Ecological Indicators 39 (2014) 94–101

Contents lists available at ScienceDirect

Ecological Indicators

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

Evaluating the efficiency of environmental monitoring programs☆

Carrie R. Levine[a,*], Ruth D. Yanai[a], Gregory G. Lampman[b], Douglas A. Burns[c], Charles T. Driscoll[d], Gregory B. Lawrence[c], Jason A. Lynch[e], Nina Schoch[f]

[a] Department of Forest and Natural Resources Management, State University of New York, College of Environmental Science and Forestry, 1 Forestry Drive, Syracuse, NY 13210, USA
[b] New York State Energy Research and Development Authority, 17 Columbia Circle, Albany, NY 12203, USA
[c] U.S. Geological Survey, 425 Jordan Road, Troy, NY 12180, USA
[d] Department of Civil and Environmental Engineering, Syracuse University, 151 Link Hall, Syracuse, NY 13244, USA
[e] U.S. Environmental Protection Agency, Clean Air Markets Division, 1200 Pennsylvania Avenue NW, Washington, DC 20460, USA
[f] Biodiversity Research Institute, 652 Main Street, Gorham, ME 04038, USA

ARTICLE INFO

Article history:
Received 27 July 2013
Received in revised form
28 November 2013
Accepted 3 December 2013

Keywords:
Uncertainty
Long-term
Lakes
Monitoring
Streams
Loons
Biomass

ABSTRACT

Statistical uncertainty analyses can be used to improve the efficiency of environmental monitoring, allowing sampling designs to maximize information gained relative to resources required for data collection and analysis. In this paper, we illustrate four methods of data analysis appropriate to four types of environmental monitoring designs. To analyze a long-term record from a single site, we applied a general linear model to weekly stream chemistry data at Biscuit Brook, NY, to simulate the effects of reducing sampling effort and to evaluate statistical confidence in the detection of change over time. To illustrate a detectable difference analysis, we analyzed a one-time survey of mercury concentrations in loon tissues in lakes in the Adirondack Park, NY, demonstrating the effects of sampling intensity on statistical power and the selection of a resampling interval. To illustrate a bootstrapping method, we analyzed the plot-level sampling intensity of forest inventory at the Hubbard Brook Experimental Forest, NH, to quantify the sampling regime needed to achieve a desired confidence interval. Finally, to analyze time-series data from multiple sites, we assessed the number of lakes and the number of samples per year needed to monitor change over time in Adirondack lake chemistry using a repeated-measures mixed-effects model. Evaluations of time series and synoptic long-term monitoring data can help determine whether sampling should be re-allocated in space or time to optimize the use of financial and human resources.

1. Introduction

Environmental monitoring is essential for detecting changes associated with biological invasions, land use change, and stressors such as air pollutants and climate change. Monitoring is also needed to evaluate the effects of past or proposed environmental policies and resource management activities (Lovett et al., 2007). Data from ongoing long-term monitoring programs can be used to assess the efficiency and effectiveness of those programs and can inform the development of other similar monitoring programs.

The intensity of monitoring needed to detect trends over space and time is determined by the natural spatial and temporal variation of the measured parameters, measurement and model error, and the acceptable Type I error rate. A parameter with high natural variation requires more intensive sampling to achieve the same statistical power (Garman et al., 2012). Periodic evaluation of monitoring programs is important because the objectives of a monitoring plan may change, the available technology improves, and the amount of data accumulates over time (Lovett et al., 2007). However, most literature on designing ecological monitoring programs is focused on the initial program design (McDonald, 2012; Urquhart, 2012). There has been less focus on strategies for adapting existing monitoring plans to maximize effectiveness and efficiency using the data already collected. In this paper, four case studies are presented to illustrate the evaluation of existing monitoring schemes.

Long-term monitoring of stream chemistry can be used to assess changes in watershed dynamics and to evaluate environmental effects of atmospheric deposition of air pollutants. In the northeastern US, acidic deposition of sulfuric and nitric acids and ammonium result in elevated leaching of nutrient cations from available soil

* Corresponding author. Present address: Department of Environmental Science, Policy, and Management, University of California, Berkeley, 130 Mulford Hall, Berkeley, CA 94720, USA. Tel.: +1 773307 2583.

E-mail addresses: crlevine@berkeley.edu (C.R. Levine), rdyanai@syr.edu (R.D. Yanai), ggl@nyserda.ny.gov (G.G. Lampman), daburns@usgs.gov (D.A. Burns), ctdrisco@syr.edu (C.T. Driscoll), glawrenc@usgs.gov (G.B. Lawrence), lynch.jason@epamail.epa.gov (J.A. Lynch), nina.schoch@briloon.org (N. Schoch).

1470-160X/$ – see front matter © 2014 The Authors. Published by Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.ecolind.2013.12.010

Table 1
Description of the four case studies and models included in this paper.

Case study	Data set includes spatial replication	Data set includes temporal replication	Model used	Parametric or subsampling approach	Research question
Stream chemistry	No	Yes	Regression	Subsampling, followed by parametric	Magnitude of detectable change in slope for one point in space
Loon tissue chemistry	Yes	No	Detectable difference (*t*-test)	Parametric	Detectable change over time based on one sample in time
Forest biomass	Yes	No	Mean	Subsampling	Magnitude of spatial uncertainty for one point in time
Lake chemistry	Yes	Yes	Repeated measures mixed effects model	Subsampling, followed by parametric	Magnitude of detectable change in model mean, including spatial and temporal variability

pools, resulting in deleterious effects on vegetation (Horsley et al., 2000; Battles et al., 2013). Acidic deposition has also had negative effects on water quality in the region by causing decreases in pH and acid-neutralizing capacity (ANC) and increases in aluminum concentrations (Driscoll et al., 2003a).

We use the long-term stream chemistry monitoring at Biscuit Brook in the Catskill Mountains, NY, to illustrate an approach for selecting a sampling plan for a single site with multiple measurements over time. Biscuit Brook has been sampled weekly for 28 years to evaluate the effects of atmospheric acid deposition on stream chemistry (Murdoch and Shanley, 2006). Using this long-term record, we subsample data sets of stream NO_3 and SO_4 concentrations to assess the effects of sampling frequency on the power to detect changes in stream chemistry.

Mercury contamination in water and plant and animal tissues is a major environmental concern. Mercury deposited on the landscape from the atmosphere can be converted into methylmercury, a neurotoxin that bioaccumulates up aquatic and terrestrial food chains and impairs animal and human health (Driscoll et al., 2007). There have been several recent studies of Hg in surface waters and in tissues of aquatic fauna (Yu et al., 2011; Turnquist et al., 2011; Simonin et al., 2006); detecting change over time will require repeating these measurements in the future. To explore sampling strategies for trend detection, we apply a simple power analysis to assess the sampling intensity and interval needed for future surveys, based on expected rates of change. We use the example of a survey of Hg concentrations in feathers and blood in the Common Loon (*Gavia immer*) in the Adirondacks (Schoch et al., 2011) to illustrate this approach.

Forests are routinely monitored to assess stand dynamics, forest health and productivity, and timber and non-timber values (Fahey and Knapp, 2007; Smith, 2002). Methods for accurately measuring forest biomass accumulation have important implications for quantifying carbon sequestration in emerging carbon markets and for evaluating alternatives to fossil fuels (Woodbury et al., 2007; Stupak et al., 2007; Raciti et al., 2012). By assessing the uncertainty due to sample size, we can determine how many plots must be sampled to achieve a desired confidence. To illustrate this approach, we use the example of the reference watershed at the Hubbard Brook Experimental Forest, NH. Because all the plots on the watershed are measured in each census, we can subsample plots using a bootstrapping approach to assess the effect of sampling intensity on the confidence in biomass estimates.

Lake ecosystems have been used to monitor the effects of environmental policies affecting acidic deposition (Driscoll et al., 2003b). The Adirondack region of NY exhibits some of the most severe impacts from acidic deposition in the US, and lake systems there have been monitored to assess recovery of these ecological systems over time (Civerolo et al., 2011). The long-term monitoring of Adirondack lake chemistry is similar to many long-term environmental monitoring programs that involve sampling many sites at regular intervals. It may be the case that the sampling regime is deficient or excessive in either space or time. We apply a repeated measures mixed-effects statistical model to data sets that simulated reduced sampling schemes in space and time.

These four case studies were chosen because they represent a variety of experimental designs common to environmental monitoring programs (Table 1). The objective of this paper is to illustrate approaches for reviewing the efficiency and effectiveness of environmental monitoring schemes. These approaches can be applied to a wide range of sampling designs. We describe best practices for evaluating monitoring programs and give examples of additional approaches that may be useful for analyzing complex long-term monitoring data.

2. Methods

2.1. General linear model error applied to stream sampling for long-term trends at Biscuit Brook

We used a general linear regression and the standard error of the slope to assess uncertainty in long-term temporal trends for Biscuit Brook. Biscuit Brook (latitude 41°59′43″, longitude 74°30′05″) is a water-quality station in the Neversink Reservoir watershed in the Catskill Mountains, NY. Biscuit Brook drains a 963-ha forested watershed with a gauging station at an elevation of 628 m (McHale and Siemion, 2010). The Biscuit Brook station has been in operation since 1983 and has been sampled weekly by the US Geological Survey since 1991. Field and laboratory methods for streamflow and water quality data collection are described by McHale and Siemion (2010).

We subsampled the weekly stream chemistry data from Biscuit Brook (1996–2003) to assess the effects of sampling intensity on the regression standard error and the ability to detect change over time. Subsampled data sets representing 50% of current effort (measurements taken every other week), 25% of current effort (measurements taken every month), or 12.5% of current effort (measurements taken every other month) were used in this analysis.

We used this approach to examine the effects of subsampling on SO_4 and NO_3 concentrations, as these are commonly monitored to assess the effects of atmospheric deposition on long-term stream chemical trends. Sulfate has shown a consistent decreasing long-term trend at Biscuit Brook with little seasonal or flow-related variability (Kerr et al., 2012). In contrast, NO_3 shows a more variable long-term trend with high seasonal and flow-related variability for the period we considered. Thus, these solutes illustrate scenarios for which long-term trends in solute concentration may be more or less difficult to detect due to within-year variation and the strength of the trend over time.

We compared the standard error of the regression across the subsampling scenarios. We also compared the significance of the slopes of the subsampling scenarios to assess the ability of reduced