Bootstrap Position Analysis for Forecasting Low Flow Frequency

By Gary D. Tasker,[1] Member, ASCE, and Paul Dunne[2]

Abstract: A method of random resampling of residuals from stochastic models is used to generate a large number of 12-month-long traces of natural monthly runoff to be used in a position analysis model for a water-supply storage and delivery system. Position analysis uses the traces to forecast the likelihood of specified outcomes such as reservoir levels falling below a specified level or streamflows falling below statutory passing flows conditioned on the current reservoir levels and streamflows. The advantages of this resampling scheme, called bootstrap position analysis, are that it does not rely on the unverifiable assumption of normality, fewer parameters need to be estimated directly from the data, and accounting for parameter uncertainty is easily done. For a given set of operating rules and water-use requirements for a system, water managers can use such a model as a decision-making tool to evaluate different operating rules.

INTRODUCTION

Forecasting drought risks for a surface-water–supply system is a difficult and necessary task for many water-supply managers. Water management decisions concerning proper storage levels in reservoirs and water-use limitations made at the beginning of a possible drought can be critical if below-normal streamflows persist over a period of a few months. Unlike extreme high flows, which may develop within a few days, extreme low flows require long periods of deficient rainfall to develop. Water-supply managers in the midst of a period of deficient rainfall may ask "what is the likelihood of streamflow falling below mandated flows within the next few months" or "how would water-use restrictions or changes in reservoir operations affect the likelihood of flows dropping below a specified level within the next few months"? Unconditional statistics, such as the 30-day, 10-year low flow, or a forecast of the most likely single hydrograph are of little use for such questions. What is needed is a forecast of the probability or risks of certain specific outcomes.

Position analysis (Hirsch 1978) is a tool that water managers can use to forecast risks associated with a specific operating plan for a basin over a period of a few months conditioned on current reservoir storages and streamflows. It can aid the water manager in deciding which plan of operation to implement by providing a means to evaluate and rank each proposed plan of operation in terms of future drought risks. Position analysis relies on the generation of a large number of possible monthly flow traces (a few months in length) that have been initialized with the current reservoir storages and current streamflows. The large number of flow traces are used to incorporate into the analysis the broad range of meteorological conditions that may occur but cannot be accurately forecast. These traces may be derived from conceptual hydrologic models driven by historical meteorological data as in the National Weather Service's extended streamflow prediction procedures (Day 1985), by historical flow records at gauging stations (Hirsch 1978), or from a stochastic model of streamflows based on the historical flow record as described in Hirsch (1981).

In this paper a stochastic model of streamflows is used to generate flow traces at multiple sites. The stochastic model generates synthetic streamflows by separating the flows into carryover components to model the serial correlation and random components called innovations. Hirsch (1981) generated traces for a single site using a random number generator to produce normally distributed innovations. However, the assumption of normality of innovations may not exploit all the information in the sample. In addition, when generating sequences at multiple sites, model specification and parameter estimation become extremely difficult tasks (Rasmussen et al. 1996). In this paper a new approach is presented in which the innovations of the multisite stochastic model of streamflows are generated by drawing at random, with replacement, from the residuals of the fitted stochastic model. Efron (1979) calls the method of randomly resampling with replacement a bootstrap sample. The advantage of bootstrapping is that it does not rely on the unverifiable assumption of normality and parameter estimation is simplified. As Efron (1982) points out, given a sample from a population, the nonparametric maximum likelihood estimate of the population is the sample itself. In the present case, the sample to be bootstrapped is the sample innovations or residuals derived from a stochastic model of observed runoff.

Random resampling methods have been used in hydrology for comparing statistical models, estimating parameter uncertainty, and comparing network design techniques. Examples of these uses using resampling methods include Tasker (1987), Woo (1989), Zucchini and Adamson (1989), and Moss and Tasker (1991). Pereira et al. (1984) and Oliveira et al. (1988) randomly resampled residuals for a multisite disaggregation model of lag-one annual flows. Cover and Unny (1986) use the bootstrap to estimate parameter uncertainty in autoregressive moving average (ARMA) models of streamflow.

Lall and Sharma (1996), Vogel and Shallcross (1996), and Sharma et al. (1997) apply the bootstrap to the flows instead of the residuals from a model to generate a long time series of flows. They thereby avoid assumptions about the form of dependence of a stochastic model. Their approach applied to a position analysis problem would be an attractive alternative to test in future studies; however, it is beyond the scope of this investigation.

The bootstrap runoff traces for a position analysis model may be used as inputs to drive a model simulation of a water-supply storage and delivery system for drought management. For a given set of operating rules and water-use requirements for a system, water managers can use such a model to forecast the likelihood of specified outcomes such as reservoir levels falling below a specified level or streamflows falling below statutory passing flows a few months ahead conditioned on the current reservoir levels and streamflows. Thus the model can be used to determine the effectiveness of proposed changes in

[1]Hydro., U.S. Geological Survey, 430 Nat. Ctr., Reston, VA 22092.
[2]Hydro., U.S. Geological Survey, 810 Bear Tavern Rd., Ste. 206, West Trenton, NJ 08628.

Note. Discussion open until May 1, 1998. To extend the closing date one month, a written request must be filed with the ASCE Manager of Journals. The manuscript for this paper was submitted for review and possible publication on March 31, 1997. This paper is part of the ***Journal of Water Resources Planning and Management***, Vol. 123, No. 6, November/December, 1997. ©ASCE, ISSN 0733-9496/97/0006-0359–0367/$4.00 + $.50 per page. Paper No. 15487.

J. Water Resour. Plann. Manage., 1997, 123(6): 359-367

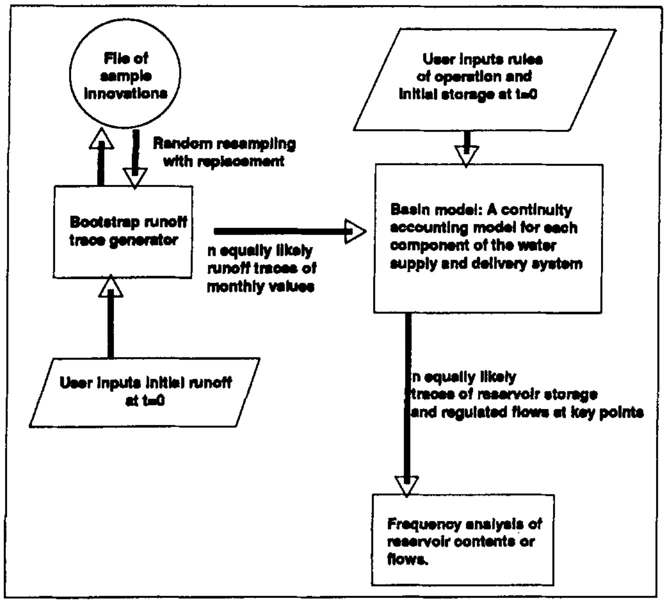

FIG. 1. Flowchart of Bootstrap Position Analysis Applied to Water-Supply Storage and Delivery System

operating rules or proposed water restrictions in reducing drought risks. The flowchart in Fig. 1 shows the connection between the bootstrap runoff traces and a basin model.

The following sections will describe the generation of multisite streamflow traces based on a stochastic streamflow model using bootstrap resampling of residuals. The bootstrap position analysis method has been applied to a surface-water–supply system in New Jersey.

BOOTSTRAP RUNOFF TRACE GENERATOR

Stochastic Flow Model

The position analysis forecasting model is driven by multiple runoff traces generated by a stochastic model of monthly runoff based on the historic streamflow records for gauged sites in the area. The stochastic runoff model used for this model is the log-transform ARMA [LT-ARMA(1,1)] cyclic model described in Hirsch (1981). Also, it is referred to as a periodic ARMA(1,1) or PARMA(1,1) model (Salas 1993). Denote the runoff in year i (i = 0, 1, 2, . . . , n) and month j (j = 1, 2, . . . , 12) as X_{12i+j}, and let $Y_{12i+j} = \log(X_{12i+j})$. Define the variable

$$Z_{12i+j} = \frac{(Y_{12i+j} - \bar{Y}_j)}{S_j} \tag{1}$$

where $\bar{Y}_j$ and S_j = sample mean and standard deviation, respectively, of the logarithms of the observed runoff values for month j. Therefore, Z_{12i+j} represents the standardized deviation for year i and month j from the mean of the LT runoff for month j.

The serial dependence is modeled as a periodic moving average process or PARMA(1,1) model of the form

$$Z_{12i+j} = \phi Z_{12i+j-1} + E_{12i+j} - \theta_j E_{12i+j-1} \tag{2}$$

where E's = independent random variables with mean of zero. The 13 parameters ϕ and θ_j are estimated by the method described in Hirsch (1979). Hirsch (1981) notes that in this model the lag-one serial correlations between the runoff values for all 12 months are exactly preserved in the long run and the lag-2 to -12 serial correlations are preserved in a least-squares sense. Other model forms and methods of parameter estimation such as the method of moments (Salas et al. 1982) or method of maximum likelihood (Vecchia 1985a,b) could be used as well. In position analysis models, focus is on generating a large number of sequences of synthetic flows a few months in length. Therefore, disaggregation approaches (Valencia and Schaake 1973) and aggregation approaches (Vecchia et al. 1983) that seek to preserve annual statistics as well as monthly statistics are not necessary.

The first and third terms on the right side of (2) can be thought of as the carryover components of the deviation from mean monthly LT runoff due to antecedent moisture and delayed runoff in a basin, whereas E_{12i+j} is the innovation or random component of the deviation from mean monthly LT runoff due to meteorological conditions in month j of year i. Breaking up the standardized runoff values into carrryover components and random components allows one to generate synthetic runoff by random resampling from the fitted model residuals. It also allows the inclusion of the effects of 90-day weather forecasts by modification of the likelihood of selecting a particular random component.

Eq. (2) may be written as

$$E_{12i+j} = Z_{12i+j} - \phi Z_{12i+j-1} + \theta_j E_{12i+j-1} \tag{3}$$

Given an observed long record of monthly runoff values Z_{12i+j}, the parameters for the model ϕ and θj, and a starting value of $E_0 = 0$ for the first month in the first year of record, a long record of residuals can be computed recursively from (3). These residuals will be approximately independent but not identically distributed in time. The seasonal differences of the residual distributions are accommodated by bootstrapping using residuals for specific months.

Single-Site Bootstrap Sequences

Starting values for Z and E are required in position analysis to generate a trace. For each trace the starting value of Z is set equal to the value of the LT standardized runoff for the present month. The starting value for E is computed recursively from (3) using observed values of Z for at least 12 months prior to the present month to overcome the effects of arbitrarily initiating the sequence with $E = 0$. With the use of the present observed value for Z and the computed value for E to start, a bootstrap trace is generated by sequentially computing Z's using a series of 12 residuals randomly drawn from the observed residuals for each month. A slight rescaling correction is needed to correct for sampling bias in the residuals. Thombs and Schucany (1990) show this correction for an ARMA(1,1) model to be $[(N_o - 4)/(N_o - 8)]^{0.5}$ where N_o is the number of observations. In the PARMA(1,1) model we take N_o to be the total number of observations divided by the number of months in a year. When generating sequences for a single site, each month of a bootstrap sequence of residuals is selected by randomly selecting a year with replacement and choosing the residual for the month for that year. This means that the selected residuals for each month in a sequence may be from a different year. By choosing residuals for specific months, estimation of parameters for a particular distribution for the innovations can be avoided.

The adequacy of a time series model is often examined by comparing statistics derived from historical observations with those derived from a time series generated by the model (Salas 1993). In Fig. 2, sample means, standard deviations, coefficients of skewness, and lag-one serial correlation coefficient based on 600 years of LT-formed simulated monthly runoff using normally distributed innovations, as in Hirsch (1981), and bootstrapped residuals are compared with the statistics for 75 years of observed data for a gauging station (Beaver Kill at Cooks Falls, N.Y.) in the upper Delaware River basin. These plots show that both the normal method and the bootstrap method reproduce the mean, standard deviation, and serial correlation fairly well, but the bootstrap reproduces the skewness

J. Water Resour. Plann. Manage., 1997, 123(6): 359-367

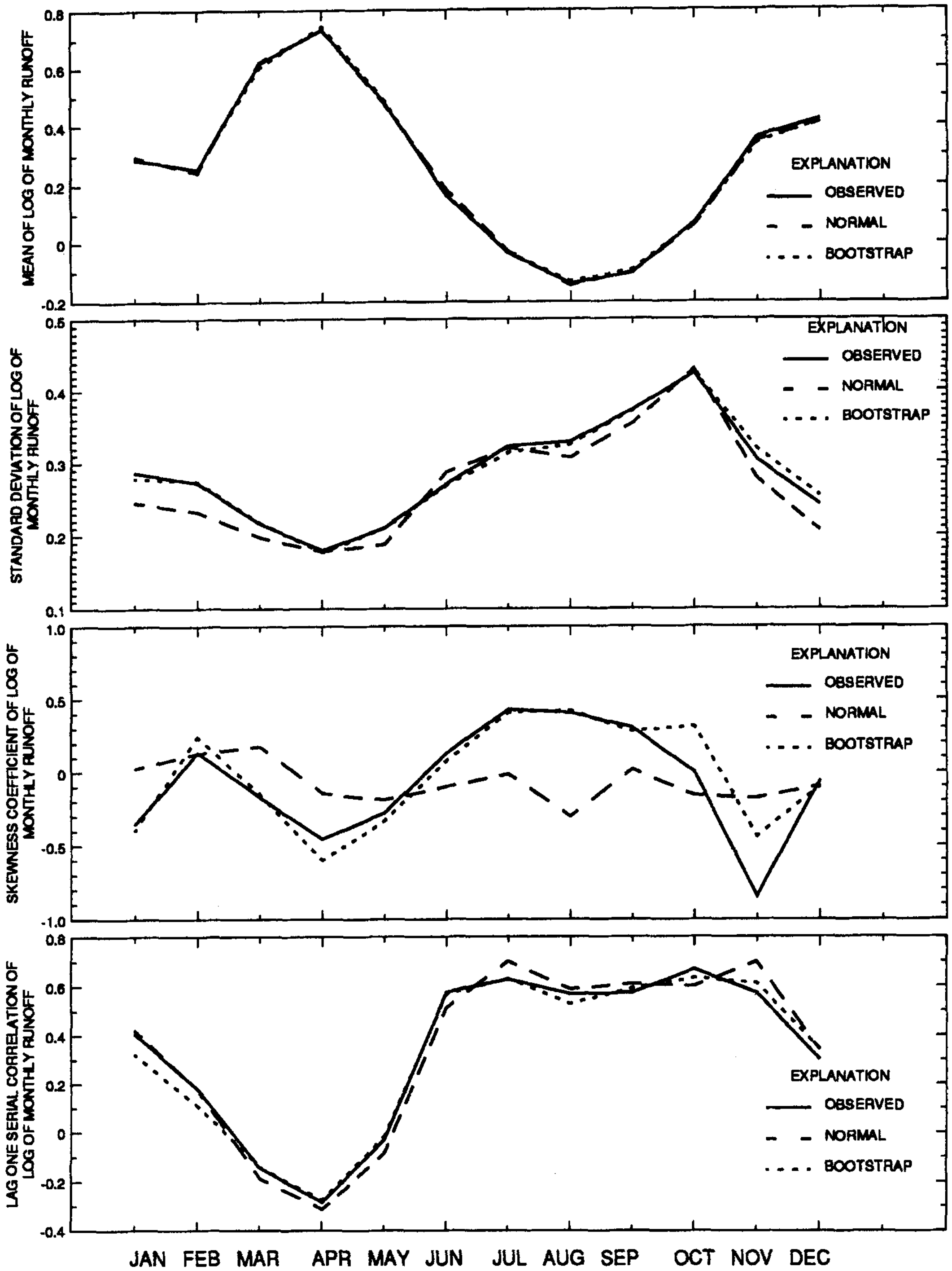

FIG. 2. Statistics for Logarithms of Monthly Runoff Based on 75 Years of Observed Record, on 600 Years of Synthetic Record with Normally Distributed Residuals, and on 600 Years of Synthetic Record with Bootstrapped Residuals

better than the normal method. The absolute value of the skewness in the observed record exceeds 0.35 five out of the twelve months. If the observations were normally distributed, coefficients as high as 0.35 would be expected no more than 2 out of 12 months. This evidence casts doubt on the normality assumption and suggests that the bootstrap method is more appropriate.

Multisite Bootstrap Sequences

Salas et al. (1980) and Loucks et al. (1981) recommend a simplified multisite ARMA model in which the temporal and spatial elements of the model are "decoupled" to reduce the number of parameters needed. Haltiner and Salas (1988) and Rasmussen et al. (1996) extend this approach to PARMA models. In their approach the model parameters are estimated independently for each site and the cross-correlation structure is preserved by generating spatially correlated innovations that are independent in time. Their approach requires estimation of the covariance matrix of innovations and the assumption of normality.

The bootstrap approach uses the same decoupled model; however, it avoids direct estimation of the covariance matrix and normality assumption by resampling contemporary observed residuals across sites. Pereira et al. (1984) use this scheme of contemporaneously bootstrapped residuals to preserve lag-zero cross-correlations of the residuals in a disaggregation model of autoregressive annual flows. This "unblocked" contemporaneous bootstrap approach works well for

J. Water Resour. Plann. Manage., 1997, 123(6): 359-367

sites with similar patterns of monthly runoff. However, in regions where monthly runoff patterns may be greatly different, a simple contemporaneous approach may not work well because of the complicated cross-correlation structure of the innovations (Salas et al. 1985). A blocked bootstrap technique (Kunsch 1989) is introduced to deal with complex cross-correlation structure of the innovations in lieu of introducing more parameters to deal with the cross-correlation structure. In the blocked bootstrapped multisite PARMA model, contemporaneous 12-month, nonoverlapping blocks of residuals are sampled with replacement to preserve the spatial covariance in the runoff. For each site the residuals for the 12 months of the same randomly selected water year are chosen.

Fig. 3 shows the effects of blocking the residuals on the lag-zero cross-correlations of the LT monthly runoff between several pairs of gauging stations in the Delaware and Raritan

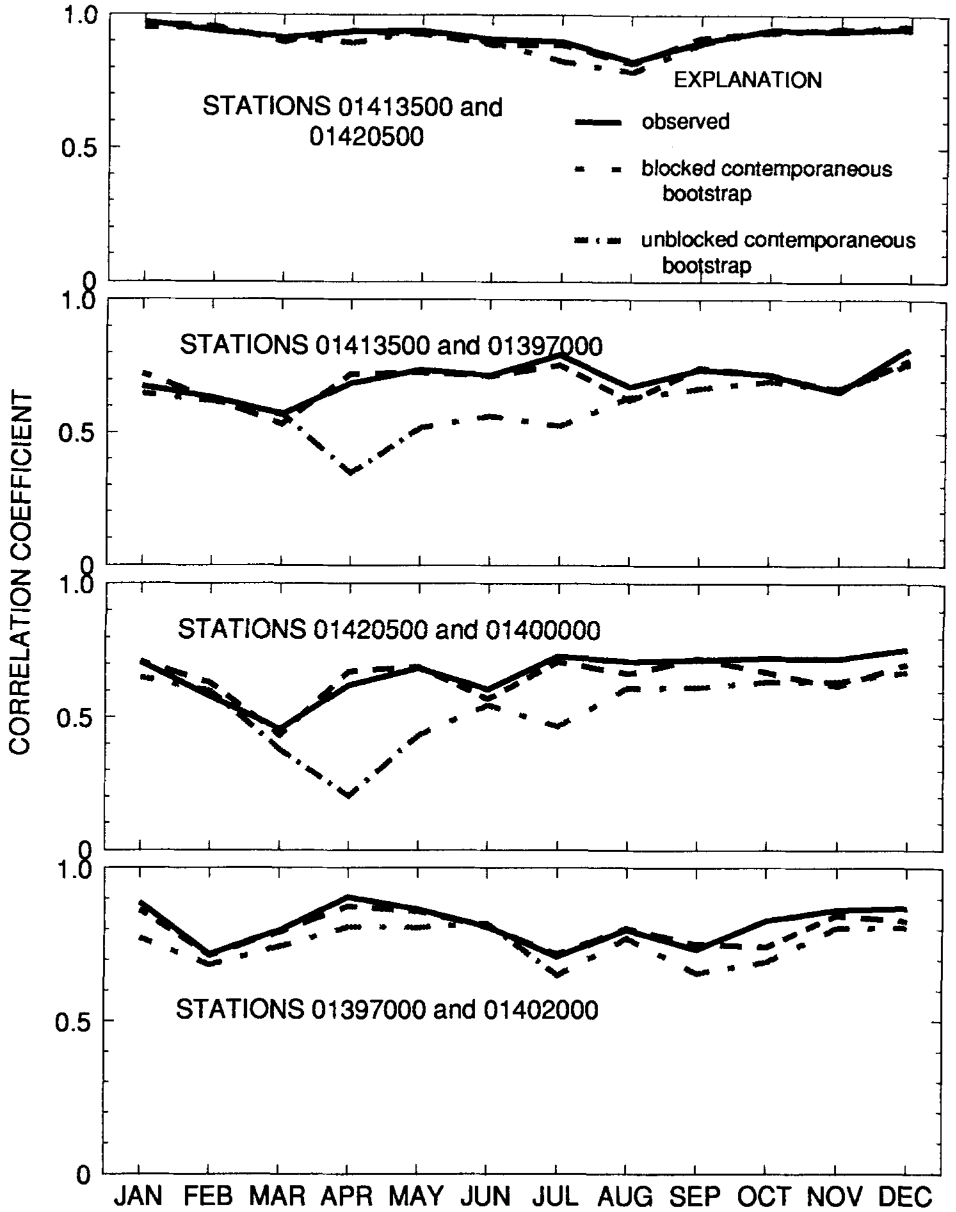

FIG. 3. Lag-Zero Cross-correlations of Logarithms of Monthly Runoff for Indicated Station Pairs

TABLE 1. Gauging Stations and Periods of Record Used to Reconstruct Unregulated Runoff for Indicated Regions of Raritan and Delaware River Basins

Station number (1)	Station name (2)	Drainage area (km^2) (3)	Period of record used (4)	Model region (5)
01396500	South Branch Raritan River near High Bridge, N.J.	169	10/1918–9/1993	South Branch Raritan
01397000	South Branch Raritan River at Stanton, N.J	381	10/1919–9/1963	South Branch Raritan
01398500	North Branch Raritan River near Far Hills, N.J.	67.9	10/1921–9/1975 10/1977–9/1993	North Branch Raritan
01399500	Lamington River near Pottersville, N.J.	85.0	10/1921–9/1993	North Branch Raritan
01400000	North Branch Raritan River near Raritan, N.J.	492	6/1923–9/1993	North Branch Raritan
01402000	Millstone River at Blackwells Mills, N.J.	668	8/1921–9/1993	Millstone-Raritan
01413500	East Branch Delaware River at Margeretteville, N.Y.	422	4/1937–9/1993	Delaware River above NYC reservoirs
01420500	Beaver Kill at Cooks Falls, N.Y.	624	10/1918–9/1993	Delaware River above and below NYC reservoirs

J. Water Resour. Plann. Manage., 1997, 123(6): 359-367

River basins (Table 1). Stations 01413500 and 01420500 are in the mountains of southern New York and generally have substantial spring runoff because of snowmelt. Stations 01397000, 01400000, and 0140200 in the Raritan River basin in central New Jersey generally experience less snowmelt runoff in the spring than the New York stations. In the unblocked bootstrap record the residuals chosen for each month at the multiple sites are contemporaneous across sites, but each month in the sequence may be chosen from a different year. The bootstrap results shown are based on 3,600 months of synthetic record. The top and bottom charts in Fig. 3 show that both the blocked and unblocked bootstrap approaches reproduce the observed cross-correlations fairly well for pairs of streams in similar hydrologic settings. The middle two charts show that only the blocked bootstrap reproduces the observed cross-correlations well when station pairs are in different hydrologic settings.

Accounting for Long-Range Weather Forecasts

Selecting the residuals at random in position analysis has another advantage aside from being able to produce many equally likely possible traces. The bootstrap model allows the user to specify the probabilities of having a drier than normal sequence, a wetter than normal sequence, and a normal sequence that coincide with the 90-day forecasts of the National Weather Service (Northeast Regional Climate Center 1996). Suppose that a 90-day forecast says that it is 5% more likely that precipitation will be below normal. Then the bootstrap selection process can be modified to make it 5% more likely to select a relatively negative 3-month sequence of innovations than a relatively positive sequence of innovations, which would make it more likely for a "dry" position analysis trace to be generated.

Accounting for Parameter Uncertainty

The effects of parameter uncertainty on synthetic streamflow generation have been clearly established (Stedinger and Taylor 1982). Cover and Unny (1986) use the bootstrap method to analyze the uncertainty in ARMA models. A similar approach is used here to account for the effects of parameter uncertainty on the PARMA model used here. The effects of parameter uncertainty and long-range weather forecasts are included by following the steps:

Step 1. From an observed N-year record of LT runoff data, monthly means and standard deviations are calculated and the standardized runoffs computed [see (1)] for each site. Parameters ϕ and θ_j are estimated and residuals calculated [see (3)] and rescaled to correct for bias.

Step 2. A bootstrap N-year record is computed for each site by randomly drawing with replacement from the residuals in step 1 $N + 2$ blocks of 12 consecutive sample innovations (contemporaneous across sites) and recursively calculating standardized runoffs for $N + 2$ years. The first 2 years of calculated values are discarded to overcome the effects of arbitrary starting values. New monthly statistics and parameters for the N-year bootstrap record are calculated and saved. A new set of residuals are calculated from the bootstrap "record" and saved.

In addition, for each residual the sum of the residual and the next two residuals is computed and the sum ranked with the all the sums for each month of the year. The residuals with ranks in the lowest third are classified as below normal. The residuals with ranks in the middle and upper thirds are classified as normal, and above normal, respectively. Thus each residual carries with it a classification (below normal, normal, or above normal) based on the sum of the residual and its two subsequent residuals. This classification scheme will be used in accounting for long-range weather forecasts. Step 2 is repeated B times. As a result one has B sets of monthly means and standard deviations, B sets of parameters ϕ and θ_j, and B sets of monthly residuals N-year in length from which to generate the position analysis runoff sequences.

Step 3. To generate one 12-month runoff sequence for the position analysis model with beginning month k, randomly choose one of the B sets of parameters and residuals. Next, based on the 90-day forecast of the Weather Service, set the probability of precipitation being below normal, normal, or above normal and randomly choose one of the three classifications accordingly. From among the N years of residuals randomly choose a 12-month sequence of residuals beginning with month k from among all the months k with the chosen classification. Finally, recursively compute a 12-month runoff sequence using specified starting values for runoff and innovation, the chosen parameters and monthly statistics, and the selected sequence of residuals. Repeat step 3 B_p times.

By generating a large number of likely sequences conditioned on the same starting values, as illustrated in Fig. 4, one may construct conditional frequency curves for runoff for a time several months ahead. Fig. 5 shows frequency curves for June runoff based on unconditional observed data and two bootstrap position analysis runs. Both runs are based on 300 bootstrap sequences started with the same relatively low March runoff. Run 1 set the 90-day weather forecast probabilities of above-normal, normal, and below-normal precipitation all at 0.333, which are the nominal probabilities for the National Weather Service 90-day outlooks. In run 2, the below-normal and above-normal precipitation probabilities were changed to 0.433 and 0.233, respectively, which are the probabilities for a "10% below-normal" 90-day outlook. The difference between the observed data curve and the position analysis curves illustrates the effects of starting the position analysis from a low runoff position. The difference between runs 1 and 2 illustrates the effects of the 90-day forecast on

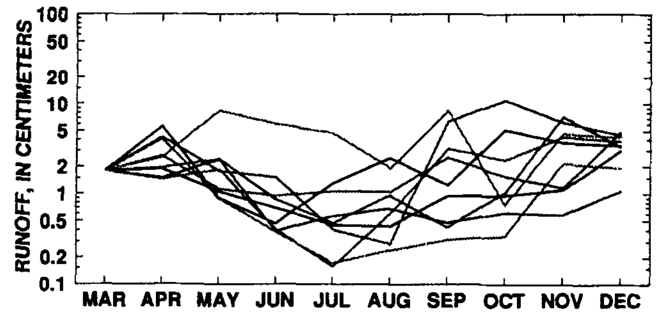

FIG. 4. Ten of 300 Flow Sequences Generated by Bootstrap Procedure for Millstone River at Blackwells Mills, N.J. Conditioned on March Runoff of 1.7 cm and Drier than Normal 90-day Weather Forecast

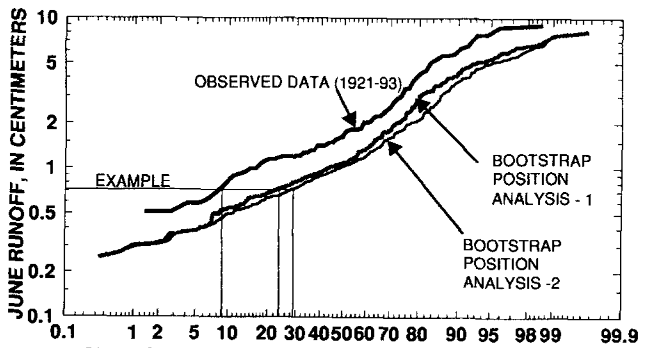

FIG. 5. Frequency Curves for June Runoff at Millstone River at Blackwells Mills, N.J. Based on Observed Data and Results for two 300 Sequence Bootstrap Position Analysis Runs

J. Water Resour. Plann. Manage., 1997, 123(6): 359-367

the frequency estimates. For example, June runoffs below about 0.7 cm would be expected to occur about 9% of the time in the long run. If the current March runoff is low, the risk of the coming June runoff of less than 0.7 cm would increase to about 24%, and if the 90-day forecast predicted 10% excess likelihood of below-normal precipitation, the risks increase further to about 30%.

In the following example, 300 (B_p = 300) flow sequences were generated for five runoff sites to drive the position analysis model. The sequences were drawn on 50 (B = 50) sets of possible parameters and 50 sets of residuals each 50 years (N = 50) in length.

APPLICATION TO WATER-SUPPLY RESERVOIR SYSTEM IN NEW JERSEY

Description of Water-Supply System

The Raritan River basin reservoir system is located in central New Jersey including Hunterdon, Mercer, Middlesex, Monmouth, Morris, and Somerset counties (Fig. 6). The reservoirs and pumping stations of the Raritan River basin reservoir system and its interconnections to the Delaware-Raritan Canal water-supply system supply potable water to central New Jersey communities. This combined system includes Spruce Run reservoir, Round Valley reservoir, Hamden pumping station, and the Delaware-Raritan Canal. The systems can easily be strained by an extended period of below-average precipitation.

Spruce Run reservoir, usable capacity of 42,000,000 m^3, is located on a tributary of the South Branch Raritan River and has a drainage area of 107 km^2. Round Valley reservoir (capacity at the spillway level of 208,000,000 m^3) is located on a branch of the South Branch Rockaway Creek and has a drainage area of 14.8 km^2. Most of the water needs to be pumped into Round Valley reservoir because of its small drainage area. Water is pumped from the South Branch Raritan River at the Hamden pumping station near Stanton, N.J. The drainage area at the pumping station is approximately 363 km^2. Primary water releases from Round Valley reservoir are made through a 2.7-m-diameter pipeline, which discharges to the South Branch Rockaway Creek.

The Delaware and Raritan Canal is an integral part of the Raritan River basin reservoir system. The canal is used to divert water from the Delaware River to be used for water supply or to help meet passing flow requirements on the Raritan River at Bound Brook below Calco Dam. The canal is con-

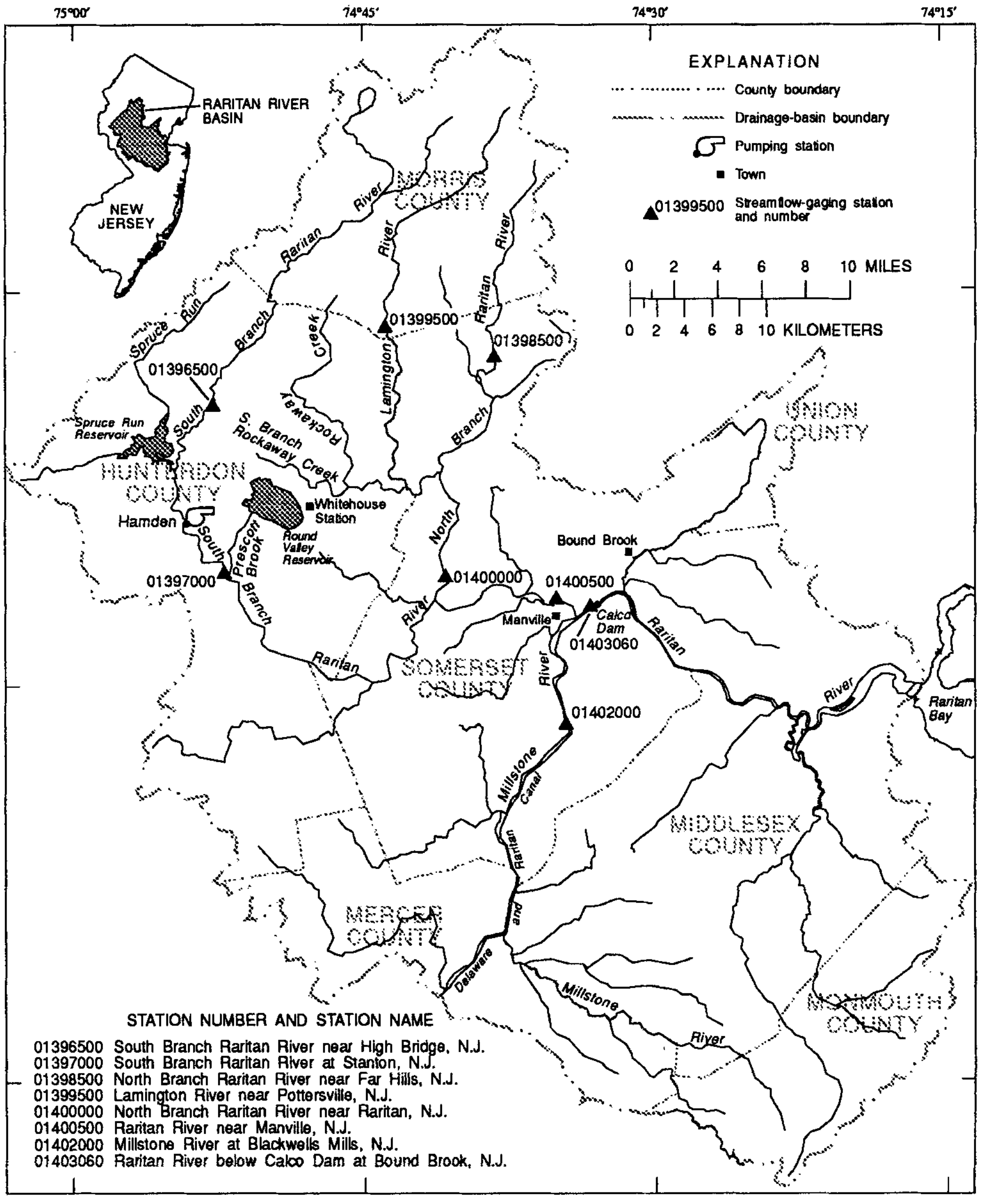

FIG. 6. Map of Raritan Basin

J. Water Resour. Plann. Manage., 1997, 123(6): 359-367

nected to the reservoir system through the Millstone River near Bound Brook. Water can be released into the Millstone River or pumped from the river into the canal. The amount of water that can be diverted from the Delaware River at any one time is determined by agreement with the Delaware River basin commission and depends upon the storage available in three large New York City reservoirs in the upper Delaware River basin.

The Raritan River basin reservoir system must meet statutory flow requirements at three locations. These requirements include 1.75 m^3/s on the south branch of the Raritan River at Stanton, 3.07 m^3/s on the Raritan River at Manville, and 3.94 m^3/s on the Raritan River at Bound Brook below Calco Dam.

Unregulated Natural Runoff for Basin

The first step in developing the position analysis model for the basin is to determine the natural unregulated monthly runoff values in the basin for a base period. U.S. Geological Survey (USGS) gauging station records in the area are used for this purpose. Table 1 shows the stations and periods of unregulated flow used. A 75-year base period of October 1918 to September 1993 was selected. Flow records were extended by regression using the MOVE.1 technique (Hirsch et al. 1993) when the period of record for unregulated flow was less than the base period. Hirsch (1982) and Vogel and Stedinger (1985) recommend the MOVE.1 technique for extending records used in simulation studies of the probability of water-supply shortage because the MOVE.1 method preserves important statistical properties of the flow rather well. With the use of these extended flow records, time series of monthly runoff values were developed for the base period for three regions in the Raritan River basin and two regions in the Delaware River basin. The three regions of the Raritan River are the south branch of the Raritan, the north branch of the Raritan, and the remaining drainage above the gauge on the Raritan at Bound Brook (which includes the Millstone River basin). The two regions in the Delaware River basin include the area above the three New York City reservoirs and the area between the reservoirs and the gauge at Montague, N.J.

Following steps 1 and 2 outlined in the foregoing paragraph, fifty 52-year bootstrap records were generated for each of the five regions and the PARMA model parameters along with means and standard deviations of the monthly runoffs were determined for each region and each record. The means and standard deviations of the parameters ϕ and θ_j based on the 50 records are shown in Table 2. After eliminating the first 2 years of each bootstrap record, there were 2,500 years of monthly innovations and 50 sets of parameters for each region from which to generate cross-correlated position analysis sequences 12 months in length for the five regions (step 3).

Basin Model

The basin model is a continuity accounting model consisting of a series of interconnected nodes. At each node the monthly inflow volume, outflow volume, and change in storage are determined and recorded for each month. Fig. 7 is a flow diagram showing the nodes (as boxes) and flow paths as lines (arrows represent inflows and outflows). The model is driven by the natural inflows into each node, which are the natural runoff traces generated by the bootstrap position analysis model. A set of operating rules control the reservoir and canal releases and pumpage to meet passing flow and withdrawal demands. The basin model runs with a set of default operating rules for releases, pumpage, and diversions that the user may change to see the effects of alternative sets of rules and evaluate their usefulness.

Operating Rules

The following is a brief description of a complex set of operating rules for the basin. The model user may modify these rules, the reservoir capacities, passing flow requirements, or maximum pumping capacities before running the model to test alternative methods for management of the resources. A detailed description of the system and operating rules can be found in Dunne and Tasker (1996).

At any time of the year, the combined storage in the New York City reservoirs may be in one of four states—normal, upper warning, lower warning, or drought. The amount of the diversions to the Raritan River via the Delaware-Raritan Canal, the amount of withdrawals allowed to New York City, and the passing flow requirements for the Montague gauge are set for each state of the reservoirs.

As long as storage in the Spruce Run reservoir in the Raritan basin is above a set recreation minimum, releases are made from Spruce Run as needed to meet the withdrawal demands for the basin and passing flow demands at Stanton, Manville, and Bound Brook. If storage falls below the recreation minimum in a given month, then releases are made only to meet demands for passing flow at Stanton until storage rises above the recreation minimum once again. The model user may

TABLE 2. Mean and Standard Deviation of Estimates of Parameters for PARMA(1, 1) Model for Indicated Regions Based on 50 Bootstrap Replications

	REGION									
	Delaware River above NYC Reservoirs		Delaware River below NYC Reservoirs		South Branch Raritan River		North Branch Raritan River		Raritan and Millstone Rivers	
Parameter (1)	Mean (2)	SD (3)	Mean (4)	SD (5)	Mean (6)	SD (7)	Mean (8)	SD (9)	Mean (10)	SD (11)
ϕ	0.43	0.09	0.43	0.08	0.74	0.06	0.66	0.07	0.58	0.09
θ_1	0.05	0.16	0.02	0.15	0.17	0.14	0.10	0.13	0.07	0.14
θ_2	0.39	0.22	0.30	0.22	0.81	0.22	0.76	0.20	0.90	0.22
θ_3	0.64	0.14	0.61	0.14	0.90	0.16	0.77	0.16	0.93	0.24
θ_4	0.67	0.15	0.77	0.12	0.25	0.13	0.21	0.14	0.23	0.24
θ_5	0.45	0.17	0.54	0.16	0.41	0.16	0.37	0.16	0.51	0.19
θ_6	−0.15	0.12	−0.16	0.11	0.12	0.10	0.00	0.12	0.03	0.13
θ_7	−0.43	0.17	−0.30	0.02	0.26	0.18	0.05	0.19	0.14	0.17
θ_8	−0.19	0.31	−0.20	0.21	0.33	0.20	0.20	0.20	0.21	0.19
θ_9	−0.24	0.16	−0.23	0.16	0.12	0.13	−0.01	0.14	0.24	0.17
θ_{10}	−0.35	0.16	−0.32	0.12	0.03	0.12	0.00	0.18	0.06	0.14
θ_{11}	−0.37	0.25	−0.31	0.21	0.24	0.18	0.19	0.15	0.21	0.13
θ_{12}	0.12	0.26	0.21	0.20	0.19	0.14	0.17	0.14	0.26	0.16

Note: SD, standard deviation.

J. Water Resour. Plann. Manage., 1997, 123(6): 359-367

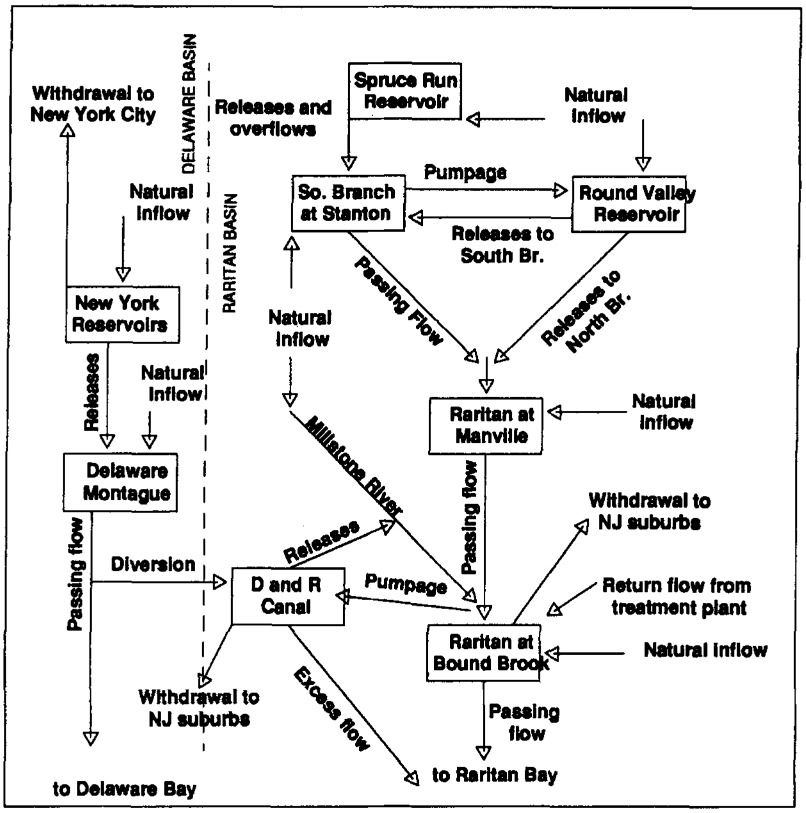

FIG. 7. Flow Diagram for Basin Model of Raritan Water-Supply System

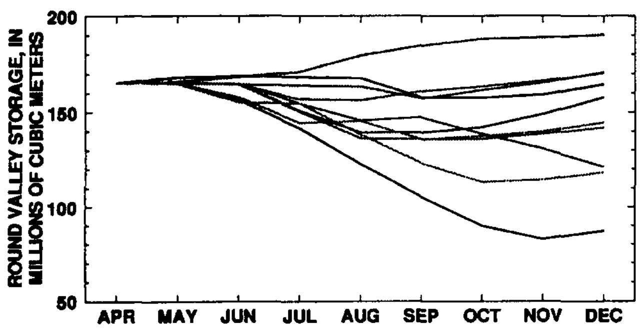

FIG. 8. Ten of 300 Bootstrap Position Analysis Sequences for Storage in Round Valley Reservoir Initialized with April Storages and Flows and Operating Rule 2

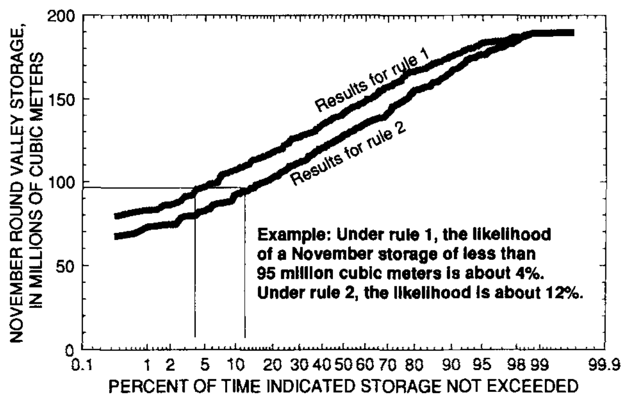

FIG. 9. Storage Frequency Curves for Round Valley Reservoir at end of November Based on 300 Bootstrap Position Analysis Sequences, Assumed Storages and Flows at End of April, and Assumed Operating Rules

change the recreation minimums, the reservoir capacity, and the minimum flow release.

Releases from Round Valley reservoir to South Branch are made to meet Stanton passing flow requirements that cannot be met by unregulated flows or by releases from Spruce Run. Releases to North Branch are made to meet withdrawal demands and passing flow requirements at Manville and Bound Brook that cannot be met by unregulated flows, by releases from Spruce Run, or by diversions from the Delaware-Raritan Canal. Inflow into Round Valley comes from natural runoff supplemented by pumped water from the Hamden pumping station above the Stanton gauge. Pumping is only done when minimum passing flows at Stanton are met and when storage in Round Valley is below target levels. The model user sets the target storage levels for each month. The model user may also change the reservoir capacity.

Example

Suppose that it is currently the end of April, all reservoirs are 80% full, April runoff is known, and the long-range weather forecast is for 5% drier than normal conditions. A water manager wishes to test the effects on Round Valley reservoir storage of changing the recreation minimum levels of Spruce Run storage. In rule 1, recreation minimum storages are set at 11,000,000 m^3. In rule 2, the recreation minimum for Spruce Run is raised to 28,000,000 m^3, which will mean greater recreation use of the reservoir. Ten of the 300 model-generated traces of Round Valley storage for rule 2 are shown in Fig. 8. A frequency analysis of all traces for rule 1 and 2 for the end of November (Fig. 9) shows the increased likelihood under rule 2 of storages in Round Valley at the end of November falling below a specified level. An analyst may wish to use this information to compare the expected increase in recreation benefits for Spruce Run with increase in pumping costs to refill Round Valley and loss in recreation benefits for Round Valley.

SUMMARY AND CONCLUSIONS

A method of random resampling of residuals in stochastic models is used to generate a large number of traces of natural

J. Water Resour. Plann. Manage., 1997, 123(6): 359-367

monthly runoff to be used in a position analysis model for a water-supply storage and delivery system. Resampling the residuals eliminates the need to make parametric assumptions about the distribution of the innovations and simplifies parameter estimation. One only assumes that the best estimate of the population of the innovations is the sample itself. Lag-zero and lag-one cross-correlations of the innovations in a multisite application are approximately preserved by having the innovations for multiple sites stay together in a contemporaneous block resampling scheme. The effects of long-range weather predictions can be included by setting the probabilities that relatively large negative or large positive innovations will be randomly drawn. The position analysis model (Hirsch 1978) is a powerful forecasting tool that allows for the quantification of risks associated with different management options. It provides the water resources manager with a means to rank alternative operating rules and evaluate risks-cost trade-offs.

ACKNOWLEDGMENTS

The writers wish to thank the New Jersey Water Supply Authority for their support for the development of the basin model described here. They also wish to thank Aldo Vecchia, Greg Schwarz, Mike Karlinger, and three anonymous referees for their helpful comments and insightful discussions of earlier drafts of this paper.

APPENDIX I. REFERENCES

Cover, K. A., and Unny, T. E. (1986). "Application of computer intensive statistics to parameter uncertainty in streamflow synthesis." *Water Resour. Bull.* 22(3), 495–507.

Day, G. N. (1985). "Extended stream flow forecasting using NWSRFS." *J. Water Resour. Plng. and Mgmt.*, ASCE, 111(2), 157–170.

Dunne, P., and Tasker, G. (1996). "Computer model of the Raritan River basin water-supply system in central New Jersey." *USGS Open-file Rep. No. 96-360.*

Efron, B. (1979). "Bootstrap methods: another look at the jackknife." *Ann. of Statistics*, 7(1), 1–26.

Efron, B. (1982). *The jackknife, the bootstrap and other resampling plans.* Society for Industrial and Applied Mathematics, Philadelphia, Pa.

Haltiner, J. P., and Salas, J. D. (1988). "Development and testing of a multivariate seasonal ARMA(1,1) model." *J Hydro.*, Amsterdam, The Netherlands, 104, 247–272.

Hirsch, R. M. (1978). "Risk analysis for a water-supply system—Occoquan Reservoir, Fairfax and Prince William counties, Virginia." *Hydro. Sci. Bull.*, 23(4), 476–505.

Hirsch, R. M. (1979). "Synthetic hydrology and water supply reliability." *Water Resour. Res.*, 15(6) 1603–1615.

Hirsch, R. M. (1981). "Stochastic hydrologic model for drought management." *J. Water Resour. Plng. and Mgmt. Div.*, ASCE, 107(2), 303–313.

Hirsch, R. M. (1982). "A comparison of four streamflow record extension techniques." *Water Resour. Res.*, 18(4), 1081–1088.

Hirsch, R. M., Helsel, D. R., Cohn, T. A., and Gilroy E. (1993). "Statistical analysis of hydrologic data." *Handbook of hydrology*, D. R. Maidment, ed. McGraw-Hill, Inc., New York, N.Y. 17.1–17.55.

Kunsch, H. R. (1989). "The jackknife and the bootstrap for general stationary observations." *Ann. of Statistics*, 17(3), 1217–1241.

Lall, U., and Sharma, A. (1996). A nearest neighbor bootstrap for resampling hydrologic time series." *Water Resour. Res.*, 32(3), 679–693.

Loucks, D. P., Stedinger, J. R., and Haith, D. A. (1981). *Water Resour. Syst. Plng. and Anal.* Prentice-Hall, Inc., Englewood Cliffs, N.J.

Moss, M. E., and Tasker, G. D. (1991). "An intercomparison of hydrological network-design technologies." *Hydrological Sci. J.*, Oxford, U.K., 36(3), 209–221.

Northeast Regional Climate Center. (1996). *Northeast climate impacts.* Northeast Regional Climate Center, Cornell University, Ithaca, N.Y.

Oliverira, C. G., Kelman, J., Pereira, M. V. F., and Stedinger, J. R. (1988). "A representation of spatial correlations in large stochastic seasonal streamflow models." *Water Resour. Res.*, 24(5), 781–785.

Pereira, M. V. F., Oliverira, G. C., Costa, C. G., and Kelman, J. (1984). Stochastic streamflow models for hydroelectric systems." *Water Resour. Res.*, 20(3), 379–390.

Rasmussen, P. F., Salas, J. D., Fagherazzi, L., Rassam, J., and Bobee, B. (1996). "Estimation and validation of contemporaneous PARMA models for streamflow simulation." *Water Resour. Res.*, 32(10), 3151–3160.

Salas, J. D. (1993). "Analysis and modeling of hydrologic time series." *Handbook of hydrology*, D. R. Maidment, ed. McGraw-Hill, Inc., New York, N.Y., 19.1–19.72.

Salas, J. D., Boes, D. C., and Smith, R. A. (1982). "Estimation of ARMA models with seasonal parameters." *Water Resour. Res.*, 18(4) 1006–1010.

Salas, J. D., Dellur, J. W., Yevjevich, V., and Lane, W. L. (1980). *Applied modeling of hydrologic time series.* Water Resources Publications, Littleton, Colo.

Salas, J. D., Tabios, Guillermo Q., and Bartolini, P. (1985). "Approaches to multivariate modeling of water resources time series." *Water Resour. Bull.*, 21(4), 683–708.

Sharma, A., Tarboton, D. G., and Lall, U. (1997). "Streamflow simulation: a nonparametric approach." *Water Resour. Res.*, 33(2), 291–308.

Stedinger, J. R., and Taylor, M. R. (1982). "Synthetic streamflow generation, 2. Effect of parameter uncertainty." *Water Resour. Res.*, 18(4), 919–924.

Tasker, G. D. (1987). "Comparison of methods for estimating low flow characteristics of streams." *Water Resour. Bull.*, 23(6), 1077–1083.

Thombs, L. A., and Schucany, W. R. (1990). "Bootstrap prediction intervals for autoregression." *J. Am. Statistical Assoc.*, 85(410), 486–492.

Valencia, D. R., and Schaake, J. C. (1973). "Disaggregation process in stochastic hydrology." *Water Resour. Res.*, 9(3) 580–585.

Vecchia, A. V. (1985a). "Maximum likelihood estimation for periodic autoregressive moving average models." *Technometrics*, 27(4), 375–384.

Vecchia, A. V. (1985b). "Periodic autoregressive-moving average (PARMA) modeling with applications to water resources." *Water Resour. Bull.*, 21(5), 721–730.

Vecchia, A. V., Obeysekera, J. T. B., Salas, J. D., and Boes, D. C. (1983). "Aggregation and estimation for low-order periodic ARMA models" *Water Resour. Res.*, 19(5), 1297–1306.

Vogel, R. M., and Stedinger, J. R. (1985). "Minimum variance streamflow augmentation procedures." *Water Resour. Res.*, 21(5), 715–723.

Vogel, R. M., and Shallcross, A. L. (1996). "The moving blocks bootstrap versus parametric time series models." *Water Resour. Res.*, 32(6), 1875–1882.

Woo, M. K. (1989). "Confidence intervals of optimum risk-based hydraulic design parameters." *Can. Water Resour. J.*, 14(2), 10–16.

Zucchini, W., and Adamson, P. T. (1989). "Bootstrap confidence intervals for design storms from exceedence series." *Hydrological Sci. J.*, Oxford, U.K., 34(1), 41–48.

APPENDIX II. NOTATION

The following symbols are used in this paper:

B = number of sets of parameters generated by bootstrap;
B_p = number of position analysis sequences generated;
E_j = innovation for month j;
N = number of years of runoff generated for each set of parameters;
S_j = standard deviation of logarithms of runoff for month j;
X_j = runoff for month j;
Y_j = log of X_j;
$\bar{Y}_j$ = mean of the logarithms of runoff values for month j;
Z_j = standardized runoff for month j; and
ϕ, θ_j = parameters of PARMA(1,1) model.