METHODOLOGY TO DERIVE WATER-QUALITY TRENDS FOR USE BY THE NATIONAL WATER SUMMARY PROGRAM OF THE U.S. GEOLOGICAL SURVEY

By Kenneth J. Lanfear and Richard B. Alexander

U.S. GEOLOGICAL SURVEY

Open-File Report 90-359

Reston, Virginia
1990

DEPARTMENT OF THE INTERIOR

MANUEL LUJAN, JR., Secretary

U.S. GEOLOGICAL SURVEY

Dallas L. Peck, Director

For further information write to:

Assistant Chief Hydrologist, Water
Assessment and Data Coordination
U.S. Geological Survey
407 National Center
Reston, Virginia 22092

Copies of this report can be
purchased from:

U.S. Geological Survey
Books and Open-File Reports
Box 25425, Federal Center, Bldg. 810
Denver, Colorado 80225

CONTENTS

FIGURES

Figure

TABLES

Table

METHODOLOGY TO DERIVE WATER-QUALITY TRENDS FOR USE BY THE NATIONAL WATER SUMMARY PROGRAM OF THE U. S. GEOLOGICAL SURVEY

By
Kenneth J. Lanfear and Richard B. Alexander

ABSTRACT

Methods were developed to examine a water-quality record and, based on the application of a detailed set of trend testing criteria, determine whether a monthly, bimonthly, or quarterly test for monotonic trend could be performed. The trend testing criteria were automated and allowed for an efficient and accurate analysis for trend in an extremely large number of water-quality records. Water-quality records for about 3,000 stations nationwide were retrieved from the U.S. Geological Survey's National Water Information System, and were evaluated as to their suitability for trend testing. From this set of records, it was determined that about 1,100 stations had at least one water-quality constituent that met the criteria for trend testing. For those water-quality records, a nonparametric test for monotonic trend, known as the seasonal Kendall test, was applied to three time periods (1970-89, 1975-89, and 1980-89). The results of these water-quality trends are intended for use in the U.S. Geological Survey's National Water Summary Program activities on stream water quality.

INTRODUCTION

The subject of the U.S. Geological Survey's (USGS) 1990-91 National Water Summary (NWS) is water-quality conditions and trends in streams of the United States, Puerto Rico, the U.S. Virgin Islands, and the Western Pacific Islands. The NWS report, planned for release during the 20th anniversary of the 1972 Clean Water Act, will discuss the results of analyses of water-quality data derived from water samples collected at about 3,000 stations during the period 1970-89. The results of these analyses, which are based on data retrieved from the USGS National Water Information System (NWIS) and on comparable data retrieved from data bases maintained by other Federal agencies and by State and local governments, will be reported on a State-by-State and national basis.

More than 50,000 trend tests were performed on a data base that consisted of water-quality constituents collected at nearly 3,000 stations; more than 200 different constituents were determined for water samples collected at one or more of these stations. Clearly, this effort required highly automated procedures for examining the data. This report documents the methodology used to extract the data from NWIS and perform the trend tests. Documentation of the software used is beyond the scope of this report.

DEVELOPMENT OF NATIONAL WATER SUMMARY DATA BASE

Station Selection

Before the NWS report on stream water quality could be undertaken, the question "Are there enough data?" had to be addressed. A sufficient number of water-quality sampling stations, well-distributed among the states, with sampling records of appropriate frequency and duration had to be found. Smith and others (1987) had examined trends in water-quality data for more than 300 stations in the USGS' National Stream Quality Accounting Network (NASQAN). However, for a State-by-State discussion of stream water quality in the NWS report, it was necessary to supplement those 300 stations.

The USGS's National Water Data Exchange (NAWDEX) data base is an index to water sampling stations in the United States. Parks (1988) converted NAWDEX into an ARC/INFO[1] format, which permitted more complicated queries than the earlier NAWDEX data base. This version of NAWDEX was searched for water-quality stations having a minimum of 5 years of contiguous annual values of streamflow and water-quality data. Of 2,927 stations that were found, 2,828 were USGS stations, including 480 NASQAN stations.

Data Retrieval

A retrieval was made of all water-quality records for 2,946 USGS stations, which consisted of the 2,828 USGS stations identified from the NAWDEX retrieval plus 118 NASQAN stations not identified in the original selection. The location of the stations in the conterminous United States is shown in figure 1. Water-quality records for these sites were retrieved from the National Water Information System (NWIS). These records reside in about 50 computers located in USGS offices nationwide.

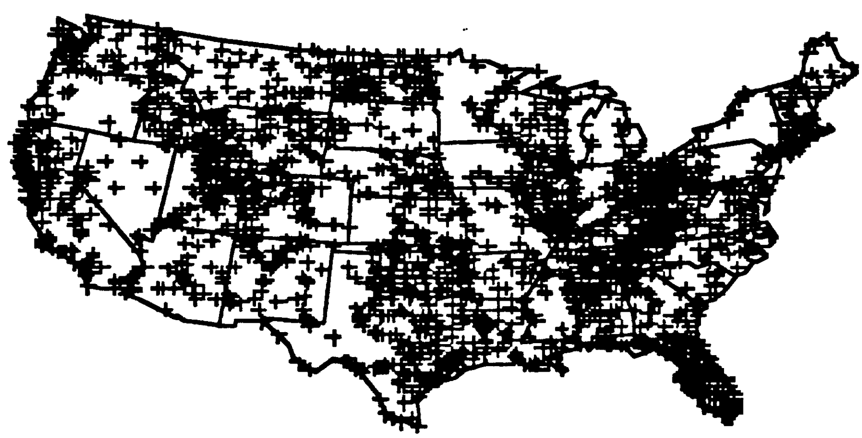

Figure 1. -- Location of surface-water stations in the conterminous United States for which water-quality records were selected for retrieval from the National Water Information System. Water-quality records were also retrieved for stations in Alaska, Hawaii, and Puerto Rico.

Station records were retrieved using existing NWIS software that resides on each USGS computer, and electronically transferred to a computer at the USGS Headquarters in Reston, Va. There, the data were reformatted and stored in an ARC/INFO data base. The information retrieved consisted of the station identification number, date, time, location, parameter code, data value, and remark code for all water-quality samples collected at the selected stations. All retrieval procedures automatically logged their operations in a computer file, so a complete processing history was recorded for every data value.

---

1 Use of trade names in this report is for identification only and does not constitute endorsement by the U.S. Geological Survey.

Additional Stations

Besides the 99 non-USGS stations identified in the original NAWDEX search and those NAWDEX stations that were not retrieved, there probably are many other stations that have data records that are suitable for trend testing. Authors of the individual NWS State sections were allowed to nominate other USGS or other agency stations for inclusion in the data base. Once nominated, such stations were processed similarly to the originally retrieved stations.

To avoid including station records that were not representative of ambient water-quality conditions in the channel cross-section or samples that were collected with techniques that could bias a trend test, all stations would be required to have water-quality records with the characteristics listed in table 1. Although nearly all USGS station records met these criteria, authors of the individual State sections were required to review the stations previously selected to ensure the stations did meet the criteria.

Table 1. -- Criteria for including stations in the National Water Summary data base

- Data must be transferrable electronically.
- Sampling site must be located by latitude and longitude, and have an assigned identification code and descriptive name. The agency collecting the data must be identified.
- Data must represent ambient surface water-quality conditions as opposed to effluent samples. That is, the sample must have been obtained from a stream and not from a discharge pipe.
- Water samples must have been collected using a method that is equivalent to standard USGS sampling methods, that is, depth-integrated samples composited to represent the cross-section of the stream. Water samples *for dissolved constituents only* may have been collected at a single point in the cross section of the stream if (1) it is known that the stream was well-mixed at the sampling point, (2) the method of sample collection is documented, and (3) any changes in the sampling methodology over time would not affect the comparability of the data.
- Water samples must have been analyzed using standardized methodology approved by USGS, or the U.S. Environmental Protection Agency. The methods used must provide data that are comparable over time.

Additional criteria were given to authors regarding the type of station data record (that is, length of record and frequency of sampling) needed to qualify for trend testing, to avoid having to process obviously unsuitable records. The trend testing methods described below, however, place no advance restrictions on the station data record and automatically check the record for suitability. Thus, authors were encouraged to nominate stations that were "borderline" cases for trend testing so that the records could be objectively evaluated as to their suitability for trend testing.

Data were successfully retrieved for 2,856 of the 3,064 stations for which retrievals were initiated. Of these, 1,138 stations, or 40 percent, (and 403 out of 509 NASQAN stations retrieved) had data for at least one constituent (excluding temperature) which were found suitable for trend testing.

METHODOLOGY FOR AUTOMATED TREND TESTING

In performing a trend test, the null hypothesis that the probability distribution of water quality for each of the seasons is unchanged over the period of record was tested. The possible outcomes were (1) the null hypothesis was rejected with some degree of confidence (alpha=0.05 for the 1990-91 NWS) and it was declared that a trend existed in the data or (2) the null hypothesis was not rejected and it was declared that a trend could not be discerned.

Not all data records are suitable for trend testing. The length of record may be too short, there may be too few values, or the sampling schedule may be biased towards certain hydrologic events such as floods. Because of the large volume of data evaluated for the NWS, automated criteria were formulated in computer software to objectively determine the suitability of a water-quality record for trend testing. Some of these criteria have been considered by others (R. M. Hirsch, R. B. Alexander, and R. A. Smith, written communication, 1989, and T. L. Schertz, R. B. Alexander, and D. J. Ohe, written communication, 1989). Additional criteria were formulated as experience was gained in applying the criteria to the data base.

Trends Testing Periods

For purposes of the 1990-91 NWS, data records for three time periods (corresponding to calendar years) were selected from the data base for testing: 1970-89, 1975-89, and 1980-89. Data records for each time period were tested separately. Records before January 1970 were retained in the data base but were not tested .

Data Preparation

Paired streamflow and concentration data are required to test for trend in flow-adjusted concentration values (see section "Flow Adjustment"). If instantaneous discharge (NWIS parameter code 61) was not available, then daily mean discharge (parameter code 60) was used.

For certain chemical constituents (primarily trace inorganic elements), "less than" or nondetected values (values reported as less than or equal to a specified reporting limit) were sometimes reported in the data base as zero and the analytical reporting limits of these data are unknown. The Seasonal Kendall (SK) trend test (see section "Trend Test" for a description of the test) requires knowledge of the reporting limits used for nondetected values. For nondetected values reported as zero, the reporting limit was estimated by searching backward in time within the data record of the station for the previous nonzero nondetected data value, and then inferring this value as the reporting limit. If no such value was found by searching backwards, the record was searched forwards through 5 years of data. Because analytical reporting limits commonly are lowered as the precision of analytical procedures is improved, preference was given to backward searches so that any error in estimating reporting limits would be towards higher reporting limits. If the situation still could not be resolved after a forward search, the zero data value was discarded.

Trend Test

A test for monotonic trend was performed with the SK test (Hirsch and others, 1982) on water-quality records that met the criteria described in the section "Trend Qualifications." This seasonally-adjusted, nonparametric test examines all possible seasonal pairs of data values and tallies the number of times the later value is higher (positive difference) or lower (negative difference) than the earlier one. That is, in a monthly SK test, each January value is compared to every other January value, February to February, and so on. If there is no trend in the record, the

the number of positive and negative differences would tend to be equal. The seasonality of data records tested for trend was determined according to the methods described in the sections "Seasonal Designation of Water-Quality Data" and "Trend Qualifications." Because the SK test incorporates comparisons of the ranks of the data, effects of outliers on trend detection are minimized. Moreover, the SK test may be applied to water-quality records with nondetected or censored data (data reported as less than a specified reporting limit) as long as a single reporting limit is selected for the record.

Censored data values in a water-quality record were adjusted before trend testing in the following manner.

(1) If fewer than 10 percent of the observations in the record were censored, the censored data values were assigned one-half the reporting limit and treated as uncensored data.

(2) If more than 10 percent or the data were censored, all data (censored and uncensored) below the highest reporting limit were considered to be at this reporting limit.

An estimate of the rate of change or the trend slope for the period of analysis was computed according to Sen (1968). The trend slope, expressed as the change in original units per year, was computed as the median of all pairwise comparisons (each paired difference is divided by the number of years separating the pair of observations). The magnitude of the trend slope estimate was not reported for water-quality constituents when more than 10 percent of the data were censored (category two described above) since the estimate is likely to be inaccurate.

Seasonal Designation of Water-Quality Constituents

Water-quality constituents commonly are influenced by seasonally varying biochemical or hydrologic processes or human activities. In investigations that have the objective of identifying possible reasons for the occurrence of trends, adjustments commonly are made in water-quality data for seasonal variability to improve the ability of a trend test to detect changes in ambient water-quality conditions. The SK test adjusts for seasonal variations in data by allowing comparisons of data values to be made only within the same season of different years. Therefore, if the seasonality is defined as monthly, then only the sample values in the same month of different years are compared to one another.

Each water-quality record was evaluated for three possible seasonal designations (monthly, bimonthly, and quarterly). Representative monthly samples were first selected, if they existed, from the data record. The vast majority of the water-quality stations analyzed were sampled no more frequently than monthly. In those few cases of more frequent sampling, the sample collected nearest the 15th of the month was selected (for example, the 15th, 16th, 14th, 17th, and so on, was the order of selection). If there were multiple values on the day closest to the 15th, the first noncensored value collected on that day was given preference; the first censored value was selected as an alternative. Bimonthly (January-February, etc.) and quarterly (December-January-February, etc.) records then were selected from the monthly records by choosing a single monthly value (if they exist) for each defined season of the year according to the order specified in Table 2.

Table 2. Order of selection of monthly data for bimonthly and quarterly data sets

SEASON	FIRST MONTH	SECOND MONTH	THIRD MONTH
Bimonthly 1	January	February	
Bimonthly 2	March	April	
Bimonthly 3	May	June	
Bimonthly 4	July	August	
Bimonthly 5	September	October	
Bimonthly 6	November	December	
Quarterly 1	January	February	December (prior year)
Quarterly 2	April	May	March
Quarterly 3	July	August	June
Quarterly 4	October	November	September

The seasonal designation that most closely conformed to the minimum sampling frequency used during the period of record was selected for trend testing. This choice of seasonality for a record was of particular importance because significant differences in the sampling frequency of a record, especially in the first and last portions of the record, can result in temporally biased trend determinations. For example, the application of a monthly SK trend test to a record where the sampling frequency changed from monthly to quarterly would bias the trend results towards the earlier portion of the time period which has a larger number of observations. That is, a greater number of data comparisons would be made among observations during the earlier portion of the period of analysis; for certain months, comparisons of observations would not be made for the entire period of analysis. For those water-quality records in which the sampling frequency changed during the period of record (typically from a higher to a lower frequency), the lowest sampling frequency (quarterly rather than monthly) was selected to avoid bias in trend testing. The methods used to select the seasonality of a record from among the three possible seasons are described in detail in the section "Trend Qualifications."

Trend Qualifications

In general, water-quality records were tested for trend only if enough data values were present during the first and last portions of the record so that an accurate assessment could be made of changes in water-quality conditions during the period of analysis. There are two criteria, described here in general terms, for evaluating the suitability of a water-quality record for trend testing. First, the data must have nearly spanned the time period of analysis. Second, for a given seasonal frequency, the beginning and ending portions of the record must have contained sufficient data such that a majority of the possible number of pairwise comparisons (as made in the SK test) were present for most of the seasons. Records with samples collected less frequently than quarterly were not tested for trend since these records do not adequately reflect seasonal variability in water quality.

For a given period of record and seasonal sampling frequency, the above criteria are stated more precisely as follows.

(1) The period of record must have spanned the period of analysis, with no more than 2 years of missing data at the beginning and end portions of the period of analysis. For a period of P years, the length of time in years between the first and last observation in this period of record must have spanned a time of at least P minus 2 years. Therefore, for the three designated periods of analysis, 1970-89, 1975-89, and 1980-89, the minimum length of the periods of record were 18, 13, and 8 years, respectively.

(2) At least one-half of the possible number of seasonal, pairwise data comparisons must have been present in the first and last thirds of the record. The seasonal, pairwise data comparisons used

for this purpose were those made in the SK test except that the data comparisons were restricted to comparisons of values in the first third of the record with values from the same season in the last third of the record[2]. The length in years of the first and last thirds of the period of record was defined as the length in years of the period of analysis divided by three and rounded to the next largest integer. In making seasonal, pairwise comparisons for a monthly time series for example, all January values in the first third of the record are compared with all January values in the last third of the record; no comparisons of are made among January values within the first or within the last third of the record. If the first and last thirds of the record are each N years long, then the maximum number of possible comparisons per season is N^2.

(3) To determine whether the record qualifies for monthly, bimonthly, or quarterly analysis, the following procedure was followed. A subset of the data was used to construct a monthly record (according to the rules given in "Seasonal Designation of Water-Quality Data"). If at least 7 of the 12 months had at least $N^2/2$ valid pairs, then the record qualified for monthly trend testing, otherwise the bimonthly record was considered. If at least 5 of 6 bimonths had at least $N^2/2$ valid pairs, then the record qualified for bimonthly trend testing, otherwise the quarterly record was constructed. If at least 3 of 4 quarters had at least $N^2/2$ valid pairs, then the record qualified for quarterly trend testing; otherwise no trend test was performed.

Flow Adjustment

Water quality data were adjusted, if possible, to lessen the effects of short-term variability in streamflow, thus improving the detection of trends in water quality. A LOcally WEighted Scatterplot Smoothing (LOWESS) procedure (Cleveland, 1979) was used to relate constituent concentrations and streamflow. A natural log transformation was made to streamflow for all constituents, and to both concentration and streamflow for the following constituents: suspended sediment, turbidity, organic carbon, fecal bacteria, all nutrient species, and biological measures. Flow-independent residuals (observed concentrations minus predicted or smoothed concentrations) were computed from the concentration-streamflow relation, and then tested for trend with the SK test.

The LOWESS procedure is a robust method of fitting a smoothed line to bivariate data. The use of distance and residual weighting functions in conjunction with weighted least squares minimizes the influence of distant and outlying points in fitting a smooth line to the data. The degree of distance weighting is controlled by adjusting the magnitude of f, the smoothing factor. Possible values for f lie in the range 0 to 1, and represent the fraction of the total number of observations in the record used in the LOWESS computation for any given discharge. A smoothing factor of 0.5 was selected for use in the automated execution of the LOWESS procedure after conducting visual examinations of LOWESS fits for a large number of water-quality and discharge records. The selected smoothing factor of 0.5 tended to give a good fit to the data without smoothing out real features of the relation or producing many abrupt local changes in slope.

No flow adjustment was made if more than 10 percent of the water-quality values were censored, if there were fewer than 25 observations, or if more than half of the water-quality values had the same flow value.

---

[2] In evaluating the suitability of a water-quality record for trend testing, no restrictions were placed on the availability of data values in the middle portion of each record, which allowed a greater number of water-quality records to be qualified for trends testing.

INTERACTIVE DISPLAY OF DATA AND TREND TEST RESULTS

There are a great many possible combinations of time periods, constituents, and flow-adjustment conditions that can be examined for trends in a State data base. An interactive display system was developed to allow users to examine water-quality records and trend results as maps, graphs, and listings on a computer terminal. Although documentation of the display software is beyond the scope of this report, key features of the system are presented here.

The display system operates through "pull down" menus, where the user points the terminal's crosshairs to choose commands. A full suite of "help" choices is available to explain all the commands in detail. Maps of water-quality sampling stations in a State showing trends for any constituent can be displayed, as shown in figure 2. The user can choose the constituent, the time period, and whether or not to show flow-adjusted trends. Default choices allow for a quick review of the data. By pointing the terminal crosshairs to any station, the computer provides the user with a scatterplot of the relation between water-quality data and time, along with a LOWESS-smoothed line, and a scatterplot of the relationship between water-quality data and flow, along with the LOWESS flow-adjustment line. Other options allow the user to see diagnostic printouts of trend calculations, list the water-quality data, or display box-and-whisker plots of water-quality data for selected stations.

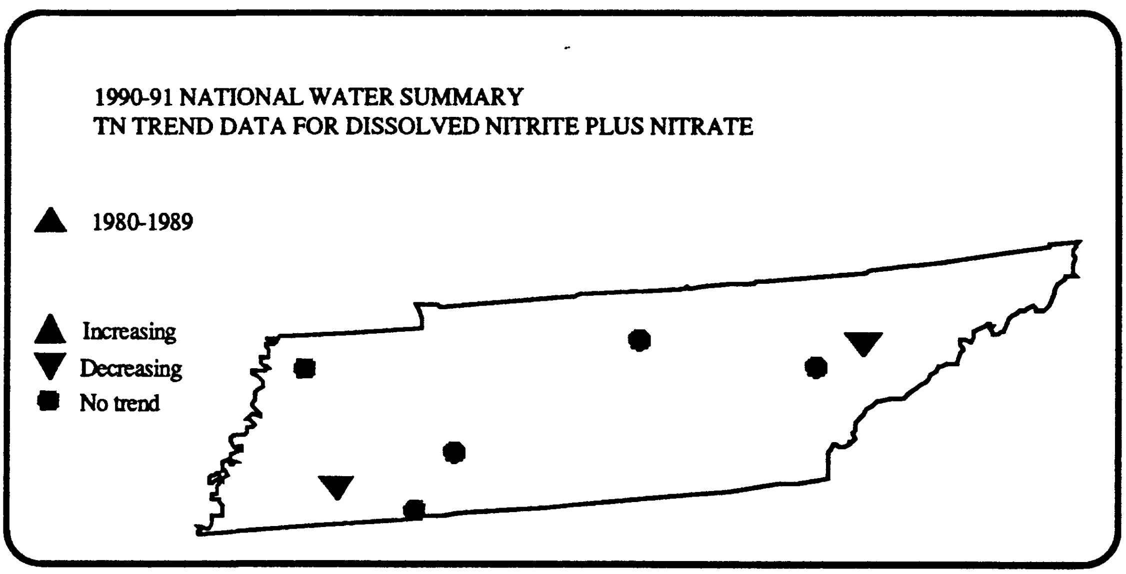

Figure 2. -- Example of a computer screen display showing trends.

A principal objective of the interactive system was for users to compose the contents of the NWS's water-quality trends illustrations on the terminal screen and to save the information for eventual automated production of the illustrations for the author's State report. The system, however, exerted a very important influence in monitoring the performance of the trend qualification algorithms and the trend test under a wide variety of conditions.

QUALITY ASSURANCE

Quality assurance procedures centered around two concerns. First, all data had to be traceable back to its source, and any modifications to the data had to be thoroughly documented. Second, the trend-qualifying and trend-testing procedures had to work correctly under a wide range of conditions.

All data were entered into the data base by automated procedures that summarized all transactions in a data log maintained on the computer. Any subsequent manual changes made to the data were also recorded in this log. Every data value was associated with a specific data-file retrieval, and all such files were saved on magnetic tapes. With these magnetic tapes and the data log file, it is possible to exactly reconstruct the NWS data base.

The software must be designed so that it correctly decides whether or not to do a monthly, bimonthly, or quarterly trend test. A key concern is that in no known case should the software perform a trend test on inappropriate data. Incorrectly rejecting an appropriate data record, while not nearly as serious as accepting an inappropriate one, must not occur so often as to make the test worthless. Initially, the software was thoroughly tested by four authors on data from the States of California, Missouri, New York, and Tennessee. These authors examined several thousand records in some detail. In cases where problems were encountered -- some stations were found, for example, with an unusual mixture of detection limits for some constituents --, the software was refined until it successfully handled all data records. Additional testing and refinement of the software occurred as the data and results were examined by other National Water Summary authors. The criteria discussed previously in this report reflect the refinement that occurred as a result of this review.

SUMMARY

Methods were developed to examine a water-quality record and, based on the application of a detailed set of trend testing criteria, determine whether a monthly, bimonthly, or quarterly test for monotonic trend could be performed. The trend testing criteria were automated and allowed for an efficient and accurate analysis for trend in an extremely large number of water-quality records. Water-quality records for about 3,000 stations nationwide were retrieved from the U.S. Geological Survey's National Water Information System, and were evaluated as to their suitability for trend testing. From this set of records, it was determined that about 1,100 stations had at least one water-quality constituent that met the criteria for trend testing. For those water-quality records, a nonparametric test for monotonic trend, known as the seasonal Kendall test, was applied to three time periods (1970-89, 1975-89, and 1980-89). The results of these water-quality trends are intended for use in the U.S. Geological Survey's National Water Summary Program activities on stream water quality.

Perhaps the most important lesson to be learned from the experience of developing an automatic trend test is that software must be very "smart" if it is to cope with the myriad ways in which water-quality data are collected and recorded. Irregular or changed sampling frequencies and outliers (which may be due to data-input errors) are not all that uncommon, even in a data base considered to be of generally high quality.

Prior to this examination, it was thought likely that the approximately 600 NASQAN stations represented the vast majority of water-quality stations with NWIS data records suitable for use in trend tests. It now appears there are nearly twice this number of stations with suitable water-quality records within the NWIS data base. The wide variety of constituents sampled at NASQAN stations were not sampled at all these stations, but many supplement the NASQAN data to present a more comprehensive national and statewide picture of water-quality trends.

REFERENCES CITED

Cleveland, W.S., 1979, Robust locally weighted regression and smoothing scatterplots: Journal of the American Statistical Association, v. 74, p. 829-836.

Hirsch, R.M., Slack, J.R., and Smith, R.A., 1982, Techniques of trend analysis for monthly water-quality data: Water Resources Research, v. 18, no. 1, p. 107-121.

Parks, Bruce, 1988, Processing data from the U.S. Geological Survey's National Water Data Exchange by use of a geographic information system, in ESRI User Conference, 8th, March 21-25, 1988, Palm Springs, Calif., Proceedings: Redlands, Calif., Environmental Systems Research Institute, 10 p.

Sen, P.K., 1968, Estimates of regression coefficient based on Kendall's tau: Journal of the American Statistical Association, v. 63, p. 1379-1389.

Smith, R.A., Alexander, R.B., and Wolman, M.G., 1987, Water-quality trends in the Nation's rivers: Science, v. 235, p. 1607-1615.