# APPENDIX 7A THE DISAGGREGATION MODEL[1]

*DARIO VALENCIA*
*JOHN C. SCHAAKE, JR.*

## THE MODEL

The disaggregation model takes the simple mathematical form

$$\mathbf{Y} = \mathbf{AX} + \mathbf{W} \tag{7A.1}$$

where **Y** is an $(n \times 1)$ vector of correlated random variables, **X** is an $(m \times 1)$ vector of correlated random variables, **A** is an $(n \times m)$ coefficient matrix, and **W** is an $(n \times 1)$ vector of correlated random variables and is independent of **X**. It is assumed that all random variables have been transformed to have zero mean. Although it may be helpful to assume that the appropriate data transformations have been made to render the data Gaussian, it is not always necessary. It can be shown that the model preserves first- and second-order moment properties regardless of the types of underlying probability distributions.

An important application of this model is to generate series of seasonal streamflow volumes from a given series of annual streamflow volumes. In this case, the vector **X** is a vector of annual streamflow volumes at $m$ different sites

[1]This appendix has been taken from Valencia and Schaake (1973).

$$\mathbf{X} = \begin{bmatrix} x_1 \\ \cdot \\ \cdot \\ \cdot \\ x_m \end{bmatrix}$$

The vector **Y** is a vector of seasonal volumes—that is, an $n \times 1$ vector, where $n = 3m$ assuming three 4-month seasons in a year.

$$\mathbf{Y} = \begin{bmatrix} y_{11} = y_1 \\ y_{12} = y_2 \\ y_{13} = y_3 \\ y_{21} = y_4 \\ y_{22} = y_5 \\ y_{23} = y_6 \\ \cdot \quad\quad \cdot \\ \cdot \quad\quad \cdot \\ \cdot \quad\quad \cdot \\ y_{m3} \quad y_n \end{bmatrix}$$

The element $y_{ij}$ denotes the streamflow volume at site $i$ during season $j$. The seasonal runoff volumes at the $m$ sites for a year are represented by the $y_{ij}$ in a **Y** vector.

Certain subsets of the $y_{ij}$ bear a simple relation to the $x_i$. Specifically,

$$x_i = \sum_{j=1}^{3} y_{ij} \tag{7A.2}$$

or, in general,

$$\mathbf{X} = \mathbf{CY} \tag{7A.3}$$

where

$$\mathbf{C} = \begin{bmatrix} 111 & 000 & \cdots & 000 \\ 000 & 111 & \cdots & 000 \\ \cdot & \cdot & \cdot & \cdot \\ 000 & 000 & \cdots & 111 \end{bmatrix}$$

## DISAGGREGATION OF ANNUAL STREAMFLOW DATA

The strategy for using the model is first to generate values of **X** for a number of years by using a fractional noise model or an autoregressive model. Then, given the generated value of **X** for a particular year, the proposed model is used to generate **Y**.

Because historically the relation between **X** and **Y** given by Eq. (7A.3) is always maintained, it is important that this relation be maintained in the generated data. In other words, the sum of the three generated seasonal volumes at any site should be exactly equal to the given annual volume. It is shown in Valencia and Schaake (1973) that this important relationship is maintained. This central property of the model guarantees that the disaggregation preserves all the statistics considered in the annual generation.

In the case of annual disaggregation, the value of each seasonal runoff volume at a particular site is computed as the sum of terms depending on the annual runoff volumes at the different sites plus a random component. The random components may be generated by the relation

$$\mathbf{W} = \mathbf{B}\mathbf{V} \tag{7A.4}$$

where **V** is a vector of $n$ independently distributed standard normal deviates and **B** is a coefficient matrix selected to preserve the proper covariance structure of **W**. In a second step, the disaggregation model could be used to disaggregate seasonal values into monthly values.

## PARAMETER ESTIMATION

The parameter estimation problem is to use historical data to estimate numerical values of **A** and **B** so that the generated values of **Y** resemble the historical values of **Y** according to an appropriate resemblance criterion. A widely used criterion is that expected means, variances, and covariances of the generated data are equal to the historical means, variances, and covariances. Expressions for the generated means, variances, and covariances in terms of the parameters **A** and **B** are derived in Valencia and Schaake (1973), resulting in a pair of equations for the matrix estimators $\hat{\mathbf{A}}$ and $\hat{\mathbf{B}}$.

## EXTENSION OF THE MODEL

Equation (7A.1) can be seen in a broader perspective as representing a whole class of models. The disaggregation model already discussed is a member of this class; other models proposed in the field of stochastic hydrology, such as the Matalas (1967) model or, in general, multivariate autoregressive models of any order, are also included in the class. The question in each case is to define in a proper way the meanings of **X** and **Y**; the model is applicable when the relevant statistics to be reproduced in the generation are those related to means, variances, and covariances.

As a result of this generalization, all the properties exhibited by the disaggregation model, such as those related to estimation and existence of parameters and to preservation of linear transformations [see Valencia and Schaake (1973)], are also valid for the other models of the class.

An illustration of the flexibility of the general model is presented in Appendix 10A, where Eq. (7A.1) is used to augment hydrologic data.

## REFERENCES

MATALAS, N. C. "Mathematical Assessment of Synthetic Hydrology." *Water Resources Research*, 3 (*4*), 1967, pp. 937–946.

VALENCIA, DARIO, and JOHN C. SCHAAKE, JR. "Disaggregation Processes in Stochastic Hydrology." *Water Resources Research*, 9 (*3*), 1973, pp. 580–585.

Tomado del libro *Applied Water Resource Systems Planning*
de David C. Major y Roberto L. Lenton
Prentice Hall International
Estados Unidos, 1979