by changes had an effective veto over decision making (Camacho 2008, Feller 2008). These powerful interests relied on residual uncertainty to argue against operational changes (Susskind et al. 2010). They also drew on burden of proof: staff members for the Western Area Power Administration, a major power interest, have called for setting a high burden of proof for any changes from the status quo (Wieringa and Morton 1996).

**Residual uncertainty fueling ongoing stakeholder battles**

Even without a dominant political player, uncertainty may allow various stakeholders to continue policy disputes after the monitoring program has been established. Most natural
resource debates are ‘wicked problems’ that involve fundamental conflicts over goals and objectives, with a wide range of goals that are politically and legally permissible (Rittel and Webber 1973, McLain and Lee 1996, Grumbine 1997).

Clear articulation of goals is essential for successful monitoring and adaptive management (Gregory et al. 2006, Williams 2011) for three reasons: first, goals help determine what the important management or regulatory questions are, and therefore how to design a monitoring program to reduce the relevant uncertainty in an efficient manner (Nichols and Williams 2006, Doremus 2011, Runge et al. 2011); second, goals help determine whether the costs of monitoring are outweighed by the benefits; and third, goals allow monitoring to support evaluation of success or failure for various management options.

A failure to resolve underlying controversy can therefore prevent monitoring from producing significant management or regulatory changes. Stakeholders still in conflict over underlying goals can point to residual uncertainty to resist unfavorable decisions (McLain and Lee 1996, Gregory et al. 2006). The Glen Canyon Dam adaptive management program again provides an example: because the U.S. Congress never provided clear guidance among conflicting goals for the management of the dam, i.e., water storage, power generation, recreational use, protection of downstream Grand Canyon resources, and protection of endangered species, it is difficult to resolve underlying uncertainty by determining what risks are more important to consider in making management changes (Zellmer and Gunderson 2009).

**Conflicting agency goals**

Agencies might be reluctant to monitor because monitoring might produce accountability for performance on environmental goals. Critics of the Chesapeake Bay restoration program argued that it overemphasized models at the expense of monitoring, both because the models were cheaper, and because the models provided more positive performance assessments (Ernst 2003). Private parties may be reluctant to conduct effective monitoring because it may demonstrate that regulatory standards require changes, reducing regulatory certainty.

The hard-to-measure nature of many environmental goals may make agencies particularly wary of monitoring ambient environmental conditions. Agencies are often assigned multiple conflicting goals; the more measurable goals are usually higher priorities (Biber 2009). As a result, agencies might systematically underperform in conducting environmental monitoring. The opacity of monitoring means that monitoring is hard to measure in terms of quality. Improvements in ambient monitoring may conflict with other agency goals: monitoring uses scarce budgetary resources, and improved monitoring might produce political or legal pressure to perform better on underperforming environmental goals that are themselves in tension with other goals.

Monitoring may be downplayed even more when monitoring or environmental performance conflict with an agency’s mission. Because of the vague goals often given to public agencies, strong performance-based incentives are usually unavailable to motivate agency employees. Public agency managers therefore might orient the agency around a ‘mission’ that employees are committed to achieving without strong incentives (Simon 1976, Wilson 1989, Besley and Ghatak 2005). Improved monitoring of environmental resources may produce information that might conflict with that mission, or use resources without advancing that mission.

The U.S. Forest Service created a ‘survey and manage’ program to determine the status of hundreds of rare, old-growth dependent species in the Pacific Northwest. The program was generally successful in providing new information about little-known species, but also was expensive and controversial. In response to timber company litigation, the U.S. Forest Service ended the program, but litigation by environmental groups eventually restored it (Marcot and Molina 2006, Molina et al. 2006). In justifying its termination of the program, the U.S. Forest Service emphasized that surveys were interfering with timber sales (U.S. Forest Service 2007), consistent with a historic orientation by the agency toward timber production (Biber 2009).

Just as an agency may be reluctant to collect monitoring data because that data might undermine various agency goals, it may also be reluctant to use monitoring data to change agency management or regulatory decisions because those changes would conflict with agency goals, or because changes would require admitting mistakes (Imperial 1999, Lee 1999, Ascher 2001, Doremus 2001).

**Importance of agency autonomy**

Less information often gives agencies political or legal leeway. Agencies use various tools to ‘stretch’ incomplete monitoring data such as indicators, proxies, extrapolation, and modeling. All of these tools require underlying assumptions that allow an agency to bury important policy conclusions. In addition, courts provide significant deference to all of these tools (Wagner et al. 2010). Although more data might improve accuracy, more data also might constrain an agency’s discretion. Agencies highly value discretion (Wilson 1989).

A U.S. Forest Service regulation required the agency to monitor indicator wildlife species populations to ensure that management activities were protecting overall species diversity and ecosystem health (Corbin 1999). Over time, the agency adapted the regulatory requirements to increase its discretion. The U.S. Forest Service interpreted those regulations as allowing it to measure suitable habitat for the relevant wildlife species and then extrapolate from habitat to
the status of the species themselves, i.e., the ‘proxy-on-proxy’ approach. This approach increased the agency’s freedom to interpret population trends and reduced the power of the data to constrain management decisions (Corbin 1999, Elzinga et al. 2001). After some court decisions rejected the proxy-on-proxy approach, the U.S. Forest Service eliminated the regulations (Biber 2011).

Professional culture of scientists
An agency’s culture may be inhospitable to monitoring because of the reluctance to conduct routine monitoring among agency scientists (Nisbet 2007, Doremus 2008). Scientists are essential to the operations of leading environmental agencies (Biber 2011). But long-term monitoring projects are often not rewarded with publication, grants, tenure, or improved professional reputation. Extended monitoring projects can be longer than the timeframes for personal professional advancement within science.

The professional skepticism of scientists toward monitoring affects how environmental agencies treat monitoring. Fellow agency scientists will askance at heavy investments in monitoring by a colleague, and political scientists have identified peers as a major influence on agency employee attitudes, motivation, and performance (Brehm and Gates 1997). This skepticism can also lead agency decision makers to avoid supporting monitoring activities. The long-term measurements of atmospheric carbon dioxide levels so essential for understanding climate change were repeatedly threatened with funding cuts because funding agencies did not see the monitoring as a scientific endeavor (Keeling 1998).

Even if collected, monitoring data may languish unused because agency scientists might also devalue the analysis and use of monitoring data. Although collection of monitoring data may be devalued because the extended periods needed for monitoring can exceed the professional career of scientists, analysis and use can occur quickly if the data has already been collected. Thus, resistance here may be derivative of the general reluctance to be involved with monitoring data and may be easier to overcome.

Timing
The extended periods of time needed to collect useful monitoring data also means that extended periods of time are required before data can be used for management decisions (Moir and Block 2001). Time lags are problematic if advance planning is required for agency decisions (Moore et al. 2011), or if the decision making agency does not outlive the monitoring program (Williams et al. 2009). Monitored resources may also not show significant, detectable changes until it is too late to prevent those changes; it might be necessary to measure other resources that can predict problems and allow timely action (Brown and Havstad 2004). Statistics can make this third problem worse. Monitoring programs are generally designed to detect statistically significant changes. This puts the burden of detecting negative trends on the monitoring system, exacerbating timing problems when monitoring programs do not detect statistically significant changes until it is too late for an effective response (Field et al. 2004).

POSSIBLE SOLUTIONS FOR THE FAILURE TO COLLECT AND USE EFFECTIVE MONITORING DATA
Solutions must resolve the difficulties of scale, continuity, opacity, and residual uncertainty. Scale requires that the institutional scale of monitoring agencies match the optimal scale of monitoring; but there may be multiple optimal monitoring scales, depending on the questions being asked. Continuity requires commitment and reliability by the agencies conducting the monitoring. Opacity makes it difficult for outside supervision to ensure effective monitoring. Successful supervision demands significant technical expertise and resources. Residual uncertainty requires recognition of how stakeholders can exploit uncertainty and obstruct the use of even effective monitoring data.

Leadership and funding
Many observers have argued that agency leadership will improve monitoring and its use (Folke 2005 et al., Walters 2007). However, great leaders are a highly contingent solution, and successful monitoring requires continuity. Even if strong leadership can be found, the structure and function of agencies are critical to success.

Another commonly proposed solution is more funding (Doremus 2008, Camacho 2009). The challenge is overcoming the political resistance to additional funding. And even with reliable funding, agencies still may not conduct effective monitoring or use data.

Collaboration among agencies
Improved collaboration among monitoring agencies could produce better use of existing information by aggregating data at larger or smaller scales than individual monitoring programs (NRC 1993). However, collaborative efforts may not address the reluctance, whether conscious or not, of agency officials to pursue effective monitoring programs.

Citizen monitoring
Individual citizens can provide monitoring data. For instance, community ‘bucket brigades’ use inexpensive technology to measure air quality (Overdevest and Mayer 2008). However, many monitoring technologies are too expensive for most volunteer groups. Volunteers may implement monitoring protocols improperly. The public or the regulator may be reluctant to use volunteer data because they believe volunteer groups have hidden agendas (O’Rourke and Macey 2003, Overdevest and Mayer 2008). Finally, volunteer organizations

http://www.ecologyandsociety.org/vol18/iss4/art68/
may not be able to maintain long-term monitoring programs because of short institutional lifespans (O’Rourke and Macey 2003).

Collaboration among stakeholders to resolve underlying policy disputes
Another alternative is to resolve fundamental conflicts that interfere with the collection and use of effective monitoring data before a monitoring program is established. Because of residual uncertainty, monitoring data can always be deconstructed, undermining its ability to resolve important policy dilemmas. If the stakes surrounding an environmental dispute are high enough, and if there are underlying disputes over goals that have not been resolved, stakeholders might draw on the inevitable uncertainty in monitoring data to contest its use in decision making.

Monitoring programs will therefore only use data to make decisions when the costs to stakeholders of initiating a conflict over the effectiveness of data exceed the benefits to stakeholders of such a conflict. Costs might exceed benefits when there are no fundamental disagreements over goals. Even when goals are contested, stakeholders might not challenge monitoring data if the use of the data would not have major policy impacts. Finally, costs may be high because the groups or individuals who collected the data are especially respected, or when the standards used to judge the data as effective are generally well accepted by important stakeholders.

The adaptive management and monitoring literature has called for incorporating collaborative processes into monitoring programs to reduce conflict and increase the possibility that monitoring can improve decision making (Lee 1999, Allen et al. 2011). There is promise and peril in this approach. There is promise if the collaborative approach succeeds in truly resolving underlying disputes in goals, or at least in resolving enough of the dispute such that there are useful policy questions for a monitoring program to address. There is peril if the collaborative process does not produce consensus, and stakeholders use the collaborative process to obstruct any policy changes, as perhaps occurred in the Glen Canyon Dam adaptive management program.

Using triggers to help resolve disputes
One specific way to establish a dispute resolution process is through triggers. Under an ex ante trigger mechanism, if certain benchmarks are met in the monitoring data, then various consequences, i.e., additional monitoring, specific management or regulatory choices, and additional stakeholder consultation and discussion, would be automatically implemented (Doremus 2001, Moir and Block 2001, Martin et al. 2009). This solution reduces the ability of stakeholders to ‘shift the goalposts’ as to acceptable uncertainty. If stakeholders must take positions up front about how the monitoring data should change decision making, the public nature of those positions might encourage accountability and reduce obstruction. Triggers also require stakeholders and agencies to articulate acceptable levels of uncertainty, nurturing a constructive dialogue about the role of uncertainty in decision making. Finally, if uncertainty is high enough, parties might negotiate behind a veil of ignorance in which they do not know how future monitoring results might support or undermine their particular goals.

Triggers cannot produce agreements when there are none to be found (Nie and Schultz 2012). Sophisticated stakeholders can manipulate the setting of triggers and may even simply refuse to agree to any triggers. Moreover, it is impossible to plan for all contingencies in environmental decision making.

Finally, a trigger system might have an ‘expiration date.’ If the monitoring program and trigger system require a long enough time to operate, the political, economic, and social landscapes may have changed enough to render triggers unenforceable (Salafsky et al. 2001).

Restructuring agencies
A change in structure of environmental agencies might increase agency incentives to conduct and use effective monitoring. An agency focused primarily on monitoring avoids the problem of conflicting agency goals that can reduce an agency’s incentives to produce effective monitoring data (Doremus 2001, Shapiro and Steinzor 2008). For instance, the U.S. Geological Survey undertakes much federal environmental monitoring, and its other tasks, primarily scientific research, do not directly conflict with that monitoring role (Biber 2011).

With a separate agency, there is the advantage of continuity, because there is a public institution, usually fairly long-lived; there is the advantage of expertise, because the agency primarily focuses on monitoring; there is an institution without conflicting goals; and an institution whose jurisdictional scope can be designed to match the optimal scale for monitoring (Biber 2011).

A large organization that combines monitoring with other goals might, in response to budget cuts, reduce monitoring budgets disproportionately to protect other goals. But if the agency only does monitoring, the agency has to maintain its monitoring budget to ensure its institutional survival. A separate monitoring agency might also develop a reputation as an unbiased provider of information (Biber 2011).

Some examples from large ecological restoration projects indicate that more independence can improve monitoring. Both the restoration efforts for the Colorado River in the Grand Canyon below the Glen Canyon Dam (NRC 2004) and for the Everglades (NRC 2010) have received praise for their monitoring programs. Both have relatively independent monitoring organizations (NRC 2008). By contrast, the monitoring for the Chesapeake Bay Program has been criticized, and its monitoring group is integrated into the

A comparison of fisheries programs is also illuminating. California, Peru, and Canada all had large, economically valuable fisheries that collapsed. In California, regulatory authority was vested with the legislature; the Fish and Game Commission enforced the regulations, and conducted much of the population monitoring (Radovich 1982). The commission’s monitoring identified potential weaknesses in the fishery, and it called for legislative action to reduce fishing pressure, although those warnings were ignored until after the fishery collapsed (Radovich 1982, McEvoy and Scheiber 1984). In Peru, monitoring was conducted by the Instituto del Mare del Peru (IMARPE). Prior to the 1970s, this was an independent agency, separate from the Ministry of Fisheries that actually set fishing quotas (Clark 1976). However, after 1970, IMARPE was merged into the Ministry (Hammergren 1981). The IMARPE’s monitoring data after 1970 was criticized as being deeply flawed, possibly significantly overestimating the size of the anchovy population (Clark 1976), while at the same time, the Ministry of Fisheries sought to increase the fishing fleet (Hammergren 1981). In Canada, monitoring and management for the Atlantic cod fishery were both conducted within the Department of Fisheries and Oceans (DFO; Hutchings et al. 1997). Through the early 1990s, DFO’s assessments were optimistic. Those assessments were primarily based on estimates of the catch from off-shore commercial fishing operations, as well as DFO’s own survey trawls; DFO ignored contrary data from in-shore fishermen, those who fished much closer to shore, using different techniques at smaller scales, (Hutchings and Myers 1994, Hutchings et al. 1997, Finlayson and McCay 1998). Critics noted these weaknesses in the DFO monitoring data at the time, but DFO rejected these criticisms (Neis 1992, Hutchings et al. 1997). Probably not coincidentally, DFO sought to expand the fishery to provide employment (Walters and Maguire 1996, Finlayson and McCay 1998). Eventually, the fishery collapsed and has been closed up till today.

But independence has costs too. Separating monitoring activities might hamper communication between the decision makers and those conducting the monitoring. Monitoring is often more effective and efficient if it is closely coordinated with the decisions it informs. A lack of coordination can produce a disconnect between management or regulatory goals and the variables or scale of a monitoring program, thereby undermining monitoring’s effectiveness (Silsbee and Peterson 1993). The separation of monitoring from decision making may also exacerbate problems of timing because the decision making agency may act without waiting for the monitoring data to be collected or analyzed.

Funding monitoring may be easier if the relevance of monitoring is clearer; uniting monitoring and use of monitoring in one agency might help make that relevance clearer. A larger, combined agency may also have more political clout and therefore a greater ability to obtain overall funding, thus offsetting the risk that the agency might disproportionately cut monitoring. Isolated, small agencies might even be eliminated, as almost happened to the FPC.

Independent monitoring agencies may decrease the use of monitoring data. Relatively effective independent monitoring agencies, i.e., the Everglades restoration program, the Glen Canyon Dam restoration program, and California’s Fish and Game Commission’s monitoring of the sardine fishery, all struggled with using data for decision making. The challenges at the Glen Canyon Dam were described above; in California, the contradictory interpretations of the monitoring data by state and federal fisheries agencies prevented any significant regulatory efforts to protect the sardine fishery (McEvoy and Scheiber 1984); and for the Everglades, progress has been extremely slow on the restoration efforts despite clear data that water quality and hydrology problems persist (NRC 2012).

If we consider the tradeoffs among reducing internal agency goal conflicts and budget pressures, coordinating decision making with monitoring, and ensuring that monitoring data is used, when is a stand-alone monitoring agency the best approach? Regulatory agencies may have fewer conflicts between monitoring and other goals than management agencies. Regulatory agencies such as the U.S. Environmental Protection Agency often have a mission to identify environmental problems that require regulatory solutions and may need monitoring data to justify new regulations legally or politically (DeMuth and Ginsburg 1986). Management agencies such as the U.S. Forest Service or Bureau of Land Management often have a mission based on development rather than environmental protection (Stewart 2001, Biber 2009); monitoring data may identify new or emerging environmental problems that might restrict proposed development activities. Thus, the benefits of an independent monitoring agency might be higher in combination with management agencies.

Certain activities may also require less coordination, and therefore be more amenable to separation, for instance, the imposition of uniform environmental standards. The monitoring program need not be tailored for individual management decisions because standards must be met. On the other hand, assessing whether a particular management option has achieved particular environmental goals may require calibrating the monitoring program to the specifics of the management option selected and the goals to be achieved.

The difference between a monitoring agency and a management or regulatory decision making agency is more a matter of degree than of kind. One solution to the tradeoffs listed above is to blur the differences between categories. For
instance, one could expand the role of independent monitoring agencies beyond data collection and analysis to include reaching conclusions about management or regulatory changes. This might prevent a decision making agency from “sitting” on monitoring data that is institutionally inconvenient, especially if there is strong public trust in the monitoring agency (Imperial 1999). Similarly, one solution to timing problems is to give the monitoring agency the power to recommend or impose delays on the decision making agency.

Expanding the scope of monitoring agencies in these ways might improve the use of monitoring data. However, involving monitoring agencies more in management and regulatory decisions may subject them to the very political and bureaucratic pressures that independence was supposed to insulate them from.

Another risk is that giving monitoring agencies greater decision making power might fragment the decision making process. If the monitoring agency has the power to order particular management decisions, then conflicts between the monitoring agency and the management agency might produce inconsistent and counterproductive management actions.

One could cabin the scope of the monitoring agency’s ability to force the use of monitoring data. For instance, the monitoring agency might be limited to producing reports on the relevance of monitoring data for decision making; although such reports may molder on a bookshelf, they may lead courts to strike down decisions that are inconsistent with the monitoring data. This, in turn, will create incentives for the decision making agency to take the data into account in the future.

Or the monitoring agency might set minimum standards for important environmental resources. This would not prevent the decision making agency from pursuing its own goals, but would ensure that the monitoring data would be considered to some extent. This is similar to consultation among federal agencies under the Endangered Species Act.

Standards have limits; they can address only some resource conflicts. Standards might not force affirmative steps to restore a resource. They also increase the transaction and litigation costs of decision making.

Thus, blurring the distinctions between decision making agencies and monitoring agencies should probably only be adopted when the monitoring data will be particularly important for decision making, for instance, when the resource being monitored has high economic, cultural, or ecological value or is highly endangered. The higher the benefit from using the monitoring data in decision making, the more worthwhile it might be to impose significant costs on the decision making process. However, it is not often clear ex ante, before a monitoring system is established, which resources are most valuable or most endangered.

CONCLUSION
It is essential that the design of any monitoring program consider how to best construct the institutional and legal framework within which the monitoring will occur. Constructing that institutional and legal framework necessarily requires a thoughtful examination of the political, economic, and social contexts of the monitoring program. Nonetheless, the monitoring literature has focused on the technical aspects of monitoring or on its justification from a cost-benefits perspective (Sayre et al. 2013).

There may not be any easy institutional, political, or legal solutions to the challenges for effective monitoring. The alternatives, however, may be even less palatable. Failure to consider these challenges increases the likelihood that a monitoring program will fail, and thus the possibility that unwelcome surprises will drive decision making. As a result, important options to resolve environmental problems might be foreclosed.

Of course, we could decide to be ready for surprises or at least be aware of their possibility (Kates and Clark 1996). If monitoring programs are sometimes unreliable because of legal, political, or institutional obstacles, then we should be prepared for them to fail. We may not wish to base all, or important components of, our regulatory or management systems solely on the existence of high-quality ambient monitoring data. We might want to make our systems resilient to the failures of monitoring. That option raises questions about currently popular management choices such as adaptive management that heavily depend on monitoring to be successful.

Responses to this article can be read online at: http://www.ecologyandsociety.org/issues/responses.php/6117

Acknowledgments:
This paper was made possible by a grant from the David and Lucille Packard Foundation. K. Lee and participants in a roundtable hosted by the Packard Foundation provided important input on prior drafts. M. Loun, J. Dimon, H. Bodwitch, L. Driscoll, J. Barandiarian, and S. Schwartz provided important research assistance. Thanks to two anonymous reviewers for helpful comments.

LITERATURE CITED


Ernst, H. R. 2003. *Chesapeake Bay blues: science, politics, and the struggle to save the bay*. Rowman and Littlefield, Lanham, Maryland, USA.


Roe, E. 1996. Why ecosystem management can’t work without social science: an example from the California


