Risk-based performance criteria for real-time reservoir operation

DONALD H. BURN, HENRY D. VENEMA, AND SLOBODAN P. SIMONOVIC

Department of Civil Engineering, The University of Manitoba, Winnipeg, Man., Canada R3T 2N2

Received February 5, 1990
Revised manuscript accepted June 21, 1990

Reliability, resiliency, and vulnerability criteria are formulated as risk-based performance measures for the evaluation of a real-time reservoir operation model. The reservoir operation model includes a multi-objective compromise programming algorithm to select, in real time, an optimal operating horizon for the reservoir operation. The utility of the risk-based performance criteria for comparing operational strategies resulting from the selection of different parameters for the compromise programming algorithm is demonstrated. Although trade-offs exist between the performance evaluators, it is shown that appropriate compromises can be reached between the conflicting modeling goals.

Key words: reservoir operation, real time, risk, optimization.

Introduction

Water resource projects have traditionally been designed and operated with the objective of optimizing an economic performance measure (i.e., maximizing net benefits, minimizing expected damages, etc.). More recently, the importance of additional performance measures in the design or operation of water resource systems has been recognized. Work by Fiering (1982a, b) and Hashimoto et al. (1982) has introduced the concepts of reliability, resiliency, and vulnerability into the water resources literature. These three concepts can be formulated mathematically and used to evaluate the performance of a water resource system. In this paper, reliability, resiliency, and vulnerability are jointly referred to as risk-based performance criteria. This terminology reflects the probabilistic nature of water resource system behaviour as a result of nondeterministic inputs and influences that are germane for virtually all real-world systems. The inclusion of risk-based performance criteria, in addition to purely economic evaluators, is vital, given the stochastic nature of the demands on water resource systems as well as the extreme variability that characterizes the hydrologic inputs.

Following the works of Fiering (1982a, b) and Hashimoto et al. (1982), there have been several applications of risk-based performance criteria for different problem areas in the water resources field. Datta and Houck (1984) describe the use of risk-based performance criteria within the context of short-term reservoir operation. They used risk-based performance criteria to compare different operating strategies reflecting the relative importance of deviations from targets for reservoir release versus reservoir storage target deviations. Their model used generated forecasts with specified error characteristics as input to a reservoir operation model. Weeraratne et al. (1986) employed reliability, resiliency, and vulnerability measures to evaluate reservoir release policies for low flow augmentation. The release policies corresponded to different target flow levels at critical points in the system. Moy et al. (1986) investigated the trade-offs between reliability, resiliency, and vulnerability, in the context of reservoir operation for water supply, using a multi-objective mathematical programming model.

The intent in the present paper is to explore potential risk-based performance criteria formulations for a novel real-time reservoir operation model (Burn and Simonovic 1988; Simonovic and Burn 1989). The reservoir operation model is unique in that the operating time horizon is an explicit real-time decision variable in the modeling package. The present work arose as part of a larger study investigating the characteristics of the above-noted modeling framework and examining the potential for linking the reservoir operation model to an expert system module. Risk-based performance criteria were sought within this exercise to assist in the evaluation of the operating policies defined by the system. Reliability, resiliency, and vulnerability will thus be used as tools to assist in the adjustment of the parameters of the real-time reservoir operation model in order to enhance the resulting operation.

One of the original contributions in the present work is in the formulation of reliability, resiliency, and vulnerability measures for a real-time reservoir operation model. Hashimoto et al. (1982) noted that all planning problems have unique characteristics that will likely necessitate the definition of appropriate, problem-specific, risk-based performance measures. As such, the measures defined by Hashimoto et al. (1982) are used herein as a point of departure for the definition of relevant criteria. The work described in this paper further differs from previous uses of reliability, resiliency, and vulnerability in that actual forecasts of reservoir inputs, obtained from real-world data, are used to drive the reservoir operation model, which is evaluated using the risk-based performance criteria.
Real-time reservoir model

Simonovic and Burn (1989) present an improved real-time reservoir operation model that explicitly utilizes the trade-off between flow forecast reliability and the performance of the reservoir operation, which is conditioned on the forecasted flows. The model package includes (i) a forecasting algorithm, (ii) a real-time reservoir operation model, and (iii) a multi-objective compromise programming algorithm. A Kalman filtering technique provides real-time forecasts with lead times from 1 to T time intervals and a forecast error variance associated with each forecast period. The forecast error variance functions as a measure of the reliability of the forecast. The reservoir optimization model is then solved for each of the T operating horizons. The optimal amount of water to be released during each time interval, in order to minimize deviations from storage and release targets (and thus minimize penalties), is determined.

The real-time reservoir operation model, imbedded in the modelling framework, employs a linear programming optimization module with the inflow forecasts acting as inputs. This module determines the reservoir releases that minimize the losses resulting from deviations from storage and release targets subject to constraints on minimum and maximum release and storage. The explicit information quantifying the losses associated with deviations from storage and release targets is in the form of piecewise linearized penalty functions, derived by considering the intended uses and operating goals of the particular reservoir (Simonovic and Burn 1989).

The multi-objective compromise programming algorithm has the dual objective of the maximization of inflow forecast reliability (represented by minimization of forecast error variance) as well as the minimization of the penalties associated with the reservoir operation. The optimal decision variable for the multi-objective algorithm is the length of the operating horizon, \( T_0 \), in the range of 1 to T time intervals, which represents the best compromise between the two objectives. The reservoir release suggested by the reservoir operation model for the optimal operating horizon is then implemented for the next time interval. The entire procedure is subsequently repeated for the following time interval when a new set of forecasts are obtained based upon the latest hydrologic information. Further details on the forecasting model development are presented in Burn and Simonovic (1988), and the complete model package is described by Simonovic and Burn (1989).

The compromise programming algorithm explicitly utilizes the trade-off between the forecast uncertainty (generally increasing with the length of the forecast period) and reservoir operation penalties (normally decreasing with the length of the reservoir operation horizon). The first objective in the compromise programming algorithm is the minimization of forecast error variance:

\[
\begin{align*}
\text{minimize} & \quad \text{FV}(	au) \\
\{ \tau \}
\end{align*}
\]

where \( \text{FV}(\tau) \) is the inflow forecast error variance for lead time \( \tau \), and \( \tau \) is the length of the operating horizon. The second objective is the minimization of a measure of the total penalties associated with the reservoir operations:

\[
\begin{align*}
\text{minimize} & \quad L(\tau) \\
\{ \tau \}
\end{align*}
\]

where \( L(\tau) \) is the reservoir penalty measure for operating horizon \( \tau \). The decision variable is the length of the operating horizon to be employed and the problem is formulated, in general, as follows:

\[
D_p(\tau) = \left[ \frac{\text{FV} - \text{FV}^{\tau}}{\text{FV}^{\tau}} \right]^p + \left[ \frac{L^\tau - L(\tau)}{L^\tau} \right]^q
\]

where \( D_p(\tau) \) is the distance from the ideal point, to be minimized; \( w_1 \) and \( w_2 \) are the weights assigning the relative importance to the two objectives such that \( w_1 + w_2 = 1 \); \( \text{FV}^{\tau} \) and \( L^\tau \) are the individual optimal values of \( T \) possible for each objective; \( \text{FV}^{\tau} \) and \( L^\tau \) are the individual worst values obtainable of \( T \) possible for each objective; \( \text{FV}(\tau) \) and \( L(\tau) \) are the results of implementing decision \( \tau \) with respect to each objective; \( \rho \) is a parameter in the range of 1 to \( \infty \); and \( \tau \) is the length of operating horizon (the decision variable).

The formulation of the multi-objective compromise programming algorithm, presented above, requires that weights be assigned to the two objectives reflecting the relative importance of each. This information is embodied in the parameters \( w_1 \) and \( w_2 \). As well, a parameter reflecting the significance of the maximal deviation from the “ideal point” must be determined. The value chosen for this parameter, \( \rho \), was not found to greatly influence the results. As can be inferred from the above formulation, the optimization problem thus obtained is discrete and is therefore evaluated for each of the \( T \) possible operating horizons. The decision variable, \( \tau \), is chosen on the basis of which value yields the minimal weighted distance, \( D_p \), to the ideal point. A practical upper bound for candidate operating horizons can be expected to arise from the limits of the forecast lead time.

The determination of appropriate relative values to assign to the weights for the two objectives in the compromise programming algorithm is an area of ongoing research. The authors intend to pursue this topic using an expert system module to aid in the selection process. A prerequisite for this activity is the development of performance measures that can be calculated for different weighting values and used to assist in the evaluation of the resulting operational strategies. Thus, alternative operational strategies can be viewed as corresponding to the real-time reservoir operation resulting from different relative weights applied to forecast reliability versus operational penalties. The intent in this paper is thus to examine risk-based performance criteria that exhibit a trend as a function of the value of the weightings.

The modeling package, outlined above, calculates a cumulative total of the penalties incurred for an entire operating season as a result of the deviations from storage and release targets in each time interval. The cumulative penalty cost was the primary criterion used to illustrate the advantage of explicitly seeking the optimum operating horizon reflecting the best compromise between forecast reliability and reservoir penalty for the \( T \) possible operating horizons (Simonovic and Burn 1989). The intent herein is to develop additional, risk-based, criteria to assist in the evaluation of reservoir operation performance. The approach taken can be regarded as an indirect means of obtaining reservoir operating policies that are reliable, resilient to failure, and comparatively invulnerable to failure. Although the risk-based attributes are important goals for reservoir operation, we feel that it is unlikely that these can be included directly into the objective function of an optimization model owing to the real-time nature of the modeling procedure.
Risks-based performance criteria

Hashimoto et al. (1982) introduced the concept of evaluating water resource system performance in terms of reliability, resiliency, and vulnerability. Reliability measures the likelihood of system failure; resiliency quantifies how quickly the system recovers from failure, and vulnerability describes the severity of failure. These performance criteria were introduced to quantify the operational risks inherent in a water resource system operation, that risk primarily involving the failure of the system to perform in a satisfactory state.

Reliability

The principal risk criterion is the probability of occurrence of an unsatisfactory, or failure, state. Reliability is then defined as the opposite of risk, simply one minus the risk. Thus, the probability of the system residing in a satisfactory state is the reliability. The formulation of this criterion is straightforward for the real-time reservoir operation model, as it merely involves evaluating, for every time interval of the operational period, whether or not the system is in a satisfactory state. However, what constitutes a satisfactory state, in the context of the real-time reservoir operation model, had to be established. Utilizing the explicit penalty functions for storage and release in the reservoir operation optimization module, a satisfactory state was defined as occurring when both release and storage levels were within specified ranges that bounded the release and storage target values. The demarcation of these ranges, defining a satisfactory state, can be effected by considering the goals of the operation of the reservoir system (i.e., flood control versus recreation, etc.). A failure state was thus defined to occur when either the storage or the release was outside the respective acceptable ranges. Reliability thus equals the number of non-failure time intervals divided by the number of time intervals in the operational period analyzed, resulting in

$$\alpha = \frac{1}{NS} \sum_{t=1}^{NS} Z_t$$

with

$$Z_t = 1 \quad \text{for} \quad X_t \in S$$

$$Z_t = 0 \quad \text{for} \quad X_t \in F$$

where $X_t$ is the state of the system in time interval $t$, $S$ and $F$ denote satisfactory and failure states, respectively, and $NS$ is the duration of the operating season. $X_t$ corresponds to a satisfactory state if the release and storage are both in the acceptable range and a failure occurs if either variable is outside the acceptable range.

As Moy et al. (1986) and Hashimoto et al. (1982) observed, the disadvantages of using only this risk criterion lie in the implicit assumptions that (1) all failures are of equal significance and are not differentiated on the basis of magnitude and consequence, and (2) failures are independent when in fact consecutive failures may be more severe than failures separated by periods where the system resides in a satisfactory state. To characterize other dimensions of a reservoir’s operational risk, other criteria describing the severity of failure, such as resiliency and vulnerability, are required.

Resiliency

Holling (1973) introduced the concept of resiliency to describe the ability of a multi-species ecological system to undergo a wide variability of species population after some severe shock and still persist with the same basic structure. Hashimoto et al. (1982) extended the concept to water resource system planning. In this context, resiliency was defined as how quickly the system returns to a satisfactory state, given that a failure has occurred.

An initial formulation of the resiliency criterion, as applied to the real-time reservoir operation model performance evaluation, was developed from the above definition. Following Hashimoto et al. (1982), a transition from a satisfactory state to a failure state can be characterized as

$$W_t = \begin{cases} 1 & \text{if } Z_t = 1 \text{ and } Z_{t+1} = 0 \\ 0 & \text{otherwise} \end{cases}$$

The average duration of a sojourn into a failure state is

$$T_f = \frac{A}{B}$$

with

$$A = \sum_{t=1}^{NS} (1 - Z_t)$$

and

$$B = \sum_{t=1}^{NS} W_t$$

where $A$ is the total time in $F$, $B$ is the number of times the system went into $F$, and $NS$ is the duration of the operational season. In the long run, as $NS$ approaches infinity, $T_f$ is the average number of days a failure is expected to persist once it has occurred. In this initial formulation of the resiliency measure, the inverse of $T_f$ is the system’s average recovery rate and is defined as the system’s resiliency, giving

$$\gamma_1 = \frac{1}{T_f}$$

Moy et al. (1986) proposed a simplified version of the resiliency criterion based solely on the maximum number of consecutive days of failure. The shorter the number of consecutive failure days, the more resilient is the system. The second formulation of the resiliency criterion utilized is simply the following:

$$\gamma_2 = \frac{1 - \text{MD}}{\text{NS}}$$

where $\text{MD}$ is the maximum number of consecutive days of failure in a season.

A third and original formulation of the resiliency criterion was also used to evaluate the model performance. Incorporating elements of both of the previously described formulations, this measure is as follows:

$$\gamma_3 = \frac{1}{\left( \frac{\text{MD}}{\text{NS} - \text{MD}} \right) \text{NF}}$$

with $\text{NF}$ the number of times the system enters a failure state in a season. Formulating resiliency in this manner can be justified intuitively. Consider, for example, the simplified measure proposed by Moy et al. (1986). This measure does not consider a system experiencing a 20 day duration failure, in a season with a total of five failures, any less resilient than a system that experiences only one 20 day failure in the entire
season. By introducing the factor of the number of failures occurring in a season, an appropriate weight can be assigned to the significance of the largest failure in the season. Since the number of failures in a season is obviously dependent on the length of the season, it is seemingly appropriate to retain the length of the season as a scaling factor. Taking the inverse of this weighted failure duration parameter is then a potentially interesting metric for defining resiliency.

**Vulnerability**

The final operational risk criterion considered is system vulnerability. This measure attempts to quantify the severity of the consequences of failure. Holling (1978) discussed the concept of endeavoring to design a system to be safe in the event of failure, or “safe-fail,” rather than “fail-safe.” Since the probability of failure can rarely be entirely eliminated, effort should be more wisely expended in attempting to minimize the consequences of failure.

Hashimoto et al. (1982) present a mathematical index of vulnerability requiring that, for each discrete failure state, the probability that the state is the most severe outcome of the set of states corresponding to that failure sojourn be assigned. The overall vulnerability is then defined as the expected maximum severity of a failure. A modified version of this vulnerability measure is employed herein. The penalty associated with the worst-case failure time interval from every failure sojourn (defined as a sequence of time intervals where deviancy from storage and (or) release targets occurred) became the vulnerability measure. The total vulnerability was calculated as the cumulative penalty for the most highly penalized time interval from each discrete failure sojourn over the entire season. Vulnerability can then be defined as

\[ V = \sum P^k \]

where \( P^k \) is the maximum penalty (unweighted) for a time interval in failure sojourn \( k \), and the summation is over all failure sojourns in the operating period.

Moy et al. (1986) offer a simplified vulnerability measure. The criterion used is simply the magnitude of the worst-case failure. The second formulation of vulnerability, in the context of the reservoir operation model, was thus the penalty cost of the consequences of failure requiring that, for each discrete failure sojourn over the entire season. Vulnerability can then be defined as

\[ V = \text{maximum} \{ P^k \} \]

where \( P^k \) is the maximum penalty (weighted) for a time interval in failure sojourn \( k \), and the summation is over all failure sojourns in the operating period.

**Modeling results**

**Reliability, resiliency, and vulnerability**

Four different historical streamflow and precipitation records (winter/summer 1967 and winter/summer 1968) were examined for the Green River basin using the real-time reservoir operation package (Simonovic and Burn 1989). For illustrative purposes, results for the winter 1967 season are presented herein. This season serves as a basis for examining the risk-based performance criteria formulated above. If another reservoir system or a different time period had been investigated, it is anticipated that the numerical results would differ, but the conclusions from the analysis would most likely still hold.

The model was executed repeatedly for the 110-day winter season, based on the flow forecasts, using the parameters \( p \), \( w_1 \), and \( w_2 \) in the multi-objective compromise programming algorithm as input variables. Since \( w_1 \) and \( w_2 \) are explicitly related (\( w_1 + w_2 = 1 \)), the decision space for these parameters is really only two-dimensional. In addition, although \( p \) could range from one to infinity, a previous heuristic search had shown that \( p \) values of 2, 4, and 200 were generally representative of the effects of this variable. Furthermore, in the context of the risk-based performance criteria, it was found that the trends in the evaluators were relatively insensitive to the \( p \) value selected. Therefore, for a value of \( p = 2 \) (a representative value), \( w_1 \) was varied between 0 and 1. About 15 points generally sufficed to depict any obvious trends present in the evaluation criteria.

Reliability was observed to display a very dramatic increase with an increase in \( w_1 \) to a plateau level, as shown in Fig. 1. Figure 1 shows all three performance measures, as described below, as a function of the weighting value, \( w_1 \). The performance measures are plotted on a normalized scale from 0 for the worst value attained on the measure to 1 for the best value. The curves plotted represent smoothed curves depicting the trend obtained from 15 individual data points. The curve for the reliability plot, displayed in Fig. 1, represents an entirely satisfactory performance measure in that it allows an explicit evaluation of the operating strategies embodied in different \( w_1 \) values.

The first formulations of the resiliency and vulnerability criteria produced results that were very difficult to interpret. Resiliency, as defined by \( y_1 \), showed a pattern as a function of \( w_1 \) that was so erratic that it could almost be characterized as pure noise (see Fig. 2). A plausible explanation for the observed behaviour may be that for this formulation, NS is not sensitive to the performance of a small number of days of reservoir operation as effected by weighting in the multi-objective compromise programming algorithm. If NS were
much larger, small fluctuations in the value of $B$ would not be as significant in determining resiliency. An alternate measure for resiliency was therefore examined.

The second formulation of resiliency, $\gamma_2$, was adapted from Moy et al. (1986). For this formulation, resiliency showed an entirely different behaviour with virtually no variation with $w_1$ at all (see Fig. 2). Since this provides very little insight into the system performance and is not at all useful for identifying preferred operating strategies, the third measure of resiliency was explored.

The third formulation of resiliency, $\gamma_3$, is an independent hybrid of the concepts presented by Hashimoto et al. (1982) and Moy et al. (1986). This measure produced results much more easily interpreted, as is shown in the plot of this measure in Figs. 1 and 2. This resiliency metric shows a definite upward trend with increases in $w_1$, to a two-stage plateau in the $w_1 = 0.3$ to 0.5 region. The results obtained using this measure of resiliency lend themselves to the development of general guidelines for selecting $w_1$.

Vulnerability, as measured by $v_1$, displays perhaps even more disturbing behaviour than was the case for the first resiliency measure (see Fig. 3). Although this measure is quite erratic, if any trends were to be inferred, they would be contradictory for the 2 years studied. The point of the larger study was, however, to obtain consistency in the trends of the evaluation criteria, regardless of the particular year studied, with the aim of developing general guidelines for the selection of $p$ and $w_1$. Thus, an alternate version of vulnerability for the reservoir operation model was attempted.

Vulnerability, as measured by $v_2$, showed a definite trend as $w_1$ was varied. This is shown in Figs. 1 and 3. It should be noted that while reliability and resiliency are both criteria that should be maximized, optimal vulnerability measures correspond to the minimum values attainable. Thus, a normalized vulnerability value of 1.0 corresponds to the lowest vulnerability score, indicating the best vulnerability response. Figure 1 does, therefore, portray trade-offs amongst the risk-based performance measures for different operating strategies.

**Discussion**

Hashimoto et al. (1982) presented observed relationships between reliability, resiliency, and vulnerability as a function of an arbitrary loss function shape parameter, $\beta$. For small values of $\beta$, Hashimoto et al. (1982) observed that high vulnerability corresponded to high reliability and resiliency. Although the problem application used different formulations of the risk-based performance criteria than were employed herein, the general relational characteristics of the criteria were similar to those obtained in this study.

From Fig. 1, it can be seen that the global minima of reliability, resiliency, and the global maxima of vulnerability all occur at low values of $w_1$. The observed trend for all of the criteria can be rationalized in a manner logically consistent with the multi-objective programming algorithm mechanism. Since $w_1$ weights the significance of forecast error variance, it is consistent to have the lowest reliabilities and resiliencies when one does not emphasize the forecast reliability in the compromise programming algorithm. Failures will tend to be frequent and may be relatively lengthy because the releases are often based on forecasts with comparatively high uncertainty. Deviations from the targets for reservoir release and storage are, however, generally slight because the minimization of operational penalties objective is emphasized. Vulnerability costs are, therefore, small and the normalized vulnerability measure is comparatively large. For high values of $w_1$, forecast reliability is emphasized in the compromise programming algorithm and the high reliability and resiliency achieved reflect this. Since the significance of operational penalties is downplayed with this weighting, large vulnerability costs can be incurred when the operating schedule is preferentially chosen to minimize uncertainty at the expense of penalty costs.
The existence of trade-offs amongst the three risk-based performance criteria was identified from the results in Fig. 1. The trade-offs imply that all three criteria cannot simultaneously attain optimal values. The nature of the existing trade-offs can be illustrated by plotting coincident values for the performance criteria in pairs, as is done in Figs. 4—6. Each of these curves displays points in a two-dimensional performance space corresponding to different values of the weighting function parameter, \( w_i \). The solid symbols in the diagrams signify observed performance value pairs, while the line segments joining the points in sequence are primarily for illustrative purposes. It should be noted that while there were 15 different \( w_i \) values selected, there are less than 15 points in each diagram because some points represent more than one \( w_i \) value (i.e., changing the weighting values would not always change all of the performance measures).

Figure 4 portrays the trade-off between reliability and resiliency. As can be seen from the graph, there is not a distinct conflict between reliability and resiliency. The best value obtained for both of these measures occurs at the same weighting value, \( w_i = 0.7 \), although this point is clearly suboptimal for the vulnerability criterion. Although there is not a traditional conflict between reliability and resiliency, the diagram in Fig. 4 does, nonetheless, elucidate inefficient combinations of these performance measures that should be avoided.

Figure 5, representing the plot for reliability and vulnerability, represents a classical trade-off curve. Both criteria cannot be optimized simultaneously, and a compromise must be reached between these two conflicting measures. There are clearly inefficient combinations of the two measures such that only a limited number of combinations (probably 3) would have to be considered if these were the only two performance measures of relevance.

The final trade-off curve, Fig. 6, shows a plot of resiliency versus vulnerability. This curve also demonstrates a classical trade-off relationship, although there are more points that would have to be considered as viable solutions if this were a two performance measure trade-off problem.

The curves presented in Figs. 4—6 represent two-dimensional portrays of a three-dimensional trade-off surface. As such, all three curves must be considered in conjunction in order to identify agreeable compromise points. However, as can be deduced from the discussion above, the set of points under consideration is not large.

Conclusions

Reliability, resiliency, and vulnerability criteria for a real-time reservoir operation model have been formulated to compare alternate operating strategies embodied in weightings assigned to forecast reliability and reservoir penalties. The risk-based performance criteria demonstrated trends, as a function of the weighting parameter value, which results in a useful set of tools for the evaluation of the real-time reservoir model performance. This research has demonstrated the need for the identification of problem-specific performance criteria. This result is in concurrence with the observations of Hashimoto et al. (1982).

The existence of trade-offs between the three risk-based performance criteria examined herein implies that the selection of a preferred operating strategy is not a trivial exercise. Any rational comparison of candidate operating strategies must therefore consider the multi-dimensional character of the operational problem, within the selection process. The trade-offs between
the evaluators presented indicate that agreeable zones where an appropriate compromise may be realizable are obtainable. The implication of this latter result is that preferred operating strategies, in terms of the risk-based performance criteria, can be selected in a systematic manner. Further work is required to formalize a procedure for this.

**References**


