

Practical Stats Newsletter for Dec 2011

Subscribe and Unsubscribe: <http://practicalstats.com/news/>

All of our past newsletters:

<http://www.practicalstats.com/news/news/bydate.html>

In this newsletter:

1. Upcoming courses in 2012
2. The Effect of Serial Correlation on the Mann-Kendall Test for Trend
3. New Book on Methods for Censored Environmental Data

1. Upcoming courses in 2012

Time Series and Forecasting

Test groups, trends, etc. with frequently collected, "real-time" data.

March 12-13, 2012

Hilton Suites Atlanta Perimeter [link for info](#)

\$895 through Feb 24th, \$995 after

\$795 if registered for both courses

Untangling Multivariate Relationships

Turn confusion into recognizable patterns.

Relate a suite of chemical compounds to a suite of bio community structure data.

March 14-15, 2012

Hilton Suites Atlanta Perimeter [link for info](#)

\$895 through Feb 24th, \$995 after

\$795 if registered for both courses

Applied Environmental Statistics

Statistics, down to earth.

Hypothesis tests, regression done right, and trend analysis.

June 25-29, 2012

Doubletree Tampa Bay-North Redington Beach

\$1395 through May 31st, \$1495 after

Our **Nondetects And Data Analysis** class will likely be offered in May-June 2012, in an as yet undisclosed location. It will reflect the new material from Helsel's February 2012 release of *Statistics for Censored Environmental Data Using Minitab and R*, the second edition of his book on interpreting nondetect data.

We'll be conducting a two-day "Environmental Stats using ProUCL" course for Health Canada in Vancouver, BC in January. Contact us if you'd like to have this taught at your site. Several other 2012 conference workshops on nondetects or environmental statistics are in the works.

You can always find our complete course listing on our "Upcoming Classes" page at http://www.practicalstats.com/new_classes/classes.html

2. The Effect of Serial Correlation on the Mann-Kendall Test for Trend

One of the newer topics in our Time Series and Forecasting course is how to adjust the nonparametric Mann-Kendall test for trend when data are serially correlated. We discussed the general effect of serial correlation on hypothesis tests in our January 2009 newsletter (available for download from our website). Measurements for streamflow and water quality, among others, are today frequently recorded using automatic methods every 15 minutes or less. Data collected this frequently often have a strong memory (positive serial correlation) from one measurement to another. This correlation violates the independence assumption of hypothesis tests, including the Mann-Kendall test for trend. The number of observations with independent information from the previous observation is fewer than the number of measurements made. Due to this serial correlation, if the sample size n is set to the number of measurements made, p -values reported by the trend test will be too small, and trends will be found more often than they actually occur. Software for adjusting the Mann-Kendall test for serial correlation is not easily available (outside of the software we provide to students in our course).

For example, Figure 1 below plots total organic carbon data over the first few months of 2008. Is there a general increase or decrease during this period? Data are collected daily, if not more frequently. The Mann-Kendall test without adjusting for serial correlation returns a p -value of 2×10^{-7} , an extremely small value, indicating a trend is present. However, the sequential concentrations have a strong carryover or memory, and so are serially correlated. The number of independent observations is much lower than the number of measurements made. Adjusting the sample size to its independent or ‘effective’ number as first recommended by Matalas and Langbein (1962) in the *Journal of Geophysical Research* (vol 67, no 9), the corrected p -value is 0.22, resulting in the conclusion of no overall trend, a much more reasonable conclusion given that the total length of record is only several months long.

More information on the correction for serial correlation, and the software to perform it, is available when you take our *Time Series and Forecasting* course. We discuss methods for first adjusting for climatic signals. We discuss the appropriate way to adjust the test for serial correlation, along with ways proposed through the years (some are called “pre-whitening”) that can actually make matters worse, inadvertently stripping out some of the trend while trying to perform a correction for serial correlation.

Hope to see you there sometime soon, maybe in Atlanta this March?

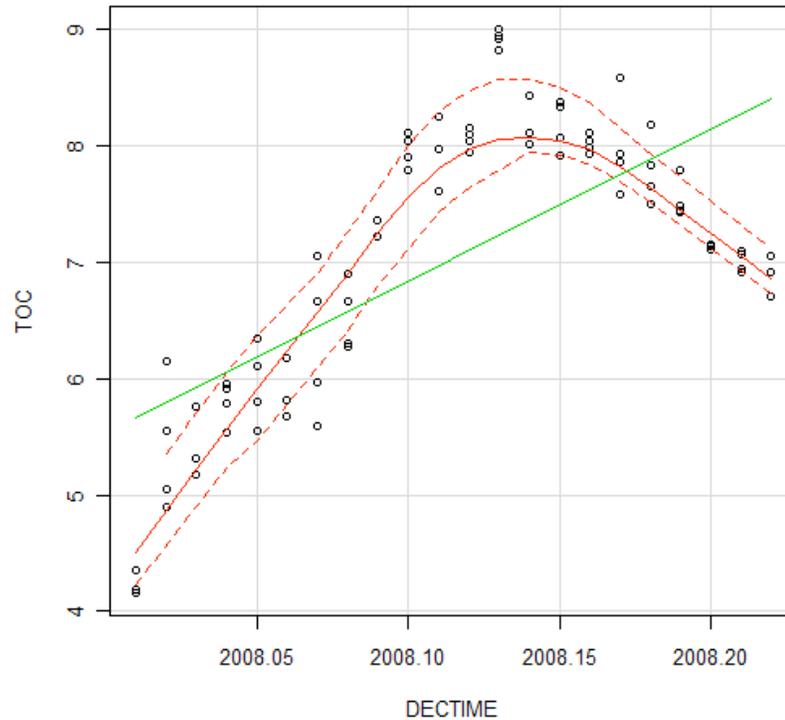


Figure 1. TOC data in the first quarter of 2008

3. New Book on Methods for Censored Environmental Data

The new, second edition of Dennis Helsel's book for handling nondetect data is out in early February 2012. The new title, *Statistics for Censored Environmental Data using Minitab and R*, doesn't have as nice of an acronym as the old one, but emphasizes that it contains methods for censored data that are not true nondetects – above the detection limit but below a reporting limit. The new edition adds info on incorporating nondetects in multivariate procedures, using interval-censored data such as values between the detection and quantitation limits, and examples using both Minitab and R software throughout the text. It includes ways to avoid common pitfalls -- the first chapter is titled “Things people do with censored data that are just wrong”. There's also an explanation of the mysterious "invasive data" that piggybacks with numbers like one-half the detection limit which could right now be affecting your database ecosystem.

<http://www.wiley.com/WileyCDA/WileyTitle/productCd-0470479884.html>

http://www.amazon.com/Statistics-Censored-Environmental-Minitab-Practice/dp/0470479884/ref=sr_1_1?ie=UTF8&qid=1324395185&sr=8-1

'Til next time,
Practical Stats (Dennis Helsel)
-- Make sense of your data