INPUT SPECIFICATIONS TO A STOCHASTIC DECISION MODEL

D.M. Clainos

Systems and Industrial Engineering Department
University of Arizona

L. Duckstein

Systems and Industrial Engineering Department
University of Arizona

T.G. Roefs

Hydrology and Water Resources Department
University of Arizona

ABSTRACT

The use of discrete conditional dependency matrices as input to stochastic decision models is examined. Some of the problems and initial assumptions involved with the construction of the above mentioned matrices are discussed. Covered in considerable detail is the transform used to relate the gamma space with the normal space. A new transform is introduced that should produce reasonable results when the record of streamflow (data) has a highly skewed distribution. Finally, the possibility of using the matrices to provide realistic inputs to a stochastic dynamic program is discussed.

Introduction

Much progress has taken place in the past decade in devising new and improved methods of determining optimal operating rules for reservoirs. By optimal operating rules for reservoirs we mean those decisions concerning the amount of water to be released during specified time periods that maximizes the utility of the releasedwater. Several different techniques have been proposed and used to model the operation of reservoirs. The most basic techniques are 1) deterministic linear programming, 2) deterministic dynamic programming, 3) stochastic linear programming, 4) stochastic dynamic programming, and 5) simulation. The purpose of this study is not to propose a new technique, but to suggest a type of input that could be used with a stochastic dynamic programming model. The type of input that is proposed is a monthly conditional probability dependency matrix that represents the probability of a certain flow occuring during the current month based on the flow that occured during the previous month. The conditional probability dependency matrix for the pair of months of January and February are graphically shown in Figure 1.

A typical matrix would consist of mxn conditional probabilities, P_{ij}, where i = 1,m and j = 1,n with the constraint that $\sum_{j=1}^{n} P_{ij} = 1$. Information could be contained from the matrix as follows: assuming we are presently in the month of February, we would know the magnitude of the flows observed

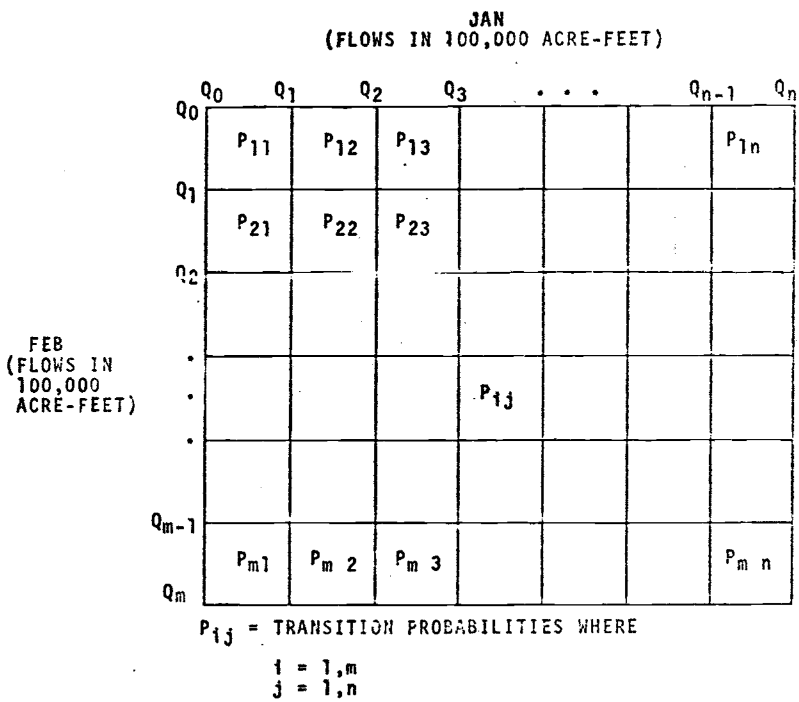


Figure 1. Conditional Probability Dependency Matrix for January-February

in January. Let us say then, for example, that for a given flow in January between Q_1 and Q_2, we might want to know the probability of experiencing a flow in February between Q_1 and Q_2. In this case, the probability that is described is P_{22}.

After the conditional probability dependency matrix is constructed for the pair of months of January and February, a similar matrix would be constructed for the pair of months of February-March. Following February-March, a matrix would be created for March-April, and so on until twelve sets of matrices are created. These tweleve sets of conditional probability dependency matrices would comprise the input to a stochastic dynamic programming model.

Loucks [1969] and Butcher [1971] have considered the use of conditional probability dependency matrices as input to stochastic dynamic programming models. Butcher used sets of matrices as input to a stochastic dynamic program in order to obtain optimal operating rules for the Watasheamu Dam, which is near the California-Nevada state line. Butcher reported that the system he studied was ergodic in nature since the final state of the system was independent of the starting state. In other words, his report is equivalent to stating that regardless what the state of the reservoir is at tht start of the computations, the steady state of the system will be independent of that starting state. This is an important observation since it is the basic assumption that Butcher made in order to determine that an optimal policy can be determined in the first place.

It was also noted by Butcher that the probabilities of the various flows in a river were dependent on the flows in the previous time period. Hence the probability of being in a given state following another given state is a fixed value. This sequence of events, or course, is known as a Markov chain and the conditional probabilities, in this case, are stationary. Hence the conditional probability dependency matrices that this study is concerned with are stationary Markov chains.

The sets of matrices that are constructed in this study are similar to those used by Butcher. However, the matrices constructed herein are more precise than any others constructed, to the best of the writers knowledge, for the purpose of obtaining more precise results.

Streamflow Synthesis Technique Selected

The stochastic synthesis technique that was selected for the purposes of this study was Beard's method [Monthly Streamflow Simulation, 1967]. This method was not selected because it was necessarily the best technique available, but because it is at least a reasonable method and has been shown to work. The reader should bear in mind that the "best" streamflow synthesis technique is mainly in the eyes of the beholder. It is extremely difficult, if not impossible, for one person to select the "best" data generation method and have all of his colleagues concur with his decision. After all, the problems that practitioners of stochastic hydrology attempt to solve are usually so complex that it is difficult to pick the best technique based on the results. At any rate, it is felt that Beard's method is satisfactory, since it had been used by the U.S. Government and the results, for practical purposes, were satisfactory [Monthly Streamflow Simulation, 1967]. Therefore, considering several factors, Beard's method is best suited for the purposes of this study. There are, however, some problems that were encountered when Beard's technique was used. The two problems are 1) the logarithmic transformation that is used to "smooth" the original streamflows, and 2) the Wilson-Hilferty transform that is used to transform gamma deviates to normal deviates and vice-versa. With respect to the first problem, Weber and Hawkins [1971] noted that the format, or organization, of the data is important when a logarithmic transformation is performed. If the data are not in a rectangular hyperbolic form, then Weber and Hawkins indicate that a logarithmic transformation will introduce a significant amount of error. The degree of error will depend on how closely the organization of the data adheres to the rectangular hyperbolic form. Beard's technique does not attempt to determine how closely the original data compares with a rectangular hyperbolic form. Unfortunately, to the best of our knowledge, no technique that utilizes a logarithmic transformation considers this particular problem. Hence, until implementable methods of working in the non-transformed domain are found, we have to resort to such methods as used here to solve practical reservoir operation planning problems.

The second problem encountered with Beard's technique was the use of the Wilson-Hilferty transform. This problem will be discussed in detail later in this paper.

Background Information

The data that was used in this study was supplied by the Hydrologic Engineering Center, U.S. Aruy Corps of Engineers at Sacramento, California. The average monthly streamflows, that constitutes the data, were observed at the Sacramento River in the vicinity of the Shasta Dam, which is in Northern California. The length of record is forty-eight years of continuous observations, commencing with 1928.

Assumptions

In order to proceed with this study, a number of assumptions were made. The first assumption was that the statistics, specifically the mean, standard deviation, coefficient of skewness, and the first order serial correlation coefficient, that were generated by Beard's routine are correct.

In the case of the mean and the variance, the assumption is completely justified. However, the assumption that the skew coefficient is correct may be questionable to some practioners of stochastic hydrology. Since the skew coefficient is a function of the third moment about the mean, many researchers believe that any measure of skewness is unreliable. According to Matalas and Benson [1968], however, the error related to the coefficient of skewness can be determined. They acknowledge that for a finite number N of observations the true value of the coefficient of skewness $\sqrt{\beta_1}$ is unknown, but that an estimate of this coefficient, denoted by $\sqrt{b_1}$, can be determined. The estimate of $\sqrt{\beta_1}$ is defined as

$$\sqrt{b_1} = \left(\frac{\sum_{i=1}^{n} (x_i - \bar{x})^3}{N}\right)\left(\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{(N - 1)}\right)^{-3/2}$$

where $\bar{x}$ is the sample mean for the hydrological variable x. To determine how good an estimate of the coefficient of skewness $\sqrt{b_1}$ is, Fisher, [1931], showed that for samples obtained from a normal distribution the standard error of $\sqrt{b_1}$ is

$$\sigma_1 = \left(\frac{6N(N - 1)}{(N - 2)(N + 1)(N + 3)}\right)^{1/2}$$

where σ_1 is independent of the central moments of x. Matalas and Benson [1968] suggested that although the above equation is based on assumption of normality, it may still be used as an indicator of the magnitude of the standard error. As far as this study is concerned, there are 48 observations (monthly flow values for forty-eight years of record). Substituting the value of N into the above equation, the standard error is approximately equal to 0.34. Since we have some measure of the magnitude of the standard error of the estimate of skewness, and we consider this error to be acceptable, it is felt that this assumption is supportable.

The assumption that the monthly serial correlation coefficient is correct may also be questionable to some since it was computed by considering only the preceding monthly flows, not all of the annual flows. However, it was again Matalas [1967], who showed that even though all preceding flows could have an influence on streamflows in a given month due to recession, a good estimate of the phenomena involved is given if it is assumed that the streamflows are serially correlated with a lag of one. It is in view of Matalas' research that it is assumed that the first order serial correlation coefficient is correct.

The second assumption that was made in this study was that the base method for the observed flow frequencies is the log-Pearson Type III distribution. This assumption was prompted by the Water Resources Council, which consists of technical staff members of the Federal departments represented in the Council and of the Tennessee Valley Authority. In 1966, the Chairman of the Water Resources Council, Stewart L. Udall, urged the adoption of the log-Pearson Type III distribution as a base method for flow frequencies for two reasons: 1) the log-Pearson Type III distribution is a sufficiently accurate representation of flow distributions, and 2) it is felt that a uniform technique for determining flow frequencies would be desirable for research agencies.

The third assumption is related to the type of reservoir that is used in this study. Beard, in his streamflow simulation routine, assumed a multi-station, multi-reservoir system. For the purposes of this study, a single-station, single reservoir system has been assumed. Since a single-station set-up has been assumed the formulas that Beard presented will be changed to some extent. The formulas, in their somewhat different form, will be discussed next.

Construction of the Conditional Probability Matrices

Beard's routine essentially consists of two parts: The statistical analysis portion and the synthetic streamflow generation portion. Since the assumption was made that only a single-station, single reservoir system would be considered in this study, Beard's procedure, which assumed a multi-station, multi-reservoir system, would be altered as follows:

1) add a small increment (one percent of the average monthly flow) to each streamflow to preclude negative infinite logarithms;
2) take the common logarithm of the result of step 1;
3) for each monthly logarithmic vector, compute the mean, standard deviation, and skew factor; the skew factor is defined as:

$$g = \frac{N \sum_{i=1}^{N} (X_i - \bar{X})^3}{(N-1)(N-2)S^3}$$

where X_i is the result of step 2, S is the standard deviation, X is the mean, and N is the number of observations;

4) the standardized gamma deviate y is then computed by

$$y = \frac{X_i - \bar{X}}{S}$$

5) the result of step 4 is then transformed to a normal deviate K using the Wilson-Hilferty transform as shown below:

$$K = \frac{6}{g}\{[(\frac{gy}{2}) + 1]^{1/3} - 1\} + \frac{g}{6}$$

where g and y are as previously defined:

6) the correlation coefficient r is then computed between the normal deviates for adjacent months:

) a uniform (0,1) random number is generated;

8) the uniform random number is transformed to a standard normal deviate;

9) a stochastic process is generated by using the following relationship

$$K_{t+1} = rK_t + (1 - r^2)^{1/2} \; Z$$

where K_t and K_{t+1} are correlated normal deviates at time t and t+1, respectively; K_1 is given K_t, where t<1 is undefined, and Z is the random normal deviate;

10) the inverse of the transform of step 5

$$y = \{[(\frac{g}{6})(K - \frac{g}{6}) + 1]^3 - 1\}\frac{2}{g}$$

is used to produce a standardized gamma deviate from a standardized normal deviate;

11) the standardized gamma deviate y is then multiplied by the standard deviation S and added to the sample mean $\bar{X}$ as shown below:

$$X = yS + \bar{X}$$

12) the antilogarithm of X is taken;

13) the small increment that was added in step 1 is subtracted.

Steps 1 through 6 comprise the analysis portion of the routine and steps 7 through 13 define the generation portion of the routine. The generation portion of the routine might, when the routine is used in a simulation mode, be repeated several hundred or thousand times. However, what concerns one here is not the use of this routine in a simulation mode, but construction of a conditional dependency matrix.

The construction of a discrete conditional probability dependency matrix between streamflows, not gamma or normal deviates, is desired for each pair of sequential months (e.g., Jan-Feb., Feb-Mar., Mar.-Apr., etc). Therefore, using the definition of conditional probability, the conditional probability $P(A|B)$ is desired where

$$P(A|B) = P(A,B)/P(B)$$

To solve the above equation for the conditional probability, one must first determine the unconditional probability of the January flows and the joint probability of the January and February flows.

Computation of the Unconditional Probabilities

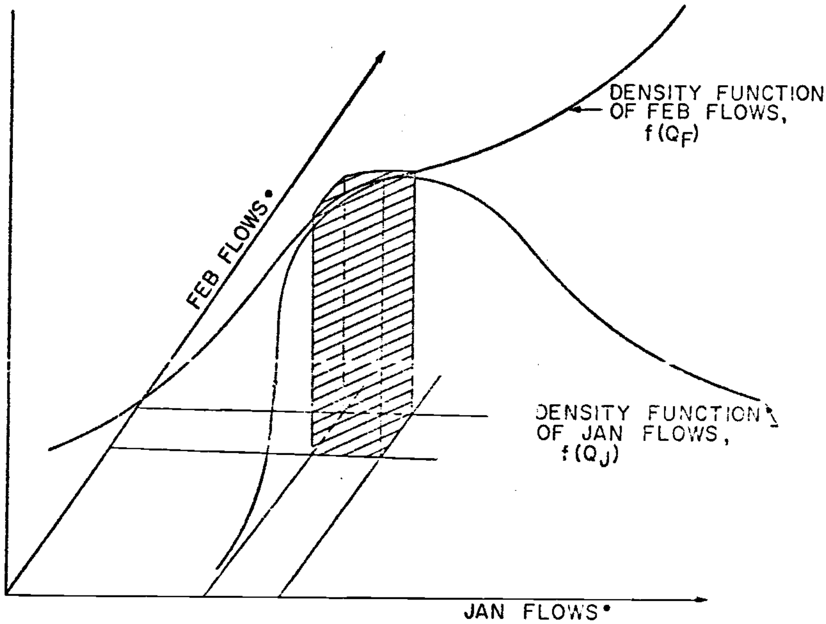
The computation of the unconditional probabilities of the prior monthly streamflows is straightforward. A procedure that could be used to determine them is as follows:

1) take the common logarithm of each streamflow in question;
2) determine the standardized gamma deviate of the value in step 1;
3) transform the standardized gamma deviate of step 2 to a standardized normal deviate;

4) then using a table of normal probability function or a numerical integration scheme and a digital computer, determine the unconditional probability of each streamflow in question.

Computation of the Joint Probabilities

Determining the joint probability, however, is not so straightforward. The joint probability is expressed by means of bivariate normal distribution. Graphically , the joint probability is represented by the shaded volume as shown in Figure 2. Of course, one could obtain tables of the bivariate normal distribution and obtain the joint probabilities if the flow intervals were long enough. But, the flow interval size selected for the study was 100,000 acre-feet, which is too small for a bivariate normal table to be of help. Unfortunately, no more accurate tables could be found. Another method for computing the joint probability, as shown in Figure 2, is to perform double numerical integration. However, considering the practical side of the problem, if a 50x50 dependency matrix were desired, it would mean that 2500 conditional probabilities, P_{ij}, would have to be computed for each pair of months. Since there are tweleve pairs of months, a total of 30,000 conditional probabilities would have to be computed. If it took five seconds to perform each double numerical integration on a digital computer, one can readily see that approximately, forty-two hours of computer time would be required. Obviously, an analysis of numerical methods must indicate another method of determining the conditional probabilities that preserves accuracy but reduces the computational time. Roefs and Clainos



VOLUME = JOINT PROBABILITY OF JAN AND FEB FLOWS

VOLUME = $P(Q_J, Q_F) = \iint f(Q_J)\, f(Q_F)\, dQ_J\, dQ_F$

* ALL FLOWS ARE EXPRESSED AS STANDARD NORMAL DEVIATES.

Figure 2. Joint Probability Density Function $P(Q_J, Q_F)$

[1971] determined that if the flow interval of the previous month, January in this case, were sufficiently small it would be represented by a single point. More specifically, the interval could be more accurately represented by the midpoint of the interval than any other point. If this numerical approximation is made, the computation is greatly reduced since only single numerical integration is required. But what about the accuracy of this method? Roefs and Clainos [1971] showed that if the interval that is represented as a point were sufficiently small then all of the probabilities of the same class would sum to unity. Expressed in more familiar form we have

$$\sum_{j=1}^{n} P_{ij} = 1 \qquad \text{as } \underset{i=1,2,\ldots,m}{n \to \infty}$$

It was determined by an iterative procedure that an interval of 100,000 acre-feet was sufficiently small to represent an interval as a point. The reader should remember, however, that the size of an interval will depende on the magnitude of monthly streamflows; there is nothing sacred about 100,000 acre-feet.

Now that both the unconditional and joint probabilities can be determined, the conditional probabilities can be computed. A computer program was devised for this purpose; it will be discussed later. The conditional probability dependency matrices were computed for twelve pairs of months. The error for each computation of a conditional probability ranged from zero to approximately twelve percent.

Throughout the preceding paragraphs, the months of January and February have been used as examples. The same procedure is used for the other twelve pairs of months.

The Problem with Transforms

In Beard's routine, the Wilson-Hilferty transform is used to transform a Pearson Type III deviate to a normal deviate, and vice-versa for the synthetic streamflow generation portion. This particular transformation is possibly the weakest part of Beard's routine. At best the transformation is an approximation. McGinnis and Sammons [1970] studied the Wilson-Hilferty transform in detail and evaluated, along with other transforms, its accuracy with respect to the coefficient of skewness that was used. Of the four transformations that were evaluated, the Fisher transform could be eliminated immediately since it produced values that were obviously inferior to the remaining transforms. The three remaining transforms were 1) the Fisher-Cornish transform, 2) the Wilson-Hilferty transform, and 3) the Severo-Zelen transform. A comparison of these three transforms evaluated when the coefficient of skewness is 1.0 or less shows no significant difference. As the coefficient of skewness is increased to 2.0, the Severo-Zelen transform seems to perform the best (relative error of approximately 0.25 percent), the Wilson-Hilferty transform was second (relative error of approximately 2.19 percent), and the Fisher-Cornish transform was third (relative error approximately 19.1 percent). However, McGinnis and Sammons point out that as the coefficient of skewness increases beyond 2.0, all of the above mentioned transforms seem to produce poor estimates of the upper and lower tails of distribution. It should be made clear, however, that at low levels of skewness $g<1.0$ there does not seem to be any significant error (relative error <.01 percent). Therefore, the question is now what does one do when the coefficient of skewness exceeds 2.0?

At this time, one should acknowledge that the above mentioned decision is not purely technical in nature, but also managerial. For example, McGinnis and Sammons [1970] gathered information on seventy-four watersheds in the state of Pennsylvania that ranged in size between 2 square miles and 200 square miles and in length of record from 10 years to 59 years. Of these, two thirds of the watersheds ahd coefficients of skewness between -0.5 and +0.5 and only seven had coefficients of skewness outside of the range of -2.0 to +2.0. Hence the water resources manager, the same as the business manager must determine what his return-on-investment will be if he elects to create a mew method for the specific prrpose of solving a small fraction of the total, or he must determine the penalty for not creating that special method. His alternatives are three-fold. First, he can do nothing and accept the errors introduced into the conditional probabilities, P_{ij}. Second, he can establish some arbitrary cut-off level of skewness; for example $g = +1.8$. This cut-off level of skewness was determined subjectively. Since McGinnis and Sammons [1970] showed that a significant amount of error was introduced by the Wilson-Hilferty transform when skewnedd exceeded 2.0, the cut-off level of skewness was arbitrarily set at 1.8. Using this method whenever the coefficient of skewness exceeded +1.8, it would automatically be truncated to + 1.8, depending on if the distribution is positively or negatively skewed. Finally, a new method, or transform, could be devised to handle the transformation when the coefficient of skewness exceeds 2.0. A decision must be made.

A New Transform

In this study it became necessary to investigate the possibility of devising a new transform because 1) one of the monthly coefficients of skewness exceeded 2.0, and 2) it is felt that when this new transform is perfected, or at least greatly improved, it will be able to transform variables from the Pearson Type III distribution to the normal distribution, and vice-versa, under highly skewed conditions. Such a transform would certainly aid water resources planners in an arid-land environment such as Arizona, where the rivers and streams are characteristically highly skewed.

Although this transform is a new technique in the sense that is was devised by this writer the idea was provided by Harter [1969] who developed the table of percentage points to the Pearson Type III distribution arranged in categories of different levels of skewness. Basically, Harter computed the percentage points of the chi-square distribution, which he then modified to percentage points of the Pearson Type III distribution. One procedure that Harter used can be described as follows:

1) Compute the probability integral $P(\chi^2;\nu) = I(u,p)$ of χ^2 with ν degrees of freedom such that

$$I(u,p) = \frac{1}{\Gamma(p+1)} \int_0^{u\sqrt{p+1}} v^p e^{-v} dv$$

where $u = \chi^2 \sqrt{2\nu}$, $p = \nu/2 = 1$, and $I(u,p)$ is the incomplete gamma function ratio;

2) Obtain the percentage points of χ^2 by performing inverse interpolation in the table of the probability integral in order to obtain the percentage point of u, then multiply by $\sqrt{2\nu}$;

3) Relate the chi-square distribution with ν degrees of freedom to the Pearson Type III distribution with mean $\mu=\nu$, standard deviation $\sigma=\sqrt{2\nu}$, and skewness $\gamma = \sqrt{8/\nu}$.

The result of Harter's study is a table of percentage points to the Pearson Type III distribution for a specified level of skewness. At this point, one could build on Harter's work and construct an algorithm suited for a digital computer that would produce standard normal deviates. The procedure is as follows:

1) the gamma deviates are computed from the streamflows;

2) given the level of skewness, the gamma deviate is related with the appropriate percentage point as determined by Harter, using linear interpolation (Lagrangian interpolation could be used if additional accuracy is required).

3) by means of areas, the percentage point of the Pearson Type III distribution is then related to the percentage point of the standard normal distribution,

4) iterative numerical integration is performed until convergence (wtihin specified error bounds) is achieved;

5) the result is a normal deviate.

The question now is how well does this "new" transform work. The transform has been subjected to two tests. The first test involved actually computing conditional probabilities when the coefficient of skewness exceeded 2.0. This occurred during the pairs of months of September-October and October-November. When the Wilson-Hilferty transform was used, unrealistic conditional probabilities were produced. The probabilities that were generated appeared to be hydrologically unreasonable because a high flow in September virtually dictated that a flow two or three times higher would occur in October and the probability of lower flows was essentially nonexistent. In addition, the sequence of the conditional probabilities caused a disturbance since a small conditional probability would be followed by a slightly higher probability, only to be next followed by a probability of zero. In other words, there was no visible trend followed by the probabilities, as one might expect, but merely an erratic sequence of numbers. These errors were so pronounced that the dependency matrix that was generated was considered to possess more noise than meaningful information. However, using the new transform, the conditional probabilities were reasonable in as much as the rows approximately equalled unity and there were no pronounced inconsistencies. Although encouraging, this test certainly does not show that the "new" transform works well.

Another test was performed. The purpose of the second test was to determine how the new transform would perform when skew coefficients less than 1.0 were encountered. In this case, the new transform could be directly compared with the Wilson-Hilferty transform. Since McGinnis and Sammons [1970] determined the accuracy of the Wilson-Hilferty transform, one could determine the accuracy of the new transform by comparing the two. The pair of months of June-July was arbitrarily selected. The coefficient of skewness for the streamflows in June and July are 0.183 and 0.240, respectively. One hundred different conditional probabilities were computed by the new transform and by the Wilson-Hilferty transform. THe results were compared. The average difference between the conditional probabilities was 1.28 percent. These results show that, at least, the new transform performs as well as the Wilson-Hilferty transform. The results of the first test indicate that the new transform performs well when the skew coefficient exceeds 2.0. Although encouraging, the new transform should undergo several additional tests before it is accepted.

In this section it has been shown that the conditional probability dependency matrices can be constructed. The actual matrices that have been constructed are too voluminous to be included in this report. However, they are available [Clainos, 1972] by special request. In the next section the possibility of using these matrices to provide realistic inputs to a stochastic dynamic program will be discussed.

Application

In order to demonstrate the application of the conditional probability dependency matrices as inputs to a stochastic dynamic programming model, such a model will be formulated in this chapter. At this point, it is important to make the assumption that only a single-purpose reservoir will be considered. This is an important assumption because with a single-purpose reservoir the economic return in each time period is a function of the water released only. Before the model is formulated, however, we shall specify what subscripts will be used to identify the different time periods involved. Basically, the formulation will start at some arbitrary time period in the future. Then each set of computations (in each time period) will look forward in time but a step backwards in time will be made in order to determine the next set of computations. The two state variables, amount of reservoir storage and inflow, will be defined as s and q, respectively. The decision variable, amount of release, is defined as r. It might be helpful, at this point, to refer to Figure 3 to determine the time relation of the computations involved in the formulation.

We are presently at some arbitrary future time period. If the storage in the reservoir at this time period is s_1 and the flow that has taken place in the preceding time period is q_2, then the value of the reservoir at this time is equal to the value of the release only. At this point, one should realize that the initial value of the reservoir is necessarily set to be equal to zero. The value of the reservoir is shown mathematically as:

$$f_1(s_1,q_2) = \max_r \; [V(r)]$$

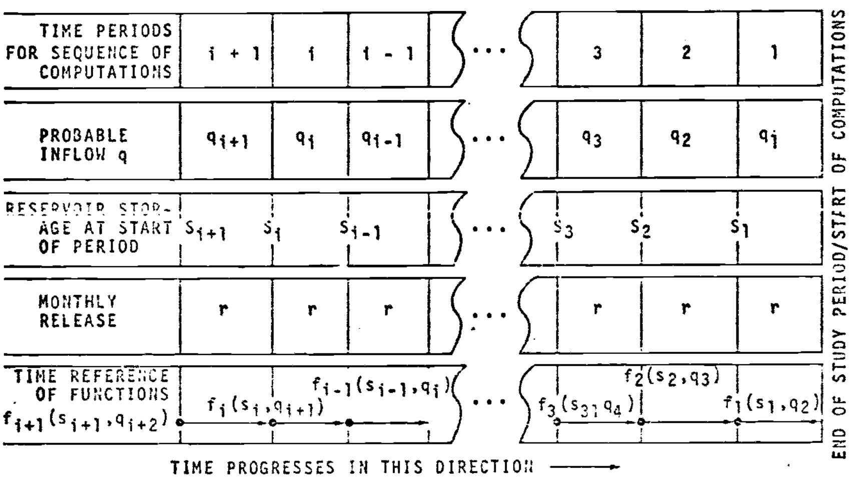


Figure 3. Time Relation of Computations

where $f_1(s_1,q_2)$ is the expected return from the optimal operation of the system in the first time period; and V(r) is the value of the release in the time period. We are now ready to take another step backwards in time. Following Bellmans' Principle of Optimality [Bellman, 1957], which is loosely defined as any optimum policy has the property that regardless of initial state or initial decision, the resulting decisions will form an optimal policy with regard to the state that resulted from the initial decision. Use will be made of one of the optimal policy determinations made for the last stage in the next stage. Stated mathematically we have,

$$f_2(s_2,q_3) = \max_{r} \left[V(r) + \sum_{q_2=0}^{\max q_2} P(q_2|q_3) \cdot f_1(s_1,q_2)\right]$$

where $\Sigma P(q_2|q_3)$ are the transition probabilities that relate the flow in the second time period, q_2, with the flow in the third time period, q_3. Also the limits of the summation of the transition probabilities range from zero to the maximum flow in the second time period. If this procedure is repeated f_3 can be determined from f_2, f_4 can be determined from f_3, and so on. This recursive procedure then can be generalized as:

$$f_i(s_i,q_{i+1}) = \max_{r} \left[V(r) + \sum_{q_i=0}^{\max q_i} P(q_i|q_{i+1}) \cdot f_{i-1}(s_{i-1},q_i)\right]$$

with the recursive relationship:

$$s_{i-1} = s_i - r + q_i$$

where $f_i(s_i,q_{i+1})$ is the expected return from the optimal operation of a system in the ith time period; s_i is the volume of storage at the start of the ith time period; q_i is the flow into the reservoir in the ith time period; and $\Sigma P(q_i | q_{i+1})$ are the transition probabilities that relate the flow in the ith time period, q_i, with the flow in the (i + 1)th time period, q_{i+1}.

Using the above described relationship, which would start in some future time period and utilize the conditional probability dependency matrices in conjunction with data from an actual reservoir, it is possible to calculate values of release for each time period as a function of the state variables, storage and previous inflow. These releases then determine the optimal operating policy of that reservoir when the policy converges. A policy is said to converge when the values of release that are used to evaluate the function $f_i(s_i, q_{i+1})$ are repeated for all values of i as i increases. In his study at the Watasheamu Dam, Butcher [1971] observed that the release policy did converge. In this case, convergence took place after 30 iterations, but Butcher computed an additional 24 iterations to insure convergence.

It is felt that if the conditional probability dependency matrices constructed in this study and actual data from a reservoir were used as input to the stochastic dynamic programming model developed in this section, one could obtain an optimal operating policy for that reservoir. This statement is based on the assumption that such a solution exists in the first place.

Conclusions

A number of conclusions can be made when creating conditional dependency matrices based on a stochastic streamflow synthesis technique. First, not all assumptions of current streamflow synthesis techniques produce results which are hydrologically reasonable. Because of the high value of the skew coefficient, the transformation from the gamma space to the normal space and back to the gamma space is not at all accurate. Hence, the result is that the conditional probabilities that are generated are clearly not accurate.

Second, double numerical integration is not computationally feasible because of the large amount of computer time required. An approximation, such as the midpoint method, is required to solve the problem of double numerical integration in a practical manner.

Third, it may be possible to use transforms efficiently when streamflows are high skewed. Although this new transform was used successfully in this study, the real value of such a transform may be realized when studying watersheds in an arid region where the coefficient of skew of the original data may be as large as 7.

Finally, it has been shown that conditional probability dependency matrices can be constructed. It has been further shown that conditional probability dependency matrices can be used as input to a stochastic dynamic programming model in order to obtain an optimal operating policy for a reservoir.

Acknowledgement. The research for this study was made possible by a research grant, #ARIZ-024, from the Office of Water Resources Research.

References

1. Bellman, Richard E., Dynamic Programming, Princeton University Press, 1957.

2. Butcher, William S., "Stochastic Dynamic Programming for Optimum Reservoir Operation", Water Resources Bulletin, Vol. 7, No. 1, February, 1971.

3. Clainos, Deme M., "Conditional Probability Dependency Matrices". Systems & Industrial Engineering Department File Report No. 972-320-G, University of ARizona, Tucson, Arizona 1972.

4. Fisher, R.A., "The Moments of the Distribution for Normal Samples of Measures of Departure from Normality", Proceedings of the Royal Society of London, Vol. 130, 1931.

5. Harter, Leon H., "A New Table of Percentage Points of the Pearson Type III Distribution", Technometrics, Vol. 11, No. 1, February 1969.

6. Loucks, D.P., "Stochastic Methods for Analyzing River Basin Systems", Research Project Technical Completion Report, OWRR project, No. C-1034, Department of Water Resources Engineering and the Water Resources and Marine Sciences Center, Cornell University, Ithica, New York, Aug. 1969.

7. McGinnis, David F., Jr. and William H. Sammons, Discussion of "Daily Streamflow Simulation", by K. Payne, W.R. Neuman and K.D. Kerri, Journal of the Hydraulics Division, Proceedings, of the A.S.C.E., Vol.96, No. HY5, May, 1970.

8. Matalas, Nicholas, C., "Time Series Analysis", Water Resources Research, Vol. 3, No. 3, Third Quarter, 1967.

9. Matalas, Nicholas C. and Manuel A. Benson, "Note on the Standard Error of the Coefficient of Skewness", Water Resources Research, Vol. 4, No. 1, February, 1968.

10. Monthly Streamflow Simulation, Computer Program 23-c*L267, Hydrologic Engineering Center, Corps of Engineers, U.S. Army, July, 1967.

11. Roefs, Theodore G. and Deme M. Clainos, "Conditional Streamflow Probability Distributions", Meeting of Arizona Section of the American Water Resources Association and the Arizona Academy of Science, Hydrology Section, Tempe, Arizona, April, 1971.

12. Weber, Jean E. and Clark A. Hawkins, "The Estimation of Constant Elasticities", The Southern Arizona Economic Journal, Vol. 38, No. 2, October, 1971.