The Encyclopedia of Water

Stochastic Simulation of Hydrosystems (SW-913)

Demetris Koutsoyiannis
Department of Water Resources, Faculty of Civil Engineering, National Technical University, Athens, Heroon Polytechneiou 5, GR-157 80 Zographou, Greece
dk@itia.ntua.gr

Introduction

Simulation is defined to be a technique to imitate the evolution of a real system by studying a model of the system (Winston, 1994, p. 23; Ripley, 1987, p. 2). The model is an abstraction, a simplified and convenient mathematical representation of the actual system typically coded and run as a computer program. If the model has a stochastic[1] element, then we have *stochastic simulation*. The term stochastic simulation sometimes is used synonymously to the *Monte-Carlo method*.

Stochastic simulation is regarded as mathematical experimentation and is appropriate for complex systems, whose study based on analytical methods is laborious or even impossible. For such systems, stochastic simulation provides an easy means to explore their behavior by answer specific 'what if' questions. Moreover, stochastic simulation can be viewed as a numerical method to solve mathematical problems in several fields like statistical inference, optimization, integration, and even equation solving. Under certain conditions stochastic simulation is more powerful than other more common numerical methods (e.g. numerical integration of high-dimensional differential equations).

---

1 The term *stochastic* is etymologized from the Greek verb *στοχάζομαι* initially meaning to aim, point, or shoot (an arrow) at a target (*στόχος* is a target). Metaphorically, the verb meant to guess or to conjecture (the target).. The modern Greek meaning is to imagine, to think, to meditate. It appears that the word *stochastic* is found in English since the 17th century with the obsolete meaning pertaining to conjecture. Its use as a scientific term is attributed to the Swiss mathematician Jakob Bernoulli (1654-1705), who in his famous masterpiece *Ars Conjectandi* realised that randomness and uncertainty are important aspects of our world and should be objects of scientific analysis. In its modern sense, *stochastic* can be regarded as synonymous to random or probabilistic, but it is mostly used for processes that indicate a mixture of structure and randomness; the term *stochastic process* was used in 1932 by A. N. Kolmogorov.

Due to their complexity, hydrosystems, including water resource systems, flood management systems, hydropower systems, etc., are frequently studied using stochastic simulation. In particular, hydrologic processes that have a dominant role in hydrosystems are often regarded as stochastic processes. Stochastic hydrology, the application of theory of stochastic processes in analysis and modeling of hydrologic processes, has offered very efficient tools in tackling a variety of hydrosystems problems, including systems identification, modeling and forecasting, hydrologic design, water resources management, and flood management.

A brief history

Synthetic streamflow records were first used early in the 20th century by Hazen (1914) in studies of water supply reliability. Their construction, however, was not based on the theory of stochastic processes, then not developed, but on merging and rescaling observed records of several streams. This early work emphasizes the need for long synthetic records and the importance of simulation in water resources technology. The foundation of stochastic hydrology followed the significant developments in mathematics and physics in the 1940s, as well as the development of computers. Specifically, it followed the establishment of the Monte Carlo method, which was invented by Stanislaw Ulam in 1946. Notably, Ulam conceived the method while playing solitaires during convalescing from an illness, in his attempt to estimate the probabilities of success of the plays. As Ulam describes the story in some remarks later published by Eckhardt (1989), "After spending a lot of time to estimate them by pure combinatorial calculations, I wondered whether a more practical method than 'abstract thinking' might not be to lay it out say one hundred times and simply observe and count the number of successful plays". Soon after the method grew to solve neutron diffusion problems by himself and other great mathematicians and physicists in Los Alamos (John von Neumann, Nicholas Metropolis, Enrico Fermi), and was first implemented on the ENIAC computer (Metropolis, 1989; Eckhardt, 1989). The 'official' history of the method began in 1949 with the publication of a paper by Metropolis and Ulam (1949).

In the field of water resources, the most significant initial steps were the works by Barnes (1954) for the generation of uncorrelated annual flows at a site from normal distribution; Maass et al. (1962) and Thomas and Fiering (1962) for the generation of flows correlated in time; and Beard (1965) and Matalas (1967) for the generation of concurrent flows at several sites.

The classic book on time series analysis by Box and Jenkins (1970) was also originated from different, more fundamental scientific fields. However, it has subsequently become very popular in stochastic hydrology. Box and Jenkins developed a classification scheme for a large family of time series models. Their classification distinguishes among autoregressive models of order p (AR(p)), moving average models of order q (MA(q)) combinations of the two, called autoregressive-moving average (ARMA(p, q)) models, and autoregressive integrated moving average (ARIMA(p, d, q)) models. However, despite making a large family, Box-Jenkins models do not cover fully the needs of hydrologic modeling, as they do not comply with some peculiarities of hydrologic and other geophysical processes. This gave rise to substantial research that resulted in numerous stochastic tools appropriate for applications in water recourses.

Utility of stochastic simulation in hydrosystems

Due to the significant uncertainties inherent to hydrosystems, among which the major is hydrologic uncertainty (related to the unknown future of inflows to hydrosystems), the estimation of a system's reliability is important for its design and operation. The reliability of a system is defined to be the probability that a system will perform the required function for a specified period of time under stated conditions (Chow et al., 1988, p. 434). Reliability is the complement of probability of failure or risk, i.e. the probability that the "loading" will exceed the "capacity". In many instances the risk can be estimated using analytical means, so stochastic simulation is not required. For example, in the design of dykes that confine a river's flow, the risk of overtopping of dykes can be estimated in a typical probabilistic manner, provided that there exists a record of floods of the river with sufficient length (some decades). The estimation procedure includes the selection of a probabilistic model (e.g. an

extreme value distribution function), the fitting of the model based on the available record, and the estimation of the probability that flood exceeds the discharge capacity of the designed river cross section (the estimation of the latter is a matter of hydraulics). Behind this procedure, there are two implicit assumptions that make the methodology appropriate for this example problem:

1. The project under study (the dykes) does not modify the natural flow regime, so that if the project were constructed many years before, the observed flow record would not be altered. Thus, the assumed probabilistic model, although fitted on past data, is still valid after the construction of the project.
2. The quantity (flood discharge) whose exceedance was assumed to be the risk is the same quantity, for which we have observed data. Thus, the probabilistic model that was constructed for this quantity can directly yield the risk.

In many cases, however, these assumptions are not valid. Let us first examine the case where the assumption 2 is untrue. For example, we may have available rainfall data, from which we can construct a probabilistic model for extreme rainfall intensity, and wish to estimate the probability of exceedance of flood discharge. In this case we can use a simple one-to-one mapping (transformation) of rainfall to discharge values (e.g. to adopt the relation known as the rational formula), so that the risk of exceeding a certain discharge level equals the risk of exceeding the corresponding rainfall level. This methodology usually incorporates serious oversimplifications and ignorance of certain factors that affect the actual hydrologic process (e.g. retention, infiltration, etc.). A more realistic methodology is to use a more detailed model that transforms a rainfall series (not each isolated value) into a discharge series, also considering all processes involved in this transformation. In this case we can use simulation to obtain a discharge series.

The assumption 1 can be untrue in many cases as well. For example, the construction of a dam will alter the flood regime at the dam and downstream, as the spillway outflow does not equal the natural inflow (attenuation occurs due to temporal flood storage). Also, the construction of a storm sewer will modify the contributing areas and flow times in the area (in

addition, it is impossible to have observed data for the sewer discharge in its design phase). Another typical example is a reservoir (see the entry Reliability Concepts in Reservoir Design – SW-776), whose storage (a quantity that determines the risk, which is the probability of emptying of the reservoir) is inexistent before the construction of the reservoir. Obviously, in all these cases where assumption 1 is not valid, assumption 2 is too not valid. Thus, we will proceed as in the previous paragraph with simulation being the most appropriate procedure to obtain a series of data values for the quantity of interest.

Even in an existing project (e.g. an existing reservoir), where the quantity of interest could be measured directly to obtain a historical record, simulation may be again necessary to assess impacts of several possible changes in the future that were not experienced in the past. For example, a change on water use (e.g. increase of water demand) and a change of land use or climate, which alters water availability, calls for simulation to estimate a series for the quantity of interest for the examined scenario.

The above discussion explains why in most studies of hydrosystems (with the exception of cases where both assumptions listed above are valid) it is necessary to perform simulation as a means for transformation of some input time series of initial quantities to some output time series of the final quantities of interest. Grace to the power of computers, the simulation methodology has greatly replaced older methodologies that used simplified one-to-one transformations. But why simulation should be stochastic?

In stochastic simulation the input time series are no longer the observed records but synthetic time series constructed by an appropriate stochastic model. An observed time series is unique and has a limited length equal to the period of observations. On the contrary, a stochastic model can produce as many time series as required and of any arbitrary length. The utility of a long length of time series becomes obvious in steady state simulations (Winston, 1994, p. 1220), when a low value of probability of failure (risk) is to be estimated. For example, in a problem where the accepted probability of failure is 1% per year, apparently several hundreds of simulated years are needed to detect a few failures. The utility of ensemble time series (as opposed to the unique observed record) becomes obvious in non-steady state problems (i.e. in terminating simulations) and in forecast problems in which the

initial conditions (present and past values of the processes of interest) are known. In these cases, stochastic simulation offers the possibility of different sample paths of the quantity of interest, instead of having a single value at a time, so that we can estimate expected values and confidence zones.

Components and solution procedure of stochastic simulation

The components and the steps followed in stochastic simulation of a hydrosystem are demonstrated in Figure 1. The entire procedure includes two main model components (marked 1 and 2 in Figure 1) and two simpler procedures (marked 3 and 4 in Figure 1). The first model component is the stochastic model of inputs, which produces a vector $\mathbf{X}(\boldsymbol{\mu}, \omega)$ of hydrological inputs (e.g. time series of rainfall, evaporation, river flow, depending on the problem studied) to the hydrosystem, where $\boldsymbol{\mu}$ is a vector that contains the parameters of hydrologic inputs (all estimated from the available records of observations) and ω denotes a sample path realization of the random variables (that is, ω can be thought of as representing the randomness in the system, e.g., all random numbers in a simulation run). At a minimal configuration, the vector of parameters $\boldsymbol{\mu}$ includes mean values, standard deviations, autocorrelations (at least for lag one) and cross-correlations (for multiple-site models).

The second component is the transformation model which takes the inputs $\mathbf{X}(\boldsymbol{\mu}, \omega)$ and produces the outputs $\mathbf{Z}(\mathbf{X}(\boldsymbol{\mu}, \omega), \boldsymbol{\lambda})$ (e.g. river flow if $\mathbf{X}$ is rainfall and evaporation, or reservoir release and storage if $\mathbf{X}$ is river flow); here the vector $\boldsymbol{\lambda}$ contains parameters of the transformation model (e.g. parameters determining the hydrological cycle in a basin and/or parameters determining the operation of a specific project like a reservoir).

The third component is a procedure that takes the outputs $\mathbf{Z}(\mathbf{X}(\boldsymbol{\mu}, \omega), \boldsymbol{\lambda})$ and determines a sample performance measure $L(\mathbf{Z}(\mathbf{X}(\boldsymbol{\mu}, \omega), \boldsymbol{\lambda}))$ of the system that corresponds to the sample realization represented by ω. This performance measure depends of the problem examined; for instance, in a design flood problem it can be the risk of exceeding a specified flood level; in a reservoir design problem it can be either the risk of emptying a reservoir, or the attained release for a stated reliability.

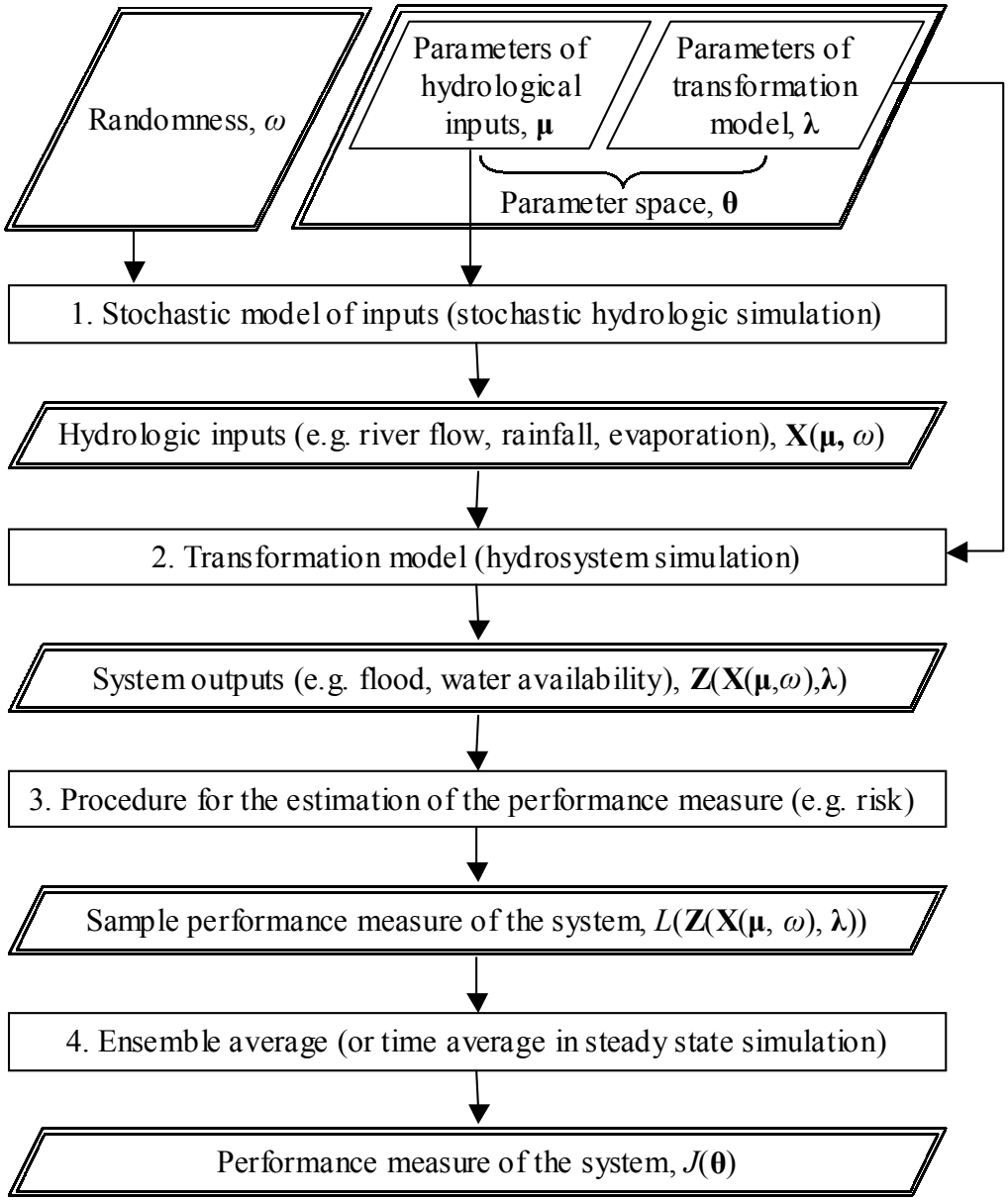

Figure 1 Schematic representation of the components and the solution procedure in hydrosystem simulation. Rectangles represent the components of the solution procedure, whereas parallelograms represent inputs and outputs to the different components.

By repeating the execution of these three components using different simulation runs, represented by different ω, we can obtain an ensemble of simulations and a sample of performance measures, from which we can estimate the true (independent of ω) performance measure of the system $J(\boldsymbol{\theta}) := E[L(\mathbf{Z}(\mathbf{X}(\boldsymbol{\mu}, \omega), \boldsymbol{\lambda}))]$, where $E[\ ]$ denotes expected value and $\boldsymbol{\theta} := (\boldsymbol{\mu}, \boldsymbol{\lambda})$. However, if the system is stationary and ergodic (in other words, if we have a steady state simulation), L will tend to J as the simulation length tends to infinity. Therefore, a single instance of the sample performance measure, estimated from a simulation with a large length,

is an adequate estimate of the true performance measure. This is the case, for instance, in a reservoir simulation with constant water demand. Conversely, if the water demand is growing in time (a common situation in practice), the simulation is no more a steady state one, and numerous runs, typically with a small length each one, must be performed to estimate the true performance measure.

In the following part of the article we will focus on the model component 1, i.e. the stochastic model of inputs.

Typical Box-Jenkins models

Let X_i denote the process of interest (e.g. rainfall or streamflow at a site) with i denoting discrete time. To generate a time series of X_i we start generating a sequence of independent identically distributed variables (iid, also known as white noise) V_i having a specified distribution function (e.g. Gaussian). This is known as generation of random numbers; a concise introduction to this topic can be found in Papoulis (1990) and a more detailed presentation can be found in Ripley (1987). If X_i can be assumed stationary (i.e. with probability distribution function that does not vary in time, which is typically the case if the time step is a year), the unstructured sequence of V_i can be converted to a structured sequence X_i by means of a recursive relationship, whose general form has been studied by Box and Jenkins (1970). From the large family of Box-Jenkins processes, the ones that have been widely used in stochastic hydrology are special cases of the equation

$$X_i = a\, X_{i-1} + a'\, X_{i-2} + b\, V_i + b'\, V_{i-1} \tag{1}$$

where a, a', b and b' are parameters that are estimated from auto-covariance properties of the process X_i. The special cases are

1. The iid or white noise or AR(0) process in which $a = a' = b' = 0$; $b \neq 0$.
2. The Markovian or AR(1) process in which $a' = b' = 0$; $a, b \neq 0$.
3. The AR(2) process in which $b' = 0$; $a, b, a' \neq 0$.
4. The ARMA(1, 1) process in which $a' = 0$; $a, b, b' \neq 0$.

The complete form of (1), known as the ARMA(2, 1) process, has been not used so frequently in hydrology. The listed special cases preserve auto-covariance properties for lag 0 (case 1) to lag 2 (cases 2 and 3); beyond these lags the model autocovariance is zero (case 1) or tends to zero exponentially (cases 2-4).

Peculiarities in stochastic representation of hydrologic processes

Simple stationary models, such as the above described, often are not the best choice in hydrologic modeling because of several peculiarities of hydrologic processes, of which the most important are discussed below.

Seasonality. When the time scale of interest is finer than annual, hydrologic processes cannot be regarded as stationary because of the effect of season of the year to the properties of the process. A simple method often used to take into account seasonality is to standardise the process X_i using seasonal values of mean and standard deviation, i.e. setting $Y_i := (X_i - \mu_i) / \sigma_i$ and assuming that Y_i is a stationary process that can be modeled for instance by (1); here μ_i and σ_i denote the mean and standard deviation, respectively, of X_i, which are assumed to vary with i in a periodic manner. This standardization approach, however, is flawed; the stationarity assumption for Y_i implies that, apart from the mean and standard deviation, other statistical properties of X_i like autocorrelation and skewness do not vary in season, which is not true. A more precise way of modelling seasonality is to assume a cyclostationary (also known as periodic) process, which is expressed as in (1) but with parameters a, b a', b', ..., and statistics of the noise variables V_i varying with i in a periodic manner.

Long-term persistence. Box-Jenkins models like (1) are essentially of short memory type, that is, their autocorrelation structure decreases rapidly with the lag time. However, the study of long historical time series of hydrological and other geophysical processes has revealed that autocorrelations may be significant for large lags, e.g. 50 or 100 years. This property is related to the tendency of streamflows to stay above or below their mean for long periods, observed for the first time by Hurst (1951), or, equivalently, to multiple time scale fluctuations of hydrologic processes (Koutsoyiannis, 2002; see also the entry SW-434 - Hydrologic Persistence and the Hurst Phenomenon). Therefore, models like (1) are proven

inadequate in stochastic hydrology, as the long-term persistence of hydrologic processes is important to be reproduced (see the entry SW-776 - Reliability Concepts in Reservoir Design).

Intermittency. At fine time scales, some hydrologic processes like rainfall and in some cases streamflow appear as intermittent processes; thus, rainfall alternates between two states, the dry (zero rainfall) and wet (positive rainfall). This is manifested in the marginal probability distribution of rainfall depth by a discontinuity at zero. Box-Jenkins processes like the one in (1) need to be truncated to represent this discontinuity; this is not so easy, nor common. To model intermittency, alternative two-state processes, like two-state Markov chains (e.g. Haan, 1977, p. 302) and point process models (e.g. Waymire and Gupta, 1981) have been proposed.

Skewness. Another peculiarity of hydrologic processes is the skewed distribution functions observed mostly in fine and intermediate time scales. This is not so common in other scientific fields whose processes are typically Gaussian. Therefore, attempts have been made to adapt standard models to enable treatment of skewness (e.g. Matalas and Wallis, 1976; Todini, 1980; Koutsoyiannis, 1999, 2000). The skewness is mainly caused by the fact that hydrologic variables are non-negative and sometimes intermittent. Therefore, a successful modeling of skewness indirectly contributes at avoiding negative values of simulated variables; however, it does not eliminate the problem and some ad hoc techniques (such as truncation of negative values) are often used in addition to modeling skewness.

Spatial variation. Hydrologic processes evolve both in time and space. Typically, time series models consider only the temporal evolution. The most precise mathematical representation of hydrologic processes can be achieved extending the indexing set of the process from one dimension (representing time) to three dimensions (one for time and two for space). However, multidimensional modeling is not easy and has been implemented only in few cases (for example in continuous time and space modeling of rainfall; e.g. Waymire et al., 1984). A midway solution, which is the most common in stochastic hydrology, is to use multivariate models, which describe the temporal evolution of the process simultaneously in a number of points. The same method can be directly used to model more than one cross-

correlated hydrological processes (e.g. rainfall and runoff) at the same location simultaneously.

In the following sections we give some characteristic examples of models that respect these peculiarities and together can deal with a large spectrum of problems in stochastic hydrologic simulation. These are:

1. The multivariate periodic autoregressive model of order 1 (PAR(1)), which reproduces seasonality and skewness but not long-term persistence.
2. A generalized multivariate stationary model that reproduces all kinds of persistence and simultaneously skewness but not seasonality.
3. A combination of the previous two cases in a multivariate disaggregation framework that can respect almost all the above listed peculiarities, with the exception of intermittency whose handling by such a type of models may be not easy.
4. The Bartlett-Lewis process that is appropriate to model rainfall, with emphasis on its intermittent character, at fine time scale, but at a single point basis only.

The multivariate PAR(1) model

Let $\mathbf{X}_s := [X_s^1, X_s^2, \ldots, X_s^n]^T$ represent a hydrologic process at a sub-annual (e.g. monthly) time scale (δ) and at n locations (the subscript T denotes the transpose of a vector or matrix). The PAR(1) model is similar to the AR(1) model but with periodically varying parameters. In multivariate setting, it is expressed by

$$\mathbf{X}_s = \mathbf{a}_s \mathbf{X}_{s-1} + \mathbf{b}_s \mathbf{V}_s \qquad (`2)$$

where $\mathbf{a}_s$ and $\mathbf{b}_s$ are $(n \times n)$ matrices of parameters and $\mathbf{V}_s$ is a vector of innovations (independent, both in time and location, random variables) with size n. The time index s can take any integer value but the parameters $\mathbf{a}_s$ and $\mathbf{b}_s$ are periodic functions of s with period $k :=$ 1 year / δ (e.g. 12 if δ is one month). This model can reproduce the following set of statistics:

1. the mean values, i.e., the k vectors $\boldsymbol{\mu}_s := E[\mathbf{X}_s]$ of size n each;

2. the variances and lag-zero cross-covariances among different locations, i.e. the k matrices $\boldsymbol{\sigma}_{ss} := \mathrm{Cov}[\mathbf{X}_s, \mathbf{X}_s] = E[(\mathbf{X}_s - \boldsymbol{\mu}_s)(\mathbf{X}_s - \boldsymbol{\mu}_s)^T]$ (with Cov[ ] denoting covariance), of size $(n \times n)$ each;
3. the lag-one auto-covariances at each location, i.e. the k vectors $\boldsymbol{\gamma}_{s,1} := [\gamma_{s,1}^1, \ldots, \gamma_{s,1}^n]^T$, where $\gamma_{s,\tau}^l := \mathrm{Cov}[X_s^l, X_{s-\tau}^l] = E[(X_s^l - \mu_s^l)(X_{s-\tau}^l - \mu_{s-\tau}^l)]$, of size n each (notice the notational identity $\gamma_{s,0}^l \equiv \sigma_{ss}^{ll}$);
4. the third moments, i.e., the k vectors $\boldsymbol{\xi}_s = \mu_3[\mathbf{X}_s] = E[(X_s^l - \mu_s^l)^3, l = 1, \ldots, n]^T$ of size n each (with $\mu_3[\,]$ denoting the third central moment of a random variable or random vector).

The model parameters $\mathbf{a}_s$ and $\mathbf{b}_s$ are typically determined by the moment estimators that are

$$\mathbf{a}_s = \mathrm{diag}(\gamma_{s,1}^l / \gamma_{s-1,0}^l, l = 1, \ldots, n) \tag{3}$$

$$\mathbf{b}_s \mathbf{b}_s^T = \boldsymbol{\sigma}_{ss} - \mathbf{a}_s \boldsymbol{\sigma}_{s-1,s-1} \mathbf{a}_s \tag{4}$$

These equations are extensions for the seasonal model of those for the stationary Markov model given by Matalas and Wallis (1976, p. 63). In an alternative estimation, a full (rather than a diagonal) matrix $\mathbf{a}_s$ can be derived, which enables preservation of the lag-one cross-covariances among different locations. However, the more parsimonious formulation in (3) is sufficient for most cases. The calculation of $\mathbf{b}_s$, given the product $\mathbf{b}_s \mathbf{b}_s^T$ from (4), is not a trivial issue. A generalized methodology to do this operation, also known as the extraction of the square root of a matrix, has been proposed by Koutsoyiannis (1999). Another group of model parameters are the moments of the auxiliary variables $\mathbf{V}_s$. The first moments (means) are obtained by

$$E[\mathbf{V}_s] = \mathbf{b}_s^{-1} (\boldsymbol{\mu}_s - \mathbf{a}_s \boldsymbol{\mu}_{s-1}) \tag{5}$$

The variances are by definition 1, i.e., $\mathrm{Var}[\mathbf{V}_s] = [1, \ldots, 1]^T$ and the third moments are obtained by

$$\mu_3[\mathbf{V}_s] = \left(\mathbf{b}_s^{(3)}\right)^{-1} (\boldsymbol{\xi}_s - \mathbf{a}_s^{(3)} \boldsymbol{\xi}_{s-1}) \tag{6}$$

where $\mathbf{a}_s^{(3)}$ and $\mathbf{b}_s^{(3)}$ denote the matrices whose elements are the cubes of $\mathbf{a}_s$ and $\mathbf{b}_s$, respectively.

A generalized multivariate stationary model respecting long-term persistence

The most difficult and often the most important task in simulating hydrologic processes at the annual scale is to reproduce long-term persistence. The Box-Jenkins processes are inappropriate for this purpose. Other types of models like fractional Gaussian noise (FGN) models and broken line models (whose comprehensive discussion can be found e.g. in Bras and Rodriguez-Iturbe, 1985) have several weak points such as parameter estimation problems, narrow type of autocorrelation functions that they can preserve, and their inability to reproduce skewness and simultaneously to perform in multivariate problems. In a recent paper (Koutsoyiannis, 2000), all these problems have been remedied and the proposed generalized methodology can perform in multivariate problems respecting all categories of statistics listed in points 1-4 of the previous section and, in addition, the auto-covariances at all locations for any lag r.

The setting of the method is stationary, rather than cyclostationary, so all statistics and parameters are not functions of time, which is reflected in the notation used. For example, the autocovariance for lag τ is denoted as $\boldsymbol{\gamma}_\tau := [\gamma_\tau^1, \ldots, \gamma_\tau^n]^T$ where $\gamma_\tau^l := \mathrm{Cov}[X_i^l, X_{i-\tau}^l]$. It is reminded that long-term persistence implies non-ignorable autocovariances for high lags (e.g. of the order 10^2-10^3). Such autocovariances can be described by a power-type (as opposed to the exponential type of ARMA processes) functions, like

$$\gamma_\tau^l = \gamma_0^l \, (1 + \kappa^l \, \beta^l \, \tau)^{-1/\beta^l} \tag{7}$$

where κ^l and β^l are constants. This generalized autocovariance structure (GAS) incorporates as special cases the exponential ARMA type structure (for $\beta = 0$) and the FGN structure (for a special combination of κ^l and β^l; see Koutsoyiannis, 2000). The constants κ^l and β^l can be estimated by fitting (7) to the sample autocovariance estimates; note that (7) can be used for lags beyond a certain lag τ_0, thus allowing the possibility to specify different values (i.e. the historical values of the sample autocovariance estimates) for smaller lags.

In each of the locations, the process X_i^l can be expressed in terms of some auxiliary variables V_i^l, uncorrelated in time i (i.e., Cov[V_i^l, V_m^k] = 0 if $i \neq m$) but correlated in different locations l for the same time i, by using

$$X_i^l = \sum_{r=-q}^{q} a_{|r|}^l \, V_{i+r}^l \tag{8}$$

This equation defines the so-called symmetric moving average (SMA) scheme. Like the conventional (backward) moving average (MA) process, the SMA scheme transforms a sequence of temporally uncorrelated variables V_i^l into a process with autocorrelation by taking the weighted average of a number of V_i^l. In the SMA process, the weights a_r^l are symmetric about a centre (a_0^l) that corresponds to the variable V_i^l. The number of the variables V_i^l that define X_i^l is $2q + 1$, where q is theoretically infinity but in practice can be restricted to a finite number, as the sequence of weights a_r^l tends to zero for increasing r. Koutsoyiannis (2000) showed that the discrete Fourier transform $s_a^l(\omega)$ of the a_r^l sequence is related to the power spectrum $s_\gamma^l(\omega)$ of the process (i.e. the discrete Fourier transform of the sequence of γ_τ^l) by

$$s_a^l(\omega) = \sqrt{2\, s_\gamma^l(\omega)} \tag{9}$$

This enables an easy and fast (utilizing the fast Fourier transform) computation of the sequence of a_r^l, even if the terms of the sequence are thousands. The computation includes the transformation of the sequence of γ_τ^l to $s_\gamma^l(\omega)$, the calculation of $s_a^l(\omega)$ from (9), and the inverse transformation of $s_a^l(\omega)$ to the sequence of a_r^l.

The auxiliary variables V_i^l have by definition unit variances, and means $E[V_i^l]$ and third central moments $\mu_3[V_i^l]$ given by

$$\left(a_0 + 2\sum_{j=1}^{s} a_j\right) E[V_i^l] = \mu^l, \qquad \left(a_0^3 + 2\sum_{j=1}^{q} a_j^3\right) \mu_3[V_i^l] = \xi^l \tag{10}$$

Their variance-covariance matrix $\mathbf{c} := \mathrm{Cov}[\mathbf{V}_i, \mathbf{V}_i]$ has elements c^{lk} that can be expressed in terms of σ^{lk} (the elements of the variance-covariance matrix $\boldsymbol{\sigma}$ of $\mathbf{X}_i$), and the sequences a_i^l and a_i^k by

$$c^{lk} = \sigma^{lk} \Big/ \sum_{r=-q}^{q} a_{|r|}^{l} \, a_{|r|}^{k} \tag{11}$$

Given the matrix **c**, the vector of variables $\mathbf{V}_i = [V_i^1, V_i^2, \ldots, V_i^n]^T$ can be generated using the simple multivariate model

$$\mathbf{V}_i = \mathbf{b}\, \mathbf{W}_i \tag{12}$$

where $\mathbf{W}_i = [W_i^1, W_i^2, \ldots, W_i^n]^T$ is a vector of innovations with unit variance independent both in time i and in location $l = 1, \ldots, n$, and **b** is a matrix with size $n \times n$ such that

$$\mathbf{b}\, \mathbf{b}^T = \mathbf{c} \tag{13}$$

The other parameters needed to completely define model (12) are the vector of mean values $E[\mathbf{W}]$ and third moments $\mu_3[\mathbf{W}]$ of W_i^l. These can be calculated in terms of the corresponding vectors of V_i^l, already known from (10), by

$$E[\mathbf{W}] = \mathbf{b}^{-1}\, E[\mathbf{V}], \qquad \mu_3[\mathbf{W}] = (\mathbf{b}^{(3)})^{-1}\, \mu_3[\mathbf{V}] \tag{14}$$

Stochastic disaggregation techniques

Seasonal models capable of reproducing the long-term persistence of hydrologic processes do not exist at present. If the timescale of interest is finer than annual and, simultaneously, respecting of long-term persistence is important, a two-scale approach is followed. A stationary stochastic model like the one described in previous section is used to generate the annual time series. These are then disaggregated into a finer time scale in a manner that periodicity and short-term memory of the process of interest is respected. Traditionally, the latter task has been tackled by the so-called disaggregation models, which were initially proposed by Valencia and Schaake (1973) and improved since then by the contribution of several researchers (for an outline of such contributions see Grygier and Stedinger, 1988, and Koutsoyiannis, 1992). These are purposely-designed models to generate a process at the finer time scale given that at the coarser one. Specifically, they do not model the process of interest in the lower-level time scale itself, but rather they are hybrid schemes using simultaneously both time scales. Sometimes (owing to nonlinear transformations of variables) these models

are not able to ensure consistency with the higher-level process. Then, adjusting procedures are necessary to restore consistency (e.g. Grygier and Stedinger, 1988; Koutsoyiannis and Manetas, 1996).

A different approach was recently proposed by Koutsoyiannis (2001), which is a generalized framework for coupling stochastic models of different time scales. This approach, couples two independent stochastic models appropriate respectively for the coarser (annual) and finer (e.g. monthly) scales using a transformation that modifies the output of the latter to become consistent with the series produced by the former model. To demonstrate this approach, we will assume that the coarse scale model is the multivariate SMA model described in the previous section, which produces annual series $\mathbf{Z}_i$, and the finer scale model is the multivariate PAR(1) model described two sections before, which produces monthly series $\mathbf{X}_s$. Consistency of the two series requires that they obey

$$\sum_{s=(i-1)\,k+1}^{i\,k} \mathbf{X}_s = \mathbf{Z}_i \tag{15}$$

where k is the number of fine-scale time steps within each coarse-scale time step (k = 12 in our example). The annual series $\mathbf{Z}_i$ are generated first. The finer-scale model is run independently of the coarser-scale one, without any reference to the known $\mathbf{Z}_i$, and produces monthly series $\widetilde{\mathbf{X}}_s$. If we aggregate the latter at the annual scale (by means of (15)) we will obtain some annual series $\widetilde{\mathbf{Z}}_i$, which will apparently differ from $\mathbf{Z}_i$. In a subsequent step, we modify $\widetilde{\mathbf{X}}_s$ thus producing $\mathbf{X}_s$ consistent with $\mathbf{Z}_i$ (in the sense that they obey (15)) without affecting the stochastic structure that characterizes $\widetilde{\mathbf{X}}_s$. For this modification we use a linear transformation $\mathbf{X}_s = \mathbf{f}(\widetilde{\mathbf{X}}_s, \widetilde{\mathbf{Z}}_i, \mathbf{Z}_i)$**,** which has been termed the coupling transformation. This is given by (Koutsoyiannis, 2001)

$$\mathbf{X}_i^* = \widetilde{\mathbf{X}}_i^* + \mathbf{h}\,(\mathbf{Z}_i^* - \widetilde{\mathbf{Z}}_i^*) \tag{16}$$

where

$$\mathbf{X}_i^* := \left[\mathbf{X}_{(i-1)k+1}^T, \ldots, \mathbf{X}_{i\,k}^T\right]^T \tag{17}$$

$$\mathbf{Z}_i^* := \left[\mathbf{Z}_i^T, \mathbf{Z}_{i+1}^T, \mathbf{X}_{(i-1)k}^T\right]^T \tag{18}$$

$$\mathbf{h} = \mathrm{Cov}[\mathbf{X}_i^*, \mathbf{Z}_i^*] \; \{\mathrm{Cov}[\mathbf{Z}_i^*, \mathbf{Z}_i^*]\}^{-1} \tag{19}$$

and $\widetilde{\mathbf{X}}_i^*$ and $\widetilde{\mathbf{Z}}_i^*$ are defined in terms of $\widetilde{\mathbf{X}}_s$ and $\widetilde{\mathbf{Z}}_i$ in a manner identical to that of the definition of $\mathbf{X}_i^*$ and $\mathbf{Z}_i^*$.

It is clarified that the vector $\mathbf{X}_i^*$ contains the monthly values of all 12 months of year i for all examined locations (e.g. for 5 locations, $\mathbf{X}_i^*$ contains $12 \times 5 = 60$ variables) and the vector $\mathbf{Z}_i^*$ contains (a) the annual values of the current year; (b) the annual values of the next year; and (c) the monthly values of the last month of the previous year (e.g. for 5 locations $\mathbf{Z}_i^*$ contains $3 \times 5 = 15$ variables). Items (b) and (c) of $\mathbf{Z}_i^*$ are included to assure that the transformation will preserve not only the covariance properties among the monthly values of each year, but the covariances with the previous and next years as well. Note that at the stage of the generation at year i the monthly values of year $i - 1$ are known (therefore, in $\mathbf{Z}_i^*$ we enter monthly values of the year $i - 1$) but the monthly values of year $i + 1$ are not known (therefore, in $\mathbf{Z}_i^*$ we enter annual values of the year $i + 1$, which are known).

The quantity $\mathbf{h}\,(\mathbf{Z}_i^* - \widetilde{\mathbf{Z}}_i^*)$ in (16) represents the correction applied to $\widetilde{\mathbf{X}}$ to obtain $\mathbf{X}$. Whatever the value of this correction is, the coupling transformation will ensure preservation of first and second order properties of variables (means and variance-covariance matrix) and linear relationships among them (in our case the additive property (15)). However, it is desirable to have this correction as small as possible in order for the transformation not to affect seriously other properties of the simulated processes (e.g. the skewness). It is possible to make the correction small enough, if we keep repeating the generation process for the variables of each period (rather than performing a single generation only) until a measure of the correction becomes lower that an accepted limit. This measure can be defined as

$$\Delta = (1 / m) \, \| \mathbf{Z}_i'^* - \widetilde{\mathbf{Z}}_i'^* \| \tag{20}$$

where $\mathbf{Z}_i'^*$ and $\widetilde{\mathbf{Z}}_i'^*$ are respectively $\mathbf{Z}_i^*$ and $\widetilde{\mathbf{Z}}_i^*$ standardized by standard deviation (i.e. $Z_i'^{*l} := Z_i^{*l} / \{\mathrm{Var}[Z_i^{*l}]\}^{1/2}$), m is the common size of $\mathbf{Z}_i'^*$ and $\widetilde{\mathbf{Z}}_i'^*$, and $\|.\|$ denotes the Euclidian norm.

Point process models

At even finer time scales like daily or hourly, the intermittency of hydrologic processes dominates and stochastic models like the ones described earlier can hardly describe it. Point process models have been the most widespread approach to represent intermittent hydrologic processes and particularly rainfall. As a representative example, we summarize here the rainfall model based on the Bartlett-Lewis process; this was chosen due to its wide applicability and experience in calibrating and applying it to several climates. Accumulated evidence on its ability to reproduce important features of the rainfall field from the hourly to the daily scale and above can be found in the literature (e.g. Rodriguez-Iturbe et al., 1987, 1988; Onof and Wheater, 1993). This type of model has the important feature of representing rainfall in continuous time; the statistical properties at any discrete time scale are directly obtained from those in continuous time and this enables model fitting combining statistics of different time scales.

The Bartlett-Lewis Rectangular Pulse model assumes that rainfall occurs in the form of storms of certain durations and each storm is a cluster of random cells, each having constant intensity during the time period it lasts. The general assumptions of the model are (Figure 2):

(1) Storm origins t_i occur following a Poisson process with rate λ (this means that time durations between consecutive storm origins, $t_i - t_{i-1}$, are independent identically distributed following an exponential distribution with parameter λ).
(2) Origins t_{ij} of cells of each storm i arrive following a Poisson process with rate β.
(3) Arrivals of each storm i terminate after a time v_i exponentially distributed with parameter γ.
(4) Each cell has a duration w_{ij} exponentially distributed with parameter η.
(5) Each cell has a uniform intensity X_{ij} with a specified distribution.

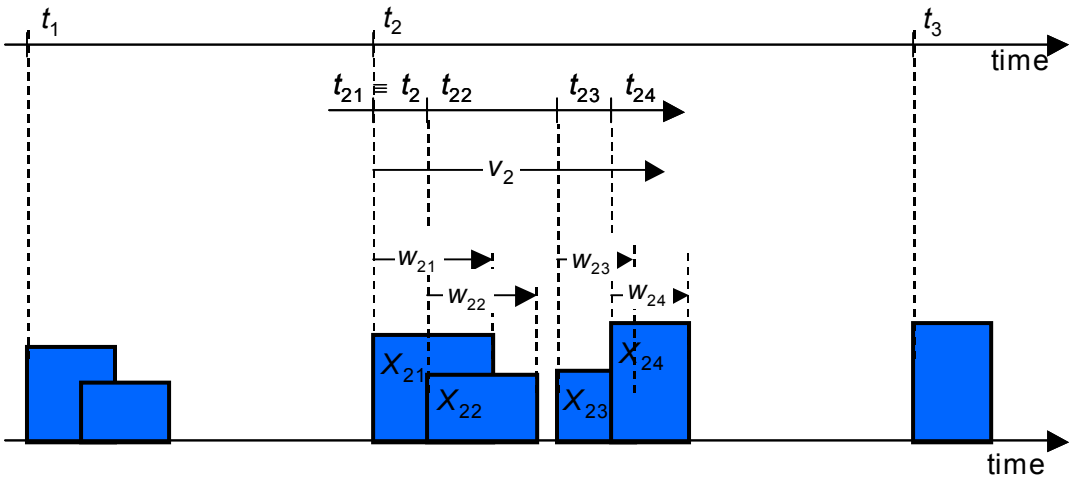

Figure 2 Explanatory sketch for the Bartlett-Lewis rectangular pulses model.

In the original version of the model, all model parameters are assumed constant. In a modified version, the parameter η is randomly varied from storm to storm with a gamma distribution with shape parameter α and scale parameter ν. Subsequently, parameters β and γ also vary so that the ratios $\kappa := \beta / \eta$ and $\varphi := \gamma / \eta$ are constant. The distribution of the uniform intensity X_{ij} is typically assumed exponential with parameter $1 / \mu_X$. Alternatively, it can be chosen as two-parameter gamma with mean μ_X and standard deviation σ_X. Thus, in its most simplified version the model uses five parameters, namely λ, β, γ, η, and μ_X and in its most enriched version seven parameters, namely λ, κ, φ, α, ν, μ_X and σ_X.

The equations of the Bartlett-Lewis model, relating the statistical properties of the rainfall process in discrete time to the model parameters, may be found in the references mentioned above. These equations serve as the basis for the model fitting.

Concluding remarks

Stochastic simulation is a powerful method, easily applicable and extremely flexible. Its main advantage is its ability to perform in complex systems describing them faithfully, without simplifying assumptions. However, it is an approximate procedure and the accuracy of its results depends on the sample size. In addition, it is a slow procedure, as the estimation error decreases inversely proportional to the square root of the simulation length (i.e. for half error we need four times greater simulation length). Today, this is not a major problem as the

progress in computer technology makes attainable even a vast simulation length in reasonable computer time.

In addition to the estimation error due to a finite simulation length, another significant source of uncertainty is always the limited historical records (based on hydrological measurements), which are used to fit probabilistic or stochastic models. This source of uncertainty, which concerns not only the simulation method but also any method, including an analytical one, is forgotten sometimes, so the following points of caution should be stressed:

- The choice of a particular stochastic model and the estimation of its parameters are always based on the available historical records, which are the only authentic source of information.
- The simulated (synthetic) hydrologic records do not replace the historical records.
- The generation of a synthetic record (with a length usually a multiple of that of historical record) does not add any information nor does it extend the historical record length.

In conclusion, the following points should be added:

- In problems that can be solved analytically (like in the example of the design of dykes discussed earlier), stochastic simulation is not the preferable method.
- Stochastic simulation becomes a powerful numerical method when a complex system is to be studied, and analytical (or other numerical) methods are not applicable or are very difficult or require oversimplifying assumptions for the system.

Acknowledgment

The review by the editor and the comments by A. Efstratiadis are gratefully appreciated.

References

Barnes, F. B., Storage required for a city water supply, *J. Inst. Eng. Australia*, 26(9) 198-203, 1954.

Beard, L. R., Use of interrelated records to simulate streamflow, *Proc. ASCE, J. Hydraul. Div.*, 91(HY5), 13-22, 1965.

Box, G. E. P., and G. M. Jenkins, *Time Series analysis, Forecasting and Control*, Holden Day, 1970.

Bras, R.L. and I. Rodriguez-Iturbe, *Random functions in hydrology*, Addison-Wesley, 1985.

Chow V.T., Maidement D.R and Mays L.W., *Applied Hydrology*. McGraw-Hill, New York, 1988.

Eckhardt, R., Stan Ulam, John von Neumann and the Monte Carlo method, in *From Cardinals to Chaos*, ed. by N. G. Cooper, Cambridge University, NY., 1989.

Grygier, J. C. and J. R. Stedinger, Condensed disaggregation procedures and conservation corrections for stochastic hydrology, *Water Resour. Res.*, 24(10), 1574-1584, 1988.

Haan, C. T., *Statistical methods in hydrology*, The Iowa State University Press, Iowa, 1977.

Hazen, A., Storage to be provided in impounding reservoirs for municipal water systems, *Trans. Amer. Soc. Civ. Eng.*, 77 1539, 1914.

Hurst, H. E., Long term storage capacities of reservoirs, *Trans. ASCE*, 116, 776-808, 1951.

Koutsoyiannis, D., A nonlinear disaggregation method with a reduced parameter set for simulation of hydrologic series, *Water Resources Research*, 28(12), 3175-3191, 1992.

Koutsoyiannis, D., Optimal decomposition of covariance matrices for multivariate stochastic models in hydrology, *Water Resources Research,* 35(4), 1219-1229, 1999.

Koutsoyiannis, D., A generalized mathematical framework for stochastic simulation and forecast of hydrologic time series, *Water Resources Research,* 36(6), 1519-1533, 2000.

Koutsoyiannis, D., Coupling stochastic models of different time scales, *Water Resources Research,* 37(2), 379-392, 2001.

Koutsoyiannis, D., The Hurst phenomenon and fractional Gaussian noise made easy, *Hydrol. Sci. J.*, 47(4), 573-596, 2002.

Koutsoyiannis, D., and A. Manetas, Simple disaggregation by accurate adjusting procedures, *Water Resources Research*, 32(7), 2105-2117, 1996.

Maass, A., M. M. Hufschmidt, R. Dorfman, H. A. Thomas, Jr., S. A. Marglin, and G. M. Fair, Design of Water Resource Systems, Harvard Univerity Press, Cambridge, Mass., 1962.

Matalas, N. C., Mathematical assessment of synthetic hydrology, *Water Resour. Res.*, 3(4), 937-945, 1967.

Matalas, N. C. and J. R. Wallis, Generation of synthetic flow sequences, in *Systems Approach to Water Management*, edited by A. K. Biswas, McGraw-Hill, New York, 1976.

Metropolis, N., The beginning of the Monte Carlo method, in *From Cardinals to Chaos*, ed. by N. G. Cooper, Cambridge University, NY., 1989.

Metropolis, N. and S. M. Ulam, The Monte Carlo method, *J. Amer. Statist. Assoc.*, 44, 335-341, 1949.

Onof, C. and H. S. Wheater, Modelling of British rainfall using a Random Parameter Bartlett-Lewis Rectangular Pulse Model, *J. Hydrol.*, 149, 67-95, 1993.

Papoulis, A., *Probability and Statistics*, Prentice-Hall, London, 1990.

Ripley, B. D., *Stochastic Simulation*, Wiley, New York, 1987.

Rodriguez-Iturbe, D. R. Cox, and V. Isham, Some models for rainfall based on stochastic point processes, Proc. R. Soc. Lond., A 410, 269-298, 1987.

Rodriguez-Iturbe, D. R. Cox, and V. Isham, A point process model for rainfall: Further developments, Proc. R. Soc. Lond., A 417, 283-298, 1988.

Thomas, H. A., and M. B. Fiering, Mathematical synthesis of streamflow sequences for the analysis of river basins by simulation, in *Design of Water Resource Systems*, by A. Maass, M. M. Hufschmidt, R. Dorfman, H. A. Thomas, Jr., S. A. Marglin, and G. M. Fair, Harvard Univerity Press, Cambridge, Mass., 1962.

Todini, E., The preservation of skewness in linear disaggregation schemes, *J. Hydrol*., 47, 199-214, 1980.

Valencia, D. and J. C. Schaake, Disaggregation processes in Stochastic Hydrology, *Water Resour. Res.*, 9(3) 211-219, 1973.

Waymire, E., and V. K. Gupta, The mathematical structure of rainfall representations, 2, A review of theory of point processes, *Water Resour. Res*. 20(10), 1453-1465.

Waymire, E., V. K. Gupta and I. Rodriguez-Iturbe, A spectral theory or rainfall intensity at the meso-β scale, *Water Resour. Res*. 17(5), 1273-1985.

Winston, W, L., *Operations Research, Applications and Algorithms*, 3rd ed., Duxbury, Belmont, 1994.