RESEARCH ARTICLE OPEN ACCESS

Synthetic Flow Generation

Chetansinh R. Vaghela*, Ajaysinh R. Vaghela**
*(Assistant professor, Mathematics department, MEFGI, Rajkot, Gujarat, India)
** (Department of Applied mechanics, The M.S. University of baroda, India)

ABSTRACT
We are given the historical data of inflow of river observed on Karjan Dam near Vadodara with sampling time interval of a month. We are required to know the inflow of the river in future time using this available historical data. That is to generate inflow of river for future time using the model namely 'Single Site Seasonal Thomas Fiering Model'. Also, we want to find the Statistical Parameters like mean, standard deviation and Correlation coefficient of generated data and compare it with historical data. Once we have the predicted inflow data we may use these prediction of inflow into a storage reservoir for planning the operation of the reservoir to ensure maximum benefit from a limited quantity of stored water.
Keywords – MATLAB, Thomas Fiering method, etc.

I. INTRODUCTION

Water is one of man's basic and precious resources. It plays a vital role in agriculture, industry, navigation and production of energy. The main problem is not often found when and where it is needed, and it may not be of good quality. So, the prediction of inflow into a storage reservoir is needed for planning the operation of the reservoir to ensure maximum benefit from a limited quantity of stored water. The development of a runoff prediction model to assess monthly and yearly inflows in advance allows an operational procedure to be formulated which can be modified on the basis of inflows observed subsequently.The monthly flow series are non-stationary and therefore complicated mathematical models are employed in their simulation. The first model that appeared in the hydrology literature for the generation of synthetic monthly flow sequences is that due to Thomas & Fiering (1962). Basically, this model is having periodic parameters, namely, the monthly means, standard deviations and the lag-zero cross correlations between successive months. The model implicitly allows for the non-stationarity observed in monthly flow data.

1.1 Single Site Seasonal Model

Any hydrologic series observed with a sampling time interval of less than year would inevitably be non-stationary in structure because of cyclic component with a period of one year introduced into it. Therefore seasonal models have to take care of this non-stationarity.

The essential procedure in generation of model is selecting correct type of model for describing the process and its parameter are to be properly estimated so that important stochastic structure noticed in the observed function is preserved in the generation model.

If the model is used to generate synthetic sequence such a sequence should be identical in a statistical sense with corresponding parameters like mean, standard deviation etc, with its historical sequence.

Apart from preserving the important statistical parameter, the model should also preserve the probability distribution of observed flows.

When a model is built to satisfy all the above conditions it can be used to generate any number of sequence and there is no difference between the generated sequence and observed sequence and they are likely to occur as the observed sequence and can be used wherever an observed sequence is used.

1.2 Quality of the synthetic flow record

A Synthetic flow record is only as good as the flow statistics used to produce it. A user of a synthetic flow record should be satisfied that the statistics comply with the following conditions:
(1)They should be adequately defined
(2)They should describe the characteristics of flow
(3)They should be modified if necessary to account for any future changes in the hydrological properties of the catchment.

1.3 Adequate definition

If the statistics of an observed record are not adequately defined, the synthetic flows recode may not adequately represent the behavior of the river.

For this reason mean monthly flows rather than mean daily flows are generated, since daily flows for a given date will show more variation from year to tear than corresponding mean monthly flows. For example, the standard error of mean flow on the first day of January will be greater than the standard error of the mean January flow.

These types of models can also be used to generate monthly inflows by using the monthly historical inflow records.

Statistics of monthly flow of the river for which a record is to be generated may be obtained by one of three methods, given in decreasing order of probable reliability:

(1) By analysis of an observed flow record for the river,

(2) By analysis of similar rivers in the region,

(3) By analysis of the monthly water balance.

We will use first Method as it is more reliable.

II. SYNTHETIC FLOW GENERATION

2.1 Construction of the model

In order to realize the two relationships required from a multiplicative ARIMA model, it is helpful to arrange the monthly flows in a two dimensional array as is shown in Following Fig. 1.

The columns of the array represent the months whereas the rows represent the years. If an observation which is notationally characterized by is considered, then it is possible to express this observation in terms of and plus a random shock. Therefore

$$\sigma = \alpha \; + \beta j + € \; i,j \qquad (1)$$

Where αj and βj are model parameters which reflect the relationships between successive months of the same year and between successive years of the same month, respectively.

Finally, €i,j is a random variable independent and .Moreover, α, βj and the variance, €j, of the random variable €i,j, are all periodic due to the periodicity in the monthly data. By considering $\alpha = 0$ and ignoring the subscript j in Equation (1), one can obtain the lag-one Markov process which is currently used in the simulation of annual flow data. On the other hand, when $\beta j = 0$, Equation (1) reduces to the Thomas-Fiering model which takes into account the relationship between successive months only.

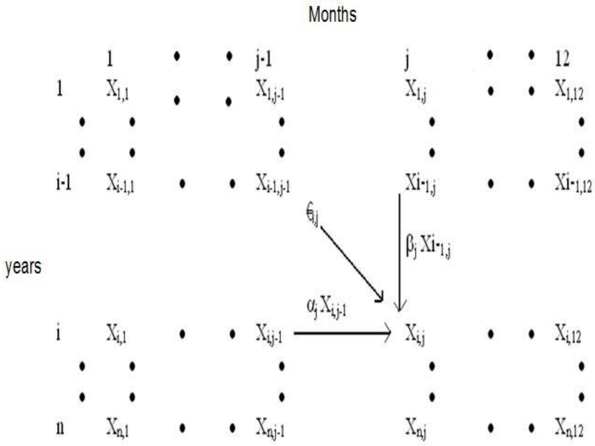


Fig 1: Two-dimensional array of monthly flows

The parameters to be preserved in Equation (1) are as follows:

(1) 12 monthly means, μ; j = 1, 2. . . 12

(2) 12monthly standard deviations, σj where j = 1, 2, 12

(3) 12 first order serial correlation coefficients, one for each month

ρj (l) ; where j = 1 , 2 , . . . , 12

(4) 12 lag-zero cross-correlation coefficients between successive

Months, ρj, j-1(0) where j = 1, 2... 12.

Thus, there are 48 parameters to be estimated from a given historic sequence of monthly flows. Apart from the means, there are 36 model parameters which must be calculated in terms of historic parameters obtained from the data available. The model parameters are α, βj and σ€j.

2.2 Calculation of model parameters

The model parameters must be estimated from given data of monthly flows in such a way that the generated synthetic sequences yield the same parameters in the long run.

In order to preserve the monthly means, it is convenient to rewrite Equation (1) as

$$xi,j - \mu j = \alpha\, j\, (xi,j\text{-}1 - \mu j\text{-}1\,) + \beta j\, (xi,j - \mu j\,) + € \; i,j \quad (2)$$

By assuming E (€ i,j) = 0 one can obtain E(xi,j) = μj , which shows that monthly means are preserved by the model. The introduction of a new random variable

$$w\, i,j = xi,j - \mu j$$

Converts Equation (2) into the following form

$$wi,j = \mu j\; wi,j\text{-}1 + \beta j\; w\; i\text{-}1,j + € \; i,j \qquad (3)$$

This last expression is quite different from Equation (1) in that Equation (3) has an expected value which is equal to zero. That is,
E(wi,j) = E(wi,j-1) = E(wi-1,j) = E(€i,j) = 0

The multiplication of Equation (3) by w i,j and then the application of the expectation operation to both sides leads to
E(w2 i,j) = μ j E(w i,j-1 w i,j) + βj E(w i-1,j w i,j) + E(€ i,j w i,j) (4)

Where
E(w2 i,j) = E(x2i,j) - μ j 2 = σj2
E(w i,j-1 w i,j) = ρj,j-1 (0) σj σj-1
E(w i-1,j w i,j) = ρj(1) σj2
and
E(€ i,j w i,j) = E(€2 i,j) = σ2€j

The substitution of these last expressions in Equation (4) yields
σj2 = αj ρj,j-1 (0) σj σj-1 + βj ρj(1) σj2 + σ2€j (5)

In order to be able to calculate correlation along years first of all it is necessary to multiply both sides of Equation (3) by wi-1,j and then the application of the expectation operation leads to

E(wi,j wi-1,j) = α E(wi,j-1 wi-1,j) + βj E(wi-1,j wi-1,j) + E(€ i,j wi-1,j)

Or in terms of moments after some algebra,
ρj(l) σj = α ρj,j-1(1) σj-1 + βjσj (6)

Where ρj,j-1(1) is the lag-one cross-correlation coefficient between successive months. Finally, by virtue of the correlation along months the following expression can be obtained:
ρj,j-1 (0) σj = ασj-1+ βj ρj,j-1(1) σj (7)

The three unknown model parameters α , βj and σ€j can be obtained from the simultaneous solution of Equations (5), (6) and (7) which leads to,

α= [(ρj,j-1 (0) - ρj(l) ρj,j-1(1))/(1 - ρ2j,j-1(1))] $\frac{\sigma_j}{\sigma_{j-1}}$ (8)

βj = ρj(1) - $\frac{\rho_{j,j-1}(0)-\rho_j(1)\rho_{j,j-1}(1)}{1-\rho^2{}_{j,j-1}(1)}\frac{\sigma_j}{\sigma_{j-1}}$ (9)

And

σ2€j = σj2 - $\frac{\sigma_j{}^2}{1-\rho^2{}_{j,j-1}(1)}$ [ρ2j,j-1(0) -2 ρj(l) ρj,j-1(0) ρj,j-1(1) + ρ2j(1)] (10)

The incorporation of these model parameters into Equation (1) gives synthetic sequences which resemble the historic sequence in terms of
μj , σj, ρj(1) and ρj,j-1(0).

As a result, ρj,j-1(0) is a dummy parameter which can be defined to have an arbitrary but convenient value and together with values of α , βj and σ2Єj. Obtained, they will still allow the preservation of σj, ρj,j-1(0) and , ρj,(1).
It is interesting to notice various special cases of equations (8), (9) and (10).
For instance, when both ρj,j-1(0) and ρj,j-1(1) are equal to zero, i.e. α = 0, then equation (9) and equation (10) become

Bj = ρj(1) (11)

and

σ€j = σj2 [1 - ρ2j(1)] (12)

Respectively. Finally, for this special case Equation (1) becomes,

$w_{i,j} = \rho_j(1)w_{i-1,j} + \sigma_j\ [1 - \rho^2{}_j(1)]^{1/2}\ €_{i\,j}$ (13)

This is the lag-one Markov process provided that the subscripts are dropped.
A second special case can be derived by assuming βj = 0. Hence, from Equation (9) it is possible to observe that,

ρj(l) = ρj,j-1(0) ρj,j-1(1)

The substitution of which into Equations (7) and (10) gives,

αj = ρj,j-1 (0) $\frac{\sigma_j}{\sigma_{j-1}}$ (14)

and

σ2€j = σj2 [1 – ρ2j,j-1 (0)] (15)

Respectively, the incorporation of these new sets of parameters into Equation (1) results in

w i,j = ρj,j-1 (0) $\frac{\sigma_j}{\sigma_{j-1}}$ w i,j-1 + σj [1 – ρ2j,j-1 (0)] ½ € i,j (16)

This is in fact the Thomas-Fiering model. A third special case is possible when αj and βj are all zero in which case the model reduces to a normal independent process.

As a conclusion, it is possible to state that the model developed here is a mixture of the stationary lag-one Markov model and the periodic Thomas-Fiering model.

By substituting $w_{i,j} = x_{i,j} - \mu_j$ in Equation (16) and simplifying,
We get the Thomas-Fiering Model in the following form.

$X(i,j) = X_j + b_j [X(i,j-1) - X_{j-1}] + Z_t S_j \sqrt{1-r_j^2}$ (A)

Where,

X_j = Mean flow in j^{th} month

S_j = Standard deviation of j^{th} month

r_j = Correlation coefficient between the flows of j^{th} and $(j-1)^{th}$ month

b_j = Regression coefficient in j^{th} season of $(j-1)^{th}$ month

$X(i,j)$ = Flow generated in j^{th} month of i^{th} year

$j = 1,2,\ldots\ldots m$

$Z_t = N(0, 1)$ at any time step t

$Q(i,j)$ =Observed flow in j^{th} month of i^{th} year with available record of n years

$X_j = 1/n \sum Q(i,j)$

$S_j = \text{sqrt}[1/n -1) \sum[Q(i,j) - X_j]^2]$

$r_j = (1/n) \sum[[\{Q(i,j) - X_j\} \{Q(i,j) - X_j\}]] / S_j S_{j-1}$

$b_j = (r_j S_j) / S_{j-1}$

III. GENERATION OF SYNTHETIC FLOW RECORD

We have written a computer programme in MATLAB for the generation of a synthetic record of mean monthly flow by the method of Thomas and Fiering (1962). The following statistics of flow for each month are required:
(1) Mean,
(2) Standard deviation,
(3) Regression coefficient,
(4) Correlation coefficient.
Regression and correlation coefficients are for the regression of one month's flow on the next months.

3.1 Input File

The data required to be input includes the Historical data of monthly inflow at Karjan for 10 years from 1994-95 to 2003-04 in Million cubic feet.

3.2 Program logic

The basic software is being developed for programming purpose in order to generate the inflow for 100 years using historical data. The input taken is the historical inflow data from the already existing input file for 10 years.
First monthly historical inflow is considered as root for the generation series. The further generation of the monthly inflow continues based on the previous generated monthly flow.

Here, using the various parameters like regression coefficient, Standard deviation, Correlation coefficient between the progressive months, the generated monthly flow is obtained from the previous generated monthly inflow. The first monthly inflow is taken as the first historical monthly inflow data.

This program gives generated monthly inflow of all 12 months for 100 years.

IV. RESULTS

Table -1**:** Comparison for Mean values:

MONTHS	Historical Mean value	Generated Mean value
JAN	1.479	2.1146
FEB	0.917	1.0861
MAR	1.094	1.3931
APR	0.581	0.6861
MAY	1.851	1.9881
JUN	88.843	94.8125
JUL	254.584	229.9699
AUG	258.472	263.8052
SEP	190.872	177.2797
OCT	32.985	33.5411
NOV	5.61	6.3891
DEC	2.352	2.4518

The graph of monthly inflow at reservoir site from historical data is superimposed with the mean

monthly inflow from the generation of 100 years as shown in following graph.
It is obvious from the graph that the results are quite matching.

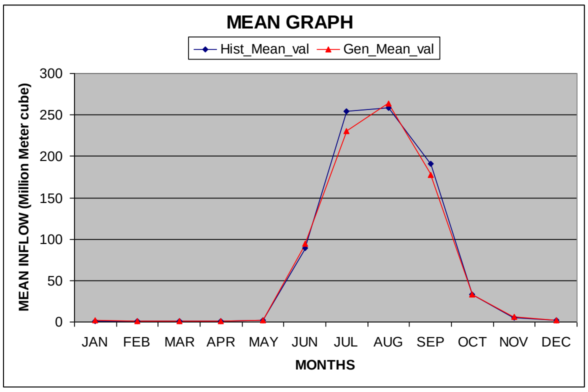


Table -2**:** Comparison for Standard Deviation values:

MONTHS	Historical Mean value	Generated Mean value
JAN	2.7722	2.1455
FEB	1.1155	0.8754
MAR	1.8515	1.4176
APR	0.5214	0.4526
MAY	3.3075	2.0864
JUN	84.1869	75.0371
JUL	226.7015	171.0252
AUG	191.7979	169.9912
SEP	159.5455	132.2647
OCT	23.33	18.6935
NOV	4.1658	3.6228
DEC	1.9424	1.7999

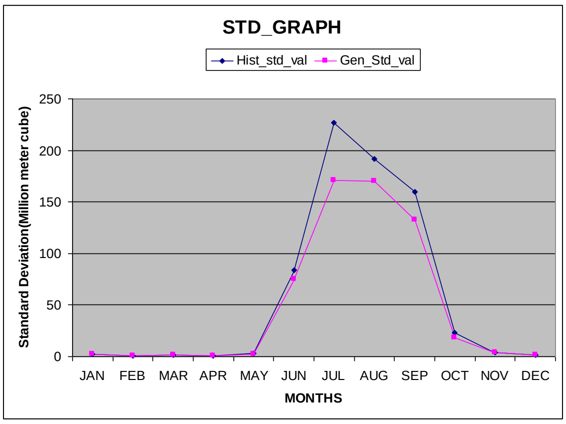
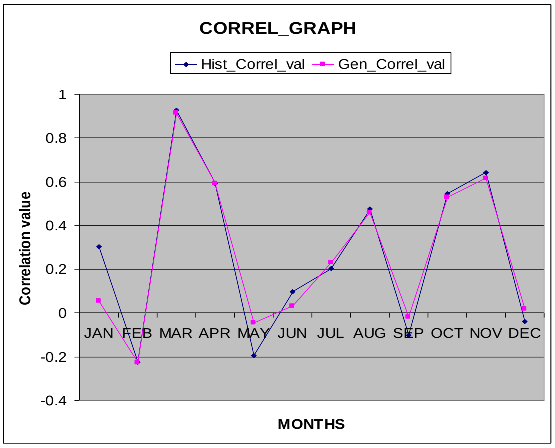


Table -3: Comparison for Correlation coefficient values:

MONTHS	Historical Mean value	Generated Mean value
JAN	0.393	0.0545
FEB	-0.225	-0.2265
MAR	0.9281	0.9126
APR	0.5922	0.5924
MAY	-0.1929	-0.0441
JUN	0.0972	0.0315
JUL	0.2054	0.2293
AUG	0.4755	0.4587
SEP	-0.1002	-0.0187
OCT	0.5468	0.5281
NOV	0.6424	0.6159
DEC	-0.0389	0.0187



V. Conclusion

The model presented has the advantage of a two-dimensional relationship between successive observations, namely, along months and along years.

The generated data results show that the model used is able to generate sequences of daily flows preserving the main parameters of a historic record.

The model is based on a series of relatively simple approximations to the recession limb, in conjunction with empirical transition probabilities for 'wet' and 'dry' days. At the time of writing this model has been applied only to the river for which data are presented here in and to one other river of very different hydrology. This and the fact that some of the empirical factors used in the model may not be applicable to other rivers should be borne in mind when proceeding to use the model elsewhere.

If we have the future inflow data we can plan for agriculture, industry, navigation and production of energy.

References

[1]. Harms, A.A. & Campbell, T.H. (1967) an extension to the Thomas- fiering model for the sequential generation of stream flow. Wat. Resour. Res. 3, 653-661

[2]. Box, G.E.P. & Jenkins, G.M. (1970) Time Series Analysis Forecasting and Control. Holden Day, San Francisco.

[3]. N.T.Kottegoda, Stochastic Water Resources Technology. University of Birmingham

[4]. Roesner, L.A. & Yevjevich, V.M. (1966) Mathematical models for time series of monthly precipitation and monthly runoff. Colorado State University, Hydrology Paper 15, Fort Collins, Colorado.

[5]. Thomas, H.A. & Fiering, M.B. (1962) Mathematical synthesis of stream flow sequences for the analysis of river basins by simulation. In: Design of Water Resources Systems.