Statistical performance indices for a hydropower reservoir
Sharad K. Jain

ABSTRACT
Commonly employed indices to assess the performance of a storage reservoir include reliability, resilience and vulnerability. Depending upon the way in which resilience and vulnerability are computed, different values and behaviour of these are obtained. This study has examined the behaviour of statistical performance indices for a hydropower reservoir located in a semi-arid region with the aim to identify those indices which give correct and complete insight of the behaviour of the reservoir. A detailed Monte Carlo simulation analysis was carried out using synthetic data to investigate the behaviour of a number of performance measures. Uncertainty of the indices was also assessed using synthetic data. While considering the sensitivity of measures, their monotonic behaviour and uncertainty of estimates, use of time and volume reliabilities and vulnerability were found to give adequate information about the performance of the project.

Key words | data generation, hydropower, Monte Carlo, reliability, resilience, vulnerability

INTRODUCTION
Time and volume reliabilities have been traditionally employed to evaluate performance of a reservoir. Hashimoto et al. (1982) advanced two more criteria: resilience and vulnerability to measure the features that are not evaluated by reliability alone. Subsequently, these criteria were used in many studies of reservoir performance such as Moy et al. (1986), Kundzewicz & Kindler (1995), Kundzewicz & Laski (1995), Vogel & Bolognese (1995), Srinivasan et al. (1999) and Vogel et al. (1999). To compute the performance indices, the basic reservoir operation data employed includes the volume of water demanded and supplied for all the time periods, number of failure periods and extent of failure in terms of its length and severity. A month or year is the most commonly used time period in studies, particularly those related to planning.

Kjeldsen & Rosbjerg (2004) explored the monotonic behaviour of reliability, resilience and vulnerability. They also addressed issues such as overlap and correlation between the estimators using synthetically generated data. Subsequently, Jain & Bhunya (2008) examined the behaviour of reliability, resilience and vulnerability for a multipurpose storage reservoir from a probabilistic interpretation of their variation. It was found that when inflows generated using long-memory models were input in simulation, there were large variations in reliability, resilience and vulnerability among the runs. In contrast, when data from short-memory models were used, the indices were confined to a narrow band.

Many indicators of reservoir performance are present and it is likely that a decision maker may use some of them without fully knowing their characteristic behaviour and may be misled. The objective of this paper is to explore the behaviour of statistical measures of performance of a hydropower reservoir such as reliability, resilience and vulnerability to identify the indices that may be used for evaluation purposes. The paper also examines uncertainty associated with these indices.

doi: 10.2166/nh.2009.072
SIMULATION EXPERIMENTS

Let the storage capacity of the reservoir be $S_{\text{max}}$. A time series of monthly inflows is available; the monthly demands are also known. The reservoir is operated following the standard linear operating policy (SLOP), graphically represented in Figure 1. Let $A_t$ represent the available water, $D_t$ the target demand and $R_t$ release (both controlled and uncontrolled) from the reservoir, where subscript $t$ represents the month. Here $A_t = S_t + I_t$ where $S_t$ represents storage content of the reservoir at the beginning of period $t$ which can vary over the range $[0, S_{\text{max}}]$. $I_t$ represents inflows during the period $t$. Mathematically, the SLOP can be expressed as:

\[
\text{If}\ A_t \leq D_t,\ R_t = A_t,
\]

\[
\text{If}\ D_t < A_t \leq S_{\text{max}} + D_t,\ R_t = D_t
\]

\[
\text{If}\ A_t > S_{\text{max}} + D_t,\ R_t = A_t - S_{\text{max}}
\]

The continuity equation for the reservoir is given by

\[
S_t + I_t - E_t - R_t = S_{t+1}
\]

where $E_t$ represents evaporation loss from the reservoir during month $t$. Evaporation loss $E_t$ is computed by multiplying the average reservoir surface area during a month with the normal depth of evaporation for the month. While operating the reservoir following SLOP, attempt is made to meet the demands to the extent possible. If after satisfying the demands and evaporation, the remaining quantity of water exceeds $S_{\text{max}}$ surplus water is spilled and the reservoir will be full at the end of the month.

Operation of a hydropower reservoir

Water is released from a hydropower reservoir to generate electricity. The amount of power generated is a function of discharge and the hydraulic head:

\[
P = 9.817QH\eta.
\]

where $P$ is the electric power generated (kW), $Q$ is the discharge through power plant ($\text{m}^3/\text{s}$), $H$ is the net head (m) and $\eta$ is the overall efficiency of the power plant expressed as a ratio. When there is no shortage of water and all the releases pass through the power plant, $Q$ will be equal to $D$ in Equation (1). Note that both $Q$ and $H$ are assumed to be constant during the period. Electric power generation depends on the volume of water passing through turbines and the effective head. Thus, the same amount of power can be produced either by releasing more water at a low head or less water at a high head. The amount of power generated over a time, or energy (expressed in kilowatt-hour or kW-hr), can be computed (Jain & Singh 2003) as:

\[
\text{KWHR} = 9.817QHT\eta.
\]

in which $\text{KWHR}$ is the hydropower generated during the period in kW-hr and $T$ is the length of period (hours). Of course, the generation of energy is limited by the installed capacity of the power plant. Further, if the values of $Q$ or $H$ are outside the specified range, the efficiency of the power plant drops rapidly and the machines cannot be operated beyond a threshold.

Performance of a hydropower reservoir may be measured in terms of volume of water supplied and required to generate the demanded amount of energy. It may also be measured in terms of the demand of energy and the quantity of energy generated. Since the volume of water supplied and energy generated are nonlinearly related, the two measures may not give identical results.

If, during a particular month, the quantum of energy generated or the volume of water supplied is less than the required quantity then a failure is said to have taken place. Depending upon the inflow and demand in the next month, the reservoir may continue to remain in the failure state or may switch to success state. When the reservoir enters the success state, the current failure state is said to be over. Clearly, a failure event has two variables: the number of months which the event lasts and the total quantity of shortfall during that event. Statistical indices that have been
developed to highlight different aspects of failure are described in the following.

STATISTICAL INDICES FOR PERFORMANCE EVALUATION

A reservoir that has success/failure in month \( t \) may remain in the same state in month \((t + 1)\) or may switch to the other state. Failures in the operation of a reservoir have many aspects: their number, extent, severity, etc. The commonly used indices to measure the performance of a reservoir are reliability, resilience and vulnerability (Hashimoto et al. 1982).

Reliability

Two reliability indices are generally followed in water resources management.

Time or occurrence-based reliability is the probability that the system state lies in the set of satisfactory states

\[ r_t = P[x(t) \in S] \]  

where \( P \) is the probability, \( x(t) \) is the system’s state in the given time period \( t \) and \( S \) is a set of satisfactory states. Time reliability can be estimated by

\[ r_t = 1 - (f_p/n); \quad 0 \leq r_t \leq 1, \quad f_p \leq n. \]  

where \( r_t \) is the estimate of time reliability and \( f_p \) is the number of failure periods out of the total periods, \( n \). Figure 2 gives a sketch of the variables involved in the computation of reliability indices.

Volume or quantity-based reliability \( r_v \) is computed as

\[ r_v = \frac{V_s}{V_d} \]  

where \( V_s \) is the volume of water supplied and \( V_d \) is the volume of water demanded during a given period. For a hydropower reservoir, \( V_s \) and \( V_d \) may also represent the quantity of energy generated and demanded. Note that volume reliability is not really a probabilistic quantity.

Let \( M \) be the total number of failure events. Moving from time step \( t \) to \((t + 1)\), the system can either remain in the same state or switch to the other state. The duration of the \( j \)th failure event is denoted by \( d_j \) and \( v_j \) is the corresponding deficit volume. The deficit volume of the failure event \( j \) is the sum of deficits for all the months of that event:

\[ v_j = \sum_{i=1}^{d_j} [D_i - R_i] \]  

Resilience

A desirable characteristic of a system would be a rapid return to a satisfactory state after a failure. Resilience describes how quickly a system is likely to recover from failure. If failures are prolonged events and system recovery is slow vis-à-vis its natural characteristics, the system design may have a serious flaw. Let \( S \) denote the set of satisfactory states and \( F \) the set of all unsatisfactory states. The resilience \( \gamma \) can then be defined as the average probability of a recovery from the failure set in a single time-step (Hashimoto et al. 1982):

\[ \gamma = P[X_{t+1} \in S | X_t \in F] = \frac{P[X_t \in F \text{ and } X_{t+1} \in S]}{P[X_t \in F]}. \]  

Alternatively, resilience may also be defined as the inverse of the mean value of the time which the system spends in an unsatisfactory state. Mean resilience may be computed as:

\[ \gamma_{\text{mean}} = \left( \sum_{i=1}^{M} d_i \right)^{-1}. \]  

Kjeldsen & Rosbjerg (2004) used another definition of resilience as the inverse of maximized duration \( d_j \).
Hence, the maximum resilience can be computed as:

\[ \gamma_{\text{max}} = \left[ \max_j \{d_j\} \right]^{-1}. \]  \hspace{1cm} (11)

According to Kundzewicz & Kindler (1995), the estimation of resilience based on maximum value is better than the mean value estimation because small insignificant events may lower the mean value. In a mixed integer programming model to optimize the operation of a water supply reservoir during critical periods, Srinivasan et al. (1999) placed more emphasis on average values where the maximum duration of continuous shortfall is likely to be minimal. Kjeldsen & Rosbjerg (2004) compared these two estimates of resilience along with the estimate using the 0.9th fractile of the empirical cumulative distribution function (CDF) of failure duration and deficit volume and concluded that no recommendation can be made based on an investigation of the degree of monotonic behaviour alone. If the values of a performance measure continuously increase or decrease with the independent variable, the measure is said to exhibit a monotonic behaviour. They advocated the use of long series of synthetic data to obtain robust estimates of resilience and vulnerability. A monotonic estimator either increases or decreases consistently with respect to the dependent variable. For example, if yield of a reservoir consistently increases with the storage volume, it will be termed as a monotonic estimator.

Vulnerability

Sometimes the consequences of the failure of a low probability event may be of large magnitude. Prior strategies should therefore be adopted to deal with the possible consequences of failures due to such events. Here, the idea of safe fail as opposed to fail safe is important. Vulnerability is a measure of the damage in a failure event and refers to the likely magnitude of a failure, if one occurs. To develop an index of system vulnerability, assume that the system performance variable \( X_t \) can take discrete values \( x_1, \ldots, x_n \). Now each discrete failure state \( x_j \in F \) is assigned a numerical indicator of the severity of that state, denoted by \( s_j \). Furthermore, let \( e_j \) be the probability that \( x_j \), corresponding to \( s_j \), is the most unsatisfactory and severe outcome that occurs in a sojourn into the set of unsatisfactory states \( F \); then \( e_j \) equals the probability that \( x_j \) (corresponding to \( s_j \)) is the most severe outcome in a sojourn in \( F \). A measure for overall system vulnerability is the expected maximum severity of a sojourn into the set of unsatisfactory states:

\[ v = \sum_{j \in F} s_j e_j. \]  \hspace{1cm} (12)

The emphasis here is not on how long failure persists but on how bad things may become.

Kjeldsen & Rosbjerg (2004) proposed a simplification of Equation (12) by considering the probability of each failure event to be equal. They estimated vulnerability as the mean value of the deficit events \( v_j \) as:

\[ V_{\text{mean}} = \frac{1}{M} \sum_{j=1}^{M} v_j. \]  \hspace{1cm} (13)

Kundzewicz & Kindler (1995) suggested that the use of a maximum event might yield a better estimate of vulnerability. Hence, maximum vulnerability can be computed by

\[ V_{\text{max}} = \max_j (v_j). \]  \hspace{1cm} (14)

There are therefore different approaches to estimate resilience and vulnerability depending upon whether the mean or the maximum value of the variable denoting failure is adopted. Importantly, one may arrive at completely different conclusions depending upon which estimator has been used. Table 1 summarizes the various indices of reservoir performance and their computational aspects.

This paper investigates the behaviour of the indices by simulation experiments using synthetically generated data.

DATA USED

Monthly flow data series for a reservoir located in a semi-arid region of India (the Dharoi reservoir) was available for 35 years. For the historical data, mean annual flow was 858 million cubic meters (MCM) and its standard deviation (SD) was 694 MCM. This reservoir is an over-year storage operated to meet irrigation and hydropower demands.
For the present study, it is assumed that the reservoir is operated to generate hydropower. The annual hydropower demand at this dam was assumed to be 41,000 MW-hr and its monthly distribution was taken to be known. As the length of the inflow data series is not adequate to obtain stable estimates of performance indices, we resorted to the use of synthetic data.

Using the observed data, statistical parameters were computed and were used to generate synthetic sequences of flows. Basically, two concepts are employed in hydrology to generate synthetic data – long memory and short memory. Hydrological time series display long-term memory (LM) or persistence which is the tendency of high events (when the observed values are larger than the mean) or low events (when the observed values are smaller than the mean) to cluster together. Thus, LM implies that the influence of a particular value of the series on the future values is felt for a long period of time and hence, in LM models, the long range dependence in the discharge series is considered. Among the models that have been proposed to consider long range dependence while generating synthetic data, the fractional Gaussian noise (fGn) model was used in this study. The correlation function, $\rho$ for the fGn processes is expressed as:

$$\rho(j, H) = 0.5|j|^{2H} - 2|j|^{2H} + |j - 1|^{2H}$$

$$\equiv H(2H - 1)|j|^{2H-2}$$  \hspace{1cm} (15)

where $j$ is the lag and $H$ is the Hurst exponent.

The symmetric moving average approach proposed by Koutsoyiannis (2002) to generate an fGn series was used to synthetically generate annual flows.

**DESIGN OF SIMULATION EXPERIMENTS**

Kjeldsen & Rosbjerg (2004) investigated the length of record required to give robust estimates of reliability, resilience and vulnerability and found that the estimates reached a constant level when a record length of 1,000 years is used. Keeping this aspect in view, synthetic annual data series of 1,500 years was generated as explained above. These annual flows were disaggregated into monthly flows by following the method of fragmentation used by
Srikanthan & McMahon (1980). A data series of 18,000 months was therefore used in simulations.

Mean and SD of generated annual series were 874 MCM and 688 MCM, respectively. In the numerical experiments, 12 values of demand factor (demand for a particular run/annual demand) were used: 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.1, 1.2, and 1.3. Further, storage corresponding to four values of storage ratios ($S_{\text{max}}$/MAR, where $S_{\text{max}}$ is the maximum storage capacity of the reservoir and MAR is mean annual runoff) were used. Values of these four ratios were: $S_1 = 0.5$, $S_2 = 0.7$, $S_3 = 0.9$ and $S_4 = 1.1$. Thus, for $S_1$, $S_{\text{max}} = 0.5 \times \text{MAR}$.

Using the generated inflow sequences, the operation of the Dharoi reservoir was simulated following SLOP. In reservoir operation studies, particularly those related to planning and using the SLOP, evaporation losses are usually ignored. However, if the reservoir is located in semi-arid or arid areas, the loss of water due to evaporation could be quite high. If this loss is ignored, one does not get realistic values of performance indices. For the Dharoi reservoir, the long-term average monthly depths of evaporation were available; the long-term average annual depth of evaporation was 2.15 m. These monthly depths of evaporation along with average water surface area of the reservoir for the particular month were used to determine evaporation from the reservoir. In the simulation studies, it was found that the consideration of evaporation loss has significant impact on different indices and hence it was not ignored.

Several performance indices were computed; these are listed in Table 1. Indices such as volume reliability and resilience can be computed either in terms of water or electric energy. In the present work, the indices were computed using both of these and the results were compared. In the following, the simulation results are discussed in detail.

**Reliability**

Figure 3 shows volume reliability as a function of demand factor for different values of storage ratio considering volume of water and electricity. Note that the energy demand for a given month was the same for all the years but the volume of water required to generate the desired quantity of energy in a month was different in different years depending upon the hydraulic head. In all the cases, volume reliabilities are seen to monotonically fall as the demand factor increases. In general, power reliabilities are higher compared to water and drop slowly compared to water reliability. Shortages are accompanied by lower head and hence, in the event of a shortage, higher volume of water is needed to generate the requisite amount of energy. Although in case of a shortage, the ratio of water supplied/demanded and electric energy generated/demanded for a particular month is the same, when these quantities are summed over the entire operating horizon, the reliabilities will be different. At higher demand factor, water-based reliabilities were found to be converging.

**Resilience**

Resilience depends upon duration of failure. Its value will therefore be the same irrespective of whether water volume
or electricity generation criteria are used. Variation of mean and maximum resilience as a function of demand factor for different values of $S_{\text{max}}/\text{MAR}$ is shown in Figures 4 and 5. Among the measures of resilience, the mean resilience was found to decrease non-monotonically when the demand factor ratio increases. Due to this behaviour, mean resilience does not appear to be a suitable index to measure performance of a hydropower reservoir. Further, the maximum resilience (Figure 5) showed a monotonically decreasing trend. However, it is insensitive to values of demand factor. Except for low value of demand factor, the curves of maximum resilience for different storage ratios are very close to each other. This makes the maximum resilience an unsuitable index for performance assessment of a hydropower reservoir.

Due to the above reasons, resilience does not appear to be a preferred performance measure for hydropower reservoirs.

Vulnerability

Figures 6 and 7 show the variation of mean and maximum vulnerability as a function of demand factor ratios for different values of storage ratios. At low values of demand factor, the curves for different storage ratios converge but as the demand factor increases, these diverge. Kjeldsen & Rosbjerg (2004) noted that mean vulnerability changes non-monotonically with draft ratio. In this study, it was found that the mean and the maximum vulnerabilities monotonically rise with increasing draft ratio. Expectedly, both mean

![Variation of mean resilience with demand factor for different values of storage ratio.](image1)

![Variation of maximum resilience with demand factor for different values of storage ratio.](image2)
and maximum vulnerability reduce as the storage ratio increases. Further, the maximum and the mean vulnerability are lower when the amount of energy generated is used to compute these compared to the situation when the indices are computed using amount of water required.

**DISCUSSION**

The manager of a power system is more likely to be concerned with reliability value determined from the values of electricity generated, while a water engineer may be more interested in a water-based index. Time reliability will be the same regardless of the quantity used to determine these.

The results from this study are in line with the observations by Kundzewicz & Kindler (1995) and Kjeldsen & Rosbjerg (2004). The estimates of resilience based on mean values of deficit volume are not monotonic, although the degree non-monotonic behaviour in the results of this study was very small for resilience. This study also demonstrated that maximum resilience shows monotonic change but is very sensitive to changes in storage ratios. Kjeldsen & Rosbjerg (2004) argued that, in view of large estimation uncertainty, use of maximum values estimated from historical time series is less appealing.

Further, Kundzewicz & Kindler (1995) and Kjeldsen & Rosbjerg (2004) found that vulnerability and resilience are highly correlated. Therefore, keeping in view the above and the principle of parsimony, it is felt that mean vulnerability along with volume reliability are likely to provide adequate insight into the performance of a hydropower reservoir.
Variability of performance indices

An important aspect of performance indices is the uncertainties associated with their estimates. An index that has a smaller uncertainty and that can be estimated with confidence using shorter data series would be preferred one which has a large uncertainty. In this study, an attempt was made to study uncertainties associated with estimates of each of the indices used earlier. To that end, 10,000 times series each of length of 1,000 years (12,000 months) were generated. Operation of the reservoir was simulated for a particular demand (0.8) and the performance indices corresponding to power were computed. This resulted in 10,000 values of six indices: time reliability, volume reliability, mean resilience, maximum resilience, mean vulnerability and maximum vulnerability. To determine uncertainty associated with estimation of each of these indices, minimum, maximum, mean, standard deviation (SD) and coefficient of variation (CV) estimates of each these indices were computed and are given in Table 2. The values arrived at in the long-term simulation (LTS) can be considered to be the asymptotic values of the indices.

It can be seen in Table 2 that time reliability, which is the probability of the reservoir state being success, is moderate. Volume reliability, which is the fraction of the demands satisfied, is also about the same. It is noted from this table that time and volume reliabilities have small values of SD and CV and can therefore be considered to be robust statistics.

Box diagrams are useful tools to depict variations in values of a variable. These were drawn to depict the uncertainties associated with each performance indicator. In the box (Figure 8), the central line shows the median value and the lines of the grey box on either side are lower and upper quartiles. Whiskers (error bars) above and below the box indicate the 90th and 10th percentiles and the black dots are 1.5 times the inter-quartile range. Also shown in the graph are the values of respective indices for the observed data (dashed line). Figure 8 shows the variation in time and volume reliabilities for LTS through the box diagram. Data given in Table 2 and through the plot shows that the values of time and volume reliabilities fall within the lower and upper quartile range, proving the robustness of these two measures. This also shows that a small amount of uncertainty is associated with their estimates, even when the sample size is small.

Regarding resilience, the average value of mean resilience (which is the inverse of the mean value of failure duration) for the LTS runs in Table 2 is 0.1149, indicating that the average failure duration is about 8.7 months. For the observed data, this duration was 6.3 months. Although shorter failure duration is desirable, these values are reasonable considering the fact that this reservoir is

<table>
<thead>
<tr>
<th>Statistical properties of various performance indices</th>
</tr>
</thead>
<tbody>
<tr>
<td>Statistic</td>
</tr>
<tr>
<td>-----------</td>
</tr>
<tr>
<td>Minimum</td>
</tr>
<tr>
<td>Maximum</td>
</tr>
<tr>
<td>Mean</td>
</tr>
<tr>
<td>SD</td>
</tr>
<tr>
<td>CV</td>
</tr>
<tr>
<td>Observed</td>
</tr>
</tbody>
</table>

Figure 8 | Box plot showing the variation in the estimates of time and volume reliabilities obtained using the Monte Carlo runs.
located in an arid zone where multi-year drought periods are not uncommon. When the maximum value criterion of resilience is used, the average value for generated data is 0.016 indicating the maximum failure duration of 62 months; the same value for observed data was 12 months. Box diagram for mean and maximum resilience is shown in Figure 9 along with the values of the indices arrived at using observed data. Higher value of failure duration is obtained in LTS, as the data are of much longer duration and are likely to have more instances of extreme events not present in the observed series of short length. Further, the duration of some extreme events might be much longer than that in the observed sample. Therefore, LTS highlights the fact that when the reservoir is operated for a very long time horizon, it may turn out to be less resilient (failures may be of longer duration) compared to results of past observed data.

The average value of mean vulnerability, which is an indicator of the mean value of deficit events, is 22.8 MCM for LTS and 17.2 MCM for the observed data. For vulnerability, smaller values are also preferred. Nevertheless, these values are reasonable considering the size of the reservoir, the demands and the inflows. Uncertainties associated with mean vulnerability are shown in Figure 10. In line with the values of resilience and following the same logic, the estimates of mean vulnerability in LTS are higher than the values obtained using the observed data. This shows that the estimates of mean vulnerability obtained using short data records are likely to be on the conservative side. Further, between the mean and the maximum value, the uncertainty in maximum vulnerability will be much larger.

To summarize, large uncertainties can be expected in the estimation of resilience and vulnerability when the maximum value criteria are employed, whereas the uncertainties associated with mean value criteria are smaller. From the point of view of uncertainty, the mean value of resilience or vulnerability should therefore be preferred.

**CONCLUSIONS**

Reservoirs are frequently designed and operated to generate hydropower. The performance of a hydropower reservoir may be evaluated either in terms of the volume of water or electricity generated. Considering the factors such as monotonic behaviour, sensitivity, estimation uncertainty and the principle of parsimony in a multi-criteria framework, volume reliability and mean vulnerability are expected to provide adequate and consistent insight into the performance of a storage reservoir and their use is recommended. Moreover, when the volume of water needed to generate energy is used to compute performance indices, lower values of reliability and higher values of vulnerability are obtained compared to the situation when
energy generated is used for these computations. From the point of view of uncertainty, the mean value of resilience or vulnerability should be preferred as these measures have smaller uncertainties compared to the maximum value criteria. It should be noted that these inferences are based on analysis of data from one reservoir; studies of reservoirs with different features are necessary before arriving at general conclusions.

ACKNOWLEDGEMENTS

The author would like to thank the anonymous reviewers for their useful comments which have helped to improve the paper.

REFERENCES


First received 14 September 2008; accepted in revised form 18 February 2009