3. Trend Monitoring Network

(1) Introduction

In a broad sense, representative water quality data from any station with a sufficient record can be used to evaluate site-specific mid- or long-term temporal changes in water quality (i.e., trends). Short-term, mid-term or long-term water quality changes are of interest in evaluating the effectiveness of mitigation measures implemented to resolve water quality problems. Sampling sites explicitly designated as trend stations are carefully selected to provide the data necessary for answering precise, formulated questions about specific water quality parameters on specific geographic scales (watershed, drainage basin, geographical region, aquatic resource type, etc.).

Central to evaluating the quality of the Commonwealth's surface waters are efforts to detect changing conditions over time. The most often asked question can be simply stated as: "Is the water getting better or worse?" Implicit in this fundamental question is the comparison of water quality during some period in the past with today's conditions. Although this question appears simple in its form, in order to scientifically measure trends the question(s) become(s)… 'Is there an increasing trend, decreasing trend, or no trend at a specific location for each individual water quality variable?' and 'How much confidence do we have in our response to the first question?' The challenge is to design a network of stations to answer these questions and to collect and analyze the data with this purpose in mind. The estimation of water quality trends facilitates predicting future adverse conditions and measuring the progress of restoration efforts.

Changes in water quality over time may indicate improvement or worsening of conditions. Water quality changes are evaluated on individual variables from specific sites that have large (long-term) data sets. To measure trends in water quality, several requirements are necessary to assure the ecological and statistical significance of the interpretation. The first requirement is related to the location of trend stations. Stations are chosen to be representative of major river segments, usually in free-flowing sections. For example, stations may be located at the fall lines (head-of-tide) or at the mouths of major rivers. For free flowing waters, stations ideally would be collocated with discharge gages. Another requirement is that sufficient data must exist to determine the significance of any trend detected. Generally, stations must have a minimum of ten years of uninterrupted monthly or bimonthly data to satisfy this requirement. As more data accumulates over a longer period of time, our ability to detect significant trends improves. Trend stations are often selected based upon their having a long, continuous history of monitoring. Some DEQ stations have been monitored for 30 years or more, and it is these sites that are good candidates for the statistical analysis of trends. As of 2013, DEQ had conducted monthly or bimonthly monitoring at 454 trend stations. Of these stations 398 had been sampled 100 or more times. Of these 398 stations 255 had records that began prior to 1986.

Trends, as discussed here, are always calculated for specific variables at individual stations. Some variables may be combined, as is the case of total nitrogen, where trend for total nitrogen is calculated from the summation of total nitrate, total nitrite, and total Kjeldahl nitrogen. Other variables, like dissolved oxygen, may have been measured using the Winkler method during the early days of monitoring but are now measured by one type or another of electronic probes - these are combined into a single variable and converted to percent saturation prior to analyzing trend.

The timing of sampling is important for various reasons. The values of water quality parameters often fluctuate considerably for different reasons and on various time scales. Biological characteristics, such as chlorophyll concentrations or the density of zooplankton, may vary in diurnal (i.e., day vs. night) or semi-diurnal (e.g., tidal) cycles. This may often be equally true of chemical parameter concentrations. Annual changes in precipitation (wetter and drier seasons), in temperature, and in land use practices may induce significant changes in chemical concentrations and certainly influence the abundance and predominant life stages of the biological components of aquatic habitats. Consequently, the timing of sample collection must be kept relatively constant from day to day, month to month and year to year, in order to separate the effects of cyclic variations over shorter time intervals from the significant long-term changes in water quality that trend analyses are designed to detect.

Although trends are initially calculated on single variables at individual stations, trend analysis results for selected parameters may subsequently be integrated and presented as characterizations for larger, more comprehensive regions or resource categories, if the site selection and monitoring methods were established with these objectives in mind. The site and parameter selection, sampling methods and frequencies, laboratory analytical methods and data reduction methods must be standardized in order to combine data from various stations for representative watershed, river basin, regional or resource characterizations or for comparisons among trends in different aquatic resource classes. Site selection criteria for trend stations also depend upon the type of aquatic resource being monitored (e.g., free-flowing streams, lakes, tidal saltwater, etc.).

(2) Purpose

Trend stations are established to provide the data for detecting and evaluating tendencies in long-term water quality changes. Data collected at trend stations serve to attain WQM Strategy Objectives: I.A (1 - 6), I.B (7), and I.C (10). The trend station network is considered to be a high priority monitoring activity. The sensitivity and consequent statistical significance of trend analyses suffer when data sets are interrupted.

(3) Monitoring Design Siting:

It is desirable to site trend stations geographically so that they are equitably distributed throughout the state, as is done to insure comprehensive geographic coverage with the Watershed Monitoring Network. One method of accomplishing this objective is to distribute trend station sites within USGS 8-digit Hydrological Units (4^{th} Order Sub-Basins) in proportion to the total land area (or total stream miles) in each unit. A summary of the USGS Sub-Basins within the state of Virginia, together with their individual land and water areas and the estimated numbers of trend monitoring stations that might be desired, is presented in the table "Virginia Land-Water Area per USGS 8-Digit HUC" [III-A-1c-1.xls]. In the table, the estimated numbers for trend monitoring sites per HUC are based on statewide totals of approximately 200, 300 and 400 permanent trend-monitoring stations, calculated on the basis of land area per HUC, and 400 statewide based on miles of free-running, non-tidal streams per HUC. That would produce an average density of approximately one trend station per 200 square miles of land surface, and one station per 125 stream miles. The numbers of sites actually designated as permanent trend stations depend upon numerous other factors, as discussed below.

In free-flowing streams and rivers, it is desirable to site trend stations in association with flow gauging stations or at locations where flow can be accurately interpolated from gauging stations in the same or in one or more adjacent drainages. In free-flowing, non-tidal waters, the volume of water (flow or discharge rate) passing the sampling site is one of the most important water quality parameters. Flow measurements are required to calculate "loadings" (the total quantity of a substance present in the water column) from the concentrations (quantity per unit volume) generally used to express water quality standards and to measure chemical water quality parameters. Chemical concentrations may vary when the volume of water in an aquatic system changes, even if the total quantity of a substance in the system remains constant. Conversely, constant concentrations when flow rates vary indicate that the total amount of pollutant in the water column is changing! Consequently, trend stations are best located adjacent to discharge gauging stations so that concentrations can be related to the total volume of water passing a site. When stream conditions at a gauging station are inadequate for other monitoring considerations (e.g., the sampling of benthic biological communities), a trend station may be sited elsewhere if the discharge rate at the site can be interpolated from one or more upstream and/or downstream gauging stations. Accurate interpolation, however, may be difficult if numerous tributary streams enter the system between the trend station and the gauging site(s). Time lags or advances in discharge volumes between the trend station and an upstream or downstream gauging station may also necessitate advance data preprocessing to insure accurate representation of flow at the point of sample collection.

In tidal systems where ebb and flow alternate on a semi-daily basis, and in lakes or reservoirs where water residence time is significant and the relative rates of flow may be negligible, total volume may be relatively constant. In such cases, the total volume of water present in the system may be as important as the discharge rate at free-flowing sites. The total loadings entering and leaving a lake through streams, and the flux of tributary waters passing through tidal systems, however, still provide important information for water quality assessment and such data should be collected whenever possible.

In addition, each trend station should be representative of easily identified, geographically-defined water bodies and should always be sited outside the mixing zones of point-source discharges.

Desirable Characteristics for Trend Station Siting:

A. Free-flowing, freshwater streams:

1. Whenever possible, stations should be sited in direct association with a flow gauge. Otherwise, stations should be near enough to one or more gauges to permit adequate interpolation of discharge at the site. When gauging is not available, a gauging device should be installed or an alternative means of flow measurement developed [23].

2. For watershed trend assessment, sites should be located near the mouth of the watershed to evaluate the loadings being discharged to the subsequent (downstream) watershed. The location of such stations may be either upstream from the outflow of one watershed, or downstream from the inflow of the subsequent watershed, but an effort should be made to minimize the number of significant tributaries that enter the gauged stream between the monitored site and the watershed boundary. On mainstem streams or rivers containing the waters from multiple upstream watersheds, sites should be located:

i. At or near the boundaries of USGS Sub-Basins (8-digit, 4th Order Hydrological Units),
ii. At or near the stream or river's fall line, when one exists, and
iii. Immediately above the freshwater head-of-tide, when one exists, with the same restrictions as those described in item 1 above.

B. Tidal Waters

For evaluating trends in tidal fresh- and saltwater tributaries, a trend station should be located near the geographical center of the tributary, and far enough from the mouth so that a minimum of open

[23] In certain cases the use of flow adjustment may be justifiably omitted from trend analyses. If trend analyses applied specifically to stream discharge data reveal no significant trend in rate of flow during the period of interest, it may be assumed that adjustment for flow would have minimal effects on the results of trends calculated for other water quality variables. When significant trends in flow are detected, however, the inclusion of flow adjustment may be extremely important. For the calculation of loadings (mass / unit time), however, discharge data are indispensable!

estuary or oceanic water is sampled at flood tide. Such samples should be representative of the tributary and not the estuary or ocean. In open estuarine areas, trend stations may be located at or immediately upstream from the stream's mouth with the open estuary or ocean, or in the mainstem of a bay/embayment, in order to evaluate estuarine water quality trends

C. All Waters

Trend stations should be located outside the mixing zone of permitted discharges and sufficiently downstream from significant tributaries to permit the complete mixing of the combined water columns. Whenever possible, sites should be located where adequate biological monitoring can be accomplished.

Additional, more specific guidance for siting trend stations and for the selection of datasets for trend analysis is provided in the document "Guidance for Siting and Identifying Trend Stations" [III-A-1c-2.doc].

(4) Parameters

Ideally, trend stations should be sampled for all parameters that are subject to water quality standards or are otherwise required to determine water quality conditions. Such parameter coverage may vary between freshwater (non-tidal) and estuarine (tidal) waters. Trend stations are currently sampled for all parameters listed in the appropriate spreadsheets of the Excel® workbook "Matrix of Parametric Coverage" [III-A-0b.xls] for TR Non-tidal or TR Tidal Trend Programs. The parameters measured at each location vary slightly, depending on whether the station is located in a free-flowing freshwater stream, reservoir, or estuary. The core parameters measured are listed below. The subset of parameters selected for trend analysis is more limited for several reasons. Some of the parameters in the table are components of others that were analyzed for trend or they are utilized simply to evaluate, qualify or calculate the values of other parameters. Some parameters have incomplete records and insufficient numbers of observations (too few data points) to calculate trends during the period of interest. For example, with the adoption of new Water Quality Standards for *Escherichia coli* and *Enterococci* bacteria, monitoring for these parameters only began in 2000. Even though Water Quality Standards for fecal coliform bacteria were phased out by 2008, fecal coliform monitoring results are still important because of lengthy records dating back to the late 1960's. Trend analyses for *E. coli* and *Enterococci* bacteria concentrations will complement those for fecal coliform bacteria, once long enough data records become available.

Of the various water-quality parameters routinely collected at these stations, the subset of nine that are generally used for trend analyses include:

(1) Bacteria (Fecal coliform and *Escherichia coli* – FCMFEC4, combination of 31616 and 31648), *
(2) Dissolved Oxygen- expressed as DO % Saturation,
(3) Total Nitrogen - TN, *
(4) Oxidized Nitrogen [24] (nitrate plus nitrite = NO_x),
(5) Total Kjeldahl Nitrogen (organic nitrogen plus ammonia = TKN),
(6) Total Phosphorus - TP, *
(7) Acidity or pH - PH,
(8) Water Temperature - TEMP, and
(9) Total Suspended Solids - TSS.*

[24] Oxidized nitrogen (NO_x) has in the past been used, in conjunction with other nitrogen species, to calculate total nitrogen! It is omitted at present because the T2 Parameter Group Code provides a direct determination of Total Nitrogen.

* Parameters followed by an asterisk in the list above are considered "Key Parameters" – see discussion in the following text.

Of the nine variables analyzed, four - Bacteria, Total Nitrogen, Total Phosphorus, and Total Suspended Solids - are considered key variables because of their significance in detecting water quality changes. Under typical ambient conditions, a declining trend (decreasing concentration) in any of these four variables is unequivocally considered a desirable trend. These four key parameters represent the most common causes of water quality deterioration or impairment. They are clear indicators of improving or deteriorating conditions based solely on changes in concentration, whereas with the other parameters the desired direction of change is subject to interpretation.

For example, decreases in dissolved oxygen may at first seem to indicate deteriorating conditions. However, if the decreasing DO is a result of decreasing concentrations of nutrients, this may be interpreted as a desirable trend. In the case of specific nitrogen species, NO_x and TKN, a trend may merely indicate a shift from one form of nitrogen to another rather than a change in total nitrogen concentration over time. In the case of pH, values at either extreme of the scale (exceedingly acidic or basic) are undesirable. Shifts in pH from either extreme (positive or negative trend) may be considered desirable depending on local conditions. Also, aquatic communities have adapted to prevalent climatic conditions over millennia. Significant temperature trends in either direction from historic values can have ecological significance on the distribution and abundance of aquatic organisms.

(5) Frequency

The number of samples required to produce a statistically reliable trend analysis is also an important consideration; it is a function of both the duration and the frequency of sampling. All DEQ water quality monitoring stations that are used for assessment purposes and the resultant 305(b) Report are currently sampled bimonthly (6 times per year), at a minimum. This provides an adequate sample size (number of observations) for short-term (2- to 6-year) assessment purposes and is generally adequate for mid- to long-term trend analyses as well. More frequent sampling may be performed if necessary, and certain parameters that demonstrate more stability in their values, or that are extremely expensive to collect and analyze, may be sampled with reduced frequencies.

At a minimum, trend stations are sampled bimonthly (6 events/yr) for most parameters, although specific substrates and/or parameter groups (i.e., toxic organics and/or metals in the water column and sediment) may only be sampled periodically (e.g., once each five or six years).

(6) Duration

Trend stations are considered to be permanent, fixed sampling sites of the Ambient Water Quality Monitoring Program.

(7) Quality Assurance Measures

Quality assurance measures associated with sample collection and analysis within this monitoring program follow agency guidance for the overall Ambient Water Quality Monitoring Program. (See Chapter IV – 'Quality Management Program' of this document.) Additional quality controls associated specifically with data screening and management prior to statistical analyses are discussed below and in the guidance documents for the specific statistical application(s) being used for analyses.

(8) Data Management

Data management procedures associated with sample collection and analysis within this monitoring program follow agency guidance for the overall Ambient Water Quality Monitoring Program. (See Chapter V – 'Data Management' of this document.) Additional data management elements, more specific to preparation of data for trend analyses, are described in more detail in the guidance documentation for the WQ3 trend analysis software discussed below.

(9) Data Analysis

As reported in DEQ's 2006 and 2012 Integrated 305(b)/303(d) Assessment Reports, the agency's trend analysis is accomplished using the Seasonal Kendall Trend Analysis with specific software developed for the agency. The decision to adopt this methodology, and the origins and development of the specific software package utilized are described below.

Background:

The Virginia '*Water Quality Monitoring, Information, and Restoration Act'* (WQMIRA), *§ 62.1-44.19:5. "Water quality monitoring and reporting"* was adopted into the Code of Virginia in 1997, and formally established the requirement for a strategy to determine water quality trends within the Commonwealth (see excerpts from WQMIRA in the discussion within section (11) "Reporting Requirements", below). Anticipating the requirements of this Act, DEQ had initiated a contract in 1996 with the Virginia Water Resources Research Center (Virginia Tech, Blacksburg, VA) to carry out an initial long-term water quality trend analysis on Virginia's waterways. Following extensive literature review, methodological selection, and data review and analysis during the ensuing two years, a final report on the study was submitted in 1998 (Zipper, et al., 1998 [III-B-3-1.pdf]; see bibliography for a complete citation, or go to http://vwrrc.vt.edu/special_reports.html and browse 1998 Reports). Various aspects of this initial trend analysis report were summarized in the agency's 2002 and 2004 305(b) Reports.

After reviewing the final report, DEQ consulted with and contracted the two primary authors, Carl Zipper, Ph.D., and Golde Holtzman, Ph.D. of Virginia Tech, to select an appropriate statistical method and to develop "user friendly" computer software that could be used by agency personnel for the detection and evaluation of water quality trends. Zipper and Holtzman further refined the methodology that had been used in the 1998 report and produced a "stand alone" computer program that was compatible with Microsoft Windows ® and could be run on a desktop PC. They named this software program "WQ2", considering it to be an improved, second generation of the methodology used in the original (1998) trend report. The method that they recommended for the analysis of long-term trends in water quality was the "Seasonal Kendall Tau" procedure, which is a non-parametric method for calculating correlations between the values of water quality parameters and the dates when they were measured. Subsequent steps calculate non-parametric regressions on the same data, with slope (rate of change) estimators and significance probabilities. The Seasonal Kendall analysis is a common technique for detecting water quality trends, and studies utilizing seasonal Kendall analysis are found in numerous peer-reviewed scientific publications. (For further information on the technique, readers may consult Hirsch et al. (1982), Hirsch et al. (1991), or Helsel and Hirsch (1992). Complete citations for these publications can be found in Part X – Bibliography of this document.) The Seasonal Kendall analysis allows the user to categorize data into specified seasonal time blocks prior to analysis; for example by month to yield 12 seasons annually or by quarter to yield four seasons annually. A general description of the method, how it functions, and how it is applied is summarized below.

Consequent to the application of WQ2 by DEQ personnel, a number of suggestions were proposed in order to improve the flexibility and efficiency of the software for agency use. Two of those improvements were (1) to provide more flexibility in selecting seasons for inclusion in the analyses and (2) the ability to automatically "batch process" multiple water quality variables at multiple sites within the same data file, rather than having to process them one by one as was required by WQ2. Zipper and Holtzman were again consulted and contracted to further refine the software. The results provided the third generation methodology in a software package named "WQ3." Although WQ3 was considerably more complex and more rapid, the statistical procedures applied in WQ3 remained fundamentally unchanged from previous versions. Its application does require the use of SAS® (Statistical Analysis Software - SAS Institute, Inc., Cary, North Carolina), however, rather than the simple Microsoft Windows ® environment of WQ2. WQ3 was utilized to generate trend analyses that were reported on in DEQ's 2006 and 2012 305(b) Reports, portions of which are summarized below.

The modified Seasonal Kendall method, as implemented in the WQ3 software currently in use by DEQ, is used to detect monotonic [25] trends in water quality over a period of time. The Kendall method utilizes *tau*, a rank-order correlation statistic, as a numeric trend indicator. The WQ3 software calculates the value of *tau* by determining the direction of change from each measured value of a water quality variable in the time series to the values of all subsequent measurements from the same season. The comparison begins with the first measurement (the oldest record) in a season, and is made sequentially with all subsequent values in time for the same season. Next, the second oldest value in time is compared in the same manner to all subsequent records in the same season. This comparison repeats until all pairs of values in the data set that occur within the same season have been compared sequentially. The seasonal trends thus characterized are subsequently integrated into a single collective trend. Because of the fact that water quality parameters tend to vary seasonally, and as a result of the frequency at which the data were collected, the monitoring data are generally analyzed in blocks of twelve (monthly) seasons. January values were only compared to January values, February values with February values, and so forth for each of the twelve months. Water quality nutrient concentrations, for example, are influenced by agricultural practices that vary seasonally. Similarly, stream flow may influence many water quality parameters, and stream flow varies with seasons. Consequently, the seasonal Kendall test compares each water-quality value only with values that occur within the same season of subsequent years. In the DEQ application of seasonal Kendall analysis, we have elected to define "seasons" as months of the year, because many data series being analyzed include monthly water quality measurements.

Because the comparisons of measured values are qualitative and only determine the direction of changes, without regard to the magnitudes of the differences in values, the statistical test is considered to be very robust. That is, its results are not unduly affected by extreme values, errors, or outliers. The analysis is also appropriate for use in data sets that have 'censored' upper and lower limits of analytical detection. For example, the variable for bacteria contains values that are recorded as less than 100 (<100), which means that the actual value lies between 0 and 99. This value (100) is defined as the lower limit of detection. Bacteria values may also be censored at the upper limit, being recorded as greater than 8000 (>8000), which means that the value could be anywhere from 8001 to infinity. All censored values at the same limit are considered to be tied and they can still be qualitatively compared with one another, with quantified values within the defined range of detection, or with censored values at the other end of the scale.

Kendall's *tau*, which may vary in value from -1.000 to +1.000, is a measure of the direction and relative uniformity (monotonicity) of the trend. The more consistent the trend is, the stronger the *tau* value will be (i.e., the further it will be from 0.000). Negative values of *tau* indicate declining trends and positive values indicate increasing trends. The *tau* value, however, does not indicate the magnitude of the trend. For instance, the dissolved oxygen (DO) values of 9.9, 9.8, 9.7, 9.6, and 9.5 mg/L would have a *tau* value of –1.000, since each value is less than all previous values; however, so would a trend of 9, 7, 6, 4, 1, which is a much more severe trend. The *tau* value indicates the consistency of directional change, not the magnitude or rate of change. Another characteristic of the method is that changes in trend direction over time (e.g.,

[25] A monotonic trend is a trend that is predominantly unidirectional. It may be linear, smoothly curved, or irregular in form. The Kendall *tau* method may fail to identify trends that reverse their predominant direction during the period of interest.

increasing trend for 10 years, followed by a decreasing trend for 10 years) may cancel out such that a trend over the full time period cannot be detected.

In addition to the *tau* statistic, WQ3 calculates the significance of the trend as a 'P-value.' The P-value is the probability of observing an equal or more extreme value of *tau* than that calculated from the data, if no real trend were present. The smaller the P-value, the more confidence we have in rejecting the 'null hypothesis' (H_0) that no real trend is present (H_0: *tau* = 0.000). A P-value = 0.05 (= 5%) would indicate that we would have a confidence of at least 95% in rejecting H_0 (Confidence = 100% - P %), and is generally interpreted to indicate statistical significance. In other words, we would have a 95% confidence that a real trend is present. WQ3 calculates two distinct P-values, an independent P-value (PVALUE) and a dependent P-value (PVALCOVS). The independent p-value is calculated under the assumption that all measured values of a specific water quality variable are completely independent of one another, while the dependent P-value includes the assumption that the water quality observations for individual months or blocks are similar in value to those of the preceding and following months. For example, the water quality in January is assumed to be more similar to the water quality in December and in February than to those in July and August. When blocking seasonal data into twelve months, the dependent p value is often a better estimate of the significance of the trend [26].

When WQ3 detects a significant trend, it also estimates a linear regression equation that can be used to characterize the observed water quality changes. This output includes estimators for a slope and a y-intercept (value of the variable when time = 0.0). The slope estimate is defined as the 'median' rate of change in value of the water quality parameter per year. An advantage of the linear regression output is that the slope and intercept estimators can then be used to estimate future water quality values for those variables with statistically significant trends. Several considerations are of importance when applying the linear regression projections: (1) the assumption that the trend will remain constant may not be valid - there is always a risk of error when regressions are projected beyond the limits of the observations that went into their estimation, and water quality management changes may cause future patterns of water quality change to vary from those of the past, (2) the assumption that the trend is linear may not be valid, and (3) the projected line represents the median point of the prediction and there is no confidence interval associated with the estimate; consequently, the point of intersection with a specified criterion or standard would represent a 50% violation rate.

It should be noted that with the Kendall Tau method a situation can occur where the Tau value indicates that a significant trend is present, but the estimated slope is still zero. This occurs most commonly with data sets where the majority of values are censored and "tied" at the lower detection limit, but occasional higher values are also observed. If there is a significant change in the frequency of uncensored values during the period of interest then a trend can be detected but, since the majority of the values are tied, the estimated (median) slope would still be zero (0.000).

With data sets as large and complex as those utilized by DEQ, a significant amount of data screening and preprocessing is required prior to the application of any trend analysis method. Data analysis involves four major steps encompassing raw data retrieval, preprocessing, analysis, and reporting. Raw data are stored in the Agency's Comprehensive Environmental Data System - Water Quality Monitoring application, CEDS-WQM. Data in CEDS-WQM include all historic water quality measurements once stored at the National Computing Center in the STORET database (under Agency Code 21VASWCB) plus all data collected by the Agency since November 1998 when STORET stopped accepting data. CEDS-WQM data span six

[26] Darken P., C. Zipper, G. Holtzman and E. Smith. 2002. Serial correlation in water quality variables: Estimation and implications for trend analysis. Water Res. Research 38:22 1-5.

decades, with the first recorded sample collection in June of 1941. For the purpose of trend analysis a subset of these data are queried, based on date of sample collection, type of station (LEVEL3 Code = ATRND - Trend Station), and the desired parameters. The queried data are then passed into SAS®, a statistical application used for subsequent preprocessing and analysis.

In addition, daily median flow values in cubic feet per second (CFS) are obtaining from the U.S. Geologic Survey as raw text files. These consist of a combination of data collected by the USGS and by DEQ's Surface Water Investigations Unit, collocated with the Virginia Department of Forestry, University of Virginia, in Charlottesville. The Virginia Department of Environmental Quality has a unique arrangement with the USGS in that the Department's flow data are considered of equal quality to the Federal data and are therefore published annually side by side with the USGS data.

Preprocessing consists of a sequence of steps that prepares the raw water quality and flow data for formal statistical analysis. The steps required to produce a data set for statistical analysis are as follows:

1) Convert sample collection date-times format to month-day-year format.
2) Join water quality data with flow data by station by day.
3) Create the dissolved oxygen (DO mg/L) variable by selecting Winkler DO or DO probe, whichever is present; Winkler takes precedence if both were measured at the same time.
4) Calculate an estimated chlorinity variable from salinity or specific conductance. DO is measured almost universally at all stations and sampling events, whereas in freshwaters specific conductance was not always measured. Where DO is measured but chlorinity is not available, the mean chlorinity value for that station is used as a substitute. The chlorinity value subsequently is used in the calculation of a dissolved oxygen percent saturation (DOSAT) variable for trend analysis. In those cases where mean chlorinity values are substituted, the specific conductance is generally very low and has very little influence on the calculated DOSAT.
5) Create a dissolved oxygen percent saturation variable (DOSAT) from the combined DO value (concentration in mg/L), temperature, and chlorinity.
6) Create a bacterial (fecal coliform) variable (BACT) by selecting MPN (31615) or MF (31616), whichever is present, with MPN taking precedence if both were measured at the same time.
7) For those cases where nitrate plus nitrite (00620) was not measured directly, calculate an oxidized nitrogen variable (NO_x) from the addition of nitrite and nitrate if both were measured.
8) Estimate missing values for total Kjeldahl nitrogen (TKN), oxidized nitrogen (NO_x) and total nitrogen (TN) under the assumption $TKN + NO_x = TN$. When any two of these variables are present the third is calculated.
9) Create corresponding remark code variables for each calculated variable. For example, in the case where TKN and NO_x were summed to produce a TN value, if both TKN and NO_x were remark coded with less than detection limit, then the corresponding remark code for TN would indicate that this measurement was also below detection limit. Field measurements of pH, temperature, DO, salinity and/or specific conductance do not have corresponding remark codes. For statistical analysis, however, a remark code field must be present with each variable. When missing, remark code variables are created for each field measurement and subsequently populated with null values.
10) Remark codes for flow from the USGS are used only to indicate an estimated (e) value when gauges malfunction, are very rare, and hardly ever co-occur with remark coded water quality data. Consequently, the estimated flow values are not used in the statistical analysis. By including a field for flow remark codes, however, trend analysis can be performed on flow itself, as a separate water quality parameter.
11) The most commonly observed remark codes are "U" - indicating the measurement was less than the detection limit, and "L" - indicating the measurement was too high to accuracy quantify. L is only

present with the bacteria variables. Because detection limits may vary among parameters, the "U" remark codes were recoded to a unique letter for each parameter. For example, all TKN "U" remark codes are converted to "K", all NO_x "U" remark codes are converted to a different letter, "X", etc. This differential remark coding permits the use of uniform but different detection limits for each parameter, for use in the next step of preprocessing - univariate analysis.

12) Univariate analysis of each variable is performed to determine suitable uniform detection limits. Over a long period of record, the same parameter may have been measured with several detection limits. This occurs because improvements in methods or instrumentation lead to improvements in detection. Regardless of these improvements, it is not possible to distinguish between a value of 0.1 and 0.01 if both are reported as "less than detection limit." In most cases the higher value (0.1 in this case) would become the uniform detection limit for that variable for the period of record. If, however, the inspection of the distribution of values in the univariate analysis reveals that the higher value (0.1) represents less than one percent of the total remark coded values, and the lower value (0.01) is much more numerous, it may be appropriate to select the uniform detection limit value that corresponds to the more frequently observed value. A manual inspection of the distribution of detection limit values for each variable is critical.
13) Once the uniform detection limit values have been selected, a new data set is produced with derived variable concentrations adjusted by the epsilon method. The epsilon method sets each remark coded variable to a new uniform value based on the uniform method detection limit.

Only after all the preprocessing is completed is the derived data set ready for statistical analysis. Among the 409 active trend stations being monitored in 2007, a number of sites had insufficient data records in 2006 for trend analyses to be performed with the WQ3 software. In 2011, for the 2012 Integrated Report, there were 436 active trend stations that contained sufficient data during the 20-year period from 1991 to 2010. The number of sites analyzed also varies among the parameters considered. Exclusion of a site from the WQ3 analyses generally results from an insufficient data record. This may be caused by its being a newly established site or by interruptions in a longer data record.

The final step in data analysis is reporting the results of all statistically significant trends (positive + and negative -) as compared to those that were not significant (no change). Although trend analyses are carried out to characterize individual sites, trend results may be integrated into basin specific or statewide summaries. Aggregated results may be summarized as percentage of total sites evaluated that have positive (increasing) trends, negative (decreasing) trends, and insignificant (no change) trends for the four key water quality parameters. An example of such a summary is presented in **Table III.B.3-1** - Relative Frequencies, by Basin and Statewide, of Trends in the Four 'Key' Water Quality Parameters, below, from the Trend Analysis Chapter 2.4 of the 2006 Integrated Report [III-B-3-1a.pdf].

Figure III.B.3-2 (below) illustrates the geographic distribution of the stream, lake, and estuarine trend monitoring stations and gauging stations utilized for the Trend Analysis Chapter 4.5 of the 2012 305(b)/303(d) Integrated Water Quality Report (IR) [III-B-3-1b.pdf].

In Chapter 4.6 of the 2012 Integrated Report [III-B-3-1c.pdf], DEQ introduced a new method of trend detection. The Integrated Water Quality (IWQ) trend analysis is a new computational procedure developed to detect regional long term trends by maximizing the amount of data used. The IWQ is a seasonally-derived nonparametric scoring procedure that was applied to various waterbody types at the watershed scale. The impetus for the creation of the IWQ was the desire to detect and explain incremental changes in water quality over time more descriptively than the traditional impaired *vs.* non-impaired dichotomy of the 305(b)/ 303(d) Integrated Reports. More stream segments and estuarine / reservoirs polygons are added to the 303(d) list of impaired waters with every two-year cycle. However, as resources have shifted towards restoration efforts in watersheds, attention has been given to the development of a new statistical approach that would capture changes in water quality at broader spatial scales. The IWQ was developed by D.H. Smith and R.E. Stewart at the Virginia Department of Environmental Quality with the encouragement and support of L. Merrill with the U.S. Environmental Protection Agency, Region 3.

Table III.B.3-1 - Relative Frequencies, by Basin and Statewide, of Trends in the Four 'Key' Water Quality Parameters (2006 Seasonal Kendall Trend Analysis for the Period 1985 – 2004)

	BACTERIA				NITROGEN				PHOSPHORUS				TOTAL SUSPENDED SOLIDS			
BASIN	+	-	NO CHANGE	COUNT	+	-	NO CHANGE	COUNT	+	-	NO CHANGE	COUNT	+	-	NO CHANGE	COUNT
POTOMAC SHENANDOAH	0%	0%	100%	34	3%	25%	72%	36	6%	9%	86%	35	3%	8%	89%	36
JAMES	0%	18%	83%	57	7%	22%	72%	46	0%	37%	63%	59	2%	8%	90%	61
RAPPAHANNOCK	0%	0%	100%	18	0%	0%	100%	11	0%	0%	100%	19	5%	10%	85%	20
ROANOKE	0%	6%	94%	16	6%	13%	81%	16	0%	20%	80%	15	0%	6%	94%	16
CHOWAN	7%	0%	93%	14	0%	25%	75%	20	0%	20%	80%	20	0%	35%	65%	20
TENNESSEE BIG SANDY	0%	33%	67%	3	0%	0%	100%	3	0%	0%	100%	2	0%	0%	100%	3
CHESAPEAKE BAY, OCEAN, SMALL COASTAL	5%	0%	95%	21	5%	5%	90%	19	5%	21%	74%	19	0%	16%	84%	19
YORK	0%	0%	100%	14	57%	0%	43%	7	0%	0%	100%	15	13%	6%	81%	16
NEW	0%	0%	100%	6	0%	0%	100%	5	0%	0%	100%	5	0%	0%	100%	5
TOTALS	1%	7%	92%	183	6%	17%	77%	163	2%	19%	79%	189	3%	11%	86%	196

Statistically significant improving trends in water quality were revealed across the Commonwealth with the modified seasonal Kendall trend analysis (see Chapter 4.5 – 2012 Integrated Report [III-B-3-1b.pdf; Internet links within that document are no longer functional]. That analysis was performed on 436 stations, specifically designated for trend analysis, over a twenty-year period. In contrast, the IWQ approach incorporates 5,776 stations of diverse types during the same time frame, allowing for a broader, watershed wide characterization. **Figure III.B.3-3**, below, illustrates the geographic distribution of stream, lake/reservoir, and estuarine monitoring stations that contributed data for the IWQ analyses.

IWQ analyses were applied to the same four key parameters (bacteria, total nitrogen, total phosphorus, and total suspended solids) addressed by the Seasonal Kendall Trend analyses. Whereas the results from the Seasonal Kendal analyses characterized individual monitoring sites (see **Figure III-B-3-4**, below), the integration of data using the IWQ permitted watershed-wide characterizations (see **Figure III-B-3-5**).

Figure III.B.3-2 - Water Quality Trend and Gage Stations Utilized in the 2012 Integrated Report
(Seasonal Kendall Trend Analysis for the Period 1991 – 2010.)

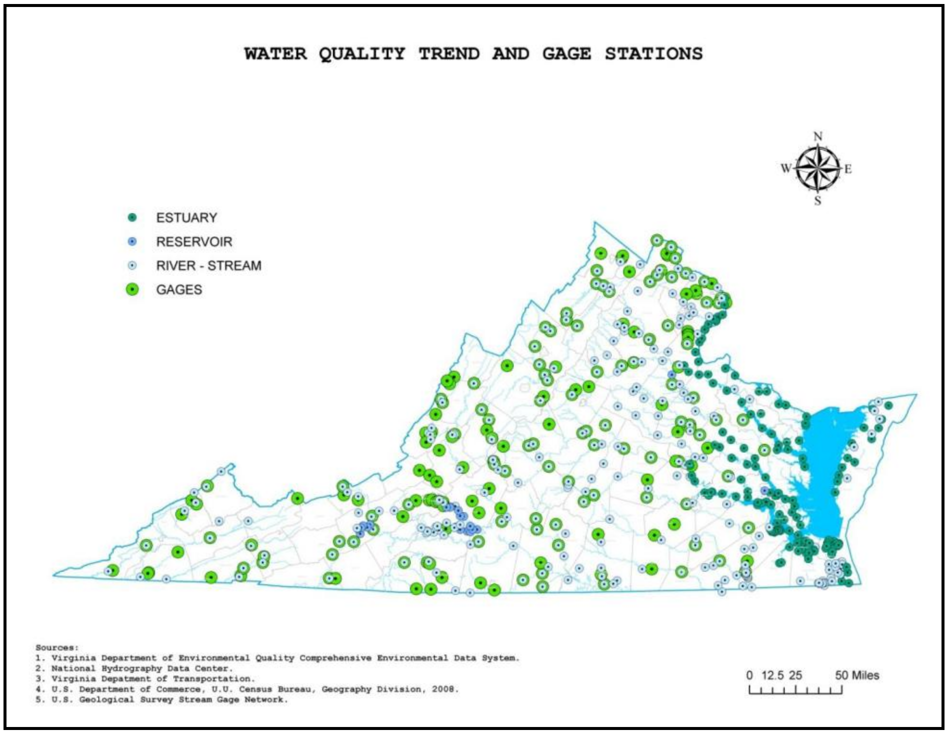
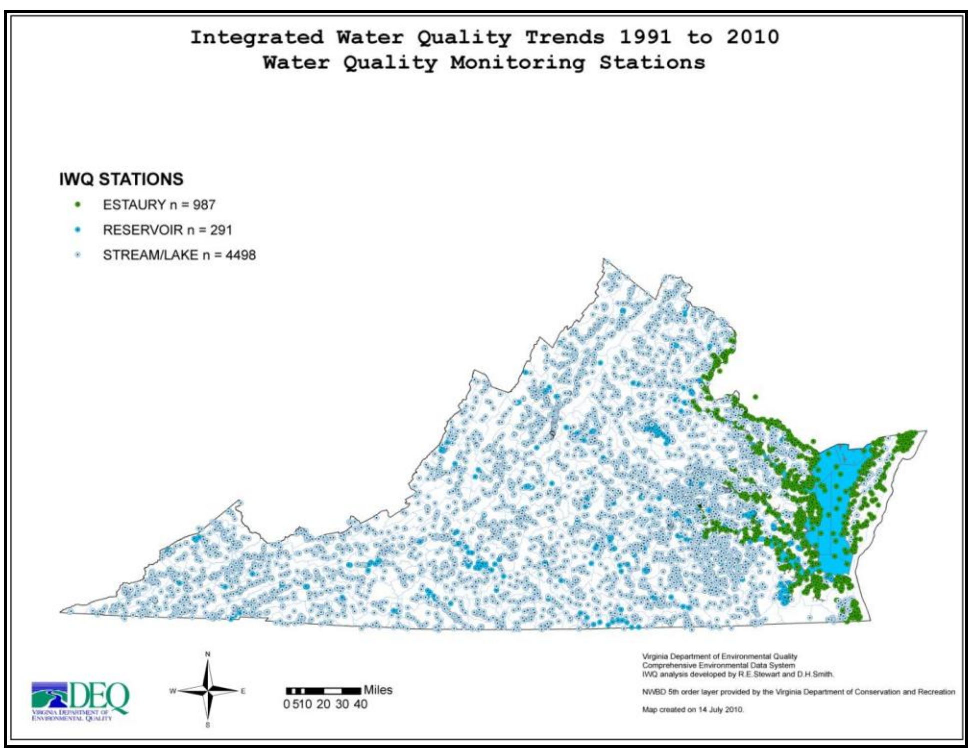


Figure III-B-3-3 – Virginia DEQ Monitoring Station Used in the IWQ Analyses.



The methodology used in calculating the IWQ is described in detail in Chapter 4.6 – Integrated Water Quality Index Results [III-B-3-1c.pdf] of DEQ's 2012 IR. In brief, a reference period is selected and characterized by calculating the upper and lower quartiles of all measurements of the parameter of interest during the reference period. In practice, the most recent ten-year period is generally selected as the reference period, because the most recent data usually undergo the most stringent quality control and are collected and analyzed using more sensitive methods than those used in the past. Calculation of the quartiles may be stratified by season, aquatic resource type (*e.g.*, free-flowing stream, lake, estuary), and by watershed. The preferred quartile - generally the first (lower concentration = higher water quality) - is identified and is assigned an arbitrary score. By convention DEQ has assigned a high score to the quartile (lower) with the highest water quality and a low score to the opposing (upper) quartile. An intermediate score is assigned to the inter-quartile range. All values of the parameter measured during the period of interest are then assigned a score, and annual mean scores are calculated. **Figure III-B-3-6** illustrates the results of an IWQ characterization of the lower (tidal) James River sub-Basin (Hydrologic Unit 02080206) from 1970 through 2008. Annual bars show the percentages of observations falling in the lower quartile (higher water quality – green), within upper quartile (lower water quality - red), and within the inter-quartile range (intermediate water quality - yellow). The curved line traces the variation in mean annual IWQ scores and the straight line is the linear regression of annual scores against years of observation (regression equation and coefficient of determination in the lower right corner).

Figure III-B-3-4 – Annual Percent Change in Flow-Corrected Bacterial Counts at Individual Trend Stations, as Revealed by the Seasonal Kendall Trend Analyses (1991 – 2010)

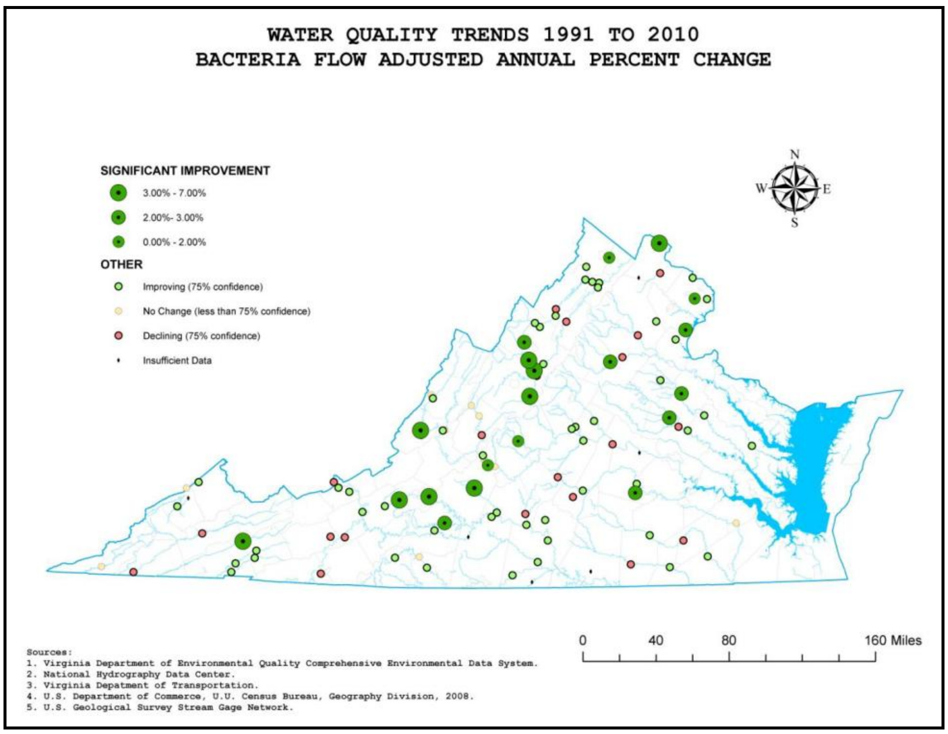
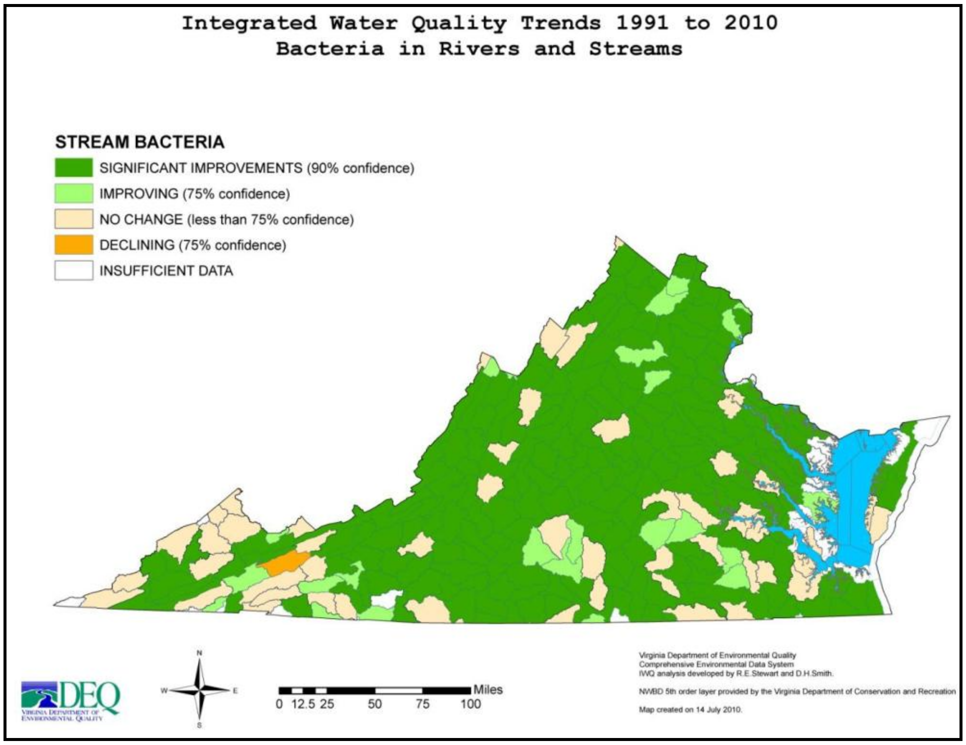


Figure III-B-3-5 –Watershed-wide Trends for Bacteria in Streams and Rivers, as Revealed by Integrated Water Quality (IWQ) Analyses (1991 – 2010)



Histograms, such as this (Figure III-B-3-6) illustrate the changes in water quality through time for the selected water body, while maps such as in Figure III.B.3-3 demonstrate the geographic distribution of water quality trends on a broader (regional or statewide) scale.

In addition to these types of summaries, a detailed discussion of the methodology and results has been included in the corresponding Chapters of the Assessment Section of the agency's 2006 and 2012 biennial 305(b)/303(d) Integrated Reports on water quality conditions to the U.S. EPA and Congress.

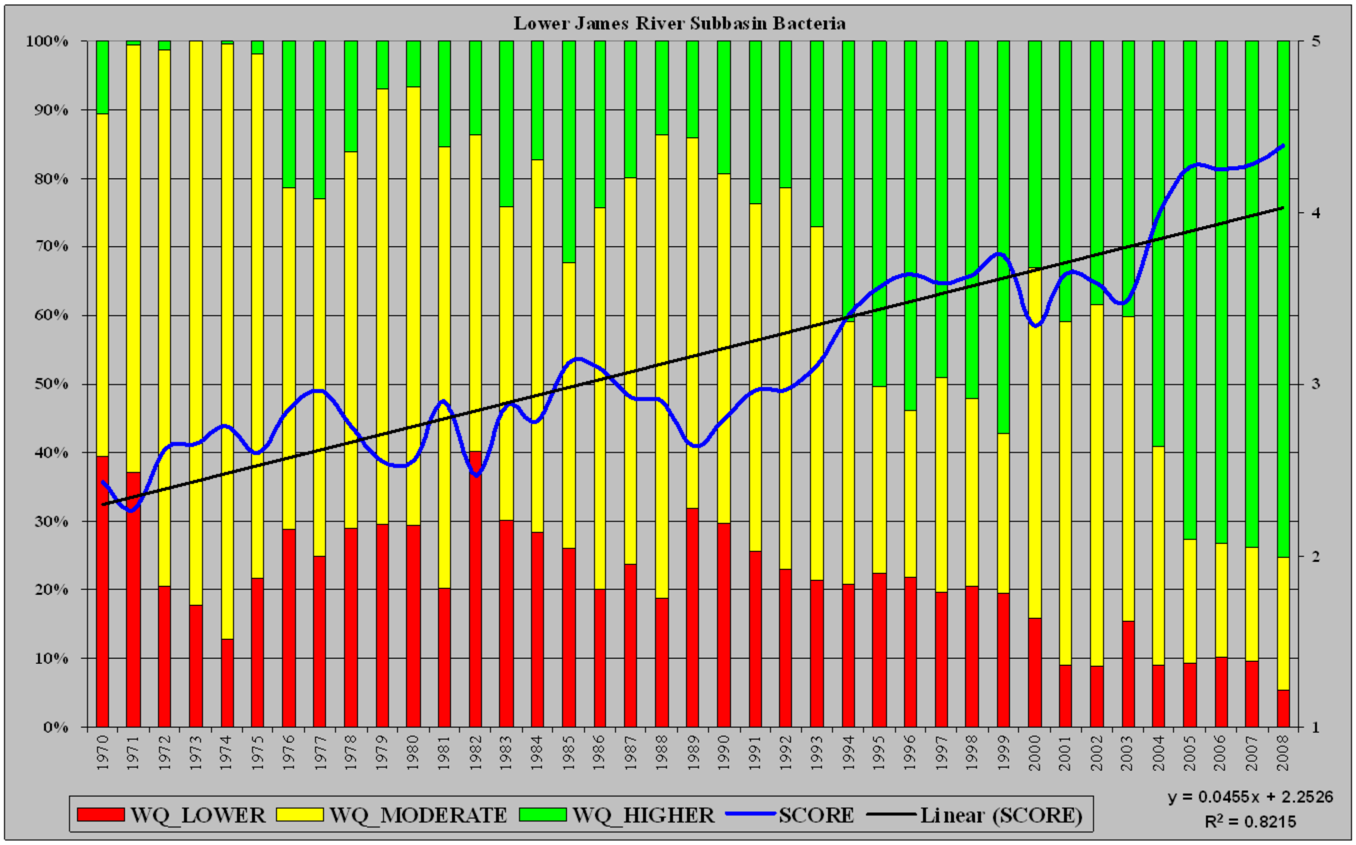
(10) Data Gaps

Future trend analysis studies may include exploratory data analyses to investigate:

1) More recent, shorter-term changes in water quality (*i.e*., non-linearity of trends),
2) Significance of trends by geographic location to include Ecoregion summaries,
3) Selection of appropriate water quality indicators,
4) Influence of frequency of sampling,
5) Influence of anthropogenic changes in the watershed on trends,
6) More in depth analysis of seasonal temperature trends, and
7) Experimentation with data blocking based on hydrologic cycles.
8) More recent, shorter-term changes in water quality (*i.e*., non-linearity of trends),
9) Significance of trends by geographic location to include Ecoregion summaries,
10) Selection of appropriate water quality indicators,
11) Influence of frequency of sampling,
12) Influence of anthropogenic changes in the watershed on trends,

13) More in depth analysis of seasonal temperature trends, and
14) Experimentation with data blocking based on hydrologic cycles.

Figure III-B-3-6 - Histogram of Annual IWQ Scores for Bacterial (fecal coliform) Counts in the Tidal Lower James River Basin (8-digit sub-basin 02080206) from 1970 to 2008. Annual bars show the percentages of observations falling in the lower quartile (higher water quality - green), upper quartile (lower water quality - red), and within the inter-quartile range (intermediate water quality - yellow). See the text above for more details.



Equitable geographic and water body type representations must be assured in order to provide meaningful statewide trend characterizations. At present, not all of the state's USGS Hydrological Units (8-digit Sub-Basins) contain representative numbers of trend stations, so redistribution will be necessary. The current distribution has resulted because many of the current trend stations are holdovers from earlier monitoring strategies that were not consistent with the new approach. Trend analyses among various water body types is important in determining if a specific class or type of water resource is experiencing significant changes in water quality. For example, we must have good representation of trend stations in our free-running streams and rivers, lakes and reservoirs, estuaries and, eventually, in wetlands.

The results of future exploratory analyses will determine the significance and magnitude of any data gaps related to the considerations above. The redistribution of stations for equal representation will occur after the first reports of trends using the new WQ3 statistical application are completed. Indeed, it is expected that the interpretation of these large-scale trend analyses will reveal information to help identify the proper course of future trend monitoring activities.

(11) Reporting Requirements

The Virginia '*Water Quality Monitoring, Information, and Restoration Act*' [I-0d.pdf] (WQMIRA) established the basis for a strategy to determine water quality trends.

§ 62.1-44.19:5. Water quality monitoring and reporting.

> *B. Monitoring shall be conducted so that it:*
>
> *1. Establishes consistent siting and monitoring techniques to ensure data reliability, comparability of data collected throughout the state, and ability to determine water quality trends within specific and easily identifiable geographically defined water segments.*

The nature of trend analysis is such that the detection of changes in trends is not expected on such a short time scale as the interval between biennial 305(b) Reports. For the future, we anticipate producing a periodic trend report in association with each six-year rotation cycle of the monitoring program and the coincident revisions of the agency's Water Quality Monitoring Strategy. Water quality trend evaluations will be summarized in the Department's corresponding 305(b) Reports to Congress and the U.S. EPA. In addition, DEQ intends to periodically produce incidental reports for data customers. Although the Trend Station Monitoring Network is well defined, trend analysis in itself is not constrained to data sets produced solely by this network. Specialized trend analysis for other suitable agency data sets will occur as needed.

(12) Periodic Review

This program, like our other monitoring programs, requires periodic evaluation to determine its effectiveness: 'Is the program meeting the needs of the data users?' 'Are the program objectives consistent with current and future needs?' and 'Are the technologies being used the best choice?' Because the WQ3 trend analysis is a relatively new addition to the Department's WQM Strategy, much time and effort will be spent reviewing and comparing the results of the 2004, 2006, and 2012 analyses and determining if subsequent changes to the program are necessary. We fully expect that station realignment and the possible modification of sampling frequencies will be the first major program changes for the next trend cycle.

(13) General Support and Infrastructure

Station locations, parameters sampled, frequency of sampling, and analytical costs for the Trend Monitoring Program can be determined from the current annual water quality Monitoring Plan (MonPlan), which can be viewed on the DEQ Water Quality Monitoring WebPages. Because equipment, personnel, and logistical support are shared across programs, precise cost estimates for these resources are not possible.

A significant amount of staff time is devoted to processing, interpreting, and reporting the data. Much of this is carried out by the Central Office Water Quality Monitoring and Assessment staff. It is estimated that several hundred man-hours are spent to retrieve suitable data sets, to preprocess the raw data, and to analyze the data via WQ3 for each statewide trend analysis report. Another hundred or more hours are devoted to the assimilation, summarization, and reporting of the WQ3 output into a format suitable to be included in the Integrated Report. Because the agency is still limited in its experience with trend analysis at such a scale, unforeseen resource requirements may arise.

The cost of maintaining the SAS statistical software application is of significance. Prior to the end of STORET at the National Computing Center, SAS was available to all data owners via a telnet-to-mainframe connection. The loss of STORET shifted the cost burden to individual users, which has been an expensive endeavor for the agency. Continued support of the SAS license will be necessary to maintain support and receive service packs containing fixes to known code problems.

(14) Plan and Schedule

A beta version of the WQ3 statistical application was delivered to DEQ in February 2004. Training by Virginia Tech staff was carried out in the spring of 2004, and by the end of 2004 agency staff had produced the first preliminary trend analysis outputs using WQ3. Following familiarization with (and a few final adjustments to) the software, more extensive data queries, data pre-processing, and statewide trend analyses were performed in 2005. Chapters summarizing the results of these trend analyses were first included in the agency's 2006 and subsequently in its 2012 305(b)/306(d) Integrated Water Quality Reports and are directly linked to this strategy document as 2006 Trend Chapter 2.4 [III-B-3-1a.pdf] and 2012 Trend Chapter 4.5 [III-B-3-1b.pdf].

Since the 2005-2006 trend analyses, DEQ has begun to address the data gaps and other issues identified above. This will require several more years of exploratory analyses and may require additional technical support from Virginia Tech, which may also include additional expenses for the agency.

DEQ has proceeded with improvements to our in-house database, CEDS-WQM, such that the DEQ and the USGS flow/discharge data can be linked to the agency's water quality data. This provides a much simpler format for flow correcting trend data (or calculating trends based on loads) and will also lead to more advanced exploration of the relationship of flow and the rate of change of flow to water quality.

With these types of advanced exploratory analyses it is difficult to plan a fixed sequence of steps for the Trend Monitoring Program, as the path is not illuminated until the exploration proceeds. We have a starting point - what is now most important is that the people working on data analyses be inquisitive, which will lead to new projects and a better understanding of the nature of water quality changes over time.

At the present time, DEQ plans to continue performing complete, repeated trend analyses on a six-year cycle, based on a sliding 20-year data window. The most recent six years would thus contribute 30% new information to the end of each 20-year trend analysis. If complete trend analyses were to be repeated on a more frequent basis (*e.g.*, for each 305(b) Report), any changes during the most recently added (two-year) time block would generally be completely masked by the longer time block that would have been included in the previous analysis. Current design considerations include one long-term analysis of the entire 20-year window, and separate analyses of the initial and final 10-year blocks for purpose of comparison and the detection of recent changes in trend tendencies. The 20-year data windows for recent and future trend analyses are included in the spreadsheet of Planned Cyclic Water Quality Monitoring Activities [VII-1d.xls] linked to this document.

Contact: For further information on the Trend Monitoring Program and Trend Analyses contact:

Roger E. Stewart II
629 East Main Street
Richmond, Virginia 23219
(804) 698-4449
Roger.Stewart@deq.virginia.gov

Donald H. Smith, PhD
629 East Main Street
Richmond, Virginia 23219
(804) 698-4429
Donald.Smith@deq.virginia.gov