Appendix 3. Methods—Flow-Adjusted Trend Analysis with Tobit Regression in the S-ESTREND System

Contents

Appendix 3. Methods—Flow-Adjusted Trend Analysis with Tobit Regression in the S-ESTREND System

By R. Edward Hickman

Flow-Adjusted Trend Analysis

U.S. Geological Survey (USGS) water-quality stations were selected for analysis of flow-adjusted trends in total nitrogen and total phosphorus based on the availability of nutrient concentration data during water years 1993–2003 and the availability of a streamflow measurement associated with each water-quality value. Trends in dissolved nitrite-plus-nitrate and dissolved ammonia were analyzed at stations in four basins selected for analysis of point source information, if sufficient water-quality data were available.

Trends were identified with the S-PLUS version of the ESTREND system (Slack and others, 2003). The ESTREND program (Schertz and others, 1991) was developed by the USGS to identify trends in water quality in streams. The program provides three methods to identify trends, and Tobit regression was the method selected to identify trends in this study.

Schertz and others (1991) recommend the amount of data required by the Tobit regression method to determine trends for periods longer than 5 years. First, a minimum of 10 detected observations are needed. Second, a minimum user-specified percentage of the total number of observations in the record must be detected observations; for this study, this minimum value was set to 20 percent. Third, a minimum of one observation per year had to be present in the beginning and ending fifths of the period. Additional requirements of (A) at least one measurement during water years 1996–2000, and (B) no more than four continuous water years with no observations, were used to screen out datasets with observations only at the beginning and end of the period.

Trend tests were conducted for periods ranging from 8 to 11 water years, depending on the availability and distribution of data. Trend test results for a nutrient at a station are reported where sufficient numbers of measurements were available to start the period of the trend test during either water years 1993, 1994, or 1995 and end the period of the test during either water years 2001, 2002, or 2003.

Relations between nutrient concentrations, year, streamflow, and season were developed with Tobit regression for each nutrient at each station with sufficient measurements. The general equation relating water quality to year, streamflow, and season is:

$$\log(C) = B_0 + B_1 * Year + B_2 * \log(Q) + B_3 \sin(2\pi t) + B_4 \cos(2\pi t) \quad (1)$$

where

C = concentration of nitrogen or phosphorus compound,
B_0 = intercept,
B_1 = coefficient for year,
$Year$ = year,
B_2 = coefficient for streamflow,
log = base-10 logarithm,
Q = streamflow, in cubic feet per second,
B_3 = first season coefficient,
sin = sine,
t = fraction of water year prior to month and day of measurement,
B_4 = second season coefficient, and
cos = cosine

The term for streamflow was included in each test. The seasonal terms were included only if (A) seasonality appeared to be important based on an examination of a plot of residuals as a function of season, and (B) if the coefficients for the seasonal terms were significant at the 0.05 level.

Trends were identified if the coefficient for year (B_1) was different than zero at the 0.05 level of significance. If B_1 was greater than zero, concentrations increased over time; if B_1 was less than zero, concentrations decreased over time. For each nutrient at each station, the magnitude of the change in concentration is reported in percent of mean per year.

Changes in the variance of the residuals over the years were considered as part of each trend test. The level of significance reported for the coefficient for year (B_1) is accurate only if the variance of the residuals is constant over time. Results of a few tests were not reported if the level of significance of coefficient for year was close to 0.05 and if the variance of the residuals did not appear to be constant during the period of the test.

The effects of outlying measurements (outliers) on the results of the trend tests also were considered. For some tests, outliers were removed from the dataset and the dataset was then retested. In some cases, therefore, results presented do not include outliers.

References

Schertz, T.L., Alexander, R.B., and Ohe, D.J., 1991, The computer program Estimate Trend (ESTREND), a system for the detection of trends in water-quality data: U.S. Geological Survey Water-Resources Investigations Report 91–4040, 72 p.

Slack, J.R., Lorenz, D.L., and others, 2003, USGS library for S-PLUS for Windows–Release 2.1: U.S. Geological Survey Open-File Report 2003–357.