Environmental Modelling & Software 73 (2015) 148–166

Contents lists available at ScienceDirect

Environmental Modelling & Software

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

A bootstrap method for estimating uncertainty of water quality trends

Robert M. Hirsch [a,*], Stacey A. Archfield [a], Laura A. De Cicco [b]

[a] U.S. Geological Survey, 432 National Center, USGS, Reston, VA 20192, USA
[b] U.S. Geological Survey, 8505 Research Way, Middleton, WI 53562, USA

A R T I C L E   I N F O

Article history:
Received 23 July 2015
Accepted 30 July 2015
Available online xxx

Keywords:
Water quality
Bootstrap
Trend
Uncertainty analysis

A B S T R A C T

Estimation of the direction and magnitude of trends in surface water quality remains a problem of great scientific and practical interest. The Weighted Regressions on Time, Discharge, and Season (WRTDS) method was recently introduced as an exploratory data analysis tool to provide flexible and robust estimates of water quality trends. This paper enhances the WRTDS method through the introduction of the WRTDS Bootstrap Test (WBT), an extension of WRTDS that quantifies the uncertainty in WRTDS-estimates of water quality trends and offers various ways to visualize and communicate these uncertainties. Monte Carlo experiments are applied to estimate the Type I error probabilities for this method. WBT is compared to other water-quality trend-testing methods appropriate for data sets of one to three decades in length with sampling frequencies of 6–24 observations per year. The software to conduct the test is in the EGRETci R-package.

Software

The statistical procedures presented here are all based on the Weighted Regressions on Time, Discharge, and Season (WRTDS) approach to water quality data analysis. The WRTDS is implemented in the EGRET (Exploration and Graphics for RivEr Trends), R-package (open source) available from the Comprehensive R Archive Network http://cran.r-project.org/web/packages/. The new software that implements the WRTDS Bootstrap Test (WBT) described in this paper is also an R-package called EGRETci, also available from the Comprehensive R Archive Network.

1. Introduction

More than 40 years after the passage of the Clean Water Act in the United States, large public investments and significant regulatory actions continue to be made in order to continue making progress towards the goals set forth in the Act (Knopman and Smith, 1993; Copeland, 2006). Public officials, land owners, and the general public express concern over perceived deterioration of water quality and seek to determine the magnitude of the impact that public and private investments and regulatory actions are having on the attainment of water quality goals (Broussard et al., 2012; National Research Council, 2011; Mehan, 2012) in order to decide about investing in further actions. On-going evaluations of the direction and magnitude of water quality trends remains an important task to support the achievement of water quality goals.

Various statistical methods have been used for more than 30 years to explore and analyze temporal trends in water quality. More recently, these methods have advanced as a result of several factors: increased lengths of consistent data sets, improvements in statistical methods, improvements in computer software and hardware, observations of a wide range of multidecadal trends in water quality, and improved understanding of watershed-based and in-channel processes affecting water quality. Examples of some of these methods include: Richards and Baker (2002), Langland et al. (2007), Ryberg et al. (2014), and Corsi et al. (2015). A part of these advancements has been the introduction of new approaches that stem from exploratory data analysis and smoothing concepts, including adaptation of locally weighted scatterplot smoothing (LOESS) (Cleveland and Devlin, 1988), and generalized additive models (GAMs) (Wood, 2006) to surface water quality data (see for example Reckhow and Qian, 1994; Langan et al., 2001; Morton and Henderson, 2008; and Hirsch et al., 2010). These methods are primarily aimed at a desire to characterize the timing, magnitude, and general nature of the trends observed.

Not surprisingly, there is an interest among many water quality professionals to have descriptions of trends be accompanied by statements of statistical significance, including confidence intervals

* Corresponding author.
E-mail addresses: rhirsch@usgs.gov (R.M. Hirsch), sarch@usgs.gov (S.A. Archfield), ldecicco@usgs.gov (L.A. De Cicco).

http://dx.doi.org/10.1016/j.envsoft.2015.07.017

on the amount of change observed (e.g. Boesch et al., 2005). This interest is very legitimate. For example, the analysis may say that the mean concentration of nitrogen at a given monitoring site has increased by 1 mg/L over the past 30 years. Recognizing that typical monitoring strategies may only sample 6 to 12 times per year, one can expect that the estimate of 1 mg/L change is highly uncertain. If the analyst can state that the 90 percent confidence interval around that value ranges from 0.9 to 1.1 mg/L this relatively narrow range of uncertainty should provide a much stronger basis for action as compared to a result which states that the 90 percent confidence interval runs from −0.5 mg/L to 2.5 mg/L. This latter result suggests that although the likely direction of change is positive, there is actually a non-trivial chance that concentrations have not increased over the 30-year period. In this case, decision-makers may be inclined to exercise more caution in committing public or private resources to remedy the situation.

Whereas the need for such confidence interval estimates and associated statements of attained significance levels is great, it is not a simple matter to provide such estimates when the method of analysis is an exploratory approach that makes very few assumptions about the statistical properties of the data. This paper delivers an approach to adding uncertainty analysis to one particular exploratory data analysis method: Weighted Regressions on Time, Discharge, and Season (WRTDS) (Hirsch et al., 2010). We use a bootstrap (Diaconis and Efron, 1983; Efron and Tibshirani, 1994) procedure to provide complimentary uncertainty information along with the graphical and numerical outputs already provided by the WRTDS method. We call this the WRTDS Bootstrap Test (WBT). This paper briefly reviews the WRTDS method, and then describes the WBT. The bootstrap procedure used here is a new type of block bootstrap designed to account for the influence of serial correlation on the test results without attempting to explicitly model the correlation structure. Modeling the serial correlation of these kinds of water quality data sets can be very problematic given the relatively sparse and often irregular sampling that is common to such data sets. The block bootstrap approach introduced here approximately preserves the serial correlation for lags on the order of weeks to months and thus achieves Type I error rates that are relatively close to the nominal Type I error rate.

The block bootstrap approach is evaluated using a set of Monte Carlo simulations to estimate the Type I error probability (probability of detecting a trend when a trend was not present) as compared to the nominal significance level for this method under the null hypothesis that water quality conditions have not changed over the period of analysis. Type II error (the probability that a trend is present but not detected), although of great importance, was not evaluated here because of the multitude of different possible manifestations of departures from the null hypothesis that are possible. These include different rates of change, step functions versus ramp functions, and trends driven by point source changes versus those driven by non-point sources. WRTDS is designed to be sensitive to a variety of different types of trend scenarios, whereas most of the more common types of trend tests assume a simple and rather rigid model of the trend. Thus it is reasonable to assume that the WRTDS method will have an advantage in terms of Type II errors for a wide range of trend scenarios, but some disadvantage when the trend scenario postulated adheres closely to the assumptions around which other tests were designed. The wide range of trend scenarios would add greatly to the complexity of this study and may not be very illuminating. Thus, we kept our inquiry to the narrower question: is the WBT test accurate in terms of Type I error? The Monte Carlo simulations are based on three different generating models for discharge and concentration that are designed to replicate the statistical properties seen in actual water quality records. The Type I error probability resulting from the WBT is compared to the Type I error probability resulting from three common alternative trend analysis procedures: these are a multiple regression approach, the Seasonal Kendall test on residuals from a flow–concentration relationship, and the Seasonal Kendall test adjusted for serial correlation. These Monte Carlo simulations are further used to provide a suggested block-length for the test. Lastly, an example data set is evaluated using the WBT and several approaches for communicating uncertainty are presented. Because the WRTDS method is fundamentally an exploratory data analysis method, software that is relatively fast and interactive is crucial to the effective use of the method. Addition of the WBT analysis to define uncertainties has the potential to slow down that rapid interactive process. The desire to obtain the uncertainty information in a timely manner motivates the particular pathway this software development follows: aimed at providing useful uncertainty information without greatly slowing the overall analytical process. Hence the WBT uses some novel approaches to maximize computational speed. An important aim of this paper is to demonstrate (through Monte Carlo testing) that these approaches do not significantly compromise the validity of the test.

2. Overview of WRTDS method

The motivations for the WRTDS method and details of its computational techniques are described in Hirsch et al. (2010) and Hirsch and De Cicco (2014); many implementation details are omitted here in the interest of brevity. New notation and explanations of the method not published previously are presented throughout Section 2 in order to provide the concepts and mathematical symbology needed to explain the uncertainty analysis presented in Section 3. The WRTDS method has been implemented within an R package, know as EGRET (Exploration and Graphics for RivEr Trends) and is available on the Comprehensive R Archive Network (CRAN) at http://cran.r-project.org/web/packages/.

Major features of WRTDS include the following:

- It can detect and describe temporal trends that may not conform to linear or quadratic functional forms.
- It is suitable for use with irregularly spaced data.
- It does not assume that the discharge versus concentration relationship has the same shape throughout the period of record.
- It does not assume that the concentration residuals are homoscedastic.
- It does not assume that the seasonal pattern remains the same over the period of record.
- It can assess both concentrations and fluxes, recognizing that the trends in each of these measures of water quality can be quite different and even of different sign.
- It can not only provide estimates of the time series of annual mean concentrations and fluxes, but also time series of “flow-normalized” mean concentrations and fluxes which integrate over the probability distribution of discharge to remove the effect of interannual streamflow variability.

2.1. WRTDS estimation of daily concentration

The WRTDS model utilizes the sampled water quality data from an individual sampling site, along with the daily mean discharge at that site for the sampling dates, to develop an estimate of the concurrent daily mean concentration given by: