Multisite ARMA(1,1) and Disaggregation Models for Annual Streamflow Generation

JERY R. STEDINGER

Department of Environmental Engineering, Cornell University, Ithaca, New York

DENNIS P. LETTENMAIER

Department of Civil Engineering, University of Washington, Seattle

RICHARD M. VOGEL

Department of Environmental Engineering, Cornell University, Ithaca, New York

Disaggregation and multisite autoregressive moving average (ARMA(1,1)) time-series models provide simple and efficient frameworks for generation of multisite synthetic streamflow sequences that exhibit long-term persistence. This paper considers multisite ARMA(1,1) models whose \( \Phi \) and \( \Theta \) matrices are diagonal; a Monte Carlo study examined the efficiency of three procedures for estimating individual \( \phi \)-\( \theta \) values for each site and two estimators of the covariance matrix of the innovations. Also included in the study was a univariate ARMA(1,1) model of the aggregate flows with a simple disaggregation algorithm to generate flows at the individual sites. In the realm of most hydrologic interest, simple diagonal multisite ARMA(1,1) models performed adequately and it is not necessary to fit the more cumbersome nondiagonal models. The disaggregation procedure coupled with an ARMA(1,1) aggregate flow model did as well as the multivariate diagonal ARMA(1,1) models.

INTRODUCTION

Autoregressive moving-average time series models have proven to be a flexible tool for use in water resources planning. The autoregressive moving average (ARMA(1,1)) model, in particular, has a physically reasonable correlation structure which can reflect the long-term persistence which may be present in geophysical time series. Several papers have discussed long-term persistence and possible physical explanations (see, for example, Klemes [1974, 1978]; Potter [1975, 1976]). Of the operational models which provide for long-term persistence, the ARMA(1,1) model has the advantage that it is consistent with at least two explanations that have been advanced. The first, advocated by Mandelbrot and Wallis [1968], was that long-term persistence is the result of long memory in hydraulic processes. However, it can be argued that this explanation is not reasonable in view of the time constants observed in the physical processes resulting in surface runoff [Moss, 1972, Klemes, 1974].

The second and we believe more plausible explanation is that long-term persistence results from shifts in the means of hydrologic processes which may be related, for instance, to climatic change. Boes and Salas [1978] have demonstrated that a univariate ARMA(1,1) structure can result from a shifting means process. Moreover, the Hurst phenomena may be just a preasymptotic behavior for relatively persistent short-memory models such as ARMA(1,1) processes [Salas et al., 1979]. O'Connell [1971, 1974, 1977] has pursued the long-term persistence basis of the model with the practical result that the ARMA(1,1) structure may be considered compatible with either explanation.

In this paper, we explore fitting methods for multivariate ARMA(1,1) models employed in two ways. The first form uses an ARMA(1,1) model structure for flows at each site. This we believe is an advantage from the standpoint of physical plausibility. Models of this kind have been proposed by Matalas and Wallis [1976] for AR(1) processes and by Salas et al. [1980] and Loucks et al. [1981] for ARMA(p,q) process. However, the question of the relative performance of the several possible parameter estimation methods has not been addressed. We also consider a univariate ARMA(1,1) model of the aggregate flows coupled with a disaggregation procedure to generate separate annual flows for each site.

\textbf{Multiple Site ARMA(1,1) Models}

The general multiple-site ARMA(1,1) model may be written

\[ X_t = \Phi X_{t-1} + V_t - \Theta V_{t-1} \] (1)

where \( X_t \) is an \( m \times 1 \) vector of normally distributed flow residuals (zero mean) in period \( t \) with covariance matrix \( S_0 \), where \( V_t \) is an \( m \times 1 \) vector of time-independent normally distributed random fluctuations with covariance matrix \( G \), and where \( \Phi \) and \( \Theta \) are \( m \times m \) coefficient matrices. A thorough discussion of such multivariate stochastic models is provided by Jenkins and Alavi [1981].

Various approaches to estimation of the parameter vectors and covariance matrices have been explored. Wilson [1973] and Tiao and Box [1981] recommend maximum likelihood estimates (MLE) of \( \Phi \), \( \Theta \), and G. Spliid [1983] presents a similar but computationally faster alternative. The difficulty with these procedures is that they both require a computationally burdensome search for the \( m^2 \) elements of the \( \Phi \) and \( \Theta \) matrices prior to estimation of \( G \) as the covariance matrix of the fitted residuals \( V_t \). Beyond the computation burden, Lettenmaier [1980] found that the single-site auto-
correlations corresponding to fitted MLE models could differ substantially from those obtained from univariate estimation at the individual sites. In this work, we also show that approximate MLE estimates of the lag-zero cross-correlations tend to be biased strongly downward, particularly when the (population) cross-correlations are high. This is troubling given that the cross-correlation of annual streamflows within the same region are generally quite high. Stedinger [1981] considered alternative estimators of the cross-correlation of AR(1) processes and obtained similar results.

Salas et al. [1980] and Loucks et al. [1981] have suggested that \( \Phi \) and \( \Theta \) be taken as diagonal matrices. Then the elements of each are essentially the parameters of univariate ARMA(1,1) models fitted to the flows at each site. \( \Gamma \) can then be estimated as the sample covariance of the fitted \( \mathbf{V} \), defined by (1). The sample elements of \( \mathbf{V} \) are produced as a byproduct of the MLE fitting process, or they can be computed directly from the fitted univariate models and the data. Initial values in the recursions can be obtained, given the univariate parameters, using the back forecasting method of Box and Jenkins [1976] to provide unconditional estimates of the residuals. Alternatively, given \( \Phi \) and \( \Theta \), \( \Gamma \) can be computed so as to reproduce the observed lag-zero covariance matrix, \( \mathbf{S}_0 \), of the flows.

O'Connell [1974, 1977] has suggested a method of moments approach to the estimation of \( \Phi \), \( \Theta \), and \( \Gamma \). His method involves substitution of the sample lag zero, one and two covariances of the \( \mathbf{X} \) process, \( \mathbf{S}_0 \), \( \mathbf{S}_1 \), and \( \mathbf{S}_2 \) into population relationships between \( \Phi \), \( \Theta \), and \( \Gamma \). In practice, the method is plagued by numerical difficulties associated with the inversion of certain matrices whose sample estimates may not be positive semidefinite [Armbruster, 1978; Puente and Deeb, 1979]. Many of these problems arise because there may be no feasible \( \Phi \) and \( \Theta \) matrices for given \( \mathbf{S}_0 \), \( \mathbf{S}_1 \), and \( \mathbf{S}_2 \). Consequently, we did not pursue this approach.

All of the methods suggested to date have some drawbacks. The estimation of \( \Phi \) and \( \Theta \) using univariate methods has the advantage that the \( m \) single-site models are then ARMA(1,1) with well known properties. This assures that the multisite model is hydrologically reasonable at each site individually and jointly. For example, the elements of the covariance matrices are given by

\[
\mathbf{S}_k = \mathbf{\Phi}^k \mathbf{S}_1 \quad k > 1
\]

and decay monotonically so long as \( 0 < \phi_i < 1 \), as should be the case for hydrologically reasonable univariate models [O'Connell, 1974].

With the assumption of diagonality for \( \Phi \) and \( \Theta \), the model fitting process consisted of two independent steps: estimation of \( \Phi \) and \( \Theta \), and estimation of \( \Gamma \) conditioned on \( \Phi \) and \( \Theta \). This resulted in a total of six parameter estimation procedure combinations. A seventh method, based on disaggregation of aggregate or summed flows to the individual sites, is discussed in the next section. The three methods for estimation of \( \Phi \) and \( \Theta \) considered here are as follows.

1. Univariate unconditional MLE: as is suggested by Box and Jenkins [1976; pp. 212–219], this involves minimization of the sum of squares of univariate residuals. The Simplex search method of Nelder and Mead [1965] was used. We have found this method to be considerably more efficient than the linearization method suggested by Box and Jenkins or the method of Rosenbrock [1960] employed by O'Connell [1974]. The MLE method was modified in certain instances when the resulting values of \( \Phi \) and \( \Theta \) were thought to be hydrologically unreasonable; these constraints are described in the experimental design section. We note that these univariate MLE's together are not equivalent to the MLE's for \( \Phi \) and \( \Theta \) which are obtained by simultaneous optimization of the joint likelihood function. Joint estimation of the model parameters may result in more efficient estimators [Nelson, 1976; Mortarty and Salamon, 1980; Umashankar and Ledoiter, 1983].

2. Univariate method of moments: this involves estimation of \( \phi_i \) by

\[
\hat{\phi}_i = \tilde{\phi}_i
\]

and computation of \( \hat{\theta}_i \) by solution of the quadratic equation

\[
\hat{\theta}_i = \frac{1 - \phi_i \phi_i - \hat{\theta}_i}{1 + \hat{\theta}_i^2 - 2\phi_i \hat{\theta}_i}
\]

yielding

\[
\hat{\theta}_i = 0.5[1 + \phi_i(1 - 2\hat{\phi}_i)][\phi_i - \hat{\phi}_i]
\]

Method of moments solutions do not exist for certain combinations of \( \hat{\phi}_i \) and \( \hat{\phi}_i \). Screening criteria, described in the experimental design section, were employed to ensure that \( \hat{\phi}_i < \phi_i < 1 \) and \( \hat{\theta}_i < \theta_i \).

3. O'Connell's method: O'Connell [1974] performed extensive Monte Carlo experiments to relate an estimator of the Hurst coefficient, \( h \) and an estimator of the lag-one correlation coefficient to parameters of the univariate ARMA(1,1) model. The Hurst coefficient can be defined by

\[
e(R_i/S_i) = (n/2)\hat{\theta}_i
\]

where \( n \) is the sequence length, \( R_i \) is the range or maximum of the cumulative departures from the mean in a sequence of length \( n \), \( S_i \) is the standard deviation of the observations, and \( h \) is a measure of the long-term persistence of a hydrologic time series. For a general description of the Hurst effect, see the works by Hurst [1965], Mandelbrot and Wallis [1968], or Hipel [1975].

O'Connell's [1974] Monte Carlo experiments provide a description of the relationship between (1) the expected value of \( K \), an estimator of \( h \); (2) the expected value of a lag-one correlation coefficient estimator, \( E[\hat{\theta}_1] \); and (3) the population parameters \( \phi \) and \( \theta \) (note that subscripts have been dropped for convenience). O'Connell's experiments were performed by fixing \( \phi \) and \( \theta \) and numerically evaluating the expected value of \( K \) and \( \hat{\theta}_1 \) for sequences of length \( n = 25, 50, \text{ and } 100 \).

O'Connell [1974, 1975] suggests that these results can be used for estimation by mapping individual sample estimates \( K \) and \( \hat{\phi}_1 \) into estimates of \( \phi \) and \( \theta \) by assuming the former equal their expected values. Although O'Connell's results might be used in this manner via a table lookup procedure for individual applications, the Monte Carlo experiments described in this paper required automation of the procedure. A family of curves was fitted to O'Connell's results. As is shown in Appendix B, there was a narrow region in which \( K \) and \( \hat{\phi}_1 \) result in feasible values of \( \phi \) and \( \theta \). Therefore we chose to relate both \( \phi \) and \( \theta \) to \( K \) alone so as to obtain a solution in the majority of cases.

Two methods were considered for estimation of \( \Gamma \), given the coefficient matrices \( \Phi \) and \( \Theta \):

1. A quasi MLE estimate of \( \Gamma \) was obtained by using \( \Phi \) and \( \Theta \) in (1) (after subtraction of the at-site sample means) to obtain estimates of the residuals \( V_i \), as described in the work.
by Box and Jenkins [1976, pp. 212–218]. Our estimate of \( G \) was then

\[
\hat{G} = \frac{1}{n} \sum_{t=1}^{n} V_i V_i^T
\]  

(5)

2. It can be shown that \( S_0, \Phi, \Theta, \) and \( G \) satisfy

\[
S_0 = \Phi S_0 \Phi^T - \Theta G \Theta^T + G - \Theta G \Theta^T
\]  

(6)

A method-of-moments estimator of \( G \) was obtained by substituting \( \Phi, \Theta \) and

\[
S_{\text{mle}} = \frac{1}{n} \sum_{i=1}^{n} X_i X_i^T
\]  

(7)

for the population values in (6). Making use of the fact that \( \Phi \) and \( \Theta \) are diagonal, the resultant estimator of \( G \) has components

\[
\hat{G}_{ij} = \frac{(S_{\text{mle}})_{ij}(1 - \hat{\theta}_i \hat{\phi}_j)}{1 - \hat{\phi}_n \hat{\phi}_j - \hat{\theta}_i \hat{\phi}_j + \hat{\theta}_i \hat{\phi}_j}
\]  

(8)

Given \( \Phi, \Theta \), this estimator of \( G \) will reproduce the observed variance and lag-zero cross-covariances of the flows.

**Aggregate Flow Model**

The seventh model represents a different generation strategy from the first six. Rather than attempting to model flows at each site individually, a univariate ARMA(1,1) model was fit to the aggregate flows using maximum likelihood estimates of \( \phi \) and \( \theta \) and the method of moments estimator of the innovation's variance, as in (8). This will be shown, in general, to be about the best univariate technique from the standpoint of parameter estimation efficiency. The fitted model was used to generate aggregate annual flows which were then divided among the sites.

The aggregate flow \( X_i^* \) was apportioned between the two sites using the simple disaggregation model described by Stedinger and Vogel [1984]. In the two-site case, that model becomes

\[
X_{1i} = \beta X_i^* + W_i
\]  

(9)

\[
X_{2i} = (1 - \beta)X_i^* - W_i
\]  

(10)

The parameter \( \beta \) and the variance of the noise sequence \( W_i \) were selected so as to reproduce the covariance of \( X_i^* \) with \( X_{1i} \) and \( X_{2i} \) as well as the variances of both \( X_{1i} \) and \( X_{2i} \). Note that (10) can be omitted because \( X_{2i} \) is also given by \( X_{1i} - X_{1i}^* \).

Without loss of generality, we assume that the \( W_i \) will be independent of the \( X_i^* \). Then, to reproduce the covariance of \( X_i^* \) and \( X_{1i} \), \( \beta \) should be

\[
\beta = \frac{\text{Cov}(X_i^*, X_{1i})}{\text{Var}(X_i^*)}
\]  

(11)

We employed the estimator

\[
\hat{\beta} = \frac{\sum_{i=1}^{n} (X_i^*, X_{1i})}{\sum_{i=1}^{n} (X_i^*)^2}
\]  

(12)

Likewise, to reproduce the observed variance of \( X_{1i} \), the variance of the generated \( W_i \) was taken to be

\[
\hat{\sigma}_w^2 = \frac{1}{n} \sum_{i=1}^{n} (X_{1i} - \beta X_i^*)^2
\]  

(13)

This choice of \( \hat{\beta} \) and \( \hat{\sigma}_w^2 \) allows reproduction of the observed covariance between \( X_i^* \) and \( X_{1i} \), and the sample variance of \( X_{1i} \); it will also reproduce the observed covariance between \( X_i^* \) and \( X_{2i} \) and the sample variance of \( X_{2i} \).

In addition, it would be desirable to have a model which would reproduce the persistence structure of the individual \( X_{1i} \) and \( X_{2i} \) series. As is discussed in the works by Lane [1980] and Stedinger and Vogel [1984], it is often impossible within a disaggregation model structure to reproduce the sample lag-one correlation of the \( X_{1i} \) series. However, a feasible and reasonable substitute is to reproduce the observed correlation of the \( W_i \) series. Thus the \( W_i \) were generated using the model

\[
W_i = \alpha W_{i-1} + V_i
\]  

(14)

with

\[
\hat{\alpha} = \frac{1}{n-2} \sum_{s=2}^{n} (X_{1i} - \beta X_i^*)(X_{s-1} - \beta X_{s-1}^*)
\]  

(15)

and independent zero-mean innovations \( V_i \) with variance

\[
\text{Var} [V_i] = (1 - \hat{\alpha}^2) \hat{\sigma}_w^2
\]  

(16)

One exception was made; should \( \hat{\alpha} \) be negative, a value of zero was substituted. We felt it was physically unreasonable for higher than expected flows in year \( t - 1 \) at site 1 to imply lower than expected flows at site 1 in year \( t \).

**Model Estimation Assessment Criteria**

Assessment measures for model performance evaluation should determine how well fitted models replicate important population characteristics. Our assessment procedure consisted of two stages. The first was to determine whether the fitted parameters, based on "historical" data, replicate population parameters, such as means, variances, autocorrelations, and cross correlations. In general, the population parameters are not known. However, this problem was avoided by using synthetically generated data (pseudo-historical) in lieu of historical sequences. A second stage (stage II) was used to determine whether flow sequences generated using the estimated parameters reproduce drought frequency and cross-drought correlation characteristics similar to sequences generated by the true model. This is especially important as the parameter estimation procedures sometimes resulted in a fitted model appreciably different from the prototype.

The stage I at-site assessment criteria included the bias and root mean square error of the fitted models' lag-one correlation coefficient and the estimated variance of flows at both sites. An additional indicator of flow persistence,

\[
\Gamma_k = \frac{1}{k \hat{\sigma}^2} \text{Var} \left[ \sum_{i=1}^{k} X_i \right]
\]  

(17)

for \( k = 4 \) and 10 was also employed. \( \Gamma_k \) would equal 1 for independent flows and increases as the modeled persistence among the \( X_i \) increases.

Because both the pseudo-historical and generated flows were normally distributed, their distribution is completely specified by their mean, variance, and autocorrelation function. Thus when examining the persistence or variance of flows over extended time periods, it is sufficient to focus on their variance, and particularly their correlations. \( \Gamma_k \) provides a measure of the normalized variance of the average flow over any \( k \)-year period; its (population) value is

\[
\Gamma_k = 1 + \frac{\sum_{j=1}^{k-1} (1 - j/k) \rho_j}{k - 1}
\]  

(18)
TABLE 1. Description of Diagonal ARMA(1,1) Models Used to Generate Flows

<table>
<thead>
<tr>
<th>Case</th>
<th>( \phi_{ii} )</th>
<th>( \theta_{ii} )</th>
<th>( \rho_{i}(i,i) )</th>
<th>( \rho_{i}(i,j) = 0.7 )</th>
<th>( \rho_{i}(i,j) = 0.9 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.20</td>
<td>0.000</td>
<td>0.200</td>
<td>0.140</td>
<td>0.180</td>
</tr>
<tr>
<td>1*</td>
<td>0.40</td>
<td>0.177</td>
<td>0.233</td>
<td>0.163</td>
<td>0.210</td>
</tr>
<tr>
<td>2</td>
<td>0.60</td>
<td>0.363</td>
<td>0.267</td>
<td>0.187</td>
<td>0.240</td>
</tr>
<tr>
<td>2*</td>
<td>0.80</td>
<td>0.572</td>
<td>0.300</td>
<td>0.210</td>
<td>0.270</td>
</tr>
<tr>
<td>3</td>
<td>0.85</td>
<td>0.617</td>
<td>0.333</td>
<td>0.233</td>
<td>0.300</td>
</tr>
<tr>
<td>3*</td>
<td>0.90</td>
<td>0.675</td>
<td>0.376</td>
<td>0.257</td>
<td>0.330</td>
</tr>
<tr>
<td>5</td>
<td>0.95</td>
<td>0.758</td>
<td>0.400</td>
<td>0.280</td>
<td>0.360</td>
</tr>
</tbody>
</table>

*Intermediate cases were used to produce figures; they do not appear in tables. Here \( \rho_{i}(i,j) \) refers to \( \rho_{i}(1,1) \) and \( \rho_{i}(2,2) \), the auto-correlations of flows at site 1 and 2, respectively; \( \rho_{i}(1,2), i \neq j \), refers to \( \rho_{i}(1,2) \) and \( \rho_{i}(2,1) \), where \( \rho_{i}(1,2) = \rho_{i}(2,1) \) and \( \rho_{i}(1,1) = \rho_{i}(2,2) \).

All of the statistics described above can be calculated after the model parameters have been estimated using stage I synthetic flows. We also calculated at stage I and II the at-site required reservoir capacity \( S \) necessary to satisfy a given hydrologist would reasonably consider fitting an ARMA(1,1) model; it was felt that in practice sequences not passing such a test would most likely be modeled as independent noise.

For each generating model, bivariate 50-year sequences were generated using the startup algorithm described in Appendix A. An initial screening was performed, and sequences were rejected for which either \( \rho_{i}(1,1) \leq 0.05 \) or \( \rho_{i}(2,2) \leq 0.05 \). Additional sequences were generated to provide a total of 1000 sequences which passed this test. The screening test was designed to limit consideration to those sequences to which a hydrologist would reasonably consider fitting an ARMA(1,1) model; it was felt that in practice sequences not passing such a test would most likely be modeled as independent noise.

Each of the seven model fitting techniques described in the previous section was applied to the 1000 generated sequences. The first six models consisted of univariate maximum likelihood, method of moments (MOM), and O'Connell's [1974] method (OCE) for univariate estimation of \( \phi_i \) and \( \theta_i \); coupled with MOM or conditional MLE (product estimate of \( \Gamma \) from fitting method was univariate MLE applied to the aggregate innovation series) for estimation of \( \Gamma \). The seventh fit was designed to limit consideration to those sequences to which a hydrologist would reasonably consider fitting an ARMA(1,1) model; it was felt that in practice sequences not passing such a test would most likely be modeled as independent noise.

All of the statistics described above can be calculated after the model parameters have been estimated using stage I synthetic flows. We also calculated at stage I and II the at-site required reservoir capacity \( S \) necessary to satisfy a given hydrologist would reasonably consider fitting an ARMA(1,1) model; it was felt that in practice sequences not passing such a test would most likely be modeled as independent noise.

For each generating model, bivariate 50-year sequences were generated using the startup algorithm described in Appendix A. An initial screening was performed, and sequences were rejected for which either \( \rho_{i}(1,1) \leq 0.05 \) or \( \rho_{i}(2,2) \leq 0.05 \). Additional sequences were generated to provide a total of 1000 sequences which passed this test. The screening test was designed to limit consideration to those sequences to which a hydrologist would reasonably consider fitting an ARMA(1,1) model; it was felt that in practice sequences not passing such a test would most likely be modeled as independent noise.


dynamic flows with a method of moments estimate of the variance, and disaggregation to the individual sites as described in the preceding section.

Additional restrictions were applied to some of the estimation procedures. For MLE and MOM estimates of \( \phi_i \) and \( \theta_i \) it was required that \( 0 < \theta_i < \phi_i < 1 \). When either of these constraints was violated we set \( \phi_i = \rho_{i}(i,i) \) and \( \theta_i = 0 \); the initial screening guaranteed \( \rho_{i}(i,i) \geq 0.05 \). For OCE, estimated values of \( \Gamma \) less than 0.670 were set to 0.670; \( \Gamma = 0.670 \) corresponds to \( E[\phi_i] = 0 \) for our ARMA(1,1) processes. This

TABLE 2. Estimates of the Mean and Root Mean Square Error of Statistics Describing the Persistence of Modeled Flows at Each Site

<table>
<thead>
<tr>
<th>Estimation Method</th>
<th>( E[\phi_i] )</th>
<th>( \text{RMSE}[\phi_i] )</th>
<th>( E[\Gamma_{14}] )</th>
<th>( \text{RMSE}[\Gamma_{14}] )</th>
<th>( E[\Gamma_{10}] )</th>
<th>( \text{RMSE}[\Gamma_{10}] )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case 1 MLE</td>
<td>0.24</td>
<td>0.11</td>
<td>1.47</td>
<td>0.28</td>
<td>1.72</td>
<td>0.57</td>
</tr>
<tr>
<td>MOM</td>
<td>0.21</td>
<td>0.11</td>
<td>1.40</td>
<td>0.25</td>
<td>1.59</td>
<td>0.45</td>
</tr>
<tr>
<td>DISAG (( \rho_{i}(1,2) = 0.7 ))</td>
<td>0.24</td>
<td>0.10</td>
<td>1.47</td>
<td>0.24</td>
<td>1.69</td>
<td>0.50</td>
</tr>
<tr>
<td>DISAG (( \rho_{i}(1,2) = 0.9 ))</td>
<td>0.24</td>
<td>0.11</td>
<td>1.48</td>
<td>0.27</td>
<td>1.71</td>
<td>0.53</td>
</tr>
<tr>
<td>True values</td>
<td>0.20</td>
<td>0.11</td>
<td>1.34</td>
<td>0.14</td>
<td>1.44</td>
<td>0.99</td>
</tr>
<tr>
<td>Case 2 MLE</td>
<td>0.31</td>
<td>0.15</td>
<td>1.74</td>
<td>0.41</td>
<td>2.47</td>
<td>1.09</td>
</tr>
<tr>
<td>MOM</td>
<td>0.25</td>
<td>0.15</td>
<td>1.55</td>
<td>0.43</td>
<td>1.94</td>
<td>1.07</td>
</tr>
<tr>
<td>DISAG (( \rho_{i}(1,2) = 0.7 ))</td>
<td>0.31</td>
<td>0.12</td>
<td>1.72</td>
<td>0.35</td>
<td>2.38</td>
<td>0.98</td>
</tr>
<tr>
<td>DISAG (( \rho_{i}(1,2) = 0.9 ))</td>
<td>0.31</td>
<td>0.13</td>
<td>1.73</td>
<td>0.38</td>
<td>2.43</td>
<td>1.05</td>
</tr>
<tr>
<td>True values</td>
<td>0.30</td>
<td>0.11</td>
<td>1.79</td>
<td>0.26</td>
<td>2.66</td>
<td>1.09</td>
</tr>
<tr>
<td>Case 3 MLE</td>
<td>0.35</td>
<td>0.19</td>
<td>1.90</td>
<td>0.59</td>
<td>3.04</td>
<td>1.88</td>
</tr>
<tr>
<td>MOM</td>
<td>0.23</td>
<td>0.24</td>
<td>1.49</td>
<td>0.78</td>
<td>1.87</td>
<td>2.46</td>
</tr>
<tr>
<td>DISAG (( \rho_{i}(1,2) = 0.7 ))</td>
<td>0.33</td>
<td>0.16</td>
<td>1.84</td>
<td>0.55</td>
<td>2.85</td>
<td>1.84</td>
</tr>
<tr>
<td>DISAG (( \rho_{i}(1,2) = 0.9 ))</td>
<td>0.34</td>
<td>0.18</td>
<td>1.87</td>
<td>0.56</td>
<td>2.95</td>
<td>1.86</td>
</tr>
<tr>
<td>True values</td>
<td>0.40</td>
<td>0.21</td>
<td>2.16</td>
<td>0.75</td>
<td>4.16</td>
<td>2.00</td>
</tr>
</tbody>
</table>

Ninety-five percent confidence intervals for \( E[\phi_i] \), \( \text{RMSE}[\phi_i] \), \( E[\Gamma_{14}] \), \( \text{RMSE}[\Gamma_{14}] \), \( E[\Gamma_{10}] \), and \( \text{RMSE}[\Gamma_{10}] \) are within ± 6, ± 5, ± 3, ± 5, ± 4, and ± 9%, respectively, of the reported values.
default value for $K$ resulted in default values of $\phi = 0.880$ and $\theta = 0.844$ which are still indicative of long-term persistence.

**MONTE CARLO RESULTS**

We consider first the stage I performance of the seven ARMA(1,1) models and parameter estimation procedures with the constraints described above. The constraints imposed are designed to improve the fidelity or precision with which the fitted models will describe the distribution of streamflow sequences one would expect to see in practice. For example, estimates such that $0 < \phi < \theta < 1$ are admissible [see Box and Jenkins, 1976, pp. 76–77, 520] but result in a negative lag-one correlation coefficients. Thus $\phi = \rho_1 \geq 0.05$ and $\theta = 0$ are almost surely better parameter estimates based on hydrologic experience with large numbers of catchments [Yevjevich, 1963]. However, if $\phi$ is near 1 and $\theta$ is not much less than $\phi$, then the occasional substitution of $(\hat{\phi}, 0)$ for $(\phi, \theta)$ can have a dramatic impact on the bias and variance of $\phi$ and $\theta$ and hence their mean square error as well. Likewise, it is important to note that for all $\phi$ and $\theta$ such that $\theta = \phi$, an equivalent model is obtained for $\phi = \theta = 0$. For these reasons the bias and mean square error of the model parameter estimates $\phi$ and $\theta$ can have little practical significance in that the relationship between these values and the correlations among the generated flows is highly nonlinear if not discontinuous; it is more appropriate and meaningful to focus on hydrologically important characteristics of the fitted models.

The characteristics considered to be most important here were the fitted variance of the flows, the fitted lag-one correlation coefficient $\rho_1$ (measuring short-term persistence), and $\Gamma_4$ and $\Gamma_{10}$ which serve as simple measures of modelled longer-term persistence. Table 2 and Figure 1 report the means and root mean square error (RMSE) of the last three statistics for the three standard cases. Also included are the corresponding statistics obtained using the aggregate-flow ARMA(1,1) model in conjunction with disaggregation. For the diagonal ARMA models the results in Table 2 are independent of the specified value of $\rho_d(1,2)$. However, $\rho_d(1,2)$ does affect the performance of the disaggregation model, though only slightly as demonstrated by our results.

Figure 1 compares the three at-site $\phi$ and $\theta$ estimation procedures along with the disaggregation model performance with $\rho_d(1,2) = 0.7$. In terms of the RMSE of flow persistence statistics such as $\rho_1$, $\Gamma_4$, and $\Gamma_{10}$, the disaggregation model generally did the best while our formulation of O'Connell's [1974] method was clearly worst. The method of moment and MLE estimation procedures as constrained generally performed almost as well, and sometimes better, than the disaggregation model. It is interesting that the relative performance of the MOM procedure improves in the low persistence case while that of the MLE procedures improves in the high persistence case. Much of the difference between the RMSE of MOM estimators and those of MLE and the disaggregation procedures is due to the large bias in MOM estimators in the high persistence cases (see the bottom panels in Figure 1).

**Estimates of $G$**

The previous discussion focused on the modeled persistence of flows at each site. With each method of estimating $\Phi$ and $\Theta$, two procedures were considered for estimating the covari-
<table>
<thead>
<tr>
<th>Method</th>
<th>$\phi$, $\Theta$</th>
<th>$G$</th>
<th>$E[\theta_1^2]$</th>
<th>RMSE[\theta_1^2]</th>
<th>$E[\theta_2^2]$</th>
<th>RMSE[\theta_2^2]</th>
<th>$E[\rho_0(1, 2)]$</th>
<th>RMSE[\rho_0(1, 2)]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case 1</td>
<td>MLE, MOM, OCL, DISAG</td>
<td>MOM</td>
<td>0.063</td>
<td>0.014</td>
<td>0.21</td>
<td>0.047</td>
<td>0.70</td>
<td>0.079</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MLE</td>
<td>0.062</td>
<td>0.014*</td>
<td>0.21</td>
<td>0.046*</td>
<td>0.68</td>
<td>0.081</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MOM</td>
<td>0.061</td>
<td>0.013</td>
<td>0.21</td>
<td>0.045</td>
<td>0.68</td>
<td>0.080</td>
</tr>
<tr>
<td></td>
<td></td>
<td>OCL</td>
<td>0.069</td>
<td>0.021</td>
<td>0.23</td>
<td>0.064</td>
<td>0.68</td>
<td>0.091</td>
</tr>
<tr>
<td></td>
<td>True values</td>
<td></td>
<td>0.0625</td>
<td>0.2125</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Case 2</td>
<td>MLE, MOM, OCL, DISAG</td>
<td>MOM</td>
<td>0.060</td>
<td>0.015</td>
<td>0.20</td>
<td>0.050</td>
<td>0.70</td>
<td>0.085</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MLE</td>
<td>0.062</td>
<td>0.017*</td>
<td>0.21</td>
<td>0.053*</td>
<td>0.66</td>
<td>0.090</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MOM</td>
<td>0.059</td>
<td>0.015</td>
<td>0.20</td>
<td>0.049</td>
<td>0.67</td>
<td>0.088</td>
</tr>
<tr>
<td></td>
<td></td>
<td>OCL</td>
<td>0.079</td>
<td>0.039</td>
<td>0.26</td>
<td>0.114</td>
<td>0.65</td>
<td>0.104</td>
</tr>
<tr>
<td></td>
<td>True values</td>
<td></td>
<td>0.0625</td>
<td>0.2125</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Case 3</td>
<td>MLE, MOM, OCL, DISAG</td>
<td>MOM</td>
<td>0.050</td>
<td>0.018</td>
<td>0.17</td>
<td>0.063</td>
<td>0.70</td>
<td>0.096</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MLE</td>
<td>0.058</td>
<td>0.026*</td>
<td>0.19</td>
<td>0.074*</td>
<td>0.64</td>
<td>0.118</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MOM</td>
<td>0.049</td>
<td>0.019</td>
<td>0.16</td>
<td>0.065</td>
<td>0.66</td>
<td>0.098</td>
</tr>
<tr>
<td></td>
<td></td>
<td>OCL</td>
<td>0.073</td>
<td>0.040</td>
<td>0.24</td>
<td>0.114</td>
<td>0.65</td>
<td>0.114</td>
</tr>
<tr>
<td></td>
<td>True values</td>
<td></td>
<td>0.0625</td>
<td>0.2125</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Here $\rho_0(1, 2) = 0.7$.

* Ninety-five percent confidence intervals for $E[\theta_1^2]$, RMSE[\theta_1^2], $E[\theta_2^2]$, RMSE[\theta_2^2], $E[\rho_0(1, 2)]$, and RMSE[\rho_0(1, 2)] are within ±3, ±7, ±3, ±7, ±1, and ±5%, respectively, of the reported values except for the values marked with asterisk whose 95% confidence intervals are within 16% of the reported values.

ance matrix $G$ of the innovations. This matrix determines the fitted variances and lag-zero correlation (or covariance) of flows at the individual sites. Table 3 reports the mean and RMSE of the fitted variances and lag-zero correlations for $\rho_0(1, 2) = 0.7$; the corresponding table for $\rho_0(1, 2) = 0.9$ may be found in the microfiche appendix.¹ (The microfiche appendix contains seven additional tables expanding upon Tables 3-5 and 8-9. The microfiche appendix also contains 12 additional figures displaying the root mean square error and bias of $\theta_1^2$, $\theta_2^2$, $\rho_0(1, 2)$, and $\Gamma_{10}$ for all four cases and $\rho_1$, $\Gamma_4$, and $\Gamma_{10}$ for the nondiagonal model simulations.) Of the six univariate fitting procedures, all those which reproduce $S_0$ will yield the same estimates of these statistics. However, the pseudo-MLE estimators of $G$ will depend on the procedure used to estimate $\Phi$ and $\Theta$.

From Table 3 and Figure 2 one can see that in terms of their ability to estimate the lag-zero cross-correlation, $\rho_0(1, 2)$, MLE procedures exhibited a substantial downward bias and also a larger RMSE than the estimators obtained using $S_0$. MOM estimators for $G$ are clearly to be preferred. The only close competitor was MOM estimates of $\phi$ and $\theta$ with MLE $G$ estimators; however, this combination makes little philosophic sense and offers no statistical advantages in terms of estimator performance. Moreover, it is clearly inferior in terms of the RMSE of $\rho_0(1, 2)$ estimators. Only MOM estimators of $G$, corresponding to reproduction of $S_0$, will be employed in the subsequent analysis. Lettenmaier [1980] and Stedinger [1981] have also documented the poor performance of MLE estimators of $G$. It should be noted that the disaggregation model also reproduces $S_0$, hence it produces $\sigma_1^2$, $\sigma_2^2$, and $\rho_0(1, 2)$ estimators equivalent to those of diagonal ARMA(1,1) models which are based upon $S_0$.

**Correlations of the Aggregate Flows**

Two approaches were employed to analyze the coherency, timing, or correspondence of streamflows, and particularly drought flows, which would be generated by the models for the two sites. The first approach is to analyze the properties of the aggregate flows that would result from use of the flow models with the various parameter estimates. Such an analysis is appropriate if the reservoir storage system to be simulated is highly interconnected so that the aggregate inflow is of paramount importance.

Table 4, analogous to Table 2, reports the mean and RMSE of $\rho_1$, $\Gamma_4$, and $\Gamma_{10}$ for the aggregate flows. Clearly, our implementation of O'Connell's [1974] method performed very poorly overall. For the other three methods, the ranking and the relative differences were dependent on the underlying parameter values. For a cross correlation $\rho_0(1, 2)$ of 0.9 (see Table in appendix) there was almost no difference in the RMSE's of $\rho_1$, $\Gamma_4$, and $\Gamma_{10}$ except in the most persistent case when the method of moments' performance deteriorated. For $\rho_0(1, 2) = 0.7$, the results were more varied. The method of

¹Supplement is available with entire article on microfiche. Order from American Geophysical Union, 2000 Florida Avenue, N.W., Washington, D.C. 20009. Package W80-001; $2.50. Payment must accompany order.
TABLE 4. Estimates of the Mean and Root Mean Square Error of Statistics Describing the Persistence of Aggregate Flows Which Would be Generated by the Fitted Models

<table>
<thead>
<tr>
<th>Method</th>
<th>ϕ, θ</th>
<th>G</th>
<th>E[\hat{ϕ}(a)]</th>
<th>RMSE[\hat{ϕ}(a)]</th>
<th>E[\hat{Γ}_4]</th>
<th>RMSE[\hat{Γ}_4]</th>
<th>E[\hat{Γ}_{10}]</th>
<th>RMSE[\hat{Γ}_{10}]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case 1</td>
<td>MLE</td>
<td>MOM</td>
<td>0.24</td>
<td>0.10</td>
<td>1.47</td>
<td>0.24</td>
<td>1.71</td>
<td>0.48</td>
</tr>
<tr>
<td></td>
<td>MOM</td>
<td>MOM</td>
<td>0.21</td>
<td>0.09</td>
<td>1.40</td>
<td>0.21</td>
<td>1.59</td>
<td>0.37</td>
</tr>
<tr>
<td></td>
<td>OCL</td>
<td>MOM</td>
<td>0.17</td>
<td>0.15</td>
<td>1.48</td>
<td>0.43</td>
<td>2.19</td>
<td>1.29</td>
</tr>
<tr>
<td></td>
<td>DISAG</td>
<td>MOM</td>
<td>0.25</td>
<td>0.11</td>
<td>1.49</td>
<td>0.29</td>
<td>1.74</td>
<td>0.59</td>
</tr>
<tr>
<td></td>
<td>True values</td>
<td></td>
<td>0.20</td>
<td>1.34</td>
<td>1.44</td>
<td>4.66</td>
<td>11.10</td>
<td>23.16</td>
</tr>
<tr>
<td>Case 2</td>
<td>MLE</td>
<td>MOM</td>
<td>0.31</td>
<td>0.12</td>
<td>1.74</td>
<td>0.34</td>
<td>2.46</td>
<td>0.92</td>
</tr>
<tr>
<td></td>
<td>MOM</td>
<td>MOM</td>
<td>0.25</td>
<td>0.13</td>
<td>1.54</td>
<td>0.38</td>
<td>1.93</td>
<td>0.96</td>
</tr>
<tr>
<td></td>
<td>OCL</td>
<td>MOM</td>
<td>0.35</td>
<td>0.21</td>
<td>2.00</td>
<td>0.63</td>
<td>3.57</td>
<td>1.80</td>
</tr>
<tr>
<td></td>
<td>DISAG</td>
<td>MOM</td>
<td>0.32</td>
<td>0.14</td>
<td>1.76</td>
<td>0.40</td>
<td>2.50</td>
<td>1.10</td>
</tr>
<tr>
<td></td>
<td>True values</td>
<td></td>
<td>0.30</td>
<td>1.79</td>
<td>1.44</td>
<td>6.66</td>
<td>22.22</td>
<td>44.44</td>
</tr>
<tr>
<td>Case 3</td>
<td>MLE</td>
<td>MOM</td>
<td>0.34</td>
<td>0.16</td>
<td>1.88</td>
<td>0.51</td>
<td>3.01</td>
<td>1.70</td>
</tr>
<tr>
<td></td>
<td>MOM</td>
<td>MOM</td>
<td>0.23</td>
<td>0.22</td>
<td>1.49</td>
<td>0.75</td>
<td>1.86</td>
<td>2.41</td>
</tr>
<tr>
<td></td>
<td>OCL</td>
<td>MOM</td>
<td>0.41</td>
<td>0.22</td>
<td>2.18</td>
<td>0.63</td>
<td>4.04</td>
<td>1.69</td>
</tr>
<tr>
<td></td>
<td>DISAG</td>
<td>MOM</td>
<td>0.35</td>
<td>0.18</td>
<td>1.91</td>
<td>0.56</td>
<td>3.06</td>
<td>1.83</td>
</tr>
<tr>
<td></td>
<td>True values</td>
<td></td>
<td>0.40</td>
<td>2.16</td>
<td>1.44</td>
<td>8.66</td>
<td>34.66</td>
<td>70.33</td>
</tr>
</tbody>
</table>

Here ρ_0(1,2) = 0.7. Ninety-five percent confidence intervals for E[\hat{ϕ}(a)], RMSE[\hat{ϕ}(a)], E[\hat{Γ}_4] and RMSE[\hat{Γ}_4], E[\hat{Γ}_{10}] and RMSE[\hat{Γ}_{10}] are within ±5, ±5, ±2, ±5, ±3, and ±7, respectively, of the reported values.

Moments went from best to worst as the persistence of the underlying model increased. Over the whole range, the disaggregation model, which explicitly models the aggregate flows, did slightly worse than a diagonal ARMA(1,1) model with MLE ϕ_i and θ_i estimators and a MOM estimator of G. The results show a curious situation where the disaggregation model, which explicitly models aggregate flows, generally yields the best at-site estimators of ϕ_i and θ_i, while the MOM estimator of G. The results show a curious situation where the disaggregation model, which explicitly models aggregate flows, generally yields the best at-site estimators of ϕ_i and θ_i, while the diagonal ARMA model framework yielded better estimators of ρ_i(a), Γ_4, and Γ_{10} for aggregate flows. We attribute this reversal, which is most apparent for a cross-site correlation ρ_0(1,2) of 0.70 as opposed to 0.90, to an averaging which occurred in the parameter estimation processes. That is, the disaggregation model to some extent yields at-site autocorrelations which are an average of the two at-site historical values: given that the ρ_i(i,ii) for i = 1,2 and various lags k are by assumption identical, such an averaging of the two estimators is advantageous here. On the other hand, the diagonal ARMA models employed individual MLE or MOM estimators of each ϕ_i and θ_i; thus the resultant values of ρ_i(a), Γ_4, and Γ_{10} were the average of the properties of these independently estimated (but not statistically independent) at-site models.

Model Assessment Measures: Cross-Site Drought Criteria

A primary consideration in multisite streamflow generation for multiple reservoir operation studies is that drought sequences at all sites occur with a distribution resembling that of the natural flows. An obvious and widely used performance measure is the fidelity of the fitted cross-correlation coefficient ρ_0(1,2) of the flows at the two sites. Persistence measures of the aggregate annual flows such as Γ_4 provide another useful measure which reflects whether the properties of the total flow are similar to those of the population. However, ρ_i(a), Γ_4, Γ_{10}, and ρ_0(1,2) alone may be insufficient to demonstrate whether or not the pseudo-historic and generated sequences are compatible in terms of drought occurrence and severity. A potentially useful approach for measuring the joint occurrence of droughts at the two sites is to evaluate the simulated simultaneous operation of two hypothetical bottomless reservoirs, one at each site.

Consider a demand level D. For a bottomless reservoir initially full, the drawdown after t periods can be calculated from:

\[ S_i(t+1) = S_i(t) - D - Q_i(t) \]

where \( Q_i(t) = X_i(t) + \mu_i \) is the inflow at the site. The sequence \( S_i(t) \) defines the cumulative impact of low flows on storage levels for demand D at site i.

One reasonable measure of multisite drought correlation is the lag-zero correlation of the deficit levels \( S_i^1 \) and \( S_i^2 \) at the two sites,

\[ \rho_d = \frac{\sum_{i=1}^{n} (S_i^1 - \bar{S}) (S_i^2 - \bar{S})}{\left[ \sum_{i=1}^{n} (S_i^1 - \bar{S})^2 \sum_{i=1}^{n} (S_i^2 - \bar{S})^2 \right]^{1/2}} \]

Another potentially useful measure focuses on the worst drought record. Let \( t^* \) be the time at which the largest \( S_i \) occurs at site i. Then a drought coincidence statistic is

\[ \gamma_T = 1 - \frac{3n}{n^2 - 1} E(t_1^* - t_2^*) \]

where \((n^2 - 1)/3n\) is the expected value of \( |t_1^* - t_2^*| \) if the drawdown sequences at the two sites are expected independent. If the drought sequences were perfectly correlated, then \( E(t_1^* - t_2^*) \) would be zero and \( \gamma_T \) would equal 1, because the critical point in both drawdown sequences would have occurred simultaneously. In general, \( \gamma_T \) will be in the range 0 ≤ \( \gamma_T \) ≤ 1 for sequences with positive cross correlation.

A final measure \( r_d \) which we have termed drought coherency, is defined as

\[ r_d = \frac{\sum_{i=t_1}^{t_2} I_i}{(t_2 - t_1)} \]

where \( (t_1, t_2) \) solves \( \sum_{i=t_1}^{t_2} I_i = \max_{t_1, t_2} (t_2 - t_1) \) such that for all \( t \in \{t_1, \ldots, t_2\} \), \( S_i^1 > 0 \) and where \( I_i = 1 \) for \( S_i^1 > 0 \) and zero otherwise. Therefore \( r_d \) is the fraction of the periods during the interval \( (t_1, t_2) \) in which site 2 is in deficit state \( S_i^1 > 0 \), where \( (t_1, t_2) \) is the longest continuous deficit period at site 1.

These drought statistics were employed to test whether concurrent drought flows at the two sites occur with the appropri-
TABLE 5. Estimates of the Mean of Drought Correlation, Concurrency, Coherency, and Required Storage for True and Fitted Models for Demand = 0.90 MAF and $\rho_d(1,2) = 0.7$

<table>
<thead>
<tr>
<th>Method</th>
<th>$\phi, \theta$</th>
<th>$G$</th>
<th>$E[\rho_d]$ (0.7)</th>
<th>$E[\gamma_T]$ (0.7)</th>
<th>$E[\gamma_s]$ (0.7)</th>
<th>$E[S]$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case 1</td>
<td>MLE</td>
<td>MOM</td>
<td>0.55</td>
<td>0.36</td>
<td>0.78</td>
<td>1.2</td>
</tr>
<tr>
<td></td>
<td>MOM</td>
<td>MOM</td>
<td>0.55</td>
<td>0.37</td>
<td>0.78</td>
<td>1.1</td>
</tr>
<tr>
<td></td>
<td>OCL</td>
<td>MOM</td>
<td>0.52</td>
<td>0.34</td>
<td>0.75</td>
<td>1.3</td>
</tr>
<tr>
<td></td>
<td>DISAG</td>
<td></td>
<td>0.59</td>
<td>0.42</td>
<td>0.78</td>
<td>1.2</td>
</tr>
<tr>
<td></td>
<td>True values</td>
<td></td>
<td>0.58</td>
<td>0.36</td>
<td>0.78</td>
<td>1.0</td>
</tr>
<tr>
<td>Case 2</td>
<td>MLE</td>
<td>MOM</td>
<td>0.53</td>
<td>0.36</td>
<td>0.75</td>
<td>1.5</td>
</tr>
<tr>
<td></td>
<td>MOM</td>
<td>MOM</td>
<td>0.54</td>
<td>0.37</td>
<td>0.76</td>
<td>1.4</td>
</tr>
<tr>
<td></td>
<td>OCL</td>
<td>MOM</td>
<td>0.53</td>
<td>0.37</td>
<td>0.73</td>
<td>1.8</td>
</tr>
<tr>
<td></td>
<td>DISAG</td>
<td></td>
<td>0.59</td>
<td>0.43</td>
<td>0.77</td>
<td>1.5</td>
</tr>
<tr>
<td></td>
<td>True values</td>
<td></td>
<td>0.55</td>
<td>0.35</td>
<td>0.77</td>
<td>1.4</td>
</tr>
<tr>
<td>Case 3</td>
<td>MLE</td>
<td>MOM</td>
<td>0.51</td>
<td>0.33</td>
<td>0.69</td>
<td>2.0</td>
</tr>
<tr>
<td></td>
<td>MOM</td>
<td>MOM</td>
<td>0.52</td>
<td>0.32</td>
<td>0.72</td>
<td>1.7</td>
</tr>
<tr>
<td></td>
<td>OCL</td>
<td>MOM</td>
<td>0.52</td>
<td>0.34</td>
<td>0.69</td>
<td>2.3</td>
</tr>
<tr>
<td></td>
<td>DISAG</td>
<td></td>
<td>0.55</td>
<td>0.36</td>
<td>0.71</td>
<td>2.1</td>
</tr>
<tr>
<td></td>
<td>True values</td>
<td></td>
<td>0.52</td>
<td>0.36</td>
<td>0.73</td>
<td>1.9</td>
</tr>
</tbody>
</table>

Ninety-five percent confidence intervals for $E[\rho_d]$, $E[\gamma_T]$, $E[\gamma_s]$, and $E[S]$ are within ±4, ±14, ±3, and ±11% of the reported values.

ate frequency and severity. Table 5 reports the expected values of the drought correlation $\rho_d$, drought coincidence $\gamma_T$, and drought coherency $\gamma_s$ statistics for a demand $D$ of 90% of the true mean annual flow. Only the expected values of these statistics calculated at stage 2 with each 50-year flow sequence generated using a unique set of fitted model parameters are reported. The observed variances are not given because they reflect both the variation in the parameter estimators and also the variability of these statistics across 50-year flow sequences generated from the respective models at stage 2. With a few exceptions, all of the models generally did about the same. The disaggregation model tended to produce higher values of $\rho_d$ and $\gamma_T$ reflecting, perhaps, higher lagged cross-correlations among the flows.

Table 5 also includes the expected value of the 50-year required storage $S$ for a demand $D$ of 90% of the true mean annual flow and a reservoir initially full. Except for the method of moments, the methods tended to over estimate $E[S]$. O’Connell’s [1974] method produced the largest values; this is consistent with the values of $\rho_d$, $\gamma_T$, and $\gamma_s$ reported in Table 2 and Figure 1.

**Nondiagonal ARMA(1,1) Models**

To this point, our attention has been restricted to ARMA(1,1) models with diagonal $\Phi$ and $\Theta$ matrices. It is important to evaluate how well diagonal models perform if the flows are actually generated by a more general model. Although the universe of possible generating models is very large, we restrict our attention to the case where $\Phi$ and $\Theta$ are symmetric with

\[
\Phi = \begin{bmatrix} \phi & \eta \phi \\ \eta \phi & \phi \end{bmatrix}, \quad \Theta = \begin{bmatrix} \theta & \eta \theta \\ \eta \theta & \theta \end{bmatrix}
\]

Further, we take $\phi > 0$, so as $\eta$ approaches zero the diagonal models considered earlier are obtained in the limit.

The assumptions of symmetry and equal diagonal elements are reasonable if streamflows arise from hydrologically similar drainage basins. These assumptions have the further practical advantage of reducing the number of degrees of freedom that need to be considered without sacrificing the essential characteristics of the problem.

**Theoretical Analysis**

Although the nondiagonal model is much more general than the diagonal model, careful consideration of the characteristics for the nondiagonal case restricts the range of parameters that may be considered hydrologically reasonable. The intrinsic behavior of the nondiagonal model can be illustrated by writing $X_t$ as a linear combination of the eigenvectors, $E_1$ and $E_2$, of $\Phi$. In particular, the $\Phi$ given in (23) has eigenvalues $\lambda = \phi(1 \pm \eta)$, with associated eigenvectors $E_1 = (1 \ 1)^T$ and $E_2 = (1 \ -1)^T$ for $\eta \neq 0$ and eigenvectors $E_1 = (1 \ 0)^T$ and $E_2 = (0 \ 1)^T$ for $\eta = 0$ corresponding to the double eigenvalue of $\lambda = \phi$.

Consider first the case with $\eta \neq 0$. Let

\[
X_t = z_t E_1 + y_t E_2
\]

Here $z_t$ and $y_t$ are univariate time series corresponding to the first and second eigenvectors. Because $\Phi$ is symmetric, $E_1$ and $E_2$ are orthogonal vectors.

Substituting (24) into (1) and decomposing the innovation terms into their contribution to each mode of variation, one finds that the time-series models for $z_t$ and $y_t$ are

\[
z_t = \phi(1 + \eta)z_{t-1} + E_1^T (V_t - \Theta V_{t-1})
\]

\[
y_t = \phi(1 - \eta)y_{t-1} + E_2^T (V_t - \Theta V_{t-1})
\]

Clearly, the terms $E_1^T (V_t - \Theta V_{t-1})$ are univariate moving average (MA) (1) processes so that the two univariate time series $z_t$ and $y_t$ are both ARMA(1,1) processes. The terms $\phi(1 \pm \eta)$ are the autoregressive coefficients that determine the long-term persistence of each series.

This analysis has demonstrated that the bivariate ARMA(1,1) stochastic process $X_t$ in (1) with (23) is a linear combination of two univariate ARMA(1,1) processes $z_t$ and $y_t$. The multivariate process can now be interpreted in terms of the statistical characteristics of its two modes of variation, corresponding to the two eigenvalues of $\Phi$. For any $X_t$ and $\eta \neq 0$,

\[
z_t = 0.5 E_1^T X_t = 0.5 (X_t^1 + X_t^2)
\]

\[
y_t = 0.5 E_2^T X_t = 0.5 (X_t^1 - X_t^2)
\]

Therefore, $z_t$ describes the variation over time of the sum of the flows at the two sites. If regional climatic patterns exhibit long-term persistence, then one would expect that persistence to be reflected in the summed flows and $\phi(1 + \eta)$ may be near.
TABLE 6. Description of Multivariate Nondiagonal ARMA(1,1) Models All of Which Have \( \rho(i,i) = 0.3 \)

<table>
<thead>
<tr>
<th>( \eta )</th>
<th>( \phi_\eta )</th>
<th>( \theta_\eta )</th>
<th>( \rho(i,1) )</th>
<th>( \rho(i,i) )</th>
<th>( \rho(i,2) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.2</td>
<td>0.667</td>
<td>0.360</td>
<td>0.154</td>
<td>0.179</td>
<td>0.063</td>
</tr>
<tr>
<td>0.0</td>
<td>0.800</td>
<td>0.572</td>
<td>0.210</td>
<td>0.240</td>
<td>0.168</td>
</tr>
<tr>
<td>0.2</td>
<td>0.667</td>
<td>0.465</td>
<td>0.246</td>
<td>0.231</td>
<td>0.204</td>
</tr>
<tr>
<td>0.4</td>
<td>0.571</td>
<td>0.392</td>
<td>0.267</td>
<td>0.232</td>
<td>0.221</td>
</tr>
<tr>
<td>0.6</td>
<td>0.500</td>
<td>0.339</td>
<td>0.282</td>
<td>0.235</td>
<td>0.231</td>
</tr>
<tr>
<td>0.8</td>
<td>0.444</td>
<td>0.298</td>
<td>0.293</td>
<td>0.239</td>
<td>0.238</td>
</tr>
<tr>
<td>1.0</td>
<td>0.400</td>
<td>0.266</td>
<td>0.300</td>
<td>0.242</td>
<td>0.242</td>
</tr>
</tbody>
</table>

Here \( \rho(i,j) \) is \( \rho_\eta(1,2) \) or \( \rho_\eta(2,1) \), where \( \rho_\eta(1,2) = \rho_\eta(2,1) \); \( \rho(i,i) \) is either \( \rho_\eta(1,1) \) or \( \rho_\eta(2,2) \), where \( \rho_\eta(1,1) = \rho_\eta(2,2) \); \( \rho(i,1) \) is \( \rho_\eta(1,2) \) or \( \rho_\eta(2,1) \), where \( \rho_\eta(1,2) = \rho_\eta(2,1) \).

unity. On the other hand, \( \phi \) describes the variation over time in the difference or imbalance in the flows at the two sites. If both sites display concurrent long-term persistence as a result of regional climatic variations, then the persistence in the flow difference sequence should be much less than that in the flow sums, since the differencing acts to filter out the low frequency variations. Thus based on hydrologic considerations, we expect that reasonable models will have the eigenvalue \( \lambda = \delta t \).

Once the \( np \) roots of this characteristic equation have been identified, the corresponding eigenvectors and univariate stochastic processes are easily identified. Box and Tiao [1977] consider examination of the eigenvalues and eigenvectors of the ARMA(1,1) model. For \( \eta > 0 \) this is equivalent to the multivariate ARMA(1,1) model's implicit aggregate flow model.

The disaggregation model in (9) and (14) employs

\[
W_i = \lambda_i X_i^1 - \beta(X_i^1 + X_i^2) = (1 - \beta)X_i^1 - \beta X_i^2
\]

If \( S_0 \) is symmetric as has been assumed throughout, with \( (S_0)_{11} = (S_0)_{22} \), then \( \beta = \frac{1}{2} \), and

\[
W_i = 0.5(X_i^1 - X_i^2)
\]

Clearly, the disaggregation model generates the sequence of differences, \( [X_i^1 - X_i^2] \), with an AR(1) model rather than the ARMA(1,1) model implicit in our non-diagonal multivariate ARMA(1,1) model. This approximation should be satisfactory if flow imbalances (or differences) exhibit relatively little long-term persistence. Thus the disaggregation model, in general, may provide a good description of flows generated by hydrologically reasonable nondiagonal but symmetric multivariate ARMA(1,1) models. However, the disaggregation model may have difficulty providing a good description of the distribution of flows generated by highly persistent diagonal ARMA(1,1) models with relatively little correlation between flows at the two sites. Such cases have not been analyzed.

It is also possible to decompose any multivariate ARMA\( (p,q) \) model in this way so as to identify independent modes of variation whose statistical behavior is determined by separate univariate ARMA\( (p,q) \) models. Consider the general multivariate model

\[
Z_t = \sum_{i=1}^{p} \Phi_i Z_{t-i} + V_t - \sum_{j=1}^{n} \Theta_j V_{t-j}
\]

One can determine the eigenvalues \( \lambda \) of the autoregressive operator by solving the \( np \)-order polynomial equation

\[
0 = \det \left( \lambda^n I - \sum_{i=1}^{n} \lambda^{n-i} \Phi_i \right)
\]

Once the \( np \) roots of this characteristic equation have been identified, the corresponding eigenvectors and univariate stochastic processes are easily identified. Box and Tiao [1977] consider examination of the eigenvalues and eigenvectors of the ARMA(1,1) model. For \( \eta > 0 \) this is equivalent to the multivariate ARMA(1,1) model’s implicit aggregate flow model.

The disaggregation model in (9) and (14) employs
TABLE 7 Estimates of the Mean and Root Mean Square Error of Statistics Describing the Persistence of Flows At Each Site When the $\Phi$ and $\Theta$ Matrices in Case 2 are Contaminated

<table>
<thead>
<tr>
<th>Method</th>
<th>$\eta$</th>
<th>$E[\rho_1]$</th>
<th>RMSE[\rho_1]</th>
<th>$E[\Gamma_4]$</th>
<th>RMSE[\Gamma_4]</th>
<th>$E[\Gamma_{10}]$</th>
<th>RMSE[\Gamma_{10}]</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLE</td>
<td>-0.2</td>
<td>0.31</td>
<td>0.13</td>
<td>1.69</td>
<td>0.35</td>
<td>2.20</td>
<td>0.83</td>
</tr>
<tr>
<td>MOM</td>
<td>-0.2</td>
<td>0.27</td>
<td>0.13</td>
<td>1.57</td>
<td>0.34</td>
<td>1.92</td>
<td>0.72</td>
</tr>
<tr>
<td>DISAG (0.7)</td>
<td>-0.2</td>
<td>0.31</td>
<td>0.11</td>
<td>1.66</td>
<td>0.28</td>
<td>2.08</td>
<td>0.62</td>
</tr>
<tr>
<td>DISAG (0.9)</td>
<td>-0.2</td>
<td>0.31</td>
<td>0.11</td>
<td>1.67</td>
<td>0.29</td>
<td>2.12</td>
<td>0.67</td>
</tr>
<tr>
<td>True values</td>
<td>-0.2</td>
<td>0.30</td>
<td></td>
<td>1.69</td>
<td></td>
<td>2.17</td>
<td></td>
</tr>
<tr>
<td>MLE</td>
<td>0.4</td>
<td>0.31</td>
<td>0.14</td>
<td>1.74</td>
<td>0.40</td>
<td>2.44</td>
<td>1.07</td>
</tr>
<tr>
<td>MOM</td>
<td>0.4</td>
<td>0.26</td>
<td>0.14</td>
<td>1.55</td>
<td>0.41</td>
<td>1.95</td>
<td>1.02</td>
</tr>
<tr>
<td>DISAG (0.7)</td>
<td>0.4</td>
<td>0.31</td>
<td>0.12</td>
<td>1.74</td>
<td>0.36</td>
<td>2.45</td>
<td>1.00</td>
</tr>
<tr>
<td>DISAG (0.9)</td>
<td>0.4</td>
<td>0.31</td>
<td>0.13</td>
<td>1.74</td>
<td>0.38</td>
<td>2.45</td>
<td>1.07</td>
</tr>
<tr>
<td>True values</td>
<td>0.4</td>
<td>0.30</td>
<td></td>
<td>1.77</td>
<td></td>
<td>2.61</td>
<td></td>
</tr>
</tbody>
</table>

Ninety-five percent confidence intervals for $E[\rho_1]$, RMSE[\rho_1], $E[\Gamma_4]$, RMSE[\Gamma_4], $E[\Gamma_{10}]$, and RMSE[\Gamma_{10}] are within $\pm 3$, $\pm 5$, $\pm 2$, $\pm 6$, $\pm 3$, and $\pm 9$% of the reported values.

*Values in parenthesis are the cross correlation of at-site flows, $p_d(1,2)$.

$S_0^{-1}(S_0 - G)$, while Tiao and Box [1981] suggest examination of other matrices.

Analysis of Model Performance

The Monte Carlo Analysis reported in Tables 1–5 and Figures 1 and 2 was repeated with a range of bivariate nondiagonal ARMA(1,1) models. In all cases, the $\Phi$ and $\Theta$ matrices were as specified in (23). For each $\eta$, the scale parameter $\phi$ was taken to be $(0.8/(1 + |\eta|))$ so that the larger eigenvalue of $\Phi$, $\phi(1 + \eta)$, will equal 0.8; this is the value of $\phi_4$ for the case 2 diagonal model. Given these values of $\phi$ and $\eta$, the scale parameter $\theta$ was selected so that the flows at each site had a lag-one correlation coefficient $\rho_4(1,1) = \rho_4(2,2) = 0.3$ corresponding to the case 2 multivariate diagonal ARMA(1,1) model. Table 6 gives the actual values of $\theta$ as well as $\rho_4(1,1)$, $\rho_4(1,2)$, and $\rho_4(2,2)$. Values of $\eta$ from $-0.2$ to $1.0$ are considered. The Monte Carlo results are summarized in Tables 7–9 analogous to Tables 2–5.

In terms of RMSE's reported in Table 7 for the fitted $\hat{\rho}_1$, $\hat{\Gamma}_4$, and $\hat{\Gamma}_{10}$, the use of a nondiagonal ARMA population model has almost no impact on the performance of the various methods as one would expect. The transition to a nondiagonal model basically affects the relationship or cross correlation of flows at the two sites rather than the correlation structure at each site.

Likewise, $\eta \neq 0$ has relatively little effect on the RMSE of various $\sigma_2^2$ and $\sigma_4^2$ estimators. However, $\eta \neq 0$ has a very large impact on the precision with which $\rho_d(1,2)$ is estimated as shown in Figure 3. Here, MLE estimators of $G$ with a structurally incorrect model and $\eta > 0$ do much worse than reproduction of the observed covariance matrix $S_0$.

In terms of the RMSE of $\hat{\rho}_4(a)$, $\hat{\Gamma}_4$, and $\hat{\Gamma}_{10}$ reported in Table 8, $\eta \neq 0$ had relatively little impact upon the ranking of the procedures which all performed about as well. (Our implementation of O'Connell's [1974] procedure was omitted because of its very poor performance earlier.) In general, MLE $\phi$ and $\theta$ estimators in a diagonal model were best with the competing procedures performing almost as well most of the time: the one exception was the RMSE of $\hat{\Gamma}_4$ when $\rho_d(1,2) = 0.7$. It is interesting to note that while the disaggregation model often yielded the larger RMSE's in Table 8, it generally had the smallest bias; this is particularly true with respect to $\hat{\rho}_4(a)$ when $\rho_d(1,2) = 0.7$. With respect to $\hat{\rho}_4(a)$ overall, there was relatively little difference in the RMSE's, while the disaggregation model yielded a substantially smaller bias for large $\eta$.

TABLE 8 Estimates of the Mean and Root Mean Square Error of Statistics Describing the Persistence of Aggregate Flows When the $\Phi$ and $\Theta$ Matrices in Case 2 are Contaminated

<table>
<thead>
<tr>
<th>Method</th>
<th>$\eta$</th>
<th>$E[\rho_4(a)]$</th>
<th>RMSE[\rho_4(a)]</th>
<th>$E[\Gamma_4]$</th>
<th>RMSE[\Gamma_4]</th>
<th>$E[\Gamma_{10}]$</th>
<th>RMSE[\Gamma_{10}]</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLE</td>
<td>-0.2</td>
<td>0.31</td>
<td>0.12</td>
<td>1.69</td>
<td>0.30</td>
<td>2.19</td>
<td>0.70</td>
</tr>
<tr>
<td>MOM</td>
<td>-0.2</td>
<td>0.27</td>
<td>0.11</td>
<td>1.57</td>
<td>0.27</td>
<td>1.92</td>
<td>0.55</td>
</tr>
<tr>
<td>DISAG (0.7)</td>
<td>-0.2</td>
<td>0.29</td>
<td>0.13</td>
<td>1.62</td>
<td>0.32</td>
<td>2.04</td>
<td>0.72</td>
</tr>
<tr>
<td>True values</td>
<td>-0.2</td>
<td>0.27</td>
<td></td>
<td>1.58</td>
<td></td>
<td>1.90</td>
<td></td>
</tr>
<tr>
<td>MLE</td>
<td>0.4</td>
<td>0.31</td>
<td>0.12</td>
<td>1.73</td>
<td>0.37</td>
<td>2.42</td>
<td>1.02</td>
</tr>
<tr>
<td>MOM</td>
<td>0.4</td>
<td>0.25</td>
<td>0.14</td>
<td>1.54</td>
<td>0.44</td>
<td>1.92</td>
<td>1.11</td>
</tr>
<tr>
<td>DISAG (0.7)</td>
<td>0.4</td>
<td>0.35</td>
<td>0.14</td>
<td>1.84</td>
<td>0.41</td>
<td>2.67</td>
<td>1.16</td>
</tr>
<tr>
<td>True values</td>
<td>0.4</td>
<td>0.33</td>
<td></td>
<td>1.87</td>
<td></td>
<td>2.84</td>
<td></td>
</tr>
<tr>
<td>MLE</td>
<td>1.0</td>
<td>0.31</td>
<td>0.13</td>
<td>1.73</td>
<td>0.41</td>
<td>2.42</td>
<td>1.10</td>
</tr>
<tr>
<td>MOM</td>
<td>1.0</td>
<td>0.25</td>
<td>0.16</td>
<td>1.54</td>
<td>0.49</td>
<td>1.92</td>
<td>1.23</td>
</tr>
<tr>
<td>DISAG (0.7)</td>
<td>1.0</td>
<td>0.37</td>
<td>0.14</td>
<td>1.88</td>
<td>0.42</td>
<td>2.76</td>
<td>1.19</td>
</tr>
<tr>
<td>True values</td>
<td>1.0</td>
<td>0.36</td>
<td></td>
<td>1.93</td>
<td></td>
<td>2.97</td>
<td></td>
</tr>
</tbody>
</table>

Here $\rho_d(1,2) = 0.7$. Ninety-five percent confidence intervals for $E[\rho_4(a)]$, RMSE[\rho_4(a)], $E[\Gamma_4]$, RMSE[\Gamma_4], $E[\Gamma_{10}]$, and RMSE[\Gamma_{10}] are within $\pm 3$, $\pm 5$, $\pm 2$, $\pm 6$, $\pm 3$, and $\pm 9$% of the reported values.
TABLE 9. Estimates of the Mean of Drought Correlation, Concurrency, Coherence, and Required Storage When the $\Phi$ and $\Theta$ Matrices in Case 2 are Contaminated (demand = 0.90 MAF and $\rho_d(1,2) = 0.7$)

<table>
<thead>
<tr>
<th>Method</th>
<th>$\phi$</th>
<th>$\theta$</th>
<th>$E[\tilde{\rho}_d]$</th>
<th>$E[\tilde{\gamma}_1]$</th>
<th>$E[\tilde{\gamma}_2]$</th>
<th>$E[\tilde{\gamma}_3]$</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLE</td>
<td>-0.2</td>
<td>0.53</td>
<td>0.36</td>
<td>0.77</td>
<td>1.3</td>
<td></td>
</tr>
<tr>
<td>MOM</td>
<td>-0.2</td>
<td>0.54</td>
<td>0.36</td>
<td>0.77</td>
<td>1.3</td>
<td></td>
</tr>
<tr>
<td>DISAG</td>
<td>-0.2</td>
<td>0.53</td>
<td>0.38</td>
<td>0.75</td>
<td>1.3</td>
<td></td>
</tr>
<tr>
<td>True values</td>
<td>-0.2</td>
<td>0.44</td>
<td>0.23</td>
<td>0.73</td>
<td>1.3</td>
<td></td>
</tr>
<tr>
<td>MLE</td>
<td>0.4</td>
<td>0.54</td>
<td>0.34</td>
<td>0.76</td>
<td>1.5</td>
<td></td>
</tr>
<tr>
<td>MOM</td>
<td>0.4</td>
<td>0.54</td>
<td>0.35</td>
<td>0.76</td>
<td>1.3</td>
<td></td>
</tr>
<tr>
<td>DISAG</td>
<td>0.4</td>
<td>0.63</td>
<td>0.44</td>
<td>0.79</td>
<td>1.6</td>
<td></td>
</tr>
<tr>
<td>True values</td>
<td>0.4</td>
<td>0.67</td>
<td>0.50</td>
<td>0.83</td>
<td>1.4</td>
<td></td>
</tr>
<tr>
<td>MLE</td>
<td>1.0</td>
<td>0.53</td>
<td>0.33</td>
<td>0.75</td>
<td>1.5</td>
<td></td>
</tr>
<tr>
<td>MOM</td>
<td>1.0</td>
<td>0.54</td>
<td>0.37</td>
<td>0.76</td>
<td>1.3</td>
<td></td>
</tr>
<tr>
<td>DISAG</td>
<td>1.0</td>
<td>0.64</td>
<td>0.48</td>
<td>0.79</td>
<td>1.6</td>
<td></td>
</tr>
<tr>
<td>True values</td>
<td>1.0</td>
<td>0.71</td>
<td>0.54</td>
<td>0.84</td>
<td>1.4</td>
<td></td>
</tr>
</tbody>
</table>

Ninety-five percent confidence intervals for $E[\tilde{\rho}_d]$, $E[\tilde{\gamma}_1]$, $E[\tilde{\gamma}_2]$, and $E[\tilde{\gamma}_3]$ are within ±4, ±14, ±2, and ±11 of the reported values.

### Conclusion

We have examined a number of ways of fitting ARMA(1,1) models to multivariate hydrologic time series by building upon simple univariate procedures. A range of statistics were employed to evaluate how well the models reproduced the persistence and variance of flows at each site and of the aggregate flow in any year as well as the correlation between concurrent flows at two sites. Moreover, in the pseudo-historical flow generation step and in the parameter estimation procedures, checks were employed to ensure that the pseudo-historical sequences resembled ones to which such procedures might be applied and that fitted models resembled hydrologically reasonable time series models.

In general, we found that the following observations were true.

1. Maximum likelihood estimators of the covariance matrix $G$ of the innovations employed in the multivariate ARMA(1,1) model were generally inferior to estimators of $G$ which reproduced the observed covariance of the flows. The MLE estimators of $G$ did a particularly poor job of reproducing the cross correlation of concurrent flows; this is consistent with results reported in the works by Stedinger [1981]. The MLE $G$ estimator’s performance deteriorated even further when flows were generated by other than the assumed ARMA(1,1) diagonal model. Our recommendation to estimate the innovations’ covariance matrix using essentially a method of moments procedure stands in contrast to suggestions by Salas et al. [1980], Loucks et al. [1981], and Tiao and Box [1981], who had assumed that MLE estimators of the innovations variances would perform well.

2. The Monte Carlo experiments allowed evaluation of the conventional univariate method-of-moments and maximum likelihood $\phi$ and $\theta$ estimators with restrictions that we added (namely that $\max(0, \tilde{\theta}_1) < \phi_1 < 1$). In general, the method of moments estimators did slightly better in terms of the RMSE of $\tilde{\rho}_1$, $\tilde{\gamma}_1$, and $\tilde{\gamma}_2$ in the low persistence case while the MLE estimators did appreciably better in the high persistence case. Thus on balance, one might view the MLE estimators as the better choice.

3. A new disaggregation modeling approach introduced by Stedinger and Vogel [1984] was employed for dividing between the two sites the aggregate flows generated with a univariate ARMA(1,1) model. In our examples the disaggregation model generally did as well or better than the bivariate diagonal ARMA(1,1) models at estimating the persistence of the flows at each site. However, the bivariate diagonal ARMA(1,1) models did slightly better than the disaggregation model at reproducing the persistence of the aggregate flow series, particularly for the smaller values of $\rho_d(1,2)$. These results are probably due in part to the particular numerical examples examined. Overall our results show that the performance of the bivariate diagonal ARMA(1,1) models with efficient parameter estimation procedures and the aggregate ARMA(1,1) model in combination with a disaggregation model reproduced the distribution of the original flow models with comparable precision. Thus the appropriate choice between the model structures in practice is likely to depend on other criteria, such as ease of parameter estimation and subsequent flow generation. In addition, if one’s goal is to obtain seasonal flows at several sites, then an aggregate flow model coupled with disaggregation would be very attractive; the aggregate annual flows could first be disaggregated to aggregate monthly flows, then the aggregate monthly flows could be apportioned among the sites [Stedinger and Vogel, 1984; Loucks et al., 1981, pp. 303–305].

4. In our initial simulation, the bivariate streamflow models were fit to time series generated using a bivariate diagonal ARMA(1,1) model. In a later section, the structure of symmetric bivariate nondiagonal ARMA(1,1) models was examined. Hydrologically reasonable models were thought to correspond to $\Phi$ matrices with positive eigenvalues and positive off-diagonal elements (corresponding to more persistent flow sums). Surprisingly, hydrologically reasonable symmetric bivariate ARMA(1,1) models generate flows whose structure resembles more that of the disaggregation model than the diagonal bivariate ARMA(1,1) model. However, in Monte Carlo experiments where flows were generated with nondiagonal ARMA models, the relative performance of the procedures changed very little, if at all.

The restricted diagonal ARMA(1,1) model performs well even if flows are not generated from such a restricted model; thus with MLE estimators of $\phi$ and $\theta$ at each site, the diagonal ARMA model should be considered as an attractive and simple operational alternative to the complete and general multivariate ARMA model. Moreover, Lettenmaier [1980] found that the full multivariate ARMA(1,1) models with maximum likelihood estimates of the parameter matrices $\Phi$ and $\Theta$ generally did a poor job of reproducing the persistence of the flows observed at each site; that difficulty was not encountered with the diagonal ARMA(1,1) model.

5. We attempted to implement O’Connell’s suggestion [O’Connell, 1974, 1977] that $\phi$ and $\theta$ be selected to reproduce in expectation sample estimators of $h$ and $\rho_1$. As Appendix B shows, this will only be feasible if estimators $K$ and $\hat{\rho}_1$ fall in a very narrow subset of $(h, \rho_1)$ space. To avoid the problem of being unable to reproduce in expectation both $K$ and $\hat{\rho}_1$, we choose to relate $\phi$ and $\theta$ to $K$ alone. While this simplification was necessary computationally, we do not believe that it made much difference in terms of the characteristics of the flows generated. In general, our implementation of O’Connell’s suggestion performed appreciably worse than the alternative estimation procedures and we recommend that it not be used.

6. Special statistics were developed to examine the joint operation of two reservoirs with the generated bivariate flows sequences. The expected values of the drought correlation, coincidence, and coherency statistics did not vary much among the models, except perhaps for the disaggregation procedure which yielded larger values of the correlation and coincidence statistics.

Stedinger et al.: Multisite Annual Flow Models
correlation coefficient \( \rho_1 \) and the Hurst coefficient \( h \) for a range of univariate ARMA(1,1) models. The results of his experiments were used here to derive expressions which relate the small sample estimator \( K \) of the Hurst coefficient \( h \) to values of \( \phi \) and \( \theta \). These expressions were used in our experiments to provide estimates for \( \phi_1 \) and \( \theta_1 \) in diagonal multivariate models as a function of \( K \).

Analysis of O’Connell’s [1974] results reveal a near-linear relationship between \( E[K|\phi, \theta] \) and \( E[\hat{\rho}_1|\phi, \theta] \) for samples of length 50 as shown in Figure B1. Figure B1 also shows that for given \( E[K] \), the range of values of \( E[\hat{\rho}_1] \) which can be obtained with ARMA(1,1) models is very limited. Seldom will it be possible to select \( \phi \) and \( \theta \) so that the expected value of \( K \) and \( \hat{\rho}_1 \) will equal the value of their sample estimators, as was originally suggested by O’Connell [1974; see also 1977]. In an attempt to implement O’Connell’s method, we were forced to relate \( \phi \) and \( \theta \) to \( K \) alone.

Further graphical analysis yielded approximate linear relationships between \( E[K] \) and \( \theta \) for each of the six values of \( \phi \) used in O’Connell’s [1974] experiments. This led to expressions of the form

\[
E[K|\phi, \theta] = A_{11}(\phi) + \theta A_{12}(\phi) \quad (B1)
\]

\[
E[\hat{\rho}_1|\phi, \theta] = A_{21}(\phi) + \theta A_{22}(\phi) \quad (B2)
\]

where the parameters \( A_{ij}(\phi), \theta \) for \( i, j = 1 \) or 2, were estimated for each of the six values of \( \phi \) (0.75, 0.80, 0.84, 0.88, 0.92, 0.96). It was found that these functions were approximately linear so our model employed

\[
A_{ij}(\phi) = a_{ij} + b_{ij}\phi \quad (B3)
\]

Substitution of (B3) into (B1) and (B2) yields a system of linear equations for \( E[K|\phi, \theta] \) and \( E[\hat{\rho}_1|\phi, \theta] \) as a function of \( \phi \) and \( \theta \), and the product \( \phi \theta \). A Newton-Raphson algorithm was used to obtain values of \( \phi \) and \( \theta \) which satisfy (B4) and (B5) for fixed values of \( E[\hat{\rho}_1] \) and \( E[K] \). For each fixed value of \( E[\hat{\rho}_1] \) a range of feasible combinations of \( \phi \) and \( \theta \) resulted, corresponding to each value of \( E[K] \). This is to be expected given the scatter depicted in Figure B1. The median value in this range was selected and a linear relationship with \( K \) observed. (See Figure B2 in microfiche appendix.) The final estimation equations for our implementation of O’Connell’s method (for \( n = 50 \)) are

\[
\phi = 0.689 + 0.285K \quad (B4)
\]

\[
\theta = 2.15 - 1.95K \quad (B5)
\]

Acknowledgments. This work was supported by grants CME 8010889 and CEE 8211730 from the U.S. National Science Foundation.

REFERENCES

Armbruster, J. T., Multisite generation of monthly flows using an autoregressive integrated moving average model and disaggregation, paper presented at American Geophysical Union Spring Meeting, AGU Miami Beach, Fla., April 1978.


correlation coefficient $\rho_1$ and the Hurst coefficient $h$ for a range of univariate ARMA(1,1) models. The results of his experiments were used here to derive expressions which relate the small sample estimator $K$ of the Hurst coefficient $h$ to values of $\phi$ and $\theta$. These expressions were used in our experiments to provide estimates for $\phi_1$ and $\theta_1$ in diagonal multivariate models as a function of $K$.

Analysis of O’Connell’s [1974] results reveal a near-linear relationship between $E[K]$ and $E[\rho_1]$ for samples of length 50 as shown in Figure B1. Figure B1 also shows that for given $E[K]$, the range of values of $E[\rho_1]$ which can be obtained with ARMA(1,1) models is very limited. Seldom will it be possible to select $\phi$ and $\theta$ so that the expected value of $K$ and $\rho_1$ will equal the value of their sample estimators, as was originally suggested by O’Connell [1974; see also 1977]. In an attempt to implement O’Connell’s method, we were forced to relate $\phi$ and $\theta$ to $K$ alone.

Further graphical analysis yielded approximate linear relationships between $E[K]$ and $\theta$ for each of the six values of $\phi$ used in O’Connell’s [1974] experiments. This led to expressions of the form

$$E[K | \phi, \theta] = A_{11}(\phi) + \theta A_{21}(\phi)$$  
$$E[\rho_1 | \phi, \theta] = A_{12}(\phi) + \theta A_{22}(\phi)$$

where the parameters $A_{ij}(\phi)$, for $i, j = 1$ or 2, were estimated for each of the six values of $\phi$ (0.75, 0.80, 0.84, 0.88, 0.92, 0.96). It was found that these functions were approximately linear so our model employed

$$A_{ij}(\phi) = a_{ij} + b_{ij}\phi$$

Substitution of (B3) into (B1) and (B2) yields a system of linear equations for $E[K | \phi, \theta]$ and $E[\rho_1 | \phi, \theta]$ as a function of $\phi$, $\theta$, and the product $\phi \theta$. A Newton-Raphson algorithm was used to obtain values of $\phi$ and $\theta$ which satisfy (B4) and (B5) for fixed values of $E[\rho_1]$ and $E[K]$. For each fixed value of $E[\rho_1]$ a range of feasible combinations of $\phi$ and $\theta$ resulted, corresponding to each value of $E[K]$. This is to be expected given the scatter depicted in Figure B1. The median value in this range was selected and a linear relationship with $K$ observed. (See Figure B2 in microfiche appendix.) The final estimation equations for our implementation of O’Connell’s method (for $n = 50$) are

$$\phi = 0.689 + 0.285K$$  
$$\theta = 2.15 - 1.95K$$

Acknowledgments. This work was supported by grants CME 8010889 and CEE 8211730 from the U.S. National Science Foundation.

REFERENCES

Armbruster, J. T., Multisite generation of monthly flows using an autoregressive integrated moving average model and disaggregation, paper presented at American Geophysical Union Spring Meeting, AGU Miami Beach, Fla., April 1978.


J. R. Stedinger and R. M. Vogel, Department of Environmental Engineering, Cornell University, Ithaca, NY 14853.

D. P. Lettenmaier, Department of Civil Engineering, University of Washington, Seattle, WA 98195.