Environmental Water Quality Data: Analyzing Change Over Time (Newsletter)*

October 5, 2012

Water quality discharge permits require periodic measurement of water quality constituent concentrations to document compliance with permit conditions. In addition to correctly including concentrations below the analytical method's detection level when describing the distribution of these concentrations there is great value for operators and environmental regulators in properly analyzing the temporal aspects of these data.

A single concentration, particularly when it exceeds a standard's threshold, lacks context and does not assess an operation's interactions with the natural environment; it is a temporal and spatial snapshot. Decisions require more understanding of the dynamics involved, and one highly useful tool is time series analysis.

Time series are analyzed to understand the past and predict the future. Forecasts assist operators and environmental regulatory managers to make informed decisions that are practical and justified. Time series analysis quantifies both the important features of the data and inherent natural variability. The main features of environmental time series are trends and periodic variations. Environmental time series frequently have irregular sampling intervals, usually because of adverse weather or water flow extremes, which requires adjustments in the statistical models used to analyze them.

There are three main approaches to modeling time series. The first is based on an assumption of a fixed seasonal pattern in addition to a trend. The trend is the focus and can be estimated by local averaging of the deseasonalized data.

The second approach describes seasonal variation and overall trend in terms of a level and slope. Both seasonal influences and long term trends are estimated by exponentially-weighted averages.

The third approach fits mathematical probability distributions to time series data. The differences between predicted and observed values is called a residual errors series. When the model closely fits the data the residual error series represents natural random variability contributing to the observed time series. If there is structure to the residual errors this information is used to improve the forecast model.

The mathematical distribution models applied to time series data include white noise (strictly random variables), random walks (highly autocorrelated data that produce trends), and autoregressive models (values at a time highly dependent upon past values, either positively or negatively).

It is comparatively easy to understand the benefits of applying time series models to historic data for operational decisions. This use is common by business managers and public policy makers. The benefits for environmental managers and regulators may be less obvious.

When a water chemistry constituent sporadically exceeds a standard threshold, how likely is this to occur in the future? Are costly remedial actions necessary to decrease the likelihood of exceeding the threshold? Time series analysis can answer these questions. Another use is evaluating rates of change of water treatment or remediation activities and identifying lag times from when actions were taken and monitoring changes are seen. These situations, and others, save time and money by supporting better informed decisions and avoiding unnecessary costly modifications of operations and facilities.