Data Analysis Protocols for the Trend Network, Version 2.0

Guidelines for Florida's Trend Monitoring Network

Division of Environmental Assessment and Restoration

Florida Department of Environmental Protection

April 2019

2600 Blair Stone Rd, MS 3560

Tallahassee, Florida 32399-2400

https://floridadep.gov/

Information Goals for the Trend Monitoring Network

The Trend Network design allows for determination of trends in individual waterbodies but may be extended to (1) the state of Florida as a whole, (2) regions of the state, and (3) large drainage basins, or drainage basin complexes (i.e., zones). The Trend Network design also allows for extrapolation of results in conjunction with the Status Network and can address statewide and large-scale (zonal) questions with a high degree of statistical confidence.

The Trend Monitoring Network consists of stations to monitor Florida's surface and groundwaters. The Surface Water Trend Monitoring Network consists of 78 fixed-location stations sited along rivers, streams, and canals throughout the state. WMS samples all stations monthly. Most stations are located at non-tidal regions, and where possible, at or near existing USGS gauging stations. At some stations located at the lower end of a watershed, the department obtains biology, chemistry, and loading data that integrates land use activities. WMS uses Surface Water Trend Network stations located at or near the Florida state boundary with Alabama and Georgia to obtain data for major waters entering Florida.

The Groundwater Trend Network consists of 51 fixed stations used to obtain chemistry and field data in confined and unconfined aquifers. WMS collects water samples quarterly at all wells in this network. Staff measure field analytes monthly at unconfined aquifer stations and quarterly at confined aquifer stations. A micro land use form, completed at all stations annually, aids in determining potential sources of contamination for groundwater resources.

WMS designed the Trend Network to accomplish five goals:

Goal 1. Statistically describe the analytes / indicators.

WMS relies on nonparametric statistics for most water quality data analyses. Nonparametric statistics do not rely on data belonging to a particular distribution. These techniques rank data from the smallest value to highest value and then use these ranks to describe and analyze the data. This process is insensitive to outliers and missing values, which are common in environmental data.

Data analysts create tables listing nonparametric descriptive statistics for each station's analytes (**Table 1**). These tables allow analysts to examine data distribution prior to analysis.

Table 1. Example of descriptive statistics for a Groundwater Trend Network well

This is a nine-column table. Column 1 lists the analyte, Column 2 lists the measurement unit, Column 3 lists the number of samples, Column 4 lists the number of values Below Detection Limits (BDLs), Column 5 lists the minimum value, Column 6 lists the first quartile value, Column 7 lists the median value, Column 8 lists the third quartile value, and Column 9 lists the maximum value.

NO3 -- nitrate; o-PO4 – ortho-phosphate; mg/L – milligrams per liter

Analyte	***Measurement Unit***	***Number of Samples***	***Number of BDLs***	***Minimum Value***	***25th Percentile Value***	***Median Value***	***75th Percentile Value***	***Maximum Value***
NO3	mg/L	30	7	0.05	0.09	1.00	2.00	10.30
o-PO4	mg/L	29	4	0.05	0.10	0.10	0.15	1.30

In **Table 1**, the reported minimum concentration value reflects the minimum detection level reported by the laboratory. Limitations of laboratory instruments can cause low concentrations of analytes to be undetected. This is referred to as data censorship. When the analyte level in a sample is lower than the minimum detection level of an instrument, the measurement is marked as BDL (below detection limit). The DEP laboratory marks censored / BDL measurements with a qualifier (**Appendix**). During analyses the data analyst replaces BDL values with either the highest reported minimum detection level of the instrument, a maximum likelihood estimation (MLE), or the minimum value in the period of record.

If a period of record is missing a data point, the statistics (median, quartiles, ranks) are based on the number of actual values in the record. For example, if the December sample is missing from a monthly dataset, the median for the year is based only on the January-November values. If missing values cause the margin of error to exceed 15%, no analyses are conducted.

Goal 2. Determine if seasonality exists in the time series data.

Seasonal cycles, such as periodic droughts and floods, can mask trends in water quality indicators. Water constituents commonly show seasonal cycles (seasonality); therefore, an effort should be made to collect at least one sample in each season, or four per year at a minimum. Statistically valid trends may be determined only after adjusting the data for seasonality. This adjustment is referred to as deseasonalizing the data.

Using the Seasonal-Kendall (S-K) test (see below, **Goal 3**) removes effects of seasonal cycles. For quarterly data, the conducted tests assume that each quarter is a season. The quarters are: (1) December–February; (2) March–May; (3) June–August; and (4) September–November.

Goal 3. Determine if linear time series trends exist. When present, determine the slopes and interpret the results.

During the conceptual development of the Trend Monitoring Network, there was a concern that a linear trend in a time series might be found for one period, while another trend might be found for a different period at the same station. For example, suppose an upward trend was found for the first half of a time series, while a downward trend was found for the second half, or vice versa. Detectable trends depend on the period for which they are analyzed.

The S-K test is conducted at each station, for each analyte. This test adjusts the dataset for seasonality when it exists. The S-K test allows for missing values. BDLs are assigned an arbitrary value equal to the highest detection level for the period of record. For this application, each test is two-sided. The null hypothesis is no monotonic linear slope is present in the concentration. The α-level for determining the significance of the results is 0.05.

If a trend exists, its direction is determined using a Sen-Slope (S-S) estimator and WMS analysts produce a time series plot of the data. Smoothing procedures make visual inspections of time series plots easier to interpret. For each trend, WMS evaluates possible reasons for an analyte trend. If the trend is believed to be human-induced, staff discuss plausible reasons for the trend. *Goal 4. Summarize the major water quality issues and recommend sampling strategies.*

Goal 4. Summarize the major water quality issues and recommend sampling strategies.

WMS summarizes the major water quality issues for each Trend Monitoring Network station using the information generated from Goals 1 through 4. This can lead to recommendations on the future sampling of both ground and surface waters.

Appendix: Qualifiers

Qualifiers are codes that flag irregular or questionable measurements. Data may be qualified for a variety of reasons described in **Table 2** (from 2018 revision of Chapter 62-160.700, F.A.C.).

Table 2. Qualifiers used by the Status and Trend Monitoring Networks, beginning October 1, 2017

This is a two-column table. Column 1 lists the qualifier code and Column 2 lists a description.

Note: Prior to October 1, 2017 the qualifier codes used were those listed in the 2014 revision of Chapter 62-160.700 F.A.C.

Code	Definition
A	Value reported is the arithmetic mean (average) of two or more determinations.
B	Results based upon colony counts outside the acceptable range. This code applies to microbiological tests and specifically to membrane filter colony counts. The code is to be used if the colony count is generated from a plate in which the total number of coliform colonies is outside the method indicated ideal range.
G	Indicates that the analyte was detected at or above the method detection limit in both the sample and the associated field collected blank, and the value of the blank is greater than 10% of the associated sample value.
I	The reported value is greater than or equal to the laboratory method detection limit but less than the laboratory practical quantification limit.
J	Estimate value. **Shall be accompanied by a detailed explanation** to justify the reason(s) for designating the value as estimated. Examples of situations in which a "J" code must be reported include: instances where a quality control item associated with the reported value failed to meet the established quality control criteria (the specific failure must be identified); instances when the sample matrix interfered with the ability to make any accurate determination; instances when data are questionable because of improper laboratory or field protocols (e.g., composite sample was collected instead of a grab sample); instances when the analyte was detected at or above the method detection limit in an analytical laboratory blank other than the method blank (such as calibration blank) and the value the blank is greater than 10% of the associated sample value; or instances when the field or laboratory calibrations or calibration verifications did not meet calibration acceptance criteria.
K	Off-scale low. The actual value is known to be less than the value given. (Only used for lab analyses.)
L	Off-scale high. The actual value is known to be greater than the value given. (Only used for lab analyses.)
N	Presumptive evidence of presence of material; component tentatively identified based on mass spectral library search or there is an indication that the analyte is present, but quality control requirements for the confirmation were not met.
O	Sampled but analysis lost or not performed.
Q	Sample held beyond the accepted holding time. Value is derived from a sample that was prepared or analyzed after the approved holding time restrictions for sample preparation or analysis.
R	Significant rain (typically in excess of ½ inch) in the past 48 hours, which might contribute to a lower or higher than normal value.
S	Secchi disk visible to bottom of waterbody. The value reported is the depth of the waterbody at the location of the Secchi disk measurement.
T	Value reported is less than the laboratory method detection limit. Value reported for informational purposes only and shall not be used in statistical analysis.
U	Indicates that the compound was analyzed for but not detected. The reported value shall be the method detection limit.
V	Indicates that the analyte was detected at or above the method detection limit in both the sample and the associated method blank and the blank value was greater than 10% of the associated sample value.
X	Indicates, when reporting results from a Stream Condition Index Analysis (LT 7200 and FS 7420), that insufficient individuals were present in the sample to achieve a minimum of 280 organisms for identification (the method calls for two aliquots of 140-160 organisms), suggesting either extreme environmental stress or a sampling error.
Y	The laboratory analysis was from an unpreserved or improperly preserved sample. The data may not be accurate.
Z	Too many colonies were present for accurate counting. Historically, this condition has been reported as "too numerous to count" (TNTC). The "Z" qualifier code shall be reported when the total number of colonies of all types is more than 200 in all dilutions of the sample tested using a membrane filter technique. When applicable to the observed test results, a numeric value for the colony count for the microorganism tested may be estimated from the highest dilution factor (smallest sample volume) used for the test and reported with the qualifier code.
!	Indicates that the reported value deviates from historically established concentration ranges.
?	Data are rejected and should not be used. Some or all of the quality control data for the analyte were outside criteria, and the presence or absence of the analyte cannot be determined from the data.