Water Resour Manage (2013) 27:3679–3696
DOI 10.1007/s11269-013-0373-5

Integrated Artificial Neural Network (ANN) and Stochastic Dynamic Programming (SDP) Model for Optimal Release Policy

Sabah S. Fayaed · Ahmed El-Shafie · Othman Jaafar

Received: 14 January 2013 / Accepted: 17 May 2013 /
Published online: 7 June 2013

Abstract Complexicity in reservoir operation poses serious challenges to water resources planners and managers. These challenges of water reservoir operation are illustrated using a simulation to aid the development of an optimal operation policy for dam and reservoir. To achieve this, a Comprehensive Stochastic Dynamic Programming with Artificial Neural Network (SDP-ANN) model were developed and tested at Sg. Langat Reservoir in Malaysia. The nonlinearity of the natural physical processes was a major problem in determining the simulation of the reservoir parameters (elevation, surface-area, storage). To overcome water shortages resulting from uncertainty, the SDP-ANN model was used to evaluate the input variable and the performance outcome of the Model were compared with the Stochastic Dynamic Programming integrated with auto-regression (SDP-AR) model. The objective function of the models was set to minimize the sum of squared deviation from the desired targeted supply. Comparison result on the performance between SDP-AR model policy with SDP-ANN model found that the SDP-ANN model is a reliable and resilience model with a lesser supply deficit. The study concludes that the SDP-ANN model performs better than the SDP-AR model in deriving an optimal operating policy for the reservoir.

Keywords Artificial neural network (ANN) · Simulation technique · Optimization technique · Stochastic dynamic programming (SDP) · Reservoir operation policy · Sg. Langat dam

1 Introduction

One of the most pressing difficulties encountered by Water Engineers and hydrologists is the determination of the relationship between reservoir parameters. The nonlinearity of physical processes results in a major problem in meeting with increasing water demand. An intense rise in population and fast growing economy is apprehended by shortage of water supply. Efficient operation of reservoirs present numerous challenges to the administration and

S. S. Fayaed (✉) · A. El-Shafie · O. Jaafar
Department of Civil and Structural Engineering, Faculty of Engineering and Built Environment, University Kebangsaan Malaysia, UKM, 43600 Bangi, Malaysia
e-mail: sabahal_fayath2006@yahoo.com

management of water resources services (Rani and Moreira 2010; Li et al. 2012). To minimize wastages resulting from inefficient water supply system, there is a need to develop an optimized operating system to aid planning and management of a complex water resources system (Jothiprakash and Shanthi 2006).

Losses resulting from inefficient system is an important factor which is taken into consideration during the design of water distribution systems. Lack of proper consideration may cause a large error which might result to greater risk. Simulation and optimization are the two powerful tools for reservoir system analysis. The quantity of water released from a reservoir are usually guided by the operating policies and are defined by the rules with respect to the targeted storage level at a specific period (Wang and Liu 2013). Most of the comprehensive study on the optimization and simulation techniques related to reservoir operations and modeling primarily used for planning purposes due to uncertainties in a physical system and a gap between theoretical model development and practices (Yeh 1985; Simonovic 1992; Wurbs 1993; Labadie 2004; Olsson et al. 2010; Sreekanth et al. 2012). This has attracted for more research on designing an optimal water resources management policy specifically for the reservoir operation in an attempt to enhance the operation system and efficient water delivery (Maqsood et al. 2005; Chang 2005; Li et al. 2006, 2008; Zarghami and Szidarovszky 2009; Lv et al. 2010; Ticlavilca and McKee 2011). Storage-surface area of a reservoir is used to calculate the losses from the lake of the dam. Auto-regression (AR) analysis is one of the oldest techniques used in the past to evaluate the relationships between parameters of a system (Shahin et al. 1993). The technique had been used in the past to determine and evaluate the relationships the parameters of the storage—surface area of a reservoir system (Ostadrahimi et al. 2012). However, the models obtained from this analysis were found to be inaccurate as a result, their simulation results cannot be relied upon. This necessitated for an alternative analytical tool that can simulate a better and more accurate result. In an attempt to optimize the performance of the reservoir system, this present study used an ANN as the analytical tool to simulate the reservoir system and the results obtained can be replicated for other reservoir systems (Chandramouli and Deka 2005; Najah et al. 2012; Kumar et al. 2013). The SDP method was considered explicit in extending the dynamic programming (DP) by incorporating the stochastic information in the systems analysis framework. This was initiated to ensure efficient handling of the reservoir operation problems with stochastic characteristics under a multistage context.

In the present study an integration of artificial neural network (ANN) and stochastic dynamic programming (SDP) models has been developed and applied to facilitate optimal operation policies for the Sg. Langat Reservoir in Malaysia. The performance of these model policies was compared with a conventional model (SDP-AR) with same objective function and same length of inflow data. The performances of the derived models were assessed using detailed performance indices which include reliability, resilience, vulnerability and cumulative penalty. The inflow into the reservoir was estimated using a Gamma distribution function. The objective of the stochastic model was to minimize the penalty function and to develop an optimal water delivery scheme from the reservoir considering the uncertainty of inflow.

1.1 Problem Statement

The efficiency and effectiveness of a reservoir largely depend on a reservoir simulation system. Different methods are presently being used for simulation among which the commonest method inefficiently to handle nonlinearity processes occurring in the reservoir system. This necessitated for a need to develop an appropriate reservoir simulation technique to ensure accurate results to keep water wastage at minimum. Simulation models used for reservoir operation are usually based on mass balance equation which depicts the hydrological behavior of the reservoir systems based on unregulated inflow operating conditions. Therefore, development of an accurate model will provide optimum results needed for optimal release policy. This study used the artificial neural network model to simulate the reservoir system to ensure a precise and more accurate results with a more reliable and robust model.

1.2 Objectives

This work examines the appropriateness of the integrated Artificial Neural Network (ANN) and Stochastic Dynamic Programming Model (SDP) in determining the reservoir operation. Obtaining a more accurate value is needed to ensure an optimal reservoir policy and is considered as the basis of the case study reported in this paper. In addition, this work provides a detailed review on the application of the SDP model with other simulation model (AR) and compares their effectiveness based on their results. This information is needed to enhance the operation and the performance reservoir system to efficient distribution system which is the primary focus of obtaining a more accurate reservoir simulation.

2 Sg. Langat Dam—Case Study

This study focuses on Sg. Langat dam as a case study. This dam is located in Malaysia at latitude 3° 12′ 43 07″ North of the Equator and longitude 101° 53′ 39.28″ East of the prime meridian in Kuala Lumpur. The dam was built in 1979 for water supply and is surrounded by a forest upstream of the Langat river at Batu 24, Hulu Langat. Its earth embankment is 2.5 MCM of the earth and crest elevation of 223.72 m with maximum height of 61 m. The type of primary spillway is gated and the crest elevation is 220.96 m. The reservoir catchment area is 41 km^2 and is as shown in Fig. 1. The dam controls and regulates the volume of water released into the Langat river during the dry season to ensure that adequate water is always available at the Sg. Langat water treatment plant at Batu 10, located at about 14 miles downstream of the dam. This dam plays a very important role in the neighborhood of Kajang and Bangi and is a major source of drinking water to the residents. The relationship between elevation, storage volume, and water surface area is fundamental information characterizing a reservoir, which is required in modeling studies. The storage versus surface area relationship is used in evaporation computations, since evaporation volumes are a function of evaporation rates and water surface area. Detailed time series plot of 15 years of historical inflow data into Sg Langat reservoir is as shown in Fig. 2.

3 Materials and Methods

3.1 Simulation Models

3.1.1 Artificial Neural Networks (ANN)

The Artificial Neural Network (ANN) is a new technique with a flexible mathematical structure which is capable of identifying the complex nonlinear relationship between input and output data when compared to other classical model techniques. The artificial neuron uses an approach based on mathematical or computational model for management and delivery of information in such a way as to interconnect and interrelate events as they

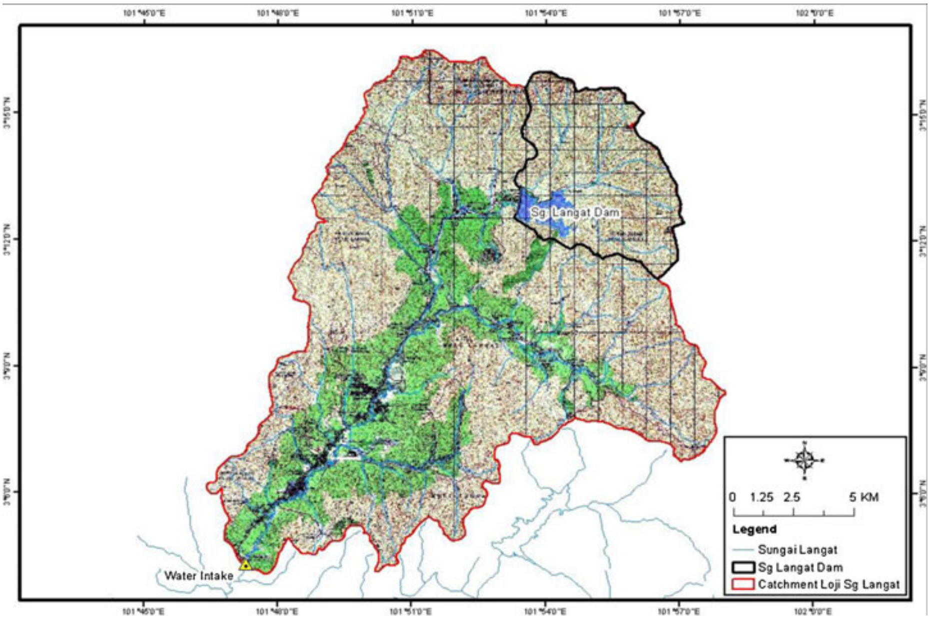

Fig. 1 Catchments area of sg. langat reservoir

happen and operate in the same way the human brain functions. Although the concept of artificial neurons was first introduced by McCulloch and Pitts (1943), research in applications of ANNs has flourished since the introduction of the back propagation training algorithm for feedforward ANNs. ANNs can be considered as a new tool in the field of prediction and forecasting (Li et al. 2010). The rules that control traditional statistical models are infrequently considered in the ANN model development process and a higher tendency exist for the use of neural network in formulating specific and appropriate solutions (Flood and Kartam 1994). In most applications, the developmental processes of a model are not

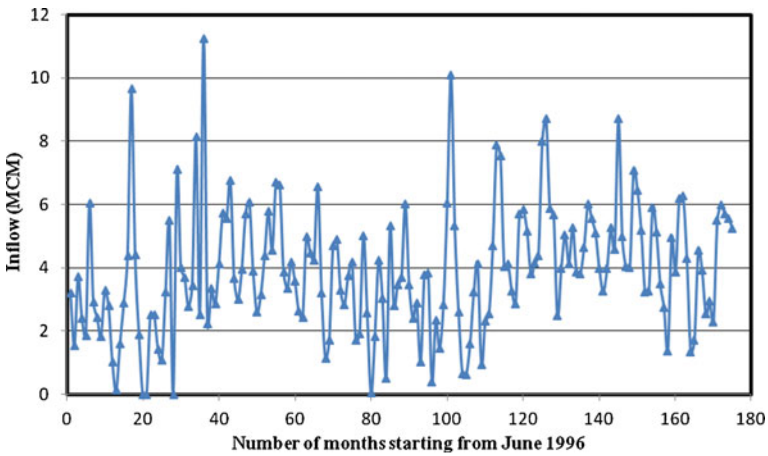

Fig. 2 Time series plot of monthly inflow into Sg Langat reservoir

properly described, making it very difficult to evaluate the best results obtained. Studies have shown that the statistical principles of ANN model building process improve the model performance (Eum and Simonovic 2010). Therefore, the present study used artificial neural network to simulate the reservoir relationship as a step towards enhancing its operation efficiency. The study further illustrates the application of radial basis function and feed forward back propagation with hyperbolic tangent neurons in the hidden layer and linear neuron in the output layer that was used for simulation (Brandão 2010). The set of all available data is separated into two disjoint sets and comprises a training set and testing set. The test set was not during the learning or training phase of the networks but to evaluate the performances of the models. In this study, 100 inputs of reservoir elevations (first 50 and last 50 inputs) were selected to provide a working order for the ANN model while the remaining 19 inputs (in between) were used to verify or test the performance of the model. The structure of storage-surface area model is 12 neurons in the input layer and 6 neurons in the hidden layer and is as shown in Fig. 3.

Feed Forward Back Propagation Neural Network (FBNN) The FFBP model contains an input layer, one or two hidden layers and an output layer in a forward multi-layer neural network. The input layer contains I nodes, the hidden layer contains J nodes and the output layer contains K nodes. Therefore, the output Z_k can be expressed as:

$$z_k = f\left(b_{ok} + \sum_{j=1}^{J} b_{jk} \cdot f\left(a_{oj} + \sum_{i=1}^{I} a_{ij} \cdot x_i\right)\right) \tag{1}$$

where function f depicts the transfer function or activation function, xi is the input quantity, a_{ij} and b_{jk} (i=1, 2,. . ., I; j=1, 2,. . ., J; k=1, 2,. . ., K) depicted the weighted values and a_{0j} and b_{0k} are the deviations. In Eq. 1, function f is a kind of mapping rule to convert neuron from weighted input to output and also is a kind of design to introduce non-linear influence into the FBNN network (Kothari and Agyepong 1996). The FBNN provide numerous transfer function but in the back propagation unit, the following general selection principles can be used: continuous function, differentiable, and monotonous non-decreasing. This study used general binary logistic sigmoid function and is defined based on the following relation:

$$f(x) = \frac{1}{1 + e^{-x}} \tag{2}$$

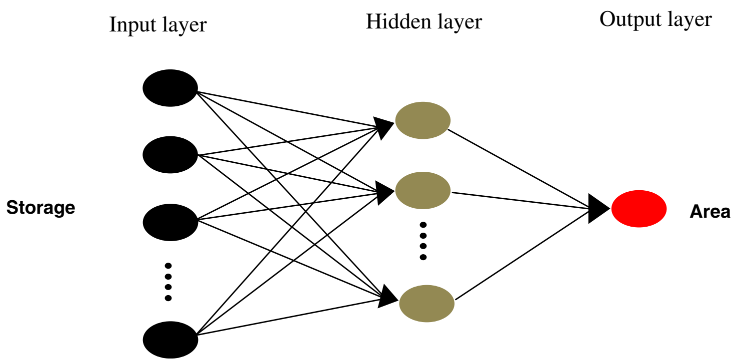

Fig. 3 Feed Forward Back Propagation Neural Network

Where the range of the value is (0, 1) and if a linear function is selected, the transfer function such as $f(x)=x$, the whole ANN structure will become the linear influence from the input layer to output layer (Barron 1993). Figure 3 shows the structure of the feed forward back propagation Neural Network.

FBNN is a supervisory learning algorithm and the optimum parameters can be obtained by adjusting the weighted values in the network (EL-Shafie and El-Manadely 2011). Optimum as used in this context implies that the squared deviation values between the network outputs z_k and the actual values or target values t_k achieved is at minimum, i.e.,

$$E = \frac{1}{2}\sum_{k=1}^{K}(z_k - t_k)^2 \tag{3}$$

In order for an ANN to generate an output vector Y=(y_1, y_2, . . ., y_p) that is as close as possible to the target vector T=(t_1, t_2, . . ., t_p), a training process also called learning were used to find the optimal weight matrices (W) and bias vectors (V), that minimizes a predetermined error function expressed as:

$$E = \sum_{P}\sum_{p}(y_i - t_i)^2 \tag{4}$$

Here, t_i is a component of the desired output T; y_i=corresponding ANN output; p=number of output nodes; and P=number of training patterns. Training is a process by which the connection weights of an ANN are adapted through a continuous process of stimulation by the environment in which the network is embedded.

Radial Basis Function Networks (RBF) The common multilayer feed forward network with one hidden layer of radial basis function network is used in the analysis presented in this study. In the radial basis function network, the input quantities are fed to input nodes and passes onto the hidden layer nodes. This adds up to the weighted input received from each layer after each neuron consisting of radial basis function centered on a point with many dimensions and comprises the predictor variable. To calculate the center of the radial, Gaussian function was used in each hidden unit depending on the distance of the input from the center and the spread of the parameter that were determined (Wang et al. 2010; Fayaed et al. 2011). The three layer stages are as shown in Fig. 4 and the present research was designed with the *newrbe* function.

RBF network was used in this research to supervised application and learning (Chen et al. 1991). During the supervision of the application, the research was provided with a set of data

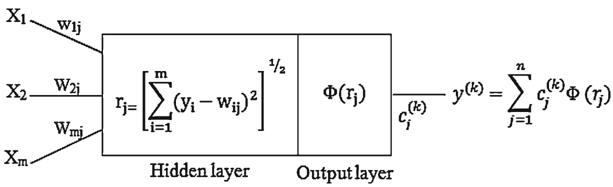

Fig. 4 Structure of Radial Basis Function Model

samples called training set for which the corresponding network outputs were determined. The most general formula for any RBF is given by:

$$\mathrm{y(x)} = \Phi\left((\mathrm{x}-\mathrm{c})^{\mathrm{T}}\mathrm{R}^{-1}(\mathrm{x}-\mathrm{c})\right) \tag{5}$$

Where Φ=activation function used, c=center; and R=metric. The term $((x-c)^{\mathrm{T}}\mathrm{R}^{-1}(x-c))$ is the distance between the input x and the center c in the metric defined by R usually the metric is Euclidean. In this case, $\mathrm{R}=r^2$ for some scalar radius r and Eq. (5) are simplifies to:

$$y(x) = \Phi\left(\frac{(x-c)^T(x-c)}{r^2}\right) \tag{6}$$

According to (Fausett 1994), the Euclidean length is represented by $\mathrm{r_j}$ which measures the radial distance between the datum vector $\mathbf{y}=(y_1, y_2, \ldots, y_m)$, and the radial center $\mathrm{Y(j)}=(w_{1j},w_{2j}, \ldots,w_{mj})$, where y_i and w_{ij}=output and weights respectively. This can be written as

$$r_j = \left\|\mathrm{y}-\mathrm{Y}^{(\mathrm{j})}\right\| = \left[\sum\nolimits_{\mathrm{i}=1}^{\mathrm{m}}\left(\mathrm{y_i}-\mathrm{w_{ij}}\right)^2\right]^{1/2} \tag{7}$$

A suitable transfer function is then applied to r_j to give

$$\Phi\left(\mathrm{r_j}\right) = \Phi\left(\left\|y-Y^{(k)}\right\|\right) \tag{8}$$

Finally the output layer (k=1) receives a weighted linear combination of $\Phi=(\mathrm{r_j})$,

$$y^{(k)} = \sum_{j=1}^{n} c_j^{(k)}\Phi\left(r_j\right) = \sum_{j=1}^{n} c_j^{(k)}\Phi\left(\left\|\mathrm{y}-\mathrm{Y}^{(\mathrm{j})}\right\|\right) \tag{9}$$

The output of each hidden unit solely depends on the distance of the input x from the center μ. During the training process, the center μ and the spread d depicted the parameters to be determined and can be deduced from the Gaussian radial function that a hidden unit is more sensitive to data points near the centre. This sensitivity can be tuned (adjusted) by controlling the spread d. The typical Gaussian radial function is as shown in Fig. 5. It can be observed that the larger the spread the less the sensitivity of the radial basis function to the input data.

3.2 Optimization Model

3.2.1 Stochastic Dynamic Programming

Stochastic dynamic programming (SDP) is an extension of DP whereby the stochastic nature of the inflows is explicitly considered during the optimization process. This is achieved by using the returns associated with each of the possible states (normally storage and inflow) in a particular period multiplied by the probability of the inflows that give rise to them, to

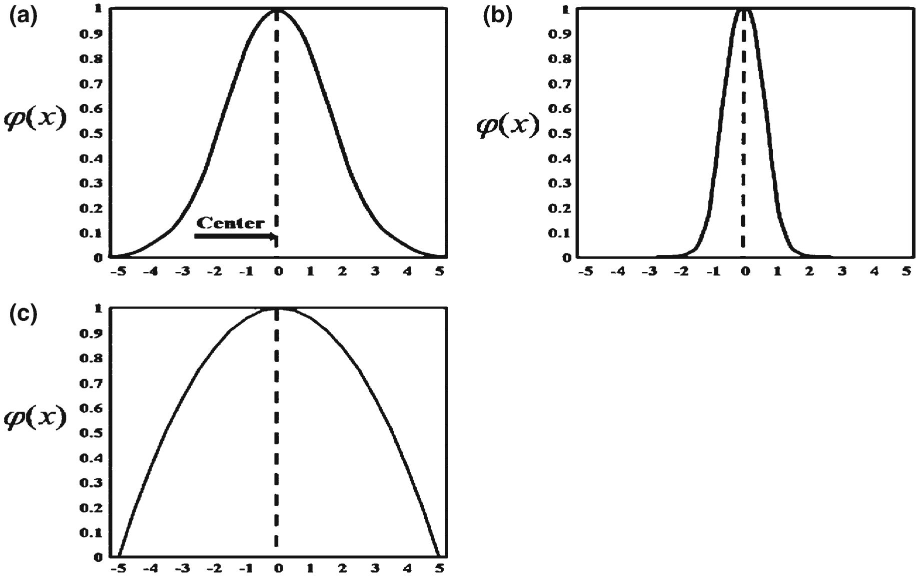

Fig. 5 Radial bases function with different levels of spread. **a** Normal spread. **b** Small spread . **c** Large spread

generate the expected value of the objective function. This probability is often referred to as transition probability and reflects the correlation between two successive inflows. The optimal operating policy designed using SDP formulation depicted a "look up table" reflects the optimal decision (the ending storage S_{t+1} or the release R_t for the time period t) under different initial state status. Dynamic Programming (DP) is useful for decision making process involving a sequence of decisions that interact with the outcomes that evolve over time (Michalland et al. 1997). In the reservoir operation system, the units of time that are considered as stages in the optimization framework comprises; water storages which is defined as the system state at a specific stage and the inflows stochastic feature which could be effectively tackled using SDP (Homayounfar et al. 2011; Eum et al. 2012). In Stochastic Dynamic Programming SDP, the randomness of the operating systems of the reservoir can be expressed as probabilities of various stages representing occurrences of different inflow levels and can be incorporated into the optimization framework at each stage (Ahmadi et al. 2010). The recursive function f of the employed SDP model is expressed as:

$$f_n^*(S_t, Q_t) = \mathrm{Min}\left[C(S_t, Q_t, R_t) + \alpha \sum_{Q_{t+1}} P_r\left(Q_{t+1} \big/ Q_t\right) \cdot f_{n-1}^*(S_{t+1}, Q_{t+1})\right] \quad (10)$$

Where

$f_n^*(.)$	Objective function
C (.)	Immediate return function
S_t	Storage at the begging of the time period t
Q_t	Inflow during the month t
R_t	Release during the time period t
α	Discount rate in the SDP algorithm

In the present study, a monthly time stepped SDP model was developed and used with the objective function of (fitness function) to minimize the squared deviation of monthly release deficit. Mathematically, the objective function is given by:

$$\text{Minimize } C(t) = \sum_{t=1}^{N} \left[\frac{R(t) - D(t)}{D(t)} \right]^2 \tag{11}$$

Where t depicts the monthly index; N is the operating horizon in the months and R(t) and D(t) represent release and demand in months t respectively. The above objective function Eq. (11) was subjected to the mass balance, storage, evaporation and surplus constraints. Storage and inflow are discretized into 60 and 100 class and are shown in Fig. 6. The first class of the inflow state variable is represented by the value zero.

Figure 7 depicted the flow chart of the proposed SDP-ANN model. The model implementation its procedural processes firstly, the operating storage volume (level) and discretized stream inflows. The SDP solution approach starts by initiating the values of the objective function at the last stage (a period in the future) to zero for each combination of the stated variables and the simulation of the reservoir system using artificial neural network (ANN). The processes continue backwards along the temporal stages. Each iteration consists of T stages to complete one annual cycle. The algorithm then reaches the end of the period storage levels for each combination of the discretized stages and stores up the average inflow of the current period. The behavior of the policy convergence after several iterations was due to the characteristics of Markov transition probability matrix which was incorporated into the recursive equation.

4 Models Evaluation Criteria

4.1 Evaluation Indicators for Simulation Models

The RBF, FBNN networks and regression models were designed to simulate the relationship between the storage and surface area of Sg. Langat reservoir. Because of the absence of definitive test to evaluate the success of each model, a multicriteria assessment was carried out. Basically, the performance of the model was evaluated by comparing the simulated data and

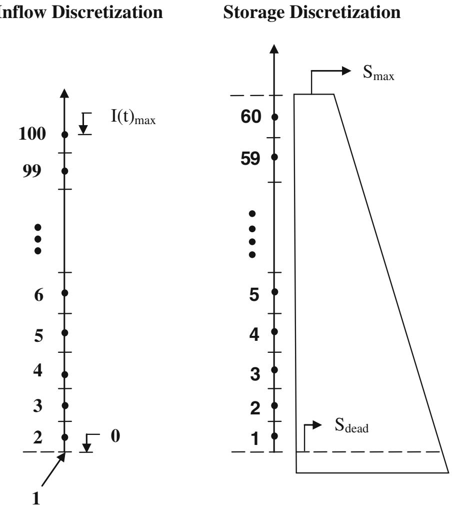

Fig. 6 Discretization used by the SDP model for each month

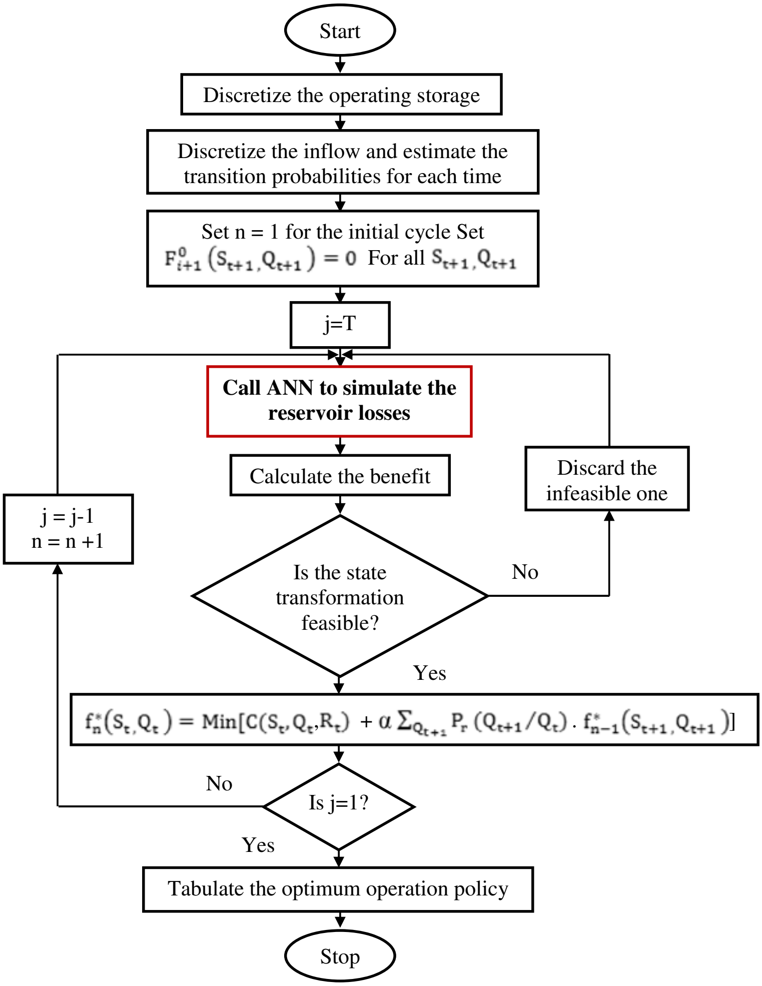

Fig. 7 Flow-chart for the solution of the proposed SDP-ANN model

actual data. The simulation of each model was evaluated using the root mean square error (RMSE), relative root mean square error (RRMSE), and the mean absolute percentage error (MAPE). Formulas for calculating RMSE, RRMSE, and MAPE were given below and are as follows:

$$\mathrm{RMSE} = \left[\frac{1}{n}\sum\nolimits_{t=1}^{n}\left(D_{a(t)}-D_{f(t)}\right)^{2}\right]^{1/2} \tag{12}$$

$$\mathrm{RRMSE} = \left[\frac{1}{n}\sum\nolimits_{t=1}^{n}\left[\frac{\left(D_{a(t)}-D_{f(t)}\right)}{D_{a(t)}}\right]^{2}\right]^{1/2} \tag{13}$$

$$MAPE = \frac{1}{n}\sum\nolimits_{t=1}^{n}\left|\frac{\left(D_{a(t)}-D_{f(t)}\right)}{D_{a(t)}}\right| \times 100\% \tag{14}$$

Generally, RMSE and RRMSE formulas used to evaluate the models were based on the result obtained by comparing the estimated error of the actual and simulated model. A model with the minimum error is considered the best choice. The result showed that MAPE was about 30 % and is considered a reasonable simulation (Johnson and King 1988). Furthermore, the analysis will be considered very accurate when the MPAE is in the range of 5–10 %. Finally, the accurate efficiency of the proposed model is measured in terms of coefficient of efficiency (CE), given by Eq. 15. Higher values of CE (maximum value is 1) indicate that the simulation performance of the model is efficient.

$$\mathrm{CE} = 1 - \frac{\sum_{t=1}^{n}\left(D_{a(t)} - D_{f(t)}\right)^2}{\sum_{t=1}^{n}\left(D_{a(t)} - \overline{D}_{a(t)}\right)^2} \tag{15}$$

Where

$D_{a(t)}$ the actual value.
$D_{f(t)}$ the simulated value.
$\overline{D}_{a(t)}$ actual mean value.
$\overline{D}_{f(t)}$ simulated mean value.
n number of forecasting periods.

4.2 Evaluation Indicators for Optimization Model

4.2.1 Reliability (R)

The most widely used and the oldest performance criterion for water resources systems is reliability and is defined by (Hashimoto et al. 1982) as:

$$\mathrm{Rel} = \mathrm{P}\{\mathrm{S} \in \mathrm{NF}\} \tag{16}$$

Where S is the state of the system variable under consideration. The most widely accepted and applied definition is occurrence reliability and can be estimated using:

$$\mathrm{Rel} = 1 - \frac{\sum_{j=1}^{m} d(j)}{T} \tag{17}$$

Where d(j) is the duration of the jth failure event, M is the number of failure events and T is the total number of time intervals.

4.2.2 Resilience (R)

Resilience entails a measure of how fast a system returns to its satisfactory state after the system performance has been distorted. Hashimoto et al. (1982) define resilience as a conditional probability:

$$\mathrm{Res} = \mathrm{P}\left\{\mathrm{S}(\mathrm{t}+1) \in \mathrm{NF} \middle| \mathrm{S}(\mathrm{t}) \in \mathrm{F}\right\} \tag{18}$$

Where S (t) is the system state variable under consideration. This definition of resilience is equivalent to the inverse of the mean value of the time the system spends in an unsatisfactory state i.e.:

$$\mathrm{Res}_1 = \left[\frac{1}{M}\sum\nolimits_{j=1}^{m} d(j)\right]^{-1} \tag{19}$$

Where again d (j) is the duration of the jth failure event and M is the total number of failure events. A study by Moy et al. (1986) defined resilience as the maximum consecutive duration a system spends in an unsatisfactory state. To conceptualize this definition with the definition of Res_1 in Eq. (19), resilience is then expressed as the inverse of the maximum duration as:

$$\mathrm{Res}_1 \left\{max_j[\mathrm{d(j)}]\right\}^{-1} \tag{20}$$

4.2.3 Vulnerability (V)

Vulnerability is a measure of the likely damage of a failure event and was defined by (Hashimoto et al. 1982) as:

$$\mathrm{Vul} = \sum\nolimits_{j\in \mathrm{F}} e(j)h(j) \tag{21}$$

Where h (j) is the most severe outcome of the jth obtained in an unsatisfactory state while *e* (*j*) is the probability of *h* (*j*) and the most severe outcome obtained in an unsatisfactory state (Hashimoto et al. 1982; Jinno et al. 1995) based on their vulnerability measure on the total water deficit experienced during the sojourning of the entire *j*th into F which explains the deficit volume. This definition is suited for reservoirs as the most severe outcome of a reservoir state is often empty, $h(j)=0$. For further simplification of the Eq. (21), both studies considered the probability of each event to be equal in that $e(1)=\ldots=e(M)=1/M$, where M is the number of failure events. Therefore, the estimated vulnerability which depicts the mean value of the deficit events $v(j)$ can be expressed as:

$$\mathrm{Vul} = \frac{1}{M}\sum\nolimits_{j=1}^{m} V(j) \tag{22}$$

5 Results and Discussions

In the present study the artificial neural network (ANN) model is used to simulate the reservoir system. The use of the ANN model was motivated by previous literature work that used regression models to simulate values of storage-surface area parameters of reservoirs. However, ANN takes into consideration its imperfection with regression models for storage-surface area model. The primary goal of the training was to minimize the error function by searching for a set of connection strengths and threshold values that cause the ANN to produce outputs that are equal or close to targets. The anticipated mean square error for the training was regulated at 10^{-4}. The criterion that was considered prior to rating FBNN module as the best was the minimum MSE obtained during the training. Meanwhile, this regulation provided a means of agitating the training protocol and to achieve a minimum MSE.

Table 1 Performance indicators of RBF, FBNN and AR models for storage—surface area relationship

	RBF	FBNN	AR
CE	0.9937	0.9048	0.4400
MAPE	0.9825	3.8793	10.200
RMSE	0.0095	0.0372	0.0930
RRMSE	0.0113	0.0421	0.1037

5.1 Storage—Surface Area Relationship

Results of the RBF, FBNN and Regression model analysis with respect to the relationship between storage and surface area for Sg. Length reservoir is as shown in Table 1. The performance of each model was indicated using CE, RMSE, RRMSE and MAPE values of each simulation method that was used. Figures 8 and 9 also illustrated the graphical results of some of the tests that were carried out on the reservoir. In a study by Johnson and King (1988), it was found that the basis for model accuracy was based on MAPE value. The values of co-efficient of efficiency, CE of each model are shown in Table 1. It can be seen that the RBF model is more efficient than the FBNN and AR models when they are subjected to testing. This was because the CE of RBF model was 0.993 which is a very high utility value and depicts high accuracy and reliability for simulating the relationship between storage and surface area. For any simulation model, simulation is considered reasonable if its MAPE value is below 30 %. The simulated model could be considered to be very accurate if the MAPE value is less than 1 %. It was on this basis that the simulation of the models used in Sg. Langat reservoir was considered to be very accurate, the MAPE value of the model was less than 10 %. In Table 1., it can be seen that the value of MAPE was 0.982 % for RBF which is less than 1 % and thus, indicated that the simulation result obtained from RBF to be very accurate. This does not apply to either FBNN or AR with 3.879 % and 10.2 % as MAPE values. It becomes obvious that RBF provided a better result than both FBNN and AR. In addition, it is known that the closer the values of RMSE and RRMSE to zero, the less the degree of error in the simulated value. As can be seen from Table 1 for all the 3 methods that were considered, the value of RMSE (0.009 km^2) and RRMSE (0.011) for RBF models were close to zero and were less than half of the value obtained from other models, thus indicated a very low error margin while those of the other methods were greater than RBF which implies a higher degree of error. This further confirmed that the RBF method gives a more accurate simulation. From the findings

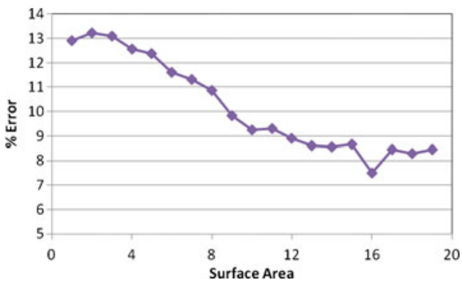

Fig. 8 Percentage error of AR for storage- surface area relationship

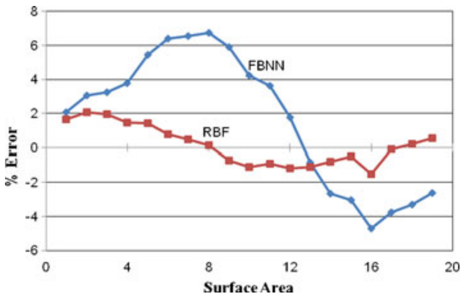

Fig. 9 Percentage error of RBF-FBNN for storage- surface area relationship

discussed above, it becomes obvious that the application of the RBF simulation method with respect to the relationship between storage and surface-area of Sg. Langat reservoir is better than the those of FBNN and AR models considering the rigorous verification processes which the methods were subjected to. The FBNN and AR models give a higher percentage error than the RBF model which has a high degree of efficiency and low values of RMSE and RRMSE.

Considering the level of error associated with different models result obtained from the different simulation methods in terms of surface area, Figures 8 and 9 showed that the model obtained using the RBF method has the least error factor and ranged between −2 to 2 (Fig. 9). The implication of this is that the minimum error value of −2 and the maximum value less than 2 obtained from FBNN depicted that the level of error in the model is slightly higher while a minimum error of −5 and the maximum was less than 7 and is as shown in Fig. 9. This shows that the AR method surpasses those of the two methods mentioned earlier. However, the range of the error values obtained are different in nature and greater in magnitude than in other simulation methods as shown in Fig. 8. In addition, the AR model shows almost a linear relationship between storage and surface area and a linear relationship with error which indicates that the model has a very high percentage of error unlike the ANN where the error is slightly regular around the axis. The RBF network shows a better performance during testing. It can be seen from the results shown in the Figures and the table that the RBF model is apparently better than the FBNN and AR models because during various series of testings the BRF model showed a consistent value of both RMSE and RRMSE. The RBF method of simulation has shown that it can efficiently handle nonlinearity processes within the reservoir. Certainly, a flexible model will consistently

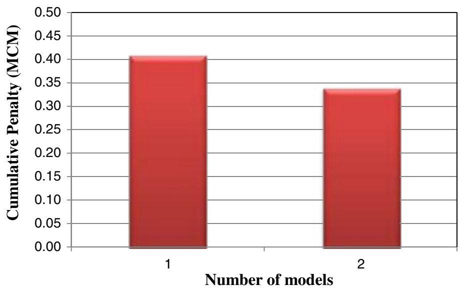

Fig. 10 Cumulative penalty for SDP-AR and SDP-ANN models

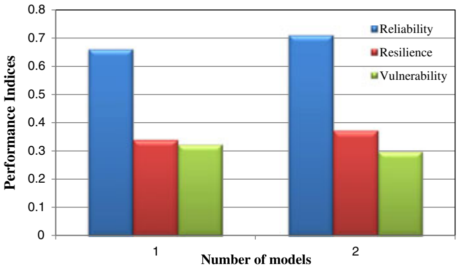

Fig. 11 Performance indices of supplying the water demand for SDP-AR and SDP-ANN models

yield a lower RMSE, RRMSE and MAPE errors. The RBF network has been noted to be flexible in modeling the storage and surface area relationships.

It therefore implies that the ANN models can be conveniently used to model hydrological process. ANN are useful and powerful tools that are commonly used to handle complex problems better than the traditional models. In the present study, the results obtained from the simulation analysis showed that the ANN provides a more accurate and reliable results than other methods and a clear relationship between storage-surface area capacity of the reservoir. This result further supports the general belief of the superiority of ANN over other traditional simulation methods used in resolving problems associated with hydrology. The present results and comparative study provided supportive evidences that the ANN method is suitable for use in simulating interrelationship than the classical regression model. The ANN model provides a useful and accurate means of solving problems in water storage. Research findings have shown that the RBF model which a form of ANN is suitable for modeling the relationships between reservoir parameters better than other forms of ANN as well as other types of modeling method such as the classical regression model. The model

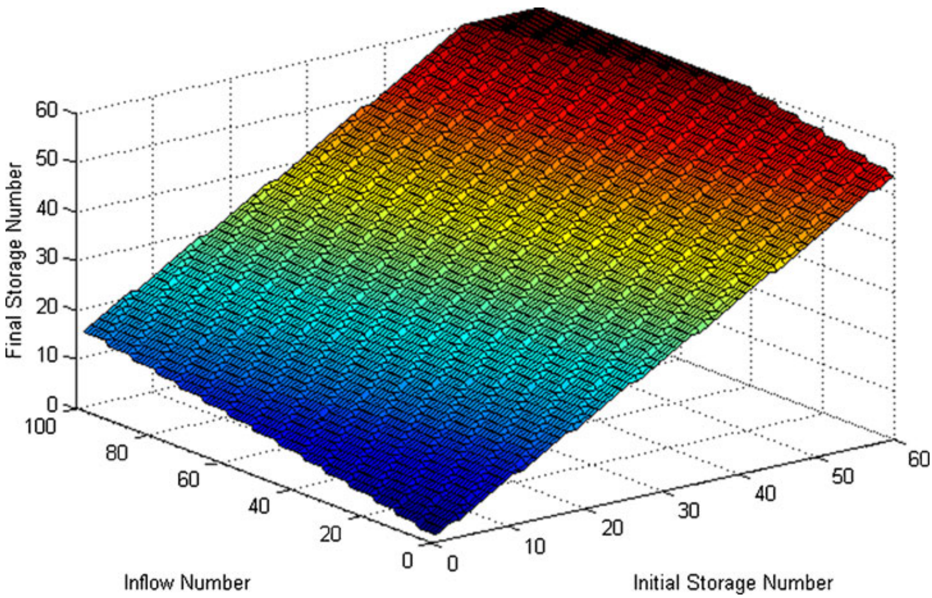

Fig. 12 Steady-state policy for February found by Model 2

could be furthermore be used to simulate the relations accurately with good compatibility between the observed and simulated values.

5.2 Stochastic Dynamic Programming with AR or ANN as a Simulation Model

The SDP based model with the AR or ANN were developed and were used to enhance the operating policies of the reservoir. The cumulative penalty of the SDP-AR and SDP-ANN models are presented in Fig. 10. The penalty function was used to measure the performance of the reservoir system as well as to identify the decision making process that contributed most to the operation of the reservoir (Karamouz and Houck 1987). The enhancement of the decision making processes was aimed at improving the supply of water to meet the targeted demand and to measure the performance of the operation based on the deviation from the targeted demand. The penalty function was used to maintain the functional value to a minimum so as to satisfactorily meet the targeted demand. Therefore the performance of a choice system should desirably improve the reliability and resiliency with less vulnerability and cumulative penalty. The present Model (model 2) encompasses the integration of artificial neural networks (ANN) and stochastic dynamic programming (SDP). The values obtained using the cumulative penalty shown in Fig. 10 were 0.39 and 0.36 for Model 1 and Model 2 respectively. The result obtained depicted that the minimum penalty value obtained from Model 2 (SDP-ANN) were comparable with alternative model. The result showed that the cumulative penalty value obtained using the Model 2 is better compared to Model 1 (SDP-AR). The gain obtained using the Model 2 was 8.3 % better that model 1.

The performance indices used to evaluate water supply demand at the downstream section of the reservoir for a 15-year simulation period were confined to reliability, resiliency, and vulnerability-used. Results showed that the reliability scores were 0.66 % and 0.71 % respectively for Model 1 and 2. This implies that using the Model 2 improved the reliability of water supply demands by 7.5 % compared with model 1. The implication of water supply shortages over 175 months of the study period was an unsatisfactory water supply for the period of 51 months during model 2 and water supply shortages for 60 months when Model1 was used. In addition, more robust and reliable results were obtained and are as shown in Fig. 11. The resilience obtained using the Model 2 provided a gain of 6.3 % compared to Model 1. This implies that model 2 were improved back to a satisfactory state faster than model 1. The vulnerability values representing the expected values of shortages of water demand were about 0.38 and 0.36 MCM in Model 1 and 2 respectively. From the vulnerability values, it was found that the Model 2 have the least value with gains of 5.5 % compared to Model 1.

The result obtained by using the reservoir operating system for each month is represented by the three prolonged relationship between storage number, monthly inflow and final storage number. This relationship represents a steady-state policy for each respective month. As shown in Fig. 12, February was used to demonstrate the shape of the relationship.

6 Conclusions

SDP-ANN model has been developed and successfully used in a case study. The optimal results obtained were compared with the SDP-AR model. Both the models are resolved with same objective function and constraints with the same length of inflow data. The nonlinearity of natural physical processes poses a major problem in determining the simulation of reservoir systems, especially when using the conventional linear methods as in the case of the auto regression method with AR. In order to overcome this problem, a non linear computational method such as ANN were used in the present study to simulate the system. The present results revealed the superiority of using ANN as a simulation model over AR model. It becomes obvious that Model 2 has the least penalty value. In addition, it can be observed that the Model 2 is more reliable, resilient and less vulnerable than the Model 1. In order to compare the accuracy of SDP-AR and that of the SDP-ANN, a comparison was made between Model 1 and Model 2. Model 2 showed an increase in reliability and resilience of 7.5 % over Model 1 (6.3 %). It can then be concluded that the SDP-ANN model performs better than the SDP-AR model in deriving the optimal operating policy for the reservoir system. In the SDP model, the stochastic nature of inflow were described using the Markov process. This assumption has proved adequate for a prolonged period (monthly and yearly inflow) for models with short frame of time (daily and weekly). However, the assumption inadequately captures the precursors of an inflow value which could significantly impact the benefits of using SDP to provide innovative approach to transform previous inflow patterns. Based on the present findings, this study suggests that PSDP-ANN model should be integrated into reservoir operation.

Acknowledgments This research was supported by a research grant to the second author by Ministry of Higher Education, University Kebangsaan Malaysia, and project leader FRGS/1/2012/TK03/UKM/02/4. The authors would like to thank the Puncak Niaga (M) SDN BHD for providing the data for Sg. Langat reservoir.

References

Ahmadi A, Karamouz M, Moridi A (2010) Robust methods for identifying optimal reservoir operation strategies using deterministic and stochastic formulations. Water Resour Manag 24(11):2527–2552
Barron AR (1993) Universal approximation bounds for superposition of a sigmoid al function. IEEE Trans Inf Theory 39:930–945
Brandão JL (2010) Performance of the equivalent reservoir modelling technique for multi-reservoir hydropower systems. Water Resour Manag 24(12):3101–3114
Chandramouli V, Deka P (2005) Neural network based decision support model for optimal reservoir operation. Water Resour Manag 19(4):447–464
Chang NB (2005) Sustainable water resources management under uncertainty. Stoch Environ Resour Risk Assess 9(2):97–98
Chen SC, Cowan FN, Grant PM (1991) Orthogonal least squares learning algorithm for radial basis function networks. IEEE Trans Neural Netw 2(2):302–309
EL-Shafie A, El-Manadely MS (2011) An integrated neural network stochastic dynamic programming model for optimizing the operation policy of Aswan High Dam. Hydrol Res 42(1):50–67
Eum H, Simonovic SP (2010) Integrated reservoir management system for adaptation to climate change: the nakdong river basin in Korea. Water Resour Manag 24(13):3397–3417
Eum H, Vasan A, Simonovic SP (2012) Integrated reservoir management system for flood risk assessment under climate change. Water Resour Manag 26(13):3785–3802
Fausett L (1994) Fundamentals of neural networks. Prentice Hall, Englewood Cliffs
Fayaed S, El-Shafie A, Jaafar O (2011) Adaptive neuro-fuzzy inference system–based model for elevation–surface area–storage interrelationships. Neural Comput & Applic 22(5):987–998
Flood I, Kartam N (1994) Neural networks in civil engineering. I: principles and understanding. J Comput Civ Eng 8(2):131–148
Hashimoto T, Loucks DP, Stedinger J (1982) Reliability, resilience and vulnerability for water resources system performance evaluation. Water Resour Res 18(1):14–20
Homayounfar M, Ganji A, Martinez CJ (2011) A novel solution for stochastic dynamic game of water allocation from a reservoir using collocation method. Water Resour Manag 25(13):3427–3444

Jinno K, Zongxue X, Kawamura A, Tajiri K (1995) Risk assessment of a water supply system during drought. Water Resour Dev 11(2):185–204

Johnson D, King M (1988) Basic forecasting techniques. Butterworth, London

Jothiprakash V, Shanthi G (2006) Single reservoir operating policies using genetic algorithm. Water Resour Manag 20:917–929

Karamouz M, Houck MH (1987) Comparison of stochastic and deterministic dynamic programming for reservoir operating rule generation. Water Resour Bull 23(1):1–8

Kothari R, Agyepong K (1996) On lateral connections in feedforward neural networks. Proc., IEEE Int. Conf. on Neural Networks, Institute of Electrical and Electronics Engineers, New York. 13–18

Kumar ARS, Goyal MK, Ojha CSP, Singh RD, Swamee PK, Nema R (2013) Application of ANN, fuzzy logic and decision tree algorithms for the development of reservoir operating rules. Water Resour Manag 27(3):911–925

Labadie JW (2004) Optimal operation of multi-reservoir systems: State-of-the-art review. J Water Resour Plan Manag ASCE 130(2):93–111

Li YP, Huang GH, Nie SL (2006) An interval-parameter multistage stochastic programming model for water resources management under uncertainty. Adv Water Resour 29:776–789

Li YP, Huang GH, Nie SL, Liu L (2008) Inexact multistage stochastic integer programming for water resources management under uncertainty. J Environ Manag 88(1):93–107

Li L, Xu H, Chen X, Simonovic SP (2010) Streamflow forecast and reservoir operation performance assessment under climate change. Water Resour Manag 24(1):83–104

Li FF, Wei JH, Fu XD, Wan XY (2012) An effective approach to long-term optimal operation of large-scale reservoir systems: case study of the three gorges system. Water Resour Manag 26(14):4073–4090

Lv Y, Huang GH, Li YP, Yang ZF, Liu Y, Cheng GH (2010) Planning regional water resources system using an interval fuzzy bi-level programming method. J Environ Inform 16(2):43–56

Maqsood I, Huang GH, Huang YF, Chen B (2005) ITOM: an interval parameter two-stage optimization model for stochastic planning of water resources systems. Stoch Environ Resour Risk Assess 19(2):125–133

McCulloch WS, Pitts W (1943) A logic calculus of the ideas immanent in nervous activity. Bull Math Biophys 5:115–133

Michalland B, Parent E, Duckstein L (1997) Bi-objective dynamic programming for trading off hydropower and irrigation. Appl Math Comput 88(1):53–76

Moy WS, Cohon JL, ReVelle CS (1986) A programming model for analysis of the reliability, resilience, and vulnerability of a water supply reservoir. Water Resour Res 22(4):489–498

Najah AA, El-Shafie A, Karim OA, Jaafar O (2012) Water quality prediction model utilizing integrated wavelet-ANFIS model with cross-validation. Neural Comput & Applic 21:833–841

Olsson O, Ikramova M, Bauer M, Froebrich J (2010) Applicability of Adapted Reservoir Operation for Water Stress Mitigation Under Dry Year Conditions. Water Resour Manag 24(2):277–297

Ostadrahimi L, Mariño MA, Afshar A (2012) Multi-reservoir operation rules: multi-swarm PSO-based optimization approach. Water Resour Manag 26(2):407–427

Rani D, Moreira MM (2010) Simulation–optimization modeling: a survey and potential application in reservoir systems operation. Water Resour Manag 24(6):1107–1138.33

Shahin M, Van Oorschot HJL, De Lange SJ (1993) Statistical analysis in water resources engineering. A. AaBalkema, Rotterdam

Simonovic SP (1992) Reservoir systems analysis: closing gap between theory and practice. J Water Resour Plan Manag 118(3):262–80

Sreekanth J, Datta B, Mohapatra PK (2012) Optimal short-term reservoir operation with integrated long-term goals. Water Resour Manag 26(10):2833–2850

Ticlavilca AM, McKee M (2011) Multivariate Bayesian regression approach to forecast releases from a system of multiple reservoirs. Water Resour Manag 25(2):523–543

Wang H, Liu J (2013) Reservoir operation incorporating hedging rules and operational inflow forecasts. Water Resour Manag 27(5):1427–1438

Wang Y, Chang JX, Huang Q (2010) Simulation with RBF neural network model for reservoir operation rules. Water Resour Manag 24(11):2597–2610

Wurbs R (1993) Reservoir system simulation and optimization models. J Water Resour Plan Manag ASCE 119(4):455–472

Yeh W (1985) Reservoir management and operation models: a state-of-the-art review. Water Resour Res 21(12):1797–1818

Zarghami M, Szidarovszky F (2009) Stochastic-fuzzy multi criteria decision making for robust water resources management. Stoch Environ Resour Risk Assess 23(3):329–339