8. Modelling Uncertainty

8 Modelling Uncertainty

Decision-makers are increasingly willing to consider the uncertainty associated with model predictions of the impacts of their possible decisions. Information on uncertainty does not make decision-making easier, but to ignore it is to ignore reality. Incorporating what is known about the uncertainty into input parameters and variables used in optimization and simulation models can help in quantifying the uncertainty in the resulting model predictions – the model output. This chapter discusses and illustrates some approaches for doing this.

1. Introduction

Water resources planners and managers work in an environment of change and uncertainty. Water supplies are always uncertain, if not in the short term at least in the long term. Water demands and the multiple purposes and services water provide are always changing, and these changes cannot always be predicted. Many of the values of parameters of models used to predict the multiple hydrological, economic, environmental, ecological and social impacts are also changing and uncertain. Indeed models used to predict these impacts are, at least in part, based on many imprecise assumptions. Planning and managing, given this uncertainty, cannot be avoided.

To the extent that probabilities can be assigned to various uncertain inputs or parameter values, some of this uncertainty can be incorporated into models. These models are called *probabilistic* or *stochastic* models. Most probabilistic models provide a range of possible values for each output variable along with their probabilities. Stochastic models attempt to model the random processes that occur over time, and provide alternative time series of outputs along with their probabilities. In other cases, sensitivity analyses (solving models under different assumptions) can be carried out to estimate the impact of any uncertainty on the decisions being considered. In some situations, uncertainty may not significantly affect the decisions that should be made. In other situations it will. Sensitivity analyses can help estimate the extent to which we need to try to reduce that uncertainty. Model sensitivity and uncertainty analysis is discussed in more detail in Chapter 9.

This chapter introduces a number of approaches to probabilistic optimization and simulation modelling. Probabilistic models will be developed and applied to some of the same water resources management problems used to illustrate deterministic modelling in previous chapters. They can be, and have been, applied to numerous other water resources planning and management problems as well. The purpose here, however, is simply to illustrate some of the commonly used approaches to the probabilistic modelling of water resources system design and operating problems.

2. Generating Values From Known Probability Distributions

Variables whose values cannot be predicted with certainty are called random variables. Often, inputs to hydrological simulation models are observed or synthetically generated values of rainfall or streamflow. Other examples of such random variables could be evaporation losses, point and non-point source wastewater discharges, demands for water, spot prices for energy that may impact the amount of hydropower production, and so on. For random processes that are stationary – that is, the statistical attributes of the process are not changing – and if there is

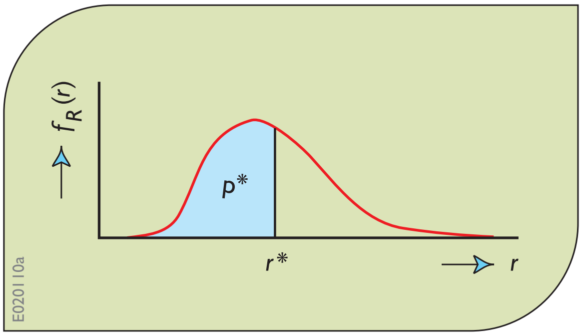

Figure 8.1. Probability density distribution of a random variable *R*. The probability that *R* is less than or equal *r** is *p**.

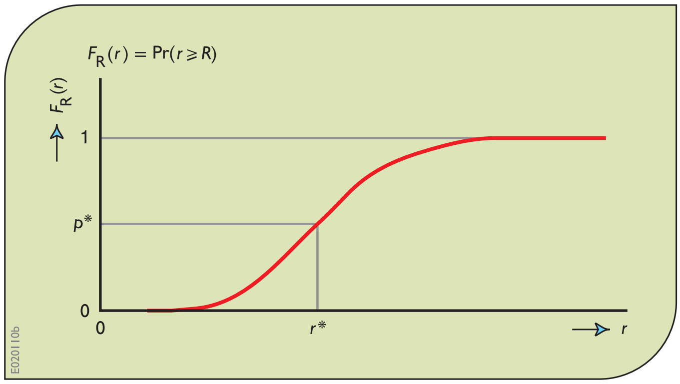

Figure 8.2. Cumulative distribution function of a random variable *R* showing the probability of any observed value of *R* being less than or equal to a given value *r*. The probability of an observed value of *R* being less than or equal to *r** is *p**.

no serial correlation in the spatial or temporal sequence of observed values, then such random processes can be characterized by single probability distributions. These probability distributions are often based on past observations of the random variables. These observations or measurements are used either to define the probability distribution itself or to estimate parameter values of an assumed type of distribution.

Let R be a random variable whose probability density distribution, $f_R(r)$, is as shown in Figure 8.1. This distribution indicates the probability or likelihood of an observed value of the random variable R being between any two values of r on the horizontal axis. For example, the probability of an observed value of R being between 0 and r^* is p^*, the shaded area to the left of r^*. The entire shaded area of a probability density distribution, such as shown in Figure 8.1, is 1.

Integrating this function over r converts the density function to a cumulative distribution function, $F_R(r^*)$, ranging from 0 to 1, as illustrated in Figure 8.2.

$$\int_0^{r^*} f_R(r)\,\mathrm{d}r = \Pr(R \leq r^*) = F_R(r^*) \tag{8.1}$$

Given any value of p^* from 0 to 1, one can find its corresponding variable value r^* from the inverse of the cumulative distribution function.

$$F_R^{-1}(p^*) = r^* \tag{8.2}$$

From the distribution shown in Figure 8.1, it is obvious that the likelihood of different values of the random variable varies; ones in the vicinity of r^* are much more likely to occur than are values at the tails of the distribution. A uniform distribution is one that looks like a rectangle; any value of the random variable between its lower and upper limits is equally likely. Using Equation 8.2, together with a series of uniformly distributed (all equally likely) values of p^* over the range from 0 to 1 (that is, along the vertical axis of Figure 8.2), one can generate a corresponding series of variable values, r^*, associated with any distribution. These random variable values will have a cumulative distribution as shown in Figure 8.2, and hence a density distribution as shown in Figure 8.1, regardless of the types or shapes of those distributions. The mean, variance and other moments of the distributions will be maintained.

The mean and variance of continuous distributions are:

$$\int r\, f_R(r)\mathrm{d}r = E[R] \tag{8.3}$$

$$\int (r - E[R])^2 f_R(r)\mathrm{d}r = \mathrm{Var}[R] \tag{8.4}$$

The mean and variance of discrete distributions having possible values denoted by r_i with probabilities p_i are:

$$\sum_i r_i p_i = E[R] \tag{8.5}$$

$$\sum_i (r_i - E[R])^2 p_i = \mathrm{Var}[R] \tag{8.6}$$

If a time series of T random variable values, r_t, from the same stationary random variable, R, exists, then the serial or autocorrelations of r_t and r_{t+k} in this time series for any positive integer k can be estimated using:

$$\hat{\rho}_R(k) = \sum_{\tau=1,T-k} [(r_\tau - E[R])(r_{\tau+k} - E[R])] \Big/ \sum_{t=1,T} (r_t - E[R])^2 \tag{8.7}$$

The probability density and corresponding cumulative probability distributions can be of any shape, not just those named distributions commonly found in probability and statistics books.

The process of generating a time sequence $t = 1, 2, \ldots$ of inputs, r_t, from the probability distribution of a random variable R where the lag 1 serial correlation, $\rho_R(1) = \rho$, is to be preserved is a little more complex. The expected value of the random variable R_{t+1} depends on the observed value, r_t, of the random variable R_t, together with the mean of the distribution, $E[R]$, and the correlation coefficient ρ. If there is no correlation (ρ is 0), then the expected value of R_{t+1} is the mean of the population, $E[R]$. If there is perfect correlation (ρ is 1), then the expected value of R_{t+1} is r_t. In general, the expected value of R_{t+1} given an observed value r_t of R_t is:

$$E[R_{t+1}|R_t = r_t] = E[R] + \rho(r_t - E[R]) \tag{8.8}$$

The variance of the random variable R_{t+1} depends on the variance of the distribution, $\text{Var}[R]$, and the lag 1 correlation coefficient, ρ.

$$\text{Var}[R_{t+1}|R_t = r_t] = \text{Var}[R](1 - \rho^2) \tag{8.9}$$

If there is perfect correlation ($\rho = 1$), then the process is deterministic, there is no variance, and $r_{t+1} = r_t$. The value for r_{t+1} is r_t. If there is no correlation – that is, serial correlation does not exist ($\rho = 0$) – then the generated value for r_{t+1} is its mean, $E[R]$, plus some randomly generated deviation from a normal distribution having a mean of 0 and a standard deviation of 1, denoted as $N(0, 1)$. In this case the value r_{t+1} is not dependent on r_t.

When the serial correlation is more than 0 but less than 1, then both the correlation and the standard deviation (the square root of the variance) influence the value of r_{t+1}. A sequence of random variable values from a multivariate normal distribution that preserves the mean, $E[R]$; overall variance, $\text{Var}[R]$; standard deviation σ, and lag 1 correlation ρ; can be obtained from.

$$r_{t+1} = E[R] + \rho(r_t - E[R]) + Z\,\sigma(1 - \rho^2)^{1/2} \tag{8.10}$$

The term Z in Equation 8.10 is a random number generated from a normal distribution having a mean of 0 and a variance of 1. The process involves selecting a random number from a uniform distribution ranging from 0 to 1, and using it in Equation 8.2 for an $N(0, 1)$ distribution to obtain a value of random number for use in Equation 8.10. This positive or negative number is substituted for the term Z in Equation 8.10 to obtain a value r_{t+1}. This is shown on the graph in Figure 8.3.

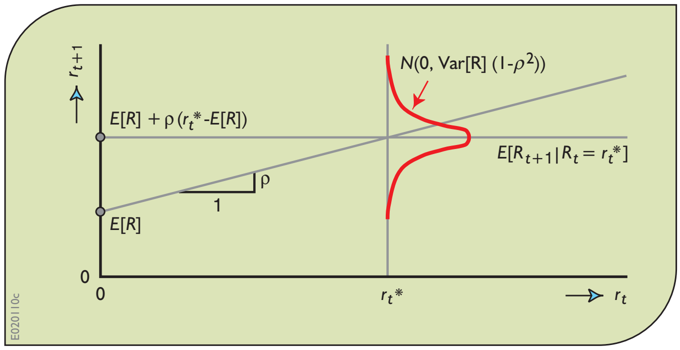

Figure 8.3. Diagram showing the calculation of a sequence of values of the random variable R from a multivariate normal distribution in a way that preserves the mean, variance and correlation of the random variable.

Simulation models that have random inputs, such as a series of r_t values, will generally produce random outputs. After many simulations, the probability distributions of each random output variable value can be defined. These can be used to estimate reliabilities and other statistical characteristics of those output distributions. This process of generating multiple random inputs for multiple simulations to obtain multiple random outputs is called Monte Carlo simulation.

3. Monte Carlo Simulation

To illustrate Monte Carlo simulation, consider the water allocation problem involving three firms, each of which receives a benefit, $B_i(x_{it})$, from the amount of water, x_{it}, allocated to it in each period t. This situation is shown in Figure 8.4. Monte Carlo simulation can be used to find the probability distribution of the benefits to each firm associated with the firm's allocation policy.

Suppose the policy is to keep the first two units of flow in the stream, to allocate the next three units to Firm 3, and the next four units to Firms 1 and 2 equally. The remaining flow is to be allocated to each of the three firms equally up to the limits desired by each firm, namely 3.0, 2.33, and 8.0 respectively. Any excess flow will remain in the stream. The plots in Figure 8.5 illustrate this policy. Each allocation plot reflects the priorities given to the three firms and the users further downstream.

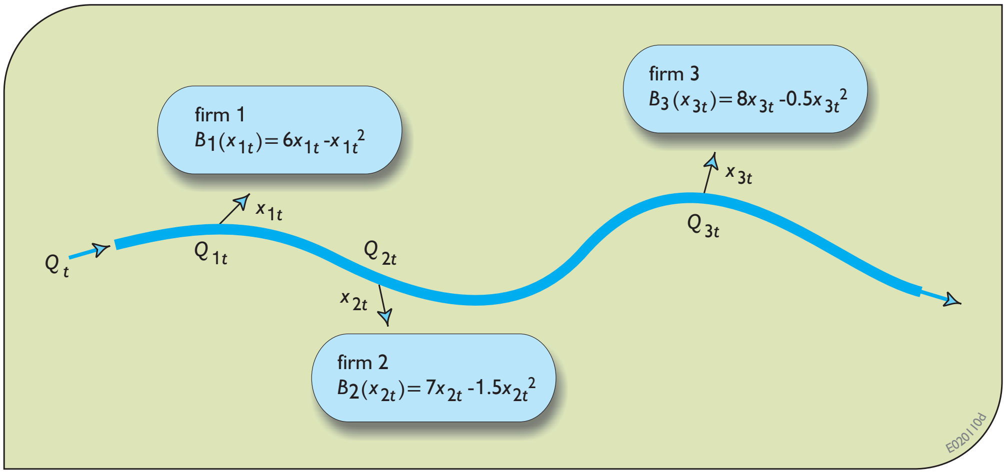

Figure 8.4. Streamflow allocations in each period t result in benefits, $B_i(x_{it})$, to each firm i. The flows, Q_{it}, at each diversion site i are the random flows Q_t less the upstream withdrawals, if any.

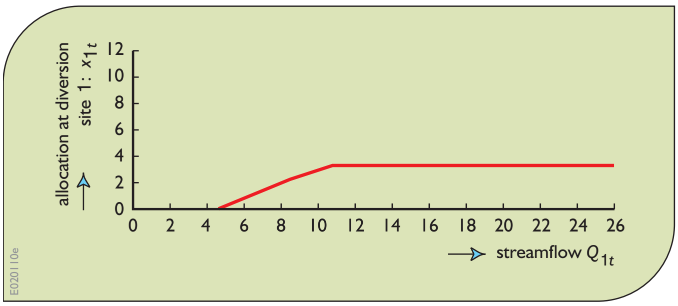

Figure 8.5a. Water allocation policy for Firm 1 based on the flow at its diversion site. This policy applies for each period t.

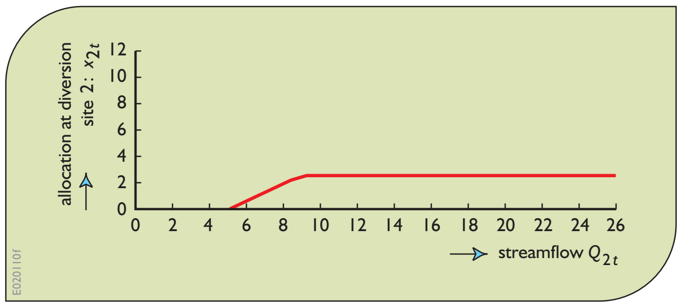

Figure 8.5b. Water allocation policy for Firm 2 based on the flow at its diversion site for that firm. This policy applies for each period t.

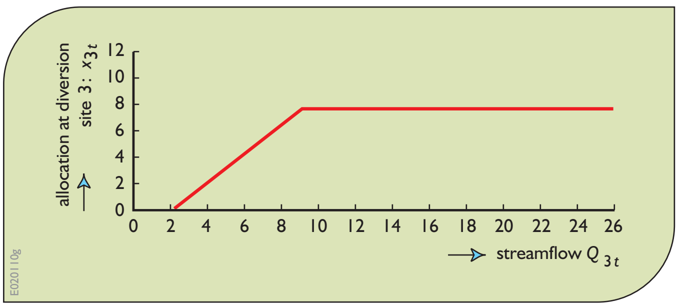

Figure 8.5c. Water allocation policy for Firm 3 based on the flow at its diversion site. This policy applies for each period t.

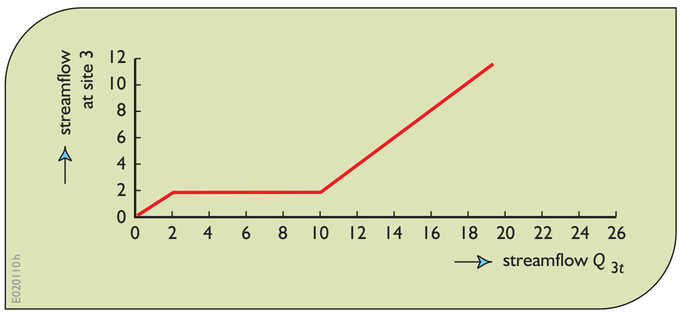

Figure 8.5d. Streamflow downstream of site 3 given the streamflow Q_{3t} at site 3 before the diversion. This applies for each period t.

A simulation model can be created. In each of a series of discrete time periods t, the flows Q_t are drawn from a probability distribution, such as from Figure 8.2 using Equation 8.2. Once this flow is determined, each successive allocation, x_{it}, is computed. Once an allocation is made it is subtracted from the streamflow and the next allocation is made on the basis of that reduced streamflow, in accordance with the allocation policy defined in Figures 8.5a – d. After numerous time steps, the probability distributions of the allocations to each of the firms can be defined.

Figure 8.6 shows a flow chart for this simulation model.

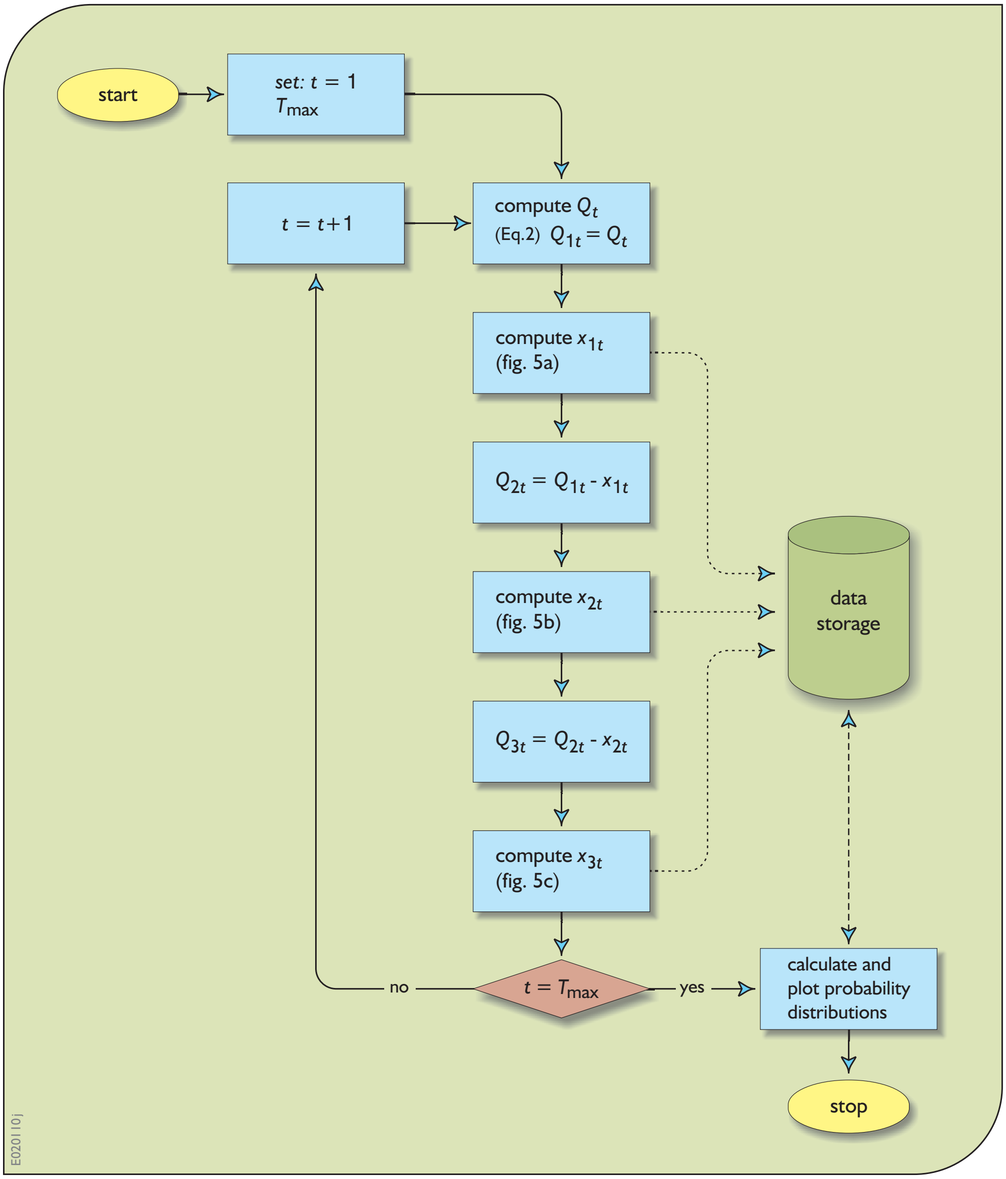

Figure 8.6. Monte Carlo simulation to determine probability distributions of allocations to each of three water users, as illustrated in Figure 8.4. The dashed lines represent information (data) flows.

Having defined the probability distribution of the allocations, based on the allocation policy, one can consider each of the allocations as random variables, X_1, X_2, and X_3 for Firms 1, 2 and 3 respectively.

4. Chance Constrained Models

For models that include random variables, it may be appropriate in some situations to consider constraints that do not have to be satisfied all the time. Chance constraints specify the probability of a constraint being satisfied, or the fraction of the time a constraint has to apply. Consider, for example, the allocation problem shown in Figure 8.4. For planning purposes, the three firms may want to set allocation targets, not expecting to have those targets met 100% of the time. To ensure, for example, that an allocation target, T_i, of firm i will be met at least 90% of the time, one could write the chance constraint:

$$\Pr\{T_i \leq X_i\} \geq 0.90 \quad i = 1, 2 \text{ and } 3 \qquad (8.11)$$

In this constraint, the allocation target T_i is an unknown decision-variable, and X_i is a random variable whose distribution has just been computed and is known.

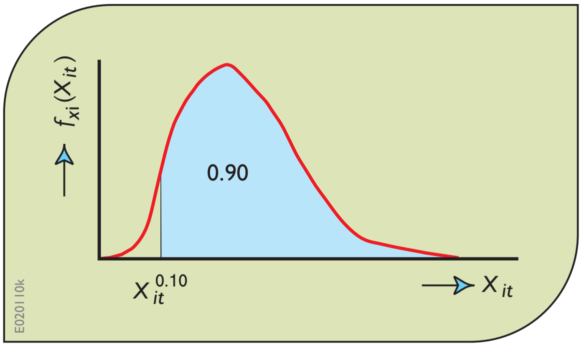

Figure 8.7. Probability density distribution of the random allocation X_i to firm i. The particular allocation value $x_{it}^{0.10}$ has a 90% chance of being equalled or exceeded, as indicated by the shaded region.

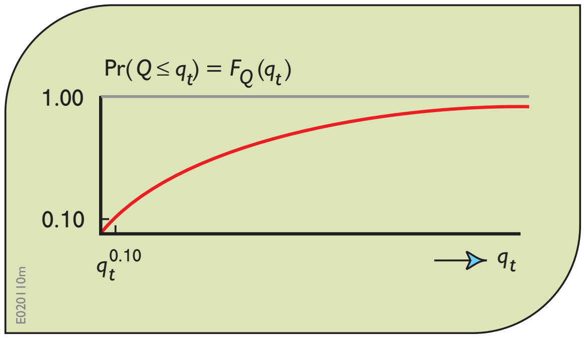

Figure 8.8. Example cumulative probability distribution showing the particular value of the random variable, $q_t^{0.10}$, that is equalled or exceeded 90% of the time.

To include chance constraints in optimization models, their deterministic equivalents must be defined. The deterministic equivalents of the three chance constraints in Equation 8.11 are:

$$T_i \le x_{it}^{0.10} \quad i = 1, 2 \text{ and } 3 \tag{8.12}$$

where $x_{it}^{0.10}$ is the particular value of the random variable X_i that is equalled or exceeded 90% of the time. This value is shown on the probability distribution for X_i in Figure 8.7.

To modify the allocation problem somewhat, assume the benefit obtained by each firm is a function of its target allocation and that the same allocation target applies in each time period t. The equipment and labour used in the firm is presumably based on the target allocations. Once the target is set, assume there are no benefits gained by excess water allocations. If the benefits obtained are to be based on the target allocations rather than the actual allocations, then the optimization problem is one of finding the values of the three targets that maximize the total benefits obtained with a reliability of, say, at least 90%.

$$\text{Maximize}\left(6T_1 - T_1^2\right) + \left(7T_2 - 1.5T_2^2\right) + \left(8T_3 - 0.5T_3^2\right) \tag{8.13}$$

subject to:

$$\Pr\{T_1 + T_2 + T_3 \le [Q_t - \min(Q_t, 2)]\} \ge 0.90 \quad \text{for all periods } t \tag{8.14}$$

where Q_t is the random streamflow variable upstream of all diversion sites. If the same unconditional probability distribution of Q_t applies for each period t, then only one Equation 8.14 is needed.

Assuming the value of the streamflow, $q_t^{0.10}$, that is equalled or exceeded 90% of the time, is greater than 2 (the amount that must remain in the stream), the deterministic equivalent of chance constraint Equation 8.14 is:

$$T_1 + T_2 + T_3 \le \left[q_t^{0.10} - \min\left(q_t^{0.10}, 2\right)\right] \tag{8.15}$$

The value of the flow that is equal to or exceeds 90% of the time, $q_t^{0.10}$, can be obtained from the cumulative distribution of flows as illustrated in Figure 8.8.

Assume this 90% reliable flow is 8. The deterministic equivalent of the chance constraint Equation 8.9 for all periods t is simply $T_1 + T_2 + T_3 \le 6$. The optimal solution of the chance-constrained target allocation model, Equations 8.8 and 8.9, is, as seen before, $T_1 = 1$, $T_2 = 1$ and $T_3 = 4$. The next step would be to simulate this problem to see what the actual reliabilities might be for various sequences of flows q_t.

5. Markov Processes and Transition Probabilities

Time-series correlations can be incorporated into models using transition probabilities. To illustrate this process, consider the observed flow sequence shown in Table 8.1.

The estimated mean, variance and correlation coefficient of the observed flows shown in Table 8.1 can be calculated using Equations 8.16, 8.17 and 8.18.

period t	flow Q_t	period t	flow Q_t	period t	flow Q_t
1	4.5	11	1.8	21	1.8
2	5.2	12	2.5	22	1.2
3	6.0	13	2.3	23	2.5
4	3.2	14	1.8	24	1.9
5	4.3	15	1.2	25	3.2
6	5.1	16	1.9	26	2.5
7	3.6	17	2.5	27	3.5
8	4.5	18	4.1	28	2.7
9	1.8	19	4.7	29	1.5
10	1.5	20	5.6	30	4.1
				31	4.8

Table 8.1. Sequence of flows for thirty-one time periods t.

$$E[Q] = \sum_{1}^{31} q_t/31 = 3.155 \tag{8.16}$$

$$\text{Var}[Q] = \sum_{1}^{31} (q_t - 3.155)^2/30 = 1.95 \tag{8.17}$$

Lag-one correlation coefficient $= \rho$

$$= \left[\sum_{1}^{30} (q_{t+1} - 3.155)(q_t - 3.155)\right] \Big/ \sum_{1}^{31} (q_t - 3.155)^2 = 0.50 \tag{8.18}$$

The probability distribution of the flows in Table 8.1 can be approximated by a histogram. Histograms can be created by subdividing the entire range of random variable values, such as flows, into discrete intervals. For example, let each interval be two units of flow. Counting the number of flows in each interval and then dividing those interval counts by the total number of counts results in the histogram shown in Figure 8.9. In this case, just to compare this with what will be calculated later, the first flow, q_1, is ignored.

Figure 8.9 shows a uniform unconditional probability distribution of the flow being in any of the possible discrete flow intervals. It does not show the possible dependency of the probabilities of the random variable

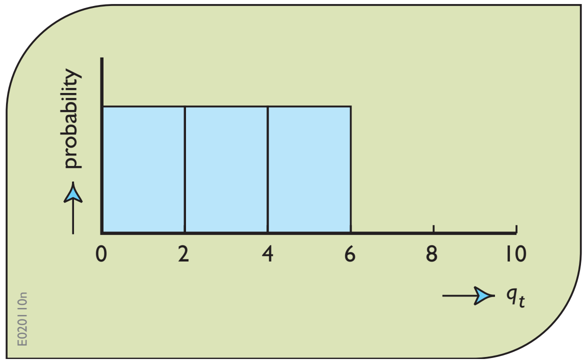

Figure 8.9. Histogram showing an equal 1/3 probability that the values of the random variable Q_t will be in any one of the three two-flow unit intervals.

flow interval in t: i	flow interval in t + 1: j = 1	2	3
1	5	4	1
2	3	4	3
3	2	2	6

Table 8.2. Matrix showing the number of times a flow in interval i in period t was followed by a flow in interval j in period $t + 1$.

value, q_{t+1}, in period $t + 1$ on the observed random variable value, q_t, in period t. It is possible that the probability of being in a flow interval j in period $t + 1$ depends on the actual observed flow interval i in period t.

To see if the probability of being in any given interval of flows is dependent on the past flow interval, one can create a matrix. The rows of the matrix are the flow intervals i in period t. The columns are the flow intervals j in the following period $t + 1$. Such a matrix is shown in Table 8.2. The numbers in the matrix are based on the flows in Table 8.1 and indicate the number of times a flow in interval j followed a flow in interval i.

Given an observed flow in an interval i in period t, the probabilities of being in one of the possible intervals j in the next period $t + 1$ must sum to 1. Thus, each number in each row of the matrix in Table 8.2 can be divided by the total number of flow transitions in that row (the sum

flow interval in t: i	flow interval in $t+1$: j 1	2	3
1	0.5	0.4	0.1
2	0.3	0.4	0.3
3	0.2	0.2	0.6

E020911b

Table 8.3. Matrix showing the probabilities P_{ij} of having a flow in interval j in period $t+1$ given an observed flow in interval i in period t.

of the number of flows in the row) to obtain the probabilities of being in each interval j in $t+1$ given a flow in interval i in period t. In this case there are ten flows that followed each flow interval i, hence by dividing each number in each row of the matrix by 10 defines the transition probabilities P_{ij}.

$$P_{ij} = \Pr\{Q_{t+1} \text{ in interval } j \mid Q_t \text{ in interval } i\} \tag{8.19}$$

These conditional or transition probabilities, shown in Table 8.3, correspond to the number of transitions shown in Table 8.2.

Table 8.3 is a matrix of transition probabilities. The sum of the probabilities in each row equals 1. Matrices of transition probabilities whose rows sum to 1 are also called stochastic matrices or first-order Markov chains.

If each row's probabilities were the same, this would indicate that the probability of observing any flow interval in the future is independent of the value of previous flows. Each row would have the same probabilities as the unconditional distribution shown in Figure 8.9. In this example the probabilities in each row differ, showing that low flows are more likely to follow low flows, and high flows are more likely to follow high flows. Thus the flows in Table 8.1 are positively correlated, as indeed has already determined from Equation 8.18.

Using the information in Table 8.3, one can compute the probability of observing a flow in any interval at any period on into the future given the present flow interval. This can be done one period at a time. For example assume the flow in the current time period $t = 1$ is in interval $i = 3$. The probabilities, $PQ_{j,2}$, of being in any of the three intervals in the following time period $t = 2$ are the probabilities shown in the third row of the matrix in Table 8.3.

The probabilities of being in an interval j in the following time period $t = 3$ is the sum over all intervals i of the joint probabilities of being in interval i in period $t = 2$ and making a transition to interval j in period $t = 3$.

$$\Pr\{Q_3 \text{ in interval } j\} = PQ_{j,3} = \sum_i \Pr\{Q_2 \text{ in interval } i\} \Pr\{Q_3 \text{ in interval } j \mid Q_2 \text{ in interval } i\} \tag{8.20}$$

The last term in Equation 8.20 is the transition probability, from Table 8.3, that in this example remains the same for all time periods t. These transition probabilities, $\Pr\{Q_{t+1} \text{ in interval } j \mid Q_t \text{ in interval } i\}$ can be denoted as P_{ij}.

Referring to Equation 8.19, Equation 8.20 can be written in a general form as:

$$PQ_{j,t+1} = \sum_i PQ_{it}P_{ij} \text{ for all intervals } j \text{ and periods } t \tag{8.21}$$

This operation can be continued to any future time period. Table 8.4 illustrates the results of such calculations for

time period t	flow interval i: 1	2	3
1	0	0	1
2	0.2	0.2	0.6
3	0.28	0.28	0.44
4	0.312	0.312	0.376
5	0.325	0.325	0.350
6	0.330	0.330	0.340
7	0.332	0.332	0.336
8	0.333	0.333	0.334

probability PQ_{it}

E020827y

Table 8.4. Probabilities of observing a flow in any flow interval i in a future time period t given a current flow in interval $i = 3$. These probabilities are derived using the transition probabilities P_{ij} in Table 8.3 in Equation 8.21 and assuming the flow interval observed in Period 1 is in Interval 3.

up to six future periods, given a present period ($t = 1$) flow in interval $i = 3$.

Note that as the future time period t increases, the flow interval probabilities are converging to the unconditional probabilities – in this example 1/3, 1/3, 1/3 – as shown in Figure 8.9. The predicted probability of observing a future flow in any particular interval at some time in the future becomes less and less dependent on the current flow interval as the number of time periods increases between the current period and that future time period.

When these unconditional probabilities are reached, PQ_{it} will equal $PQ_{i,t+1}$ for each flow interval i. To find these unconditional probabilities directly, Equation 8.21 can be written as:

$$PQ_j = \sum_i PQ_i P_{ij} \quad \text{for all intervals } j \qquad (8.22)$$

Equation 8.22 (less one) along with Equation 8.23 can be used to calculate all the unconditional probabilities PQ_i directly.

$$\sum_i PQ_i = 1 \qquad (8.23)$$

Conditional or transition probabilities can be incorporated into stochastic optimization models of water resources systems.

6. Stochastic Optimization

To illustrate the development and use of stochastic optimization models, consider first the allocation of water to a single user. Assume the flow in the stream where the diversion takes place is not regulated and can be described by a known probability distribution based on historical records. Clearly, the user cannot divert more water than is available in the stream. A deterministic model would include the constraint that the diversion x cannot exceed the available water Q. But Q is a random variable. Some target value, q, of the random variable Q will have to be selected, knowing that there is some probability that in reality, or in a simulation model, the actual flow may be less than the selected value q. Hence, if the constraint $x \leq q$ is binding, the actual allocation may be less than the value of the allocation or diversion variable x produced by the optimization model.

If the value of x affects one of the system's performance indicators, such as the net benefits, $B(x)$, to the user, a more accurate estimate of the user's net benefits will be obtained from considering a range of possible allocations x, depending on the range of possible values of the random flow Q. One way to do this is to divide the known probability distribution of flows q into discrete ranges, i – each range having a known probability PQ_i. Designate a discrete flow q_i for each range. Associated with each specified flow q_i is an unknown allocation x_i. Now the single deterministic constraint $x \leq q$ can be replaced with the set of deterministic constraints $x_i \leq q_i$, and the term $B(x)$ in the original objective function can be replaced by its expected value, $\sum_i PQ_i \cdot B(x_i)$.

Note, when dividing a continuous known probability distribution into discrete ranges, the discrete flows q_i, selected to represent each range i having a given probability PQ_i, should be selected so as to maintain at least the mean and variance of that known distribution as defined by Equations 8.5 and 8.6.

To illustrate this, consider a slightly more complex example involving the allocation of water to consumers upstream and downstream of a reservoir. Both the policies for allocating water to each user and the reservoir release policy are to be determined. This example problem is shown in Figure 8.10.

If the allocation of water to each user is to be based on a common objective, such as the minimization of the total sum, over time, of squared deviations from pre-specified target allocations, each allocation in each time period will depend in part on the reservoir storage volume.

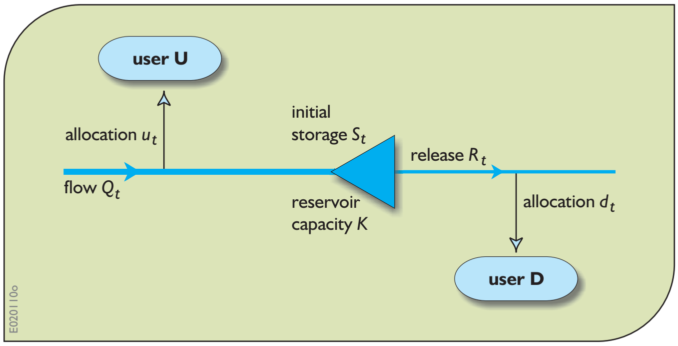

Figure 8.10. Example water resources system involving water diversions from a river both upstream and downstream of a reservoir of known capacity.

Consider first a deterministic model of the above problem, assuming known river flows Q_t and upstream and downstream user allocation targets UT_t and DT_t in each of T within-year periods t in a year. Assume the objective is to minimize the sum of squared deviations from actual allocations, u_t and d_t, and their respective target allocations, UT_t and DT_t in each within-year period t.

$$\text{Minimize} \sum_{t}^{T} \{(UT_t - u_t)^2 + (DT_t - d_t)^2\} \quad (8.24)$$

The constraints include:

a) Continuity of storage involving initial storage volumes S_t, net inflows $Q_t - u_t$, and releases R_t. Assuming no losses:

$$S_t + Q_t - u_t - R_t = S_{t+1} \quad \text{for each period } t, \; T+1 = 1 \quad (8.25)$$

b) Reservoir capacity limitations. Assuming a known active storage capacity K:

$$S_t \leq K \quad \text{for each period } t \quad (8.26)$$

c) Allocation restrictions. For each period t:

$$u_t \leq Q_t \quad (8.27)$$

$$d_t \leq R_t \quad (8.28)$$

Equations 8.25 and 8.28 could be combined to eliminate the release variable R_t, since in this problem knowledge of the total release in each period t is not required. In this case, Equation 8.25 would become an inequality.

The solution of this model, Equations 8.24 – 8.28, would depend on the known variables (the targets UT_t and DT_t, flows Q_t and reservoir capacity K). It would identify the particular upstream and downstream allocations and reservoir releases in each period t. It would not provide a policy that defines what allocations and releases to make for a range of different inflows and initial storage volumes in each period t. A backward-moving dynamic programming model can provide such a policy. This policy will identify the allocations and releases to make based on various initial storage volumes, S_t, and flows, Q_t, as discussed in Chapter 4.

This deterministic discrete dynamic programming allocation and reservoir operation model can be written for different discrete values of S_t from $0 \leq S_t \leq$ capacity K as:

$$F_t^n(S_t, Q_t) = \min\{(UT_t - u_t)^2 + (DT_t - d_t)^2 + F_{t+1}^{n-1}(S_{t+1}, Q_{t+1})\}$$

The minimization is over all feasible u_t, R_t, d_t:

$$u_t \leq Q_t$$

$$R_t \leq S_t + Q_t - u_t$$

$$R_t \geq S_t + Q_t - u_t - K$$

$$d_t \leq R_t$$

$$S_{t+1} = S_t + Q_t - u_t - R_t \quad (8.29)$$

There are three variables to be determined at each stage or time period t in the above dynamic programming model. These three variables are the allocations u_t and d_t and the reservoir release R_t. Each decision involves three discrete decision-variable values. The functions $F_t^n(S_t, Q_t)$ define the minimum sum of squared deviations given an initial storage volume S_t and streamflow Q_t in time period or season t with n time periods remaining until the end of reservoir operation.

One can reduce this three decision-variable model to a single variable model by realizing that, for any fixed discrete pair of initial and final storage volume states, there can be a direct tradeoff between the upstream and downstream allocations, given the particular streamflow in each period t. Increasing the upstream allocation will decrease the resulting reservoir inflow, and this in turn will reduce the release by the same amount. This reduces the amount of water available to allocate to the downstream use.

Hence, for this example problem involving these upstream and downstream allocations, a local optimization can be performed at each time step t for each combination of storage states S_t and S_{t+1}. This optimization finds the allocation decision-variables u_t and d_t that

$$\text{minimize}(UT_t - u_t)^2 + (DT_t - d_t)^2 \quad (8.30)$$

where

$$u_t \leq Q_t \quad (8.31)$$

$$d_t \leq S_t + Q_t - u_t - S_{t+1} \quad (8.32)$$

This local optimization can be solved to identify the u_t and d_t allocations for each feasible combination of S_t and S_{t+1} in each period t.

Given these optimal allocations, the dynamic programming model can be simplified to include only one discrete decision-variable, either R_t or S_{t+1}. If the decision variable S_{t+1} is used in each period t, the releases R_t in those periods t do not need to be considered. Thus the dynamic programming model expressed by Equations 8.29 can be written for all discrete storage volumes S_t from 0 to K and for all discrete flows Q_t as:

$$F_t^n(S_t, Q_t) = \min\{(UT_t - u_t(S_t, S_{t+1}))^2 + (DT_t - d_t(S_t, S_{t+1}))^2 + F_{t+1}^{n-1}(S_{t+1}, Q_{t+1})\}$$

The minimization is over all feasible discrete values of S_{t+1},

$$S_{t+1} \leq K \tag{8.33}$$

where the functions $u_t(S_t, S_{t+1})$ and $d_t(S_t, S_{t+1})$ have been determined using Equations 8.30 – 8.32.

As the total number of periods remaining, n, increases, the solution of this dynamic programming model will converge to a steady or stationary state. The best final storage volume S_{t+1} given an initial storage volume S_t will probably differ for each within-year period or season t, but for a given season t it will be the same in successive years. In addition, for each storage volume S_t, streamflow, Q_t, and within-year period t, the difference between $F_t^{n+T}(S_t, Q_t)$ and $F_t^n(S_t, Q_t)$ will be the same constant regardless of the storage volume S_t and period t. This constant is the optimal, in this case minimum, annual value of the objective function, Equation 8.24.

There could be additional limits imposed on storage variables and release variables, such as for flood control storage or minimum downstream flows, as might be appropriate in specific situations.

The above deterministic dynamic programming model (Equation. 8.33) can be converted to a stochastic model. Stochastic models consider multiple discrete flows as well as multiple discrete storage volumes, and their probabilities, in each period t. A common way to do this is to assume that the sequence of flows follow a first-order Markov process. Such a process involves the use of transition or conditional probabilities of flows as defined by Equation 8.20.

To develop these stochastic optimization models, it is convenient to introduce some additional indices or subscripts. Let the index k denote different initial storage volume intervals. These discrete intervals divide the continuous range of storage volume values from 0 to the active reservoir capacity K. Each S_{kt} is a discrete storage volume that represents the range of storage volumes in interval k at the beginning of each period t.

Let the following letter l be the index denoting different final storage volume intervals. Each $S_{l,t+1}$ is a discrete volume that represents the storage volume interval l at the end of each period t or equivalently at the beginning of period $t + 1$. As previously defined, let the indices i and j denote the different flow intervals, and each discrete flow q_{it} and $q_{j,t+1}$ represent those flow intervals i and j in periods t and $t + 1$ respectively.

These subscripts and the volume or flow intervals they represent are illustrated in Figure 8.11.

With this notation, it is now possible to develop a stochastic dynamic programming model that will identify the allocations and releases that are to be made given both the initial storage volume, S_{kt}, and the flow, q_{it}. It follows the same structure as the deterministic models defined by Equations 8.30 through 8.32, and 8.33.

To identify the optimal allocations in each period t for each pair of feasible initial and final storage volumes S_{kt} and $S_{l,t+1}$, and inflows $q_{it,}$ one can solve Equations 8.34 through 8.36.

$$\text{minimize } (UT_t - u_{kit})^2 + (DT_t - d_{kilt})^2 \tag{8.34}$$

where

$$u_{kit} \leq q_{it} \quad \forall\ k, i, t. \tag{8.35}$$

$$d_{kilt} \leq S_{kt} + q_{it} - u_{kit} - S_{l,t+1} \quad \forall \text{ feasible } k, i, l, t. \tag{8.36}$$

The solution to these equations for each feasible combination of intervals k, i, l, and period t defines the optimal allocations that can be expressed as $u_t(k, i)$ and $d_t(k, i, l)$.

The stochastic version of Model 8.33, again expressed in a form suitable for backward-moving discrete dynamic programming, can be written for different discrete values of S_{kt} from 0 to K and for all q_{it} as:

$$F_t^n(S_{kt}, q_{it}) = \min\left\{(UT_t - u_t(k, t))^2 + (DT_t - d_t(k, i, l))^2 + \sum_j P_{ij}^t F_{t+1}^{n-1}(S_{l,t+1}, q_{j,t+1})\right\}$$

The minimization is over all feasible discrete values of $S_{l,t+1}$

$$S_{l,t+1} \leq K$$

$$S_{l,t+1} \leq S_{kt} + q_{it} \tag{8.37}$$

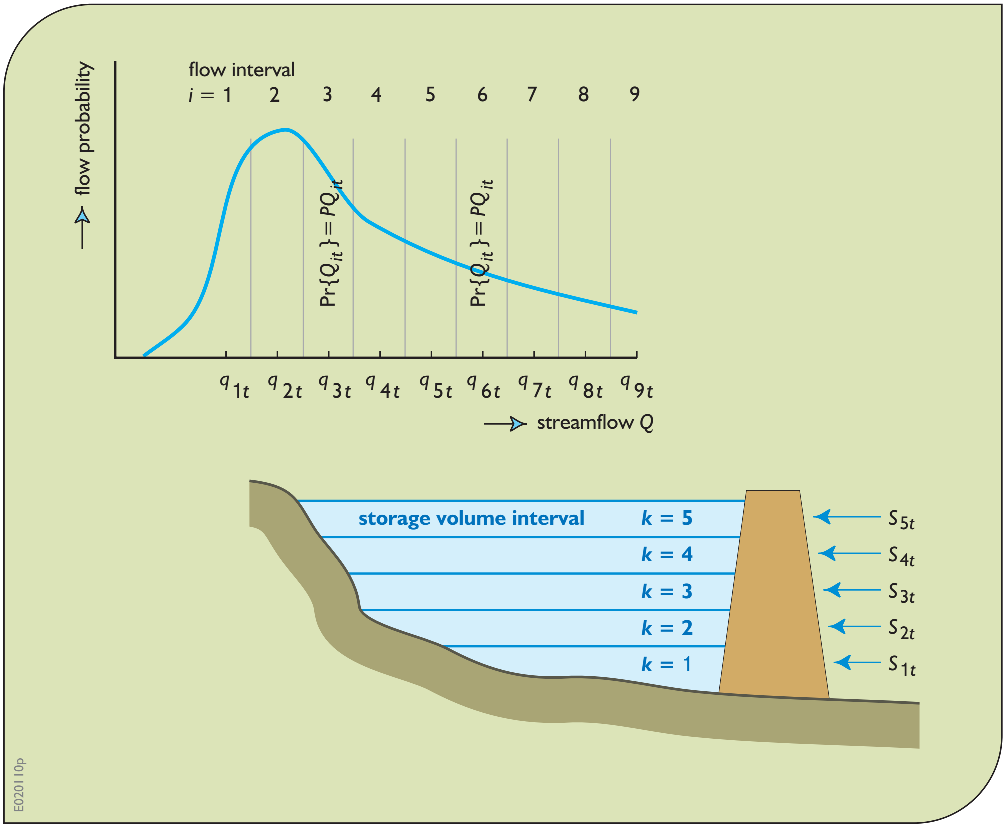

Figure 8.11. Discretization of streamflows and reservoir storage volumes. The area within each flow interval *i* below the probability density distribution curve is the unconditional probability, PQ_{it}, associated with the discrete flow q_{it}.

Each P^t_{ij} in the above recursive equation is the known conditional or transition probability of a flow $q_{j,t+1}$ within interval j in period $t+1$ given a flow of q_{it} within interval i in period t.

$$P^t_{ij} = \Pr\{\text{flow } q_{j,t+1} \text{ within interval } j \text{ in } t+1 \mid \text{flow of } q_{it} \text{ within interval } i \text{ in } t\}$$

The sum over all flow intervals j of these conditional probabilities times the $F^{n-1}_{t+1}(S_{l,t+1}, q_{j,t+1})$ values is the expected minimum sum of future squared deviations from allocation targets with $n-1$ periods remaining given an initial storage volume of S_{kt} and flow of q_{it} and final storage volume of $S_{l,t+1}$. The value $F^n_t(S_{kt}, q_{it})$ is the expected minimum sum of squared deviations from the allocation targets with n periods remaining given an initial storage volume of S_{kt} and flow of q_{it}. Stochastic models such as these provide expected values of objective functions.

Another way to write the recursion equations of this model, Equation 8.37, is by using just the indices k and l to denote the discrete storage volume variables S_{kt} and $S_{l,t+1}$ and indices i and j to denote the discrete flow variables q_{it} and $q_{j,t+1}$:

$$F^n_t(k,i) = \min_l \left\{ (UT_t - u_t(k,t))^2 + (DT_t - d_t(k,i,l))^2 + \sum_j P^t_{ij} F^{n-1}_{t+1}(l,j) \right\}$$

such that $S_{l,t+1} \le K$

$$S_{l,t+1} \le S_{kt} + q_{it} \qquad (8.38)$$

The steady-state solution of this dynamic programming model will identify the preferred final storage volume $S_{l,t+1}$ in period t given the particular discrete initial storage volume S_{kt} and flow q_{it}. This optimal policy can be expressed as a function ℓ that identifies the best interval l given intervals k, i and period t.

$$l = \ell(k, i, t) \qquad (8.39)$$

All values of l given k, i and t, defined by Equation 8.39, can be expressed in a matrix, one for each period t.

Knowing the best final storage volume interval l given an initial storage volume interval k and flow interval i, the optimal downstream allocation, $d_t(k, i)$, can, like the upstream allocation, be expressed in terms of only k and i in each period t. Thus, knowing the initial storage volume S_{kt} and flow q_{it} is sufficient to define the optimal allocations $u_t(k, i)$ and $d_t(k, i)$, final storage volume $S_{l,t+1}$, and hence the release $R_t(k, i)$.

$$S_{kt} + q_{it} - u_t(k, i) - R_t(k, i) = S_{l,t+1} \quad \forall\ k, i, t$$
$$\text{where } l = \ell(k, i, t) \tag{8.40}$$

6.1. Probabilities of Decisions

Knowing the function $l = \ell(k, i, t)$ permits a calculation of the probabilities of the different discrete storage volumes, allocations, and flows. Let

PS_{kt} = the unknown probability of an initial storage volume S_{kt} being within some interval k in period t;

PQ_{it} = the steady-state unconditional probability of flow q_{it} within interval i in period t; and

P_{kit} = the unknown probability of the upstream and downstream allocations $u_t(k, i)$ and $d_t(k, i)$ and reservoir release $R_t(k, i)$ in period t.

As previously defined,
P^t_{ij} = the known conditional or transition probability of a flow within interval j in period $t+1$ given a flow within interval i in period t.

These transition probabilities P^t_{ij} can be displayed in matrices, similar to Table 8.3, but as a separate matrix (Markov chain) for each period t.

The joint probabilities of an initial storage interval k, an inflow in the interval i, P_{kit} in each period t must satisfy two conditions. Just as the initial storage volume in period $t+1$ is the same as the final storage volume in period t, the probabilities of these same respective discrete storage volumes must also be equal. Thus,

$$\sum_j P_{l,j,t+1} = \sum_k \sum_i P_{kit} \quad \forall l, t \tag{8.41}$$

where the sums in the right hand side of Equation 8.41 are over only those combinations of k and i that result in a final volume interval l. This relationship is defined by Equation 8.39 ($l = \ell(k, i, t)$).

While Equation 8.41 must apply, it is not sufficient. The joint probability of a final storage volume in interval l in period t and an inflow j in period $t+1$ must equal the joint probability of an initial storage volume in the same interval l and an inflow in the same interval j in period $t+1$. Multiplying the joint probability P_{kit} times the conditional probability P^t_{ij} and then summing over all k and i that results in a final storage interval l defines the former, and the joint probability $P_{l,j,t+1}$ defines the latter.

$$P_{l,j,t+1} = \sum_k \sum_i P_{kit} P^t_{ij} \quad \forall l, j, t \quad l = \ell(k, i, t) \tag{8.42}$$

Once again the sums in Equation 8.42 are over all combinations of k and i that result in the designated storage volume interval l as defined by the policy $\ell(k, i, t)$.

Finally, the sum of all joint probabilities P_{kit} in each period t must equal 1.

$$\sum_k \sum_i P_{kit} = 1 \quad \forall t \tag{8.43}$$

Note the similarity of Equations 8.42 and 8.43 to the Markov steady-state flow Equations 8.22 and 8.23. Instead of only one flow interval index considered in Equations 8.22 and 8.23, Equations 8.42 and 8.43 include two indices, one for storage volume intervals and the other for flow intervals. In both cases, one of Equations 8.22 and 8.42 can be omitted in each period t since it is redundant with that period's Equations 8.23 and 8.43 respectively.

The unconditional probabilities PS_{kt} and PQ_{it} can be derived from the joint probabilities P_{kit}.

$$PS_{kt} = \sum_i P_{kit} \quad \forall k, t \tag{8.44}$$

$$PQ_{it} = \sum_k P_{kit} \quad \forall i, t \tag{8.45}$$

Each of these unconditional joint or marginal probabilities, when summed over all their volume and flow indices, will equal 1. For example,

$$\sum_k PS_{kt} = \sum_i PQ_{it} = 1 \tag{8.46}$$

Note that these probabilities are determined only on the basis of the relationships among flow and storage intervals as defined by Equation 8.39, $l = \ell(k, i, t)$ in each period t, and the Markov chains defining the flow interval transition or conditional probabilities, P^t_{ij}. It is not necessary to know the actual discrete storage values representing those intervals. Thus assuming any relationship among the storage volume and flow interval indices, $l = \ell(k, i, t)$ and a knowledge of the flow interval transition probabilities P_{ij}^{t}, one can determine the joint probabilities P_{kit} and their marginal or unconditional probabilities PS_{kt}. One does not need to know what those storage intervals are to calculate their probabilities.

Given the values of these joint probabilities P_{kit}, the deterministic model defined by Equations 8.24 to 8.28 can be converted to a stochastic model to identify the best storage and allocation decision-variable values associated with each storage interval k and flow interval i in each period t.

$$\text{Minimize} \sum_k \sum_i \sum_{\tau}^{T} P_{kit}\{(UT_t - u_{kit})^2 + (DT_t - d_{kit})^2\} \tag{8.47}$$

The constraints include:

a) Continuity of storage involving initial storage volumes S_{kt}, net inflows $q_{it} - u_{kit}$, and at least partial releases d_{kit}. Again assuming no losses:

$$S_{kt} + q_{it} - u_{kit} - d_{kit} \geq S_{l,t+1} \quad \forall k, i, t$$
$$l = \ell(k, i, t) \tag{8.48}$$

b) Reservoir capacity limitations.

$$S_{kit} \leq K \quad \forall k, i, t \tag{8.49}$$

c) Allocation restrictions.

$$u_{kit} \leq q_{it} \quad \forall k, i, t \tag{8.50}$$

More detail on these and other stochastic modelling approaches can be found in Faber and Stedinger (2001); Gablinger and Loucks (1970); Huang et al. (1991); Kim and Palmer (1997); Loucks and Falkson (1970); Stedinger et al. (1984); Su and Deininger (1974); Tejada-Guibert et al. (1993 1995); and Yakowitz (1982).

6.2. A Numerical Example

A simple numerical example may help to illustrate how these stochastic models can be developed without getting buried in detail. Consider two within-year periods each year. The random flows Q_t in each period t are divided into two intervals. These flow intervals are represented by discrete flows of 1 and 3 volume units per second in the first period, and 3 and 6 volume units per second in the second period. Their transition probabilities are shown in Table 8.5.

Q_i flow in t = 1 \ Q_j flow in t = 2	3	6
1	0.6	0.4
3	0.3	0.7

Q_i flow in t = 2 \ Q_j flow in t = 1	1	3
3	0.7	0.3
6	0.2	0.8

E020829a

Table 8.5. Transition probabilities for two ranges of flows in two within-year periods.

Assuming equal within-year period durations, these three discrete flow rates are equivalent to about 16, 47 and 95 million volume units per period.

Assume the active storage volume capacity K in the reservoir equals 50 million volume units. This capacity can be divided into different intervals of storage. For this simple example, assume three storage volume intervals represented by 10, 25 and 40 million volume units. Assume the allocation targets remain the same in each period at both the upstream and downstream sites. The upstream allocation target is approximately 2 volume units per second or 30 million volume units in each period. The downstream allocation target is approximately 5 volume units per second or 80 million volume units in each period.

With these data we can use Equations 8.34 – 8.36 to determine the allocations that minimize the sum of squared deviations from targets and what that sum is, for all feasible combinations of initial and final storage volumes and flows. Table 8.6 shows the results of these optimizations. These results will be used in the dynamic programming model to determine the best final storage volumes given initial volumes and flows.

With the information in Tables 8.5 and 8.6, the dynamic programming model, Equation 8.38 or as expressed in Equation 8.51, can be solved to find the optimal final storage volumes, given an initial storage volume and flow. The iterations of the recursive equation, sufficient to reach a steady state, are shown in Table 8.7.

$$F_t^n(k, i) = \min_l \left\{ SD_{kil} + \sum_j P_{ij}^t F_{t+1}^{n-1}(l, j) \right\}$$

$$\text{such that} \quad S_{l,t+1} \leq K$$

$$S_{l,t+1} \leq S_{kt} + Q_{it} \tag{8.51}$$

Table 8.6. Optimal allocations associated with given initial storage, S_k, flow, Q_i, and final storage, S_l, volumes. These allocations u_{ki} and d_{kil} minimize the sum of squared deviations, $DS_{kil} = (30 - u_{ki})^2 + (80 - d_{kil})^2$, from upstream and downstream targets, 30 and 80 respectively, subject to $u_{ki} \leq$ flow Q_i, and $d_{kil} \leq$ release $(S_k + Q_i - u_{ki} - S_l)$.

E020829b

initial storage S_k	flow Q_i	final storage S_1	interval indices k, i, l	upstream allocation u_{ki}	down-stream allocation d_{kil}	sum squared deviation SD_{kil}
10	16	10	1, 1, 1	0.0	16.0	4996.0
10	16	25	1, 1, 2	0.0	1.0	7141.0
10	47	10	1, 2, 1	0.0	47.0	1989.0
10	47	25	1, 2, 2	0.0	32.0	3204.0
10	47	40	1, 2, 3	0.0	17.0	4869.0
10	95	10	1, 3, 1	22.5	72.5	112.5
10	95	25	1, 3, 2	15.0	65.0	450.0
10	95	40	1, 3, 3	7.5	75.5	1012.5
25	16	10	2, 1, 1	0.0	31.0	3301.0
25	16	25	2, 1, 2	0.0	16.0	4996.0
25	16	40	2, 1, 3	0.0	1.0	7141.0
25	47	10	2, 2, 1	6.0	56.0	1152.0
25	47	25	2, 2, 2	0.0	47.0	1989.0
25	47	40	2, 2, 3	0.0	32.0	3204.0
25	95	10	2, 3, 1	30.0	80.0	0.0
25	95	25	2, 3, 2	22.5	72.5	112.5
25	95	40	2, 3, 3	15.0	65.0	450.0
40	16	10	3, 1, 1	0.0	46.0	2056.0
40	16	25	3, 1, 2	0.0	31.0	3301.0
40	16	40	3, 1, 3	0.0	16.0	4996.0
40	47	10	3, 2, 1	13.5	63.5	544.5
40	47	25	3, 2, 2	6.0	56.0	1152.0
40	47	40	3, 2, 3	0.0	47.0	1989.0
40	95	10	3, 3, 1	30.0	80.0	0.0
40	95	25	3, 3, 2	30.0	80.0	0.0
40	95	40	3, 3, 3	22.5	72.5	112.5

This process can continue until a steady-state policy is defined. Table 8.8 summarizes the next five iterations. At this stage, the annual differences in the objective values associated with a particular state and season have come close to a common constant value.

While the differences between corresponding F_t^{n+T} and F_t^n have not yet reached a common constant value to the nearest unit deviation (they range from, 3475.5 to 3497.1 for an average of 3485.7), the policy has converged to that shown in Tables 8.8 and 8.9.

Given this operating policy, the probabilities of being in any of these volume and flow intervals can be determined by solving Equations 8.42 through 8.45. Table 8.10 shows the results of these equations applied to the data in Tables 8.5 and 8.8. It is obvious that if the policy from Table 8.9 is followed, the steady-state probabilities of being in storage Interval 1 in Period 1 and in Interval 3 in Period 2 are 0.

Multiplying these joint probabilities by the corresponding SD_{kit} values in the last column of Table 8.6 provides the annual expected squared deviations, associated with the selected discrete storage volumes and flows. This is done in Table 8.11 for those combinations of k, i, and l that are contained in the optimal solution as listed in Table 8.9.

The sum of products of the last two columns in Table 8.11 for each period t equals the expected squared deviations in the period. For period $t = 1$, the expected

Table 8.7. First four iterations of dynamic programming model, Equations 8.51, moving backward in successive periods n, beginning in season $t = 2$ with $n = 1$. The iterations stop when the final storage policy given any initial storage volume and flow repeats itself in two successive years. Initially, with no more periods remaining, $F_1^0(k, i) = 0$ for all k and i.

E020829c

period t = 2, n = 1

storage & flow k, i		$SD_{kil} + \Sigma_j P_{ij}^t F_{t+1}^{n-1}(l, j)$	$F_t^n(k, i)$	optimal l
1,2	$l = 1$	1989.0 + 0	1989.0	1
	$l = 2$	3204.0 + 0		
	$l = 3$	4869.0 + 0		
1,3	$l = 1$	112.5 + 0	112.5	1
	$l = 2$	450.0 + 0		
	$l = 3$	1012.0 + 0		
2,2	$l = 1$	1152.0 + 0	1152.0	1
	$l = 2$	1989.0 + 0		
	$l = 3$	3204.0 + 0		
2,3	$l = 1$	0.0 + 0	0.0	1
	$l = 2$	112.5 + 0		
	$l = 3$	450.0 + 0		
3,2	$l = 1$	544.5 + 0	544.5	1
	$l = 2$	1152.0 + 0		
	$l = 3$	1989.0 + 0		
3,3	$l = 1$	0.0 + 0	0.0	1,2
	$l = 2$	0.0 + 0		
	$l = 3$	112.5 + 0		

E020829d

period t = 1, n = 2

storage & flow k, i		$SD_{kil} + \Sigma_j P_{ij}^t F_{t+1}^{n-1}(l, j)$	$F_t^n(k, i)$	optimal l
1,1	$l = 1$	4996.0 + 0.6 (1989.0) +0.4 (112.5) = **6234.4**	6234.4	1
	$l = 2$	7141.0 + 0.6 (1152.0) +0.4 (0.0) = 7832.2		
	$l = 3$	infeasible ---- = ---		
1,2	$l = 1$	1989.0 + 0.3 (1989.0) +0.7 (112.5) = **2664.45**	2664.5	1
	$l = 2$	3204.0 + 0.3 (1152.0) +0.7 (0.0) = 3549.6		
	$l = 3$	4869.0 + 0.3 (544.5) +0.7 (0.0) = 5032.35		
2,1	$l = 1$	3301.0 + 0.6 (1989.0) +0.4 (112.5) = **4539.4**	4539.4	1
	$l = 2$	4996.0 + 0.6 (1152.0) +0.4 (0.0) = 5687.2		
	$l = 3$	7141.0 + 0.6 (544.5) +0.4 (0.0) = 7467.7		
2,2	$l = 1$	1152.0 + 0.3 (1989.0) +0.7 (112.5) = **1827.45**	1827.5	1
	$l = 2$	1989.0 + 0.3 (1152.0) +0.7 (0.0) = 2334.6		
	$l = 3$	3204.0 + 0.3 (544.5) +0.7 (0.0) = 3367.35		
3,1	$l = 1$	2056.0 + 0.6 (1989.0) +0.4 (112.5) = **3294.4**	3294.4	1
	$l = 2$	3301.0 + 0.6 (1152.0) +0.4 (0.0) = 3992.2		
	$l = 3$	4996.0 + 0.6 (544.5) +0.4 (0.0) = 5322.7		
3,2	$l = 1$	544.5 + 0.3 (1989.0) +0.7 (112.5) = **1219.95**	1220.0	1
	$l = 2$	1152.0 + 0.3 (1152.0) +0.7 (0.0) = 1497.6		
	$l = 3$	1989.0 + 0.3 (544.5) +0.7 (0.0) = 2152.35		

(contd.)

Table 8.7. Concluded.

E020829e

storage & flow k, i	period t = 2, n = 3	$SD_{kil} + \Sigma_j P_{ij}^t F_{t+1}^{n-1}(l, j)$	$F_t^n(k, i)$	optimal l
1,2	l = 1	1989.0 + 0.7 (6234.4) +0.3 (2664.5) = 7152.4	6929.8	2
	l = 2	3204.0 + 0.7 (4539.4) +0.3 (1827.5) = **6929.8**		
	l = 3	4869.0 + 0.7 (3294.4) +0.3 (1219.9) = 7541.1		
1,3	l = 1	112.5 + 0.2 (6234.4) +0.8 (2664.5) = 3490.0	2647.3	3
	l = 2	450.0 + 0.2 (4539.4) +0.8 (1827.5) = 2819.8		
	l = 3	1012.5 + 0.2 (3294.4) +0.8 (1219.9) = **2647.3**		
2,2	l = 1	1152.0 + 0.7 (6234.4) +0.3 (2664.5) = 6315.4	5714.8	2
	l = 2	1989.0 + 0.7 (4539.4) +0.3 (1827.5) = **5714.8**		
	l = 3	3204.0 + 0.7 (3294.4) +0.3 (1219.9) = 5876.1		
2,3	l = 1	0.0 + 0.2 (6234.4) +0.8 (2664.5) = 3378.4	2084.8	3
	l = 2	112.5 + 0.2 (4539.4) +0.8 (1827.5) = 2482.3		
	l = 3	450.0 + 0.2 (3294.4) +0.8 (1219.9) = **2084.8**		
3,2	l = 1	544.5 + 0.7 (6234.4) +0.3 (2664.5) = 5707.9	4661.1	3
	l = 2	1152.0 + 0.7 (4539.4) +0.3 (1827.5) = 4877.8		
	l = 3	1989.0 + 0.7 (3294.4) +0.3 (1219.9) = **4661.1**		
3,3	l = 1	0.0 + 0.2 (6234.4) +0.8 (2664.5) = 3378.4	1747.3	3
	l = 2	0.0 + 0.2 (4539.4) +0.8 (1827.5) = 2369.8		
	l = 3	112.5 + 0.2 (3294.4) +0.8 (1219.9) = **1747.3**		

E020829f

storage & flow k, i	period t = 1, n = 4	$SD_{kil} + \Sigma_j P_{ij}^t F_{t+1}^{n-1}(l, j)$	$F_t^n(k, i)$	optimal l
1,1	l = 1	4996.0 + 0.6 (6929.8) +0.4 (2647.3) = **10212.8**	10212.8	1
	l = 2	7141.0 + 0.6 (5714.8) +0.4 (2084.8) = 11403.8		
	l = 3	infeasible ---		
1,2	l = 1	1989.0 + 0.3 (6929.8) +0.7 (2647.3) = **5921.1**	5921.1	1
	l = 2	3204.0 + 0.3 (5714.8) +0.7 (2084.8) = 6377.8		
	l = 3	4869.0 + 0.3 (4661.1) +0.7 (1747.3) = 7490.5		
2,1	l = 1	3301.0 + 0.6 (6929.8) +0.4 (2647.3) = **8517.8**	8517.8	1
	l = 2	4996.0 + 0.6 (5714.8) +0.4 (2084.8) = 9258.8		
	l = 3	7141.0 + 0.6 (4661.1) +0.4 (1747.3) = 10636.6		
2,2	l = 1	1152.0 + 0.3 (6929.8) +0.7 (2647.3) = **5084.1**	5084.1	1
	l = 2	1989.0 + 0.3 (5714.8) +0.7 (2084.8) = 5162.8		
	l = 3	3204.0 + 0.3 (4661.1) +0.7 (1747.3) = 5825.5		
3,1	l = 1	2056.0 + 0.6 (6929.8) +0.4 (2647.3) = **7272.8**	7272.8	1
	l = 2	3301.0 + 0.6 (5714.8) +0.4 (2084.8) = 7563.8		
	l = 3	4996.0 + 0.6 (4661.1) +0.4 (1747.3) = 8491.6		
2,2	l = 1	544.5 + 0.3 (6929.8) +0.7 (2647.3) = 4476.6	4325.8	2
	l = 2	1152.0 + 0.3 (5714.8) +0.7 (2084.8) = **4325.8**		
	l = 3	1989.0 + 0.3 (4661.1) +0.7 (1747.3) = 4610.5		

Table 8.8. Summary of objective function values $F_t^n(k, i)$ and optimal decisions for stages $n = 5$ to 9 periods remaining.

storage & flow k, i	t = 2, n = 5 $F_t^n(k, i)$	l^*	t = 1, n = 6 $F_t^n(k, i)$	l^*	t = 2, n = 7 $F_t^n(k, i)$	l^*	t = 1, n = 8 $F_t^n(k, i)$	l^*	t = 2, n = 9 $F_t^n(k, i)$	l^*
1,1			13782.1	1			17279.2	1		
1,2	10691.7	2	9345.9	1	14217.7	2	12821.4	1	17708.3	2
1,3	5927.7	3			9381.5	3			12861.3	3
2,1			12087.1	1			15584.2	1		
2,2	9476.7	2	8508.9	1	13002.7	2	11984.4	1	16493.2	2
2,3	5365.2	3			8819.0	3			12298.7	3
3,1			10842.1	1			14339.2	1		
3,2	8377.7	3	7750.7	2	11903.7	3	11226.1	2	15394.3	3
3,3	5027.7	3			8481.5	3			11961.2	3

E020829g

sum of squared deviations are 1893.3 and for $t = 2$ they are 1591.0. The total annual expected squared deviations are 3484.3. This compares with the expected squared deviations derived from the dynamic programming model, after 9 iterations, ranging from 3475.5 to 3497.1 (as calculated from data in Table 8.8).

These upstream allocation policies can be displayed in plots, as shown in Figure 8.12.

The policy for reservoir releases is a function not only of the initial storage volumes, but also of the current inflow, in other words, the total water available in the period. Reservoir release rule curves such as shown in Figures 4.16 or 4.18 now must become two-dimensional. However, the inflow for each period usually cannot be predicted with certainty at the beginning of each period. In situations where the release cannot be adjusted during the period as the inflow becomes more predictable, the reservoir release policy has to be expressed in a way that can be followed without knowledge of the current inflow. One way to do this is to compute the expected value of the release for each discrete storage volume, and show it in a release rule. This is done in Figure 8.13. The probability of each discrete release associated with each discrete river flow is the probability of the flow itself. Thus, in Period 1 when the storage volume is 40, the expected release is $46(0.41) + 56(0.59) = 52$. These discrete expected releases can be used to define a continuous range of releases for the continuous range of storage volumes from 0 to full capacity, 50. Figure 8.13 also shows the hedging that might take place as the reservoir storage volume decreases.

Another approach to defining the releases in each period in a manner that is not dependent on knowledge of the current inflow, even though the model used assumes this, is to attempt to define either release targets with constraints on final storage volumes, or final storage targets with constraints on total releases. Obviously, such policies will not guarantee constant releases throughout each period. For example, consider the optimal policy shown in Table 8.9. The releases (or final storage volumes) in each period are dependent on the initial storage and current inflow. However, this operating policy can be expressed as:

- If in period 1, the final storage target should be in interval 1. Yet the total release cannot exceed the flow in interval 2.
- If in period 2 and the initial storage is in interval 1, the release should be in interval 1.
- If in period 2 and the initial storage is in interval 2, the release should be in interval 2.
- If in period 2 and the initial storage is in interval 3, the release should equal the inflow.

This policy can be followed without any forecast of current inflow. It will provide the releases and final storage volumes that would be obtained with a perfect inflow forecast at the beginning of each period.

E020829h

period t	initial storage volume and flow interval k	i	final storage volume interval l
1	1	1	1
1	1	2	1
1	2	1	1
1	2	2	1
1	3	1	1
1	3	2	2
2	1	2	2
2	1	3	3
2	2	2	2
2	2	3	3
2	3	2	3
2	3	3	3

Table 8.9. Optimal reservoir policy $l = \ell(k, i, t)$ for the example problem.

E020829j

unconditional probabilities PQ_{it} of flow intervals i in the 2 time periods t

$PQ(1, 1) = 0.4117647$	$PQ(2, 2) = 0.4235294$
$PQ(2, 1) = 0.5882353$	$PQ(3, 2) = 0.5764706$

unconditional probabilities PS_{kt} of storage intervals k in the 2 time periods t

$PS(1, 1) = 0.0000000$	$PS(1, 2) = 0.5388235$
$PS(2, 1) = 0.4235294$	$PS(2, 2) = 0.4611765$
$PS(3, 1) = 0.5764706$	$PS(3, 2) = 0.0000000$

joint probabilities P_{kit} of storage volume intervals k and flow intervals i in the 2 time periods t

$P(1, 1, 1) = 0.0000000$	$P(1, 2, 2) = 0.2851765$
$P(1, 2, 1) = 0.0000000$	$P(1, 3, 2) = 0.2536471$
$P(2, 1, 1) = 0.2964706$	$P(2, 2, 2) = 0.1383529$
$P(2, 2, 1) = 0.1270588$	$P(2, 3, 2) = 0.3228235$
$P(3, 1, 1) = 0.1152941$	$P(3, 2, 2) = 0.0000000$
$P(3, 2, 1) = 0.4611765$	$P(3, 3, 2) = 0.0000000$

Table 8.10. Probabilities of flow and storage volume intervals associated with the policy as defined in Table 8.9 for the example problem.

Table 8.11. The optimal operating policy and the probability of each state and decision.

initial storage S_k	flow Q_i	final storage S_1	interval indices k, i, l	time period	optimal allocation u_{kit}	d_{kit}	sum squared deviations SD_{kit}	joint probability P_{kit}
10	16	10	1,1,1	1	0.0	16.0	4996.0	0.0
10	47	10	1,2,1	1	0.0	47.0	1989.0	0.0
25	16	10	2,1,1	1	0.0	31.0	3301.0	0.2964706
25	47	10	2,2,1	1	6.0	56.0	1152.0	0.1270588
40	16	10	3,1,1	1	0.0	46.0	2056.0	0.1152941
40	47	25	3,2,2	1	6.0	56.0	1152.0	0.4611765
							sum =	**1.0**
10	47	25	1,2,2	2	0.0	32.0	3204.0	0.2851765
10	95	25	1,3,2	2	7.5	57.5	1012.5	0.2536471
25	47	25	2,2,2	2	0.0	47.0	1989.0	0.1383529
25	95	40	32,3	2	15.0	65.0	450.0	0.3228235
40	47	40	3,2,3	2	0.0	47.0	1989.0	0.0
40	95	40	3,3,3	2	22.5	72.5	112.5	0.0
							sum =	**1.0**

E020829k

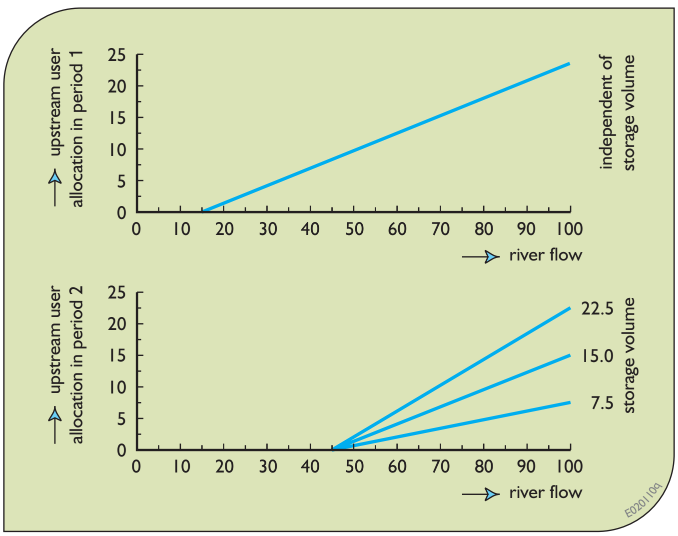

Figure 8.12. Upstream user allocation policies. In Period 1 they are independent of the downstream initial storage volumes. In Period 2 the operator would interpolate between the three allocation functions given for the three discrete initial reservoir storage volumes.

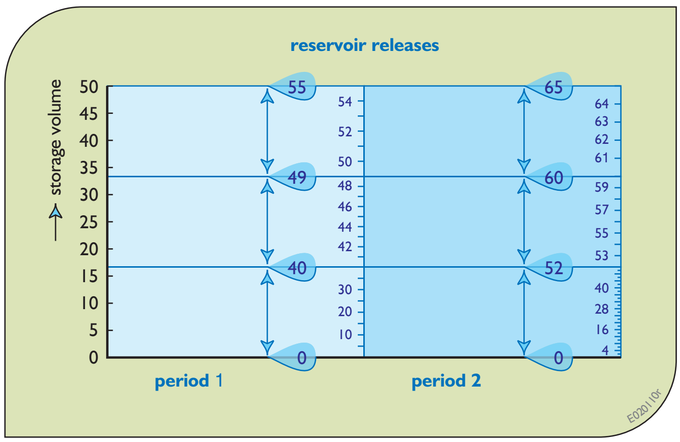

Figure 8.13. Reservoir release rule showing an interpolated release, increasing as storage volumes increase.

Alternatively, in each period t one can solve the model defined by Equation 8.37 to obtain the best decision for the current and a sequence of future periods, taking into account all current information regarding the objectives and possible inflow scenarios and their probabilities. The actual release decision in the current period can be the expected value of all these releases in this current period. At the beginning of the next period, the model is updated with respect to current initial storage and inflow scenarios (as well as any changes in objectives or other constraints) and solved again. This process continues in real time. This approach is discussed in Tejada-Guibert et al. (1993) and is the current approach for providing release advice to the Board of Control that oversees the releases from Lake Ontario that govern the water levels of the lake and the St. Lawrence River.

These policies, and modifications of them, can be simulated to determine improved release rules.

7. Conclusions

This chapter has introduced some approaches for including risk in optimization and simulation models. The discussion began with ways to obtain values of random variables whose probability distributions are known. These values, for example streamflows or parameter values, can be inputs to simulation models. Monte Carlo simulation involves the use of multiple simulations using these random variable values to obtain the probability distributions of outputs, including various system performance indicators.

Two methods were reviewed for introducing random variables along with their probabilities into optimization models. One involves the use of chance constraints. These are constraints that must be met, as all constraints must, but now with a certain probability. As in any method there are limits to the use of chance constraints. These limitations were not discussed, but in cases where chance constraints are applicable, and if their deterministic equivalents can be defined, they are probably the only method of introducing risk into otherwise deterministic models that do not add to the model size.

Alternatively, the range of random variable values can be divided into discrete ranges. Each range can be represented by a specific or discrete value of the random variable. These discrete values and their probabilities can become part of an optimization model. This was demonstrated by means of transition probabilities incorporated into both linear and dynamic programming models.

The examples used in this chapter to illustrate the development and application of stochastic optimization and simulation models are relatively simple. These and similar probabilistic and stochastic models have been applied to numerous water resources planning and management problems. They can be a much more effective screening tool than deterministic models based on the mean or other selected values of random variables. But sometimes they are not. Clearly if the system being analysed is very complex, or just very big in terms of the number of variables and constraints, the use of deterministic models for a preliminary screening of alternatives prior to a more precise probabilistic screening is often warranted.

8. References

FABER, B.A. and STEDINGER, J.R. 2001. Reservoir optimization using sampling SDP with ensemble streamflow prediction (ESP) forecasts. *Journal of Hydrology*, Vol. 249, Nos. 1–4, pp. 113–33.

GABLINGER, M. and LOUCKS, D.P. 1970. Markov models for flow regulation. *Journal of the Hydraulics Division, ASCE,* Vol. 96, No. HY1, pp. 165–81.

HUANG, W.C.; HARBO, R. and BOGARDI, J.J. 1991. Testing stochastic dynamic programming models conditioned on observed or forecast inflows. *Journal of Water Resources Planning and Management*, Vol. 117, No.1, pp. 28–36.

KIM, Y.O.; PALMER, R.N. 1997. Value of seasonal flow forecasts in bayesian stochastic programming. *Journal of Water Resources Planning and Management*, Vol. 123, No. 6, pp. 327–35.

LOUCKS, D.P. and FALKSON, L.M. 1970. Comparison of some dynamic, linear, and policy iteration methods for reservoir operation. *Water Resources Bulletin*, Vol. 6, No. 3, pp. 384–400.

STEDINGER, J.R.; SULE, B.F. and LOUCKS, D.P. 1984. Stochastic dynamic programming models for reservoir operation optimization. *Water Resources Research*, Vol. 20, No. 11, pp. 1499–505.

SU, S.Y. and DEININGER, R.A. 1974. Modeling the regulation of lake superior under uncertainty of future water supplies. *Water Resources Research*, Vol. 10, No. 1, pp. 11–22.

TEJADA-GUIBERT, J.A.; JOHNSON, S.A. and STEDINGER, J.R. 1993. Comparison of two approaches for implementing multi-reservoir operating policies derived using dynamic programming. *Water Resources Research*, Vol. 29, No. 12, pp. 3969–80.

TEJADA-GUIBERT, J.A.; JOHNSON, S.A. and STEDINGER, J.R. 1995. The value of hydrologic information in stochastic dynamic programming models of a multi-reservoir system. *Water Resources Research*, Vol. 31, No. 10, pp. 2571–9.

YAKOWITZ, S. 1982. Dynamic programming applications in water resources. *Water Resources Research*, Vol. 18, No. 4, pp. 673–96.

Additional References (Further Reading)

BASSSON, M.S.; ALLEN, R.B.; PEGRAM, G.G.S. and VAN ROOYEN, J.A. 1994. *Probabilistic management of water resource and hydropower systems*. Highlands Ranch, Colo., Water Resources Publications.

BIRGE, J.R. and LOUVEAUX, F. 1997. *Introduction to stochastic programming*. New York, Springer-Verlag.

DEGROOT, M.H. 1980. *Optimal statistical decisions*. New York, McGraw-Hill.

DEMPSTER, M.A.H. (ed.). 1980. *Stochastic programming*. New York, Academic Press.

DORFMAN, R.; JACOBY, H.D. and THOMAS, H.A. Jr. 1972. *Models for managing regional water quality*. Cambridge, Mass., Harvard University Press.

ERMOLIEV, Y. 1970. *Methods of stochastic programming*. Moscow, Nauka. (In Russian.)

ERMOLIEV, Y. and WETS, R., (eds.). 1988. *Numerical techniques for stochastic optimization*. Berlin Springer-Verlag.

HEYMAN, D.P. and SOBEL, M.J. 1984. *Stochastic models in operations research. Vol. 2: Stochastic optimization*. New York, McGraw-Hill.

HILLIER, F.S. and LIEBERMAN, G.J. 1990. *Introduction to stochastic models in operations research*. New York, McGraw-Hill.

HOWARD, R.A. 1960. *Dynamic programming and Markov processes*. Cambridge, Mass., MIT Press.

HUFSCHMIDT, M.M. and FIERING, M.B. 1966. *Simulation techniques for design of water-resource systems*. Cambridge, Mass., Harvard University Press.

LOUCKS, D.P.; STEDINGER, J.S. and HAITH, D.A. 1981. *Water resource systems planning and analysis*. Englewood Cliffs, N.J., Prentice Hall.

MAASS, A.; HUFSCHMIDT, M.M.; DORFMAN, R.; THOMAS, H.A. Jr.; MARGLIN, S.A. and FAIR, G.M. 1962. *Design of water-resource systems*. Cambridge, Mass., Harvard University Press.

PRÉKOPA, A. 1995. *Stochastic programming*. Dordrecht, Kluwer Academic.

RAIFFA, H. and SCHLAIFER, R. 1961. *Applied statistical decision theory*. Cambridge, Mass., Harvard University Press.

REVELLE, C. 1999. *Optimizing reservoir resources*. New York, Wiley.

ROSS, S.M. 1983. *Introduction to stochastic dynamic programming*. New York, Academic Press.

SOMLYÓDY, L. and WETS, R.J.B. 1988. Stochastic optimization models for lake eutrophication management. *Operations Research*, Vol. 36, No. 5, pp. 660–81.

WETS, R.J.B. 1990 Stochastic programming. In: G.L. Nemhauser, A.H.G. Rinnoy Kan and M.J. Todd (eds.), *Optimization: handbooks in operations research and management science*, Vol. 1, pp. 573–629. Amsterdam, North-Holland.

WURBS, R.A. 1996. *Modeling and analysis of reservoir system operations*. Upper Saddle River, N.J., Prentice Hall.