Environmetrics; 1992; 3(4): 385–411

On Kendall's Tau as a Test of Trend in Time Series Data

A. H. El-Shaarawi[1] and *Stefan P. Niculescu*[2]

Abstract

Kendall's tau is often used to detect the presence of trend in environmental time series data. Assuming stationarity, a general expression for the variance of the associated S score is derived. The results are specialized to the cases of MA(1) and MA(2). Asymptotic normality of tau is established. It is shown that the variance of tau is strongly affected by the assumption of statistical dependence. A seasonal model with non-zero correlations between successive seasons and years is considered. Simulation results indicate that the effects of departures from distributional assumptions are less important when compared to the effects of departures from the independence assumption.

KEY WORDS: Kendall's tau; S score; MA(1); MA(2); water quality.

1. Introduction

Methods of nonparametric trend analysis such as those based on Kendall's coefficient of rank correlation tau are widely used to test for the presence of monotonic trends in environmental time series data. Tau was proposed independently by Mann (1945) and Theil (1950) as a simple easy-to-compute distribution-free trend test for a sequence of independent observations. Ferguson (1965) suggested combining a number of individual trend tests into a single test for detecting the overall trend in a number of independent data sets. This was used later by Hirsch, Slack and Smith

[1,2] Canada Centre for Inland Waters, National Water Research Institute, Rivers Research Branch, 867 Lakeshore Road, P.O. Box 5050, Burlington, Canada L7R 4A6.

1180–4009/92/040385–27$18.50

Received 10 September 1992
Revised 1 December 1992

(1982) to test for trends in seasonal data and the test was named Seasonal Kendall's Tau (SKT). Since it is hard to imagine that the assumption of independence is valid for the majority of seasonal time series data, Zetterqvist (1988) considered the case of m-autocorrelated seasonal time series with m smaller than half the number of seasons. In all of the above, it was always assumed that the observations within each season are independent and only autodependence occurs among data in different seasons. This paper extends the previous works on Kendall's tau in two directions. The first deals with the case of nonseasonal data with significant lag 1 and lag 2 [MA(1),MA(2)] autocorrelations, and the second considers seasonal data with non-zero correlations between successive seasons and successive years. In these two cases, test statistics are derived and their performances are evaluated by simulations. The results indicate that the proposed tests are almost nonparametric once the autocorrelations are accounted for in the test. An application concludes the paper.

2. A MOVING AVERAGE EXTENSION OF THE NON-SEASONAL MODEL

Let x_i be the ith observation $(i = 1, \ldots, n)$ and consider the model

$$x_i = \alpha + \beta i + y_i \ , \qquad i = 1, \ldots, n \tag{1}$$

where α and β are unknown parameters and y_i is a stationary time series with autocorrelation function $\rho(t) = \mathrm{Corr}(y_{t+s}, y_s)$ for all $t, s \in Z = \{0, \pm 1, \pm 2, \ldots\}$. Consider the hypothesis $H_0 : \beta = 0$ of no trend. The most convenient technique used to test H_0 is to compute Theil's statistic

$$C(n) = \frac{\tau(n)}{[\mathrm{Var}\,\tau(n)]^{1/2}} = \frac{S(n)}{[\mathrm{Var}\,S(n)]^{1/2}} \tag{2}$$

where

$$S(n) = \frac{1}{2} n(n-1)\,\tau(n) = \sum_{1 \le j < i \le n} U_{ij} \ ,$$

$U_{ij} = \mathrm{sgn}(x_i - x_j)$, and $\mathrm{sgn}(\Theta) = -1$ for $\Theta < 0$; 0 for $\Theta = 0$, and 1 for $\Theta > 0$. If the y_i's are independent with a common continuous distribution function,

$$\mathrm{Var}[S(n)] = \frac{1}{18} n(n-1)(2n+5) = V_0(n) \tag{3}$$

and the test based on $C(n)$ is the classical Kendall's Tau Test (KT). To study the behaviour of $C(n)$ for dependent y_i's, a general expression for $\mathrm{Var}[S(n)]$ is required. We shall assume x_0 has a continuous distribution function in order to avoid non-zero probability ties. Because of the symmetry of the distribution of $y_i - y_j$ all the odd moments of S are zeros. We deduce

$$\mathrm{Var}[S(n)] = \sum_{\substack{1\le j<i\le n\\ 1\le k<l\le n}} \mathrm{Cov}(U_{ij}, U_{lk}) = \sum_{\substack{1\le j<i\le n\\ 1\le k<l\le n}} E(U_{ij}U_{lk}) \tag{4}$$

where E refers to expectation, and for all $i, j, l, k \in \{1, \ldots, n\}$, $j < i$, $k < l$:

$$E(U_{ij}U_{lk}) = 4\,\mathrm{pr}(\gamma_{ijlk}) - 1 \tag{5}$$

where

$$\gamma_{ijlk} = \{X_i - X_j > 0,\ X_l - X_k > 0\}\ .$$

For reason of convenience we shall restrict our investigation to the case in which $(y_t)_{t\in Z}$ is an MA(1) or MA(2) process.

2.1 The MA(1) Case

Let $\{a_i,\ i \in Z\}$ be a sequence of i.i.d. random variables with zero mean and unit variance. The associated invertible MA(1) processes are of the form:

$$y_t = a_t + \Theta a_{t-1}\ , \qquad t \in Z \tag{6}$$

where $\Theta \in (-1, 1)$. The autocorrelation function of $\{y_t,\ t \in Z\}$ is

$$\rho(0) = 1\ , \qquad \rho(\pm 1) = \frac{\Theta}{1+\Theta^2}\ , \qquad \rho(s) = 0\ , \qquad s \ge 2\ .$$

Suppose $i, j, l, k \in \{1, \ldots, n\}$ are arbitrarily fixed such that $j < i$ and $k < l$. Let $\Im_1 = \{i-1, i\} \cup \{j-1, j\} \cup \{l-1, l\} \cup \{k-1, k\} = \{i_1, \ldots, i_m\}$, say, where $m = \mathrm{Card}\ \Im_1$ and $i_s < i_t$ for all $1 \le s < t \le m$. Consider the transform:

$$\begin{aligned}\Psi : \boldsymbol{u} = (u_{i_1}, \ldots, u_{i_m}) \to \boldsymbol{v} = (&u_i - x_j + \Theta(u_{i-1} - u_{j-1}), u_l - u_k \\ &+\Theta(u_{l-1} - u_{k-1}), u_{i_3}, \ldots, u_{i_m})\ .\end{aligned} \tag{8}$$

It follows that,

$$f_{(y_i - y_j, y_l - y_k, a_{i_3}, \ldots, a_{i_m})}(\boldsymbol{v}) = |\mathrm{J}_\Psi|\ f_{(a_{i_1}, \ldots, a_{i_m})}(\Psi^{-1}(\boldsymbol{v}))\ , \qquad \boldsymbol{v} \in \mathbf{R}^m \tag{9}$$

where $|\mathrm{J}_\Psi|$ is the determinant of the Jacobian of the transform Ψ and f, indexed by a random vector, stands for the probability density function of that vector. Consequently

$$\begin{aligned}&f_{(y_i - y_j, y_l - y_k)}(u, v) \\ &= \int_{\mathbf{R}^{m-2}} |\mathrm{J}_\Psi|\, f_{(a_{i_1}, \ldots, a_{i_m})}\left(\Psi^{-1}(u, v, w_{i_3}, \ldots, w_{i_m})\right) dw_{i_3} \ldots dw_{i_m}\end{aligned} \tag{10}$$

which could be used to evaluate (5) and to derive $\mathrm{Var}[S(n)]$ after solving the associated combinatorial problems involved in (4). Under the assumptions on the model the following holds true

$$\mathrm{Var}[S(2)] = 1 \tag{11}$$
$$\mathrm{Var}[S(3)] = 3 + 4K_1 + 2K_7 \tag{12}$$
$$\mathrm{Var}[S(4)] = 6 + 8K_1 + 4K_2 + 2K_3 + 2K_6 + 4K_7 + 2K_8 \tag{13}$$
$$\mathrm{Var}[S(5)] = \tfrac{32}{3} + 12K_1 + 12K_2 + 4K_3 + 4K_4 + 4K_6 + 6K_7 + 4K_8 + 4K_9 \tag{14}$$

and for $n \geq 6$

$$\begin{aligned}\mathrm{Var}[S(n)] = {} & \tfrac{1}{18}(2n^3 - 9n^2 + 43n - 48) + 4(n-2)K_1 + 2(n-3)(n-2)K_2 \\ & + 2(n-3)(K_3 + K_6 + K_8) + 2(n-4)(n-3)(K_4 + K_9) \\ & + \tfrac{2}{3}(n-5)(n-4)(n-3)K_5 + 2(n-2)K_7 \ .\end{aligned} \tag{15}$$

The constants K_1 to K_9 are defined as follows

$$\begin{array}{lll} K_1 = 4\,\mathrm{pr}(\gamma_{2131}) - 1 & K_2 = 4\,\mathrm{pr}(\gamma_{3141}) - 1 & K_3 = 4\,\mathrm{pr}(\gamma_{4231}) - 1 \\ K_4 = 4\,\mathrm{pr}(\gamma_{5241}) - 1 & K_5 = 4\,\mathrm{pr}(\gamma_{6241}) - 1 & K_6 = 4\,\mathrm{pr}(\gamma_{3241}) - 1 \\ K_7 = 4\,\mathrm{pr}(\gamma_{3221}) - 1 & K_8 = 4\,\mathrm{pr}(\gamma_{4321}) - 1 & K_9 = 4\,\mathrm{pr}(\gamma_{5231}) - 1 \ . \end{array}$$

Let us consider for example the case $n \geq 6$. The relative position of all indices $i, j, l, k \in \{1, \ldots, n\}$ such that $j < i$ and $k < l$ is described by the sets

$$\begin{array}{lll} A = \{1 \leq j = k < i = l \leq n\} & E = \{1 \leq k < l = j < i \leq n\} & J = \{1 \leq j < k < l < i \leq n\} \\ B = \{1 \leq j = k < i < l \leq n\} & F = \{1 \leq k < l < j < i \leq n\} & L = \{1 \leq j = k < l < i \leq n\} \\ C = \{1 \leq k < j < l < i \leq n\} & G = \{1 \leq j < k < i = l \leq n\} & K = \{1 \leq k < j < i = l \leq n\} \\ D = \{1 \leq k < j < i < l \leq n\} & H = \{1 \leq j < k < i < l \leq n\} & M = \{1 \leq j < i = k < l \leq n\} \\ & N = \{1 \leq j < i < k < l \leq n\} & \end{array}$$

Hence (4) can be expressed as

$$\mathrm{Var}[S(n)] = \sum_A + \cdots + \sum_N \ .$$

From the stationarity assumption it follows that,

$$\sum_B = \sum_G = \sum_L = \sum_K , \quad \sum_C = \sum_H , \quad \sum_D = \sum_J , \quad \sum_E = \sum_M , \quad \sum_F = \sum_N ,$$

and consequently that,

$$\mathrm{Var}[S(n)] = \sum_A + 4\sum_B + 2\left(\sum_C + \sum_D + \sum_E + \sum_F\right) \ . \tag{16}$$

In our case,

$$\sum_A = \tfrac{1}{2}n(n-1) \ , \tag{17}$$

$$\sum_B = (n-2)K_1 + \tfrac{1}{2}(n-3)(n-2)(K_2 - K_{10}) + \tfrac{1}{18}(n-4)(n-3)(n-2) \ , \tag{18}$$

$$\begin{aligned}\sum_C = &(n-3)K_3 + \tfrac{1}{2}(n-4)(n-3)K_4 + \tfrac{1}{6}(n-5)(n-4)(n-3)K_5 \\ &+ (n-4)(n-3)K_9 \ ,\end{aligned} \tag{19}$$

$$\begin{aligned}\sum_D = &(n-3)K_6 + \tfrac{1}{2}(n-4)(n-3)K_4 + \tfrac{1}{3}(n-5)(n-4)(n-3)K_5 \\ &+ (n-4)(n-3)K_{11} \ ,\end{aligned} \tag{20}$$

$$\sum_E = (n-2)K_7 + (n-3)(n-2)K_{10} - \tfrac{1}{18}(n-4)(n-3)(n-2) \ , \tag{21}$$

$$\sum_F = (n-3)K_8 - \tfrac{1}{6}(n-5)(n-4)(n-3)K_5 - (n-4)(n-3)K_{11} \ , \tag{22}$$

where $K_{10} = 4\,\mathrm{pr}(\gamma_{4221}) - 1$ and $K_{11} = 4\,\mathrm{pr}(\gamma_{3251}) - 1$. Hence (15) follows by summing up (17) to (22). Constants K_{10} and K_{11} do not play any role in the final result. For completeness we illustrate the derivation of (17) to (22) by considering the typical case of (18). In order to capture the interactions between variables let us decompose B into four disjoint subsets B_1, B_2, B_3 and B_4 where

$$\begin{aligned} B_1 &= \{1 \le j = k < i = j+1 < l = j+2 \le n\} \\ B_2 &= \{1 \le j = k < i = j+1 < j+3 \le l \le n\} \\ B_3 &= \{1 \le j = k < j+2 \le i < l = i+1 \le n\} \\ B_4 &= \{1 \le j = k < j+2 \le i < i+2 \le l \le n\} \ . \end{aligned}$$

Observe that

$$\begin{aligned} 4\,\mathrm{pr}(\gamma_{ijlk}\,|\,(i,j,l,k) \in B_1) - 1 &= K_1 \\ 4\,\mathrm{pr}(\gamma_{ijlk}\,|\,(i,j,l,k) \in B_2) - 1 &= -K_{10} \\ 4\,\mathrm{pr}(\gamma_{ijlk}\,|\,(i,j,l,k) \in B_3) - 1 &= K_2 \\ 4\,\mathrm{pr}(\gamma_{ijlk}\,|\,(i,j,l,k) \in B_4) - 1 &= 1/3 \end{aligned} \tag{23}$$

and

$$\begin{aligned} \mathrm{Card}\, B_1 &= n-2 \\ \mathrm{Card}\, B_2 = \mathrm{Card}\, B_3 &= \tfrac{1}{2}(n-2)(n-3) \\ \mathrm{Card}\, B_4 &= \tfrac{1}{6}(n-2)(n-3)(n-4) \ . \end{aligned} \tag{24}$$

Now (18) follows by combining (23) and (24). Expressions (19), (20) and (22) accommodate four distinct indices. In order to capture the interactions between variables each of the sets C, D and F will be decomposed into eight disjoint subsets by analogy with the case of B and the proof is based on the same idea as the proof of (18). Result (17) follows immediately from the properties of the sgn function. A different proof of (15) based on mathematical induction is presented in the Appendix.

REMARK 1: If we replace Θ by zero in the MA(1) process we obtain $K_1 = K_2 = 1/3$, $K_7 = -1/3$, and $K_3 = K_4 = K_5 = K_6 = K_8 = K_9 = 0$. Consequently (15) reduces to (3), the variance when the x's are independent.

EXAMPLE 1 — MA(1) PROCESS WITH $a_0 \sim \mathsf{N}(0,1)$: Under this assumption it follows :

$$X_l - X_k \sim \mathsf{N}(0, 2[1-\rho(l-k)](1+\Theta^2)) \tag{25}$$

for all $l, k \in \{1, \ldots, n\}$, $k < l$. It is well known (see Cramér 1971, page 290) that:

$$\begin{aligned} \operatorname{pr}(\gamma_{ijlk}) &= \frac{1}{4} + \frac{1}{2\pi} \arcsin[\operatorname{Corr}(y_i - y_j, y_l - y_k)] \\ &= \frac{1}{4} + \frac{1}{2\pi} \arcsin \frac{\rho(i-l) + \rho(j-k) - \rho(j-l) - \rho(i-k)}{2\sqrt{[1-\rho(i-j)][1-\rho(l-k)]}} \end{aligned} \tag{26}$$

for all $i, j, l, k \in \{1, \ldots, n\}$, $j < i$, $k < l$. Consequently:

$$\begin{aligned} K_1 &= \frac{2}{\pi} \arcsin \frac{1}{2} \sqrt{\frac{1+\Theta^2}{1-\Theta+\Theta^2}} \\ K_2 &= \frac{2}{\pi} \arcsin \frac{1+\Theta+\Theta^2}{2(1+\Theta^2)} \\ K_3 &= K_5 = \frac{2}{\pi} \arcsin \frac{\Theta}{2(1+\Theta^2)} \\ K_4 &= \frac{2}{\pi} \arcsin \frac{\Theta}{1+\Theta^2} \\ K_6 &= \frac{2}{\pi} \arcsin \frac{\Theta}{\sqrt{(1+\Theta^2)(1-\Theta+\Theta^2)}} \\ K_7 &= -\frac{2}{\pi} \arcsin \frac{(1-\Theta)^2}{2(1-\Theta+\Theta^2)} \\ K_8 &= -\frac{2}{\pi} \arcsin \frac{\Theta}{2(1-\Theta+\Theta^2)} \\ K_9 &= 0 \ . \end{aligned}$$

The detailed equations of Var[$S(n)$] will result by replacing these values in (12) to (15).

REMARK 2: Consider the proof of Theorem 2.1 in Zetterqvist (1988) where the assumption $m < [(b-1)/2]$ is replaced by $m = b = 1$ and n is large enough to ensure $n^{1-q} > 2$ for a given $(1/2) < q < 1$. The main blocks of the proof remain valid and consequently as $n \to \infty$:

$$C(n) \Longrightarrow \mathsf{N}(0,1) \ . \tag{27}$$

Via Slutsky's Theorem (Theorem 5.2.5 in Chow and Teicher 1978) this is equivalent to

$$\left(\frac{1}{9} + \frac{2}{3}K_5\right)^{-1/2} n^{-3/2} S(n) \Rightarrow \mathsf{N}(0,1) \ . \tag{28}$$

as $n \to \infty$. For a given significance level $\alpha \in (0,1)$ this leads to the following asymptotically equivalent MA(1)-extensions of the Kendall's Tau Test:

T_1: with significance level α reject H_0 if

$$|C(n)| > \lambda_{\alpha/2} \ ;$$

and

T_2: with significance level α reject H_0 if

$$\left|\tfrac{1}{9} + \tfrac{2}{3}K_5\right|^{-1/2} n^{-3/2} \, |S(n)| > \lambda_{\alpha/2} \ ,$$

where $\phi(\lambda_{\alpha/2}) = 1 - \alpha/2$ and where ϕ denotes the standard Gaussian cumulative distribution function.

REMARK 3: As demonstrated by equations (12)–(15) without the assumptions of independence Var[$S(n)$] is dependent on the distribution of the noise and the nature of the statistical dependence between successive values. To estimate the effects of departure from independence on Var[$S(n)$], Monte Carlo simulation experiments were conducted to estimate the coefficients K_1 to K_9 of equations (12)–(15) for Normal and non-Normal data. For pre-specified values of Θ and noise distributions (Normal, Laplace and Cauchy), 500 time series each of 1000 values were generated. The coefficients were estimated by averaging their values over the 500 runs and these were inserted into equation (15) to produce the estimates of Var[$S(n)$] for the MA(1) process. The ratios Var[$S(n)$]$/V_0(n)$ were computed for samples of size 10, 15, 20, 25, and 30 from various distributions. The results are presented in Table 1. Included in this table are the exact ratios for the Normal distribution which demonstrate the accuracy of the simulation. The results

		θ	−0.99	−0.75	−0.50	−0.25	0	0.25	0.50	0.75	0.99
$n = 10$	N(0,1)	T	0.372	0.397	0.499	0.705	1.000	1.312	1.550	1.623	1.639
	N(0,1)	S	0.373	0.398	0.500	0.704	1.003	1.317	1.556	1.677	1.710
	Laplace	S	0.378	0.394	0.470	0.662	1.001	1.378	1.614	1.711	1.717
	Cauchy	S	0.388	0.393	0.441	0.573	1.006	1.529	1.713	1.769	1.729
$n = 15$	N(0,1)	T	0.271	0.300	0.419	0.660	1.000	1.352	1.612	1.701	1.720
	N(0,1)	S	0.270	0.298	0.418	0.659	1.004	1.357	1.618	1.745	1.781
	Laplace	S	0.282	0.304	0.392	0.617	1.002	1.414	1.661	1.768	1.780
	Cauchy	S	0.300	0.305	0.363	0.513	1.007	1.553	1.741	1.800	1.784
$n = 20$	N(0,1)	T	0.216	0.248	0.376	0.636	1.000	1.374	1.647	1.747	1.770
	N(0,1)	S	0.215	0.244	0.374	0.634	1.005	1.379	1.653	1.784	1.822
	Laplace	S	0.231	0.255	0.349	0.592	1.002	1.436	1.688	1.801	1.814
	Cauchy	S	0.253	0.257	0.320	0.479	1.007	1.568	1.759	1.819	1.813
$n = 25$	N(0,1)	T	0.182	0.215	0.349	0.621	1.000	1.388	1.669	1.778	1.803
	N(0,1)	S	0.180	0.210	0.346	0.618	1.005	1.393	1.676	1.809	1.850
	Laplace	S	0.199	0.225	0.321	0.577	1.002	1.450	1.705	1.823	1.836
	Cauchy	S	0.224	0.227	0.294	0.457	1.007	1.578	1.773	1.833	1.831
$n = 30$	N(0,1)	T	0.159	0.193	0.331	0.610	1.000	1.397	1.684	1.799	1.826
	N(0,1)	S	0.157	0.186	0.326	0.608	1.006	1.403	1.692	1.827	1.869
	Laplace	S	0.177	0.204	0.302	0.566	1.001	1.459	1.716	1.839	1.851
	Cauchy	S	0.204	0.206	0.276	0.442	1.007	1.585	1.782	1.842	1.843

Table 1. MA(1) time series: values of the ratio $\mathrm{Var}[S(n)]/V_0(n)$ (T = Theory, S = Simulation).

show that the use of $V_0(n)$ will badly over or under estimate $\mathrm{Var}[S(n)]$ when the assumption of independence is violated. The effects depend on the nature of dependence and the distribution of the noise. As can be seen the type of distribution used to generate the data will have very little impact on the results as compared to that of ignoring the dependence among successive values.

REMARK 4: Both tests T_1 and T_2 are large sample tests and consequently before using them we have to investigate how far the left hand terms in (27) and (28) are from their limiting distributions. To answer this question a simulation experiment was performed. For a pre-specified distribution for the noise and pre-specified values of n and Θ, 500 sample data sets $\{y_1, \ldots, y_n\}$ of the form (6) were generated and the associated $S(n)$ scores were computed. They were then standardized by dividing them by the square root of the theoretical variance of S given by equation (15). The Q–Q plot of the resulting sample data was drawn in order to compare them with $\mathrm{N}(0,1)$.

Figure 1a–b gives this plot when the noise is $\mathsf{N}(0,1)$ and $\Theta = 0.5$ for $n = 10$ and respectively $n = 20$. For the same values of Θ and n, the results when the noise variables were generated from the Laplace and Cauchy distributions are illustrated in Figure 1c–d, respectively 1e–f. The results show that T_1 can be safely used for $n \geq 20$. A similar study shows that T_2 can be safely used for $n \geq 50$.

REMARK 5: To investigate the power of T_1, assume $\alpha = 0.05$ and the distribution of the noise is $\mathsf{N}(0,1)$. For $n = 20$ and each Θ in the set $\{-0.99, -0.75, -0.5, -0.25, 0, 0.25, 0.5, 0.75, 0.99\}$ the following experiment was performed. 500 sample data sets of the form $\{y_1, \ldots, y_n\}$ were generated by Monte Carlo simulation using model (6). For prescribed values of β each of the sets $\{x_1, \ldots, x_n\}$ associated with $\{y_1, \ldots, y_n\}$ was tested for trend using T_1. Its empirical power is given by the proportion of cases in which the presence of the trend was accepted. Table 2 summarizes the results. The values obtained by using the classical Kendall's Tau Test were included in order to detect the effect of neglecting the dependence between successive values. The results indicate that T_1 preserves the significance level and almost maintains the same power as Kendall's Tau Test.

2.2 THE MA(2) CASE

Let $\{a_i,\ i \in Z\}$ be a sequence of i.i.d. random variables with zero mean and unit variance. The associated invertible MA(2) process is of the form

$$X_t = a_t + \Theta_1 a_{t-1} + \Theta_2 a_{t-2} \ , \qquad t \in Z \tag{29}$$

where

$$\Theta_1 + \Theta_2 > -1 \ ; \qquad \Theta_1 - \Theta_2 < 1 \ ; \qquad -1 < \Theta_2 < 1 \ . \tag{31}$$

The corresponding autocorrelation function is given by

$$\begin{aligned} \rho(0) &= 1 \\ \rho(\pm 1) &= \frac{\Theta_1(1+\Theta_2)}{1+\Theta_1^2+\Theta_2^2} \\ \rho(\pm 2) &= \frac{\Theta_2}{1+\Theta_1^2+\Theta_2^2} \\ \rho(\pm s) &= 0 \ , \qquad s \geq 3 \ . \end{aligned} \tag{32}$$

Assume $i, j, l, k \in \{1, \ldots, n\}$ to be arbitrarily fixed and such that $j < i$ and $k < l$. Let $\Im_2 = \Im_1 \cup \{i-2, j-2, l-2, k-2\} = \{i_1, \ldots, i_m\}$ say, where

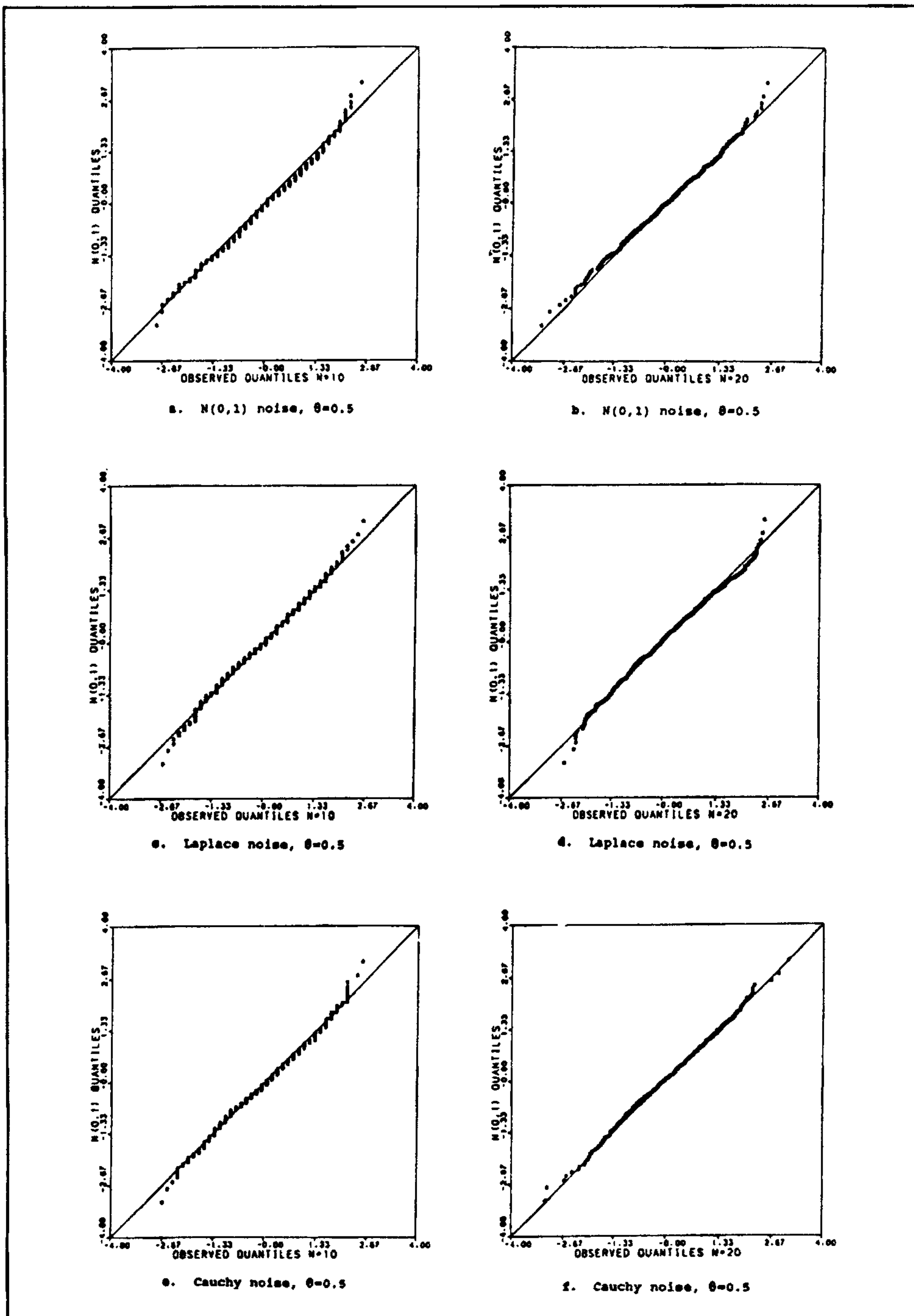

N(0,1) QUANTILES
OBSERVED QUANTILES N=10
a. N(0,1) noise, θ=0.5
N(0,1) QUANTILES
OBSERVED QUANTILES N=20
b. N(0,1) noise, θ=0.5
N(0,1) QUANTILES
OBSERVED QUANTILES N=10
c. Laplace noise, θ=0.5
N(0,1) QUANTILES
OBSERVED QUANTILES N=20
d. Laplace noise, θ=0.5
N(0,1) QUANTILES
OBSERVED QUANTILES N=10
e. Cauchy noise, θ=0.5
N(0,1) QUANTILES
OBSERVED QUANTILES N=20
f. Cauchy noise, θ=0.5

Figure 1.

β:	−0.25		−0.2		−0.15		−0.1	
$\Theta \downarrow$	T_1	KT	T_1	KT	T_1	KT	T_1	KT
−0.99	1.000	0.996	1.000	0.972	0.996	0.706	0.926	0.228
−0.75	1.000	1.000	1.000	0.988	1.000	0.882	0.962	0.400
−0.50	1.000	1.000	1.000	1.000	1.000	0.946	0.928	0.546
−0.25	1.000	1.000	1.000	0.998	0.986	0.948	0.802	0.610
0	1.000	1.000	0.996	0.996	0.938	0.938	0.598	0.598
0.25	0.996	1.000	0.976	0.992	0.798	0.898	0.458	0.602
0.50	0.970	0.994	0.894	0.966	0.638	0.806	0.360	0.552
0.75	0.888	0.976	0.774	0.908	0.564	0.760	0.298	0.490
0.99	0.854	0.952	0.670	0.830	0.448	0.672	0.212	0.428

β:	−0.05		0		−0.05	
$\Theta \downarrow$	T_1	KT	T_1	KT	T_1	KT
−0.99	0.456	0.014	0.048	0	0.414	0.012
−0.75	0.450	0.026	0.036	0	0.460	0.028
−0.50	0.438	0.088	0.058	0.004	0.388	0.060
−0.25	0.286	0.134	0.040	0.014	0.268	0.142
0	0.240	0.240	0.056	0.056	0.220	0.220
0.25	0.138	0.226	0.040	0.098	0.164	0.232
0.50	0.120	0.244	0.036	0.118	0.156	0.198
0.75	0.098	0.262	0.044	0.134	0.090	0.242
0.99	0.090	0.230	0.040	0.136	0.094	0.220

β:	0.10		0.15		0.20		0.25	
$\Theta \downarrow$	T_1	KT	T_1	KT	T_1	KT	T_1	KT
−0.99	0.930	0.186	0.996	0.730	1.000	0.968	1.000	0.992
−0.75	0.940	0.342	0.998	0.904	1.000	0.994	1.000	0.998
−0.50	0.922	0.510	1.000	0.942	1.000	1.000	1.000	1.000
−0.25	0.778	0.572	0.978	0.936	1.000	0.994	1.000	1.000
0	0.634	0.634	0.924	0.924	0.992	0.992	1.000	1.000
0.25	0.526	0.668	0.770	0.864	0.968	0.992	0.986	0.998
0.50	0.348	0.554	0.680	0.836	0.896	0.958	0.958	0.990
0.75	0.308	0.496	0.506	0.722	0.772	0.892	0.852	0.948
0.99	0.212	0.416	0.448	0.664	0.662	0.848	0.780	0.918

Table 2. Empirical powers of T_1 and of the Kendall's Test (KT) for the non-seasonal MA(1) model described in Remark 5 (n=20, α=0.05).

$m = \operatorname{Card} \Im_2$, and $i_s < i_t$ for all $1 \le s < t \le m$. The problem of computing (5) is solved by (10) where ψ is the transform

$$\begin{aligned} \psi : u = (u_{i_1}, \dots, u_{i_m}) \to v = &(u_i - u_j + \Theta_1(u_{i-1} - u_{j-1}) \\ &+ \Theta_2(u_{i-2} - u_{j-2}), u_l - u_k \\ &+ \Theta_1(u_{l-1} - u_{k-1}) \\ &+ \Theta_2(u_{l-2} - u_{k-2}), u_{i_3}, \dots, u_{i_m}) \ . \end{aligned} \tag{33}$$

Using the procedure described in the previous section it follows that:

$$\operatorname{Var}[S(3)] = 3 + 4L_1 + 2L_7 \tag{34}$$

$$\operatorname{Var}[S(4)] = 6 + 8L_1 + 4L_2 + 2L_6 + 4L_7 + 2L_8 + 4L_{10} + 4L_{12} \tag{35}$$

$$\begin{aligned} \operatorname{Var}[S(5)] = &10 + 12L_1 + 8L_2 + 4L_3 + 2L_4 + 4L_6 + 6L_7 + 4L_8 + 4L_9 + 8L_{10} \\ &+ 4L_{11} + 8L_{12} + 4L_{14} + 4L_{17} + 4L_{25} + 2L_{35} + 2L_{36} + 2L_{37} \end{aligned} \tag{36}$$

$$\begin{aligned} \operatorname{Var}[S(6)] = &15 + 16L_1 + 12L_2 + 6L_3 + 4L_4 + 4L_5 + 6L_6 + 8L_7 + 6L_8 + 8L_9 \\ &+ 12L_{10} + 8L_{11} + 12L_{12} + 4L_{13} + 12L_{14} + 4L_{15} + 4L_{16} \\ &+ 8L_{17} + 4L_{23} + 4L_{24} + 8L_{25} + 4L_{30} + 2L_{31} + 4L_{32} + 2L_{33} \\ &+ 4L_{35} + 4L_{36} + 4L_{37} \end{aligned} \tag{37}$$

$$\begin{aligned} \operatorname{Var}[S(7)] = &\tfrac{65}{3} + 20L_1 + 16L_2 + 8L_3 + 6L_4 + 8L_5 + 8L_6 + 10L_7 + 8L_8 \\ &+ 12L_9 + 16L_{10} + 12L_{11} + 16L_{12} + 12L_{13} + 24L_{14} + 12L_{15} \\ &+ 12L_{16} + 12L_{17} + 8L_{18} + 12L_{23} + 12L_{24} + 12L_{25} + 2L_{26} + 2L_{27} \\ &+ 2L_{28} + 8L_{30} + 4L_{31} + 8L_{32} + 4L_{33} + 4L_{34} + 6L_{35} + 6L_{36} + 6L_{37} \end{aligned} \tag{38}$$

$$\begin{aligned} \operatorname{Var}[S(8)] = &\tfrac{92}{3} + 24L_1 + 20L_2 + 10L_3 + 8L_4 + 12L_5 + 10L_6 + 12L_7 \\ &+ 10L_8 + 16L_9 + 20L_{10} + 16L_{11} + 20L_{12} + 24L_{13} + 40L_{14} \\ &+ 24L_{15} + 24L_{16} + 16L_{17} + 24L_{18} + 4L_{20} + 2L_{21} + 4L_{22} + 24L_{23} \\ &+ 24L_{24} + 16L_{25} + 4L_{26} + 4L_{27} + 4L_{28} + 2L_{29} + 12L_{30} + 6L_{31} \\ &+ 12L_{32} + 6L_{33} + 6L_{34} + 8L_{35} + 8L_{36} + 8L_{37} \end{aligned} \tag{39}$$

and for $n \ge 9$

$$\begin{aligned} \operatorname{Var}[S(n)] = &\tfrac{1}{18}(2n^3 - 21n^2 + 139n - 240) + 2(n-2)(2L_1 + L_7) \\ &+ 2(n-3)(2L_2 + L_3 + L_6 + L_8 + 2L_{10} + 2L_{12}) \\ &+ 2(n-4)(L_4 + 2L_9 + 2L_{11} + 2L_{17} + 2L_{25} + L_{35} + L_{36} + L_{37}) \\ &+ 2(n-5)(2L_5 + 2L_{30} + L_{31} + 2L_{32} + L_{33} + L_{34}) \\ &+ 2(n-6)(L_{26} + L_{27} + L_{28}) \\ &+ 2(n-4)(n-3)L_{14} + 2(n-5)(n-4)(L_{13} + L_{15} + L_{16} + L_{23} + L_{24}) \\ &+ 4(n-6)(n-5)L_{18} + (n-7)(n-6)(2L_{20} + L_{21} + L_{29}) \\ &+ \tfrac{2}{3}(n-8)(n-7)(n-6)L_{19} + \tfrac{2}{3}(n-7)(n-6)(n-5)L_{22} \ . \end{aligned} \tag{40}$$

The constants L_1 to L_{37} are given by

$$L_i = 4\,\mathrm{pr}(A_i) - 1 \quad , \qquad i = 1,\ldots,37 \quad ,$$

where

$$\begin{array}{lllll}
A_1 \;= \gamma_{3121} & A_2 \;= \gamma_{4131} & A_3 \;= \gamma_{4231} & A_4 \;= \gamma_{5241} & A_5 \;= \gamma_{6241} \\
A_6 \;= \gamma_{4132} & A_7 \;= \gamma_{3221} & A_8 \;= \gamma_{4321} & A_9 \;= \gamma_{5231} & A_{10} = \gamma_{4221} \\
A_{11} = \gamma_{5132} & A_{12} = \gamma_{4121} & A_{13} = \gamma_{6131} & A_{14} = \gamma_{5141} & A_{15} = \gamma_{6141} \\
A_{16} = \gamma_{6441} & A_{17} = \gamma_{5131} & A_{18} = \gamma_{7251} & A_{19} = \gamma_{9471} & A_{20} = \gamma_{6381} \\
A_{21} = \gamma_{8461} & A_{22} = \gamma_{5281} & A_{23} = \gamma_{6231} & A_{24} = \gamma_{6251} & A_{25} = \gamma_{5431} \\
A_{26} = \gamma_{7531} & A_{27} = \gamma_{7351} & A_{28} = \gamma_{5371} & A_{29} = \gamma_{8351} & A_{30} = \gamma_{6421} \\
A_{31} = \gamma_{6341} & A_{32} = \gamma_{5361} & A_{33} = \gamma_{4361} & A_{34} = \gamma_{6431} & A_{35} = \gamma_{5421} \\
A_{36} = \gamma_{4251} & A_{37} = \gamma_{5331} \;.
\end{array}$$

REMARK 6: The case of MA(1) processes is obtained by replacing Θ_2 by zero. Then $L_i = K_i$, $i \in \{1,\ldots,10\}$, $L_{11} = -L_{25}$, $L_{12} = -L_{10}$, $L_{14} = L_2$, $L_{19} = L_{20} = L_{21} = L_{26} = L_{27} = L_{28} = L_{29} = L_{30} = L_{33} = L_{35} = 0$, $L_{13} = L_{15} = -L_{16} = L_{17} = -L_{37} = 1/3$, $L_{24} = L_{36} = K_4$, $L_{18} = L_{22} = -L_{31} = L_{32} = -L_{34} = K_5$, $L_{23} = K_9$. Consequently (34)–(40) reduce to (12)–(15). If both Θ_1 and Θ_2 are zero then (40) reduces to (3).

EXAMPLE 2 — MA(2) PROCESS WITH $a_0 \sim \mathsf{N}(0,1)$: Under this assumption

$$y_l - y_k \sim \mathsf{N}\{0, 2[1-\rho(l-k)](1+\Theta_1^2+\Theta_2^2)\} \tag{41}$$

for all $l, k \in \{1,\ldots,n\}$, $k < l$. The values of all coefficients L_1 to L_{37} can be computed using (26). The detailed equations of $\mathrm{Var}[S(n)]$ will then follow by replacing them in (34) to (40).

REMARK 7: The asymptotic normality of $C(n)$ follows using the technique in Zetterqvist (1988). This leads to the following asymptotically equivalent MA(2)-extensions of Kendall's Tau Test:

T_3: with significance level α reject H_0 if

$$|C(n)] > \lambda_{\alpha/2} \quad ;$$

and

T_4: with significance level α reject H_0 if

$$\left| \left[\frac{1}{9} + \frac{2}{3}(L_{19} + L_{22})\right]^{-1/2} n^{-3/2} S(n) \right| > \lambda_{\alpha/2} \quad ,$$

where $\phi(\lambda_{\alpha/2}) = 1 - \alpha/2$.

REMARK 8: The effects of Θ_1 and Θ_2 on $\mathrm{Var}[S(n)]$ were investigated by Monte Carlo simulation. The coefficients L_1 to L_{37} in (34)–(40) were estimated for various combinations of Θ_1 and Θ_2 by the same simulation technique used for estimating K_1 to K_9 in the MA(1) case. Table 3 gives the ratios $\mathrm{Var}[S(n)]/V_0(n)$ for samples of size 10, 15, 20, 25, and 30 . The exact ratios for the Normal distribution are included in order to demonstrate the accuracy of the simulation. As can be seen from Table 3, the type of distribution used to generate the data has less impact on the results as compared to that of ignoring dependence among successive values.

		Θ_1 =	0.50	0.50	0.50	0.50
		Θ_2 =	−0.49	0	0.50	0.99
$n = 10$	N(0,1)	T	0.965	1.550	1.980	2.046
	N(0,1)	S	0.965	1.556	2.036	2.078
	Laplace	S	0.954	1.614	2.117	2.098
	Cauchy	S	0.932	1.713	2.317	2.178
$n = 15$	N(0,1)	T	0.887	1.612	2.142	2.223
	N(0,1)	S	0.886	1.618	2.194	2.254
	Laplace	S	0.864	1.661	2.265	2.277
	Cauchy	S	0.808	1.741	2.444	2.358
$n = 20$	N(0,1)	T	0.844	1.647	2.238	2.325
	N(0,1)	S	0.843	1.653	2.284	2.352
	Laplace	S	0.816	1.688	2.346	2.374
	Cauchy	S	0.745	1.759	2.507	2.447
$n = 25$	N(0,1)	T	0.817	1.669	2.301	2.392
	N(0,1)	S	0.816	1.676	2.342	2.415
	Laplace	S	0.786	1.705	2.398	2.435
	Cauchy	S	0.709	1.773	2.545	2.501
$n = 30$	N(0,1)	T	0.798	1.684	2.345	2.439
	N(0,1)	S	0.798	1.692	2.382	2.459
	Laplace	S	0.766	1.716	2.434	2.478
	Cauchy	S	0.684	1.782	2.570	2.536

Table 3. MA(2) time series: values of the ratio $\mathrm{Var}[S(n)]/V_0(n)$ (T = Theory, S = Simulation).

REMARK 9: We repeated in the MA(2) case the experiment described in Remark 4. The Q-Q plots of the resulting sample data were drawn in order to compare them with $\mathsf{N}(0,1)$. Figure 2a–c refers to the case $n = 20$, $\Theta_1 = 0.5$, $\Theta_2 = -0.49$, and the distribution of the noise is respectively $\mathsf{N}(0,1)$, Laplace and Cauchy. In case the two parameters Θ_1 and Θ_2 have opposite signs T_3 can be used safely for $n \geq 20$. In case both parameters Θ_1 and Θ_2 have the same sign we have to double the value of n in order to safely use T_3. We recommend the use of T_4 only in case $n \geq 100$.

3. A Moving Average Extension to the Seasonal Model

Let $x_{(j-1)b+i}$ be the observation from season i year j, $i = 1,\ldots,b$ $j = 1,\ldots,n$. We will consider the following model:

$$x_{(j-1)b+i} = \alpha_i + \beta[(j-1)b+i] + y_{(j-1)b+i} \quad , \quad i = 1,\ldots,b \;\; j = 1,\ldots,n \tag{42}$$

where $\alpha_1,\ldots,\alpha_b$ and β are unknown parameters and y_t is a stationary error process such that

$$y_t = a_t + \Theta a_{t-1} + \theta a_{t-b} \quad , \qquad t \in Z \tag{43}$$

with $(a_t)_{t\in Z}$ i.i.d. $\mathsf{N}(0,1)$ random variables. For season i let us consider the associated $S_i(n)$ score

$$S_i(n) = \sum_{1\leq s<t\leq n} \operatorname{sgn}(x_{(t-1)b+i} - x_{(s-1)b+i})$$

and let

$$S(n) = \sum_{i=1}^{b} S_i(n) \quad . \tag{44}$$

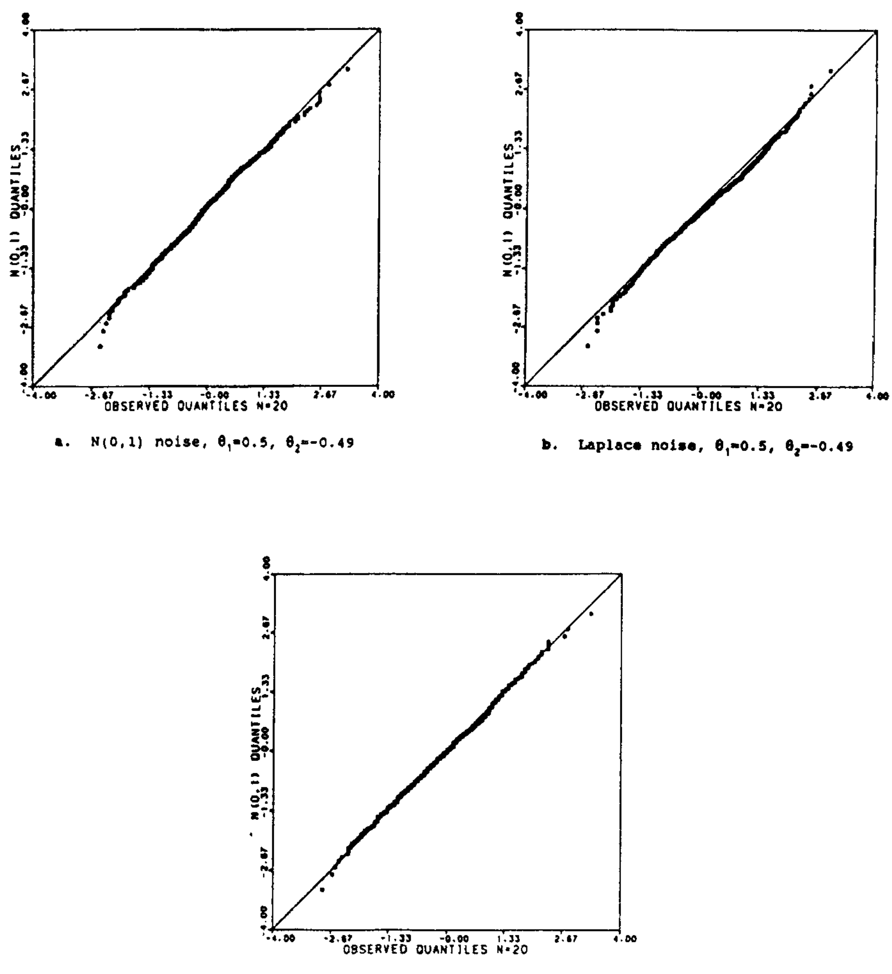
Consider the hypothesis $H_0 : \beta = 0$ of no trend. Under H_0 and the assumptions on the model,

$$S_i(n) = \sum_{1\leq s<t\leq n} \operatorname{sgn}(y_{(t-1)b+i} - y_{(s-1)b+i})$$

$$E[S(n)] = \sum_{i=1}^{b} E[S_i(n)] = 0$$

and

$$\begin{aligned}\operatorname{Var}[S(n)] = &\sum_{i=1}^{b} \operatorname{Var}[S_i(n)] + 2\left\{\sum_{i=1}^{b-1} \operatorname{Cov}[S_i(n), S_{i+1}(n)]\right.\\ &\left. + \operatorname{Cov}[S_1(n), S_b(n)]\right\} \\ =&b \operatorname{Var}[S_1(n)] + 2(b-1)\operatorname{Cov}[S_1(n), S_2(n)] + 2\operatorname{Cov}[S_1(n), S_b(n)] \quad .\end{aligned} \tag{45}$$



a. N(0,1) noise, θ_1=0.5, θ_2=-0.49

b. Laplace noise, θ_1=0.5, θ_2=-0.49

c. Cauchy noise, θ_1=0.5, θ_2=-0.49

Figure 2.

After solving the combinatorial problems involved in the computation we deduce

$$
\begin{aligned}
\mathrm{Var}[S_1(n)] = & \frac{1}{18}(2n^3 - 9n^2 + 43n - 48) \\
& + \frac{4}{3\pi}(n-3)(n^2 - 9n + 23) \arcsin \frac{\theta}{2(1+\Theta^2+\theta^2)} \\
& + \frac{4}{\pi}(n-3)(n-4) \arcsin \frac{\theta}{1+\Theta^2+\theta^2} \\
& + \frac{4}{\pi}(n-2)(n-3) \arcsin \frac{1+\theta+\Theta^2+\theta^2}{2(1+\Theta^2+\theta^2)} \\
& + \frac{8}{\pi}(n-2) \arcsin \sqrt{\frac{1+\Theta^2+\theta^2}{2[1+2\Theta^2+(1-\theta)^2+\theta^2]}} \\
& + \frac{4}{\pi}(n-2) \arcsin \frac{2\theta - 1 - \Theta^2 - \theta^2}{1+2\Theta^2+(1-\theta)^2+\theta^2} \\
& + \frac{4}{\pi}(n-3) \arcsin \frac{\theta\sqrt{2}}{\sqrt{(1+\Theta^2+\theta^2)[1+2\Theta^2+(1-\theta)^2+\theta^2]}} \\
& - \frac{4}{\pi}(n-3) \frac{\theta}{1+2\Theta^2+(1-\theta)^2+\theta^2} \quad ,
\end{aligned} \tag{46}
$$

$$
\begin{aligned}
& \mathrm{Cov}[S_1(n), S_2(n)] \\
& = \frac{2}{3\pi}(n-1)(n-2)(n-3) \arcsin \frac{\Theta}{2(1+\Theta^2+\theta^2)} \\
& \quad + \frac{2}{3\pi}(n-1)(n-3)(n-5) \arcsin \frac{\Theta\theta}{2(1+\Theta^2+\theta^2)} \\
& \quad + \frac{1}{\pi}(n-1)(n-2) \arcsin \frac{\Theta}{1+\Theta^2+\theta^2} \\
& \quad + \frac{2}{\pi}(n-2)(n-3) \arcsin \frac{\Theta(1+\theta)}{2(1+\Theta^2+\theta^2)} \\
& \quad + \frac{1}{\pi}(n-2)(n-3) \arcsin \frac{\Theta\theta}{1+\Theta^2+\theta^2} \\
& \quad - \frac{2}{\pi}(n-3) \arcsin \frac{\Theta\theta}{1+2\Theta^2+(1-\theta)^2+\theta^2} \\
& \quad + \frac{2}{\pi}(n-1) \arcsin \frac{\Theta(2-\theta)}{1+2\Theta^2+(1-\theta)^2+\theta^2} \\
& \quad - \frac{2}{\pi}(n-2) \arcsin \frac{\Theta}{1+2\Theta^2+(1-\theta)^2+\theta^2} \\
& \quad + \frac{4}{\pi}(n-2) \arcsin \frac{\Theta(1+\theta)}{\sqrt{2(1+\Theta^2+\theta^2)[1+2\Theta^2+(1-\theta)^2+\theta^2]}}
\end{aligned} \tag{47}
$$

$$+\frac{4}{\pi}(n-2)\arcsin\frac{\Theta(1-\theta)}{\sqrt{2(1+\Theta^2+\theta^2)[1+2\Theta^2+(1-\theta)^2+\theta^2]}}$$
$$+\frac{4}{\pi}(n-3)\arcsin\frac{\Theta\theta}{\sqrt{2(1+\Theta^2+\theta^2)[1+2\Theta^2+(1-\theta)^2+\theta^2]}}$$
$$+\frac{2}{\pi}(n-2)\arcsin\frac{\Theta(2\theta-1)}{1+2\Theta^2+(1-\theta)^2+\theta^2} \quad ,$$

and

$$\begin{aligned}
&\mathrm{Cov}[S_1(n), S_b(n)] \\
&=\frac{2}{3\pi}(n-1)(n-2)(n-3)\arcsin\frac{\Theta\theta}{2(1+\Theta^2+\theta^2)} \\
&\quad+\frac{2}{3\pi}(n-1)(n-3)(n-5)\arcsin\frac{\Theta}{2(1+\Theta^2+\theta^2)} \\
&\quad+\frac{1}{\pi}(n-1)(n-2)\arcsin\frac{\Theta\theta}{1+\Theta^2+\theta^2} \\
&\quad+\frac{1}{\pi}(n-2)(n-3)\arcsin\frac{\Theta}{1+\Theta^2+\theta^2} \\
&\quad+\frac{2}{\pi}(n-2)(n-3)\arcsin\frac{\Theta(1+\theta)}{2(1+\Theta^2+\theta^2)} \\
&\quad+\frac{2}{\pi}(n-1)\arcsin\frac{\Theta(2\theta-1)}{1+2\Theta^2+(1-\theta)^2+\theta^2} \\
&\quad+\frac{2}{\pi}(n-2)\arcsin\frac{\Theta(2-\theta)}{1+2\Theta^2+(1-\theta)^2+\theta^2} \qquad (48) \\
&\quad-\frac{2}{\pi}(n-2)\arcsin\frac{\Theta\theta}{1+2\Theta^2+(1-\theta)^2+\theta^2} \\
&\quad+\frac{4}{\pi}(n-2)\arcsin\frac{\Theta(\theta+1)}{\sqrt{2(1+\Theta^2+\theta^2)[1+2\Theta^2+(1-\theta)^2+\theta^2]}} \\
&\quad+\frac{4}{\pi}(n-2)\arcsin\frac{\Theta(\theta-1)}{\sqrt{2(1+\Theta^2+\theta^2)[1+2\Theta^2+(1-\theta)^2+\theta^2]}} \\
&\quad+\frac{4}{\pi}(n-3)\arcsin\frac{\Theta}{\sqrt{2(1+\Theta^2+\theta^2)[1+2\Theta^2+(1-\theta)^2+\theta^2]}} \\
&\quad-\frac{2}{\pi}(n-3)\arcsin\frac{\Theta}{1+2\Theta^2+(1-\theta)^2+\theta^2} \quad .
\end{aligned}$$

The detailed expression for $\mathrm{Var}[S(n)]$ follows by replacing (46)–(48) in (45). This can be expressed in a compressed form as:

$$\mathrm{Var}[S(n)] = b\,f(\Theta,\theta)n^3 + o(n^3) \quad ,$$

where

$$f(\Theta,\theta) = \frac{1}{9} + \frac{4}{3\pi}\left[\arcsin\frac{\Theta}{2(1+\Theta^2+\theta^2)} + \arcsin\frac{\theta}{2(1+\Theta^2+\theta^2)} + \arcsin\frac{\Theta\theta}{2(1+\Theta^2+\theta^2)}\right] .$$

REMARK 10: Now consider the proof of Theorem 2.1 in Zetterqvist (1988) where the assumption $m < [(b-1)/2]$ is replaced by $m = b$ and n is assumed to be large enough to ensure $[nb]^{1-q} > 2b$ for a given $(1/2) < q < 1$. The main block of the proof remains valid and consequently the random variable $C(n)$ associated with (44) will be asymptotically $\mathsf{N}(0,1)$. This leads to the following extension of the Seasonal Kendall's Tau Test

T_5: with significance level α reject H_0 if

$$|C(n)| > \lambda_{\alpha/2} .$$

For large n, via Slutsky's Theorem, this is equivalent to:

T_6: with significance level α reject H_0 if

$$\left|[b\,f(\Theta,\theta)]^{-1/2} n^{-3/2} S(n)\right| > \lambda_{\alpha/2} .$$

REMARK 11: For $\Theta = 0$ let $z_t = y_{(t-1)b+1} = a_{(t-1)b+1} + \theta a_{(t-2)b+1}, t \in Z$. The process $(z_t)_{t\in Z}$ will be MA(1). Let us denote the associated S score by $S_z(n)$. From (46) we deduce:

$$\begin{aligned}\mathrm{Var}[S_z(n)] =& \frac{1}{18}(2n^3 - 9n^2 + 43n - 48) \\ &+ \frac{4}{3\pi}(n-3)(n^2 - 9n + 23)\arcsin\frac{\theta}{2(1+\theta^2)} \\ &+ \frac{4}{\pi}(n-4)(n-3)\arcsin\frac{\theta}{1+\theta^2} \\ &+ \frac{4}{\pi}(n-3)(n-2)\arcsin\frac{1+\theta+\theta^2}{2(1+\theta^2)} \\ &+ \frac{8}{\pi}(n-2)\arcsin\frac{1}{2}\sqrt{\frac{1+\theta^2}{1-\theta+\theta^2}} \\ &- \frac{4}{\pi}(n-2)\arcsin\frac{(1-\theta)^2}{1(1-\theta+\theta^2)} \\ &+ \frac{4}{\pi}(n-3)\arcsin\frac{\theta}{\sqrt{(1+\theta^2)(1-\theta+\theta^2)}} \\ &- \frac{4}{\pi}(n-3)\arcsin\frac{\theta}{2(1-\theta+\theta^2)} .\end{aligned} \quad (49)$$

REMARK 12: For $\Theta = 0$, from (43) it follows $\mathrm{Cov}[S_1(n), S_2(n)] = 0$ as expected from the independence of $S_1(n)$ and $S_2(n)$. For $\Theta = \theta = 0$, $(z_t)_{t\in Z}$ is a sequence of i.i.d. random variables and (49) reduces to (3). When $\theta = 0$, (47) and (48) give:

$$\begin{aligned}\mathrm{Cov}[S_1(n), S_2(n)] =& \frac{1}{\pi} n(n-1) \arcsin \frac{\Theta}{1+\Theta^2} \\ &+ \frac{2}{3\pi}(n-3)(n-2)(n-1) \arcsin \frac{\Theta}{2(1+\Theta^2)}\end{aligned}$$

and

$$\begin{aligned}\mathrm{Cov}[S_1(n), S_b(n)] =& \frac{1}{\pi}(n-1)(n-2) \arcsin \frac{\Theta}{1+\Theta^2} \\ &+ \frac{2}{3\pi}(n-1)(n^2-8n+9) \arcsin \frac{\Theta}{2(1+\Theta^2)} \quad ,\end{aligned}$$

which agree with Zetterqvist (1988) when $m = 1$.

REMARK 13: The test T_5 converges faster to the $\mathsf{N}(0,1)$ than T_6. To illustrate this phenomenon a simulation experiment was performed. For prespecified Θ and θ, $b = 12$ and $n = 12$, 500 sample data sets $\{y_1, \ldots, y_{nb}\}$ using (43) were generated and the associated S scores were computed. Figure 3a–b contains their Q-Q plots for $\Theta = 0.3$, $\theta = -0.5$, and $\Theta = 0.3$, $\theta = 0.5$, respectively. The speed of convergence is faster when Θ and θ have opposite signs. Figure 4a–c refers to the statistic $[bf(\Theta,\theta)]^{-1/2} n^{-3/2} S$ for $\Theta = 0.3$, $\theta = -0.5$ and $n = 12$, 50, 100. For $n \geq 50$ the distribution of T_6 is well approximated by the $\mathsf{N}(0,1)$. The empirical power of T_5 was also investigated using the methods in Remark 5 and the results are summarized in Table 4. The empirical power of the Seasonal Kendall's Tau Test was included in the table to illustrate the effects of neglecting the dependence between successive values.

REMARK 14: Let us consider the model (42) in the general case of $(a_t)_{t\in Z}$ i.i.d., i.e. without the $\mathsf{N}(0,1)$ restriction on the noise. As we seen in Remark 10, $C(n)$ is asymptotically $\mathsf{N}(0,1)$. For $s \in \{1,2,b\}$ let

$$\begin{aligned}g_s =& \tfrac{1}{6}(\delta^s_{3151} + \delta^s_{5131} + \delta^s_{5153} + \delta^s_{5351} + \delta^s_{5331} + \delta^s_{3153} + \delta^s_{6241} + \delta^s_{4162} \\ &+ \delta^s_{6341} + \delta^s_{4163} + \delta^s_{5163} + \delta^s_{6351} + \delta^s_{4261} + \delta^s_{6142} + \delta^s_{4361} + \delta^s_{6143} \\ &+ \delta^s_{5361} + \delta^s_{6153} + \delta^s_{6421} + \delta^s_{2164} + \delta^s_{6431} + \delta^s_{3164} + \delta^s_{6531} + \delta^s_{3165}) \quad ,\end{aligned}$$

where

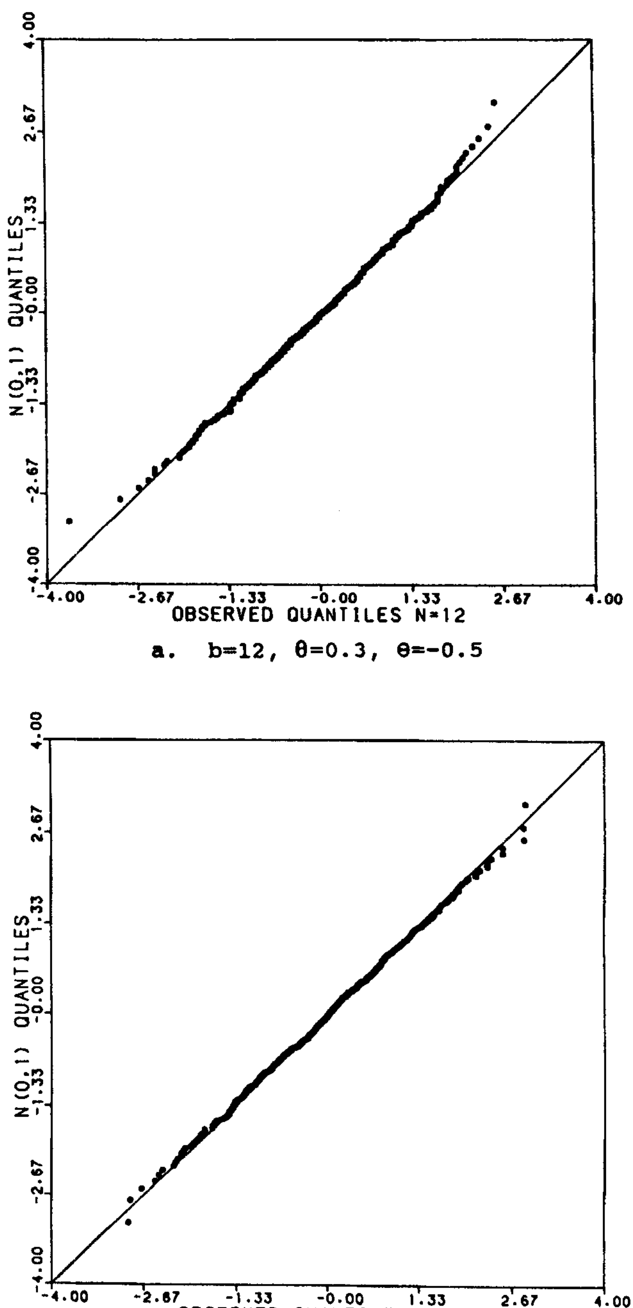
$$\begin{aligned}\delta^1_{ijlk} =& 4\,\mathrm{pr}\left[y_{(i-j)b+1} - y_{(j-1)b+1} > 0, y_{(l-1)b+1} - y_{(k-1)b+1} > 0\right] - 1 \\ \delta^2_{ijlk} =& 4\,\mathrm{pr}\left[y_{(i-j)b+1} - y_{(j-1)b+1} > 0, y_{(l-1)b+2} - y_{(k-1)b+2} > 0\right] - 1 \\ \delta^b_{ijlk} =& 4\,\mathrm{pr}\left[y_{(i-1)b+1} - y_{(j-1)b+1} > 0, y_{lb} - y_{kb} > 0\right] - 1 \quad .\end{aligned}$$

β:	−0.05		−0.04		−0.03		−0.02	
$\Theta \downarrow$	T_5	SKT	T_5	SKT	T_5	SKT	T_5	SKT
−0.99	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.960
−0.75	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.998
−0.50	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.982
−0.25	1.000	1.000	1.000	1.000	1.000	1.000	0.988	0.962
0	1.000	1.000	1.000	1.000	1.000	1.000	0.936	0.936
0.25	1.000	1.000	1.000	1.000	0.976	0.990	0.818	0.904
0.50	1.000	1.000	0.998	1.000	0.936	0.978	0.664	0.824
0.75	1.000	1.000	0.988	0.996	0.876	0.952	0.528	0.742
0.99	0.990	0.988	0.944	0.984	0.780	0.904	0.428	0.656

β:	−0.01		0		0.01	
$\Theta \downarrow$	T_5	SKT	T_5	SKT	T_5	SKT
−0.99	0.996	0.200	0.046	0.000	1.000	0.400
−0.75	0.990	0.048	0.032	0.000	0.996	0.056
−0.50	0.882	0.226	0.050	0.000	0.854	0.019
−0.25	0.600	0.316	0.058	0.008	0.612	0.382
0	0.416	0.416	0.044	0.044	0.452	0.452
0.25	0.306	0.420	0.056	0.104	0.316	0.438
0.50	0.214	0.366	0.052	0.170	0.240	0.402
0.75	0.182	0.364	0.048	0.138	0.174	0.350
0.99	0.156	0.304	0.066	0.196	0.142	0.306

β:	0.02		0.03		0.04		0.05	
$\Theta \downarrow$	T_5	SKT	T_5	SKT	T_5	SKT	T_5	SKT
−0.99	1.000	0.970	1.000	1.000	1.000	1.000	1.000	1.000
−0.75	1.000	0.988	1.000	1.000	1.000	1.000	1.000	1.000
−0.50	1.000	0.986	1.000	1.000	1.000	1.000	1.000	1.000
−0.25	0.998	0.974	1.000	1.000	1.000	1.000	1.000	1.000
0	0.948	0.948	1.000	1.000	1.000	1.000	1.000	1.000
0.25	0.796	0.872	0.980	0.988	1.000	1.000	1.000	1.000
0.50	0.672	0.812	0.952	0.976	1.000	1.000	1.000	1.000
0.75	0.568	0.738	0.866	0.962	0.976	0.998	0.998	1.000
0.99	0.394	0.634	0.808	0.926	0.960	0.994	0.996	1.000

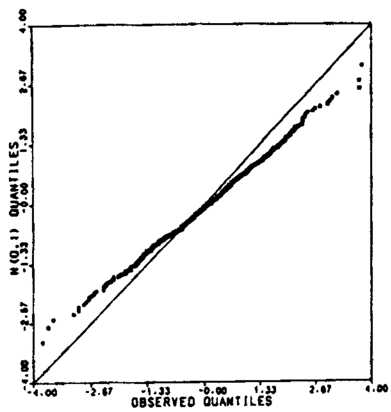
Table 4. Empirical powers of T_5 and of the Seasonal Kendall's Test (SKT) for the model in Remark 13 (Θ=0.3397, θ=0, b=12, n=12, α=0.05).



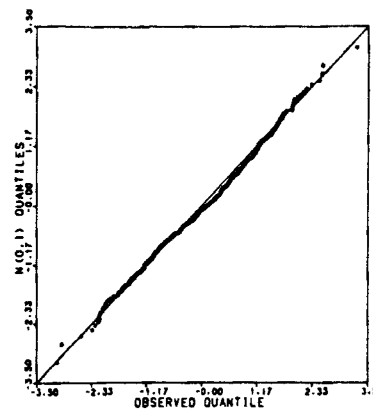
a. b=12, θ=0.3, Θ=-0.5

b. b=12, θ=0.3, Θ=0.5

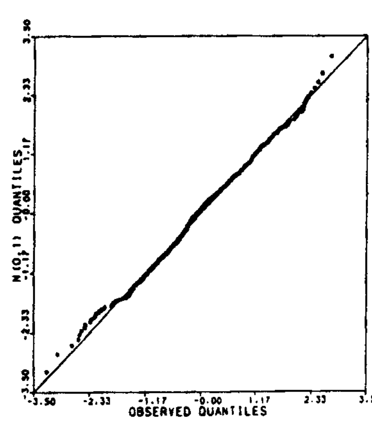
Figure 3.



a. n=12, b=12, θ=0.3, θ=-0.5



b. n=50, b=12, θ=0.3, θ=-0.5



c. n=100, b=12, θ=0.3, θ=-0.5

Figure 4.

It follows that:

$$\mathrm{Var}[S(n)] = n^3\,[bg_1 + 2(b-1)g_2 + 2g_b] + o(n^3) \ .$$

Under the assumption that the distribution of the noise is known, all the δ's can be computed using Monte Carlo simulation. For a given significance level $\alpha \in (0,1)$ and large n this leads to the following extension of T_6:

T_7: with significance level α reject H_0 if

$$\left|[b\,g(\Theta,\theta)]^{-1/2}\, n^{-3/2} S(n)\right| > \lambda_{\alpha/2} \ ,$$

where

$$g(\Theta,\theta) = bg_1 + 2(b-1)g_2 + 2g_b \ .$$

The performance of T_7 as compared with T_5 can be investigated using the methods in Remark 13.

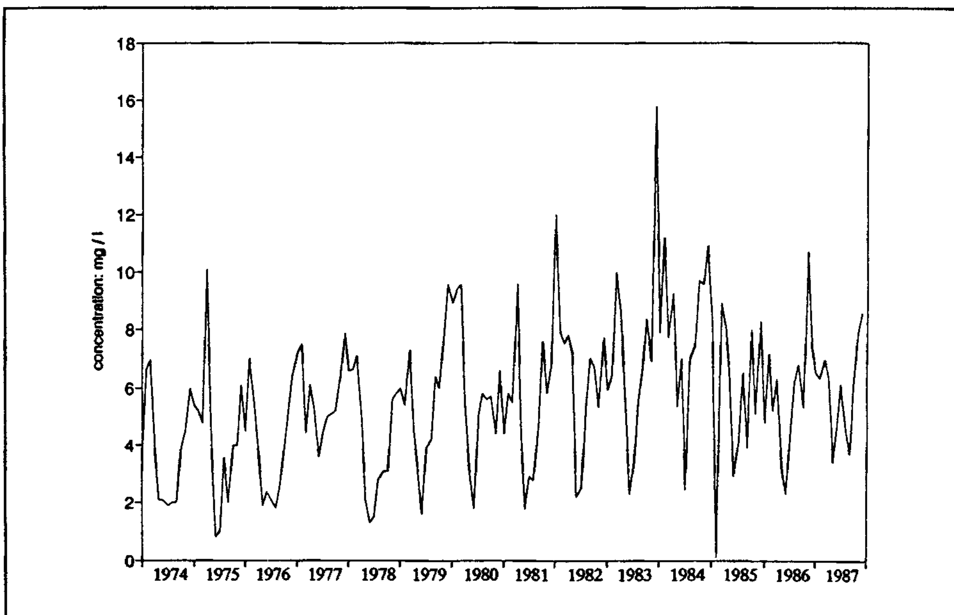


Figure 5. Dissolved chloride in the South Saskatchewan River at location AK0001.

4. A Case Study

We consider, as an example, the dissolved Chloride monthly data (Figure 5) for years 1974 to 1987 at location AK0001 on the South Saskatchewan River. First, the seasonal effect was removed by subtracting the seasonal mean. This residual series was modelled adequately by the MA(2) model (29) with $\Theta_1 = 0.3034$, $\Theta_2 = 0.1851$, and Gaussian noise. The overall observed value of Kendall's S score is $S(168) = 4384$. From equation (40) $\mathrm{Var}[S(168)] = 1005.087^2$. It follows that $C(168) = 4.3618$, and using T_3 we accept the presence of the trend at significance level 0.05. If we neglect the dependence structure of data, from (3) we obtain $\mathrm{Var}[S(168)] = 729.044^2$. Hence $C(n) = 6.013$ and, using Kendall's Tau Test, we reject H_0 at the same significance level. We cannot rely on the Kendall's Test for Trend unless the independence of the data is satisfied. Note that the presence of the trend was also recognized by polynomial regression. The next step was to group the data according to the seasons. Note that the 14 values available for each season are independent and consequently equation (3) is valid for computing the variance of S. We obtained

$$\begin{array}{llllll} S_1 = -2 & S_2 = 14 & S_3 = 27 & S_4 = 22 & S_5 = 17 & S_6 = 41 \\ S_7 = 42 & S_8 = 48 & S_9 = 42 & S_{10} = 48 & S_{11} = 41 & S_{12} = 46 \ . \end{array}$$

Hence

$$\begin{array}{llll} C_1 = -0.109 & C_2 = 0.766 & C_3 = 1.478 & C_4 = 1.204 \\ C_5 = 0.930 & C_6 = 2.244 & C_7 = 2.299 & C_8 = 2.627 \\ C_9 = 2.299 & C_{10} = 2.267 & C_{11} = 2.244 & C_{12} = 2.518 \ . \end{array}$$

The above values show the absence of trend for the first five months while the remaining seven months show an increasing trend at significance level 0.05. These results agree with the conclusion that an overall increasing trend is present in the Chloride data. If we ignore the dependence between seasons in the model (42) then, from (44), it follows that $S(n) = 386$ and consequently that $C(n) = 6.1$. The Seasonal Kendall'S Tau Test for Trend only apparently confirms the presence of a seasonal common trend. The result is misleading because of the glaring violation of the basic assumptions upon which this test was based.

REFERENCES

Box, G.E.P. and G.M. Jenkins (1970). *Time Series Analysis: Forecasting and Control.* Holden-Day, San Francisco.

Chow, Y.S. and H. Teicher (1978). *Probability Theory: Independence, Interchangeability, Martingales.* Springer-Verlag, New York.

Cramér, H. (1971). *Mathematical Methods of Statistics.* Princeton Univ. Press, Princeton N.J.

Dietz, E.J. and T.J. Killeen (1981), "A nonparametric multivariate test for monotone trend and its pharmaceutical applications". *Journal of the American Statistial Association* **76**, 169–174.

Ferguson, G.A. (1965). *Nonparametric Trend Analysis.* McGill Univ. Press, Montreal.

Hirsch, R.M., J.R. Slack and R.A. Smith (1982), "Techniques of trend analysis for monthly quality data". *Water Resources Research* **18**, 107–121.

Hirsch, R.M. and J.R. Slack (1984). "A nonparametric trend test for seasonal data with serial dependence". *Water Resources Research* **20**, 727–732.
Kendall, M.G. (1938). "A new measure of rank correlation". *Biometrika* **30**, 81–93.
Mann, H.B. (1945), "Non-parametric test against trend". *Econometrica* **13**, 245–259.
Theil, H. (1950), "A rank invariant method of linear and polynomial regression anlysis I". *Nederl. Akad. Wetensch. Proc.* **53**, 386–392.
Zetterqvist, L. (1988), "Asymptotic distribution of Mann's test for trend for m-dependent seasonal observations". *Scandinavian Journal of Statistics* **15**, 81–95.

APPENDIX — PROOF OF (15) BY INDUCTION

Let us assume $n \geq 6$. Then

$$S(n+1) = S(n) + \sum_{j=1}^{n} U_{n+1,j} \quad ,$$

and consequently

$$\begin{aligned}\mathrm{Var}[S(n+1)] &= \mathrm{Var}[S(n)] + 2\sum_{j=1}^{n} E[S(n)U_{n+1,j}] + E\left[\left(\sum_{j=1}^{n} U_{n+1,j}\right)^2\right] \\ &= \mathrm{Var}[S(n)] + 2\Sigma_1 + \Sigma_2 \quad , \qquad \text{say.}\end{aligned} \tag{50}$$

After a tedious computation we deduce:

$$\begin{aligned}\Sigma_1 =& K_1 + (n-2)K_2 + K_3 + 2(n-3)K_4 + (n-3)(n-4)K_5 \\ &+ K_6 + K_7 + K_8 + 2(n-3)K_9 + (n-2)K_{10}\end{aligned}$$

and

$$\begin{aligned}\Sigma_2 &= \sum_{j=1}^{n} \mathrm{Var}(U_{n+1,j}) + 2\sum_{1\leq j<i\leq n} E(U_{n+1,i}U_{n+1,j}) \\ &= n + 2\sum_{1\leq j\leq n-2} E(U_{n+1,j}U_{n+1,j+1}) + 2\sum_{1\leq j<j+2\leq i\leq n-1} E(U_{n+1,i}U_{n+1,j}) \\ &\quad + 2\sum_{1\leq j\leq n-2} E(U_{n+1,n}U_{n+1,j}) + 2E(U_{n+1,n}U_{n+1,n-1}) \\ &= n + 2(n-2)K_2 + \tfrac{1}{3}(n-3)(n-2) - 2(n-2)K_{10} + 2K_1 \quad .\end{aligned}$$

The validity of (15) for $n = 6$ is routine. Let us assume (15) is valid for some $n > 6$. If we replace Σ_1 and Σ_2 in (50) we obtain:

$$\begin{aligned}\mathrm{Var}[S(n+1)] =& \tfrac{1}{18}(2n^3 - 3n^2 + 31n - 12) + 4(n-1)K_1 \\ &+ 2(n-2)(n-1)K_2 + 2(n-2)(K_3 + K_6 + K_8) \\ &+ 2(n-3)(n-2)K_4 + \tfrac{2}{3}(n-4)(n-3)(n-2)K_5 \\ &+ 2(n-1)K_7 + 2(n-3)(n-2)K_9\end{aligned}$$

which is equal to the value obtained by replacing directly n by $n+1$ in (15). The proof is concluded.