A Simulation Technique of Runoff at Multistations

Kiyoshi HOSHI* and Isao YAMAOKA*

(Received April 28, 1973)

Abstract

This paper presents a mathematical model for generating hydrologic data when the structure of a hydrologic time series is characterized by the Markov process; the system inputs are the lag-zero and lag-one cross correlations. The procedure to estimate parameters is based on criterion in which vectors are mutually orthogonal. The adoption of an orthogonality facilitates the evaluation of regional differences of hydrologic characteristics over a large area where hydrologic events are far from homogeneous, and independent effects of the derived parameters are presented which improve on the shortcomings of the regression technique.

The system proposed herein ultimately leads to the solution of the eigenvalues and corresponding eigenvectors of a real symmetric matrix. Therefore, the computation procedure is simple and useful in practical hydrologic engineering.

The inverse transformation of a model equation produces properly correlated synthetic sequences of hydrologic data over any specified time length for any number of stations.

The model equation was applied for generating sequences of monthly runoff in the Ishikari River basin, Hokkaido, Japan. The results indicate that the statistical properties of the simulated data are in good agreement with those of the observed records and that the proposed technique would be useful for planning of water-resource systems in the simulation process.

The numerical analysis of this study was carried on the FACOM 230-60 system at the computing center of Hokkaido University.

1. Introduction

A simulation technique which generates synthetic sequences is particularly valuable since the complexity of water resources planning problems has increased. The purpose of this simulation is to use the derived data as inputs to subsequent hydrologic analysis, and to apply them to the determination and variability of the expected responses in water-resource system.

There is a need for a wider range of regional evaluation of the quantities of available water in place of local control to meet rapidly increasing future demands for water. For the purposes of the study which motivate this work, the ability to generate a synthetic sequence at a single station is of small importance. Whereas an adequate synthesis requires data to be simulated simultaneously at several stations, where their series are related to each other in an appropriate manner. The generation of such series is the subject of this paper.

When interest is limited to the simultaneous syntheses of runoff at multistations, the mathematical features of two models which have been proposed for generating synthetic sequences are examined. Matalas[1] presents a multi-site Markov chain model in which lag-one serial correlation coefficients and lag-zero cross

* Department of Civil Engineering, Hokkaido University, Sapporo, Japan.

correlations among the stations are introduced. His technique is basically equivalent to the least squares criterion in terms of estimation of the parameters, which however results in failing to evaluate independent effects of the obtained parameters and regional differences of hydrologic events over a wide area. Fiering[2] introduces a multivariate technique of principal component analysis which is based on finding all orthogonal or independent variates in the given data. His method preserves the lag-zero cross-correlation coefficients of the data matrix, but it does not preserve the lag-one serial correlation coefficients. Therefore, Fiering's approach fails to produce the structure of time dependence such as a first-order Markov process in the synthesized runoff data.

The study described in this paper is concerned with the combination of the above two methods, and demonstrates that the multivariate statistical analysis and simulation could produce satisfactorily accurate results with a minimum of time and effort expended.

2. Notation

The letter symbols adopted for use in this study are defined as follows and are assembled alphabetically:

$\boldsymbol{a}_i$ = structure vector and $(1\times n)$ matrix
$\boldsymbol{g}_i$ = ith composite variate and $(1\times N)$ matrix
$\boldsymbol{G}=(n\times N)$ matrix
$\boldsymbol{I}$ = identity matrix
m = sample mean
n = number of stations
N = number of observations
$\boldsymbol{0}$ = null vector
Q_i = simulated streamflow at the station i
r_{ij} = correlation coefficient between streamflows at stations i and j
$\boldsymbol{R}$ = lag-zero cross-correlation coefficients matrix
$\boldsymbol{R}_{21}$ = lag-one serial correlation coefficients matrix
t = random normal deviate with zero mean and unit variance
T = matrix transpose as a superscript
$tr.\ (\boldsymbol{Y})$ = trace equal to the sum of the diagonal elements of a matrix $\boldsymbol{Y}$
$\boldsymbol{w}_i$ = weight vector and $(1\times n)$ matrix
$\boldsymbol{W}=(n\times n)$ matrix
X = observed streamflow
$\boldsymbol{Y}^{-1}$ = inverse matrix of $\boldsymbol{Y}$
$\boldsymbol{Z}_k$ = standardized data array at n stations in the kth time period and $(n\times N)$ matrix
z = standard normal deviate of streamflow
λ_i, θ = Lagrange multipliers
σ = standard deviation
γ = skewness coefficient
ε = random number on a gamma distribution with zero mean and unit variance

3. Theoretical Background for the Analysis

The presence of time dependence, or persistence in a hydrologic time series implies repetition of information given by previous streamflows. Recognition of this property of hydrologic phenomena requires an adequate simulation model which at least should have the lag-zero and lag-one cross-correlation coefficients among streamflows at given stations.

Before the multivariate procedure is discussed, the historical record dealt with in this study is assumed to be transformed to a normal standard variate with zero mean and unit variance, and interest is limited to the time period of a month.

The system described in this paper is presented schematically in Fig. 1.

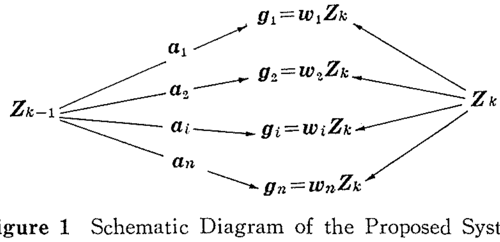

Figure 1 Schematic Diagram of the Proposed System

In this figure, $\boldsymbol{Z}_{k-1}$ and $\boldsymbol{Z}_k$ are the flows during the $(k-1)$ st and k th time period, respectively, $\boldsymbol{g}_i$ is the i th composite variate transformed from the flow in the k th time period, and $\boldsymbol{a}_i$ is defined as a structure vector herein, serving as a measure of the degree of association between streamflows in the $(k-1)$ st and the k th time periods.

The procedure to estimate weight vectors from the available records is subject to the conditions that composite variates should be mutually orthogonal and each element of a structure vector should have a correlation coefficient as high as possible in each composite variate.

As indicated in Fig. 1, a model for generating simultaneous synthetic sequences of monthly streamflow at the stations is defined as

$$\boldsymbol{G}=\boldsymbol{W}\boldsymbol{Z}_k \tag{1}$$

where

$$\boldsymbol{G}=[\boldsymbol{g}_1\boldsymbol{g}_2\boldsymbol{g}_3 \cdots \boldsymbol{g}_n]^T \tag{2}$$

$$\boldsymbol{W}=[\boldsymbol{w}_1\boldsymbol{w}_2\boldsymbol{w}_3 \ldots \boldsymbol{w}_n]^T \tag{3}$$

$$\boldsymbol{g}_i=\boldsymbol{w}_i\boldsymbol{Z}_k \tag{4}$$

Before the detailed solution of the model equation is examined, several important properties with respect to Eq. (4) are presented. Since $\boldsymbol{g}_i$ should be standardized with unit variance by definition, postmultiplying Eq. (4) by $\boldsymbol{g}_i^T/N$ gives

$$\boldsymbol{w}_i\boldsymbol{R}\boldsymbol{w}_i^T=1 \tag{5}$$

where

$$\boldsymbol{R}=\boldsymbol{Z}_k\boldsymbol{Z}_k^T/N \tag{6}$$

To obtain a measure of the degree of association between $\boldsymbol{g}_i$ and $\boldsymbol{Z}_{k-1}$, postmultiplying Eq. (4) by $\boldsymbol{Z}_{k-1}^T$ gives

$$\boldsymbol{a}_i=\boldsymbol{w}_i\boldsymbol{R}_{21} \tag{7}$$

where

$$\boldsymbol{a}_i=\boldsymbol{g}_i\boldsymbol{Z}_{k-1}^T/N \tag{8}$$

$$\boldsymbol{R}_{21}=\boldsymbol{Z}_k\boldsymbol{Z}_{k-1}^T/N \tag{9}$$

$\boldsymbol{a}_i$ is a structure vector consisting of correlation coefficients between $\boldsymbol{g}_i$ and $\boldsymbol{Z}_{k-1}$ $(-1 \leq \boldsymbol{a}_i \leq 1)$. The structure vector is a useful parameter to account for the structure of dependence between monthly flows in the $(k-1)$ st and the k th time periods.

The sum of the squared elements of a structure vector indicates the total contribution of the previous month flows to the composite variate transformed from the current month flows at the stations. This sum is given by

$$V_i=\boldsymbol{a}_i\boldsymbol{a}_i^T=\boldsymbol{w}_i\boldsymbol{R}_{21}\boldsymbol{R}_{21}^T\boldsymbol{w}_i^T \tag{10}$$

To extract regional differences of hydrologic events over a wide area from the available records, an orthogonal condition where two composite variates are uncorrelated is introduced. This expression is given by

$$\boldsymbol{g}_i\boldsymbol{g}_j^T/N=\boldsymbol{w}_i\boldsymbol{R}\boldsymbol{w}_j^T=0 \quad (i\neq j) \tag{11}$$

3-1 Solution of the First Composite Variate

The procedure to estimate a weight vector in the first composite variate is to choose $\boldsymbol{w}_1$ in such a way as to make V_1 a maximum under the condition of Eq. (5). This criterion is written as follows:

$$2T_1=\boldsymbol{w}_1\boldsymbol{R}_{21}\boldsymbol{R}_{21}^T\boldsymbol{w}_1^T-\lambda_1(\boldsymbol{w}_1\boldsymbol{R}\boldsymbol{w}_1^T-1) \tag{12}$$

Then, set the partial derivative of this new function T_1 with respect to each element of $\boldsymbol{w}_1$ equal to zero, that is,

$$(\boldsymbol{R}_{21}\boldsymbol{R}_{21}^T-\lambda_1\boldsymbol{R})\boldsymbol{w}_1^T=\boldsymbol{O} \tag{13}$$

Premultiplying Eq. (13) by $\boldsymbol{w}_1$ with use of Eqs. (5) and (10) gives

$$V_1=\lambda_1 \tag{14}$$

The maximization of Eq. (10) under the condition of Eq. (5) leads to the system for the solution of an eigenvalue (λ_1) and corresponding eigenvector ($\boldsymbol{w}_1$) of the real symmetric matrices of $\boldsymbol{R}_{21}\boldsymbol{R}_{21}^T$ and $\boldsymbol{R}$. Equation (14) implies that an eigenvalue of λ_1 is precisely V_1, the quantity which is to be maximized. In other words, V_1 is equal to the largest one of the roots of the characteristic equation.

For convenience of interpretation, Eq. (13) is transformed to the system for the solution in terms of the structure vector of $\boldsymbol{a}_1$. Substituting Eq. (7) into Eq. (13), and then premultiplying by the inverse of $\boldsymbol{R}$ gives

$$\boldsymbol{R}^{-1}\boldsymbol{R}_{21}\boldsymbol{a}_1^T-\lambda_1\boldsymbol{w}_1^T=\boldsymbol{O} \tag{15}$$

where $\boldsymbol{R}^{-1}$ denotes the inverse matrix of $\boldsymbol{R}$. Since the lag-zero cross-correlation matrix of $\boldsymbol{R}$ is nonsingular, it has an inverse.

Premultiplying Eq. (15) by the transpose of $\boldsymbol{R}_{21}$ gives

$$(\boldsymbol{R}_{21}^T\boldsymbol{R}^{-1}\boldsymbol{R}_{21}-\lambda_1\boldsymbol{I})\boldsymbol{a}_1^T=\boldsymbol{O} \tag{16}$$

Equation (16) is a commonly used expression for the solution of an eigenvalue and corresponding eigenvector of a real symmetric matrix of $\boldsymbol{R}_{21}^T\boldsymbol{R}^{-1}\boldsymbol{R}_{21}$.

A weight vector in the first composite variate is derived from Eqs. (7) and (16) as follows:

$$\boldsymbol{w}_1=\boldsymbol{a}_1\boldsymbol{R}_{21}^T\boldsymbol{R}^{-1}/\lambda_1 \tag{17}$$

3-2 Solution of Other Composite Variates

The procedure to estimate the weight vectors in the other composite variates presupposes that the quantity of Eq. (10) should be maximized under the conditions of Eqs. (5) and (11). The expressions necessary for obtaining the weight vector of $\boldsymbol{w}_j$ ($j>i$) in Eq. (4) corresponding to Eqs. (12) and (13) are derived as follow:

$$2T_j=\boldsymbol{w}_j\boldsymbol{R}_{21}\boldsymbol{R}_{21}^T\boldsymbol{w}_j^T-\lambda_j(\boldsymbol{w}_j\boldsymbol{R}\boldsymbol{w}_j^T-1)-2\theta\boldsymbol{w}_i\boldsymbol{R}\boldsymbol{w}_j^T \quad (j>i) \tag{18}$$

$$\boldsymbol{R}_{21}\boldsymbol{R}_{21}^T\boldsymbol{w}_j^T-\lambda_j\boldsymbol{R}\boldsymbol{w}_j^T-\theta\boldsymbol{R}\boldsymbol{w}_i^T=\boldsymbol{O} \quad (j>i) \tag{19}$$

Premultiplying Eq. (19) by $\boldsymbol{w}_j$ and using Eqs. (5), (10) and (11) yield

$$V_j=\lambda_j=\boldsymbol{a}_j\boldsymbol{a}_j^T \tag{20}$$

Premultiplying Eq. (19) by $\boldsymbol{w}_i$ ($i\neq j$) yields

$$\theta=\boldsymbol{w}_i\boldsymbol{R}_{21}\boldsymbol{R}_{21}^T\boldsymbol{w}_j^T=\boldsymbol{a}_i\boldsymbol{a}_j^T \quad (i\neq j) \tag{21}$$

Assuming that $i=1$ and $j=2$, then substituting Eq. (13) into Eq. (21) yields

$$\theta = \boldsymbol{a}_1 \boldsymbol{a}_2^T = \lambda_1 \boldsymbol{w}_1 \boldsymbol{R} \boldsymbol{w}_2^T = 0 \tag{22}$$

In general, Eq. (22) holds for the successive pairs of i and j. Therefore, Eq. (19) reduces to

$$\boldsymbol{R}_{21} \boldsymbol{R}_{21}^T \boldsymbol{w}_j^T - \lambda_j \boldsymbol{R} \boldsymbol{w}_j^T = \boldsymbol{O} \tag{23}$$

Equation (23) is explicity identical with Eq. (13) in a mathematical notation.

An expression corresponding to Eq. (16) is also derived as follows:

$$(\boldsymbol{R}_{21}^T \boldsymbol{R}^{-1} \boldsymbol{R}_{21} - \lambda_i \boldsymbol{I}) \boldsymbol{a}_j^T = \boldsymbol{O} \tag{24}$$

From Eqs. (20) and (24) it is clear that λ_j and $\boldsymbol{a}_j$ are the j th largest eigenvalue, and its associated eigenvector of the symmetric matrix $\boldsymbol{R}_{21}^T \boldsymbol{R}^{-1} \boldsymbol{R}_{21}$, respectively. By the same token as in the first composite variate, a weight vector of the j th composite variate is written as follows:

$$\boldsymbol{w}_j = \boldsymbol{a}_j \boldsymbol{R}_{21}^T \boldsymbol{R}^{-1} / \lambda_j \tag{25}$$

The advantage of the technique proposed herein lies in the fact that all the weight vectors of Eq. (4) are obtained directly from the successive eigenvalues and their associated eigenvectors of the original matrix $\boldsymbol{R}_{21}^T \boldsymbol{R}^{-1} \boldsymbol{R}_{21}$. The elements of each structure vector in Eq. (24) must be normalized to satisfy the relation of Eq. (20), in such a way that they are divided by the square root of the sum of their squares and then multiplied by $\sqrt{\lambda_j}$.

Since $\boldsymbol{R}_{21}^T \boldsymbol{R}^{-1} \boldsymbol{R}_{21}$ is a real symmetric matrix, the eigenvalues are all real, positive, and distinct. And the properties with respect to the structure vectors and the weight vectors are summarized as follow:

$$\boldsymbol{a}_i \boldsymbol{a}_j^T = \begin{Bmatrix} \lambda_i & i=j \\ 0 & i \neq j \end{Bmatrix} \quad (i, j = 1, 2, 3, \dots, n) \tag{26}$$

and

$$\boldsymbol{w}_i \boldsymbol{R} \boldsymbol{w}_j^T = \begin{Bmatrix} 1 & i=j \\ 0 & i \neq j \end{Bmatrix} \quad (i, j = 1, 2, 3, \dots, n) \tag{27}$$

It is noted that Eqs. (26) and (27) represent special instances of an orthogonality property. The orthogonality is a useful index for extracting from the available data regional differences of hydrologic characteristics over a fairly large area, and for evaluating independent effects of the structure vectors and the weight vectors in the model equation which would improve on the shortcomings of a multiple regression approach.

A property with respect to the eigenvalue λ_i is derived as follows:

$$\sum_{i=1}^{n} \lambda_i = \mathrm{tr.}\ (\boldsymbol{R}_{21}^T \boldsymbol{R}^{-1} \boldsymbol{R}_{21}) \tag{28}$$

The λ_i means the variance of each composite variate, and the sum of the variances of all n composite variates is equal to the trace. Since the trace is the total variance to be accounted for, the cumulative sum of the eigenvalues divided by the trace is the proportion of variance accounted for by the resulting composite variates.

3-3 Simulation Approach

By Eqs. (24) and (25), the weight vectors of all composite variates are estimated from the historical records of monthly runoff during the $(k-1)$ st and the k th time periods. From Eq. (27) the row vectors of $\boldsymbol{W}$ in Eq. (1) are linearly independent, so that the determinant of $\boldsymbol{W}$ is different from zero and therefore $\boldsymbol{W}$ has an inverse. Premultiplying Eq. (1) by $\boldsymbol{W}^{-1}$ gives

$$\boldsymbol{Z}_k = \boldsymbol{W}^{-1} \boldsymbol{G} \tag{29}$$

Equation (29) is a simultaneous simulation model of monthly runoff in the k th time period, in which the information of the lag-zero and lag-one cross correlations among the stations is employed. Numerical procedures are quite simple and consist only of generating random numbers of G because n composite variates are mutually orthogonal (uncorrelated) by Eq. (11).

4. Practical Application

In order to test the validity of the model equation developed in the preceding section, the upper basin of the Ishikari River in Hokkaido was selected. Figure 2 shows the location of the selected basin and four water gauging stations with the station numbers assigned. Table 1 gives the drainage area of the chosen stations.

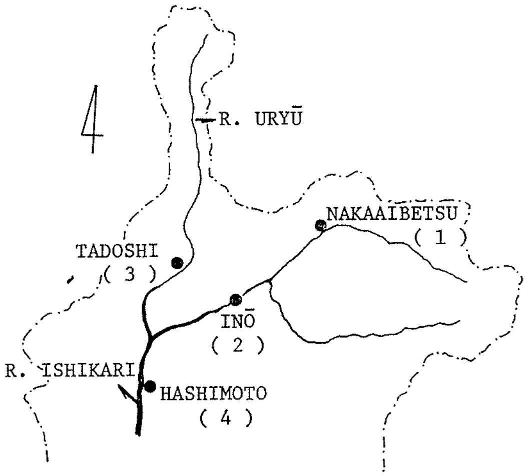

Figure 2 Location Map of Selected Basin and Stations

Table 1 Adopted Stations

No.	Water Gauge	Drainage Area [km^2]
1	NAKAAIBETSU	1082.5
2	INŌ	3378.6
3	TADOSHI	1015.5
4	HASHIMOTO	5460.0

In the practical study, the application for the syntheses of monthly streamflow was made from April through December because records are unavailable or inadequate during the winter months.

Simultaneous records of monthly total flow are available for a period of 19 years (1951–1970). To preserve the first three moments in frequency distribution of the synthesized monthly flow and also to eliminate the negative flows that occur occasionally in the generating process, the following transformation by Log Pearson Type III distribution[3] was applied to yield normal flow variates:

$$z=\frac{6}{\gamma}\left\{\left[\frac{\gamma}{2}\left(\frac{\log_e X-m}{\sigma}\right)+1\right]^{1/3}-1\right\}+\frac{\gamma}{6} \tag{30}$$

where m, σ, and γ are mean, standard deviation, and skewness cofficient of the monthly flow logarithm ($\log_e X$), respectively.

Some results by the multivariate technique are presented. Table 2 gives the statistical quantities of monthly streamflow in terms of structure vectors and eigenvalues in four composite variates.

In this table, the column vector gives an element of a structure vector in each composite variate, and the variance of each composite variate (sum of squares in each column) is shown in the next to last row. The last row gives the percentage contribution of each composite variate to the total variance of monthly flow where each eigenvalue is divided by the trace of $\boldsymbol{R}_{21}^T\boldsymbol{R}^{-1}\boldsymbol{R}_{21}$.

Some characteristics with respect to the regional and seasonal variations of

Table 2-A
Structure Vectors and Contributions of Four Composite Variates (May)

	a_1	a_2	a_3	a_4
1	−.384	−.261	−.077	.068
2	−.140	.518	.183	.035
3	.293	−.334	.302	.014
4	.681	.103	−.136	.039
λ_i	.716	.459	.149	.008
$100\,\lambda_i/\alpha$	53.75	34.45	11.18	0.06

$$\alpha=\sum_{i=1}^{4}\lambda_i=1.332$$

Table 2-B
Structure Vectors and Contributions of Four Composite Variates (October)

	a_1	a_2	a_3	a_4
1	.759	.275	−.026	−.003
2	.675	−.176	−.043	−.017
3	.624	−.100	−.050	.020
4	.634	−.044	.125	.002
λ_i	1.823	.119	.021	.001
$100\,\lambda_i/\alpha$	92.82	6.05	1.06	0.05

$$\alpha=\sum_{i=1}^{4}\lambda_i=1.964$$

monthly streamflow are:

(a) The contribution of the first component to the total variance is the largest in every month, indicating a general factor of runoff characteristics among the stations. However, it shows a low value (54%) in May, while it has a higher value (93%) in October.

(b) A structure vector of the first composite variate has large positive element in October, causing λ_1 to give a large contribution to the total variance. This also means that all elements of a matrix of $\boldsymbol{R}_{21}^T\boldsymbol{R}^{-1}\boldsymbol{R}_{21}$ are positive. On the other hand, during the spring, a structure vector has a wide range of values in the first composite variate, resulting in λ_1 giving a small contribution.

(c) During the spring, the second and third composite variates show higher contributions since elements of the structure vectors are large at the specific stations.

(d) From the above results, no regional differences of runoff are recognized during the summer and autumn seasons because precipitation is uniformly distributed over a wide area. On the other hand, runoff during the spring season reveals the regional variations inherent in the individual basins because snowmelt is influenced by the retarded precipitation stored in the form of snowcover and released with the thawing temperature.

Monthly streamflows at 4 stations were simultaneously synthesized by Eq. (29), in which the elements of $\boldsymbol{W}$ were estimated from the available records of 19 years by use of Eqs. (24) and (25). In the practical study, the mutually orthogonal composite variates were somewhat skewed even though monthly flow was approximately normalized by a Log Pearson Type III transformation of Eq. (30). To overcome this situation in the simulation process, normal random numbers on a gamma distribution were generated as follows[1]:

$$\varepsilon=\frac{2}{\gamma}\left(1+\frac{\gamma t}{6}-\frac{\gamma^2}{36}\right)^3-\frac{2}{\gamma} \tag{31}$$

where γ is a skewness coefficient in the i th composite variate of Eq. (4).

The standardized synthetic data of monthly flows were simulated using an estimation of $\boldsymbol{W}$ and random numbers of $\boldsymbol{G}$ on Eq. (31) as defined in Eq. (29), and then the inverse transform by solving Eq. (30) for X must be used to calculate the synthetic monthly flows in raw form.

An evaluation of a multi-site simulation model was made by comparing all

Table 3 Statistics of the Observed and Svnthesized Monthly Streamflow

Month	Station	NAKAAIBTSU m	NAKAAIBTSU σ	NAKAAIBTSU γ	INŌ m	INŌ σ	INŌ γ
Apr.	obs.	1820.5	457.8	−0.495	8720.6	2157.0	0.532
	sim.	1789.7	571.7	0.648	8686.5	2381.9	0.935
May	obs.	4146.0	1158.8	0.944	8428.8	2911.9	0.878
	sim.	3846.7	1060.0	0.967	7521.1	2822.7	1.310
Jun.	obs.	2342.7	669.3	1.230	4901.6	1210.4	0.909
	sim.	2297.1	654.6	1.120	4867.3	1101.8	0.550
Jul.	obs.	1530.5	363.7	0.538	3606.6	1184.0	0.526
	sim.	1579.1	367.2	0.759	3611.5	1572.4	1.580
Aug.	obs.	1920.1	949.9	1.130	5742.5	3090.7	0.643
	sim.	1734.9	668.6	0.835	5256.7	2569.0	1.200
Sep.	obs.	1381.1	492.6	−0.110	3992.5	1326.8	0.436
	sim.	1408.1	608.5	1.320	3982.8	1404.9	1.280
Oct.	obs.	1254.6	436.0	0.643	3825.8	1907.7	2.040
	sim.	1199.8	459.0	0.730	3545.9	1694.7	1.910
Nov.	obs.	1046.5	332.9	−0.148	3411.5	1113.3	0.936
	sim.	1030.8	408.2	1.180	3317.9	900.2	0.528
Dec.	obs.	752.7	198.9	−0.576	2445.9	521.4	0.191
	sim.	739.0	208.9	−0.012	2350.5	452.1	0.193

Table 3 (Cont.) Statistics of the Observed and Synthesized Monthly Streamflow

Month	Station	TADOSHI m	TADOSHI σ	TADOSHI γ	HASHIMOTO m	HASHIMOTO σ	HASHIMOTO γ
Apr.	obs.	4002.3	1504.2	−0.038	18259.4	4346.4	0.221
	sim.	4319.7	3088.0	1.720	19138.2	6360.1	1.490
May	obs.	3963.9	2410.9	0.974	15464.7	5483.1	−0.238
	sim.	3383.3	2151.7	1.270	16329.6	9136.8	2.360
Jun.	obs.	795.1	320.4	1.100	6199.5	1515.5	−0.358
	sim.	810.6	275.4	1.240	7074.7	1783.6	0.718
Jul.	obs.	859.9	542.8	2.510	4975.5	2350.8	1.530
	sim.	957.0	561.9	1.550	5641.3	2899.2	1.500
Aug.	obs.	1270.0	820.8	0.965	8689.2	4929.6	0.574
	sim.	1260.2	889.4	1.920	8625.9	4326.7	1.180
Sep.	obs.	943.9	395.3	0.666	6095.5	2353.3	0.743
	sim.	938.7	409.0	0.469	6338.7	2501.9	1.470
Oct.	obs.	972.8	372.7	0.394	5612.1	2448,3	1.890
	sim.	929,6	354.6	0.491	5563.9	2524.5	1.900
Nov.	obs.	1123.5	471.6	0.587	5640.7	1920.6	1.050
	sim.	1153.1	458.4	0.956	5624.3	1480.4	0.422
Dec.	obs.	869.6	396.9	0.505	4204.3	1190.7	1.100
	sim.	778.0	445.6	1.850	4392.9	1078.1	0.798

m; Mean of Monthly Flow [m^3/sec]
σ; Standard Deviation of Monthly Flow [m^3/sec]
γ; Skewness Coefficient of Monthly Flow
obs.; Observed Data
sim.; Simulated Data

Table 4 Correlation Coefficients of Monthly Streamflow

Month	Corre. Co.	r_{12}	r_{13}	r_{14}	r_{23}	r_{24}	r_{34}
Apr.	obs.	.313	.088	.157	−.176	.322	.404
	sim.	.386	.007	.305	−.143	.534	.266
May	obs.	.912	.653	.622	.668	.652	.775
	sim.	.902	.634	.659	.727	.743	.595
Jun.	obs.	.860	.329	.481	.366	.634	.480
	sim.	.846	.333	.632	.313	.800	.294
Jul.	obs.	.740	.495	.522	.675	.771	.853
	sim.	.714	.699	.686	.893	.943	.916
Aug.	obs.	.898	.662	.691	.784	.829	.827
	sim.	.878	.628	.871	.752	.909	.831
Sep.	obs.	.860	.637	.793	.793	.902	.787
	sim.	.850	.572	.776	.695	.932	.673
Oct.	obs.	.836	.602	.870	.573	.844	.626
	sim.	.692	.567	.686	.776	.948	.772
Nov.	obs.	.693	.493	.552	.651	.806	.618
	sim.	.589	.355	.502	.408	.825	.515
Dec.	obs.	.263	−.079	−.499	.254	.181	.221
	sim.	.259	−.384	−.490	.103	.190	.543

r_{ij}; Correlation Coefficient between Flows at Stations i and j

essential statistics of the observed records with corresponding statistics of the simulated data for 90 years of operation.

The generated monthly streamflows at four stations were numerically tested on the continuity conditions to be satisfied :

$$Q_2 > Q_1 \text{ and } Q_4 > Q_2 + Q_3 \tag{32}$$

where numbers correspond to the station numbers in Fig. 2.

Table 3 gives the means (m), standard deviations (σ) (m^3/sec), and skewness coefficients (γ) of the observed (obs.) and simulated (sim.) monthly flows at four stations. Table 4 gives the correlation matrices of their data among the stations. The statistical properties of the observed records are reasonably consistent with those derived from the simulated data.

The skewness coefficients of the observed records would not resemble those of the synthesized data due to the small sample size of monthly observations. Since the higher-order moments in frequency distribution are sensitive to the sample size, it is possible to preserve skewness coefficients if the records for a longer period are available.

5. Conclusion

The multivariate technique is presented for the analysis and synthesis of runoff where the system inputs are lag-zero and lag-one cross correlations. An outstanding feature of the model equation is that the original records are initially transformed to a set of orthogonal composite variates, and the resulting coefficients measure the degree of dependence between the composite variate and the given independent variables. In contrast to the least-squares method, the proposed technique provides a means for evaluating independent effects of the parameters.

The inverse transform of the model equation provides a means for synthesizing correlated hydrologic sequences over any specified time length for any number of stations.

Acknowledgment

The authors wish to express their appreciation to the Hokkaido Development Bureau at Sapporo which has very kindly supplied the valuable data used in this study.

References

1) Matalas, N. C.: Mathematical Assessment of Synthetic Hydrology, Water Resources Res., Vol. 3, No. 4, p. 937–945, 1967.
2) Fiering, M. B.: Multivariate Technique for Synthetic Hydrology, J. Hydraulic Div., ASCE, HY 5, p. 43–60, 1964.
3) Beard, L. R.: Use of Interrelated Records to Simulate Streamflow, J. Hydraulic Div., ASCE, HY 5, p. 13–22, 1965.