Journal of Hydrology, 150 (1993) 61–80

0022-1694/93/$06.00 © 1993 – Elsevier Science Publishers B.V. All rights reserved

[4]

Non-parametric trend analysis of water quality data of rivers in Kansas

Yun-Sheng Yu[a,]*, Shimin Zou[a], Donald Whittemore[b]

[a]*Department of Civil Engineering, University of Kansas, Lawrence, KS 66045, USA*
[b]*Kansas Geological Survey, Lawrence, KS 66045, USA*

(Received 17 November 1992; accepted 25 January 1993)

Abstract

Surface water quality data for 15 sampling stations in the Arkansas, Verdigris, Neosho, and Walnut river basins inside the state of Kansas were analyzed to detect trends (or lack of trends) in 17 major constituents by using four different non-parametric methods. The results show that concentrations of specific conductance, total dissolved solids, calcium, total hardness, sodium, potassium, alkalinity, sulfate, chloride, total phosphorus, ammonia plus organic nitrogen, and suspended sediment generally have downward trends. Some of the downward trends are related to increases in discharge, while others could be caused by decreases in pollution sources. Homogeneity tests show that both station-wide trends and basin-wide trends are non-homogeneous.

Introduction

One of the long-term goals of the monitoring program of stream water quality is to detect changes or trends in pollution levels over time and to identify, describe, and explain the major factors that affect trends in water quality. A vast amount of surface water quality data has been collected in Kansas by the Kansas Department of Health and Environment and the USGS. Analysis of the lower Kansas River data is being conducted by the USGS researchers (Jordan and Stamer, 1991), as part of the National Water-Quality Assessment Program (NAWQA) (Hirsch et al., 1988).

The purpose of this study was to apply non-parametric methods of trend analysis to the monthly surface water quality data for the upper and lower Arkansas River inside Kansas, the Neosho River, the Verdigris River, and the

* Corresponding author.

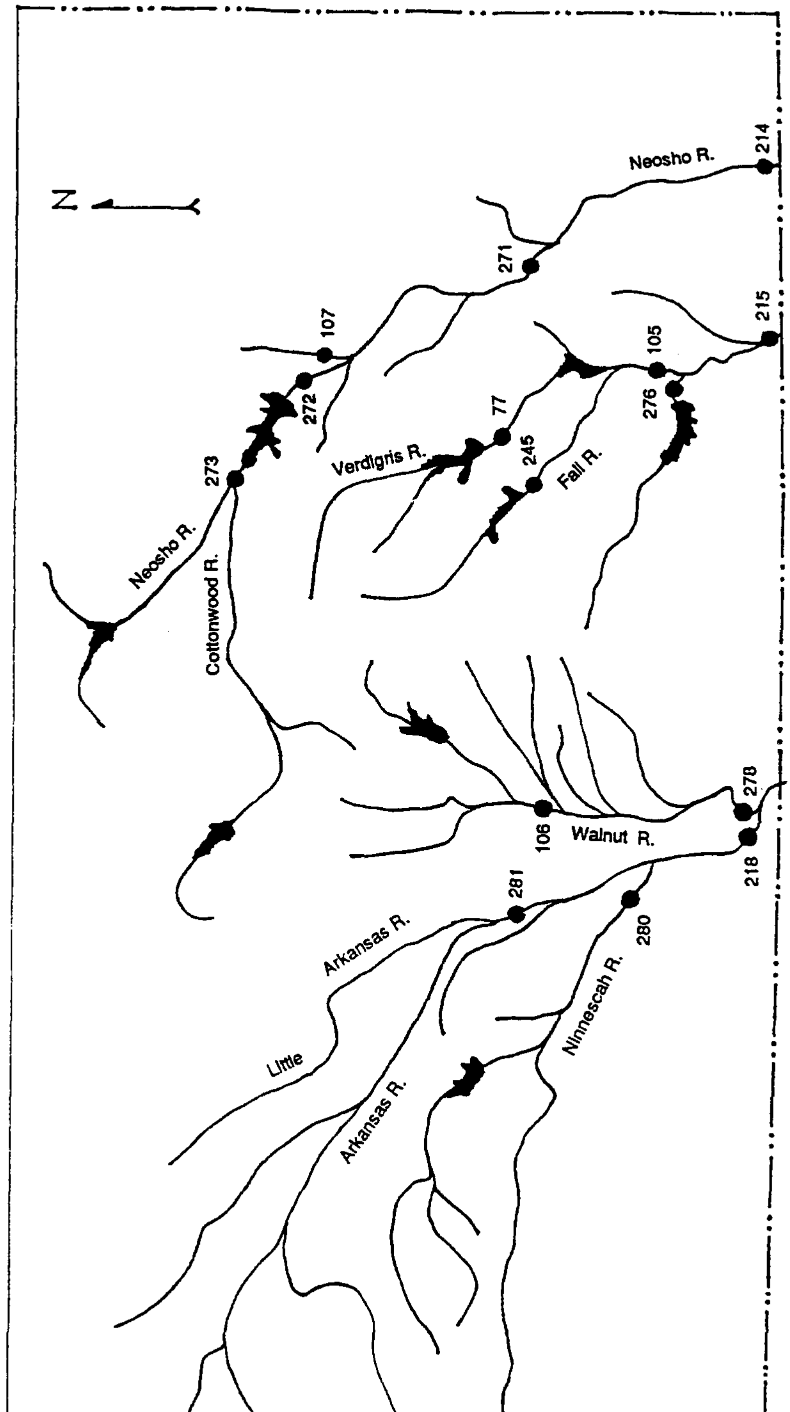

Fig. 1. Study area and locations of sampling stations.

Walnut River to detect water quality trends (or lack of trends) and the homogeneity of trends. Exploratory data analyses were conducted to define the characteristics of the data set for selecting the trend analysis procedures. Four different non-parametric trend tests were selected to detect linear trends of water quality. They are the Mann–Kendall test, the seasonal Kendall test, Sen's T test, and the Van Belle and Hughes test. A comparison of these four methods based on their power for different sample sizes shows that for sample sizes larger than 9 years, all four methods have the same power. The trend (or lack of trend) for each individual constituent has been determined by using the Mann–Kendall test and the seasonal Kendall test. The magnitudes of linear trends are computed by using Sen's estimator. The homogeneity of trend direction at multiple stations and, in different seasons, is also tested by using the Van Belle and Hughes method.

Data preparation and exploratory analyses

Most of the data used in this study were collected by the Kansas Department of Health and Environment. The locations and identification numbers of the 15 sampling stations are shown in Fig. 1. After the initial compilation of the data, 17 constituents, as listed in Table 1, were selected for exploratory analyses. These constituents are part of the target variables chosen in the pilot

Table 1
The selected 17 water-quality constituents

Name of variable	Code
Specific conductance	00095
pH	00400
Discharge	00061
Temperature	00010
Sulfate	00410
Chloride	00940
Total phosphorus	00665
Calcium	00915
Magnesium	00925
Dissolved oxygen	00300
Ammonia + organic nitrogen	00625
Total hardness	00900
Dissolved solids	70300
Suspended sediment	80154
Alkalinity	00410
Potassium	00935
Sodium	00930

National Water-Quality Assessment Program (Hirsch et al., 1988). The period of records used in this analysis was from December 1975 through November 1989. Monthly time series for each of the 17 constituents were obtained at every sampling station. If two or more observations were available in a month, then the arithmetic mean of these values was taken as the value for the month. Missing values were checked and filled by the average value of surrounding observations.

The following preliminary data analyses for all water quality constituents at every sampling station were conducted: (1) time series plot; (2) Box-and-Whisker plot; (3) summary statistics; (4) distribution test; (5) time-dependence test; (6) seasonality test; (7) flow relatedness test. Most of these analyses can be conducted by using available commercial software for statistics, for example, SPSS (Norusis, 1988) or SAS. The results are presented in detail elsewhere (Yu et al., 1991), and the following is a brief summary.

The time series plot provides a visual impression of the degree of trend, periodicity, and outliers, and perhaps it may suggest clues on assumption violation. Seasonal (monthly) and annual Box-and-Whisker plots were prepared to detect seasonal changes in the distribution, and the annual plot together with the time series plot gives an indication of time trend. Summary statistics include sample means and standard deviations.

The Kolmogrov–Smirnov goodness-of-fit test was used to test the null hypothesis of normality for each time series. Those rejected in normal distribution tests at a significance level of 0.05 were tested again using a log–normal distribution hypothesis. At the significance level of 0.05, the normality distribution hypothesis was rejected for more than 40% of the time series. Among those failing the normality distribution test, about 40% would also fail the log–normal distribution test at the same significance level.

The time-dependence test and seasonality test of the data determine, respectively, the correlation among observations of water quality constituents and cyclic variations in the time series. A sample of correlation matrix computed from data for 16 constituents is shown in Table 2. Eighty percent of all values of correlation coefficient in Table 2 are statistically significant using a two-tailed test at a significance level of 0.01. The computed autocorrelation function generally shows a lag-one coefficient lying outside the 95% confidence interval which shows the serial correlation of the time series.

The seasonality of the time series was tested by using the Kruskal–Wallis method (Conover, 1980). The results indicate that the hypothesis of no seasonal variation was rejected for about 45% of the data tested at a significance level of 0.05. It is well known that concentrations of water quality constituents are correlated with river discharge. Flow adjustment for trend detection has been suggested by Hirsch et al. (1982). However, flows in the rivers of the

Table 2
Correlation matrix for Station 281

ALKAL	AMORG	CHLOR	DISLD	DISOX	PTSUM	SPCND	SULFT	TOHRD	TOPHS	CLCIM	DISCH	MGNES	SDIUM	SEDMT	TEMPT
1.000															
0.441	1.000														
0.781	0.222	1.000													
0.865	0.270	0.940	1.000												
0.239	−0.017	0.347	0.282	1.000											
−0.201	0.237	−0.189	−0.174	−0.215	1.000										
0.850	0.263	0.948	0.978	0.280	−0.187	1.000									
0.562	0.144	0.386	0.647	−0.026	−0.052	0.064	1.000								
0.862	0.203	0.709	0.883	0.182	−0.241	0.855	0.849	1.000							
0.208	0.758	0.037	0.077	−0.123	0.348	0.057	0.001	−0.013	1.000						
0.892	0.212	0.720	0.867	0.234	−0.296	0.844	0.746	0.965	−0.006	1.000					
−0.600	−0.346	−0.539	−0.550	−0.021	0.020	−0.581	−0.350	−0.507	−0.191	−0.510	1.000				
0.708	0.148	0.578	0.788	0.062	−0.162	0.764	0.937	0.950	−0.035	0.864	−0.424	1.000			
0.810	0.249	0.980	0.976	0.308	−0.169	0.968	0.489	0.773	0.081	0.772	−0.544	0.658	1.000		
−0.668	−0.362	−0.571	−0.601	−0.121	0.064	−0.637	−0.463	−0.608	−0.131	−0.583	0.542	−0.561	−0.559	1.000	
−0.282	−0.198	−0.198	−0.203	−0.674	0.413	−0.206	0.049	0.229	0.090	−0.302	−0.034	−0.119	0.177	0.090	1.000

study region are regulated by reservoirs, diversions, and consumptive uses during the period of the trend analysis. Therefore, flow adjustment as suggested by Hirsch et al. (1991) should not be used.

The results of the exploratory analyses generally indicate that the water quality data for the study area are non-normal, serially correlated, censored with values below detection limits, and with seasonal fluctuations. Therefore, non-parametric methods, which are procedures for testing an hypothesis whereby the test does not depend upon the form of the underlying distribution of the null hypothesis of no trend, are used in this study. These non-parametric methods are more flexible than parametric methods and can handle these characteristics of time series more easily (Van Belle and Hughes, 1984; Berryman et al., 1988).

Methods of analysis

Two of the most common trends are step and monotonic trends (Lettenmaier, 1976; Gilbert, 1987). A step trend is an instantaneous change in the mean level at one point in time. This may result from a relatively abrupt permanent change in land use or the construction of a waste treatment plant. A prior hypothesis of a time of change is required for the step trend (Hirsch, 1988). A monotonic trend is simply an increase or decrease in the mean level over time. This may result, for example, from a change in nutrient loading from agricultural runoff. The problem of selecting appropriate methods for trend analysis of water quality time series has received considerable attention. Since there is no prior hypothesis of a time of change for the river basin and concurrent records from a variety of stations are being analyzed, the monotonic trend procedures are most appropriate (Hirsch et al., 1991) for this study.

Recently, several non-parametric tests for linear trends in water quality have been proposed by a number of investigators (e.g. Hirsch et al., 1982; Hirsch and Slack, 1984; Van Belle and Hughes, 1984; Lettenmaier et al., 1991; and others). Performances of different non-parametric methods have been compared by using Monte Carlo simulation (Van Belle and Hughes, 1984; Taylor and Loftis, 1989; Hirsch et al., 1991; Loftis et al., 1991). Although each method has its own limitations, the four methods outlined below are generally accepted as appropriate for detecting linear trends of water quality data. All four methods apply rank transformation to the data (i.e. the original data are replaced by their ranks in the procedure) and then use the usual parametric procedure. These methods are briefly described as follows.

The Mann–Kendall (MK) test

The Mann–Kendall test (Mann, 1945; Kendall, 1975) uses the deseasonalized data to test the null hypothesis of the randomness of data against trend. According to Mann, the null hypothesis H_0 states that the deseasonalized data $(x_1, x_2, \ldots, x_n)$ are a sample of n independent and identically distributed random variables. The alternative hypothesis (H_1) of a two-sided test is that the distribution of x_i and x_j are not identical for all $i, j \leqslant n$ with $i \neq j$.

The test statistic S is defined as

$$S = \sum_{i>j} \operatorname{sgn}(x_i - x_j) \tag{1}$$

where

$$\operatorname{sgn}(\theta) = \begin{bmatrix} 1 & \text{if}\, \theta > 0 \\ 0 & \text{if}\, \theta = 0 \\ -1 & \text{if}\, \theta < 0 \end{bmatrix} \tag{2}$$

Mann showed that under H_0 the distribution of S is normal in the limit as $n \to \infty$. Given the possibility that there may be ties in the values of x, Kendall obtained the mean and variance of S, under the assumption of no trend, as

$$E(S) = 0 \tag{3}$$

and variance

$$\operatorname{Var}(S) = \left[n(n-1)(2n+5) - \sum_t t(t-1)(2t+5) \right] / 18 \tag{4}$$

where t is the extent of any given tie (number of xs involved in a given tie) and $\sum_t$ denotes the summation over all ties. Both Mann and Kendall derived the exact distribution of S for $n \leqslant 10$ and showed that even for $n = 10$ the normal approximation is excellent, provided one computes the standard normal variate Z by using the following equation

$$Z = \begin{bmatrix} \dfrac{S-1}{[\operatorname{Var}(S)]^{\frac{1}{2}}} & \text{if}\, S > 0 \\ 0 & \text{if}\, S = 0 \\ \dfrac{S+1}{[\operatorname{Var}(S)]^{\frac{1}{2}}} & \text{if}\, S < 0 \end{bmatrix} \tag{5}$$

Thus, in a two-sided test for trend at a significance level of α, the H_0 should be rejected if $|Z| \geqslant z_{\alpha/2}$, where $F_N(z_{\alpha/2}) = 1 - \alpha/2$, F_N being the standard normal cumulative distribution function. A positive value of S indicates an 'upward trend' and a negative value of S indicates a 'downward trend'.

This procedure is particularly useful since missing values are allowed and the data need not conform to any particular distribution. Also, data reported as trace or less than the detection limit can be used by assigning them a common value smaller than the smallest measured value in the data set because the Mann–Kendall test and the seasonal Kendall test described in the next section use only the relative magnitudes.

The seasonal Kendall (SK) test

The seasonal Kendall test can be used for time series with seasonal variation, having missing values, tied values, or values below the limit of detection, and does not require normality of the time series. Hirsch et al. (1982) and Taylor and Loftis (1989) provide assessments of this test.

This test is intended to assess the randomness of a data set $\mathbf{X} = (X_1, X_2, ..., X_{12})$ and $\mathbf{X}_i = (x_{i1}, x_{i2}, ..., x_{in_i})$, where X is a matrix of the entire monthly data over n years for a single constituent at a sampling station. The matrix X consists of 12 column vectors $\boldsymbol{X_1}$ through a$\boldsymbol{X_{12}}$ (one for each month), and each subsample $\boldsymbol{X_i}$ contains the n_i annual values for month i. Specifically, the hypotheses are H_0: $\boldsymbol{X}$ is a sample of independent random variables (x_{ij}) and that $\boldsymbol{X_i}$ is a subsample of independent and identically distributed random variables for $i = 1, 2, ..., 12$; H_1: the subsample is not distributed identically. The test statistic is a sum of the Mann–Kendall statistic computed for each month (season). For the month i

$$S_i = \sum_{j>k} \mathrm{sgn}(X_{ij} - X_{ik}) \tag{6}$$

Under H_0 the subsample $\boldsymbol{X_i}$ satisfies the null hypothesis H_0 of Mann's test. From the Mann–Kendall procedure, one has the expectation (under the null hypothesis of no trends),

$$E(S_i) = 0 \tag{7}$$

and variance

$$\mathrm{Var}(S_i) = \left[n_i(n_i - 1)(2n_i + 5) - \sum_{t_i} t_i(t_i - 1)(2t_i + 5)\right] / 18 \tag{8}$$

and the distribution of S_i is normal in the limit as $n_i \to \infty$ (t_i is the extent of a given tie in month i). Furthermore, the grand statistic S^* can be defined as $S^* = \sum_i S_i$ and its expectation,

$$E(S^*) = \sum_i E(S_i) = 0 \tag{9}$$

and variance

$$\mathrm{Var}(S^*) = \sum_i \mathrm{Var}(S_i) + \sum_{i \neq j} \mathrm{Cov}(S_i, S_j) \tag{10a}$$

S^* has an approximate normal distribution. If the covariance terms are neglected, then

$$\mathrm{Var}(S^*) = \sum_i \mathrm{Var}(S_i) = \sum_i [n_i(n_i - 1)(2n_i + 5) - \sum_{t_i} t_i(t_i - 1)(2t_i + 5)]/18 \tag{10b}$$

In a later version, Hirsch and Slack (1984) include the covariances. However, if the errors are uncorrelated then the seasonal Kendall procedure without the covariance term performs well (Taylor and Loftis, 1989). Next, compute the standard normal deviate (zero mean and unit variance)

$$Z = \begin{bmatrix} \dfrac{S^* - 1}{[\mathrm{Var}(S^*)]^{\frac{1}{2}}} & \text{if } S^* > 0 \\ 0 & \text{if } S^* = 0 \\ \dfrac{S + 1}{[\mathrm{Var}(S^*)]^{\frac{1}{2}}} & \text{if } S^* < 0 \end{bmatrix} \tag{11}$$

For a two-sided test at a selected level of significance α, the null hypothesis of no trend is rejected if the absolute value of Z is greater than $z_{\alpha/2}$.

Sen's T test

In 1980, Farrell, following Sen (1968a), proposed an aligned rank test for detecting monotonic trends. On the basis of asymptotic relative efficiencies (Fleiss, 1981, Chapter 10), the aligned rank test is more powerful than the intrablock procedures (Van Belles and Hughes, 1984). This procedure is distribution free and not affected by seasonal fluctuations (Van Belle and Hughes, 1984). The computational procedures are as follows: let x_{ij} represent the measurement in the year i and month j for $i = 1, 2, ..., n$ and $j = 1, 2, ..., 12$ for a constituent at a sampling station.

(1) Compute the average for the month j, $X_{\cdot j} = (\sum_i x_{ij})/n$ and the average for

the year i, $X_{i\cdot} = (\sum_j X_{ij})/12$. Subtract the monthly average from each of the corresponding months in the n years of data to remove seasonal effects, i.e. calculate $X_{ij} - X_{\cdot j}$ for $i = 1, 2, ..., n$ and $j = 1, 2, ..., 12$.

(2) Rank all the differences from 1 to $12n$ to obtain the matrix (R_{ij}), where R_{ij} = rank of $(X_{ij} - X_{\cdot j})$ among the $12n$ values of differences. If t ties occur, the average of the next t ranks is assigned to each of the t tied values.

(3) The ranks for each year are averaged, i.e. $R_{i\cdot} = (\sum_j R_{ij})/12$ and likewise for each month, $R_{\cdot j} = (\sum_i R_{ij})/n$.

(4) Calculate the following statistic

$$T = \left[\frac{12m^2}{n(n+1)\sum_{i,j}(R_{ij} - R_{\cdot j})^2}\right]^{1/2} \left[\sum_{i=1}^{n}\left(i - \frac{n+1}{2}\right)\left(R_{i\cdot} - \frac{nm+1}{2}\right)\right] \tag{12}$$

where $m = 12$.

As n becomes large the distribution of T tends toward normality with mean 0 (under the null hypothesis of no trend) and unit variance. The statistic test is to reject the hypothesis of no trend if the absolute value of T exceeds a prespecified percentile of the normal distribution. Van Belle and Hughes (1984) showed by Monte Carlo simulation that the normal approximation for the statistic T was reasonable even for small samples.

Van Belle and Hughes χ^2 test for homogeneity of trends

All the tests discussed include an implicit assumption that the trend is homogeneous between seasons. When the trend is heterogeneous between seasons, for example, an upward trend in one season and a downward trend in another, an overall test of trend direction and slope estimator will be misleading. A preliminary test for homogeneity of trend direction is, therefore, necessary before any of these tests can be conducted.

The test procedure for homogeneity in seasonal trends at a given station developed by Van Belle and Hughes (1984) involves the computation of the following statistic.

$$\chi^2_{\text{homog}} = \chi^2_{\text{total}} - \chi^2_{\text{trend}} = \sum_{i=1}^{m} Z_i^2 - m\bar{Z}^2 \tag{13}$$

where

$$Z_i = \frac{S_i}{[\text{Var}(S_i)]^{1/2}}$$

S_i is the Mann–Kendall statistic, computed from the data collected over n years for the season i, and

$$\bar{Z} = \frac{1}{m}\sum_{j=1}^{m} Z_i \tag{14}$$

If χ^2_{homog} exceeds the α level critical value for the chi-square distribution with $(m-1)$ degrees of freedom (d.o.f.), the null hypothesis of homogeneous seasonal trends over time (trends in the same direction and of the same magnitude) must be rejected. In that case the seasonal Kendall test and slope estimate are not meaningful, and it is best to use the Mann–Kendall test for each individual season. If χ^2_{homog} does not exceed the α level critical value, then the calculated value for χ^2_{trend} is referred to the chi-square distribution with 1 d.o.f. to test for a common trend in all seasons. This procedure can be extended to test the homogeneity of trends of different constituents at a station or homogeneity of basin-wide trend of a constituent.

When no seasonal cycles are present in the data set, the above procedure can be used to test the homogeneity of trend direction at multiple stations in a basin by replacing the number of months m by the number of stations k in Eqs. (13) and (14). In this case, if χ^2_{homog} exceeds the α level critical value, the H_0 of homogeneity of stations in the basin is rejected. However, the Mann–Kendall test for trend at each station may still be used. If χ^2_{homog} does not exceed the α level critical value, then the statistic χ^2_{trend} is referred to the chi-square distribution with 1 d.o.f. to test the null hypothesis that the (common) trend direction is not significantly different from zero.

In the previous section, the χ^2_{homog} statistic was used to test the homogeneity of trend direction in different seasons at a given sampling station or at multiple stations when no seasonality is present in the data. The same procedure can also be used to test the homogeneity of trend direction at different stations when seasonality is present as follows.

Let x_{ijl} represents the observation at station l in year i and month j, where $i = 1, 2, ..., n, j = 1, 2, ..., m$ and $l = 1, 2, ..., k$; n, m and k are number of years of records, number of months (seasons) in a year, and number of stations, respectively. The first step is to compute the Mann–Kendall statistic for each season at each station. Let S_{jl} denote this statistic for the season j at the station l. Then compute

$$Z_{jl} = \frac{S_{jl}}{[\text{Var}(S_{jl})]^{1/2}}, \quad j = 1, 2, ..., m, l = 1, 2, ..., k \tag{15}$$

Next, compute the mean over k stations for the season j,

$$\bar{Z}_{j\cdot} = \frac{1}{k}\sum_{l=1}^{k} Z_{jl}, \quad j = 1, 2, ..., m \tag{16}$$

the mean over m seasons for station l,

$$\bar{Z}_{\cdot l} = \frac{1}{m}\sum_{j=1}^{m} Z_{jl}, l = 1, 2, ..., k \tag{17}$$

the grand mean over all km station–season combinations,

$$\bar{Z}_{\cdot\cdot} = \frac{1}{km}\sum_{j=1}^{m}\sum_{l=1}^{k} Z_{jl} \tag{18}$$

Then, compute the chi-square statistics according to equations shown in Table 3 in the following order:

(1) χ^2_{total}, χ^2_{trend}, χ^2_{station}, and χ^2_{season}.
(2) $\chi^2_{\text{homog}} = \chi^2_{\text{total}} - \chi^2_{\text{trend}}$.
(3) $\chi^2_{\text{station-season}} = \chi^2_{\text{homog}} - \chi^2_{\text{station}} - \chi^2_{\text{season}}$.

The acceptance or rejection of the hypothesis can then be determined by

Table 3
Equations for testing trend homogeneity using the Van Belle and Hughes method (1984)

Chi-square statistics	Degree of freedom
$\chi^2_{\text{total}} = \sum_{j=1}^{m}\sum_{l=1}^{k} Z^2_{jl}$	km
$\chi^2_{\text{homog}} = \sum_{j=1}^{m}\sum_{l=1}^{k} Z^2_{jl} - km\bar{Z}^2_{\cdot\cdot}$	$km - 1$
$\chi^2_{\text{season}} = k\sum_{j=1}^{m} \bar{Z}^2_{j\cdot} - km\bar{Z}^2_{\cdot\cdot}$	$m - 1$
$\chi^2_{\text{station}} = m\sum_{l=1}^{k} \bar{Z}^2_{\cdot l} - km\bar{Z}^2_{\cdot\cdot}$	$k - 1$
$\chi^2_{\text{station-season}} = \sum_{j=1}^{m}\sum_{l=1}^{k} Z^2_{jl}$	$(k-1)(m-1)$
$k\sum_{j=1}^{m} \bar{Z}^2_{j\cdot} - m\sum_{l=1}^{k} \bar{Z}^2_{\cdot l} + km\bar{Z}^2_{\cdot\cdot}$	
$\chi^2_{\text{trend}} = km\bar{Z}^2_{\cdot\cdot}$	1

comparing the computed values of χ^2_{station}, χ^2_{season}, and $\chi^2_{\text{station-season}}$ with the α level critical values in the standard chi-square table with $(k-1)$, $(m-1)$, and $(k-1)(m-1)$ d.o.f., respectively.

If all three tests are non-significant, meaning homogeneous trends in all three tests, refer χ^2_{trend} to the chi-square distribution with 1 d.o.f. to test for global trends. If χ^2_{season} is significant (non-homogeneous seasonal trend), but χ^2_{station} is not (homogeneous basin-wide trend), then test for a different trend direction in each season by computing the m seasonal statistic

$$k\bar{Z}^2_{j\cdot} \quad j = 1, 2, ..., m \text{ seasons} \tag{19}$$

and refer each to the α level critical value of the chi-square distribution with 1 d.o.f.

If χ^2_{station} is significant (non-homogeneity of stations), but χ^2_{season} is not (homogeneity in seasonal trend), then test for trend at each station by computing the k station statistics

$$m\bar{Z}^2_{\cdot l} \quad l = 1, 2, ..., k \text{ stations} \tag{20}$$

and refer to the α level critical value of the chi-square distribution with 1 d.o.f.

If both χ^2_{station} and χ^2_{season} are significant (non-homogeneity in both seasonal and station trends) or if $\chi^2_{\text{station-season}}$ is significant (non-homogeneous in station–season interaction), then the χ^2 trend test should not be done. The only meaningful trend tests in that case are those for individual station–season combinations. These tests are made by referring each Z_{jl} statistic to the α level critical value of the standard normal distribution, as discussed earlier. For these individual Mann–Kendall tests or seasonal Kendall tests, the Z_{jl} should be recomputed so as to include the correction for continuity.

Note that if the number of stations is taken as k water quality constituents at one station and the water quality constituent replaces station in Table 3 and all previous equations, then the above procedure can be used to test the homogeneity of trend directions for different constituents at a given sampling station. This test is referred to as a station-wide trend test and the previous one is called the basin-wide trend test.

Sen's estimator of slope

If a linear trend is present, the true slope (change per unit time) can be estimated by using a simple non-parametric procedure developed by Sen (1968b). This procedure is not greatly affected by gross data errors or outliers and can be used for records with missing values. The computational procedures are as follows.

First, compute the slope estimates of N pairs of data, Q,

$$Q = (x_j - x_k)/(j - k) \tag{21}$$

where x_j and x_k are data values at times j and k, respectively, with $j > k$. The median of these N values of Q is Sen's estimator of slope. If there is only one datum in each time period, then $N = n(n-1)/2$, where n is the number of time periods. The median of the N slope estimates is obtained in the usual way, that is, the N values of Q are ranked by $Q_{[1]} \leqslant Q_{[2]} \leqslant \ldots \leqslant Q_{[N-1]} \leqslant Q_{[N]}$ and

$$\text{Sen's estimator} = \begin{bmatrix} Q_{[(N+1)/2]} & \text{if } N \text{ is odd} \\ \frac{1}{2}[Q_{(N/2)} + Q_{[(N+2)/2]}] & \text{if } N \text{ is even} \end{bmatrix} \tag{22}$$

A generalization of Sen's estimator of slope is the seasonal Kendall slope estimator which computes the individual N_i slope estimates for the month (season) i for $i = 1, 2, \ldots, 12$

$$Q_i = (x_{ij} - x_{ik})/(j - k) \tag{23}$$

where, x_{ij} is the datum for the month i of the year j, and x_{ik} is the datum for the month i of the year k, where $j > k$. Rank the $\sum_i N_i = N'$ individual slope estimates, and find their median which is the seasonal Kendall slope estimator. A $100(1 - \alpha)\%$ two-sided confidence interval about the true slope may be obtained by the non-parametric technique given by Sen (1968b).

Results and discussions

Comparison of power of non-parametric tests

The comparison of power of non-parametric tests is usually conducted by means of Monte Carlo simulation studies under the assumption that the data have some trend patterns or ideal distributions. The characteristics of the real-world data may vary significantly for different river basins, different water quality constituents, different sampling frequencies, and different record lengths. Therefore, in this study, a comparison of the relative power of the four different procedures is made by using the monthly water quality data up to 15 years.

For each selected sample size, $n = 3$, 5, 9, and 15 years, the percentage of rejection is plotted vs. significance level for each of the four tests. The results for $n = 9$ are shown in Fig. 2. The results for other n are not presented. The results for the Mann–Kendall test with deseasonalized data (MKD) and the seasonal Kendall test (SK) for $n = 3$ and 5 years, respectively, show that at the same significance level, the percentage of rejection for MKD is higher than

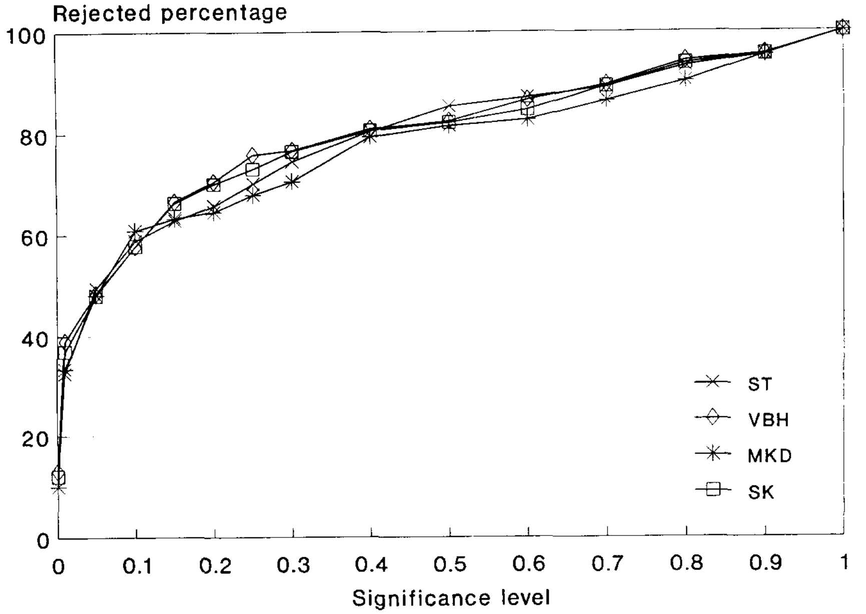

Fig. 2. Comparison of methods of test for $n = 9$.

that for SK. The Mann–Kendall deseasonalized method is apparently more powerful than the seasonal Kendall method for small sample sizes. For $n = 9$, there are no significant differences in relative power between these methods. This is also true for $n = 15$ years. In practice, the power at significance level lower than 0.1 is of interest. A close examination of the test results shows (Yu et al., 1991) that the P values of ST and MKD are close and those of VBH and SK are close. The reason is that both ST and MKD are aligned rank methods and asymptotically more powerful than intrablock methods such as SK (Van Belle and Hughes, 1984).

Trend tests for individual constituents

Before these tests were conducted, the homogeneity of seasonal trends was tested because, as stated previously, the MKD, SK, and Sen's T tests all implicitly assume that seasonal trends are homogeneous. The test results showed evidence of non-homogeneity in only one set of data, i.e. that for the specific conductance at Station 214, where the rejection was significant at $\alpha = 0.05$. The homogeneity hypothesis was acceptable for all the other data tested.

For all tests, the significance level is $\alpha = 0.05$. The numerical results for all 17 constituents are presented elsewhere (Yu et al., 1991). Numerical results

Table 4
Results of trend tests for specific conductance, pH, and ammonia + organic nitrogen by seasonal Kendall test and Sen's slope

Station no.	Specific conductance		pH		Ammonia + organic nitrogen	
	P value	Sen's slope (unit year^{-1})	*P* value	Sen's slope (unit year^{-1})	*P* value	Sen's slope (unit year^{-1})
77	0.000	−5.8	0.212	−0.075	0.000	−0.01
105	0.626	−4.9	0.000	−0.10	0.752	0.00
106	0.176	−22.5	0.442	−0.02	0.011	−0.01
107	0.226	4.4	0.643	0.00	0.354	0.00
214	0.000	−10.6	0.199	−0.05	0.000	−0.01
215	0.000	−8.8	1.000	0.00	0.000	−0.02
218	0.150	−12.4	0.512	0.00	0.000	−0.02
245	0.090	−1.9	0.038	−0.10	0.000	−0.01
271	0.001	−8.7	0.002	−0.19	0.000	−0.02
272	0.004	−5.7	0.322	−0.04	0.000	−0.01
273	0.196	−5.0	0.005	0.10	0.000	−0.03
276	0.000	−6.3	0.560	0.00	0.000	−0.01
278	0.003	−22.9	0.041	0.05	0.000	−0.01
280	0.164	−10.6	0.105	0.10	0.000	−0.01
281	0.118	−13.6	1.000	0.00	0.007	−0.04

obtained from the SK tests for three constituents, i.e. specific conductance, pH, and ammonia plus organic nitrogen, are shown in Table 4 for illustrative purposes. In Table 4, if the *P* value is less than or equal to 0.05, then the corresponding value of Sen's slope is statistically significant. Otherwise, it is statistically insignificant and no trend could be detected. The test results for all constituents at each and every station are summarized in Table 5 in which the detected trends are represented by arrows with an upward arrow for upward trend and a downward arrow for a downward trend. An asterisk designates no detected trend.

Table 4 shows that for the specific conductance at a significance level of 0.05, Stations 77, 214, 215, 271, 272, 276, and 278 all exhibit downward trends decreasing at a rate of about 6 units per year or larger. Only Station 107 shows a positive Sen's slope (Table 4). This value, however, is not statistically significant at the level of 0.05. The other seven stations show no statistically significant trends at $\alpha = 0.05$.

For the pH value, Stations 105, 245, and 271 show statistically significant downward trends with Sen's slopes between −0.10 and −0.19. However, Stations 273 and 278 indicate upward trends with Sen's slopes equal to 0.10 and 0.05, respectively. No statistically significant trends could be detected at

Table 5
A summary of results of trend analysis for 15 stations in the Arkansas and the Neosho river basins

Sta.	Variables																
	Ca	Mg	K	Na	Cl	TH	SC	DS	pH	SU	AK	DO	N	P	SS	T	Q
77	*	*	↓	↓	↓	*	↓	↓	*	*	*	*	↓	*	↓	*	*
105	*	*	*	↓	↓	*	*	↓	↓	*	*	↓	*	*	*	*	↑
106	*	*	↓	↓	↓	*	*	*	*	*	*	*	↓	*	*	*	↑
107	↓	↑	*	↑	↑	*	*	*	*	*	*	*	*	↓	↓	*	*
214	↓	↓	↓	↓	↓	↓	↓	↓	*	↓	↓	*	↓	*	*	*	*
215	*	*	↓	↓	↓	*	↓	↓	*	*	↓	*	↓	↓	*	*	*
218	*	↑	↓	↓	↓	*	*	*	*	*	*	*	↓	*	*	*	↑
245	*	*	↓	↓	↓	*	*	↓	↓	*	*	↓	↓	*	*	*	*
271	↓	↓	↓	*	↓	↓	↓	↓	↓	↓	↓	↓	↓	*	*	*	*
272	↓	*	↓	↓	↓.	↓	↓	↓	*	*	*	*	↓	*	↓	↑	*
273	*	*	↓	↓	↓	*	*	*	↑	*	*	↑	↓	↓	*	↑	*
276	↓	*	↓	↓	↓	↓	↓	↓	*	*	↓	*	↓	*	*	↓	*
278	↓	↓	↓	↓	↓	↓	↓	↓	↑	↓	↓	*	↓	*	*	*	↑
280	*	*	↓	*	↓	*	*	*	*	*	*	*	↓	*	*	*	↑
281	*	*	↓	↓	↓	*	*	↓	*	*	*	↓	↓	↓	↓	↓	*

Symbols used in Table 5: Ca, calcium; Mg, magnesium; K, potassium; Na, sodium; Cl, chloride; TH, total hardness; SC, specific conductance; DS, dissolved solids; SU, sulfate; AK, alkalinity; DO, dissolved oxygen; P, total phosphorus; N, ammonia + organic nitrogen; SS, suspended solids; T, temperature; Q, discharge. * No detected trend. ↑ Upward trend. ↓ Downward trend.

other stations. For ammonia plus organic nitrogen, except Stations 105 and 107, all other stations shown downward trends with Sen's slopes varying from −0.01 to −0.04.

Table 5 indicates that calcium, total hardness, specific conductance, sulfate, alkalinity, ammonia plus organic nitrogen, total phosphorus, and suspended solids generally show either no trends or downward trends. Potassium, sodium and chloride show strong homogeneity in trends with the exception of no trend for sodium at Station 280. The trends are overwhelmingly downward except that Station 107 has upward trends. This station is located on an isolated branch of the Neosho River near Burlington, Kansas (Fig. 1). The estimated Sen's slopes for potassium vary within a small range of −0.06 and −0.13 while those for sodium are between −0.20 and −4.78. Sen's slopes for chloride vary widely from −0.92 to −13.68 units per year. At Station 107, a positive slope of 0.57 units per year was detected for chloride. This upward trend may be due to some localized effects.

No definitive trend pattern could be detected for magnesium, pH, dissolved oxygen and temperature. Sen's slope estimates for dissolved solids at Station 106 on the Walnut Creek and Station 278 on the Elk River below Elk City Lake are substantial with −23.83 units per year and −14.94 units per year, respectively (Yu et al., 1991).

The Kendall tests for the discharge show no detected trends at most stations. However, upward trends at Stations 105, 106, 207, 218, 278, and 281 are noted. The estimated Sen's slopes (Yu, et al., 1991) at Station 105 on the Verdigris River, Stations 106 and 278 on the Walnut, and Station 218 on the Arkansas River at Arkansas City range from 0.0374 to 1.430 c.m.s. year^{-1}. The maximum value for discharge occurs at Station 218 on the Arkansas River. These upward trends might be due to the effects of upstream reservoir releases. The regulated releases would certainly affect the distribution of discharge over the period of data analysis. Therefore, flow adjustment of constituent concentration was not included (Hirsch et al., 1991).

Station-wide and basin-wide trend tests

The trends or the lack of trends shown in Table 5 generally indicate the heterogeneous nature of trends of water quality constituents with the exception of a group of ions. To establish formally the station-wide and basin-wide trend non-homogeneity, it is necessary to conduct some statistical tests by using the Van Belle and Hughes test at a significance level of 0.05. Data of 10 selected constituents collected at 10 stations from December 1975 to November 1989 were used to test the homogeneity of station-wide trend and basin-wide trends. Ten variables, namely, alkalinity, ammonia plus organic nitrogen, chloride, dissolved solids, dissolved oxygen, potassium, specific conductance, sulfate, total hardness, and total phosphorus. The results are tabulated elsewhere (Yu et al., 1991) and are not listed here.

The station-wide trend tests include homogeneity test of trends of individual constituents at a given station. The basin-wide trend tests include homogeneity tests for seasonal, station, station–season trends for a specific water quality constituent at all stations. In general, the trends (or lack of trends) of constituents at a given station are different and the trend of a given water quality constituent also varies from station to station. At the significance level of 0.05, the Van Belle and Hughes test indicates neither homogeneity of station-wide trends for all constituents nor homogeneity of basin-wide trends of a particular constituent.

Conclusions

The results of non-parametric trend analysis for data collected at 15 fixed stations for 17 water quality constituents lead to the following general conclusions.

The four different non-parametric methods have practically the same power at a statistical significance level of 0.05 for the record length equal to or greater than 9 years.

With the exception of a small group of constituents characterizing the ionic strength of water, the detected trends of other constituents at the same station are non-homogeneous. Likewise, the trends of a given constituent vary from station to station throughout the river basins.

A general pattern of downward trends of the major ions and dissolved solids, two nutrients, and suspended solids is detected. This conspicuous pattern apparently indicates that the overall water quality of these rivers has been improving. The improving water quality probably reflects the effects of the general reduction of non-point sources of pollution in recent years.

Acknowledgment

This paper was supported financially in part by the Department of Interior, US Geological Survey, through the Kansas Water Resources Research Institute, Project No. 14-08-0001-G1563.

References

Berryman, D., Bobee, B., Cluis, D. and Haemmerli, J., 1988. Nonparametric tests for trend detection in water quality time series. Water Resour. Bull., 24(3): 545–556.

Conover, W.J., 1980. Practical Nonparametric Statistics. John Wiley & Sons, New York, 229 pp.

Farrel, R., 1980. Methods for Classifying Changes in Environmental Conditions, Tech. Rep. VRF-EPA7. 4-FR80-1. Vector Research Inc., Ann Arbor, MI.

Fleiss, J.L., 1981. Statistical Methods for Rates and Proportions, 2nd edn., John Wiley & Sons, New York.

Gilbert, R.O., 1987. Statistical Methods for Environmental Pollution Monitoring. Van Nostrand Reinhold, New York.

Hirsch, R.M., 1988. Statistical methods and sampling design for estimating step trends in surface-water quality. Water Resour. Bull., 24(3): 493–503.

Hirsch, R.M. and Slack, J.R., 1984. A nonparametric trend test for seasonal data with serial dependence. Water Resour. Res., 20(6): 727–732.

Hirsch, R.M., Slack, J.R. and Smith, R.A., 1982. Techniques of trend analysis for monthly water quality data. Water Resour. Res., 18(1): 107–121.

Hirsch, R.M., Alley, W.M. and Wilber, W.G., 1988. Concepts for a National Water-Quality Assessment Program. U.S., Geol. Surv., Circ. No. 1021.

Hirsch, R.M., Alexander, R.B. and Smith, R.A., 1991. Selection of methods for the detection and estimation of trends in water quality. Water Resour. Res., 27(5): 803–813.

Jordan, P.R. and J.K. Stamer, 1991. Surface Water-Quality Assessment of the Lower Kansas River Basin, Kansas and Nebraska: Analysis of Available Data through 1986. US Geological Survey, Water-Supply Paper 2352-B (Provisional Draft).

Kendall, M.G., 1975. Rank Correlation Methods, 4th edn. Charles Griffin, London, 6 pp.

Lettenmaier, D.P., 1976. Detection of trends in water quality data from records with dependent observations. Water Resour. Res., 12(5): 1037–1046.

Lettenmaier, D.P., Hooper, E.R., Wagoner, C. and Fans, K.B., 1991. Trends in stream quality in the continental United States, 1978–1987. Water Resour. Res., 27(3): 327–339.

Loftis, J.C., Taylor, C.H. and Chapman, P.L., 1991. Multivariate tests for trend in water quality. Water Resour. Res., 27(7): 1419–1429.

Mann, H.B., 1945. Non-parametric tests against trend. Econometrica, 33: 245–259.

Norusis, M.J., 1988. SPSS, User's Manuals. SPSS.

Sen, P.K., 1968a. On a Class of Aligned Rank Order Tests in Two-Way Layouts, Ann. Math. Statist., No. 39, pp. 1115–1124.

Sen, P.K., 1968b. Estimates of the Regression Coefficient Based on Kendall's tau, J. Am. Stat. Assoc., No. 67, pp. 1379–1389.

Taylor, C.H. and Loftis, J.C., 1989. Testing for trend in lake and ground water quality time series. Water Resour. Bull., 25(4): 715–726.

Van Belle, G. and Hughes, J.P., 1984. Nonparametric tests for trend in water quality. Water Resour. Res., 20(1): 127–136.

Yu, Y.-S., Zou, S. and Whittemore, D.O., 1991. Trend analysis of surface-water quality of the upper and lower Arkansas River basin in Kansas. Contribution No. 291, Kansas Water Resources Research Institute.