2012

Surprises and Insights from Long-Term Aquatic Data Sets and Experiments

Walter K. Dodds  
Division of Biology, Kansas State University

Christopher T. Robinson  
Swiss Federal Institute of Aquatic Science and Technology

E. E. Gaiser  
Department of Biological Sciences, Southeast Environmental Research Center, Florida International University,  
gaisere@fiu.edu

Gretchen J.A. Hansen  
Center for Limnology, University of Wisconsin—Madison

Heather Powell  
National Ecological Observatory Network

See next page for additional authors

Follow this and additional works at: http://digitalcommons.fiu.edu/fce_lter_journal_articles

Recommended Citation

This material is based upon work supported by the National Science Foundation through the Florida Coastal Everglades Long-Term Ecological Research program under Cooperative Agreements #DBI-0620409 and #DEB-9910514. Any opinions, findings, conclusions, or recommendations expressed in the material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. This work is brought to you for free and open access by the FCE LTER at FIU Digital Commons. It has been accepted for inclusion in FCE LTER Journal Articles by an authorized administrator of FIU Digital Commons. For more information, please contact dcc@fiu.edu, jkrefft@fiu.edu.
Surprises and Insights from Long-Term Aquatic Data Sets and Experiments


Long-term research on freshwater ecosystems provides insights that can be difficult to obtain from other approaches. Widespread monitoring of ecologically relevant water-quality parameters spanning decades can facilitate important tests of ecological principles. Unique long-term data sets and analytical tools are increasingly available, allowing for powerful and synthetic analyses across sites. Long-term measurements or experiments in aquatic systems can catch rare events, changes in highly variable systems, time-lagged responses, cumulative effects of stressors, and biotic responses that encompass multiple generations. Data are available from formal networks, local to international agencies, private organizations, various institutions, and paleontological and historic records; brief literature surveys suggest much existing data are not synthesized. Ecological sciences will benefit from careful maintenance and analyses of existing long-term programs, and subsequent insights can aid in the design of effective future long-term experimental and observational efforts. Long-term research on freshwaters is particularly important because of their value to humanity.

**Keywords:** data, lakes, streams, freshwater, long term

Freshwaters are crucial for human survival, central to ecological structure and function, and arguably the most valuable habitats on Earth (MA 2005). Freshwaters integrate across entire watersheds: Small streams are the primary interface between terrestrial and aquatic habitats, and larger freshwater systems reflect cumulative terrestrial influences, as well as transformations within the freshwater system. Lakes have a longer turnover time than streams and integrate longer biophysical processes. Therefore, freshwaters serve as important sentinels of environmental change across multiple temporal and spatial scales (Williamson et al. 2008). Freshwaters are under siege, which generates further impetus to understanding the long-term consequences of environmental changes that endanger freshwater flora and fauna.

Sweeping anthropogenic changes can complicate ecological forecasting, because the systems involved have no historical analogue (Milly et al. 2008). Mechanistic understanding (Dodds 2009), improvements in computing, and interdisciplinary approaches (Clark et al. 2001) facilitate ecological forecasting. The use of these approaches in forecasting can be strengthened through the application of both observational data sets and ecological experiments long enough to cover multiple generations of diverse organisms in their natural habitats over widely variable environmental conditions.

Although short-term ecological observations and experiments provide valuable insights about ecological processes, they are unable to support predictions of some system-wide environmental changes (Schindler 1998, Lindenmayer and Likens 2010). High temporal variation can make short-term environmental records susceptible to errors of interpretation, especially if the conditions were anomalous during the period of record. With proper analytical and experimental approaches, long-term data can be used to address important ecological issues, even with little replication across systems (Carpenter et al. 1998). Long-term data are useful in answering ecological and evolutionary questions that encompass broad temporal and spatial frameworks and can add context to short-term experiments. Long-term studies are particularly important with respect to slow ecological processes, rare or episodic phenomena, highly variable processes, and subtle or complex phenomena and, therefore, can...
play an important role in formulating and testing ecological theory (Franklin 1988, Lindenmayer and Likens 2010).

The benefits of long-term research have been discussed previously (e.g., Strayer et al. 1986, Franklin 1988, Magnuson 1990, Hobbie et al. 2003, Lindenmayer and Likens 2010). Now researchers have an opportunity to examine even-longer-term dynamics with newly accessible sources of long-term data. Advances in the digitization of existing data and unprecedented computing ability further enhance ecological study. Many local and national agencies and international groups now provide their environmental data online, although most ecological data are probably not readily available (Reichman et al. 2011). Tools for dealing with complex data series over time, including predicting rare events, are also developing rapidly (e.g., Frei and Schär 2010).

In the present article, we document how experiences from prior long-term studies can help researchers design the long-term monitoring and experiments of the future and how large amounts of long-term ecological data are available and ready for synthesis. We contend that long-term monitoring and experimentation can provide unique insights and can lead to surprising results in ecological research. Ecology has been a field of study for less time than the life span of some organisms, so it makes sense that we are just beginning to appreciate the utility of long-term data and experiments. We focus here on freshwaters, given their importance to society and the widespread creation of freshwater data sets by governmental agencies and in order to constrain the breadth of our analyses. We address some challenges in long-term monitoring (particularly those not addressed in prior publications) and explore the contribution of long-term experimental approaches in advancing ecological understanding. We review six examples in which long-term experimentation and monitoring have provided unexpected ecological insights and then discuss useful types of long-term data, including some that are not commonly used by ecologists. We examine the challenges associated with long-term monitoring and experimentation, including how the value of data sets changes over time. Finally, we look to the future of long-term research with respect to new initiatives and technological advances.

First, we consider definitions and interpretations of long term with respect to ecological studies (box 1). The most ecologically relevant definition encompasses a period of at least several generations of organisms and time periods long enough to catch rare, strongly controlling events and sustained trends in variable habitats. In this context, freshwaters are excellent study systems because most organisms in these ecosystems have relatively short generation times. Therefore, studies of a decade or more are often long enough to observe most demographically related responses. Longer-term studies, however, are necessary to describe climate-related responses, particularly as they relate to in uences on hydrology and biogeochemistry (see Lindenmayer and Likens 2010).

**Box 1. What is long term with respect to ecological data?**

The meaning of long term is observer specific (Strayer et al. 1986); most long-term ecological studies cover only a sliver of time from a paleoecological point of view. Long term might be considered to be an arbitrary double or triple the length of the typical grant or graduate program. Long term is relative and context dependent; to be relevant to the process of interest, the data record must adequately capture cycles of interactions among drivers and response variables. For many studies, the length of the data record determines the usefulness of the data and the questions that can be evaluated (see Lindenmayer and Likens 2010).

Practically, long term should be long enough to answer the question at hand. "Long enough" should not presuppose relative stability in the long-term or baseline conditions under a changing climate (Milly et al. 2008). Many environmental drivers act in a long-sustained manner (press as opposed to pulse). With time-sensitive unreplicated measurements (e.g., hydrology in a unique watershed), long enough to answer the question might mean long enough to gain the statistical power to find a meaningful response (to detect statistically significant responses to presses).

A biological basis for defining long term could be the generation times for study organisms or collection over a long enough period for drivers to act on the ecosystem processes of interest (Strayer et al. 1986, Lindenmayer and Likens 2010). If evolutionary time is of concern, many generations are needed.

A physical basis for long term, or at least a basis built on external drivers, requires a study period long enough to include the periodicity of environmental conditions and drivers. This characterization of ecosystem rates should be long enough to catch rare but strong events, as well as the disturbance return time following such events. Long-term studies need to encompass the natural variability and periodicity of both biophysical and environmental cycles. The periodicity of climate cycles, from several years (El Niño Southern Oscillation) to decades (Pacific Decadal Oscillation), influences the magnitude and timing of precipitation and temperature, which serve as drivers for many aquatic systems. The identification of trends over time can be confounded by shifts in climatic cycles or extreme events. Routine monitoring over a longer time period with relatively low sampling frequency can catch rare events and allow assessment of their importance (e.g., Banner et al. 2009).

We therefore define a long-term ecological data set as data that are measured systematically through time using standardized methods that allow for the elucidation of ecological system responses (e.g., linear, lag, threshold, regime shift) to drivers, disturbances (e.g., presses or pulses), recovery from disturbances, and relevant interactions for a given hypothesis.
**Surprises from long-term data**

In the present article, we provide six case studies illustrating the importance of long-term data for understanding the role of disturbance in structuring freshwater ecosystems under various conditions. Ecological disturbances have been broadly characterized temporally as acute *pulses* or as chronic *presses* (Bender et al. 1984). Long-term data are important for better understanding slow directional presses and for observing statistical significance in variable processes. Sometimes, the result from a 5-year study could be completely opposite the result of a longer-term approach (see Lindenmayer and Likens 2010). Finally, long-term observation increases the probability of finding rare stochastic-pulse, time-lagged, or threshold responses (Franklin 1988).

We consider these six detailed studies and some less detailed studies to answer the following questions: Is a study period of more than 5 years required to see an effect (e.g., to boost statistical power in a variable system)? Were there surprise time lags or nonlinear responses? Could the data characterize the effects of long-term trends in slowly changing abiotic drivers? Was the result a surprise relative to prior research in the area, and was this result obtained because of long-term research?

**Case study 1: A surprising delayed effect of an invasive species on an aquatic food web in a northern lake.** Rusty crayfish (*Orconectes rusticus*) are an invasive species that affects all trophic levels of aquatic ecosystems (e.g., Lodge et al. 1994, Wilson et al. 2004). Rusty crayfish were first observed in low abundances in Trout Lake, Vilas County, Wisconsin (46 degrees 13 minutes north, 89 degrees 32 minutes west), in 1979 and dispersed throughout the lake over the subsequent decades (Wilson et al. 2004). Crayfish and aquatic macrophyte abundances were surveyed annually starting in 1981 by the North Temperate Lakes Long Term Ecological Research (LTER) Network site. Concurrently, rusty crayfish introduced to other lakes had reached high densities and displaced native congeners. Short studies showed that rusty crayfish grazed macrophytes and could reduce plant standing stocks and species richness (Magnuson et al. 1975, Lodge and Lorman 1987).

Contrary to expectations, there was no decline in macrophytes in Trout Lake even after 14 years. The rusty crayfish population steadily increased through 1995, with no apparent effects on macrophyte biomass (figure 1). Eventually, a statistically significant decrease in macrophyte biomass was observed; analysis of the relationship between macrophyte density and year revealed a breakpoint in 1996 (determined by piecewise regression, *p* < .05), but identification of this breakpoint was not possible until after 1999, at which point three consecutive years with macrophyte abundance less than the running 95% confidence interval were observed. This lake demonstrated a potential threshold effect of crayfish on macrophyte biomass, a time-lagged effect, or a combination of the two. Without long-term data, researchers and resource managers could have concluded erroneously that invasive crayfish had few impacts and were not a cause for concern in Trout Lake or other recently invaded lakes.

The surprise in this case study was that 10 years was not sufficient to identify the long-term effects that had been predicted from cross-system comparisons and small-scale experiments. Long-term observation was necessary in part because year-to-year variance could have masked trends but also because there was an apparent time-lagged or threshold response to increases in the invasive species’ biomass (i.e., a delayed response to a press disturbance).

**Case study 2: A long delay in the response of an arctic river food web to phosphorus fertilization.** A long-term experiment to evaluate the impacts of increased nutrient loads began in 1983 on the oligotrophic Kuparuk River at the Toolik Lake LTER site (68 degrees 38 minutes N, 149 degrees 38 minutes W) in Alaska. Each summer, phosphorus levels were increased tenfold (from 0.03 to 0.3 micromoles per liter) to evaluate ecosystem and community responses in this tundra river. On the basis of experimental nutrient additions in nearby streams for up to 4 years, phosphorus addition was expected to stimulate production by periphyton and higher trophic levels (Slavik...
et al. 2004). Surprisingly, almost a decade was required for the full response to this press nutrient addition.

In the first 4 years of enrichment (1983–1986), benthic chlorophyll, invertebrate density, and fish growth increased as was expected (Peterson et al. 1993) and no obvious changes were observed in bryophyte cover. Unexpectedly, after nine seasons of fertilization, a sharp increase in bryophyte biomass was observed, prompting quantitative annual sampling of bryophyte biomass. This is an example of adaptive monitoring in response to a surprising observation (Lindenmayer and Likens 2009).

Within 2 years of sampling (the minimum for statistical comparison), bryophyte biomass in the fertilized reach significantly exceeded that in the reference reach (figure 2).

Given the doubling times of the bryophyte biomass calculated from the first 4 years of sampling in the fertilized reaches, 4 years of growth would be necessary for a significant difference to be evident in the reference reaches. Therefore, the apparent time lag in bryophyte response was not an artifact of simply sampling long enough to increase statistical power and to detect a slowly growing bryophyte population. Rather, it represented a lag in the response to nutrient addition.

Total bryophyte cover increased from 5% to 50% in the fertilized reach, and the dominant moss type changed from Schistidium agassizii to Hygrohypnum spp. Moss replaced epilithic diatoms as the main primary producer, net primary production also nearly tripled, and there was an elevenfold increase in larval chironomid density in the fertilized reach (Slavik et al. 2004).

Given the observed time lag, only long-term nutrient additions could establish the importance of chronic nutrient enrichment in this arctic stream. The surprise was an extreme time lag in the response of aquatic mosses to a continued fertilization (a press disturbance). This is also an example of flexibility in long-term monitoring to capture unanticipated results.

Case study 3: Explaining the increased drying of a highly variable prairie stream from long-term data trends. An analysis of discharge data from a US Geological Survey gauging station on Kings Creek at Konza Prairie Biological Station (39°05' N, 96°35' W), in Kansas, indicates that long periods with no surface ow are becoming more common in this already intermittent stream (figure 3a). The trend would not be evident if only the last 15 years of the data set were used. A regression analysis also indicates a significant decrease in discharge over time (p < .05).

Analyses of local long-term precipitation and temperature records revealed no trends that should cause more frequent drying (figure 3b, 3c), because neither precipitation nor the ratio of precipitation to pan evaporation changed significantly. However, a 14% expansion of woody riparian vegetation occurred over the time of the study (Briggs et al. 2005). Coupling discharge and vegetation data gives the current best explanation for the discharge decrease. The surprise is the increase in the occurrence of days without ow. The observation of long-term ow decrease has led to a whole-watershed woody riparian vegetation removal experiment to test the hypothesis that the observed expansion of woody riparian vegetation is responsible for decreasing stream discharge and increasing the number of days without ow.

Case study 4: Multidecadal climate cycles drive trends in transparency in a Florida lake. An analysis of long-term meteorological, hydrological, and limnological data from Lake Annie at the Archbold Biological Station (27°11' N, 81°20' W), in
south central Florida, revealed a strong interaction between climate pressures and water clarity. Data collected over 25 years (figure 4) showed strong correlations among rainfall, lake and groundwater stage, and water transparency. In years of high rainfall punctuated by frequent tropical storms, transparency was reduced because of an increased runoff of colored dissolved organic material into the lake from the watershed. During stormy years, the mixed layer was deeper and the bottom waters were colder (Gaiser et al. 2009a).

The 5-year running mean of Secchi depth (greater depths indicate more lake clarity) illustrates how perceived trends depend on the time period chosen. Secchi depth was initially stable, but after 10 years, the running mean was outside the previous 95% confidence interval, and a sharp decrease in Secchi depth was observed. In the last 5 years, there was a trend toward greater depth, which indicates that the pattern could be cyclical.

The strong correlation between lake stage (depth) and transparency in the modern record of Lake Annie was used to create a long-term prediction of alternate transparency states from historic depth measurements. This prediction corresponded to anecdotal reports of lake color from local visitors who recalled alternations between stained and clear conditions in the past. An examination of a 75-year record of weather and lake stage data revealed a connection among rainfall, lake stage, and the Atlantic Multidecadal Oscillation (AMO, the deviation in Atlantic sea surface temperature with 30–40-year cycles). In south central Florida, warm AMO periods show an increased frequency of tropical storms and a cumulative increase in the water table, which indicates that lake transparency should oscillate with the same period as the AMO (Gaiser et al. 2009b).

In this case, long-term data linked changes in rainfall driven by cyclical climate patterns to oscillations in lake level and water clarity. Alternations in lake clarity could influence the relative importance of internal and external carbon sources to the lake food web. The short-term variation in this particular lake could be small enough that 5–10 years of observation could be misinterpreted to represent stable conditions or a unidirectional trend. This case illustrates the value of extending predictions into the past in order to validate inferred long-term changes. This research also demonstrates that patterns that appear to be presses could be viewed as repeated pulses (the AMO cycle) over longer periods of time.

Case Study 5. A flood on Mack Creek unexpectedly increased cutthroat trout populations. A major flood (which has a 35–50-year recurrence interval) occurred in Mack Creek in early February 1996 and was expected to have negative effects on fish populations; surprisingly, this pulse had the opposite effect. Mack Creek is a steep, third-order stream in
the H. J. Andrews Experimental Forest (44°12′ N, 122°09′ W), in Oregon, which owns through a 500-year-old conifer forest, then through a young conifer forest that was harvested in 1964. The maximum annual stream flow was two- to threefold greater than most other years during this period (figure 5), with water levels over 1 meter above the bankfull-channel stage. The flood moved boulders and large logs, scoured deep pools, and created extensive gravel deposits. The densities of cutthroat trout (Oncorhynchus clarkii) were predicted to be decreased by this flood, since such floods can injure fish or wash them downstream (Matthews 1998).

Surprisingly, the fish populations in reaches through old-growth and young forests did not decline in the summer after the 1996 flood. Adult trout survived and the number of young-of-the-year trout dramatically increased, which resulted in a substantial population increase (figure 5). Across the 23-year record of fish populations, trout population densities (particularly of young fish) after the flood were the highest measured. Fish densities stayed high for several years; the especially high densities of young fish successfully translated into increased numbers of adult fish, and the population persisted through another flood about 80% as large in 2000. Retrospective analyses indicated that the ratio of young-of-the-year to adult trout was positively correlated with the maximum discharge, particularly in the old-growth reach—a result that would not be apparent without decades of data.

In Mack Creek, the surprise is that large floods (pulse disturbances) actually have a positive effect on recruitment of young-of-the-year trout. Continued long-term monitoring also indicated that the harvested reach exhibited a greater variance than the old-growth reach (figure 5). This data set allows researchers to examine theories of fish responses to a large and infrequent disturbance and to follow the pulse of recruitment over sequential years.

Case Study 6: Extreme hurricane damage to a tropical forest does not decrease watershed nitrogen retention. In much of the tropics near coastal areas, hurricanes, typhoons, and cyclones can have profound effects on forested watersheds, with damage to virtually all trees and with changes in forest structure that can persist for many decades (Lugo 2008). Researchers expected that the complete deforestation of a tropical forest by a hurricane in Puerto Rico would lead to an extended “leakiness,” represented by an increase in stream nitrogen export. Prior theory suggested that tropical forests should be “leaky” with respect to nitrogen because they are often nitrogen-rich, with large stores of nitrogen in soils and plant biomass (e.g., Chestnut et al. 1999), and primary productivity is seldom nitrogen limited.

In the tropical forest streams of the Luquillo Mountains in Puerto Rico (18°19′ N, 65°45′ W), inputs of a year’s worth of detritus in one day and extensive hurricane damage to vegetation led to a sharp but short-lived increase in stream nitrate levels (figure 6). The subsequent rapid decline occurred despite the relatively high soil nitrogen content, large nitrogen input from detritus, and a lack of nitrogen limitation of tree growth. Nitrate returned to near background concentrations (albeit slightly elevated and with increased variance) within 2 years (Schaefer et al. 2000).

The short-term response of the nitrogen-rich Puerto Rican forests to hurricanes has
Abbreviations: L, liters; N, nitrogen; µg, micrograms.

It is vital to fully understand the responses of the ecosystem from before these unpredictable disturbance events were impacts on in-stream nitrate quantities. Long-term records from before these unpredictable disturbance events were vital to fully understand the responses of the ecosystem.

Figure 6. Monthly average stream nitrate values and 1-year running mean from the Quebrada Sonadora in the Luquillo Long Term Ecological Research site between 1986 and 1994. In 1989, Hurricane Hugo caused complete defoliation in the river’s watershed with subsequent impacts on in-stream nitrate quantities. Long-term records from before these unpredictable disturbance events were vital to fully understand the responses of the ecosystem.

Surprises from additional case studies. Long-term data sets can be a repeated source of surprise (see Lindenmayer et al. 2010), which further accentuates the importance of maintaining and revisiting them. Starting in 1940, researchers at Lake Windermere, England, began collecting data that eventually allowed long-term records of zooplankton abundance to be tied to long-term variation in Atlantic sea surface temperatures (George and Harris 1985) and revealed that changes in zooplankton were more strongly related to climate than to the top-down fish effects. In Lake Myvatn, Iceland, midge populations (Diptera: Chironomidae) have been monitored since 1977. After 30 years of monitoring and many publications, researchers concluded that complex fluctuations in midge densities could be explained by alternative dynamic states in which the initial conditions dictated whether midge populations would fluctuate around a fixed point or enter a period of high-amplitude cycles (Ives et al. 2008). In Lake Washington, a classic example of community recovery due to control of nutrient pollution, some of the changes in zooplankton community structure first attributed to decreasing nutrients were later shown to be due to thermal shifts resulting from progressive warming of the lake (Winder et al. 2009).

The examples from these case studies demonstrate the value of long-term data from freshwater ecosystems in understanding multiple drivers of change operating over a range of temporal and spatial scales with respect to pulse or press disturbances. However, these case studies represent only a small portion of data that are available to inform management and provide scientific insight.

What is out there? Types and sources of long-term data

Data have been collected and archived over periods ranging from decades to centuries, and some examples of long-term freshwater data sets and their sources are presented in table 1. This table is not comprehensive but illustrates the range of sources and the resulting scientific publications that can arise from those sources.

Scientific research. Many of our longest data records in aquatic ecology were initiated by individual limnologists. Examples include the 44-year record from Lake Washington, begun by W. T. Edmondson in the 1950s, that provided evidence for the efficacy of sewage diversion to allow long-term recovery from eutrophication (Edmondson 1991); the career-long (more than 30 years) attention paid to Lawrence Lake, Michigan, by R. G. Wetzel that revealed—among other things—a long-term reduction in dissolved organic carbon coincident with decreasing regional precipitation (Wetzel 2001); the long-term monitoring of Lake Tahoe by the research team of Charles Goldman started in the 1960s, which revealed eutrophication following development in the watershed and a shift from nitrogen to phosphorus limitation in the 1980s (Goldman et al. 1993); and historical evaluations of many Swiss lakes, including Lake Constance, which documented how reduced phosphorus inputs since the 1970s altered the genetic architecture of Daphnia populations (Brede et al. 2009).

Additional long-term limnological data sets have emerged from consortia at universities and government agencies. Experimental forests and biological field stations can be treasure troves of long-term aquatic data (table 1). Long-term stream biogeochemical responses to forest practices and other disturbances have occurred at US Forest Service experimental forests (e.g., the Hubbard Brook Experimental Forest; Likens et al. 1996), leading to an understanding of the watershed-level effects of acidification and deforestation. Archives of historical data are often recorded by the continuous succession of researchers, station visitors, and classes at a site. For example, ecological data from over 100 years of research are now being archived and standardized by the Swiss National Park (www.swissnationalpark.ch).
of scientific research networks, focusing attention on the importance of long-term studies (Hanson 2008). For example, long-term monitoring stations in Europe have been used by researchers to examine global environmental changes in freshwaters and the rapid recession of glaciers in the Alps (Schädler and Weingartner 2010).

**Monitoring.** National and regional government agencies employ long-term monitoring programs to detect change

<table>
<thead>
<tr>
<th>Category</th>
<th>Example sources</th>
<th>Years of record</th>
<th>Variables</th>
<th>Purpose</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>US National Park Service (more than 270 sites; <a href="http://science.nature.nps.gov/im">http://science.nature.nps.gov/im</a>)</td>
<td>1993–present</td>
<td>C P</td>
<td>R M</td>
<td>Childers et al. 2003</td>
</tr>
<tr>
<td></td>
<td>US Experiment Forests and Ranges (81 sites; <a href="http://www.fs.fed.us/research/efr">www.fs.fed.us/research/efr</a>)</td>
<td>1908–present</td>
<td>B C P</td>
<td>R</td>
<td>Likens et al. 1996</td>
</tr>
<tr>
<td></td>
<td>Florida Department of Environmental Protection (<a href="http://www.dep.state.fl.us/mainpage/programs/water.htm">www.dep.state.fl.us/mainpage/programs/water.htm</a>)</td>
<td>1979–present</td>
<td>B C P</td>
<td>M</td>
<td>Redfield 2000</td>
</tr>
<tr>
<td>Surprise sources of data</td>
<td>Citizen science monitoring programs, other local historical records (e.g., <a href="http://www.icewatch.ca">www.icewatch.ca</a>, <a href="http://lakewatch.ifas.ufl.edu">http://lakewatch.ifas.ufl.edu</a>)</td>
<td>Varies, 1986–present</td>
<td>B C P</td>
<td>M E</td>
<td>Magnuson et al. 2000, Canfield et al. 2002</td>
</tr>
<tr>
<td>Universities, research institutes and consortia, museums, data repositories</td>
<td>Cary Institute of Ecosystem Studies (<a href="http://www.ecostudies.org/emp_data_list.html">www.ecostudies.org/emp_data_list.html</a>)</td>
<td>1991–present</td>
<td>B C P</td>
<td>R</td>
<td>Pace et al. 1998</td>
</tr>
<tr>
<td></td>
<td>HYDRODB (a consortium combining data from many sources; <a href="http://www.fsl.orst.edu/climhly">www.fsl.orst.edu/climhly</a>)</td>
<td>Varies, 1980–present</td>
<td>P</td>
<td>R</td>
<td>Jones et al. 2012</td>
</tr>
<tr>
<td></td>
<td>Long Term Ecological Research Network (26 sites; <a href="http://www.lternet.edu">www.lternet.edu</a>)</td>
<td>1980–present (many have pre–1980 data)</td>
<td>B C P</td>
<td>R E</td>
<td>The present article and many cited herein</td>
</tr>
<tr>
<td></td>
<td>Royal Ontario Museum, Canada (data not online; <a href="http://www.rom.on.ca/collections/index.php">www.rom.on.ca/collections/index.php</a>)</td>
<td>1925–present</td>
<td>B</td>
<td>R E</td>
<td>Lavinie and Campenau 2010</td>
</tr>
<tr>
<td></td>
<td>The Canadian Barcode of Life Network (<a href="http://www.bolnet.ca/index.php">www.bolnet.ca/index.php</a>)</td>
<td>N/A</td>
<td>B</td>
<td>R</td>
<td>Rivera and Currie 2009</td>
</tr>
<tr>
<td></td>
<td>Biological field stations (more than 100 sites; <a href="http://www.obfs.org">www.obfs.org</a>, <a href="http://www.obfs.org/fslink.html">www.obfs.org/fslink.html</a>)</td>
<td>Varies, 1941–present</td>
<td>B C P</td>
<td>R E</td>
<td>Gaiser et al. 2009a</td>
</tr>
<tr>
<td></td>
<td>University of Wisconsin—Madison Center for Limnology (<a href="http://limnology.wisc.edu/Lake_Information.php">http://limnology.wisc.edu/Lake_Information.php</a>)</td>
<td>1911–present</td>
<td>B C P</td>
<td>R</td>
<td>Magnuson et al. 2006</td>
</tr>
</tbody>
</table>

**Note:** The “Years of record” column indicates the range of dates when data were collected. The data were not always continuously or regularly collected in the time period. The “Variables” column indicates whether biological (B), chemical (C), or physical (P) data (or a combination of those types) were collected. The “Purpose” column indicates the original intention of the data-collection program: R, research; M, monitoring, conservation, and management; E, education; G, regulatory. The “Reference” column lists an example of a peer-reviewed publication that used the data source or type. This table is meant to illustrate the breadth of the data sources available and how they can be used in scientific analysis but is by no means exhaustive. Many of the data sets span multiple institutions.
and to effectively manage aquatic resources (table 1). For example, the US Geological Survey’s National Water-Quality Assessment program data have revealed long-term trends in US water quality (Smith et al. 1987). The US Environmental Protection Agency also conducts continental-scale evaluations of aquatic resource quality through the Environmental Monitoring and Assessment Program, which includes periodic resampling to assess trends over time. The European Environment Agency coordinates a highly integrated program called the National Surface Water Quality Monitoring Network, which has 32 member countries using standardized and harmonized methods, information management, and reporting. State and regional environmental protection agencies often use long-term environmental observations to guide aquatic resource protection. Although findings from these long-term regional monitoring programs have directly in uenced policy changes (Forsius et al. 2003), the funding for continuous monitoring programs is often tenuous.

Freshwater scientists are fortunate that water quality and aquatic biota are routinely monitored using roughly similar methods. Routine and comparable biological monitoring is less available across terrestrial and marine ecosystems. Macroinvertebrate sampling is used to create indices of biotic integrity in many nations; therefore, comparable biological data collected over long periods of time over broad geographic scales are becoming available for freshwaters.

Some very-long-term freshwater data have been collected by citizen scientists. For example, communities have been recording ice-off dates in northern lakes for centuries (Magnuson et al. 2000). Citizen monitoring programs are growing in popularity, can be long lasting and reliable (Canfield et al. 2002), and may be an important source for long-term aquatic data. Formal citizen-based stream- and lake-monitoring programs leverage peoples’ attraction to and interest in freshwater lakes. Such programs depend on volunteers who are trained in basic techniques and who collect regular data or samples near their homes. Over time, these records allow the public to understand directional changes to local freshwater resources that may be so slow that they would not otherwise be noticed.

Currently, new initiatives are pulling together long-term data, stimulated by access to an increasing amount of high-quality data and a better technological ability to collate and synthesize disparate data sets. Agencies have shifted from paper to online data distribution (http://water.usgs.gov/data), and synthetic online databases across disparate data sets include HYDRODB (www.fsl.orst.edu/climhy) and the Hydrologic Information System (http://his.cuahsi.org) (table 1).

Paleolimnological and archived data. In addition to direct evaluations of long-term data, retrospective studies can also be conducted using archival or preserved materials. Long-term lake histories have been reconstructed from the gut contents of fish preserved in museums (Lavoie and Campeau 2010), remnant biofilms on macrophytes from herbarium collections (Denys 2009), and often from paleolimnology. The use of sediment cores to evaluate data preserved in lake sediment strata has become an established practice for investigating limnological and climatic changes. Lake sediments preserve layers of deposited materials and can be analyzed to reconstruct past limnological parameters such as pH, temperature, lake level, salinity, the duration of ice cover, nutrients, and general climatic trends that provide insight into global climate change (e.g., Smol and Cumming 2000). A prime example of paleolimnology informing management is related to acidification by atmospheric deposition in North America and Europe. Although some claimed that the lakes could have been naturally acidic, dated sediment cores indicated a sharp decrease in acid-sensitive planktonic species that coincided with industrial atmospheric loading.

Paleolimnology can elucidate linkages among climate, lakes, and the surrounding terrestrial ecosystems and can allow investigators to extend the climate record to before the earliest instrumental records (around 100 years in North America and 300 years in Europe; Smol and Cumming 2000). Climatic extremes with strong ecological effects can be observed in sediment records from the Holocene and coincide with the collapse of ancient civilizations. For example, a core taken from Lake Chichancanab, Mexico, made a direct link between climate and the Mayan civilization’s deterioration (Hodell et al. 1995). Such records could provide context for current rapid climate change. Sediment cores are now being used to reconstruct environmental conditions in lakes and to test the evoeolutionary responses of past and present biotic assemblages. For example, Daphnia populations can be characterized by resistant resting stages that are from different core layers and that are related to toxins detected in cores (Kerfoot et al. 2004).

Costs and benefits of maintaining long-term data sets. Long-term data are expensive to collect, and researchers and managers are periodically asked to evaluate the need for continuing data collection (Lindenmayer and Likens 2010). There are at least four possible relationships between data set length (cost) and scientific value (benefit). We will initially assume that the cost of data collection does not change over time (ignoring monetary inflation). A unit of scientific value could be roughly defined as the amount of information required in order to form the basis of a paper in the peer-reviewed literature.

First, the simplest relationship for the increase in knowledge through time would be linear (i.e., the amount of information per paper remains fixed), and the relative weight of the scientific payoff as opposed to continued investment is straightforward. One might expect a linear relationship in which a data set provides basic background information for a larger or several related smaller funded programs.

Second, the scientific value of research may increase asymptotically through time. A data set could reach a point at which further study of the system would not lead to new
understanding and at which there would be little benefit to extending the period of record.

Third, the scientific value of an entire data set can decrease drastically over time, which may lead to a sharp decrease in the value of each unit of science per cost. For example, the value of long-term data can be compromised by undocumented changes in methodologies or technological changes without metadata to document what changes were made. Important pieces of data or metadata can be lost. Value can also decrease over time when older methods are shown to be invalid. To guard against this loss of value, long-term research methods need to be standardized so that observations over time are comparable. Standardization requires that new techniques and instruments be analogous before older sample protocols are discontinued, ideally through a substantial period of methodological overlap. Reliable and consistent metadata practices, as well as data storage and archiving, are essential to avoid losing important long-term records. Furthermore, archiving samples and voucher specimens offers valuable resources for future investigators and protection against any potential loss of value.

The fourth relationship, which we think is the most likely, is a stair-step process in which information is gained with periodic and rapid increases in knowledge. Some of the factors that can contribute to these steps in knowledge include the occurrence of a rare but ecologically important event that leads to a new set of dynamics not previously observed, the crossing of a threshold that leads to a regime shift (Carpenter 2003), novel ideas generated from new researchers associated with projects, advances in technology, and new stressors to the system (e.g., invasive species).

Although we assumed stable costs of data collection over time in the previous discussion, technological changes increase the amount of information that can be quickly collected at minimal cost. Sensors and instruments that include automatic recording devices and can be deployed can increase the amount of information that can be quickly collected at minimal cost. Sensors and instruments that include automatic recording devices and can be deployed as networks can collect data at high frequency in a spatially explicit manner. Increasingly sophisticated information management systems that include data archiving, discovery, and access (Brunt and Michener 2009) are becoming invaluable to making long-term data available and useful to the scientific community and the public. These technological advances can increase the value of existing data sets at a more modest cost per unit of data collection, manipulation, and analysis.

**Future needs, opportunities, and conclusions.** The impetus behind maintaining long-term data sets may come down to little more than a sense of moral obligation to society. The scientific benefits from the foresight and continued efforts of individuals who initiate and maintain long-term data sets are clearly delineated in this article. We are in the debt of these individuals, and continuing long-term data sets is a way to pay the debt forward to future scientists.

Several other needs are apparent. One need is to further expand the ability to link disparate data sets across spatial and temporal scales. This linkage could be particularly powerful if long-term ecological data were paired with socioeconomic data to couple data on environmental and social trends. Spatially and temporally specific data sets in social and ecological realms could facilitate such coupling, but these linkages will require substantial investment in cyberinfrastructure.

Another need is the explicit recognition of the importance of taxonomy in long-term ecological data. Funding proper taxonomy and the collection and curation of voucher specimens may be viewed as extraneous and old fashioned by some. However, the long-term usefulness of ecological data sets can be greatly increased by treating taxonomy as essential metadata. As was discussed above, curated samples and specimens can allow revisitation of data sets that might otherwise be obsolete or lost.

Finally, many areas of the globe are underrepresented, with much of the available long-term data originating in developed countries. On a positive note, the International Long Term Ecological Research Network, founded in 1993, now involves 38 networks in 32 countries spread over five continents.

One area in which large gains can be made is in connecting existing data sets and synthesizing the data in publically available forms that are searchable on the Internet. While compiling the sources in table 1, it became apparent that many data sets are available for downloading but that it is likely that few of them have ever been synthesized. The results of simple searches of bibliographic databases using the data set names as keywords suggest that many have not been the subject of a well-cited paper in the scientific peer-reviewed literature. Therefore, there exists a tremendous opportunity to use existing long-term data sets, in conjunction with newly planned long-term programs, to increase our understanding of aquatic ecosystems. This opportunity is often hampered by remaining difficulties in dealing with disparate data sets taken at different temporal scales (e.g., days in one set, months in another). Such data from multiple sources are not easily combined, and problems of attribution may need to be resolved. Software that facilitates such compilations, including metadata transfer, would be immensely valuable for researchers interested in long-term ecological trends in freshwaters or elsewhere. Proper attribution of data sources, rewards for sharing data, and documentation of the work ow used to process long-term data are all areas that need development (Reichman et al. 2011).

Future long-term initiatives are being developed. Two of these initiatives include the Global Lake Ecological Observatory Network (GLEON; http://gleon.org) and the Stream Experiment and Observatory Network (STREON; www.neoninc.org/science/experiments) experiment under the National Ecological Observatory Network (NEON). GLEON is an international monitoring network that is designed to characterize temporal trends in lakes (Hanson 2008). NEON is planning a network of 106 sites, designed to observe terrestrial and aquatic ecosystems over 30 years. NEON serves
as a platform for large-scale experiments, such as STREON: a long-term nutrient enrichment and consumer removal experiment that will be coordinated across NEON aquatic sites. The groups of scientific experts in both the GLEON and the STREON came up with very similar lists of abiotic and biotic parameters to measure, indicating that aquatic scientists are generally unified in what should be included in long-term monitoring efforts.

Several of the present authors have been leaders in GLEON and STREON and have extensive experience with LTER Network sites. On the basis of this experience, we recommend the following to guide planning for future long-term monitoring: Focus monitoring programs on the basic drivers of ecological change that act across levels of ecological organization, from populations to ecosystems. Methods of ecological monitoring are converging and are common across many investigators and systems; base the monitoring framework on specific hypotheses or ecological questions, because this will make the long-term effort more likely to withstand scrutiny and modification (Lindemayer and Likens 2010). If the plans to monitor must be scaled back, the initial hypotheses can be referred to and the system redesigned to focus on the priority questions of interest (Lindemayer and Likens 2009). Archive data with detailed metadata so that others can understand how, when, and where the data were collected and analyzed. Use prior experience with long-term monitoring networks to help guide future efforts. And finally, institute opportunities to use the wide array of available long-term data in compilations of existing, ongoing, and newly collected data.

The community of freshwater researchers involved in long-term studies has risen to some of the challenges that Brown (1994) listed for long-term research: developing non-experimental methods for large-scale research; incorporating information across sources and disciplines; standardizing and controlling data quality; developing new statistical tools; integrating, synthesizing, and modeling data; and incorporating human dimensions. Here, we document the progress on these six challenges: First, an analysis of long-term data across monitoring networks has allowed for nonexperimental methods to be applied to the evaluation of large-scale long-term trends (see, e.g., the references in table 1). Second, the revolution in digital data availability and international communication has allowed for incorporation of information across sources and disciplines. For example, linking sea temperature data to lake stage records explains local historical hydrologic patterns (case study 4), and historical data on lake ice-out times have been used as an indicator of global warming (Magnuson et al. 2000). Third, efforts continue to standardize and control data quality, with an increased emphasis on digital archives and high-quality metadata. Fourth, statistical tools continue to improve for analyzing long time series. Fifth, integration, synthesis, and modeling has expanded: More sites are collecting more data, and our ability to apply climate models to existing long-term patterns for forecasting into the future has increased. Finally, incorporating human dimensions has improved: Ecologists continue to analyze data sets with the goal of understanding human effects on the environment (e.g., Redfield 2000, Banner et al. 2009).

Freshwater ecosystems are under tremendous pressures everywhere from the global, regional, and local impacts of humans. Long-term data on freshwaters and other ecosystems are now more essential and valuable than ever—fortunately, at just the time when more data are becoming available. It is an exciting time to be a researcher and to help create a legacy for the next generation of environmental scientists who will further take advantage of the wealth of long-term data.

Acknowledgments
We appreciate all the individuals who helped collect the data presented in this article. We thank the Long Term Ecological Research (LTER) Network office for enabling our group to meet in Estes Park in 2009 and the National Science Foundation (NSF) for funding (grant no. 0919443 to WKD), and support was also provided by the following NSF LTER grants: no. 1026843 at the Toolik Lake site, no. 0822700 at the North Temperate Lakes site, no. 0823341 at the Konza Prairie site, no. 0620910 at the Luquillo Experimental Forest, no. 0823380 at the H. J. Andrews Experimental Forest, no. 0832652 at the Niwot Ridge site. The Lake Annie data were supported by Archbold Biological Station. This is contribution no. 12–437–J from the Kansas Agricultural Experiment Station.

References cited
Articles


**Walter K. Dodds (wkdodds@ksu.edu)** is a University Distinguished Professor in Biology at Kansas State University, in Manhattan. Christopher T. Robinson is a senior research scientist at the Swiss Federal Institute of Aquatic Science and Technology, in Dübendorf. Evelyn E. Gaiser is an associate professor of biology at Florida International University, in University Park; a research associate at Archbold Biological Station, in Highlands County, Florida; and lead principal investigator of the Florida Coastal Everglades Long Term Ecological Research Program, part of the US Long Term Ecological Research Network. Gretchen J. A. Hansen is a PhD candidate at the University of Wisconsin–Madison’s Center for Limnology. Heather Powell is the assistant director of aquatic sciences at the National Ecological Observatory Network in Boulder, Colorado. Joseph M. Smith is currently a postdoctoral fellow in the Division of Biology at Kansas State University, in Manhattan. Nathaniel B. Morse is a PhD student in the Natural Resources and Earth Systems Science program at the University of New Hampshire, in Durham. Sherri L. Johnson is a supervisory research ecologist with US Forest Service’s Pacific Northwest Research Station, in Portland, Oregon, and lead scientist for the H. J. Andrews Experimental Forest, in Blue River, Oregon. Stanley V. Gregory is a professor in Department of Fisheries and Wildlife at Oregon State University, in Corvallis. Tissa Bell is a PhD student in microbiology at Montana State University, in Bozeman; she completed her master’s at the Institute of Arctic and Alpine Research at the University of Colorado, in Boulder. Timothy K. Kratz is the director of the Center for Limnology’s Trout Lake Station at the University of Wisconsin–Madison; he is currently serving as a program director in the Division of Environmental Biology at the National Science Foundation, in Arlington, Virginia. William H. McDowell holds a presidential chair in the Department of Environmental Sciences and is director of the New Hampshire Water Resources Research Center at the University of New Hampshire, in Durham.

---

**2012 Tyler Prize**

**Dr. John H. Seinfeld**
Louis E. Nohl Professor & Professor of Chemical Engineering
California Institute of Technology

**Dr. Kirk R. Smith**
Professor, Global Environmental Health
University of California, Berkeley

The Tyler Prize was established in 1973 by the late John and Alice Tyler as an international award honoring the achievements in environmental science, environmental health and energy of worldwide importance conferring great benefit on humanity. The Tyler Prize consists of a cash award of $200,000 and a gold Tyler Prize medallion.

For additional information on the 2012 award or nomination procedures for 2013 contact:
Andrea Brown, Administrator, The Tyler Prize
Phone (213) 740-9768 • Fax (213) 740-1313
Email: tylerprize@usc.edu • Website: www.tylerprize.usc.edu

The Tyler Prize is administered by The University of Southern California

---

The Tyler Prize Executive Committee announces the awarding of the 2012 Tyler Prize for Environmental Achievement on its thirty-ninth anniversary to Professor John H. Seinfeld for his pioneering research leading to understanding of the origin, chemistry, and evolution of aerosols in the atmosphere and Professor Kirk R. Smith for his pioneering work at the intersection of environmental science and global public health, including establishing that indoor air pollution in the rural sectors of developing countries outweighs outdoor air pollution in cities as a global health hazard.

http://www.che.caltech.edu/faculty/seinfeld_j/
http://ehs.sph.berkeley.edu/people/smithk.htm

**Recent Laureates**
2011: Professor Maria R. Berenbaum for her scientific contributions in understanding the coevolution of herbivorous insects and plants, elucidating the physiological and genetic basis for coevolution, for her application of these concepts to agricultural practices, and for sharing her insights on the role of insects in our ecosystems.

(www.tylerprize.usc.edu/pastlaureates.html)

**Tyler Prize Executive Committee**
The Tyler Prize is governed and awarded by the independent Tyler Prize Executive Committee.
(www.tylerprize.usc.edu/executivecommittee.html)

Dr. Owen T. Lind, Chair, Baylor University
Dr. Rosina M. Bierbaum, University of Michigan
Ms. Margaret Carley-Carlson, O.C., World Economic Forum Global Water Agenda Council
Dr. Alan P Covitch, University of Georgia
Dr. Robert A. Frosch, Harvard University and Woods Hole Oceanographic Institution
Dr. Arturo Gómez-Pompa, University of California, Riverside and Universidad Veracruzana
Dr. Judith E. McDowell, Woods Hole Oceanographic Institution
Ms. Julia Marton-Lefèvre, (IUCN) International Union for Conservation of Nature
Dr. Ralph Mitchell, Harvard University
Dr. Cornelius W. Sullivan, University of Southern California

---

www.biosciencemag.org
100 YEARS
Covering science for a century
Subscribe Today • 800.282.0444 • subs@amsci.org • www.americanscientist.org