Journal of Civil Engineering (IEB), 33(1) (2005) 43-54

Journal of Civil Engineering IEB

Disaggregation model for synthetic stream-flow generation

M. Shahjahan Mondal[1] and Saleh A. Wasimi[2]

[1]Institute of Water and Flood Management
Bangladesh University of Engineering and Technology, Dhaka 1000, Bangladesh
[2]Faculty of Informatics and Communication, Central Queensland University, Rockhampton, Queensland 4702, Australia

Received on 29 May 2005

Abstract

To capture the complexity of a water resources system, synthetic data generation is an essential component. Frequently, the data generation is done on an annual basis and disaggregated to smaller time scales. A generalised disaggregation framework is presented to generate seasonal stream-flows from any annual autoregressive process. A new periodic disaggregation scheme is proposed for further disaggregation into sub-seasonal flows from seasonal flows generated with a periodic autoregressive (PAR) model of any order. The new model preserves the first and second moments and has been applied to the Ganges river at Farakka in India for generation of decadal (10-day) flows from monthly flows; the 10-day period being the discrete time interval identified in the Ganges Water Treaty. The results demonstrate that the proposed coupled modelling scheme works very well and provides a flexible choice in synthetic hydrology.

Keywords: Disaggregation, synthetic data, the Ganges river flow.

1. Introduction

Synthetic data is widely used in water resource systems planning and management all over the world, especially for reservoir sizing and operation. Planning activities based on observed historic data, which suffers from sampling variability, and optimisation or simulation models of deterministic type are often of limited use for detailed study of the operation and performance of a complex water system, such as the Ganges river system, which is inherently stochastic in nature. To provide insight and guidance on how systems should be designed and operated, stochastic simulation or optimisation needs to be carried out. Stochastic models, which use synthetic data as inputs, are most flexible, powerful and widely used tools for planning and analysis of complex water resources systems (Loucks et al., 1981). Stochastic simulation models, in particular, can deal with a lot of complexities associated with systems planning and are able to solve highly non-linear relationships and constraints. Stochastic simulation models need generation of input data, which are usually of seasonal time resolution, with appropriately constructed statistical models of the system.

Seasonal input data, such as monthly stream-flow, can be generated either using a seasonal time series model directly, or using an annual time series model coupled with a disaggregation model. However, for generation of sub-seasonal (10-day) data from seasonal (monthly) data, disaggregation models are mandatory. Seasonal models such as Seasonal Autoregressive Integrated Moving Average (Box et al., 1994) model and Periodic Autoregressive (PAR, Hipel and McLeod, 1994) model have been widely used for forecasting and generation of hydrologic variables. McLeod and Hipel (1978) and Thompstone et al. (1987) have demonstrated that these models are capable of preserving both short- and long-term important historical statistics of hydrologic variables, particularly stream-flow records. Recently, Mondal and Wasimi (2005a) have developed a PAR model for monthly forecasting and generation of the Ganges river flow at Farakka in India. They generated 200 synthetic traces of historic length to demonstrate that their model is capable of preserving important theoretical and historical statistics. However, the model cannot be used directly in water resources planning for the Ganges delta within Bangladesh since its share of the Ganges water during the months of January-May is on decadal (10-day) basis as per the latest Ganges Water Treaty (GWT) of 1996. Direct development of a decadal model is constrained by the fact that decadal data for the whole year is not available. Therefore, a suitable mathematical framework is needed to generate decadal flows from already generated monthly flows. It is worth mentioning that the discharge data of the Ganges are considered as classified information by both Bangladesh and Indian governments because of long time disputes over sharing of the Ganges water, and as such, the data are not available at the measured one-day time resolution. The data available to the authors are monthly mean values derived from daily values. Fortunately, decadal data for five months (January-May) only are available in a research report by Colombi (1999), and another two months (November-December) could be collected through other sources. These data have been used in this study to disaggregate each month’s flow into its three decadal flows.

The paper is organised such that, in next Section, we briefly review the disaggregation models used for river-flow generation. We also describe the mathematical procedures of seasonal to sub-seasonal disaggregation. Thereafter, we apply the disaggregation model to the Ganges river for generation of decadal flows from monthly flows. We draw some conclusions of the study in the final Section.

2. Disaggregation models

2.1 Background information

Disaggregation is a mathematical technique for down-scaling coarser temporal or spatial levels into finer levels. Valencia and Schaake (1973) introduced a disaggregation model (hereafter referred to as VS model) that has become popular in stochastic hydrology (Tao and Delleur, 1976; Srikanthan, 1979). Among the attributes of this model, which is also known as the basic model, are the preservation of the first- and second-order moment properties at both the coarser and finer levels and the additive property: The aggregation of the generated finer level values always yields the coarser level value. However, the model proceeds in such a way that although the data inside a given coarser level preserve the statistics for all levels of aggregation, they are linked with the past finer level only through the statistics at the coarser level (Mejia and Rousselle, 1976). For example, the correlation coefficient between the last season and the first season of consecutive years is not preserved explicitly.

Mejia and Rousselle (1976, hereafter referred to as MR) attempted to overcome this shortcoming by including an extra term in the original VS model. However, Lane (1982) discovered an inconsistency in the extended MR model and found that none of the intended moments is preserved if MR parameter estimation equations are used. Salas et al. (1988) mentioned that the problem of an inconsistent casual structure in the basic VS model is not corrected by the additional term in the extended MR model, and the problem of excessive number of parameters ensues. Lane (1982), Stedinger and Vogel (1984), Lane and Frevert (1990), and Lin (1990a, b) widely discussed the impact of this moment inconsistency and suggested possible remedies.

Hoshi and Burges (1979) also attempted to overcome the shortcoming of the basic VS model by simultaneously disaggregating two successive annual events. However, their disaggregation scheme introduces some distortion between the summed inverse transformed individual seasonal flows and the annual amount from which they are disaggregated. In a later discussion on disaggregation techniques, Hoshi and Burges (1980) expressed their opinions on Lane's model as the best overall model, and recommended it for general use.

Santos and Salas (1992) have suggested a stage-wise disaggregation scheme to reduce the number of parameters involved in direct disaggregation. However, the method is based on the assumption that the upper level model is autoregressive of order one [AR(1)], which may not be true after the first stage for seasonal to sub-seasonal disaggregation. Disaggregation models of Lane (1982) and Lin (1990a, b) are also based on the assumption that the annual model is AR(1). In some cases, this assumption may not be satisfied. For instance, the annual flow of the St. Lawrence river at Ogdensburg, New York, modelled by Hipel and McLeod (1994) is AR(3) with the second parameter being zero. Furthermore, Lin's model has an unnecessarily excessive number of parameters and requires huge computer memory, and Lane's condensed model does not preserve the additive property of the generated data.

We intend to extend the MR disaggregation scheme so that it becomes suitable for an annual autoregressive process of any order. Necessary equations of parameter estimation for higher order processes are derived to bring in moment consistency and to preserve additive property of the generated data. The technique for disaggregation of an annual event into its seasonal events is addressed first in details, and then seasonal to sub-seasonal disaggregation procedure is outlined very briefly. It is to be mentioned here that another disaggregation technique known as the method of fragments (Srikanthan and McMahon, 1982; Porter and Pink, 1991; Maheepala and Perera, 1996), has been proposed for hydrologic and water resources applications, which is not covered here as the approach is not based on sound mathematical principles and does not work as well as the method discussed below.

2.2 Model Formulation and Parameter Estimation

Let X_t be the annual value of a variable to be partitioned into s sub-divisions of Y_t. X_t has been generated with an appropriate AR model. It is assumed that both variables are normally distributed with zero mean. The MR disaggregation model can be written as:

$$Y_t = AX_t + B\varepsilon_t + CY_{t-1} \tag{1}$$

where, Y_t is a column vector of dimension s containing seasonal values of year t, X_t is a single-element vector containing annual value of year t, ε_t is a column vector of dimension s containing independent standard normal variables, and Y_{t-1} is a column vector whose dimension (P) depends on the number of Y's from immediately preceding X included for preservation. A, B and C are parameter matrices of dimension $s \times 1$, $s \times s$ and $s \times P$, respectively. The difference between the MR and VS models is that the last term of the former in equation (1) is not included in the latter. Lane's condensed model differs from the MR model in that both B and C are diagonal matrices and Y_{t-1} has the same dimension as Y_t. Also, there is a difference in the actual elements of Y_{t-1} between the two models.

For moment estimation of the parameters A, B and C, equation (1) is post-multiplied with X_t^T, Y_{t-1}^T and Y_t^T and after that expected values are taken to obtain, respectively,

$$\gamma_{XY}^T = A\gamma_{XX} + C\gamma_{XY}^T(1) \tag{2}$$
$$\gamma_{YY}(1) = A\gamma_{XY}(1) + C\gamma_{YY} \tag{3}$$
$$\gamma_{YY} = A\gamma_{XY} + BB^T + C\gamma_{YY}^T(1) \tag{4}$$

where, γ_{UV} is the covariance of the general vectors U_t and the transpose of V_t, and $\gamma_{UV}(1)$ is the covariance between U_t and the transpose of V_{t-1}. The superscript T indicates the transpose of a matrix. In equations (2) to (4), there are five mathematical moments: γ_{XX}, γ_{XY}, $\gamma_{XY}(1)$, γ_{YY} and $\gamma_{YY}(1)$. However, the moment γ_{XX} is completely specified by the annual model. So, there are four moments for estimation from three equations, which creates the potential problem of moment inconsistency. These equations were originally proposed by Mejia and Rousselle, and if used without any adjustment, none of the four moments will be preserved (Lane, 1982; Lin, 1990a, b). If we closely look at the model in equation (1), we find that there is no direct link between X_t and Y_{t-1} in the model. Thus, the moment $\gamma_{XY}(1)$ can be blamed for the inconsistency (Lane, 1982) and needs to be corrected. This can be done using an annual autoregressive model:

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \quad \cdots \quad \cdots \quad + \phi_p X_{t-p} + a_t \tag{5}$$

where, ϕ_k is the AR coefficient at lag k. a_t is a stochastic random shock component with zero mean, constant variance and no serial correlation (white noise).

Post-multiplying equation (5) by Y_{t-1}^T and thereafter substituting $Y_{t-k} = AX_{t-k} + B\varepsilon_{t-k} + CY_{t-k-1}$ from equation (1) successively for $k = 1, 2, \cdots\cdots, p-1$

into those terms in which the time indices of X and Y are not identical, and finally taking expected values we obtain:

$$\gamma^*_{XY}(1) = \phi_1 \gamma_{XY} + \sum_{i=2}^{p} \phi_i \sum_{j=1}^{i-1} \gamma_{XX}(i-j) A^T + \sum_{i=2}^{p} \phi_i \gamma_{XY} C^T \tag{6}$$

The superscript * is introduced to indicate corrected estimate. This equation is an original derivation in this study. The equation can be used for any annual AR process including a constrained process. If the order of the annual model is one, the second and third terms of equation (6) disappear and the equation becomes identical to Lane's correction for AR(1) process. Use of equations (2), (3), (4) and (6) results in mathematically consistent moments. However, as $\gamma_{XY}(1)$ has been modified, the generated seasonal data no longer maintains the additive property. Lane has suggested further adjustment for $\gamma_{YY}(1)$ to preserve additivity:

$$\gamma^*_{YY}(1) = \gamma_{YY}(1) + \gamma_{YX} \gamma^{-1}_{XX} [\gamma^*_{XY}(1) - \gamma_{XY}(1)] \tag{7}$$

Using these adjusted $\gamma^*_{XY}(1)$ and $\gamma^*_{YY}(1)$, in place of $\gamma_{XY}(1)$ and $\gamma_{YY}(1)$ respectively, in equations (2) to (4), we can obtain mathematically consistent parameter estimates, which will preserve the additive property of generated seasonal data and reproduce historical moments γ_{XY} and γ_{YY}.

The parameter estimation technique for a higher order AR process differs from a simple AR(1) process in that direct solution is not possible since equation (6) contains the parameters A and C. To obtain the parameter values, the following iterative stages of estimation are required:

(i) Estimate $\gamma_{XY}(1)$ from historic data and take it as a preliminary value of $\gamma^*_{XY}(1)$;

(ii) Obtain $\gamma^*_{YY}(1)$ from equation (7) using $\gamma^*_{XY}(1)$ and then obtain A and C from equations (2) and (3);

(iii) Using the values of A and C of step (2), obtain $\gamma^*_{XY}(1)$ from equation (6); and

(iv) Repeat steps (2) to (3) until convergence occurs in $\gamma^*_{XY}(1)$.

After estimation of A and C, the solution of B for a given BB^T may be obtained from equation (4) with either principal component analysis (Anderson, 1960) or Cholesky decomposition (Healy, 1968).

Seasonal to sub-seasonal disaggregation model can be developed without any prior knowledge of annual model in a similar way assuming that seasonal data follows a PAR process. This further disaggregation process is briefly outlined below:

Let $Y_{t,s}$ be the value of a variable for year t and season s and it has to be partitioned into more than one sub-seasonal values of $Z_{t,s}$. The seasonal to sub-seasonal disaggregation model for season s can be written as:

$$Z_{t,s} = A_s Y_{t,s} + B_s \varepsilon_{t,s} + C_s Z_{t,s-1} \tag{8}$$

The notations have their usual meaning and are easily understood from the context. Post-multiplications of equation (8) with $Y_{t,s}^T$, $Z_{t,s-1}^T$ and $Z_{t,s}^T$ and thereafter taking mathematical expectation results in, respectively,

$$\gamma_{YZ}^{(s)T} = A_s \gamma_{YY}^{(s)} + C_s \gamma_{YZ}^{(s)T}(1) \tag{9}$$

$$\gamma_{ZZ}^{(s)}(1) = A_s \gamma_{YZ}^{(s)}(1) + C_s \gamma_{ZZ}^{(s-1)} \tag{10}$$

$$\gamma_{ZZ}^{(s)} = A_s \gamma_{YZ}^{(s)} + B_s B_s^T + C_s \gamma_{ZZ}^{(s)T}(1) \tag{11}$$

where, $\gamma_{UV}^{(s)}$ and $\gamma_{UV}^{(s)}(1)$ are, respectively, the lag-0 and lag-1 periodic covariance between $U_{t,s}$ and the transpose of $V_{t,s}$. Note the difference in the last term between equations (3) and (10). To maintain mathematical consistency in the generated moments, periodic model for season s is used:

$$Y_{t,s} = \phi_1^{(s)} Y_{t,s-1} + \phi_2^{(s)} Y_{t,s-2} + \cdots \cdots + \phi_p^{(s)} Y_{t,s-p} + a_{t,s} \tag{12}$$

where, $\phi_k^{(s)}$ is the AR coefficient for season s at lag k. $a_{t,s}$ has the same property as a_t mentioned earlier.

Post-multiplying equation (12) by $Z_{t,s-1}^T$ and substituting $Z_{t,s-k} = A_{s-k} Y_{t,s-k} + B_{s-k} \varepsilon_{t,s-k} + C_{s-k} Z_{t,s-k-1}$ from equation (8) successively for $k = 1, 2, \cdots\cdots, p$ into those terms in which the time indices for Y and Z are not identical, and finally taking expected values we obtain:

$$\gamma_{YZ}^{(s)*}(1) = \phi_1^{(s)} \gamma_{YZ}^{(s-1)} + \sum_{i=2}^{p} \phi_i^{(s)} \sum_{j=1}^{i-1} \gamma_{YY}^{(s-j)}(i-j) A_{s-j}^T + \sum_{i=2}^{p} \phi_i^{(s)} \gamma_{YZ}^{(s-i)} C_{s+1-i}^T \tag{13}$$

This equation is applicable for a periodic AR model of any order including constrained model. For preservation of additivity, the following adjustment can be made:

$$\gamma_{ZZ}^{(s)*}(1) = \gamma_{ZZ}^{(s)}(1) + \gamma_{ZY}^{(s)} \gamma_{YY}^{(s)-1} [\gamma_{YZ}^{(s)*}(1) - \gamma_{YZ}^{(s)}(1)] \tag{14}$$

If a given season s has a PAR model of order k, and preceding seasons $s-1, s-2, \cdots \cdots$ have models of order $k-1, k-2, \cdots \cdots, 1$ respectively, then the disaggregation parameters can be obtained directly without iteration. For example, if the model for the month of May is PAR(2) and for April is PAR(1), then the parameters for the disaggregation model of May can be obtained directly using A and C of already known April model. If the model for May is PAR(3), for April is PAR(1) or PAR(2), and for March is PAR(1), then the disaggregation parameters for May can also be obtained directly using A and C of April and March, which are already known.

3. Application of the proposed disaggregation model to the Ganges river for generation of decadal flows from monthly flows

The Ganges river, which flows through China, India, Nepal and Bangladesh, is one of the largest rivers in the world. Its total length is about 2510 km and the catchment area is about one million km^2. The river has a vital role in shaping the socio-economic conditions of Bangladesh and India, and any scheme to model the river-flow can be extremely beneficial. The disaggregation scheme developed in this paper is applied to the Ganges at Farakka in India. The selection of Farakka station is due to its most downstream location with natural flow. Any station downstream of Farakka has regulated flow mainly controlled by the GWT for sharing January-May flows between Bangladesh and India. Farakka is located about 17 km upstream from the Bangladesh-India border and significant diversion of the Ganges water from Farakka began on April 21, 1975. Inflow data at Farakka is available for the period October 1948 to December 2000. Monthly flow generated in Mondal and Wasimi (2005a) with PAR model is disaggregated here into decadal flows, and only dry season flow (November-May) is disaggregated as that flow alone is currently considered important for water resources planning and management in Bangladesh.

The strategy of generating annual flow first, and disaggregating into decadal flows subsequently, was not adopted in this study because the Ganges annual flows are weakly auto-correlated and strongly cross-correlated with a number of variables (see Mondal and Wasimi, 2005b, and the references therein). Logically therefore, generation of annual flow would require that all cross-correlated variables be generated with their appropriate models. This complicates the data generation process and involves a huge number of parameters, which is usually avoided in practice. Furthermore, finding suitable time series models for some predictors may not be possible. For example, the autocorrelation function of NIÑO 3.4 sea surface temperature, which has correlation with the Ganges annual flow, does not decay quickly and shows no identifiable pattern even after non-seasonal and/or seasonal differencing though its time-sequence plot appears to be stationary. Also, no standard technique is presently available to initialise the required seasonal values to start the data generation process in a disaggregation framework. Preservation of season-to-season flow characteristics is more important than annual characteristics for planning and management of water resources in the Ganges river basin as it is an intra-year system. Intra-year systems generally refill each year (McMahon and Mein, 1986). Furthermore, both analytical and simulation results from research in the last two decades (Vecchia et al., 1983; Obeysekera and Salas, 1986; Bartolini et al., 1988; Lin and Lee, 1992) suggest that significant gain in parameter estimation can be achieved using seasonal data and their model, rather than using annual data and its model.

To develop decadal models from seasonal models and without using the annual model, the periodic parameter matrices A_s, B_s and C_s for each month were estimated using equations (13), (14), (9), (10) and (11). For each month, flow of the last decade of immediately preceding month was considered in the vector $Z_{t,s-1}$ of equation (8). To preserve the additivity property, logarithmic decadal data ($Z_{t,s}$ and $Z_{t,s-1}$) were multiplied with corresponding number of days in a decade to obtain monthly data ($Y_{t,s}$) before estimating the model parameters. The estimated parameters for December to May were, respectively:

$$A_D = \begin{bmatrix} 0.2805 \\ 0.3561 \\ 0.3634 \end{bmatrix} \quad B_D = \begin{bmatrix} 0.3371 & 0 & 0 \\ 0.0195 & 0.1633 & 0 \\ -0.3566 & -0.1633 & 0 \end{bmatrix} \quad C_D = \begin{bmatrix} 0.2104 \\ -0.0733 \\ -0.1371 \end{bmatrix}$$

$$A_J = \begin{bmatrix} 0.2705 \\ 0.3333 \\ 0.3962 \end{bmatrix} \quad B_J = \begin{bmatrix} 0.5276 & 0 & 0 \\ -0.1137 & 0.2398 & 0 \\ -0.4139 & -0.2398 & 0 \end{bmatrix} \quad C_J = \begin{bmatrix} 0.1688 \\ -0.0071 \\ -0.1617 \end{bmatrix}$$

$$A_F = \begin{bmatrix} 0.2461 \\ 0.4040 \\ 0.3499 \end{bmatrix} \quad B_F = \begin{bmatrix} 0.4113 & 0 & 0 \\ -0.0925 & 0.2703 & 0 \\ -0.3188 & -0.2703 & 0 \end{bmatrix} \quad C_F = \begin{bmatrix} 0.2804 \\ -0.1257 \\ -0.1547 \end{bmatrix}$$

$$A_M = \begin{bmatrix} 0.2591 \\ 0.3437 \\ 0.3972 \end{bmatrix} \quad B_M = \begin{bmatrix} 0.4658 & 0 & 0 \\ -0.0049 & 0.2326 & 0 \\ -0.4609 & -0.2326 & 0 \end{bmatrix} \quad C_M = \begin{bmatrix} 0.3380 \\ -0.0723 \\ -0.2657 \end{bmatrix}$$

$$A_A = \begin{bmatrix} 0.2262 \\ 0.3424 \\ 0.4314 \end{bmatrix} \quad B_A = \begin{bmatrix} 0.5385 & 0 & 0 \\ -0.0392 & 0.2961 & 0 \\ -0.4993 & -0.2961 & 0 \end{bmatrix} \quad C_A = \begin{bmatrix} 0.3407 \\ -0.0273 \\ -0.3134 \end{bmatrix}$$

$$A_M = \begin{bmatrix} 0.2101 \\ 0.3394 \\ 0.4505 \end{bmatrix} \quad B_M = \begin{bmatrix} 0.5946 & 0 & 0 \\ 0.1460 & 0.4904 & 0 \\ -0.7406 & -0.4904 & 0 \end{bmatrix} \quad C_M = \begin{bmatrix} 0.3081 \\ -0.0746 \\ -0.2335 \end{bmatrix}$$

As the measured flow for the last decade of October was not available, the VS model was used for disaggregation of November flow. The parameter matrices for November were:

$$A_N = \begin{bmatrix} 0.4026 \\ 0.3209 \\ 0.2765 \end{bmatrix} \quad B_N = \begin{bmatrix} 0.5000 & 0 & 0 \\ -0.1383 & 0.1773 & 0 \\ -0.3617 & -0.1773 & 0 \end{bmatrix}$$

From the values of the parameter matrices above, it can be seen that the sum of the three elements of vector A_s is always one, the sum of each column elements of matrix B_s is always zero, and the sum of the three elements of vector C_s is always zero. These results indicate that the proposed periodic disaggregation model is capable of preserving the additive property of the generated sub-seasonal data.

Independent standard normal variables, $\varepsilon_{t,s}$'s, required in the disaggregation process were generated using random number seeds different from those used in monthly flow generation in Mondal and Wasimi (2005a). The 200 synthetic traces of monthly flow generated with the PAR model in Mondal and Wasimi were then disaggregated into decadal flow using above estimated parameter values and generated standard normal random numbers. Important statistics of the disaggregated decadal flow were computed for each sequence and the average of each statistic was obtained from the 200 values. They are presented in Table 1 along with their historical values. The first, second and third row entries for each month in columns two and three in the table are the correlations between the total flow of a month and its first, second and third decadal flows, respectively. The first, second and third row entries in columns four and five are the correlations between the first and second, the second and third, and the third and first decadal flows of a month, respectively.

It is seen from the table that the proposed periodic disaggregation model can preserve the correlation between concurrent monthly and decadal flows and between one decadal flow to another of a month quite well. The decadal flow for November was not properly generated as the VS model was used. Also for December, a PAR(1) model was used for estimation of disaggregation parameters for this month instead of the full model which has three significant lags (see Mondal and Wasimi, 2005a). The use of the full model requires the availability of decadal data for June-October. Therefore, decadal data for November-December were not properly generated. This may have some effects on the generated data for January to May as each month was linked to one or two earlier months in the disaggregation parameter estimation algorithm. This may have caused some of the differences between the observed and generated correlations in the table. In spite of this limitation, the proposed disaggregation scheme has worked very well.

Table 1
Some important correlations of the generated and observed concurrent monthly-decadal and decadal-decadal flows of the Ganges river at Farakka

Month	Monthly-Decadal Correlation		Decadal-Decadal Correlation	
	Observed	Generated	Observed	Generated
	0.989	0.990 (.985 .995)	0.978	0.980 (.970 .991)
Nov	0.996	0.997 (.995 .999)	0.984	0.987 (.979 .995)
	0.985	0.987 (.980 .994)	0.951	0.957 (.936 .979)
	0.980	0.975 (.961 .989)	0.972	0.964 (.943 .984)
Dec	0.995	0.993 (.989 .997)	0.961	0.949 (.921 .976)
	0.973	0.963 (.943 .983)	0.912	0.885 (.826 .944)
	0.945	0.941 (.909 .973)	0.915	0.908 (.860 .957)
Jan	0.989	0.988 (.982 .994)	0.951	0.949 (.924 .973)
	0.964	0.963 (.945 .981)	0.836	0.823 (.741 .911)
	0.959	0.955 (.930 .980)	0.924	0.918 (.874 .963)
Feb	0.988	0.987 (.980 .994)	0.951	0.947 (.917 .976)
	0.966	0.963 (.943 .983)	0.868	0.857 (.784 .929)
	0.969	0.968 (.950 .987)	0.958	0.957 (.931 .982)
Mar	0.994	0.993 (.989 .997)	0.954	0.953 (.927 .978)
	0.967	0.967 (.950 .985)	0.881	0.880 (.816 .943)
	0.930	0.927 (.891 .963)	0.908	0.904 (.856 .952)
Apr	0.988	0.987 (.980 .995)	0.909	0.908 (.861 .956)
	0.933	0.932 (.897 .966)	0.747	0.740 (.627 .853)
	0.925	0.926 (.889 .964)	0.906	0.907 (.858 .956)
May	0.970	0.970 (.952 .988)	0.857	0.857 (.774 .939)
	0.938	0.938 (.901 .975)	0.757	0.761 (.642 .880)

Note: The values within parentheses are 95% confidence limits.

Decadal inflows at Farakka were generated first at transformed unit of natural logarithm of zero periodic mean. The inverse process of demeaning and transformation was followed thereafter to obtain flows in original units. To further check the adequacy of the proposed disaggregation scheme, generated decadal inflow at Farakka was divided between Bangladesh and India according to the GWT of 1996. Bangladesh share thus obtained was assumed to be available to Bangladesh without any modification in-between Farakka and Bangladesh-India land border, a distance of about 17 km. The assumption is realistic in view of the GWT, which assures in Article-III that the water released to Bangladesh at Farakka shall not be reduced downstream by India, except for reasonable uses not exceeding 5.66 m^3/s. For each year and each sequence of available flows to Bangladesh, the quantity of water that can be constantly supplied throughout the dry season was estimated. This estimation included a possible reservoir on the Ganges within Bangladesh with a capacity of 909 Mm^3 (WARPO, 2001). Average maximum possible constant supply for each year was estimated from 200 sequences, and is given in Figure 1 along with the 95% confidence limits. For comparison, possible constant supply that could be maintained if historical sequence was used is also shown.

It is seen from the figure that use of observed and generated data resulted in similar constant supply values. At an exceedence probability of 80 percent, the maximum possible constant supply value was 77.84 Mm^3 per day when observed data were used and 78.02 Mm^3 when generated data were used. These results support the argument that the data generation-disaggregation process was appropriate.

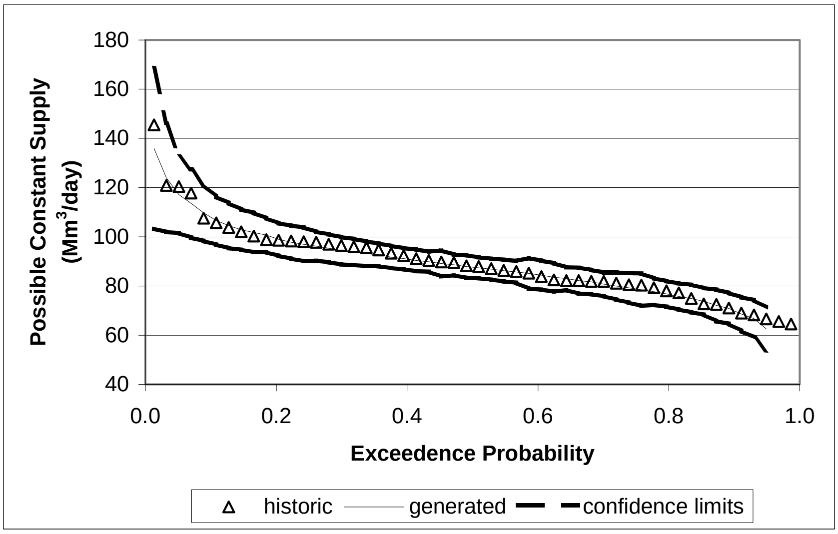

Fig. 1. Maximum possible constant supply throughout the dry season from the Ganges barrage

4. Conclusions

Over the past, many disaggregation schemes have been proposed. Many such schemes failed to preserve the autocorrelation structure of the process. In this paper we investigate into disaggregation and sub-disaggregation of annual stream-flows and present mathematical techniques to preserve the correlation structure even for processes which are of higher order than AR(1). Specifically, we have shown that when seasonal flows are generated with PAR models instead of with a disaggregation model, mathematically consistent moment equations can be developed for generation of sub-seasonal flows from seasonal flows. The proposed model is capable of preserving the first and second moments between concurrent upper- and lower-level flows and between lower-level flows themselves. The periodic version of the model has been used to generate 200 historic-length synthetic traces of decadal (10-day) flows for the Ganges river from PAR model generated monthly flows. The comparison between observed and generated flow statistics and possible reservoir supplies corroborated the fact that the proposed coupled modelling scheme works well.

References

Anderson, T. W. (1960). “An introduction to multivariate statistical analysis”, John Wiley and Sons, New York.

Bartolini, P., Salas, J. D., and Obeysekera, J. T. B. (1988). Multivariate Periodic ARMA(1,1) Processes, *Water Resources Research*, 24(8), 1237-1246.

Box, G. E. P., Jenkins, G. M., and Reinsel, G. C. (1994). Time series analysis: Forecasting and control, Prentice Hall, Englewood Cliffs, New Jersey.

Colombi, J. S. (1999). Farakka inflows and the Ganges Water Treaty. Topic Paper No. 7a, Water Resources Planning Organisation, Dhaka.

Healy, M. J. R. (1968). “Algorithm AS6: Triangular Decomposition of a Symmetric Matrix”, *Journal of the Royal Statistical Society, Series C (Applied Statistics)*, 17, 195-197.

Hipel, K. W., and McLeod, A. I. (1994). Time series modelling of water resources and environmental systems, Elsevier, Amsterdam, The Netherlands.

Hoshi, K., and Burges, S. J. (1979). Disaggregation of streamflow volumes. *Journal of the Hydraulics Division, ASCE*, 105(HY1), 27-41.

Hoshi, K., and Burges, S. J. (1980). Disaggregation of streamflow volumes. *Journal of the Hydraulics Division, ASCE*, 106(HY6), 1127-1128.

Lane, W. L. (1982). "Corrected parameter estimates for disaggregation schemes." Statistical analysis of rainfall and runoff, V. P. Singh, ed., Water Resources Publications, Littleton, Colorado, USA.

Lane, W. L., and Frevert, D. K. (1990). "Applied stochastic techniques (LAST personal computer package version 5.2), User's Manual." Earth Sciences Division, U.S. Bureau of Reclamation, Denver, Colorado.

Lin, G. F. (1990a). Corrected parameter estimates for stage disaggregation, *Chinese Institute of Engineers Journal*, 13(4), 405-416.

Lin, G. F. (1990b). Parameter estimation for seasonal to subseasonal disaggregation, *Journal of Hydrology*, 120(1-4), 65-77.

Lin, G. F., and Lee, F. C. (1992). An aggregation-disaggregation approach for hydrologic time series modelling. *Journal of Hydrology*, 138(3-4), 543-557.

Loucks, D. P., Stedinger, J. R., and Haith, D. A. (1981). Water resource systems planning and analysis, Prentice-Hall, Englewood Cliffs, N.J.

Maheepala, S., and Perera, B. J. C. (1996). "Monthly hydrologic data generation by disaggregation. *Journal of Hydrology*, 178(1), 277-291.

McLeod, A. I., and Hipel, K. W. (1978). *Simulation procedures for Box-Jenkins models*. Water Resources Research, 14(5), 969-975.

McMahon, T. A., and Mein, R. G. (1986). River and reservoir yield, Water Resources Publications, Littleton, Colorado.

Mejia, J. M., and Rousselle, J. (1976). Disaggregation models in hydrology revisited, *Water Resources Research*, 12(2), 185-186.

Mondal, M. S., and Wasimi, S. A. (2005a). Generating and forecasting monthly flows of the Ganges river with PAR model. *Journal of Hydrology*, (under review).

Mondal, M. S., and Wasimi, S. A. (2005b). "Periodic transfer function-noise model for forecasting. *Journal of Hydrologic Engineering, ASCE,* (in press).

Obeysekera, J. T. B., and Salas, J. D. (1986). Modeling of aggregated hydrologic time series. *Journal of Hydrology*, 86, 197-219.

Porter, J. W., and Pink, B. J. (1991). A method of synthetic fragments for disaggregation in stochastic data generation. Int. Hydrol. and Water Resour. Symposium, The Institution of Engineers, Australia, Perth, W.A., Australia.

Salas, J. D., Delleur, J. W., Yevjevich, V., and Lane, W. L. (1988). Applied modeling of hydrologic time series, Water Resources Publications, Fort Collins, Colorado, USA.

Santos, E. G., and Salas, J. D. (1992). Stepwise disaggregation scheme for synthetic hydrology. *Journal of Hydraulic Engineering, ASCE*, 118(5), 765-784.

Srikanthan, R. (1979). Sequential generation of monthly streamflows. *Journal of Hydrology*, 38, 71-80.

Srikanthan, R., and McMahon, T. A. (1982). Stochastic generation of monthly streamflows. *Journal of the Hydraulics Division, ASCE*, 108(HY3), 419-441.

Stedinger, J. R., and Vogel, R. M. (1984). Disaggregation procedures for generating serially correlated flow vectors. *Water Resources Research*, 20(1), 47-56.

Tao, P. C., and Delleur, J. W. (1976). Multistation, multiyear synthesis of hydrologic time series by disaggregation. *Water Resources Research*, 12(6), 1303-1312.

Thompstone, R. M., Hipel, K. W., and McLeod, A. I. (1987). "Simulation of monthly hydrological time series." Advances in Statistical Sciences, Volume IV, Stochastic Hydrology, I. B. MacNeill and G. J. Umphrey, eds., D. Reidel Publishing Co., Dordrecht, The Netherlands, 57-71.

Valencia, R. D., and Schaake Jr., J. C. (1973). "Disaggregation processes in stochastic hydrology. *Water Resources Research*, 9(3), 580-585.

Vecchia, A. V., Obeysekera, J. T. B., and Salas, J. D. (1983). Aggregation and estimation for low order periodic ARMA models. *Water Resources Research*, 19(5), 1297-1306.

WARPO. (2001). "Options for the Ganges dependent area." Final Report, Volume 2: Ganges barrage pre-feasibility study, National Water Management Plan Project, Water Resources Planning Organization, Dhaka.