Dr. Ashok N. Shahane Ph.D.

ROLE OF SYNTHETIC TIME SERIES IN HYDROMETEOROLOGICAL DATA ANALYSIS[1]

Ashok N. Shahane, Daniel Thomas and Paul Bock[2]

ABSTRACT: The use of synthetic time series (artificially simulated time series with specific and useful properties built into them) to increase the confidence in the statistical parameters of limited hydrometeorological time series is the subject matter of this paper. By constructing fourteen synthetic time series, a sensitivity analysis is performed to assess the net effect of nonstationarity, number of lags and small sample size on estimated spectral densities. Similarly, the effects of the harmonic-removal procedure on the resulting residual series and the confidence limits in cross spectral analysis are examined in the light of synthetic time series analysis. These analyses clearly indicate the useful supplemental role of synthetic time series in data analysis.

NATURE OF THE PROBLEM

Most of the modern environmental data processing systems use various advance statistical techniques to analyze the available bank of data. Among many others, the basic purposes of applying such methods is to provide statistical information which, in turn, can be used in practice for:

1) designing sampling intervals (Gunnerson, 1966),
2) validating the interrelationships between the systems parameters (Rodriquez, 1967; Rodriquez and Yevjevich, 1968),
3) interpreting the response of meteorological systems (Panofsky and Brier, 1968),
4) analyzing the atmospheric and terrestrial branches of the hydrologic cycle (Chow and Kareliotis, 1970; Kareliotis and Chow, 1972; Roesner and Yevjevich, 1966; Shahane, 1973).

While trying to extract such practical information with the help of these sophisticated statistical techniques, these techniques are looked upon as black boxes. As a result, there seems to be an increasing tendency to emphasize the interpretation task of statistical

[1]Paper No. 75039 of the *Water Resources Bulletin.* Discussions are open until December 1, 1975.

[2]Respectively, Environmental Systems Engineer, Central and Southern Florida Flood Control District, West Palm Beach, Florida 33402; Assistant Professor of Statistics, Department of Statistics, University of Connecticut, Storrs, Connecticut 06268 and; Professor of Hydrology and Water Resources, Civil Engineering Department, University of Connecticut, Storrs, Connecticut 06268.

parameters (provided by the black boxes) rather than the understanding of the basic statistical assumptions of the black boxes. This seems to be grounds for a possible confrontation between applied scientists and statisticians. A typical debate between these two groups of disciplines centers on the basic difference in the fundamental philosophy and approach of looking at statistical methods. Applied scientists claim that they are required to make decisions at a particular point in time based on the available set of data and thus their effort is directed towards applying all the possible statistical techniques (conventional as well as advanced) to extract the maximum decisive information from the available data which may be limited or sufficient. In other words, their basic planning philosophy seems to be that "something is better than absolute nothing." On the other hand, statisticians examine the available data firt and then depending on the characteristics of the data, sample size and the assumptions involved with different statistical theories, they select the appropriate technique for analyzing the available data. However, considering the fact that most of the time the data collection step is completely independent of the data analysis procedure, the collected data may not fulfill the basic theoretical statistical assumptions. In such cases, a statistician may be unable to analyze such data to provide any statistical inference confirming the other philosophy that "absolute nothing is better than something possibly incorrectly derived from a limited data base." Recently, the authors have gone through the above dilemmatic debate while analyzing some hydrometeorological data of the United States and the main purpose of this paper is to demonstrate a simple methodology to achieve a golden mean between the above two extreme viewpoints.

RESEARCH PROCEDURE

In an attempt to explore the characteristic behavior of the hydrometeorological components of the United States, the available five years monthly observations of atmospheric moisture transport (period May 1958 to April 1963) were first used to generate the atmospheric transport $(\nabla \cdot Q)$, precipitable water (ΔW), precipitation (P), runoff (R), evapotranspiration (E), and change in storage (ΔS) components of the atmospheric and terrestrial branches of the subcontinental hydrologic cycle. Although various hydrologic adjustments and comparisons strongly reflect the adequacy of our adopted methodology (Shahane, 1973), the data provides only sixty discrete values for the hydrometeorological time series. It is to be noted, however, that the available and generated values constitute a unique set of data based on the largest number of hydrometeorological observations so far. Therefore, the dilemmatic situation, similar to that mentioned in the previous section, arises when one wants to apply statistical methods to analyze such a relatively short hydrometeorological data set. More specificially, the dilemma relates to the following questions.

1) Is it feasible to apply sophisticated techniques like autocorrelation, spectral and cross-spectral techniques to explore the internal characteristics of the limited hydrometeorological data set?

From the previous discussion, we know that a statistician will probably answer the above question by "No," and an applied scientist by "Yes." Thus, the second question immediately follows:

2) Is it possible to devise a simple methodology to satisfy both professionals by performing a sensitivity analysis of some kind?

To provide answers in these directions, the available hydrometeorological time series were first examined rougly in terms of the magnitude of the individual numbers and the kind of periodicity associated with them. Based on these basic characteristics, some time series were artificially formed with different magnitudes for the sixty numbers (but in the observed range of the available hydrometeorological series) and with a known period built into them. Such time series are known as synthetic time series. Since, in the hydrometeorological study reported by Shahane (1973), it is assumed that the generating process of the hydrometeorological components is the sum of a deterministic and a random part, for comparative evaluation of data processing techniques, synthetic time series were also formed with known deterministic and random parts. Such time series are shown in table 1 and 2. Table 1 includes eight synthetic time series with a periodic deterministic part only. Basically these series have different magnitude levels but the same 12 month periodicity. Two series (No. 5 and 8) are geometric series; one with abnormal range and the other with the range found in our hydrometeorological data. Other series in table 1 were formed systematically with different ranges except series No. 7 which was formed with a 12 month period having different within-year variances. Table 2 depicts six synthetic time series with both deterministic and random parts built into them. The periodic deterministic part is shown in column 1. To this deterministic part, a random part with different variances is added to form the remaining five synthetic time series as given in columns 2, 3, 4, 5 and 6. The random parts are generated from normal random number tables.

While constructing these synthetic time series, there can be an opposing argument that for obtaining a comparable and rational picture by synthetic time series, many combinations of different numbers are required. However, the more realistic counter-argument to the above point is that it is possible to include adequately all the possible basic periodic and random characteristics of hydrometeorological time series in a small number of synthetic time series. It is indeed possible to form many other synthetic time series with different properties. However, they will have different basic properties which may not be applicable to the hydrometeorological time series in question.

RESULTS AND DISCUSSION

Spectral density analysis is widely used as a powerful data processing tool to detect the nature of the deterministic part of a stationary time series. However, the basic classification of stationary versus nonstationary seems to be inadequately agreed upon by hydrologists and statisticians. Roesner and Yevjevich (1966) have mathematically expressed stationarity as

$$E(X_t) = \mu = \text{constant} \qquad \text{(first order stationary)}$$

$$\begin{aligned} E(X_t X_{t+L}) &= f[(t+L)-t] \\ &= \sigma^2 + \mu^3 \\ &= \text{constant} \qquad \text{(second order stationary)} \end{aligned}$$

$$\begin{aligned} E[X_t X_{t+L_1} X_{t+L_2}] \\ &= g[(t+L_1)-t, (t+L_2)-t] \\ &= g(L_1, L_2) \\ &= \text{constant} \qquad \text{(third order stationary)} \end{aligned}$$

TABLE 1. Eight Synthetic Periodic Time Series (with Only Deterministic Part) Used in Comparison of Frequency Responses.

			Synthetic Time Series				
1	2	3	4	5	6	7	8
0.1	5	10	-1	1	-10	7.50	1.25
0.2	6	20	-2	2	-8	5.40	1.55
0.3	7	30	-3	4	-6	11.90	1.93
0.4	8	40	-4	8	-4	2.00	2.41
0.5	9	50	-5	16	-2	0.80	3.00
0.6	10	60	-6	32	0	0.01	3.74
0.7	11	70	-7	64	2	-0.31	4.66
0.8	10	80	-8	128	4	0.50	5.82
0.9	9	90	-9	256	6	7.50	7.25
1.00	8	100	-10	512	8	7.50	9.03
1.1	7	110	-11	1024	10	0.80	11.25
1.2	6	120	-12	2048	12	11.00	11.02
0.1	5	10	-1	1	-10	7.50	1.25
0.2	6	20	-2	2	-8	5.40	1.55
0.3	7	30	-3	4	-6	11.90	1.93
0.4	8	40	-4	8	-4	2.00	2.41
0.5	9	50	-5	16	-2	0.80	3.00
0.6	10	60	-6	32	0	0.01	3.74
0.7	11	70	-7	64	2	-0.31	4.66
0.8	10	80	-8	128	4	0.50	5.82
0.9	9	90	-9	256	6	7.50	7.25
1.0	8	100	-10	512	8	7.50	9.03
1.1	7	110	-11	1024	10	0.80	11.25
1.2	6	120	-12	2048	12	11.00	11.02
0.1	5	10	-1	1	-10	7.50	1.25
0.2	6	20	-2	2	-8	5.40	1.55
0.3	7	30	-3	4	-6	11.90	1.93
0.4	8	40	-4	8	-4	2.00	2.41
0.5	9	50	-5	16	-2	0.80	3.00
0.6	10	60	-6	32	0	0.01	3.74
0.7	11	70	-7	64	2	-0.31	4.66
0.8	10	80	-8	128	4	0.50	5.82
0.9	9	90	-9	256	6	7.50	7.25
1.0	8	100	-10	512	8	7.50	9.03
1.1	7	110	-11	1024	10	0.80	11.25
1.2	6	120	-12	2048	12	11.00	11.02
0.1	5	10	-1	1	-10	7.50	1.25
0.2	6	20	-2	2	-8	5.40	1.55
0.3	7	30	-3	4	-6	11.90	1.99
0.4	8	40	-4	8	-4	2.00	2.41
0.5	9	50	-5	16	-2	0.80	3.00
0.6	10	60	-6	32	0	0.01	3.74
0.7	11	70	-7	64	2	-0.31	4.66
0.8	10	80	-8	128	4	0.50	5.82
0.9	9	90	-9	256	6	7.50	7.25
1.0	8	100	-10	512	8	7.50	9.03
1.1	7	110	-11	1024	10	0.80	11.25
1.2	6	120	-12	2048	12	11.00	11.02
0.1	5	10	-1	1	-10	7.50	1.25
0.2	6	20	-2	2	-8	5.40	1.55
0.3	7	30	-3	4	-6	11.90	1.93
0.4	8	40	-4	8	-4	2.00	2.41
0.5	9	50	-5	16	-2	0.80	3.00
0.6	10	60	-6	32	0	0.01	3.74
0.7	11	70	-7	64	2	-0.31	4.66
0.8	10	80	-8	128	4	0.50	5.82
0.9	9	90	-9	256	6	7.50	7.25
1.0	8	100	-10	512	8	7.50	9.03
1.1	7	110	-11	1024	10	0.80	11.25
1.2	6	120	-12	2048	12	11.00	11.02

TABLE 2. Five Synthetic Time Series
(with Their Deterministic and Random Parts) Used in Comparison of Frequency Responses.

Periodic Deterministic Part	Synthetic Time Series with Normal Random Component				
	Variance=0.9 Mean=0	Variance=2.25 Mean=0	Variance=4.42 Mean=0	Variance=3.16 Mean=0	Variance=1.8 Mean=0
1	2	3	4	5	6
5	5.3216	5.5085	5.7119	5.6027	5.6100
6	5.7419	5.6235	5.4729	5.5538	5.5490
7	7.2400	7.3795	7.5313	7.4498	7.4554
8	7.7316	7.5755	7.4060	7.4949	7.4906
9	9.5455	9.8625	10.2070	10.0220	10.0350
10	10.0369	10.0585	10.0819	10.0693	10.0702
11	11.0265	11.0420	11.0588	11.0498	11.0504
10	9.5931	9.3565	9.0991	9.2373	9.2278
9	8.0229	7.4550	6.8370	7.1687	7.1460
8	7.6794	7.4930	7.2902	7.3991	7.3916
7	7.1412	8.8045	9.5263	9.1389	9.1654
6	5.2259	4.7760	4.2864	4.5492	4.5312
5	3.7991	3.0695	2.2973	2.7118	2.6834
6	6.2447	6.3870	6.5418	6.4587	6.4644
7	8.6023	9.2535	10.5469	10.0030	10.0402
8	7.8094	7.6985	7.5779	7.6429	7.6382
9	9.7684	10.2150	10.7010	10.4401	10.4580
10	10.2665	10.4219	10.5901	10.4996	10.5058
11	9.6861	8.9225	8.0915	8.5375	8.5070
10	11.2900	12.0400	12.8560	12.4180	12.4480
9	9.1423	9.2250	9.3150	9.2667	9.2700
8	8.3149	8.4980	8.6972	8.5691	8.5976
7	7.1736	7.2745	7.3843	7.3255	7.3294
6	4.9584	4.3530	3.6942	7.0470	4.0236
5	5.9382	6.4835	7.0769	6.7584	6.7802
6	5.0058	4.4280	3.7992	4.1367	4.1136
7	5.7791	5.0695	4.2973	4.7118	4.6834
8	6.9280	6.7970	6.6268	5.9909	5.9660
9	9.1936	7.7250	7.2150	7.4887	7.4700
10	11.1735	11.8555	12.5977	12.1993	12.2266
11	10.9118	10.8605	10.8047	10.8346	10.8326
10	10.2597	10.4111	10.5754	10.4872	10.4932
9	10.0891	10.7220	11.4108	11.0354	11.0664
8	7.8568	7.7750	7.6829	7.7316	7.7282
7	6.8757	6.8035	6.7249	7.7671	6.7642
6	4.4413	3.5355	2.5497	3.0788	3.0426
5	5.4202	5.6645	5.9303	5.7876	5.7974
6	6.0037	6.0060	6.0084	6.0071	6.0072
7	6.8947	6.8335	6.7669	6.8027	6.8002
8	8.9373	9.4820	10.0748	9.7566	9.7784
9	9.1281	9.2025	9.2835	9.2400	9.2430
10	9.5950	9.3595	9.1033	9.2408	9.8314
11	9.6311	8.8355	7.9697	8.4344	8.4026
10	9.0039	9.8425	9.7795	9.8133	9.8110
9	8.6301	8.4150	8.1810	8.3066	8.2980
8	6.6121	5.8055	4.9377	5.3988	5.3666
7	5.9760	5.5770	4.8496	5.1794	5.1568
6	6.2552	6.4035	6.5649	6.4782	6.4842
5	5.6887	6.0890	6.5246	6.2908	6.3068
6	5.2031	4.7400	4.2360	4.5065	4.4880
7	7.4809	7.7605	8.0647	7.9014	9.9126
8	8.2742	8.4335	8.6069	8.5138	8.5202
9	10.3795	11.1900	12.0660	11.5958	11.6280
10	10.8225	11.3000	11.8207	11.5415	11.5606
11	10.9822	10.9970	10.9958	11.9966	10.9964
10	11.4847	12.3475	12.3865	12.7825	12.8170
9	9.2020	9.3195	9.4475	7.3787	9.3834
8	9.3481	10.1315	10.9841	10.5265	10.5578
7	6.8093	6.6985	6.5779	6.6417	6.6382
6	6.6053	6.9570	7.3398	7.1343	7.1484

where

X_t = the value of the observed variable at time t,

$E(X_t)$ = the expected value,

f and g = functions,

L, L_1, L_2 = time lags,

μ = population mean,

σ^2 = population-variance of X_t.

For hydrometeorological data, for example, first order stationary implies that the expected monthly value of January streamflow or precipitation is the same as the expected value of any other month (say July) streamflow or precipitation. According to this statistical definition of stationarity, most of the hydrometeorological data with seasonal variations are nonstationary. Although the purpose of checking for stationarity is to make sure that the covariance structure of the data does not change substantially with time due to artifically imposed conditions, it seems that hydraulic engineers have a different way of looking at stationarity in their hydrologic data. Many such efforts by Chow (1970), Wastler (1969), first looked for the presence of trend by visual comparison or by a suitable testing procedure. If trend was not observed, then the data was presumed to be first order stationary. If trend was detected, the original time series was made "stationary" (first order) by removing the observed trend. The transformed time series was then subjected to spectral analysis which assumes stationarity. However, the statistical theory of spectral analysis requires the time series to be second order stationary also. In engineering applications, where statistical results are interpreted in light of physical phenomenon, if the assumption of second order stationarity is artificially satisfied by some further transformation of the data, the physical interpretation of the resulting statistical parameters becomes weak (although the statistical analysis is more valid from the statistician's point of view). In other words, the dilemma is that if the sample record satisfies the second order stationarity condition, then and then only, can the spectral density technique be applied, whereas from the hydrologic viewpoint, if an effort is made to satisfy the above statistical condition, physical interpretation is lost. To get around this dilemma in an engineering way, the effects of violation of the statistical assumptions on decisive parameters were studied. During such an effort, the net effect of nonstationarity, number of lags and small sample size on the estimated spectral densities (variances at different frequencies) was attempted and spectral densities for the synthetic time series for different lag numbers (i.e., m = 6, 12, 18, 24 and 30) were computed. Some of these values are given in the figures 1 and 2. It can be seen from these figures that the shape of the spectrum plot for a nonstationary series is indeed different than the conventional one (a high spike at the significant frequency) and it seems that the usual thumb rule of requiring the maximum number of lags in the spectral analysis to be less than one tenth of the sample record (i.e., m = 6) is applicable to the hydrometeorological time series which are similar to the given synthetic time series. Furthermore, since the shape of the variance-spectrum plot for the synthetic time series is similar to the hydrometeorological time series as reported by Shahane (1973), it appears that the variations in the statistical parameters are due to the inherent characteristics of the observed and generated hydrometeorological time series and not due to the violation of the basic statistical assumptions underlying the spectral density analysis. Precisely, this is the type of answer needed for the second question mentioned in the previous section.

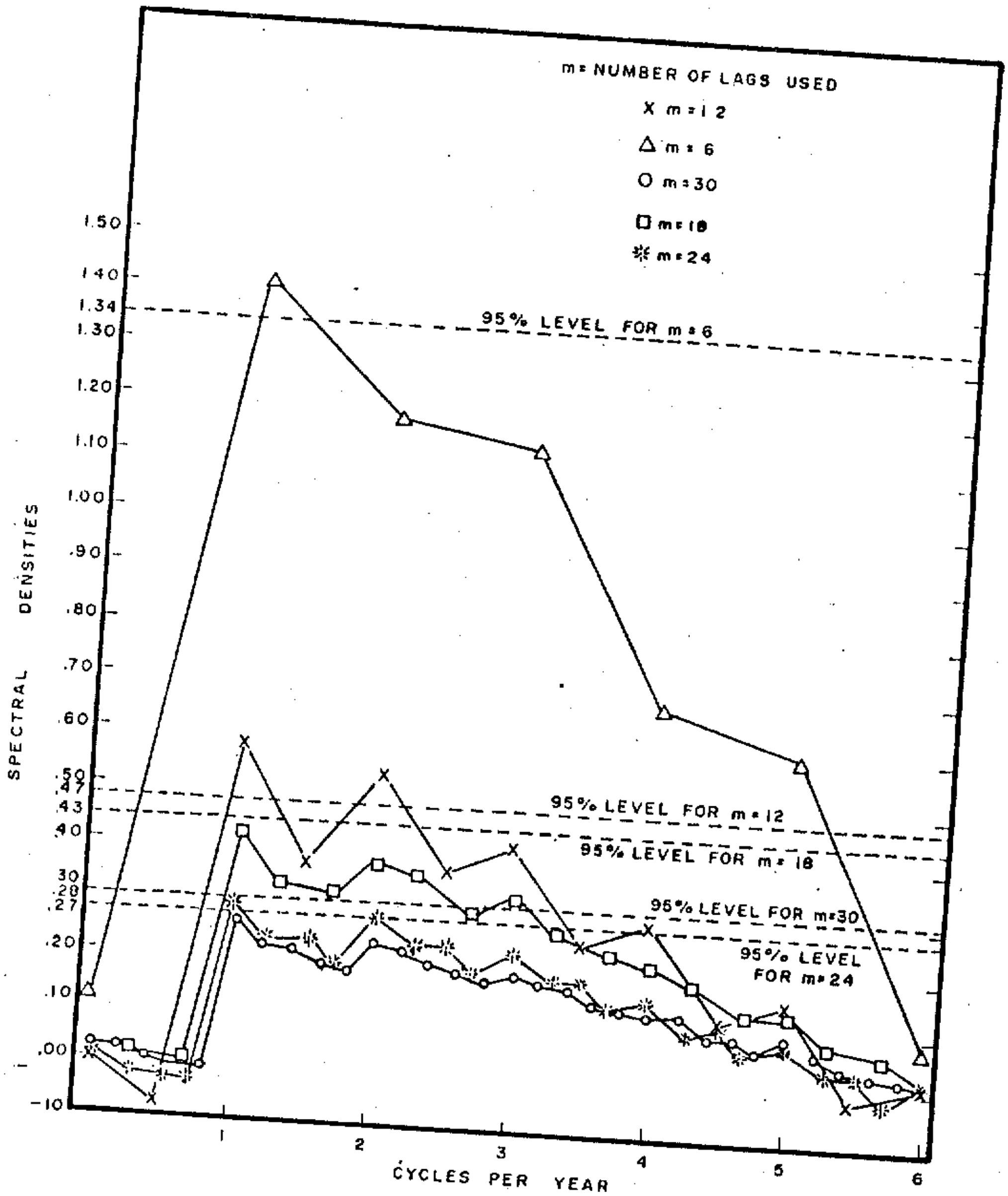

Figure 1. Effects of Number of Lags on Spectral Densities of Synthetic Time Series No. 1 of Table 2.

To investigate the effects of trend removal procedures on residual series (series formed by subtracting the deterministic part from the original series), spectral density has been suggested by some investigators (Malhotra, 1969; Roesner and Yevjevich, 1966). If spectral values for the residual series are not significant for the lags (for which significant spectral values are observed for the original series) then it indicates the adequacy of removing a fixed function of time from the original series and the resulting residual series

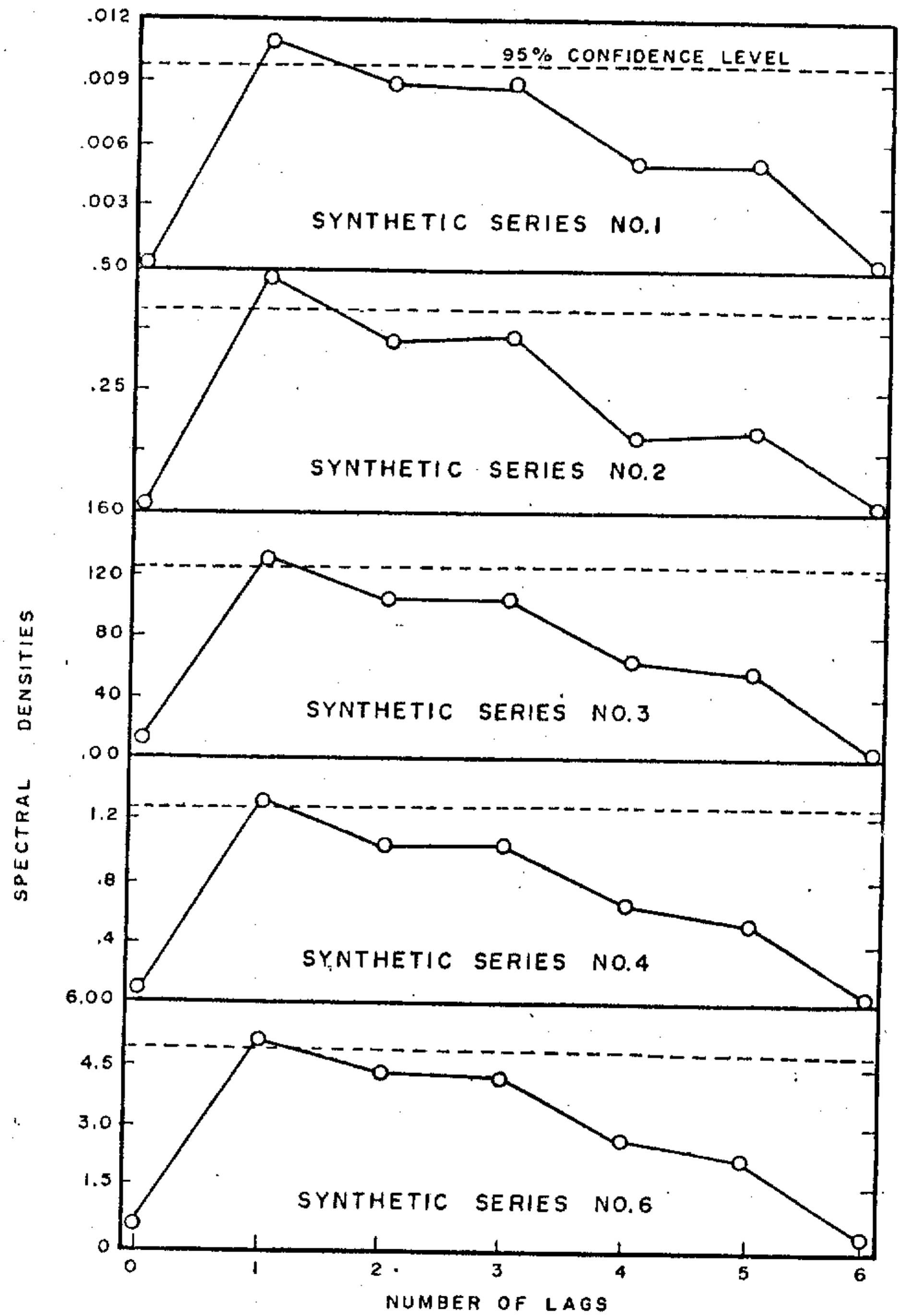

Figure 2. Standardization of Spectral Density Technique
by Synthetic Time Series Given in Table 2.

can be further analyzed as a stationary process. Although such a method looks promising, it seems to be invalid for our hydrometeorological time series because of the fact that the spectral densities show artificial periodicity in the residual part. This is demonstrated by the analysis of synthetic time series by Shahane (1973). If four harmonics are removed from an original synthetic time series, then it is observed that residual series become random noise and spectral densities at various lags lie well below significant values. However, spectral density analysis on these residual series of synthetic time series (i.e., after removing four harmonics from the original synthetic time series) indicate the presence of periodicity. Since hydrologic residual series basically represent a noise (which is known to be a nonperiodic type), this type of ambiguous observation of periodicity may be due to (1) inadequacy of spectral density technique for residual analysis because of small sample size or (2) the introduction of artificial periodicity while removing four harmonics (maybe more than required) from the original series. To investigate the validity of either of these two points, again, synthetic time series are analyzed. Five synthetic time series (with known deterministic and random parts) are subjected to a run test. The number of runs of a particular time series can be instrumental in assessing the trend, periodic and random properties of the time series. Therefore, the run test was first applied to the five original synthetic time series. Then it was applied to each of these synthetic series after removal of one harmonic. The results of each run test are given in table 3. From this table, it is observed that systematic harmonic removal does not appear

TABLE 3. Effects of Harmonics Removal on Number of Runs for Synthetic Time Series.*

Description	Number of Runs for Synthetic Time Series with Periodic and Normal Random Component Having Different Variances of Random Part					95% Confidence Interval for No. of Runs
	Variance =4.42	Variance =0.90	Variance =3.16	Variance =1.80	Variance =2.25	
Five Original Synthetic Series	17	14	16	16	16	22-39
Synthetic Series After Removal of First Harmonics	36	33	36	36	36	22-39
Removal of Second Harmonics	38	33	36	38	36	22-39
Removal of Third Harmonics	38	36	36	36	36	22-39
Removal of Fourth Harmonics	34	36	34	34	31	22-39

*Synthetic time series are No. 2, 3, 4, 5 and 6 of Table 2.

to introduce any kind of artificial periodicity in the residual series. Thus, out of the two possible reasons for observing periodicity in residual series (as mentioned above), the inadequacy of the spectral density analysis seems to be more valid than the other. Such inadequacy of spectral analysis seems to be related to the inadequate estimation of the confidence level for the spectral density (which is computed from a formulation based on large sample theory). From the above discussion, it can be summarized that the spectral density analysis has previously been proposed also for checking the adequate removal of the periodic deterministic part from the original series. However, in our case with small sample size, spectral density seems to be inadequate for such a purpose. This is another result which is obtained merely by analyzing synthetic time series and thus shows the simple and important role of synthetic time series in the data analysis.

Like any other statistical technique, an important point of cross-spectral analysis is related to the estimation of 95% confidence limits for sample coherences and phase angles. Among these two parameters, significant values of coherence indicate qualitatively the dependence structure between the hydrometeorological parameters in question. Whereas sophisticated analysis (requiring large sample size) on phase angles can estimate quantitatively the dependence between hydrologic components. For our sample size, it seems better, statistically, to emphasize coherence confidence limits rather than evaluating the significance values of phase angles. To estimate such a confidence limit for coherence, there are two formulations available in the literature. According to the formulation given by Goodman, the approximate formula for the limiting coherence at probability level p is

$$= 1-p^{1/(d.f.-1)}$$

where d.f. = degrees of freedom

$$= \frac{2N - m/2}{m}$$

N = sample size

m = lag number (6 in our case)

From the table provided by Panofsky and Brier (1958), in our case with 20 degrees of freedom, and 95% limit of coherence is 0.38. This means that the chances are 1 in 20 that a coherence of 0.38 or less will be found by accident. If the coherence values of the residual series of a particular data set are higher than this value of 0.38 for most of the significant frequencies, then dependence between those parameters is ascertained and vice versa.

Another approach is proposed by Granger and Hatanaka (1964). In this approach, instead of computing degrees of freedom, the ratio $\frac{N}{m}$ is computed and from the tables provided by the above group, a 95% limiting coherence is estimated. In our case, with N = 60 and m = 6, the limiting value of coherence with 95% level of confidence is given 0.73, whereas Goodman's approach provided a value of 0.38. Therefore, a question now arises as to "which of these two approaches is better and is to be selected?" It seems that the answer to this question may be again provided by the analysis of synthetic time series. If one looks at the coherences of the original and residual series of synthetic time series (given in tables 4 and 5), it is noted that the coherences for most of the combinations of lag 0 and 1 have values greater than 0.38. Considering the fact that the residual series of the synthetic time series are normal variates, there is likely to be no dependence between these residual series. Therefore, limiting coherence of 0.73 given by Granger and Hatanaka seems to be more realistic and convincing than Goodman's estimates. Thus, once again synthetic time series are used to select out the appropriate confidence limits for the relatively limited hydrometeorological time series.

CONCLUSIONS

Based on the discussion presented in previous sections coupled with an additional detailed analysis reported by Shahane (1973), the following conclusions can be drawn:

1) Although the use of spectral analysis to investigate the effects of trend removal procedure on residual series is suggested by Malhotra (1969), Roesner and Yevjevich (1966), it is found to be inadequate for our hydrometeorological time series because spectral density shows artificial periodicity in the residual part of the time series. This can be clearly demonstrated by the analysis of synthetic time series.

2) The results of the synthetic time series analysis in investigating the comparative response to some of the critical statistical points are encouraging. Use of such series especially in spectral and cross-spectral analysis for selecting the proper number of lags, smoothing technique and confidence limits is beneficial.

3) Sensitivity analysis performed on statistical variables of hydrometeorological and synthetic time series reveals that the proper number of lags and Hamming-Tukey weights are more important in cross-spectral analysis than in spectral density analysis as given by Shahane (1973). This is due to the unrealistic output of

TABLE 4. Coherences and Phase Angles of Original ~~Six~~ Synthetic Time Series Given in Table 1.

Pair of Synthetic Time Series	Coherences* and Phase Angles** for Six Lags						
	0	1	2	3	4	5	6
3 and 2	0.949 0.03	0.878 0.02	0.055 0.00	0.013 0.03	0.010 0.12	0.015 0.06	0.007 0.07
3 and 5	0.909 0.00	0.846 0.00	0.066 0.00	0.026 0.00	0.023 0.04	0.040 0.01	0.021 0.01
3 and 4	0.886 -0.05	0.828 -0.05	0.069 -0.03	0.029 -0.03	0.029 0.06	0.052 0.05	0.026 0.05
3 and 8	9.812 -0.23	0.633 -0.21	0.049 0.11	0.018 -0.27	0.019 0.14	0.033 -0.16	0.178 -0.16
4 and 2	0.880 0.07	0.817 0.07	0.055 0.02	0.015 0.07	0.013 0.04	0.023 0.00	0.010 0.00
4 and 5	0.869 0.04	0.808 0.04	0.072 0.03	0.035 0.03	0.035 -0.03	0.067 -0.04	0.034 -0.04
4 and 1	0.801 0.19	0.769 0.18	0.037 0.11	0.003 0.18	0.000 0.05	0.001 5.75	0.001 5.86
2 and 5	0.906 -0.03	0.839 -0.03	0.055 0.01	0.014 -0.04	0.011 -0.06	0.019 -0.04	0.008 -0.04
6 and 2	9.907 0.24	0.619 0.22	0.037 -0.12	0.009 0.43	0.007 -0.14	0.014 0.29	0.058 0.23
6 and 5	9.561 0.240	0.627 0.210	0.052 -0.130	0.022 0.29	0.025 -0.13	0.043 0.20	0.250 0.180
6 and 4	9.300 0.21	0.620 0.180	0.055 -0.18	0.025 0.25	0.033 -0.09	0.056 0.22	0.323 0.21
6 and 1	9.480 0.31	0.561 0.31	0.022 0.08	0.002 0.93	0.001 -5.28	0.001 5.71	0.032 5.87

* Upper Values Represent Coherences and
**Lower Values Represent Phase Angles in Months, ~~for Synthetic Time Series No. 8 in Table 1.~~

coherence values in the absence of the above two key factors (proper number of lags and Hamming-Tukey weights).

4) Using synthetic time series coupled with the run test, Spearman's T test, Cox and Stuart tests, a conventional procedure of formulating a mathematical model can be modified and illustrated for the atmospheric divergence ($\nabla \cdot Q$) time series of a large eastern region as shown by Shahane (1973).

5) As demonstrated in the previous sections, synthetic time series can be effectively used in

   a) detecting the erratic (or otherwise) behavior of the statistical parameters when the basic data does not follow the basic assumptions underlying many advanced and applied statistical techniques, and

   b) increasing confidence in the final estimates of statistical parameters (for example, autocorrelation, spectral and cross-spectral parameters) of nonstationary hydrometeorological time series.

TABLE 5. Coherences of Residual Series of Synthetic Time Series given in Table 1

Pair of Synthetic Time Series	Coherences* and Phase Angles** for Six Lags						
	0	1	2	3	4	5	6
3 and 2	0.527 0.07	0.457 0.06	0.243 -0.04	0.280 0.05	0.294 0.09	0.429 0.06	0.185 0.05
3 and 5	0.542 0.03	0.506 0.03	0.370 0.000	0.462 0.020	0.386 0.03	0.554 0.03	0.257 0.03
3 and 4	0.533 0.00	0.497 0.00	0.384 0.02	0.505 0.00	0.442 0.07	0.611 0.08	0.271 0.07
3 and 8	1.065 -0.03	0.384 -0.01	0.389 0.01	0.390 -0.07	0.589 0.00	0.439 0.00	0.741 0.00
4 and 2	0.495 0.07	0.415 0.07	0.221 -0.02	0.277 0.05	0.301 0.00	0.446 -0.05	0.190 -0.05
4 and 5	0.511 0.02	0.460 0.03	0.337 0.03	0.451 0.01	0.395 -0.04	0.588 -0.05	0.268 -0.04
4 and 1	0.002 2.13	0.002 3.27	0.000 -5.43	0.000 3.61	0.000 5.99	0.004 5.78	0.005 5.87
2 and 5	0.500 -0.04	0.421 -0.03	0.215 0.40	0.256 -0.03	0.266 0.00	0.411 0.00	0.183 0.00
6 and 2	0.960 0.09	0.311 0.07	0.229 -0.07	0.217 0.15	0.404 0.06	0.325 0.02	0.517 0.02
6 and 5	1.009 0.06	0.350 0.05	0.347 -0.02	0.350 0.10	0.528 0.03	0.431 0.03	0.732 0.03
6 and 4	0.989 0.04	0.343 0.02	0.356 -0.05	0.380 0.08	0.603 0.06	0.472 0.08	0.766 0.07
6 and 1	0.005 2.15	0.001 3.27	0.000 -5.29	0.000 4.12	0.000 -5.73	0.003 5.98	0.012 5.99

* Upper Values Represent Coherences and
**Lower Values Represent Phase Angles in Months ~~for Synthetic Time Series No. 8 in Table 1~~.

ACKNOWLEDGMENTS

The work upon which this paper is based was supported by funds provided by the United States Department of the Interior, Office of Water Resources Research through the Institute of Water Resources, University of Connecticut, as authorized under the Water Resources Act of 1964, PL 88-379. The University of Connecticut Research Foundation also provided for a short-term final publication and research. In this context, Dr. W. C. Kennard, Director of the Institute of Water Resources and Dr. Hugh Clark, Dean of Graduate School of the University of Connecticut, are gratefully acknowledged.

LITERATURE CITED

Chow, Ven T. and S. J. Kareliotis, 1970. Analysis of Stochastic Hydrologic Systems. Water Resources Research, Vol. 6, No. 6, pp. 1569-1582.

Granger, C. W. J., 1964. Spectral Analysis of Economic Time Series. Princeton University Press, Princeton, New Jersey.

Gunnerson, C. G., 1966. Optimizing Sampling Intervals in Tidal Estuaries. Journal of the Sanitary Engineering Division, ASCE, Vol. 92, No. SA2.

Kareliotis, S. J. and V. T. Chow, 1972. Analysis of Residual Hydrologic Stochastic Processes. Journal of Hydrology, Vol. 15, No. 2, February, pp. 113-130.

Malhotra, G. P., 1969. Hydrologic Cycle of North America Formulated by the Analysis of Water Vapor Transport Data. A dissertation submitted in partial fulfillment of the requirement for the degree of Doctor of Philosophy at the University of Connecticut, Storrs, Connecticut.

Panofsky, H. A. and G. W. Brier, 1958. Some Applications of Statistics to Meteorology. University Park, Pennsylvania, Mineral Industries Extension Services, Pennsylvania State University, p. 224.

The Rand Corporation, 1966. A Million Random Digits with 100,000 Normal Deviates. The Free Press, New York, Collier-MacMillan Limited, London, p. 600.

Rodriquez, E., 1967. The Application of Cross-Spectral Analysis to Hydrologic Time Series. Hydrology Paper No. 24, Colorado State University, Fort Collins, Colorado.

Rodriquez, I. and V. Yevjevich, 1968. The Investigation of Relationship Between Hydrologic Time Series and Sunspot Numbers. Hydrology Paper No. 26, Colorado State University, Fort Collins, Colorado.

Roesner, L. A. and V. Yevjevich, 1966. Mathematical Models for Time Series of Monthly Precipitation and Monthly Runoff. Hydrology Paper No. 15, Colorado State University, Fort Collins, Colorado.

Shahane, A. N., 1973. Characteristics Behavior of the Components of the Hydrologic Cycle of the United States. A published dissertation submitted to the University of Connecticut as a partial fulfillment for the requirement of the degree of Doctor of Philosophy, p. 547.

Wastler, T. A., 1969. Spectral Analysis. U.S. Department of the Interior Publication, CWT-3, December, p. 99.