Drought Monitoring with Indexed Sequential Modeling

Kim, Hung Soo* · Yoon, Yong Nam**

ABSTRACT : The simulation techniques of hydrologic data series have developed for the purposes of the design of water resources system, the optimization of reservoir operation, and the design of flood control of reservoir, etc. While the stochastic models are usually used in most analysis of water resources fields for the generation of data sequences, the indexed sequential modeling (ISM) method based on generation of a series of overlapping short-term flow sequences directly from the historical record has been used for the data generation in the western USA since the early of 1980s. It was reported that the reliable results by ISM were obtained in practical applications. In this study, we generate annual inflow series at a location of Hong Cheon Dam site by using ISM method and autoregressive, order-1 model (AR(1)), and estimate the drought characteristics for the comparison aim between ISM and AR(1).

1. Introduction

Generations of synthetic data series are generally needed for sizing reservoirs, for determining the risk of failure (or reliability) of water supply for irrigation systems, and that of dependable capacities of hydroelectric systems, for planning studies of future reservoir operation, for planning capacity expansion of water supply systems, and for similar applications (Salas et al., 1980). Stochastic models have been mostly used for the generation of synthetic data series which represents the same statistical properties as the historical data. In practical applications, the indexed sequential modeling (ISM) method proposed and practiced by the U. S. Bureau of Reclamation has been used for the data generation in several agencies of the USA (House and Ungvari, 1983, and Kendall and Dracup, 1991). The illustrations of this method can be found in House and Ungvari (1983), and the applications of ISM model can be also found in the literatures (Labadie, et al., 1987, and Kendall and Dracup, 1991). Labadie et al. (1987) used this model for the analysis of project dependable on the hydropower capacity, and Kendall and Dracup (1991) used it for the comparative analysis of the generated streamflows in terms of ISM method and the first order autoregressive (AR(1)) model.

* Ph. D. Student, Dept. of Civil Engineering, Colorado State Univ., Fort Collins, Colo.
* * Professor, Dept. of Civil Engineering, Korea University

The drought may be defined as the deficit of moisture in the sense of hyrologic cycle of water balance but the definition of drought has continually been a stumbling block for drought monitoring and analysis. One can see the reviews of drought definitions in the literatures (Dracup et al., 1980, and Wilhite & Glantz, 1985). Also Lee et al.(1986) studied the frequency analysis of multiyear drought durations, and Zelenhasic and Salvai (1987) derived the drought characteristics in terms of the supremum theory of a random number of random variables.

This study uses the runs theory as a means to define hydrologic drought characteristics. Runs analysis defines the drought based on the stationary time series (Yevjevich, 1967 and Frick et al., 1990). Distinct events of flow deficit in a time series were presented on a time axis by introducing a constant truncation level. The duration of such flow deficit events could thus be obtained from the time axis. In this study the synthetic inflow time series at Hong Cheon Dam site is generated in terms of ISM method and AR(1) model, and the drought characteristics are monitored by runs analysis.

2. ISM and AR(1)

ISM method is that the synthetic flow series can be generated directly from the historical data by overlapping original historical sequences. The historical data can be divided into the several data sets having some short interval, and these divided ISM sequences generate the synthetic data series by overlapping ISM sequences according to the indices to be explained in the following sentences. Let T be the total number of data, n is ISM sequences, and I is an index of ISM. Then the number of the generated data N is defined as follows

$$N=\frac{nT}{I} \tag{1}$$

For easy discussion, we assume a time series $\{x_1, x_2 \cdots\cdots\cdots, x_{10}\}$ and the definition is shown in Fig. 1.

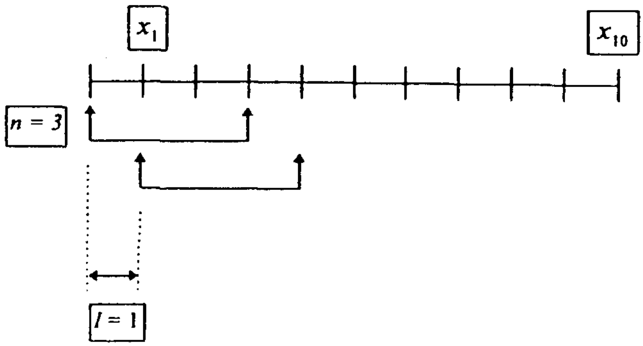

Fig. 1. Definition of ISM method

Therefore, in this case, $T=10$, and assume $I=1$ and $n=3$. Then the total number of generated series N is equal to 30. That is, the generated time series is

$$\{x_1, x_2, x_3, x_2, x_3, x_4, x_3, x_4, x_5, \cdots\cdots, x_{10},, x_1, x_2\}$$

Therefore one can conjecture the generated time series may be $N=15$ in the case of $I=2$. If one takes $n=5$ and $I=1$, N is equal to 50. Thus one can generate the data sequences as many as one needs.

According to Kendall and Dracup (1991), the generated streamflows by ISM method will preserve the statistical properties of the historical data series because of the generation property only by historical data but, for periodic data such as monthly or daily data, there are some issues on a correlation and periodicity between data. Generated data series by ISM represent a cycle of the critical period with respect to a selected indexed sequence and truncation level. A fundamental defect with this method is that the critical period observed in the historic record of the past is generally found to be the worst that has ever occurred in any of the indexed sequences. The only way an indexed sequence produces a more severe critical period is that the beginning and end of the historic record is a low flow period.

AR(1) is the most popular model of time series simulation and forecasting in hydrology. Some models such as autoregressive-moving average (ARMA) type models and threshold time series models were developed for preserving persistence of time series, but these models have serious problems concerned with parameter estimation (Salas et al., 1980, and Tong, 1990). Therefore the simple AR(1) model, which has a less parameter estimation, for generating streamflows in monthly reservoir simulations is enough and many investigators have proved the validation of AR(1) model (Bras and Rodriguez-Iturbe, 1985 and Kendall and Dracup, 1991). AR(1) model can be generally written as

$$x_t=\mu+\phi_1 (x_{t-1}-\mu)+\varepsilon_t \tag{2}$$

In this model, a stationary data series is normally distributed with mean μ and variance σ^2. The parameter ϕ_1 is estimated as the lag-1 serial correlation of the flows, ρ_1. The ε_t is the uncorrelated series which is independent of x_t, and it is also normally distributed with mean zero and variance $\sigma_\varepsilon^2=\sigma^2(1-\phi_1^2)$.

3. Drought Characteristics

Consider the time series $\{x_t\}$, $t=1,2, \cdots, N$ and let us define the negative run length, the negative run sum and the run intensity as the drought related statistics (Salas et al., 1980). A negative run is defined whenever the flow x_t, is less than the demand level Q which will be assumed to be a constant equal to a fraction of the overall mean of the time series, i.e., $Q = K\bar{x}$ where K is between greater than or equal to zero and less than or equal to one. Thus, a deficit $D_t = Q - x_t$, occurs whenever $x_t < Q$, otherwise $D_t = 0$. The first negative run length $RL(1)$ is defined as the length of the first period during which all successive runs are negative as shown in Fig. 2, and the first negative run sum $RS(1)$ is that of the deficits D_t during the first period of successive negative runs $RL(1)$.

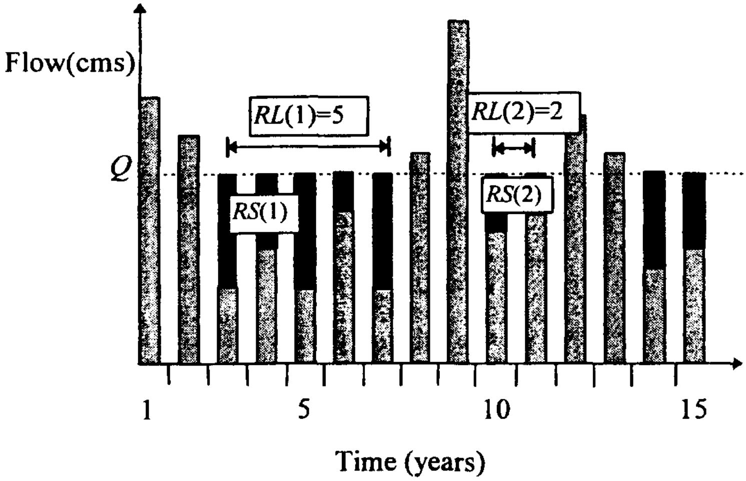

Fig. 2. Diagram of drought characteristics

In general there are several runs in a given time series. If there are m runs of run lengths $RL(1)$, $\cdots$, $RL(m)$ and run sums $RS(1)$, $\cdots$, $RS(m)$, the longest drought(RL), maximum deficit(RS) and its intensity(RI) can be defined as follows

$$RL = \max\{RL(1), \cdots, RL(m)\} \tag{2}$$

$$RS = \max\{RS(1), \cdots, RS(m)\} \tag{3}$$

$$RI = \max\left\{\frac{RS(1)}{RL(1)}, \cdots, \frac{RS(m)}{RL(m)}\right\} \tag{4}$$

The most important characteristics out of three are a run length (duration) and a run sum (deficit), and these two statistics are often chosen for the validation of the time series models. For the time series at more than one site, the period of negative runs on the streamflows at all sites can be used for the joint negative runs. Therefore once the negative runs are obtained from all the sites, the run characteristics can be computed as an univariate case (Bayazit, 1981).

4. Application

4.1 Drought Characteristics of a Historical Record

We use the inflow series of 22 years (1971~1992) at Hong Cheon Dam site to perform the drought analysis and plot the time series in Fig. 3 (Korea Electric Power Co, 1993). It shows more serious droughts in the years of 1976~1977, and 1982~1983 than those of other years. The drought characterisics for a historical record of 22 years are obtained according to its truncation levels, i. e., 100%, 90%, 80%, and 70% of the mean as the demand levels. Results of the 22-year drought characteristics are shown in table-1.

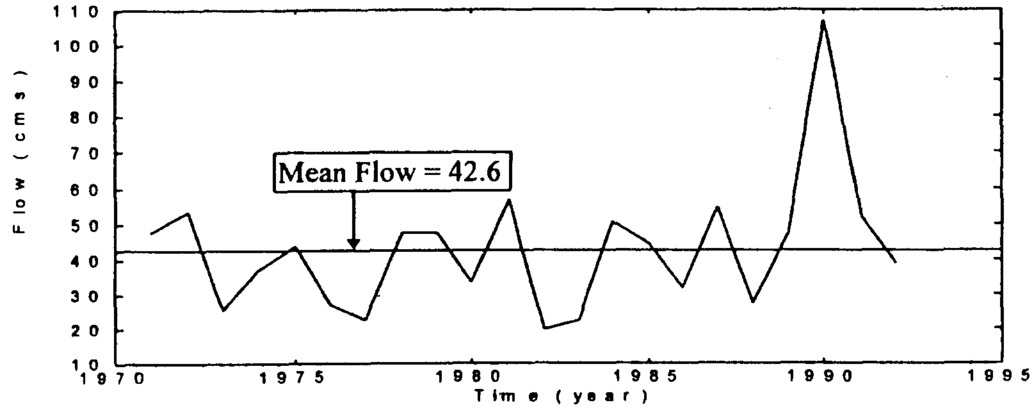

Fig. 3. Historical series of annual inflow for Hong Cheon Dam site

Table 1. Drought characteristics for a historical inflow series

	RL(years)				RS(cms-year)				RI(cms)			
truncation level(%)	100	90	80	70	100	90	80	70	100	90	80	70
return period (22yrs)	3	3	3	3	66.15	53.37	40.60	27.82	22.05	17.79	13.53	9.27

4. 2 Modeling and generation of drought characteristics

Application on drought characteristics was performed for the annual inflow time series at a location of Hong Cheon Dam site. However, since the historical data is one of many realizations, one cannot say the estimates from a historical data are not biased. Therefore one may have to use more reliable methods for the estimates of drought characteristics. Stochastic modeling and simulation has been widely used for the derivation of drought characteristics in a closed form. In this study a simple AR(1) model of Eq. (2) is assumed and tested to represent the underlying inflow data.

Model fitting

It was estimated the skewness value of 0.2 for the historical data in its logarithmic transform which can be considered as a log-normal distribution, and the mean and the standard deviation of log-transformed annual inflow series were obtained as 3.7-cms and 0.4-cms respectively. The lag-1 serial correlation, ρ_1 =0.064, was obtained, and it can be used as ϕ_1 in Eq. (2). From the suggestion of Snedecor and Cochran which the normality of time series can be tested by the skewness (Salas et al., 1980), the log-transformed annual inflow series showed its normality. For the choice of the model, the PACF(partial autocorrelation function) and the AIC (Akaike Information Criteria) are used (Salas et al., 1980). Then the residuals are examined for independence and normality of time series.

The results of PACF and AIC show order-1 of AR model. That is, PACF had a cutoff after lag-1 for log-transformed annual inflow series, and AIC represented its minimum value of -59.96 in order-1, and -57.96 for order-2 and -55.96 for order-3. We used the Porte Manteau lack of a fit test for testing independence of residuals, and obtained the statistic, Q_p=6.97, for testing independence of residuals from the Porte Manteau lack of fit test and this is less than $\chi^2_{0.95,\nu}$=11.1, where ν is the number of degrees of freedom and ν=5 is used in this study. Hence, it was considered to be adequate for the annual inflow series of Hong Cheon Dam site. The normality of residuals need not be performed because we concluded the time series is normal in the above paragraph. The correlogram of Hong Cheon Dam site annual inflow series is plotted in Fig. 4 to confirm AR(1) model fitted to its logarithm, and the autocorrelations between the generated-AR(1) and the theoretical-AR(1) are well defined. Therefore, AR(1) model is expressed as an appropriate model of this annual inflow series.

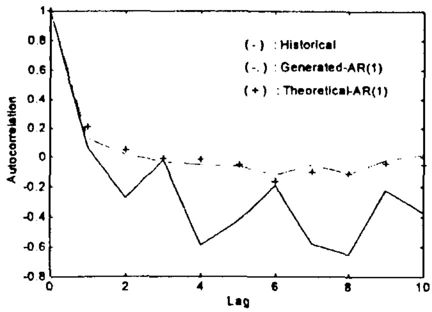

Figure 4. Historic and model autocorrelations of annual inflow for Hong Cheon Dam site

Comparison of drought characteristics between AR(1) and ISM

Alternatively ISM method is used to generate the synthetic inflow series, and to compare the drought characteristics with AR(1) model. Here the drought characteristics by AR(1) are obtained for the cases of 100, 500, and 1000 traces. The mean values for the historical data set and for the generated inflow series by AR(1) are shown in Table 2. The mean values of AR(1) are the averaged ones of the generated inflow series for each trace. Table 3 is the drought characteristics for AR(1)-generated inflow series. The drought characteristics are not significantly different for each trace.

Table 2. Mean values of historical inflow and AR(1)-generated inflow

	Historical	AR(1)				
return period (yrs)	22	50	100	150	200	500
mean(cms)	42.58	41.99	41.75	41.71	41.68	41.58

Table 3. Drought characteristics for AR(1)-generated inflow series

Table 3(a) The longest run length (RL) (years)

Trace	100				500				1000			
truncation level (%)	100	90	80	70	100	90	80	70	100	90	80	70
return period (year) 50	6	5	7	4	7	6	4	4	7	6	4	3
100	8	6	5	4	9	7	5	4	9	7	5	4
150	9	7	6	4	9	7	6	4	9	7	6	4
200	10	8	6	5	10	8	6	5	10	8	6	5
500	12	9	7	5	12	9	7	5	12	9	7	5

Table 3(b) Maximum run sum (RS) (cms-year)

Trace	100				500				1000			
truncation level (%)	100	90	80	70	100	90	80	70	100	90	80	70
return period (year) 50	92.32	65.04	44.17	30.08	93.62	64.00	43.51	29.36	92.41	63.39	43.22	29.33
100	108.2	74.48	50.99	34.65	111.8	75.77	51.27	34.13	111.8	75.62	50.89	34.31
150	119.5	80.50	53.75	36.30	121.8	81.83	55.19	36.69	121.8	81.44	54.76	36.88
200	127.9	85.19	57.09	38.87	128.5	85.67	57.67	38.32	128.5	85.29	57.22	38.45
500	157.1	103.4	67.44	43.94	153.2	99.65	66.10	43.98	150.9	99.06	65.39	43.72

Table 3(c) Maximum run intensity (RI) (cms)

Trace	100				500				1000			
truncation level (%)	100	90	80	70	100	90	80	70	100	90	80	70
return period (year) 50	21.80	19.14	15.88	12.51	21.32	18.47	15.39	12.15	21.36	18.57	15.50	12.22
100	23.26	20.05	16.76	13.6	23.25	20.22	17.01	13.62	23.3	20.41	17.19	13.86
150	24.92	21.56	18.22	15.1	24.43	21.26	18.04	14.71	24.54	21.48	18.25	14.85
200	25.51	22.16	19.13	15.87	25.21	22.09	18.76	15.46	25.19	22.07	18.76	15.35
500	27.49	24.14	20.78	17.31	27.36	24.03	20.57	17.1	27.22	23.92	20.44	16.93

As indicated by Kendall and Dracup (1991), the indexed sequence can make the most severe critical period when the beginning and end of the historic record is low flow periods. The drought characteristics using ISM method may be produced by the index $I=1$ and the historical drought period. For example, the longest run length for the historical annual inflow, $RL=3$ years. Thus $I=1$ and $n=3$ may represent the most severe drought characteristics for 22-year data. However, ISM cannot produce the most severities for the long term return periods because of the limitation for the critical period of the historical inflow series itself.

The modification of ISM may produce the most severe critical period that the generated data can exhibit. If one generates 66 inflow series in terms of $I=1$ and $n=3$ from 22-year inflow data set, it will produce $RL=5$ years. Therefore one can consider the historical longest run length $RL=$ 3-year as an ISM sequence and obtain the longest run length $RL=5$ years for 66-year return period. Successively, the synthetic annual inflow series is generated from 66-year data. That is, the66-year data may generate the maximum data $66\times5=330$ for the most severe critical drought of 330-year return period. The reason why this modification of ISM works for the most critical drought characteristics is that $I=1$ and $n=RL$ produce the most critical period of the generated data series without loosing the property of ISM which is overlapping historical data. Thus we modify the original ISM method to obtain the most critical drought characteristics. If one wants to obtain the drought characteristics for long term return periods, one can use 330-year data. In this study we generate 660-year data set from 330-year inflow series by using $n=2$ and $I=1$. The generated mean values by ISM and the modified ISM are shown in Table 4. The drought characteristics by the ISM are shown in table-5 and it does not represent the drought characteristics. We tried several ISM sequences and indices to obtain the reasonable results. However the results are similar as shown in Table 5.

Table 4. Mean values of ISM-generated inflow and the modified ISM-generated inflow
( ) is for the modified ISM-generated flow

Return period 〈yrs〉	66(66)	110(132)	220(198)	484(660)
Mean values 〈cms〉	42.58(42.58)	42.58(42.29)	42.58(42.29)	42.58(42.58)

Table 5. Drought characteristics for ISM-generated flow series

	RL(years)				*RS*(cms-year)				*RI*(cms)			
Truncation level (%)	100	90	80	70	100	90	80	70	100	90	80	70
return period (year) 66	5	5	5	5	109.3	88.02	66.73	45.44	22.98	18.72	14.72	10.21
110	4	4	4	3	75.13	58.09	41.06	27.82	22.98	18.72	14.47	10.21
220	4	4	3	3	66.15	53.37	40.6	27.82	22.05	17.79	14.47	10.21
484	5	5	4	3	75.01	53.72	40.6	27.82	22.98	18.72	14.47	10.21

Table 6. Drought characteristics of the modified ISM-generated flow series

	RL(years)				*RS*(cms-year)				*RI*(cms)			
Truncation level (%)	100	90	80	70	100	90	80	70	100	90	80	70
return period (year) 66	5	5	5	5	109.3	88.02	66.73	45.44	22.98	18.72	14.72	10.21
132	9	9	9	9	193.1	155.0	116.9	78.85	22.69	18.46	14.24	10.01
198	11	11	11	11	235.6	189.1	142.6	96.06	22.69	18.46	14.24	10.01
660	17	17	17	17	368.3	295.9	223.5	151.1	22.98	18.72	14.47	10.21

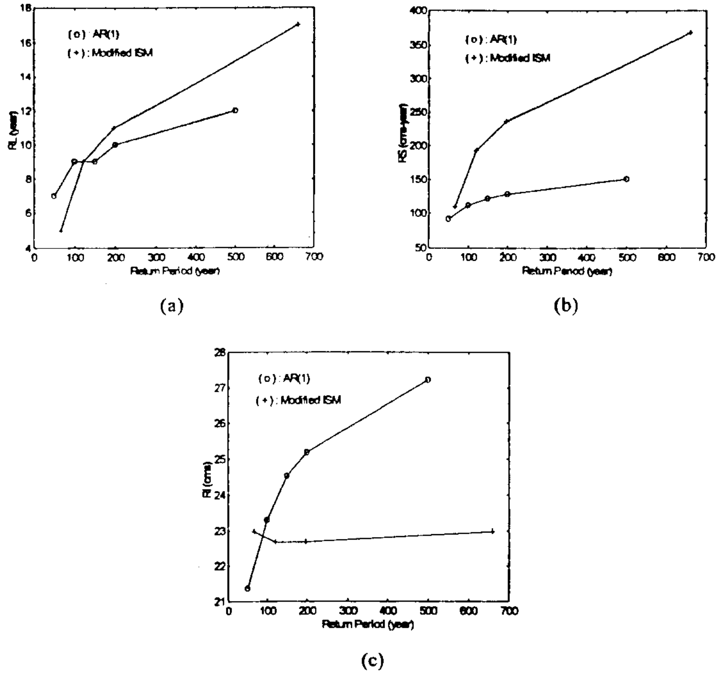

Fig. 5. Drought characteristics between AR(1) and modified ISM
(a) run length (b) run sum (c) run intensity

As one can see in Table 5, ISM method does not exhibit the drought characteristics as the return period is increased. Alternatively, the modified ISM is to represent the most critical drought characteristics which can be produced by the historical record, and results in Table 6. The comparison between AR(1) model and the modified ISM method is shown in Fig. 5. In this figure, AR(1) is for the 1000-trace and for the 100% truncation level, and the modified ISM method is for the 100% truncation level.

It is shown that RL and RS by the modified ISM are more critical than those of AR(1) as the return period is increased (Fig. 5). Specially, RS has large differences between the modified ISM and AR(1) with the return period. These differences may be due to the generation technique. As mentioned, the modified ISM produces the most critical drought characteristics by the historical record itself but AR(1) is using stochastically uncorrelated probabilistic distribution by statistical properties such as mean, standard deviation, and skewness etc. Therefore, AR(1) generates the synthetic data series from a linear relationship between one step past data series and probability distribution by statistical properties, and thus, AR(1) may be interpreted as the generation of the drought characteristics for the most probable situation. We distinguish AR(1) from the modified ISM as the most probable and the most critical situation of the drought characteristics. However, since the data generation with AR(1) is performed by a random manner, it may produce more critical droughts than the modified ISM for a relatively small return period. For instance, Fig. 5(a) shows RL=6-year for 50-year return period of AR(1), and RL= 5-year for 66-year return period of the modified ISM.

As the return period is increased, the RI by AR(1) is increased (Fig. 5). This means the probable situation of AR(1) makes the greater RI during a return period. In RI of the modified ISM, it shows that the increasing rates of RS and RL are almost the same during the return period, and this represents the proportional relationship between RS and RL. Thus, the modified ISM may be a more reasonable sense because RS and RL increase with the similar rates during the return period.

5. Conclusions

We compared the AR(1) model with the modified ISM method for determining drought characteristics. The modified ISM method produces more severe drought characteristics than AR(1) model as the return period increases. However, the original ISM method did not produce the drought characteristics well because of its limitation using the historical data set, even though the ISM method has been widely used for the generations of synthetic streamflows in the western USA. Therefore, we suggest the conceptually modfied ISM method for monitoring the most critical drought characteristics. The extremal drought conditions by the modified ISM may be used for the purposes of water supply strategies and compared with other models in practical applications. However, it remains some statistical arguments such as the correlations between the droughts of the historical data and those of the generated data even though the original ISM method does not also have any statistical debates.

Acknowledgment

We really thank to two anonymous reviewers. The indications and the comments of two reviewers were responsible for the significant improvements in the manuscript.

References

Bayazit, M., Distribution of Joint Run-Lengths of Bivariate Markov Processes, Journal of Hydrology, 50, pp 35-43, 1981.

Bras, R. L. and I. Rodriguez-Iturbe, Random Functions and Hydrology, Addison-Wesley Publ. Co., San Francisco, California, 1985.

Dracup, J. A., K. S. Lee and E. G. Paulson, On the definition of droughts, Water Resources Research, 16(2), pp 297-302, 1980.

Frick, D. M., D. Bode and J. D. Salas, Effect of Drought on Urban Water Supplies : I. Drought

Analysis, Journal of Hydraulic Engineering, Vol. 116, No. 6, June 1990.

House, P. M. and J. L. Ungvari, Indexed Sequential Modeling of Hydropower System Operations, Waterpower 83 : International Conference on Hydropower, Sept., 18-21, 1983.

Kendall, D. R. and J. A. Dracup, A Comparison of Index-Sequential and AR(1) Generated Hydrologic Sequences, Journal of Hydrology, 122, pp 335-352, 1991.

Labadie, J. W., D. G. Fontane, G. Q. Tabios III and N. F. Chou, Stochastic Analysis of Dependable Hydropower Capacity, Journal of Water Resources Planning and Management, Vol. 113, No. 3, May 1987.

Lee, K. S., J. Sadegehipour and J. A. Dracup, An Approach for Frequency Analysis of Multiyear Drought Durations, Water Resources Research, 22(5), pp 655-662, 1986.

Salas, J. D., J. W. Delleur, V. Yevjevich and W. L. Lane, Applied Modeling of Hydrologic Time Series, Water Resources Publications, Littleton, Colorado, 1980.

Tong, H., Non-linear Time Series : A Dynamical System Approach, Clarendon Press Oxford, 1990.

Wilhite, D. A. and M. H. Glantz, Understanding the Drought Phenomenon : The Role of Definitions, Water International, 10, pp 111-120, 1985.

Yevjevich, V., An Objective Approach to Definitions and Investigations of Continental Hydrologic Droughts, Hydrology Paper No. 23, Colorado State Univ., Fort Collins, Colorado, 1967.

Zelenhasic, E. and A. Salvai, A Method of Streamflow Drought Analysis, Water Resources Research, 23(1), pp 156-168, 1987.