

Detection of Nonstationarity in Hydrologic Time Series Author(s): A. Ramachandra Rao and G. H. Yu Source: *Management Science*, Vol. 32, No. 9 (Sep., 1986), pp. 1206–1217 Published by: <u>INFORMS</u> Stable URL: <u>http://www.jstor.org/stable/2631546</u> Accessed: 04/11/2014 21:31

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at http://www.jstor.org/page/info/about/policies/terms.jsp

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



INFORMS is collaborating with JSTOR to digitize, preserve and extend access to Management Science.

http://www.jstor.org

# DETECTION OF NONSTATIONARITY IN HYDROLOGIC TIME SERIES\*

# A. RAMACHANDRA RAO AND G. H. YU

#### School of Civil Engineering, Purdue University, West Lafayette, Indiana 47907

Detection of changes in hydrologic time series due to intervention by man or natural causes is an important problem. Although intervention analysis has been used in the recent past to analyze nonstationary hydrologic time series, the necessity to specify a model of change and an initial time at which the time series has started to change are obvious disadvantages of intervention analysis. An alternative to intervention analysis is a method which is based on spectral characteristics and an exponential moving average model.

The basic objective of the research discussed in the present paper is to test this alternative method. The model is tested by using synthetic uncorrelated and correlated data with step and gradual changes as well as by using real hydrologic time series. The sensitivity of the model to different parameters is also explored. The alternative model is found to be quite accurate in detecting changes in hydrologic time series.

(NONSTATIONARITY; HYDROLOGY; STOCHASTIC MODELS; TIME SERIES)

### 1. Introduction

Hydrologic time series often exhibit nonstationary behavior. Variables such as flow and rainfall values stay above or below the mean for long periods. This phenomenon of high values following high values and vice versa is called persistence in hydrology and has received considerable attention. The relationship between persistence and surface water reservoir storage design is also a subject of active research in hydrology (Klemes 1974).

The statistical significance of these excursions above and below the mean has been investigated by a statistical technique called intervention analysis (Hipel et al. 1975). In intervention analysis the times at which the excursions above or below the mean commence must be specified and this introduces subjectivity into the analysis. Also, a model for the type of change must be prescribed in intervention analysis. It is obviously not possible a priori to specify whether furture changes in a hydrologic variable are gradual or abrupt or whether they are increasing or decreasing. Consequently, intervention analysis is best suited for analysis of nature of changes after the changes have occurred.

A method in which assumptions and limitations such as those mentioned above are absent would be needed to detect changes in the stochastic characteristics of time series. The objective of research presented in this paper is to investigate the initiation of excursions of hydrologic variables above or below the mean by using a method based on evolutionary spectral analysis and which is adaptive to changes in data characteristics. The basic idea of this method used in the present study is to keep track of the changes in the spectral characteristics of data. When these changes are significant, the parameters of an exponential moving average model are updated. Thus the method automatically provides information about points of significant changes in the level of the time series and at the same time updates the parameters of an exponential moving average. This method was proposed by Rao and Shapiro (1970).

Although this method is promising, several questions about the selection of parameters of the method, the sensitivity of the method in detecting changes which may vary

1206

0025-1909/86/3209/1206\$01.25 Copyright © 1986, The Institute of Management Sciences

<sup>\*</sup> Accepted by Ambar G. Rao; received July 15, 1985. This paper has been with the authors 2 months for 1 revision.

from small to large and other related aspects have not been explored. These aspects are analyzed in the present study. After finding the method to be satisfactory, it is used to analyze observed hydrologic time series. Annual rainfall, monthly particulate concentration and river flow time series are analyzed and the results are presented. The results from this method are also compared to those obtained by intervention analysis.

This paper is organized as follows. Theoretical aspects of the method are discussed in §2. The data used in the study are discussed in §3. In §4, the results of data analysis are discussed. A summary of the study and a set of conclusions are presented in §5.

### 2. Theoretical Aspects

Let us consider a sequence of random variables  $X(t_i)$ , i = 1, 2, ..., N. An exponential smoothing method proposed by Rao and Shapiro (1970) to detect points of significant change in the  $X(t_i)$  series is used in this study. In this method, the smoothing constant  $\alpha_i$  is determined by using the information contained in the changes in frequency components of successive spectral estimators. The idea behind this method is that whenever there is a significant change in a time series such a change must manifest itself in a statistically significant difference in successive spectral ordinates which in turn is reflected in changes in  $\alpha_i$ . Once the  $\alpha_i$  values are computed, the forecast  $Y(t_i)$  of  $X(t_i)$  is estimated by

$$Y(t_i) = \alpha_i X(t_i) + (1 - \alpha_i) Y(t_{i-1}).$$
(1)

Consequently, the forecasts obtained by using this procedure are adaptive. When there is an indication of significant change in the system, forecasts reflect this change.

The power spectrum of  $X(t_i)$  is estimated by using estimated autocovariances and the Hanning window with weights  $(\frac{1}{4}, \frac{1}{2}, \frac{1}{4})$  to smooth the spectral estimates (Jenkins and Watts 1969). The successive spectral ordinates are computed with a moving time window of fixed length L. For example, the spectrum of the subseries  $X_1, X_2, \ldots, X_L$ is first computed, followed by computation of spectrum of subseries  $X_2, X_3, \ldots, X_{L+1}$ , and  $X_3, X_4, \ldots, X_{L+2}$ , and so on. Therefore, for a series  $X(t_i)$ ,  $i = 1, 2, \ldots, N$  and a window of length L, there are N - L + 1 spectral estimates computed by using these subseries which overlap by L - 1 samples. For each subseries, M + 1 spectral ordinates are considered. In the following computations the logarithm of spectral estimates  $F_{i,k}$ is used instead of the spectral estimates  $f_x(k\Delta t)$  as the logarithmic transformation stabilizes the spectral estimates. Let  $F_{i,k} = \log_e [f_x(k\Delta f)]$  where  $i = 1, 2, \ldots, N - L$ + 1 and  $k = 0, 1, \ldots, M$ . The differences in the spectra of different subseries are computed by using moving averages of  $F_{i,k}$ .

Let  $\delta_{i,k}$  be the difference between L successive moving averages of  $F_{i,k}$  as shown below:

$$\delta_{i,k} = \frac{1}{L} \left( F_{i-L+1,k} + F_{i-L+2,k} + \dots + F_{i,k} \right) - F_{i,k}$$
 and (2)

$$\Delta_i = \max_k |\delta_{i,k}|. \tag{3}$$

If  $\Delta_i$  is large compared to its standard deviation, then it is concluded that the time series has significantly changed. The relationship between  $\alpha_i$  and  $\Delta_i$  is defined as follows:

$$\alpha_{i} = \text{Max} (0.1, \text{ Min} (e^{\beta_{i}} - 1, 1)) \text{ where}$$
  

$$\beta_{i} = b + c(\Delta_{i}/\sigma)^{2} \text{ and}$$

$$b + \gamma_{1}^{2}c = 0.67, \quad b + \gamma_{2}^{2}c = 0.095.$$
(4)

| Values of b and c Corresponding to Different Levels of $P(\Delta_i < x)$ |                 |                  |                  |  |  |  |  |
|--|-----------------|------------------|------------------|--|--|--|--|
|  | <i>P</i> = 0.90 | <i>P</i> = 0.95  | <i>P</i> = 0.99  |  |  |  |  |
| b<br>c   | 0.072<br>0.097  | 0.0761<br>0.0816 | 0.0816<br>0.0576 |  |  |  |  |

TABLE I

In (4),  $\sigma$  is the standard deviation of  $\delta_{ik}$ , b and c constants,  $\gamma_1$  and  $\gamma_2$  represent values of  $\Delta_i/\sigma$  which are used to detect significant changes in the time series. If changes in the time series are significant, then at those times  $\alpha_i$  is set equal to 0.95; otherwise it is set at 0.1.  $\gamma_2$  represents the values of  $\Delta_i/\sigma$  below which  $\alpha_i$  is allowed to remain equal to 0.1.  $\gamma_1$  and  $\gamma_2$  values are determined by  $P(\Delta_i < \chi)$  for specified values of  $\chi$ . Once  $\gamma_1$  and  $\gamma_2$  are determined, the constants b and c can be determined by using (4). In the present study, the three probability levels and the corresponding b and c values shown in Table 1 are used. The computational scheme is shown in Figure 1. The details of the method are found in Rao and Shapiro (1970).



FIGURE 1. Computational Scheme.

# 3. Data Used in the Study

Two types of synthetic data are used in the present study. The first type of synthetic series used in this study is made up of two parts. Each part has 50 observations and both of them are generated by using normally distributed random numbers with different means and variances. The mean levels in each part, however, are different. The second type of synthetic data is made up of correlated time series with step increases and decreases.

Three types of observed data which are suspected to be nonstationary are used in this study. The first of these is the annual precipitation time series measured at Eastport, New Bedford, Woodstock, Fitchburg, Taunton, New Haven, Moorestown

| STATION        | INTERVENTION<br>DATE | YEARS  | NO. OF<br>OBSERVATIONS | AVERAGE<br>IN.               | VARIANCE                     | SKEWNESS<br>COEFFICIENT      | MAXIMUM<br>IN. | MINIMUM<br>IN. | CHANGE<br>IN. X |
|----------------|----------------------|--|------------------------|------------------------------|------------------------------|------------------------------|----------------|----------------|-----------------|
| EASTPORT       | 1877,<br>1945        | 1874–1876<br>1877–1944<br>1945–1976<br>1874–1976 | 3<br>68<br>32<br>103   | 48.7<br>38.1<br>42.7<br>39.8 | 44.9<br>67.5<br>48.3<br>67.6 | 2.75<br>0.79<br>0.47<br>0.52 | 64.53          | 22.84          | -28<br>+11      |
| NEW BEDFORD    | 1906                 | 1814-1905<br>1906-1976<br>1814-1976              | 92<br>71<br>163        | 46.7<br>41.6<br>44.51        | 56.8<br>61.7<br>59.8         | 0.86<br>0.67<br>0.50         | 71.52          | 21.87          | -12             |
| WOODSTOCK, VT. | 1925                 | 1870-1925<br>1925-1976<br>1870-1976              | 55<br>52<br>107        | 38.0<br>40.0<br>39.0         | 20.8<br>32.2<br>27.2         | 0.68<br>-0.09<br>0.29        | 53.11          | 26.99          | +5              |
| FITCHBURG      | 1925                 | 1865-1924<br>1925-1976<br>1865-1976              | 60<br>52<br>112        | 41.25<br>44.4<br>42.7        | 49.7<br>53.6<br>54.0         | 0.45<br>-0.04<br>0.22        | 59.6           | 27.45          | 7.1             |

TABLE 2. Characteristics of Real Data Used in the Study

TABLE 2 (Continued)

| STATION      | INTERVENTION<br>DATE | YEARS  | NO. OF<br>DATA        | AVERAGE<br>IN.               | VARIANCE<br>IN. <sup>2</sup> | SKEWNESS<br>COEFFICIENT      | MAXIMUM<br>IN. | MINIMUM<br>IN. | CHANGE<br>IN Z |
|--------------|----------------------|--|-----------------------|------------------------------|------------------------------|------------------------------|----------------|----------------|----------------|
| TAUNTON      | 1923                 | 1875-1922<br>1923-1976<br>1875-1976              | 48<br>54<br>102       | 47.2<br>44.3<br>45.6         | 39.6<br>50.3<br>47.3         | 0.85<br>-0.11<br>0.16        | 67.23          | 28.81          | -6.5           |
| NEW HAVEN    | 1946                 | 1873-1945<br>1946-1976<br>1873-1976              | 73<br>31<br>104       | 46.2<br>42.2<br>45.0         | 44.8<br>42.8<br>47.6         | 0.35<br>-0.10<br>0.20        | 60.26          | 27.68          | -9.5           |
| MOORESTOWN   | 1891 .<br>1912       | 1864–1890<br>1891–1911<br>1912–1976<br>1864–1976 | 27<br>21<br>65<br>113 | 44.0<br>47.9<br>43.2<br>44.2 | 23.0<br>51.6<br>35.9<br>39.0 | 0.14<br>0.34<br>0.22<br>0.39 | 60.95          | 29.23          | +8.1<br>-1.1   |
| WEST CHESTER | 1920                 | 1860-1919<br>1920-1976<br>1860-1976              | 60<br>57<br>117       | 49.9<br>45.0<br>47.5         | 52.0<br>53.2<br>58.7         | 0.20<br>0.09<br>0.11         | 64.0           | 30.86          | -11            |

and West Chester. These were collected by Potter (1976) and used in his study of nonstationarity of these series. Detailed statistical information about the entire and segmented series are given in Table 2. The intervention dates in Table 2 are those given by Potter (1976).

The monthly averages of total suspended particulates monitored in Chicago underwent a drastic change in January 1970 when the law banning the use of high sulfur coal came into effect in Chicago. This is the second type of data series used in this study. The mean, variance and skewness of the series before the intervention date January 1970, are 155.1, 1042 and 0.36 respectively, while after the intervention date these statistics are respectively 100.7, 414.0 and 0.260. There are 168 observations in this series covering the period between January 1964 and December 1977.

The monthly flows in Colorado river in second-ft measured during January 1914 and December 1969 are the third type of series used herein. Colorado river flows has been known to exhibit nonstationary behavior. The mean, variance and skewness of these 672 observations are 16000,  $3.5 \times 10^8$  and 2.5 respectively. The change in mean observed around 1921 is considered to be significant (Yevjevich and Jeng 1969).

The change in suspended particulate series monitored in Chicago came into existence because of the law which banned the use of high sulfur coal in Chicago. The changes in Colorado river flows are due to variations in climatic conditions. These changes in climatic conditions cause annual variation in snowpack in the Rockies, which in turn cause variations in Colorado river flows. Nonstationarity in Colorado river flows has led to problems in the past about distribution of Colorado river flows for use in Colorado, Utah, Arizona and California.

Climatic change, and more specifically changes in rainfall, is of interest in management of water resources. Although hydrologists have been treating hydrologic and climatic time series as stationary, nonstationarity in hydrologic time series has been proposed as a possible explanation of the Hurst phenomenon. Potter (1976) has demonstrated the nonstationarity in rainfall series. These changes in rainfall series may be attributed to small, but significant changes in climatic conditions which produce the rainfall. If the nonstationarity in the rainfall series can be demonstrated by other methods, such as that developed by Rao and Shapiro (1970), then it would add strength to the hypothesis of nonstationarity in these time series. Consequently these precipitation series are analyzed.

# 4. Results and Discussion

### (1) Synthetic Data

All the numerical experiments are based on series with 100 observations  $X(t_i)$ , i = 1, 2, ..., 100. The first part of the series for i = 1, 2, ..., 50 is normally distributed

| Window Lengths* |       |           |        |    |    |  |
|-----------------|-------|-----------|--------|----|----|--|
|                 | Lengt | h of Time | Window | L  |    |  |
|                 | 18    | 24        | 30     | 36 | 40 |  |
|                 | 24    | 29        | 51     | 51 | 51 |  |
|                 | 48    | 51        | 80     | 86 |    |  |
|                 | 51    | 74        | 89     |    |    |  |
|                 | 68    |           |        |    |    |  |

 TABLE 3

 Break Points Indicated to be Significant for Different

 Window Lengths\*

\* Actual break point is at 51. As the window lengths increase, the number of break points which are statistically significant decreases.

with  $N(\mu_1, \sigma_1^2)$ , whereas the second part for  $i = 51, 52, \ldots, 100$  is distributed as  $N(\mu_2, \sigma_2^2)$ . The following five types of experiments are conducted by using synthetic data. (a) Changing the Length of the Window L. This study is based on series whose  $\mu_1, \sigma_1^2, \mu_2$  and  $\sigma_2^2$  values are 10.0, 9.0, 20.0 and 9.0 respectively, and the probability p and moving average length L' are 0.99 and 3 respectively. In Table 3 the results of



FIGURE 2. The First and Second Spectral Components for AVG1 = 10, AVG2 = 20, STN1 = STN2 = 3 and L = 40.

| Probability Levels (L = 36) |   |  |  |  |  |  |  |
|-----------------------------|---|--|--|--|--|--|--|
| The Point of Change         |   |  |  |  |  |  |  |
| <i>p</i> = 0.99             | 51, 86                                  |  |  |  |  |  |  |
| <i>p</i> = 0.95             | 51, 86, 87, 100                         |  |  |  |  |  |  |
| <i>p</i> = 0.90             | 50, 51, 58, 70, 86, 87, 88, 93, 99, 100 |  |  |  |  |  |  |

#### TABLE 4

Change Points Selected to be "Significant" Corresponding to Different Probability Levels (L = 36)

changing the length of time window L are shown. All these series have a step increase in mean value at  $X_{51}$ . From the results shown in Table 3, it is obvious that L = 40would give the best time window. This model indicates a significant change at  $\chi_{86}$  also because the spectra of the two successive subseries  $\{X_{50}, \ldots, X_{85}\}$  and  $\{X_{51}, \ldots, X_{86}\}$ are significantly different. The series  $\{X_{50}, \ldots, X_{85}\}$  includes the point  $X_{50}$  which belongs to the set before the change was introduced  $(N(10, 3^2))$ , while series  $\{X_{51}, \ldots, X_{86}\}$  entirely belongs to  $N(20, 3^2)$ . The first spectral component  $F_{i,0}$  for L = 40 is shown in Figure 2. The bell shape of Figure 2 shows that the first and second spectral components significantly change at  $X_{51}$  and  $X_{86}$ .

(b) Changing the Probability p. These numerical experiments are based on generated series, the first part of which is  $N(10, 3^2)$  and the second part is  $N(20, 3^2)$ . L equal to 36 is chosen. The change points computed in the series for probabilities p equal to 0.99, 0.95 and 0.90 are shown in Table 4. In interpreting the results in Table 4, it must be remembered that the higher the probability of  $\Delta_i$  exceeding a specific value, the more accurate is the result. As discussed above,  $X_{86}$  is a spurious change point and hence the only significant change in the series occurs at  $X_{51}$ . A probability level of 0.99 is preferred.

(c) The Sensitivity of the Model. The sensitivity of the model in detecting the change in mean when the means of two parts of the series are close to each other is an important aspect. This sensitivity is analyzed by using the length of time window L equal to 40, probability p equal to 0.99 and the first part of series with  $N(20, 3^2)$ . The results obtained by this study when the average of the second part of the series is changed with  $\sigma_2^2$  equal to  $3^2$  are shown in Table 5. From the results in Table 5, we can see that when  $\mu_2$  is equal to 22, 20 and 18 the method does not indicate any point at which there is a significant change. The entire series is normally distributed with  $N(20, 3^2)$  where  $\mu_2$  is equal to 20. It is obvious that there is no significant change in this case. The conclusion of this study is that the method can detect changes in time series and is not overly sensitive to these changes.

(d) Changing the Moving Average Length. A series generated by using  $N(10, 3^2)$  is chosen for the first part and another series generated by  $N(20, 3^2)$  for the second part of series is used in the present study. This study is also based on L = 36 and p = 0.99. The points of change detected by the model for different L' values are given in Table 6. The best choice of L' in this step increase case is 2. As L' is increased the

| Effects of Changing the Differences in the Men Values, $\mu_1 = 20$ |   |    |      |      |      |    |  |  |  |  |
|---|---|----|------|------|------|----|--|--|--|--|
|   | The Average of the Second Series $\mu_2 \cdot (\mu_1 = 20)$ |    |      |      |      |    |  |  |  |  |
| μ <sub>2</sub>  | 30  | 25 | 22   | 20   | 18   | 15 |  |  |  |  |
| Significant<br>Change at  | 51  | 51 | None | None | None | 51 |  |  |  |  |

TABLE 5

| 1213 |  |
|------|--|
|------|--|

|             | The Moving Average Length L' |    |    |    |    |    |  |
|-------------|------------------------------|----|----|----|----|----|--|
|             | 2                            | 3  | 4  | 5  | 6  | 7  |  |
| Significant | 51                           | 51 | 51 | 51 | 51 | 51 |  |
| Changes     |                              | 86 | 71 | 71 | 71 | 71 |  |
| Are Noted   |                              |    | 86 | 86 | 86 | 86 |  |
| At These    |                              |    | 87 | 87 | 87 | 87 |  |
| Points      |                              |    |    | 88 | 88 | 88 |  |

 TABLE 6

 Effects of Changing the Moving Average Length L'

\* As L' increases from 2 to 7, the number of points at which significant changes are noted also increases. There is only one significant change in the series at 51.

effects of averaging also increase. For a given spectrum the longer the length of moving average used in the model, larger would be the number of points at which the significant changes are noted. Therefore, it is important to keep L' as small as possible.

(e) Numerical Experiments with Correlated Data Series. These synthetic series consist of first 50 data points which obey the autoregressive model  $y_{t1} = \phi_1 y_{t1-1} + \eta_{t1}$  and the next 50 data points which obey autoregressive model  $y_{t2} = \phi_1 y_{t2-1} + \eta_{t2}$ .  $\eta_{t1}$  and  $\eta_{t2}$  are normally distributed with N(AVG1, STN1) and N(AVG2, STN2), respectively. The values of parameters L, L', M and p are set to be 36, 3, 6 and 0.99, respectively. For data with increasing step change, 72 cases are studied. First of all, AVG1 is set to be 10, and AVG2 is set to be 10, 12, 14, 16, 18 and 20, respectively. The standard deviations (STN1, STN2) are set at (2.0, 2.0), (2.0, 1.0) and (1.0, 1.5), while AR(1) coefficients  $\phi_1$  are set at 0.4, 0.2, -0.2 and -0.4, respectively. The significance of changes is tested in each case and the results are given in Table 7 for (STN1, STN2) = (2.0, 2.0), in Table 8 for (STN1, STN2) = (2.0, 1.0), and in Table 9 for (STN1, STN2) = (1.0, 1.5).

The results given in Tables 7–9 indicate that the method can successfully detect the significant changes in correlated data series with step increases. The higher the AVG2 is, over AVG1, more accurately can the model detect the changes. However, the spurious change at t = 86 is also detected sometimes due to the length of the window L = 36 which is used in the present study. Also, according to this primary study, the results of detecting change do not depend very much on  $\phi_1$ , STN1 and STN2. As expected, in general, the method does not detect any change when AVG1 = AVG2.

|      | Points of Change |                |                 |                 |  |  |  |  |  |  |
|------|------------------|----------------|-----------------|-----------------|--|--|--|--|--|--|
| AVG2 | $\phi_1 = 0.4$   | $\phi_1 = 0.2$ | $\phi_1 = -0.2$ | $\phi_1 = -0.4$ |  |  |  |  |  |  |
| 10   | None             | 85             | None            | None            |  |  |  |  |  |  |
| 12   | 51, 64           | 51, 64         | 64, 65          | 62, 65          |  |  |  |  |  |  |
| 14   | 51               | 51             | None            | 65              |  |  |  |  |  |  |
| 16   | 51, 86           | 51, 86         | 51, 86          | 86              |  |  |  |  |  |  |
| 18   | 51,86            | 51,86          | 51, 86          | 51, 86          |  |  |  |  |  |  |
| 20   | 51, 86           | 51,86          | 51, 86          | 51, 86          |  |  |  |  |  |  |

TABLE 7Experiments with Correlated Series with Step Increase. Actual Change PointIs at t = 51. AVG1 = 10, (STN1, STN2) = (2.0, 2.0)

|   | Points of Change |                |                |                 |                 |  |  |  |  |
|---|------------------|----------------|----------------|-----------------|-----------------|--|--|--|--|
|   | AVG2             | $\phi_1 = 0.4$ | $\phi_1 = 0.2$ | $\phi_1 = -0.2$ | $\phi_1 = -0.4$ |  |  |  |  |
|   | 10               | None           | None           | None            | None            |  |  |  |  |
|   | 12               | 51, 80         | 80             | 65, 80          | 65, 80, 85      |  |  |  |  |
| I | 14               | 51, 86         | 51, 86         | 86              | 65, 86          |  |  |  |  |
|   | 16               | 51, 86         | 51, 86         | 51, 86          | 86              |  |  |  |  |
|   | 18               | 51, 86         | 51, 86         | 51, 86          | 51, 86          |  |  |  |  |
| I | 20               | 51,86          | 51, 86         | 51, 86          | 51, 86          |  |  |  |  |

TABLE 8

Experiments with Correlated Series with Step Increase. Actual Change Point Is at t = 51. AVG1 = 10, (STN1, STN2) = (2.0, 1.0)

#### (2) Application to Observed Data

(a) Annual Precipitation Series. These series are classified as "nonhomogeneous" or "probably nonhomogeneous" by Potter (1976), who analyzed them by using intervention analysis. In this study, L' = 3 and p = 0.99 are used. The results from the present analysis are summarized in Table 10. Based on these results, two stations (Woodstock, Vt., and New Haven) have only one, four stations (Eastport, New Bedford, Fitchburg and Moorestown) have two and two stations (Taunton and West Chester) have three significant change points in them. Potter concluded that two stations (Eastport and Moorestown) have two significant change points in the series, and the remaining stations have only one significant change point.

(b) Monthly Chicago Suspended Particle Data. Rao and Padmanabhan (1983) used intervention analysis and the one-month ahead forecast errors to test the significance of changes and found that the change in the series in the affected period is statistically significant. Saavedra-Cuadra and Rao (1983, 1984) also found that the law banning the use of high sulfur coal has definitely altered the series, although the effects of intervention were first felt after a lag of two years.

In the present study, the model with L = 60, p = 0.99 and L' = 3 is used. The dates of significant change detected by the model are February, March 1970, January 1973 and May 1975. The present model indicates the date of the first significant change to be February 1970, instead of January 1970 when the law banning use of high sulfur coal came into effect. This is because the particulate concentration did not decrease immediately after the law burning high sulfur coal came into effect. However, the model detects significant changes in the next two months. As Rao and Padmanabhan (1983) have noted, the series is seasonal with an increasing and a decreasing trend

|                                  | Points of Change                                   |  |  |  |  |  |  |  |  |  |
|----------------------------------|--|--|--|--|--|--|--|--|--|--|
| AVG2                             | $\phi_1 = 0.4$                                     | $\phi_1 = 0.2$                                   | $\phi_1 = -0.2$                                    | $\phi_1 = -0.4$  |  |  |  |  |  |  |
| 10<br>12<br>14<br>16<br>18<br>20 | None<br>51<br>51, 86<br>51, 86<br>51, 86<br>51, 86 | 85<br>51<br>51, 86<br>51, 86<br>51, 86<br>51, 86 | None<br>65<br>51, 86<br>51, 86<br>51, 86<br>51, 86 | None<br>64, 65<br>51, 86<br>51, 86<br>51, 86<br>51, 86 |  |  |  |  |  |  |

TABLE 9Experiments with Correlated Series with Step Increase. Actual Change PointIs at t = 51. AVG1 = 10, (STN1, STN2) = (1.0, 1.5).

|             | Eastport | New Bedford | Woodstock, VT. | Fitchburg    |  |  |  |
|-------------|----------|-------------|----------------|--------------|--|--|--|
| L           | 18       | 36          | 36             | 36           |  |  |  |
| Significant | 1950     | 1898        | 1951           | 1915         |  |  |  |
| Changes at  | 1960     | 1904        |                | 1961         |  |  |  |
| c .         | 1945*    | 1906*       | 1925*          | 1925*        |  |  |  |
|             | 1877*    |             |                |              |  |  |  |
|             | Taunton  | New Haven   | Moorestown     | West Chester |  |  |  |
| L           | 30       | 24          | 18             | 23           |  |  |  |
| Significant | 1929     | 1945        | 1902           | 1889         |  |  |  |
| Changes at  | 1931     |             | 1963           | 1890         |  |  |  |
| -           | 1971     |             |                | 1971         |  |  |  |
|             | 1923*    | 1946*       | 1891*, 1942*   | 1920*        |  |  |  |

TABLE 10 Years at Which Significant Changes in Rainfall are Detected

\* Dates at which Potter (1976) considered significant changes have occurred are marked with asterisks.

before intervention, and the seasonality is completely changed after January 1970. The reasons for the model detecting significant changes in particulate concentration at January 1973 and May 1955 are not clear. The series after January 1970 is not as stable as it was before January 1970, especially in its seasonal characteristics and this is probably why changes were detected on January 1973 and May 1955. At any rate, the ban on using the high sulfur coal has very strongly changed the suspended particulate concentration in Chicago.

(c) Monthly Mean Discharge of Colorado River at Compact Point. This series is nonhomogeneous and is studied by Yevjevich and Jeng (1969). In the present study, a model with L = 200, p = 0.99 and L' = 3 is used. The dates at which the series changed significantly are given in Table 11. The results in Table 11 indicate that all the significant changes occur in May which is a "wet" month due to snow melt in the Rockies or in September which is a "dry" month. Most of these changes occur in May, and there are no significant changes indicated between 1950 and 1954. Yevjevich and Jeng (1969) concluded that depletion continued slowly during 1930-1950 and rapidly during 1954–1957. These results in Table 11 are in good agreement with the conclusions of Yevjevich and Jeng (1969).

# 5. Summary and Conclusions

The method proposed by Rao and Shapiro (1970) can successfully detect changes in time series. However, the parameters used for detecting changes must be carefully

| Dates of Significant Changes |            |          |            |
|------------------------------|------------|----------|------------|
| May 1931                     | May 1933   | May 1937 | Sept. 1938 |
| May 1940                     | Sept. 1943 | May 1945 | May 1946   |
| May 1949                     | Sept. 1949 | May 1955 | May 1958   |
| May 1961<br>May 1968         | May 1962   | May 1964 | May 1966   |

TABLE 11



FIGURE 3. Monthly Discharges of Colorado River at Compact Point. Triangles Indicate Intervention Dates of Significant Changes Detected by the Present Study.

selected. The parameters which must be carefully chosen in this method are the length of the time window L, the probability of exceeding the observed value p and the moving average length L'. The effects of changes in time series are clearly reflected in the structure of the first and second spectral components, and result in higher  $\alpha$  values. Based on the results presented herein, the following conclusions may be arrived at:

1. The length of time window L must be long enough so that one can get an unbiased spectrum, but not so long that changes are lost in averaging.

2. More accurate results are obtained by choosing a higher p value.

3. For a series with sharp step changes, choosing a small moving average length L' is much better than choosing a large L' value.

4. The computed values of  $\alpha_i$  will not change very much when the standard deviation values STN1 and STN2 are changed for specified AVG1 and AVG2 values.

5. Changes in correlated data are also successfully detected by this method.

The following conclusions are presented from the analysis of real data sequences in this study.

1. The annual precipitation sequences measured at Eastport, New Bedford, Woodstock Vt., Fitching, Taunton, New Haven, Moorestown and West Chester are nonhomogeneous (nonstationary) series.

2. If the number of observations before change is too small, as in the case of Eastport precipitation data, then the significance of the change cannot be established. This is because the spectra are smoothed with a time window.

3. In general, if the change is greater than about 10%, then the method can successfully detect it.

4. The changes in the series are detected after a slight lag. For example, the changes in particulate concentrations in Chicago are detected with a lag of one month.

5. The monthly Colorado River series has stable spectral components. However, the present method successfully detects changes in this time series.

#### References

HIPEL, K. W., W. C. LENNOX, T. E. UNNY AND A. I. MELEOD, "Intervention Analysis in Water Resources," Water Resource Res., 11, 6 (1975), 885–861.

JENKINS, G. M. AND D. G. WATTS, Spectral Analysis and its Applications, Holden-Day, San Francisco, 1969.

KLEMES, V., "The Hurst Phenomenon: A Puzzle?," Water Resource Res., 10, 4 (1974), 675-688.

POTTER, K. W., "Comment on the Hurst Phenomenon: A Puzzle? by V. Klemes," *Water Resource Res.*, 11, 2 (1975), 373–374.

——, "Evidence for Nonstationarity as Physical Explanation of the Hurst Phenomenon," *Water Resource Res.*, 12, 5 (1976), 1049–1052.

RAO, A. R. AND G. PADMANABHAN, "Intervention Analysis of Total Suspended Particulate Data from Chicago," *Atmospheric Environment*, 17, 3 (1983), 563-572.

RAO, A. G. AND A. SHAPIRO, "Adaptive Smoothing Using Evolutionary Spectra," Management Sci., 17, 3 (November 1970), 208–218.

SAAVEDRA-CUADRA, I. I. AND A. R. RAO, "An Investigation of Changes in Particulate Concentrations in Chicago," Tech. Rept., CE-HSE-83-7, Sch. of Civil Eng., Purdue Univ., 1983.

----- AND -----, "An Investigation of Changes in Air Particulate Concentrations in Chicago," J. Environmental Systems, 14 (1984), 269–290.

YEVJEVICH, V. AND R. I. JENG, "Properties of Non-Homogeneous Hydrologic Series," *Hydrology Paper* 32, Colorado State Univ., Fort Collins, Col., 1969.