Maximising water supply system yield subject to multiple reliability constraints via simulation-optimisation

John G Ndiritu
School of Civil and Environmental Engineering, University of the Witwatersrand, Private Bag 3, WITS 2050, South Africa

Abstract

The realistic incorporation of reliability into the optimisation of reservoir system design and operation remains a particularly difficult task after decades of research. While most of this research has worked with methods based on linear or dynamic programming, little has been done to find out how well the problem could be handled by a simulation model linked to an optimisation model (SO model). Water supply systems have to satisfy different demands that each require various levels of reliability and these need to be incorporated in analyses for efficient system design and operation. This study presents an approach for determining the reservoir sizes and monthly operating rules that maximise the yield of a water supply system subject to multiple reliability constraints of supply and reservoir storage. A behaviour analysis model linked to a genetic algorithm is applied and the constraints are implemented using multiplicative penalties. This approach is found to deal with multiple reliability constraints realistically and effectively with multiple runs clearly identifying the active and the redundant constraints. The long computation times are, however, a drawback for the approach and suggestions to reduce these are suggested.

Keywords: multiple reliability constraints, reservoirs, yield, simulation, optimisation

Notation

\( a_{i,k} \) surface area of reservoir \( i \) in month \( j \) of year \( i \)
\( AS_{i,j} \) vector of the direct diversions to supply \( i \)
\( as_{i,k} \) direct diversion from reservoir \( i \) in month \( j \) of year \( i \)
\( c_{i,j} \) capacity of reservoir \( i \)
\( d_{i,j} \) demand for reservoir \( i \) in month \( j \)
\( d_{t} \) annual demand for reservoir \( i \)
\( ev_{i,j,k} \) average Symon’s pan evaporation for reservoir \( i \) in month \( j \)
\( mpres_{i} \) specified maximum probability that restrictions of level \( i \) should occur
\( mprst_{n} \) specified maximum probability that a reservoir state lower than \( n \) should occur
\( N \) total number of years of simulation
\( NEV_{i,j} \) vector of net evaporation losses in month \( j \) of year \( i \)
\( nev_{i,j,k} \) net evaporation loss from reservoir \( i \) in month \( j \) of year \( i \)
\( nres_{i} \) number of times that restrictions of the \( k \)th level are imposed in the simulation period
\( nrst_{n} \) number of times the reservoir is in a state lower than the \( n \)th state
\( pr_{i,k} \) \( k \)th percentage supply of the demand \( d_{i,j} \)
\( Q_{i,j} \) vector of the inflows in month \( j \) of year \( i \)
\( q_{i,j,k} \) inflow to reservoir \( i \) in month \( j \) of year \( i \)
\( ra_{i,j,k} \) point rainfall at reservoir \( i \) in month \( j \) of year \( i \)
\( RF_{i,j,k} \) vector of regulated flows from the upper to the lower reservoir in month \( j \) of year \( i \)
\( r_{f_{i,j,k}} \) regulated flow from reservoir \( i \) in month \( j \) of year \( i \)
\( r_{l_{i,j,k}} \) \( k \)th operating rule curve value for reservoir \( i \) for month \( j \)
\( rs_{i,j,k} \) \( n \)th storage state for reservoir \( i \)
\( rsm_{i,j,k} \) parameter for optimising reservoir state variability for reservoir \( i \) in month \( j \)
\( S_{i,j} \) vector of the initial storage volumes at the beginning of month \( j \) of year \( i \)
\( s_{i,j,k} \) storage volume of reservoir \( i \) in month \( j \) of year \( i \)
\( SO \) Simulation-optimisation
\( SP_{i,j} \) vector of spill in month \( j \) of year \( i \)
\( sp_{i,j,k} \) spill from reservoir \( i \) in month \( j \) of year \( i \)
\( tc_{i} \) total system capacity
\( TR \) total runoff in the simulation period.
\( tr_{i,j} \) \( k \)th operating rule curve value for the total storage state for month \( j \)
\( ts_{i,j} \) total storage in month \( j \) of year \( i \)
\( w_{e} \) weighting parameter for obtaining the rule curves of total storage
\( WRYM \) Water Resources Yield Model
\( Y \) ratio of total actual supply (yield) to total runoff in simulation

Introduction

The construction of large-scale reservoir systems has declined significantly in many parts of the world for various reasons. In South Africa, most of the suitable sites for large reservoirs have already been exploited. Cui and Kuczera (2003) mention constraints on further development and the limited availability of funds in Australia. Labadie (2004) states that the construction of large-scale water storage projects is at a virtual standstill in the US and other developed countries and points out the increasing...
mobilisation of opposition to large storage projects in developing countries. As water demand is continuously on the increase in most parts of the world, it is essential that the efficiency in operating and managing existing systems be maximised. Efficient operation and management needs to be based on realistic assessment in which simplifications and assumptions are minimised and as much detail as required is included. According to Cui and Kuczera (2003), simulation models applying multiple sets of synthetic data (in addition to the historical set) are widely accepted as the most realistic approaches. Reservoir system optimisation research in the last 4 decades has, however, been dominated by methods based on linear programming (LP) and dynamic programming (DP). The application of LP and DP usually involves various assumptions and encounters difficulties.

In a recent state-of-the-art review of optimal reservoir system operation, Labadie (2004) mentions in particular the difficulty to realistically incorporate hydrological uncertainties. He further indicates that the assumptions that these approaches make and their complexity are some of the factors that have hindered the practical implementation of reservoir system optimisation. In South Africa, the Water Resources Yield Model (WRYM), which is based on LP (network flow programming) is used extensively although there are indications that the complexity of WRYM poses difficulties to users and potential users. Stochastic LP models tend to large problems that call for the application of Benders decomposition (e.g. Jacobs et al. (1995)) or interior point approaches (e.g. Seifi and Hipel (2001)). Many LP approaches, including a recently developed approach for dealing with reservoir reliability (Re Velle, 1999), do not incorporate the real-life practice of applying curtailment rules. The conversion of probability constraints into deterministically equivalent constraints in chance-constrained LP models (Re Velle et al., 1969) and also in reliability programming models (Simonovic and Marino, 1980) disregards the shape of the inflow probability distributions in the sense that two distinctly different distributions could provide the same deterministically equivalent inflow at the chosen reliability level. Loucks and Dorfman (1975) demonstrated that chance-constrained models can result in conservative operating rules. For the stochastic analysis of multiple reservoir systems using DP, it is often assumed that inflows into the system are uncorrelated to maintain computational tractability (Yeh (1985), Fletcher and Ponnambalam (1998)). Kelman et al. (1990) proposed Sampling Stochastic Dynamic Programming (SSDP) - a method that applies multiple streamflow sequences thereby maintaining correlations. As with many other DP models, the SSDP suffers from the ‘curse of dimensionality’.

Water supply systems often serve demands that require different levels of supply reliability and efficient system operation needs to incorporate these. Reservoir storage state reliabilities can also be incorporated if the reliability levels of utilisations that require water storage in reservoirs are specified. This study investigates the application of a simulation-optimisation (SO) approach to determine the reservoir sizes and monthly operating rules of a water supply reservoir system that maximise yield incorporating:

- Multiple reliability constraints of water supply and reservoir storage state
- Realistic supply curtailment rules.

In addition to the operation rules, yield also depends on how the available water is distributed throughout the year. An analysis of the effect of optimising the distribution of demand on yield is therefore included. Hydrological and reservoir characteristics data are obtained from a system of two reservoirs located in a semi-arid area of South Africa.

The genetic algorithm (GA), a search method that is easier to link to simulation models than LP and DP approaches is selected for optimisation. Powell’s conjugate direction method, a non-linear optimisation method also conducive to the simulation-optimisation approach is also used to optimise one of the 8 cases analysed. This serves to compare the GA with Powell’s method. This study does not consider economic aspects and the objective is akin to the reservoir design and operating models reported by Loucks et al. (1981) and the analysis by Dandy et al. (1997). The use of these models is not uncommon, especially where reliable economic data are not available and not easily obtainable. A simple objective function that maximises the total system yield as applied by Dandy et al. (1997) is used here. The optimisation of system design (sizing) is included in addition to the operating rule problem as the adequacy of system components is dependent on how the system will be operated. This implies that the system sizing and the operating problem should be handled together and not separately as is the case traditionally. For an existing system, the system sizes are set to the existing values and the design and operating problem reduces to the operating problem.

**System simulation**

Simulation was carried out using historical data at a monthly time interval and not using synthetically generated sequences. The system consists of two reservoirs in series and Fig. 1 is a schematic of the main problem components. The upper dam, Rust de Winter has a catchment area of 1 145 km² and a mean annual runoff of 19.8 mm. The incremental area for the downstream reservoir Mkombo is 2 578 km² and the mean annual runoff from the incremental area is 3.9 mm. The historical mean point runoffs at Rust de Winter and Mkombo are 605 and 243 mm respectively. Second-order polynomials were found to fit better than the classical power-law model and were therefore used to model the area–capacity relationships for the two reservoir sites. The maximum possible capacities were taken as 27.1 Mm³ for Rust de Winter and 205 Mm³ for Mkombo - the live storages of the existing reservoirs. The South African Department of Water Affairs and Forestry (DWAF) provided 77 years of monthly runoff and point rainfall data at the two sites. DWAF
also provided the monthly average Symon’s pan evaporation data which were factored by 0.85 to obtain reservoir evaporation rates. Rust de Winter is denoted as Reservoir 1 and Mkombo as Reservoir 2.

After some trial runs, it was found reasonable to assume the two reservoirs to be half-full at the start of simulation. Mass balance was then carried out at each time step as described by Eq. (1); all symbols are defined and listed in the notation.

$$\begin{align*}
S_{ij} &= \left[ \begin{array}{cc}
1 & 0 \\
0 & 1 \\
1 & 0 \\
0 & 1 \\
0 & 0
\end{array} \right] + \left[ \begin{array}{cc}
1 & 0 \\
0 & 1 \\
0 & 1 \\
0 & 0 \\
0 & 0
\end{array} \right] Q_{ij} + \left[ \begin{array}{cc}
0 & 0 \\
1 & 0 \\
0 & 1 \\
0 & 0 \\
0 & 0
\end{array} \right] NEV_{ij},
\end{align*}$$

(1)

where:

$S_{ij}$ is the vector of the initial storage volumes at the beginning of month $j$ of year $i$

$Q_{ij}$ is the vector of the incremental inflows in month $j$ of year $i$

$NEV_{ij}$ is the vector of net evaporation losses in month $j$ of year $i$

$RF_{ij}$ is the vector of regulated flows from the upper to the lower reservoir assuming no transmission losses nor abstractions

$AS_{ij}$ is the vector of the direct diversions to supply

$SP_{ij}$ the vector of spill volumes.

Equation (1) applies to month 1 to 11 of the hydrological year ($j = 1$ to 11). For $j = 12$, $(i,j+l)$ is replaced with $(i+1,l)$.

The net evaporation losses were obtained as:

$$nev_{ij} = 0.5[a_{ij} + a_{ij+1,j} \cdot \left(0.85ev_{ij} - ra_{ij,k}\right)]$$

(2)

where:

$a_{ij}$ is the surface area of reservoir $k$ at the beginning of month $j$ of year $i$

$ev_{ij}$ is the average Symon’s pan evaporation rate for reservoir $k$ in month $j$

$ra_{ij,k}$ is the point rainfall at reservoir $k$ in month $j$ of year $i$.

The direct diversion to supply $as_{ij}$ and the regulated flow from Reservoir 1 to 2, $rf_{ij}$ depend on the storage state of the individual reservoirs, the total storage state of the system, the monthly demand and the month of the year. Three sets of shortage (operating) rule curves were specified: One for the total system storage and one each for the individual reservoirs. Each set consisted of three curves giving four supply zones which is not uncommon of the practice in South Africa.

Equations (3) to (5) describe the determination of the direct diversions to supply from the two reservoirs and the controlled release from Reservoir 1 to 2.

- When the overall storage and individual reservoirs are in the same supply zone, that level of direct diversion is provided with no controlled release from Reservoir 1 to 2
- If the upper reservoir is in a lower zone than that of total storage, the diversion to supply from Reservoir 1 is based on its zone
- If the lower reservoir is in a lower zone than that of the total storage and the upper reservoir is in the zone of the total storage or a higher one, then a controlled release is made from Reservoir 1 to Reservoir 2. This release equals the extra demand that Reservoir 2 would fail to supply for being in a lower zone than the total storage. Once the controlled release is made, Reservoir 2 provides a direct diversion corresponding to the zone of total storage.
- If no controlled release is made and Reservoir 2 is in a lower zone than the total storage, Reservoir 2 provides a direct

diversion based on its supply zone

- If a reservoir is in a zone higher than that of the total storage, the direct diversion is based on the supply zone of total storage.

$$\begin{align*}
& \text{if } \text{trl}_{i,j,k} \leq trl_{i,j} \text{ and } rlf_{i,j,k} \leq rf_{i,j} \text{ then } \text{as}_{i,j,k} = \frac{pr_{i,j}}{100} \text{ and } \text{dpras}_{i,j,k} = \frac{dpr_{i,j}}{100},
\end{align*}$$

(3)

$$\begin{align*}
& \text{if } \text{trl}_{i,j,k} > trl_{i,j} \text{ and } rlf_{i,j,k} < rf_{i,j} \text{ then } \text{as}_{i,j,k} = \frac{pr_{i,j}}{100} \text{ and } \text{dpras}_{i,j,k} = \frac{dpr_{i,j}}{100},
\end{align*}$$

(4)

$$\begin{align*}
& \text{if } \text{trl}_{i,j,k} > trl_{i,j} \text{ and } rlf_{i,j,k} > rf_{i,j} \text{ then } \text{as}_{i,j,k} = \frac{pr_{i,j}}{100} \text{ and } \text{dpras}_{i,j,k} = \frac{dpr_{i,j}}{100}
\end{align*}$$

(5)

where:

$trl_{i,j,k}$ is the $h$th operating rule curve value for the total storage state for month $j$

$ts_{i,j,k}$ is the total storage at the beginning of month $j$ of year $i$, ($ts_{i,j,k} = S_{i,j,k} + S_{i+1,j}$)

$s_{i,j,k}$ is the storage volume in reservoir $k$ at the beginning of month $j$ of year $i$

$tc$ is the total system capacity ($tc = c_1 + c_2$)

$c_1$ and $c_2$ are the capacities of Reservoir 1 and 2 respectively

$rlf_{i,j,k}$ is the $p$th operating rule curve value for reservoir $k$ for month $j$

$pr_{i,j,k}$ is the $p$th percentage supply of the demand $d_{i,j}$ for reservoir $k$ in month $j$.

For the cases where the distribution of demand was specified and not optimised, $d_{i,j}$ was obtained as a constant percentage of the annual demand $d_{c}$.

The width of all supply zones for all months was restricted to a minimum of 0.1 of storage capacity as shown in Eq. (6):

$$\begin{align*}
& \text{if } \text{trl}_{i,j,k} \leq rlf_{i,j,k} - 0.1 \text{ and } \text{trl}_{i,j,k} \leq rlf_{i,j,k} - 0.1
\end{align*}$$

(6)

The reservoir spilled a quantity $sp_{i,j,k}$ if the computed $s_{i,j,k}$ exceeded $c_{i}:

$$\begin{align*}
& \text{if } s_{i,j,k} > c_{i} \text{ then } \text{sp}_{i,j,k} = s_{i,j,k} - c_{i}
\end{align*}$$

(7)

If the computed $s_{i,j,k}$ was less than the direct diversion to supply was reduced according to Eq. (8):

$$\begin{align*}
& \text{if } s_{i,j,k} < 0 \text{ then } \text{as}_{i,j,k} = \frac{\text{as}_{i,j,k} + s_{i,j,k}}{100},
\end{align*}$$

(8)

The net evaporation loss was reduced as follows:

$$\begin{align*}
& \text{if } \text{as}_{i,j,k} < 0 \text{ then } \text{nev}_{i,j,k} = \frac{\text{nev}_{i,j,k} + \text{as}_{i,j,k}}{10},
\end{align*}$$

(9)

The net evaporation loss (Eq. (2)) in a time interval depends on the surface area at the beginning and at the end of the time interval. Since the surface area at the end of the time interval itself depends on this evaporation loss, an iterative approach was applied to obtain the two quantities and achieve mass balance at each time step. Convergence was assumed when the change in net evaporation loss was less than 1% from the previous iteration. Two iterations were found adequate in most time steps to achieve this.

Eight scenarios, Cases 1 to 8 were analysed. The distribu-
tion of the demand throughout the year was assumed constant in Cases 1, 2, 7 and 8 whereas in Cases 3, 4, 5 and 6, it was allowed to vary and was optimised. A more or less constant distribution was used for Cases 2 and 8 to reflect municipal/domestic/industrial supply. For Cases 1 and 7, a more variable distribution that reflects irrigation supply was used. Higher reliabilities of supply were imposed for Cases 2, 4, 6 and 8 than for Cases 1, 3, 5 and 7 again reflecting the higher reliabilities applied for municipal/domestic/industrial supply than for irrigation supply. The operating rule curve values for the total storage \( r_{l,j} \) were obtained as simple linear functions of the corresponding values for the individual reservoirs for Cases 1 to 4 as described in Eq. (10) with the weighting parameter \( w \) in Eq. (10) being optimised. For Cases 5, 6, 7 and 8, the rule curves for total storage were taken to be independent of the individual reservoir rule curves and were optimised.

\[
tr_{l,j} = we \times rl_{l,j,1} + (1 - we) \times rl_{l,j,2}
\]  

(10)

Constant reservoir storage states were applied to determine the storage state reliabilities for Cases 1, 2, 7 and 8 and are given as \( rs \) in Table 1. For Cases 3, 4, 5 and 6, the storage states were allowed to vary on a monthly basis but to average the corresponding values of \( rs \). Twelve parameters for each reservoir denoted by \( rs_{n,k} \) in Table 3 were used to model and optimise this variability. Table 2 summarises the details of the eight cases.

### Table 1

<table>
<thead>
<tr>
<th>Level</th>
<th>Cases 1, 3, 5 and 7</th>
<th>Cases 2, 4, 6 and 8</th>
</tr>
</thead>
<tbody>
<tr>
<td>( pr_{1,1} ), ( pr_{1,2} ) (%)</td>
<td>( pr_{1,1} ), ( pr_{1,2} ) (%)</td>
<td>( mpres, ) (%)</td>
</tr>
<tr>
<td>1</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>2</td>
<td>80</td>
<td>50</td>
</tr>
<tr>
<td>3</td>
<td>60</td>
<td>80</td>
</tr>
<tr>
<td>4</td>
<td>30</td>
<td>50</td>
</tr>
</tbody>
</table>

#### Table 2

Details of analysed cases and number of decision variables

<table>
<thead>
<tr>
<th>Case</th>
<th>Demand distribution</th>
<th>Reliability levels</th>
<th>Reservoir state</th>
<th>Total storage rule curve</th>
<th>Decision variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Constant</td>
<td>Lower</td>
<td>Constant</td>
<td>Linear function</td>
<td>77</td>
</tr>
<tr>
<td>2</td>
<td>Constant</td>
<td>Higher</td>
<td>Constant</td>
<td>Linear function</td>
<td>77</td>
</tr>
<tr>
<td>3</td>
<td>Optimised</td>
<td>Lower</td>
<td>Optimised</td>
<td>Linear function</td>
<td>123</td>
</tr>
<tr>
<td>4</td>
<td>Optimised</td>
<td>Higher</td>
<td>Optimised</td>
<td>Linear function</td>
<td>123</td>
</tr>
<tr>
<td>5</td>
<td>Optimised</td>
<td>Lower</td>
<td>Optimised</td>
<td>Independent</td>
<td>158</td>
</tr>
<tr>
<td>6</td>
<td>Optimised</td>
<td>Higher</td>
<td>Optimised</td>
<td>Independent</td>
<td>158</td>
</tr>
<tr>
<td>7</td>
<td>Constant</td>
<td>Lower</td>
<td>Constant</td>
<td>Independent</td>
<td>112</td>
</tr>
<tr>
<td>8</td>
<td>Constant</td>
<td>Higher</td>
<td>Independent</td>
<td>Constant</td>
<td>112</td>
</tr>
</tbody>
</table>

### Optimisation methods

#### The genetic algorithm

The genetic algorithm (GA) is a population-based optimisation method that works on the principle of survival of the fittest as hypothesised in Darwin’s theory of evolution. Since its original development as attributed to Holland (1975), the GA has been applied extensively in optimisation. In water distribution network design, the GA has obtained lower cost solutions than other approaches (Dandy et al., 1996). Halhal et al. (1997), Dandy and Engelhardt (2001) and Wu and Simpson (2001) report effective GA applications to water distribution system rehabilitation. Srivastava et al. (2002) applied a GA to optimise the best management practice of a watershed under agricultural use to minimise water quality degradation. Karpouzos et al. (2001) used a multi-population GA to solve the inverse problem in hydrogeology and Hilton and Culver (2000) optimised groundwater remediation design using a GA. The GA has also been applied to automatic calibration of simulation models with varying levels of success (Wang (1991), Kuczera (1997), Ndiritu and Daniell (1999)).

Several reservoir optimisation analyses using the GA have been undertaken recently. Oliviera and Loucks (1997) used the GA to estimate the effective operating policies for multi-reservoir systems. They used simple example systems of two reservoirs in series or in parallel and for single purposes: Water supply or hydropower generation. Their findings suggest that the GA may be a practical way of estimating the operating policy of complex reservoir systems where general rules of thumb are unlikely to be efficient. Wardlaw and Sharif (1999) evaluated the GA for optimal reservoir system operation using the 4-reservoir and the 10-reservoir problem. The GA obtained solutions very close to the known global optimum (between 99.5 and 100%) for the two problems. Sharif and Wardlaw (2000) applied the GA to a real world multi-reservoir multi-purpose problem and concluded that the GA provides solutions that are close to the global optimum. Their study compared the GA with Discrete Differential Dynamic Programming (DDDP) and the two methods gave close results. Cui and Kuczera (2003), however, still preferred the GA for computationally demanding optimisations involving multiple synthetic data sets as it can be easily adapted to parallel computing unlike the SCE-UA. Otero et al. (1995) applied the GA to optimise the managed runoff to an estuary and Huang et al. (2002) linked the GA with stochastic dynamic programming to deal with the curse of dimensionality of DP. Cai et al. (2001) used the GA to deal with the non-linearities of a complex water resource management model whilst optimising the linear components by LP. Ndiritu (2003) used a multi-population GA to obtain the reservoir sizes and operation rules that minimise the penalty due to non-supply of water and non-utilisation of reservoirs due to low storage levels.

The GA possesses a number of favourable features. A GA code once developed is easily transferable to any problem. The GA can work with complicated forms of objective functions and does not require discretisation of the search domain, unlike most DP approaches. Further-
more, the GA requires no feasible trial state trajectories and does not require linearisation of any problem component. The GA can be easily linked to simulation models that reservoir operators accept and Labadie (2004) considers this a great advantage as it would enhance the practical implementation of reservoir system optimisation techniques.

The basic GA consists of the process of randomly generating a coded population of feasible solutions, followed by repetitions of performance evaluation, selection, crossover and mutation until the specified termination criteria are met. Numerous approaches of implementing each of the processes are in use and are often adapted to suit the problem at hand. The common coding methods include binary, gray and real number coding. A real number coded GA was found more effective than a binary coded one and was adopted for analysis.

In real number coding, a population of the decision variables of the optimisation problem is initially obtained randomly within the specified search space. The performance of each member of the population (i.e. set of decisions) is obtained by simulating the system under these decisions. The selection of the individuals used to generate the new population is based on the performance. Two commonly used methods are proportionate and tournament selection. Tournament selection has been found to converge faster that proportionate selection (Simpson and Goldberg (1994)) and has been used here. Crossover follows selection and several approaches have been used including single point, multiple point and uniform crossover. Wu and Chow (1995) favour multiple-point crossover, while Savic and Walters (1997) report better performance with uniform crossover than single point and multiple point crossovers. With a large number of crossover positions, multiple-point crossover tends to converge to uniform crossover. Multiple-point crossover is adopted herein. Mutation is implemented by changing a small proportion (probability of mutation) of randomly selected variables by small amounts. These amounts were randomly selected within a specified fraction of the search range. Each operation of the GA contributes to optimisation. Selection helps to maintain the higher performing members and to weed out those of low performance. Crossover provides a means of searching for better combinations of the valuable information in the high performing members. Mutation helps to maintain diversity and to lead the search to new regions of the search space. The processes are repeated until the specified termination criteria are met.

An automatic search range modification procedure implemented on a binary coded GA (Ndiritu and Danielli (2001)) was included in the real coded GA used here. This procedure allows the GA to use two search ranges: An inner one that can shift (hill climb) and reduce in size to fine tune the search and an outer one that specifies the search range limits.

**Powell’s conjugate direction method**

According to Rao (1998), Powell’s conjugate gradient method (Powell, 1964) is the most frequently used direct non-linear optimisation method. Angelo et al. (1999) used Powell’s method for freeway control and Mohamed and Jang (1991) used the method for structural optimisation. Stephane and Pascal (2000) applied Powell’s method for shape optimisation of a wheel motor and James et al. (1995) applied Powell’s method to learning neural control of unknown dynamic systems. Chiang et al. (2003) undertook a numerical comparison and found Powell’s method to outperform the Nelder-Mead simplex method and a quasi-Newton method. In spite of its widespread application, Powell’s method has only found limited use in water resource optimisation. Simonovic (1987) applied Powell’s method to derive quadratic reservoir operating rules and Tanakamaru and Burges (1996) used a multi-start Powell’s method for catchment model calibration. Connell et al. (1999) applied Powell’s method to optimise the management of water movement in irrigation bays. Powell’s method has the advantage of not requiring computation of derivatives and can thus work with non-differentiable objective functions. This was particularly advantageous for this study because multiplicative penalties that create discontinuities in the response surface were applied to implement reliability constraints as explained in the next section. Powell’s method is point-based and starts optimising for each decision variable individually while holding the others constant - the coordinate direction search. A pattern direction is then obtained as the difference between the original point and the point after completion of the coordinate direction searches. The next round of searches is similar to the first one with the replacement of the first coordinate direction search with the first pattern direction search. The replacement of coordinate searches with pattern searches after each round of search continues until all the coordinate directions have been replaced. A new cycle of search similar to the first one then starts from the new location. The cycles of search are repeated until convergence is obtained or other termination criteria reached.

Powell’s method was selected on the basis of its simplicity and its distinctly different search approach to that of the GA. It is, however, recognised that other methods that do not require computation of derivatives such as Nelder and Mead’s simplex method, the Metropolis algorithm or the SCA-UA could have been applied.

Powell’s method requires the specification of the step length to apply in all search directions. The magnitudes of the decision variable ranges of the optimisation problem studied here were highly varied and specifying the step length as an absolute value would have resulted in a highly inefficient search. The step length was therefore specified as a proportion of the decision variable search range. The initial point to commence Powell’s search was obtained as the arithmetic mean of randomly selected points from the initial search range as used with the GA (Table 3), that gave the highest objective function value. The number of random points was subjectively selected as the number of decision variables to optimise.

**Optimising the system**

The optimisation aimed at maximising the total yield (actual supply) from the reservoirs whilst meeting specified reliability levels of supply and reservoir storage states. The objective function was specified as the ratio of total supply to total inflow during the simulation (Eq. (11)). The constraints were implemented by reducing the objective function value to a negligible value (1.0×10⁻⁶) in case of any violation (Eq. (12)). Because the storage state was allowed to vary on a monthly basis for Cases 3, 4, 5 and 6, an additional set of constraints requiring the storage state in any month not to exceed the capacity of the reservoir was imposed and implemented in a similar way as the other constraints. The initial parameter ranges given in Table 3 were chosen with the aim of starting the search in a constraint non-violating region. This was deemed necessary after it was found that the GA often failed to recover from initial constraint violations. The initial search ranges of the operating rule curves were thus selected to ensure supply zone widths of at least 0.1 of reservoir capacity and the initial search ranges for the demands were also set at low values. The search range limits were selected.
TABLE 3  Parameter ranges used in optimisation

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Units</th>
<th>Initial range</th>
<th>Search range limits</th>
</tr>
</thead>
<tbody>
<tr>
<td>$tr_{l,i,j}$, $r_{l,i,j,k}$</td>
<td>None</td>
<td>0.71 0.81</td>
<td>0.05 0.9</td>
</tr>
<tr>
<td>$tr_{l,i,j}$, $r_{l,i,j,k}$</td>
<td>None</td>
<td>0.49 0.59</td>
<td>0.05 0.9</td>
</tr>
<tr>
<td>$r_{sm}$, $rl_{i,k}$</td>
<td>None</td>
<td>0.28 0.38</td>
<td>0.05 0.9</td>
</tr>
<tr>
<td>$c_l$</td>
<td>Mm $^3$</td>
<td>5 10</td>
<td>0.01 27.1</td>
</tr>
<tr>
<td>$c_2$</td>
<td>Mm $^3$</td>
<td>20 50</td>
<td>0.01 205</td>
</tr>
<tr>
<td>$d_1$</td>
<td>Mm $^3$/yr</td>
<td>0.001 0.01</td>
<td>0.001 8.95</td>
</tr>
<tr>
<td>$d_2$</td>
<td>Mm $^3$/yr</td>
<td>0.001 0.01</td>
<td>0.001 15.20</td>
</tr>
<tr>
<td>$d_3$</td>
<td>Mm $^3$/m</td>
<td>0.0001 0.001</td>
<td>0.0001 8.95</td>
</tr>
<tr>
<td>$d_4$</td>
<td>Mm $^3$/m</td>
<td>0.0001 0.001</td>
<td>0.0001 15.20</td>
</tr>
<tr>
<td>$r_{sm}$, $rl_{i,k}$</td>
<td>None</td>
<td>0.8 1.20</td>
<td>0.2 1.25</td>
</tr>
<tr>
<td>we</td>
<td>None</td>
<td>0.2 0.8</td>
<td>0.001 0.999</td>
</tr>
</tbody>
</table>

\[
\text{Maximize } Y = \frac{\sum \sum a_{s,t}x_{s,t}}{TR} \tag{11}
\]

where: 

$TR$ is the total runoff in the simulation period.

\[
\text{if } \frac{nres_i}{12N} > mpr_{res} \quad \text{or} \quad \frac{nrst_i}{12N} > mpr_{rst} \quad \text{then} \quad Y = 1.0 \times 10^{-10} \tag{12}
\]

where:

$nres_i$ is the number of times that restrictions of the $i^{th}$ level are imposed in the simulation period (12N months)

$mpr_{res}$ is the specified maximum probability that restrictions at this level should occur

$nrst_i$ is the number of times the reservoir is in a state lower than the $n^s$ state specified as $rs_{s,t}$ in Table 1

$mpr_{rst}$ is the specified maximum probability that a reservoir state lower than this should occur.

To use the GA, one needs to select suitable optimisation parameters. General guidelines on GA optimisation parameter selection may not be useful when dealing with specific optimisation problems. Also, many formulations of the GA exist where no guidelines may be available. The trial and error approach was used whereby multiple randomly initialised optimisations gave an indication of the adequacy of the parameters. If highly varied results were obtained, it indicated that GA solutions were trapped in local optima far from the global and if multiple runs gave reasonably close results, the GA parameters were considered reasonable. A study on the response surface characteristics in reservoir optimisation (Cui and Kuczera (2003)) implies that this test is not sufficient as multiple runs may give similar performance whilst being trapped in large flat regions of the response surface that may be far from the global optimum. However, in the absence of a practical foolproof approach, this method was adopted. After some trials, the following optimisation parameters were adopted: A population size and number of crossover positions equal to the number of parameters; a tournament size equal to half the number of parameters; a crossover probability of 1.0; a mutation probability of 0.05 and a magnitude of mutation in the range of 0 to 0.1 of the search range. Optimisation was terminated if the improvement in the objective function was less than 1% in 300 generations or if the number of generations reached 6 000. The former criterion was found to be the basis of termination in all runs. Each case was run 5 randomly initialised times. This required about 8 h for Cases 1 and 2, 24 h for Cases 3 and 4, about 40 h for Cases 5 and 6 and about 30 h for Cases 7 and 8 on a computer with a 1GHz processor.

Due to the long computation times, it was decided to compare the GA with Powell’s method for Case 1 only. The scaled step length of Powell’s method was found to impact on performance with a smaller step giving better results but using a larger step for function evaluations. After some trial runs, a step length of 0.05 of the limiting decision variable search ranges given in Table 3 was adopted.

Results and discussion

Optimised yields, demands and capacities

Figure 2 presents the yields, demands and capacities obtained for all five runs of each of the 8 cases. Figure 2a indicates that it is unlikely that the GA ever obtained the global optimum and became trapped in clearly inferior solutions in a few runs. The following discussion is based on the best yields obtained for each case. These have been highlighted (filled in black) in Fig. 2 and Table 4 provides additional details. For both reliability levels, the cases applying the simple linear dependence operating rule with optimisation of the monthly distribution of the demand (Cases 3 and 4) gave higher yields than the rest. Comparing Case 3 with Case 5 and Case 4 with Case 6, it is evident that the independent rule gave slightly inferior results than the simple linear dependence rule though it required 35 more decision variables. This is considered a result of the more difficult optimisation due to the additional parameters. The observed poorer consistency of optimised decision variables for Cases 5 and 6 compared with Cases 3 and 4 is a further indication of this. Cases 7 and 8, which applied the independent rule with constant demand distributions obtained yields close to Cases 1 and 2 that applied the simpler linear dependence rule. Allowing the distribution of monthly demand to vary improved the yield as the cases where demand distribution was optimised (3, 4, 5 and 6) gave higher yields than corresponding cases with constant demand distributions (1, 2, 7 and 8). Figures 2b and c show the high variability of optimised demands and capacities for each reservoir and also illustrate a considerable degree of correlation. Figure 3 presents the variation of system yield with total system capacity for all the runs of all 8 cases and demonstrates the expected relationship between yield and capacity for the two reliability levels.

Available on website http://www.wrc.org.za
ISSN 0378-4738 = Water SA Vol. 31 No. 4 October 2005
ISSN 1816-7950 = Water SA (on-line)
Optimised operating rules and demand distribution

Consistent operating rule curves were obtained with multiple runs of each case with the curves for Reservoir 1 being more consistent than for Reservoir 2. Figure 4 shows the operating rule curves for the best run of Case 1 using the GA and Powell’s method. The storage states (specified as the ratio of volume of water to reservoir capacity) define the operating (shortage) rule. The curves for 30 and 60% restriction levels optimised to the lowest values set in the optimisation (0.05 and 0.15) for many of the months indicating higher yields were likely if these limits were lowered.

The optimisation of demand distribution was effected by having the 12 monthly demands (d_{j,k} in Table 3) as decision variables. The distributions from the different runs were fairly consistent for Cases 3 and 4 and less so for Cases 5 and 6 which used the independent operating rule curve method. Figure 5 shows the demand distribution for Case 3. This case consistently obtained highly variable monthly demands which would not be realistic for domestic, municipal or industrial water supplies but could be for irrigation supply. For real-life situations, constraints on the demand distribution can be easily incorporated to prevent unrealistic results. Using the maximum yields obtained for Cases 1 to 4, (which used the linear dependence rule curve approach) optimising demand distribution increased yields by 13.7% and 17.0% for the lower and the higher reliability cases respectively. Optimising the distribution of demand may therefore be a worthwhile component of water resource system optimisation.

Demand distribution optimisation could for example be incorporated into a crop scheduling problem for maximising economic benefits and/or food production.

Reliability constraints

A graphical comparison of the constraint probabilities and the probabilities obtained in the optimisations for Case 3 is presented in Fig. 6. Each slanting line connects the optimised probability

<table>
<thead>
<tr>
<th>TABLE 4: Optimised yields, demands, capacities and volumetric reliabilities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case</td>
</tr>
<tr>
<td>------</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>1*p</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>5</td>
</tr>
<tr>
<td>6</td>
</tr>
<tr>
<td>7</td>
</tr>
<tr>
<td>6</td>
</tr>
</tbody>
</table>

nog: number of generations
vol. rel.: volumetric reliability = average annual direct diversion/(\( a_d + a_{d_2} \))

Note: The results of the 5 runs in each case are random and the lines only serve to highlight the correlations
(lower value) to the corresponding constraint (maximum allowed) probability given in Table 1. Figure 6 clearly shows the active constraints where there is a close match between the optimised probability and the constraint probability. These are: the 100% (unrestricted) supply and the 60% supply (restriction level 2) constraints for Reservoir 1 and all the reservoir storage state constraints for Reservoir 2 giving a total of 5 active constraints. The number of active probability constraints for Cases 1, 2, 4, 5, 6, 7 and 8 obtained similarly were 3, 1, 4, 5, 3, 3 and 3 respectively. Although the multiple runs obtained different yields, it seems they were consistent in locating the active probability constraints. It is, however, possible that more active constraints would have been identified had a more thorough optimisation been carried out.

**Comparison of genetic algorithm and Powell’s method**

The 3rd row of Table 4 presents the results of the best run of Powell’s method. Powell’s method obtained a yield closely matching that of the GA. For Case 1, the population size used was 77 and each generation thus included 77 simulations. The equivalent number of generations for Powell’s as given in Table 4, was thus obtained by dividing the number of simulations by 77. The average yield obtained from the 5 runs of Powell’s method and the GA was 0.3000 and 0.3124 respectively while the average number of simulations for the 5 runs of Powell’s method and the GA was 125 893 and 158 902 respectively. The GA obtained slightly superior results but required a larger number of simulations. It is expected that Powell’s method would have obtained improved results with a smaller scaled step than the 0.05 of the search range applied in the analysis but this would have required a larger number of simulations. The operating rule curves for the highest yielding runs of Case 1 applying the GA and Powell’s method are presented in Fig. 4. The relatively straighter rule curves for the 2nd and 3rd levels obtained by Powell’s method in comparison to those of the GA are considered as evidence of its superior fine-tuning ability. Decision variables from the

![Figure 4](http://www.wrc.org.za)

*Operation rule curves from GA and Powell’s method*
5 runs of Powell’s method were also found to be more consistent than those obtained from the GA indicating they had a higher tendency to lead to the same location of the search space. The GA, however, obtained slightly better results suggesting it generally searched more widely than Powell’s method. These observations can be attributed to the fundamentally different search approaches. Powell’s method searches in specific directions until no improvement is obtained before searching in another direction whereas the GA is not confined to any directions at any stage of the search. Other direct search optimisation methods that are easily agreeable with a simulation-optimisation approach such as the Shuffled Complex Evolution (Duan et al. (1992)) or simulated annealing (Kirkpatrick et al. (1983), Teegavarapu and Simonovic (2002)) would be expected to optimise the problem studied here effectively.
Comparison of simulation-optimisation (SO) with the water resources yield model (WRYM)

The water resources yield model (WRYM) has been used extensively for reservoir system yield analysis and optimisation in South Africa since 1985. A comparison of the simulation-optimisation (SO) approach applied here and the WRYM is therefore imperative. A qualitative comparison based on the description of the WRYM by Basson et al. (1994) is first made and is then followed by a numerical comparison. For large reservoir systems, the usual approach used by the WRYM is to subdivide the system into subsystems whose characteristics are easier to understand. This same approach could be applied to the SO method and the two-reservoir system applied here does not therefore suggest an upper limit of the capabilities of the approach.

The SO method incorporates multiple reliability constraints of reservoir storage state which the WRYM does not. With the SO method it is possible to specify reliabilities of storage state or water level related utilisation (e.g. some forms of recreation) as part of the analysis of the system.

The WRYM uses the concept of a ‘balanced system’ that enables the system of several reservoirs to be treated as a single reservoir subjected to a target draft. Using the long-term yield characteristics, a target draft that will satisfy multiple reliability constraints can be obtained. The WRYM however does not provide a method to check that the reliability constraints for a specific demand obtaining water from a specific reservoir will be met. The SO approach applied here does this readily. After optimising system releases via network flow programming using WRYM, the inter-reservoir operating rules still need to be obtained by simulation. This is a manual and iterative approach that matches the reservoir trajectories and supplies from the simulation as closely as possible to those obtained from the optimisation. In contrast, the SO method applied here undertakes the optimisation and the derivation of the inter-reservoir operating rules simultaneously and this process is automated. It is however recognised that the inter-reservoir operating rules from the SO method would need to be checked and may need refinement before adopted.

The reservoir system studied here had been previously analysed by the WRYM (DWAF 2000) thereby enabling a quantitative comparison. The historical firm yields obtained by the WRYM (excluding instream flow requirements) were 4.5 Mm³/yr and 4.2 Mm³/yr for Reservoir 1 and Reservoir 2 respectively giving a total yield of 8.7 Mm³/yr. This total yield of 8.7 Mm³/yr is well within the yields obtained from the various scenarios by the SO model as Fig. 3 and Table 4 show. With the absence of reliability constraints of reservoir storage state, the yields of the SO analysis should be higher than the firm historical yield obtained by the WRYM as firm yield represents a reliability of 100% while the SO analyses used lower reliability levels. The SO approach obtained higher total yields for all cases of the lower reliability level but lower ones for all cases of the higher reliability level. The reason for the lower total yields for these cases was found to be the need to satisfy the level 1 storage state reliability constraint for Reservoir 2.

The simulation-optimisation (SO) approach took much longer computation times than the WRYM model typically for a problem of similar size and would need to improve its efficiency tremendously before consideration for routine practical application. The SO approach however obtains directly the inter-reservoir operation rules in an automated manner unlike with the WRYM where the operating rules have to be obtained by a manual iterative process. The extra time needed to perform this task should be considered in making a fair comparison of time requirements. It is noteworthy that the manual derivation of the inter-reservoir operating rules engages the analyst throughout while the automated approach frees the analyst while the optimisation is happening.

The long computation times of the SO approach were partly the result of the refined computation of net evaporation losses which resulted in at least one repetition of the mass balance computation at each time step. Time could be saved by invoking this refinement after the optimisation has almost achieved convergence and not from the start. Computation time could also be reduced by lowering the number of variables to optimise. This could be done by defining the operating rule curves using parsimonious functions that use fewer variables than the 12 used to define each operating rule curve. The convergence criteria used in the optimisation was also strict with termination happening after an increase in yield of less than 1% in 300 generations. A less strict convergence criterion could be used especially with a reduction in the number of variables to optimise. It is also likely that other optimisation methods such as the SCE-UA could be more efficient than the GA and may therefore reduce computation times.

Conclusions and recommendations

This study has demonstrated the determination of the reservoir capacities and operating rules that maximise the yield of a water supply reservoir system subject to multiple reliability constraints of supply and reservoir storage by linking a simulation to an optimisation model. Applying simulation and multiplicative penalties has enabled easy and realistic incorporation of multiple reliability constraints. Unlike many stochastic reservoir system optimisation methods, the approach applied here obviates the need to fit inflows into probability distributions. This approach has also enabled the incorporation of realistic operating rules reflecting the practice of supply restrictions during drought periods. The rules specify the percentage of the monthly demand to release from each reservoir depending on the total system storage, the storage in the individual reservoir and the month of the year. The rules also allow for controlled release from upstream to downstream reservoirs.

Although most of the optimisation has been carried out using a GA, it has been demonstrated that Powell’s conjugate direction method could be as effective as and more efficient than the GA. It is expected that other direct search optimisation methods that are easily amenable to simulation-optimisation could effectively optimise the problem studied here without additional assumptions.

The simulation-optimisation approach gave yields comparable to the South African Water Resources Yield Model (WRYM) and has the advantage of obtaining inter-reservoir operation rules automatically. A significant drawback of the simulation-optimisation approach is the long computation times. In this study, the computation times ranged from 8 to 40 h for 5 runs of various scenarios and considerably longer periods would be required if a more rigorous analysis of hydrological uncertainty using stochastically generated data was carried out. Suggestions to reduce the computation times were: i) defining operating rule curves using parsimonious functions of few variables, ii) applying refined computation of net evaporation losses only when optimisation is close to convergence, iii) using a more relaxed convergence-based termination criterion.
and iv) trying out other optimisation methods. Cui and Kuczera (2003) highlight the problem of long computation times and propose that such analyses be handled by super computers or by parallel computation. Methods that are easily adaptable to parallel computing such as the GA may then hold an advantage over others.

Acknowledgements

The author wishes to thank the staff of the Hydrology Division of the South African Department of Water Affairs and Forestry (DWAF) for providing the hydrological data used in this work.

References


Available on website http://www.wrc.org.za
ISSN 0378-4738 = Water SA Vol. 31 No. 4 October 2005
ISSN 1816-7950 = Water SA (on-line)


