Enhancements to Simulation- and Discretization-Free Explicit Stochastic Reservoir Operation Optimization Method

S. Jamshid Mousavi[1], Alcigeimes B. Celeste[2], Kumaraswamy Ponnambalam[3], and Ximing Cai[4]

[1]Department of Civil and Environmental Engineering, Amirkabir University of Technology, Iran.

[2]Department of Civil Engineering, Federal University of Sergipe, Brazil.

[3]Department of Systems Design Engineering, University of Waterloo, Canada.

[4]Department of Civil and Environmental Engineering, University of Illinois at Urbana-Champaign, USA.

Corresponding Author: Kumaraswamy Ponnambalam (ponnu@uwaterloo.ca)

Key Points:

- A new implementation of the FP method for storage systems leads to a vectorized unconstrained optimization problem solved 27 times faster.
- Newly-derived moments of deficits and spills lead to an even better optimality of the FP model in the case of nonlinear objective functions.
- FP results are compared to stochastic dynamic programming and novel two-stage stochastic programming results show its overall superiority.

Abstract

The Fletcher-Ponnambalam (FP) method is an explicit stochastic optimization method for design and operations management of storage systems. It has been applied successfully in many real-world operations optimization problems (for example, the Great Lakes system and the Parambikulam-Aliyar project) and groundwater management problems. The FP method faces no curse of dimensionality unlike stochastic dynamic programming (SDP) and no need for scenarios generation as in implicit stochastic programming (ISP) methods. The paper introduces a novel implementation for the FP method by removing the need for nonlinear constraints and by decreasing the number of decision variables to just one third of its original value, significantly reducing solving time (~27 times faster than the original formulation). Additionally, new expressions derived for first and second moments of both reservoir release deficit and spill terms and the already-derived expression for second moments of reservoir storage are incorporated into the new formulation enabling the FP method to reach an improved optimality for a nonlinear objective function. The enhanced procedure is applied to solving a water reservoir operation optimization problem for a major dam in Brazil. The result comparisons made with SDP, two-stage stochastic programming and ISP along with a thorough analysis of release operation policies for both non-Gaussian correlated and Gaussian independent inflows prove the optimality of this highly numerically efficient and convenient-to-use FP method.

1 Introduction

Most complex natural phenomena modeled by means of the systems concept are affected by the presence of unpredictable variables. This is the case of storage systems governed by the mass balance equation in which the input⁄output are stochastic processes. Storage systems analysis is encountered in many areas today, e.g., warehouse management, energy management, ATM cash machines, etc. Water resource systems planning and management in view of uncertain hydrology and changing climate is another mature field dealing with these problems for the past several decades. This paper presents an improved stochastic reservoir operation optimization method.

Explicit (ESP) and implicit (ISP) stochastic programming (optimization) techniques are recognized to be efficient tools for identifying optimal planning and operating strategies for multipurpose multireservoir systems under uncertainty (Alizadeh et al., 2018; Archibald & Marshall, 2018; Pan et al., 2015; Fayaed et al., 2013; Nagy et al., 2002; Celeste & Billib, 2009; Labadie, 2004). ESP incorporates probabilistic inflow models directly into the optimization formulation. In the practice of reservoir systems operation optimization under uncertainty, stochastic dynamic programming (SDP) based models are typically the optimization approach of choice. SDP finds steady-state operating policies by means of a discretization scheme of reservoir inflows and storage (Loucks & van Beek, 2017). The need for discretization in multiple state variables discrete SDP results in the so-called “curse of dimensionality”. In this context, several modifications and enhancements of the traditional SDP formulation have been introduced (Ponnambalam & Adams, 1987; Turgeon, 1981; Adams & Ponnambalam, 1994; Ponnambalam & Adams, 1996; Mousavi et al, 2004; Saadat & Asghari, 2017). ISP methods on the other hand, applies perfect-forecast deterministic optimization to operate the reservoir for several equally possible inflow scenarios and then examines the set of optimal results in order to define release policies. The main inconvenience of ISP especially for use in multireservoir systems is the need for many inflow scenarios and deterministic optimization problems to be solved, which may turn to be laborious. It also requires proper post-processing methods to infer general operation rules from the optimization results (Labadie, 2004; Mousavi et al. 2007; Alizadeh et al, 2014). Additionally, Cai et al. (2003) compared the results of a two-stage model and an ISP model and demonstrated the possible bias with the ISP model for when the number of scenarios are limited.

Fletcher and Ponnambalam (FP) (1996) introduced a new discretization-free explicit stochastic optimization method that incorporates indicator functions into the reservoir mass balance equation in order to deal with storage bounds and to find statistical moments of storage together with probabilities of deficit and spill. The FP method requires neither discretization of state variables nor generation of inflow scenarios to deal with uncertainty, making it fast to easily address multireservoir problems without facing the curse of dimensionality. The most recent version (Fletcher & Ponnambalam, 2008) of the FP method using S-type linear decision rules rather than the original (Fletcher & Ponnambalam, 1996) open-loop constant release policy has been applied successfully to single and multireservoir systems (Mahootchi & Ponnambalam, 2013; Ganji & Jowkarshorijeh, 2012; Mahootchi et al., 2010) and has been adapted to groundwater management problems (Joodavi et al., 2017) and to other storage systems such as energy storage systems (Ponnambalam et al., 2010) and warehouse systems (Mahootchi et al., 2012).

Although the FP method is well established, there is still room to improve its performance in both its formulation and its computer implementation. This paper introduces a novel implementation that formulates the FP method into an entirely unconstrained optimization problem with a drastic reduction in the number of decision variables that is easily implemented in a vectorized form facilitating its coding in numerical matrix computing environments such as MATLAB or Octave. Additionally, newly derived equations for reservoir release deficit and spill terms as well as information already derived from the FP method (second moments of storage) will be fully used for the improvement of the model formation. In this regard, most applications of the FP method so far have adopted zeroth-order Taylor series expansion of the expected value of the objective function. Thus, in spite of expressions estimating first and second moments of reservoir storage have been available, only first moments have been used in the objective function formulation in most applications (one exception being Mahootchi et al. (2010) where a risk part using second moments of storage was included for the original linear objective function), and second moments have been left just for comparison purposes with sample second moments calculated by simulating the policies derived by the FP optimization model. That has been one reason for a gap between the simulated objective function values and those estimated by the FP optimization. Analyzing this gap, we show that for a nonlinear quadratic objective function it is important to include the second moments of storage in the objective function to accurately estimate the expected value of the objective function. Moreover, in the FP model applications so far, there are expressions derived for probabilities of containment, deficit and spill, but not for the moments of deficit and spill. Those derived probabilities are also not utilized in the objective function and are left just for comparison purposes with simulation results. This study presents new explicit expressions for the moments of deficit and spill that are incorporated in the model's objective function, resulting in a much more accurate evaluation of the objective function compared to when they are not included.

The significance of the proposed enhancements and the performance of the new model implementation is assessed by applying it to the operation of a real-world system in Brazil, the very large Sobradinho reservoir. Moments and variances of storage, deficit and spill variables as well as probabilities of containment, deficit and spill terms found by a single run of the new vectorized version of the FP method are compared with those obtained by Monte Carlo simulations of monthly reservoir operations for many different scenarios including for a scenario of over 1,000 years. A comprehensive analysis of the derived optimal policies is also made by comparing the results of the enhanced FP model with those of SDP, two-stage stochastic programming (TSP) and ISP models for different types of operating policies and both non-Gaussian correlated and independent Gaussian inflows.

2 Models and Methods

This section presents the basic FP method, new modifications to a quadratic objective function that provide a better accuracy, the directions for vectorized implementation of the method, new time complexity, and extensions to other nonlinear objective functions and multireservoir systems. In order to compare the results of the FP method, we present briefly the two stage stochastic programming method which also allowed us to test the LDR policies used in FP method with other more general policies and inflow scenarios. Readers are referred to other literature for descriptions of SDP which is also used to compare the results.

2.1 The FP Method

The main function of a water supply reservoir is to accumulate water in periods of high flows in order to regulate streamflows and to meet demands to the greatest extent possible during dry seasons. A general equation that describes the mass balance of a water supply reservoir may be written as follows

$$S_t = S_{t-1} + I_t - U_t - Sp_t + \delta_t \qquad (1)$$

where S_t and S_{t-1} represent the reservoir storage at (the end of) time periods t and $t-1$, respectively; I_t is the (net) natural inflow into the reservoir during the time period t; and U_t is the (proposed) total release from the reservoir in time t. Sp_t represents the spill when the reservoir is full; while δ_t is defined as the storage deficit when the reservoir storage goes below the minimum active storage with the proposed release (U_t). In the actual operation, U_t is usually reduced to a level to make the storage to stay within the minimum storage bound.

According to Figure 1, we assume the storage S_t to be bounded by lower (S_t^{min}) and upper (S_t^{max}) limits. Let $\hat{S}_t = S_{t-1} + I_t - U_t$ denote the *projected storage volume*, i.e., the storage at the end of time t if the proposed U_t is released and the final storage is contained or remains within the bounds, i.e., $S_t^{min} \leq \hat{S}_t \leq S_t^{max}$ (in this case, $S_t = \hat{S}_t$).

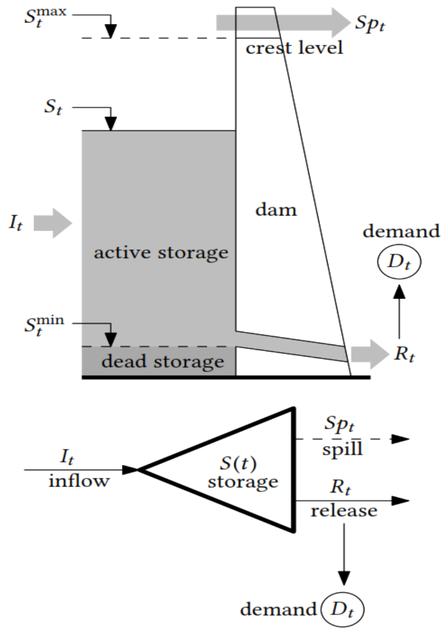

Figure 1. Representation of a reservoir system showing the variables used to represent its dynamics

When the projected storage is not contained, there will be either (but not simultaneously) spill (Sp_t) or deficit (δ_t). If releasing U_t causes the projected storage to violate the upper bound (i.e., $\hat{S}_t > S_t^{max}$), then the excess water must spill from the reservoir. In this case, the spill variable Sp_t denotes the volume of spill so that the final storage becomes $S_t = S_t^{max}$; the amount of spill will be $Sp_t = \hat{S}_t - S_t^{max}$. Alternatively, when $\hat{S}_t < S_t^{min}$ then U_t cannot be fully met and there will be a release deficit δ_t, a situation that requires an alternative release $R_t \leq U_t$ so that the final storage becomes at least S_t^{min}. In this case, the amount of deficit will be $\delta_t = S_t^{min} - \hat{S}_t$ and the actual release will be $R_t = U_t - \delta_t$. Note that both Sp_t and δ_t are nonnegative quantities. Consequently, the actual total outflow r_t from the system is:

$$r_t = \begin{cases} (U_t - 0) + 0 = U_t & if\ containment \\ (U_t - \delta_t) + 0 = U_t - \delta_t = R_t & if\ deficit \\ (U_t - 0) + Sp_t = U_t + Sp_t & if\ spill \end{cases} \qquad (2)$$

In the FP method, the dynamics of a reservoir system taking all the above situations into account is written as

$$S_t = (S_{t-1} + \bar{I}_t + \eta_t - U_t) \cdot \mathbb{1}_{[S_t^{min}.S_t^{max}]}(\hat{S}_t) + (S_t^{min}) \cdot \mathbb{1}_{[-\infty.S_t^{min}]}(\hat{S}_t) + (S_t^{max}) \cdot \mathbb{1}_{[S_t^{max}.+\infty]}(\hat{S}_t) \qquad (3)$$

in which the inflow is now split into $I_t = \bar{I}_t + \eta_t$, the mean inflow $\bar{I}_t$ plus a zero-mean random component η_t with variance $Var(\eta_t)$. The notation $\mathbb{1}_{[.]}(\hat{S}_t)$ denotes the indicator (characteristic) function with the following properties:

$$\mathbb{1}_{[S_t^{min}.S_t^{max}]}(\hat{S}_t) := \begin{cases} 1 & \text{for } S_t^{min} \leq \hat{S}_t \leq S_t^{max} \text{ (containment)} \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

$$\mathbb{1}_{[-\infty.S_t^{min}]}(\hat{S}_t) := \begin{cases} 1 & \text{for } \hat{S}_t < S_t^{min} \text{ (deficit)} \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

$$\mathbb{1}_{[S_t^{max}.+\infty]}(\hat{S}_t) := \begin{cases} 1 & \text{for } \hat{S}_t > S_t^{max} \text{ (spill)} \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

Therefore, at a given time, only one of three indicator functions can have a value of 1 and others must be zero. Continuity equation (5) can be simplified if we write the projected reservoir release in the form of an S-type *linear decision rule* (ReVelle et al., 1969), hence the proposed release is a function of current storage unlike in Fletcher and Ponnambalam (1996,1998):

$$U_t = S_{t-1} + k_t \tag{7}$$

Equation (7) is an important assumption in the proposed FP model. Therefore, we present in section 4.1 the results of a thorough analysis conducted for assessing how this simple linear decision rule (LDR) performs compared to other methods, benefiting from more sophisticated release policies, and a policy-free ISP method with a large number of scenarios which are the same used in simulation for comparison purposes.

By applying the above LDR, the projected storage volume becomes

$$\hat{S}_t = \bar{I}_t + \eta_t - k_t \tag{8}$$

and substituting $U_{t'}$ from equation (9) into equation (5) in order to eliminate S_{t-1} yields

$$S_t = (\bar{I}_t + \eta_t - k_t) \cdot \mathbb{1}_{[S_t^{min}.S_t^{max}]}(\hat{S}_t) + \left(S_t^{min}\right) \cdot \mathbb{1}_{[-\infty.S_t^{min}]}(\hat{S}_t) + (S_t^{max}) \cdot \mathbb{1}_{[S_t^{max}.+\infty]}(\hat{S}_t) \tag{9}$$

Thus, if we square the above equation, the terms containing the products of different indicator functions will disappear resulting in the final expression below

$$S_t^2 = (\bar{I}_t + \eta_t - k_t)^2 \cdot \mathbb{1}_{[S_t^{min}.S_t^{max}]}(\hat{S}_t) + \left(S_t^{min}\right)^2 \cdot \mathbb{1}_{[-\infty.S_t^{min}]}(\hat{S}_t) + (S_t^{max})^2 \cdot \mathbb{1}_{[S_t^{max}.+\infty]}(\hat{S}_t) \tag{10}$$

where the indicator functions squared yield only binary outcomes and hence, for simplicity, not shown as squared.

Taking expectation of equations (9) and (10) enables the derivation of expressions for storage first and second statistical moments, respectively. For the assumption of Gaussian statistically independent random inflows, such expressions are presented below (see Fletcher and Ponnambalam (2008) for their detailed derivations; see Mahootchi et al. (2010) for removing the Gaussian assumption to arbitrary non-gaussian inflows modelled by the Kumaraswamy distribution):

$$\mathbb{E}(S_t) = \frac{\bar{I}_t - k_t}{2}[erf(UB) - erf(LB)] - \sqrt{\frac{Var(\eta_t)}{2\pi}}[exp(-UB^2) - exp(-LB^2)] + \frac{S_t^{min}}{2}[1 + erf(LB)] + \frac{S_t^{max}}{2}[1 - erf(UB)] \quad (11)$$

$$\mathbb{E}(S_t^2) = \frac{(\bar{I}_t - k_t)^2}{2}[erf(UB) - erf(LB)] + 2(\bar{I}_t - k_t)\sqrt{\frac{Var(\eta_t)}{2\pi}}[exp(-UB^2) - exp(-LB^2)] - \sqrt{\frac{Var(\eta_t)}{2\pi}}\left\{[S_t^{max} - (\bar{I}_t - k_t)]\, exp(-UB^2) - \left[S_t^{min} - (\bar{I}_t - k_t)\right] exp(-LB^2)\right\} + \frac{Var(\eta_t)}{2}[erf\,(UB) - erf(LB)] + \frac{\left(S_t^{min}\right)^2}{2}[1 + erf(LB)] + \frac{(S_t^{max})^2}{2}[1 - erf(UB)] \quad (12)$$

where $\mathbb{E}$ denotes the expectation operator, η_t is a zero-mean random variable following a Gaussian distribution of the form $N(0, Var(\eta_t))$, $LB = \frac{S_t^{min} - (\bar{I}_t - k_t)}{\sqrt{2Var(\eta_t)}}$, $UB = \frac{S_t^{max} - (\bar{I}_t - k_t)}{\sqrt{2Var(\eta_t)}}$, and erf is the error function (see Appendix B for more details). It is important to note that the right-hand side of equations (11) and (12) are only a function of k_t, the decision variable, and other known values and hence is easily evaluated and are not considered in constraints as in the original formulation the FP method.

2.2 Implementation of the FP Method

Since the new implementation of the FP method explicitly accounts for the role of deficit and spill terms to model a quadratic objective function exactly, we first derive the expressions for these terms.

2.2.1 New expressions for the moments of deficit and spill

As defined previously, the actual total reservoir outflow $r_t = U_t + Sp_t - \delta_t$ is the proposed release U_t accounting for either deficit or spill. The deficit and spill terms can be determined as follows:

$$\delta_t = \left(S_t^{min} - \hat{S}_t\right) \cdot \mathbb{1}_{[-\infty, S_t^{min}]}\left(\hat{S}_t\right) = \left[S_t^{min} - (\bar{I}_t - k_t) - \eta_t\right] \cdot \mathbb{1}_{[-\infty, S_t^{min}]}\left(\hat{S}_t\right) \quad (13)$$

$$Sp_t = \left(\hat{S}_t - S_t^{max}\right) \cdot \mathbb{1}_{[S_t^{max}, +\infty]}\left(\hat{S}_t\right) = [(\bar{I}_t - k_t) - S_t^{max} + \eta_t] \cdot \mathbb{1}_{[S_t^{max}, +\infty]}\left(\hat{S}_t\right) \quad (14)$$

Squaring the above expressions yields

$$\delta_t^2 = \left[S_t^{min} - (\bar{I}_t - k_t)\right]^2 \cdot \mathbb{1}_{[-\infty.S_t^{min}]}\left(\hat{S}_t\right) - 2\eta_t\left[S_t^{min} - (\bar{I}_t - k_t)\right] \cdot \mathbb{1}_{[-\infty.S_t^{min}]}\left(\hat{S}_t\right) + \eta_t^2\, \mathbb{1}_{[-\infty.S_t^{min}]}\left(\hat{S}_t\right) \tag{15}$$

$$Sp_t^2 = [(\bar{I}_t - k_t) - S_t^{max}]^2 \cdot \mathbb{1}_{[S_t^{max}.+\infty]}\left(\hat{S}_t\right) + 2\eta_t[(\bar{I}_t - k_t) - S_t^{max}] \cdot \mathbb{1}_{[S_t^{max}.+\infty]}\left(\hat{S}_t\right) + \eta_t^2 \cdot \mathbb{1}_{[S_t^{max}.+\infty]}\left(\hat{S}_t\right) \tag{16}$$

Taking expectation of all above equations enables the derivation of expressions for first and second statistical moments of deficit and spill. Such expressions are presented below and their detailed derivation is given in Appendix A.

$$\mathbb{E}(\delta_t) = \left[S_t^{min} - (\bar{I}_t - k_t)\right] \cdot \int_{-\infty}^{S_t^{min}-(\bar{I}_t-k_t)} f_{\eta_t}(\eta_t)d\eta_t - \int_{-\infty}^{S_t^{min}-(\bar{I}_t-k_t)} \eta_t f_{\eta_t}(\eta_t)d\eta_t \tag{17}$$

$$\mathbb{E}(\delta_t^2) = \left[S_t^{min} - (\bar{I}_t - k_t)\right]^2 \cdot \int_{-\infty}^{S_t^{min}-(\bar{I}_t-k_t)} f_{\eta_t}(\eta_t)d\eta_t - 2\left[S_t^{min} - (\bar{I}_t - k_t)\right] \cdot \int_{-\infty}^{S_t^{min}-(\bar{I}_t-k_t)} \eta_t f_{\eta_t}(\eta_t)d\eta_t + \int_{-\infty}^{S_t^{min}-(\bar{I}_t-k_t)} \eta_t^2 f_{\eta_t}(\eta_t)d\eta_t \tag{18}$$

$$\mathbb{E}(Sp_t) = [(\bar{I}_t - k_t) - S_t^{max}] \cdot \int_{S_t^{max}-(\bar{I}_t-k_t)}^{+\infty} f_{\eta_t}(\eta_t)d\eta_t + \int_{S_t^{max}-(\bar{I}_t-k_t)}^{+\infty} \eta_t f_{\eta_t}(\eta_t)d\eta_t \tag{19}$$

$$\mathbb{E}(Sp_t^2) = [(\bar{I}_t - k_t) - S_t^{max}]^2 \cdot \int_{S_t^{max}-(\bar{I}_t-k_t)}^{+\infty} f_{\eta_t}(\eta_t)d\eta_t + 2\,[(\bar{I}_t - k_t) - S_t^{max}] \int_{S_t^{max}-(\bar{I}_t-k_t)}^{+\infty} \eta_t f_{\eta_t}(\eta_t)d\eta_t + \int_{S_t^{max}-(\bar{I}_t-k_t)}^{+\infty} \eta_t^2 f_{\eta_t}(\eta_t)d\eta_t \tag{20}$$

where $f_{\eta_t}(\eta_t)$ is the probability density function of inflow random component η_t. Based on equations (19) and (20), analytical expressions for the first and second moments of deficit and spill have been derived and presented in Appendix B for Gaussian inflows. An important aspect to notice is that all expressions become a function of the known inflow moments, system storage bounds S_t^{max} and S_t^{min}, and the LDR parameters k_t, which are the only decision variables to optimize.

2.2.2 Optimization problem formulation

A monthly stochastic reservoir operation optimization problem may be formulated with the objective of minimizing the expected value of the sum of squared deviations between releases and demands:

$$minimize \quad Z = \mathbb{E}\left[\sum_{t=1}^{12}(r_t - D_t)^2\right] = \mathbb{E}\left[\sum_{t=1}^{12}(U_t + Sp_t - \delta_t - D_t)^2\right] \tag{21}$$

in which D_t is the target demand for the month t. Adding terms for storage targets and minimizing the sum of deviations in (21) is possible and has been dealt with in Fletcher and Ponnambalm (1998); also see Section 2.2.4 later for other general nonlinear objective functions. The inclusion of deficit and spill terms in the objective function means that both water supply and flood control are important for the operation (if the only objective is water supply, then R_t may be used instead of r_t and the objective function becomes $Z = \mathbb{E}\left[\sum_{t=1}^{12}(R_t - D_t)^2\right] =$

$\mathbb{E}\left[\sum_{t=1}^{12}(U_t - \delta_t - D_t)^2\right]$. Currently, release bound constraints are not considered during optimization because the objective function penalizes both spills and deficits.

The assumed LDR is $U_t = S_{t-1} + k_t$; therefore, the objective function becomes

$$Z = \mathbb{E}\left[\sum_{t=1}^{T=12}[S_{t-1} + Sp_t - \delta_t + (k_t - D_t)]^2\right] \tag{22}$$

which can be developed to

$$Z = \sum_{t=1}^{12}[\mathbb{E}(S_{t-1}^2) + 2(k_t - D_t)\cdot\mathbb{E}(S_{t-1}) + (k_t - D_t)^2 + \mathbb{E}(\delta_t^2) - 2\cdot\mathbb{E}(S_{t-1}\cdot\delta_t) - 2(k_t - D_t) \cdot\mathbb{E}(\delta_t) + \mathbb{E}(Sp_t^2) + 2\cdot\mathbb{E}(S_{t-1}\cdot Sp_t) - 2\cdot\mathbb{E}(Sp_t\cdot\delta_t) + 2(k_t - D_t)\cdot\mathbb{E}(Sp_t)] \tag{23}$$

Since for any time period t, either Sp_t or δ_t is zero, then $\mathbb{E}(Sp_t\cdot\delta_t) = 0$. Assuming S_{t-1} to be independent of both δ_t and Sp_t, the objective function finally becomes

$$Z = \sum_{t=1}^{12}[\mathbb{E}(S_{t-1}^2) + 2(k_t - D_t)\cdot\mathbb{E}(S_{t-1}) + (k_t - D_t)^2 + \mathbb{E}(\delta_t^2) - 2\cdot\mathbb{E}(S_{t-1})\cdot\mathbb{E}(\delta_t) - 2(k_t - D_t)\cdot\mathbb{E}(\delta_t) + \mathbb{E}(Sp_t^2) + 2\cdot\mathbb{E}(S_{t-1})\cdot\mathbb{E}(Sp) + 2(k_t - D_t)\cdot\mathbb{E}(Sp_t)] \tag{24}$$

The assumption on independence of S_{t-1} and both δ_t and Sp_t was indeed verified using simulation results, and it was found that corresponding Spearman correlation coefficients that measure nonlinear dependence better were very low for the case studied. We have also provided some insight on the validity of this assumption for other systems in Section Final Remark. As all first and second moments in the above expression are dependent only on k_t, so is the objective function Z. Consequently, the vector of decision variables of the final optimization problem is $\boldsymbol{k} = \{k_1, \cdots, k_{12}\}^\top$, i.e., one value of k_t for each month of the year, $t = 1$ (January) through $t = 12$ (December). The symbol $\top$ represents the vector transpose operator.

The value of k_t in $U_t = S_{t-1} + k_t$ may be negative (storage has enough water to meet proposed release) or positive (storage needs additional water to meet proposed release). Thus, the decision vector $\boldsymbol{k}$ may be unbounded, i.e., $-\infty \leq \boldsymbol{k} \leq +\infty$. Since the only decision variables are the elements of vector $\boldsymbol{k}$, which is unbounded, we face an unconstrained nonlinear optimization problem. This formulation can be easily vectorized as detailed in Appendix C.

Once the optimal values of the LDR parameters $k_1, \cdots, k_{12}$ are found, the monthly values of first and second moments (variances) of storage, deficit and spill variables can be calculated by the derived expressions presented earlier. Furthermore, the probabilities of containment ($\mathbb{P}_t^{\text{con}}$), deficit ($\mathbb{P}_t^{\text{def}}$) and spill ($\mathbb{P}_t^{\text{sp}}$) for the projected storage $\hat{S}_{t'}$ are simply the expected values of the three indicator functions in equation (5) as presented in the FP method (Fletcher and Ponnambalam, 2008) whose expressions are also known (see also appendices A and B).

The assumed LDR can be used as a guide to operate the reservoir. For a given initial storage S_{t-1} and inflow I_t, a total release of $U_t = S_{t-1} + k_t$ is proposed. The actual total outflow for that month, $r_t = U_t + Sp_t - \delta_t$, can then be decided by checking the mass balance to identify whether spill or deficit should be triggered.

Note that in previous applications of the FP method, a zeroth-order Taylor series expansion of the objective function has been used where neither second moments of storage nor deficit and spill terms have been used leading to the following approximation:

$$
\begin{aligned}
Z_1 = \mathbb{E}\left[\sum_{t=1}^{12}(U_t - D_t)^2\right] &\approx \sum_{t=1}^{12}[\mathbb{E}(U_t) - D_t]^2 \\
&= \sum_{t=1}^{12}[\mathbb{E}(S_{t-1} + k_t) - D_t]^2 \\
&= \sum_{t=1}^{12}[\mathbb{E}(S_{t-1}) + (k_t - D_t)]^2 \\
&= \sum_{t=1}^{12}\left[\left(\mathbb{E}(S_{t-1})\right)^2 + 2(k_t - D_t)\cdot\mathbb{E}(S_{t-1}) + (k_t - D_t)^2\right]
\end{aligned}
\tag{25}
$$

In the above equation, only the first moment of storage is needed. $\mathbb{E}(S_{t-1})$ has already been estimated by equation (11) considering storage bounds using indicator functions. However, a more exact objective function estimate from equation (24), if spill and deficit terms are omitted, is

$$
Z_2 = \sum_{t=1}^{12}[\mathbb{E}(S_{t-1}^2) + 2(k_t - D_t)\cdot\mathbb{E}(S_{t-1}) + (k_t - D_t)^2] \tag{26}
$$

Function Z_2 requires the second moment of storage $\mathbb{E}(S_{t-1}^2)$, given in equation (12). The difference between equations (26) and (25) (objective functions Z_2 and Z_1) is simply equal to $\mathbb{E}(S_{t-1}^2) - \left(\mathbb{E}(S_{t-1})\right)^2 = Var(S_{t-1})$, i.e. the variance of the initial storage. Therefore, as expected, the larger the variance of the monthly storage, the greater the error in the zeroth-order Taylor approximation will be.

2.2.3 Reduction in computing time

It was noted in Section 2.2.2 that, at any given time t, all necessary expressions can be calculated as simply as a function of decision vector $\boldsymbol{k}$ leading to vectorization possibilities shown in Appendix C. The speed up one gets is a function of the software that we use, but

MATLAB ® allows for such vectorized calculations and executes much faster than nonvectorized equivalents.

However, by removing the constraints, the time to solve significantly decreases. Even in a linear programming case, the time complexity is $O(NV^3)$ where NV is the number of variables to be optimized. Here (i) we remove the $O(NV)$ constraints that was used to define the first and second moments in the original FP formulation (Fletcher and Ponnambalam, 2008)), and (ii) because of (i), the number of variables becomes 1/3 of the original number of variables as the expressions for the moments are simply calculated and defined only by the decision vector $\boldsymbol{k}$. The time to solve now reduces to $O((\frac{1}{3}NV)^3)$; therefore, the speed up of this current formulation compared to the original formulation is at least 27 times.

2.2.4 Other nonlinear objective functions

The derivation for the quadratic objective function in equation (21) produced equation (24) which required only the first and second order moments of any decision and storage variables; all of them are available in explicit analytical forms in the FP method. On the other hand, for other functions that are nonlinear but not quadratic, it is possible to use the First-order Second Moment Taylor-Series methods to approximate such objective functions as they only need the first and second order moments of the required variables such as release and water level (say hydraulic head for hydropower operations) that is related to storage nonlinearly) and are available. One can also use the approximations suggested in Loucks and van Beek (2017, page 504), which is simpler as it linearizes the equations around the mean values of variables and uses only the zeroth order Taylor-Series terms. The advantage of these approaches is that the point at which linearization is done is at the mean values of the storage and release variables that are available and are continuously updated as the optimization proceeds. The problems we solve are nonconvex so there is no global optimality guarantee, but the use of Monte Carlo simulations help us validate the accuracy of the estimates between the FP method and the corresponding simulation results.

Apart from the possibility explained above, we can easily use an extended form of the objective function in which deviations from both target releases (water demands) and target storages ($Starg$) are included as $\mathbb{E}\left[\sum_{t=1}^{12}((r_t - D_t)^2 + (S_{t-1} - Starg_t)^2)\right]$. This objective function accounts for both release- and storage-dependent purposes such as navigation, recreation, and hydropower operations in many practical real-world problems. Note that all we need in the above function are the newly derived first and second moments of release, including spill and deficit, and the already derived moments of storage variables in the FP model. Additionally, we have probabilities of spills and deficits that can be utilized in the model formulation for risk-based operation or other specific planned purposes. To the best of our knowledge, there is no other explicit optimization model available where such terms and information are available with high accuracies as shown in the results in many Figures in this paper.

2.2.5 Extension to multireservoir systems

The FP method extended to multireservoir systems still has the linear time complexity thus avoiding any curse of dimensionalities. The derivation of the means and variances of storage states of multireservoir systems using the model of Fletcher and Ponnambalam (2008) has been already presented in Mahootchi et al. (2010) solving a five-reservoir system for both Gaussian and non-Gaussian inflows. However, the objective function in that work was linear and the second moments of spills and deficits were not included in water balance equations. In order to extend the FP method and the vectorized implementations presented here to multireservoir systems or extending the previous multireservoir systems method to consider other objective functions is now possible. The only changes needed are in the objective function and the use of the moments of spills and deficits as presented here and is left for the future. The use of the linear decision rule removes the dependence of releases on the storage volumes and hence the multireservoir expressions are much simpler than in Fletcher and Ponnambalam (1998) that solved the operations optimization problem of the Great-Lakes system considering five of the lakes using standard operating policy.

2.3 Two-stage Programing (TSP)

The open loop constant-release policy (no direct dependence on the storage state) and the S-type linear decision rule (dependent on the storage state) are the policies used in previous (in 1996) and current FP (since 2008) models, respectively. These decision rules make the model formulation tractable so as to derive analytical expressions for different variables of interest. Inflows were also assumed to be normally distributed and statistically independent, ignoring serial (persistence) correlations. To assess how these assumptions impact the performance of the proposed FP model, we compare it with other methods including SDP in which a more general state-dependent policy is available and the TSP, as the implicit stochastic optimization counterpart of the FP, and ISP. The TSP method can also be used to easily account for a variety of operating policies and to consider non-Gaussian serially-correlated inflow time series as TSP can be implemented under any inflow scenarios. Such comparisons have been made in Mahootchi et al. (2010 and 2012) presenting good comparable results for all the three methods; however, the original objective function was linear where the current extension to quadratic objective function was not needed as well as deficits and spills and comparisons with general SQ-type and policy-free policies were not considered.

Following is the TSP model's formulation to compare with the FP method description above. The formulation implemented here for random inflow scenarios uses the fan-type as against the tree-type scenarios. Although both fan and tree types give good results (Séguin et al. 2017), fan-type scenario generations are more commonly used in water resources as Monte Carlo simulations:

$$min\, Z = min\left\{\frac{1}{N}.\sum_{\mathrm{i}=1}^{N}\sum_{t=1}^{T=12}\left[\left(U_t - D_\mathrm{t} + \delta_t^i - Sp_t^i\right)^2\right]\right\} \qquad (27)$$

Subject to the constraint set for each scenario:

$$S_t - S_{t-1} + U_t + \delta_t^1 - Sp_t^1 = I_t^1 \qquad \text{for } \; t = 1, \ldots . T$$

$$S_t - S_{t-1} + U_t + \delta_t^2 - Sp_t^2 = I_t^2 \qquad \text{for } \; t = 1, \ldots . T$$

.

.

.

$$S_t - S_{t-1} + U_t + \delta_t^N - Sp_t^N = I_t^N \quad \text{for } t = 1, \ldots. T$$

$$S_0 = S_T \tag{28}$$

where N is the number of scenarios (years). Other variables are already defined in the FP method description. In this original TSP model that we have named it later Model TSP1, storage variables (S_t) and the proposed release variables (U_t) are first-stage variables and do not change with the scenario number (year). However, since for each scenario or sample (one year with 12 months), inflow to the reservoir in a season (month) is different, the second-stage variables of surplus (Sp_t^i) and deficit (δ_t^i) are added to the balance equations of each scenario i to keep the model feasible. We also examine other versions of TSP depending on the variability of storage variables over different scenarios and the release operating policies adopted.

We can set a specific type of release operating policy in the TSP model by replacing U_t with the form or equation of that policy. For example, if equation $U_t = S_{t-1} + k_t$ is added to the above formulation, we will have a TSP model equipped by the S-type release operation policy, the same policy employed in the FP model.

Another well-known stochastic optimization method we use to compare the FP model with is the SDP, which works based on Bellman's principle of optimality and solves a recursive form of the objective function for different discrete values of the state variables vector within the state-decision space. We do not present here the SDP model formulation of the problem as SDP is a well-documented approach (see for example Vedula and Mujumdar (2005) or Loucks and van Beek (2017)). Note that SDP faces the curse of dimensionality in multireservoir problems due to discretization of the state variables.

3. Case Study and Data

One of the most important hydropower plants in Brazil's grid is the one located at the Sobradinho dam within the São Francisco River Basin (SFRB), in the southeast and northeast regions of the country (Figure 2). The reservoir holds approximately the same volume of water as the Three Gorges dam in China, which of course has a much higher flow capacity due to being built on the Yangtze river. The SFRB covers an area of approximately 640,000 km² (7.5% of the Brazilian National Territory), extending over six Brazilian states. Fifty-eight percent of the SFRB includes part of the so-called *Polygon of Droughts* (semiarid region), characterized by critical periods of prolonged droughts as a result of low rainfall and high evapotranspiration (Agência Nacional de Águas, 2015).

Figure 2. Map of the São Francisco River Basin and location of the Sobradinho dam.

The Sobradinho reservoir provides multiyear regulation of the São Francisco river with a minimum flow of 2,060 m³/month, allowing the full utilization of other hydroelectric plants located downstream. The 34-billion-cubic-meter capacity of the Sobradinho reservoir floods an area of 4,241 km², with a length of 327.5 km at the 393.5-meter water level (design flood), forming the largest artificial lake in Latin America (Oliveira Dantas, 2005). Hydropower production plays an important role in the SFRB. We, however, assume the Sobradinho reservoir is to be used only for water supply to test the new version of the FP model, and hence the assumption of constant demand. Note that this is not a limitation of the method as shown in equation (26), and it only facilitates focusing on the main purpose of this study. As explained in Section 2.2.4, terms penalizing deviations from target storage levels can easily be added to the objective function of the proposed FP model to consider other storage-related purposes such as recreation and hydropower operations.

In order to satisfy the real-world system requirement of constant releases (see above), a constant demand equal to 80% of the mean annual flow was specified as D_t. First, the FP method was run in order to find the best LDR parameters together with the estimation of first and second moments as well as probabilities of containment, deficit and spill. Later, a 1,000-year monthly reservoir operation implementing the derived LDR operating policy was carried out (as explained in section 2.2.2). From this Monte Carlo simulation, values of moments and frequencies were calculated to be compared with those already found by the FP method. The 1,000-year monthly inflow scenario was synthetically generated from historical records provided by government-established National Electric System Operator (ONS – Operador Nacional do Sistema Elétrico) for the period 1931–2015 (85 years).

4 Results and Discussion

In the original FP model, in the objective function, neither the second moment of storages nor the first and second moments of spill and deficit have been incorporated into the model's formulation so far as done here in equation (26); however, zeroth Taylor Series expansion has been used in the objective function, for example in Fletcher and Ponnambalam (1996) for solving the Great Lakes storage and release target operations optimization.

In the application to the case study, the role played by each new element is presented separately in Appendices D and E. Here we discuss the results of full equation (26) or what is called Model 4 in Appendix E. Initially we assume that the inflows at each month in Sobradinho follow a Gaussian distribution. That means that the 1,000-year monthly inflow scenario is generated from Gaussian-distributed random numbers with same mean and standard deviation as historical records to simulate the reservoir operation. This simulation is conducted using the optimal policies obtained by the FP method (optimal k_t values) to assess how optimal solutions of the discretization-free FP method perform under different conditions and assumptions, for example inclusion or not inclusion of first or higher moments of storage, deficit and spill variables in its formulation. We present results corresponding to the actual inflow data from Sobradinho later in Section 4.4.

4.1. Analysis of Release Operating Policies

Simple release policies of S-type, $U_t = S_{t-1} + k_t$, and open-loop, $U_t = k_t$, have been used in the current and previously-developed FP models, respectively. One can argue that such simple operation policies may not be efficient enough. Therefore, the question to be assessed here is whether these simple policies affect the FP Model's performance drastically, compared to other stochastic optimization models such as SDP employing more sophisticated, nonlinear, state-dependent policies. This comparative analysis of various operating policies with FP results is new. We compare in this section the proposed FP model with SDP, TSP and policy-free ISP approaches. The reason is that we can easily develop different versions of TSP or ISP accounting for different operation policies from the simplest constant-release policy to a policy-free model. Therefore, comparison of TSP and ISP models, as the implicit stochastic counterpart of the FP model, with the FP model when their difference is only in their release operating polices can quantify what impact using those simple linear policies will have on the performance of the FP model. To do so, the following five alternative models are tested:

1) The original TSP (TSP1-open/TSP1-S-type) in which a constant-release open loop/S-type policy is considered, respectively. For a typical year with 12 seasons (months), the number of release decision variables in TSP1 is 12 with additional 12+1=13 storage variables; these are called the first-stage variables and do not vary from one scenario to another. To these 2×12× N additional surplus and deficit decision variables are added, where N is the number of scenarios (years).
2) TSP2-(open/S-type) considers reservoir storage volumes to vary over both seasons and scenarios (years). It means that in addition to 12 constant release decision variables (which are now the only first-stage variables), 12× N +1 storage volumes are also decision variables to be optimized (so now these are second stage variables). This allows for storage variances to be non-zero like the FP method when the second moments of storages are accounted for. In other words, the FP model is the explicit stochastic equivalent version of the TSP2 model.
3) TSP3 is similar to TSP2 in which a more general complete release rule called general SQ-Type, $r_t = S_{t-1} + kk_t \times I_t - k_t$, is employed. Traditional SQ-Type policy, where kk_t=1 for all months, has already been used in chance-constrained programming (Loucks

1980, Mousavi et al. 2010). Therefore, here the variables kk_t and k_t are the first-stage variables.

4) The last one, the implicit stochastic optimization (ISP), allows all release variables to vary both over different seasons and scenarios (years). It is a policy-free model in terms of release rules (Mousavi et. al 2010) in which time series of releases are among unknown decision variables. In this model, deficits or surpluses are part of total releases, so no need to define and consider them as separate variables. The importance of this model is because it provides the best possible objective function value that can ever be reached as it does not impose any additional constraint (release policies) on the TSP optimization model, and it benefits from having perfect foresight on future inflows. Therefore, any other model utilizing even a very sophisticated nonlinear state-dependent release policy cannot perform better than this model, and its global optimum objective function value will be the upper bound of the best possible objective function value. Therefore, comparison of TSP1, TSP2, TSP3, and FP models with such a policy-free ISP model will show what impact the release operation policies used in each of them can have on the optimality of their solutions.

The number of variables of each method and sample CPU times for the FP, SDP, and TSP2 methods are presented in Table 1 and Table 3, respectively. It is clear that when the number of reservoirs increases, the number of variables in the FP method increases linearly (see also Mahootchi et al. (2010) for solving a five reservoir problem with FP method) while other methods face the curse of dimensionality and cannot be solved.

One important point for TSP1, TSP2, and TSP3 models is that if we don't make them forced to activate surplus/deficit variables (second-stage variables) only if the end-of-month storage volume reaches the upper/lower bound of the reservoir storage volume, then they will be exactly the same as ISP because of the freedom of surplus and deficit variables to take any arbitrary values in the balance equations. Additionally, in each period, simultaneously spill and deficit terms cannot be nonzero. To account for these requirements, three additional penalty terms were added to the objective function of the TSP models as follows where Z is the same as in equation (27):

$$Minimize\ Z' = Z + W_1 \times \sum_{i=1}^{N}\sum_{t=1}^{12}\left[Sp_t^i \times (S_t^{max} - S_t^i))\ \ \right] + W_2 \times \sum_{i=1}^{N}\sum_{t=1}^{12}\left[\delta_t^i \times (S_t^i - S_t^{min}))\ \ \right]$$
$$+ W_3 \times \sum_{i=1}^{N}\sum_{t=1}^{12}\left[Sp_t^i \times \delta_t^i\right] \qquad (29)$$

The second and the third terms in the above formula ensure spill (surplus) and deficit variables, i.e. $Sp_t^i\ or\ \delta_t^i$, are not triggered until $S_t^i = S_t^{max}$ and S_t^{min} , respectively, and the last term guarantees the spill and deficit terms do not take positive values concurrently. Our experiments showed that $W_1 = W_2 = W_3 = 1$ worked well. Table 1 presents the results in terms of objective function values (both in simulation and optimization) for all the models. To be fair and focus only on the role of operations policies, we have calculated SDP transition probabilities using a 125-year synthetic Gaussian inflow series. This is because other models' results being reported are also for Gaussian inflows. Later in the next section we present the SDP model results for correlated non-Gaussian historical inflows. Note that CPU time reported for the SDP method corresponds to NI =7 inflow classes (resulted in the best obj. function in optimization/simulation), $NS = 30$ discrete storages, and $Niter$ =10 cycles to reach steady-state conditions.

Table 1: Comparison of FP, SDP, TSP, and ISP models for different operation policies

Model	Description	No of decision variables	Release Operation Policy	Obj. function value in optimization		Obj. function value in simulation		
				Sample size (N) for TSP		Sample size		
				55	125	55	1,000	125
FP	ESO	12	S-type	27.08		29.22	27.33	27.87
SDP	ESO	NS=30, NI= 7, and Niter=10	State depended policies as $R^*_t(S_t, I_t)$	27.32		28.65	26.95	27.39
TSP1-open	ISO	25+24×N	open loop	29.62	28.299	30.74	29.53 (by 55) 29.15(by 125)	29.26
TSP1-Stype	ISO	25+24×N	open loop	29.62	28.30	30.15	28.86 (by 55) 28.55(by 125)	28.92
TSP2-open	ISO	13+36×N	Open loop	28.82	30.10 (stopped after 200000 iterations)	28.82(from 55) 30.88(from 125)	27.37(from 55) 29.99(from 125)	30.10(from 125) 27.72(from 55)
TSP2-Stype	ISO	13+36×N	S type	28.95	28.34	28.95(from 55) 29.30 (from 125)	27.40 (from 55) 27.99(from 125)	28.34(from 125) 27.84(from 55)
TSP3	ISO	25+36×N	General s-q type	28.65	27.55	28.65(from 55) 28.72(from 125)	27.16(from 55) 27.10 (from 125)	27.55(from 125) 27.59(from 55)

ISP	ISO	24×N+1	Free policy	27.80	26.63	27.80(only from 55)	-	26.63(only from 125)

From above results, one can see that as expected the best objective function value is that of policy-free ISP model (26.63), and the differences among the models' solutions both in optimization and the simulation are between 1-15%, and the worst is TSP1-open (ignoring the unfinished TSP2-open). Additionally, the TSP2-Stype's (and TSP2-open's) objective function value is ~8 % worse than the best possible result that can ever be achieved which is that of ISP. These results clearly indicate that simple open-loop or S-type release policies employed in original or current extended FP models perform quite well (the difference in the long term simulation with the best ISP policy is 2.6%) and close to the best possible state-dependent more sophisticated release policies of SDP. The FP's open loop policy is slightly better than FP's S-Type policy but leads to more complicated expressions, especially for mutireservoir systems (Fletcher and Ponnambalam, 1996) and is not clear that it is worth losing simplicity in practice. Therefore, the concern about using simple optimal release rules in the proposed extended FP model is not really important at least for the problem approached, which is a long-term optimal reservoir operation planning problem. On the other hand, FP can solve multireservoir problems very fast, while most of the other methods have to use other approximations even to solve multireservoir problems. The approximations are either in modeling the system, e.g. in the aggregation method of Turgeon (1981) and Ponnambalam and Adams (1987, 1996) as explicit stochastic programming (ESP) methods, or by using a reduced number of scenarios in ISP, which also produces suboptimal solutions.

4.2 Performance Assessment for Correlated Inflows

In this section, we show the application of the proposed formulation and implementation of the proposed FP model to the Sobradinho reservoir system without assuming that the synthetic inflows used in simulation follow a Gaussian distribution as in Section 4.1. This is because another concern with the proposed FP model is that of assuming serially independent Gaussian inflows. Of course FP model is not restricted to only Gaussian inflows and can easily applied by other distributions such as Kumaraswamy distribution (Mahootchi et al., 2010). However, it is yet to be extended to cases considering serial and cross correlations. Therefore, in this section we want to assess how significant the role of such simplification would be compared to models accounting for inflows persistence such as SDP.

Now, the 1,000-year monthly inflow scenario for simulation is synthetically generated by the Method of Fragments (Svanidze, 1980) trying to preserve the actual inflow structure of the historical records. Figure S1 included in the supporting information shows comparison of mean and standard deviation of historical inflow records against synthetic scenario values, indicating that historical monthly means and standard deviations were properly preserved in the generated scenario.

The final equations for storage⁄deficit⁄spill moments as well as those for probabilities were still derived assuming normality of inflows for each month of the year (January–December). Therefore, Lilliefors tests for normality (Lilliefors, 1967) were performed for each month in the inflow records. Figure S2 included in the supporting information shows the results from the tests together with normality plots, indicating that normality is reasonable only for

January, October, November and December. Inflow data for all other eight months were rejected to follow a Gaussian distribution.

After running the vectorized FP model optimization with input data from the Sobradinho reservoir, the following results (shown in Tables S1 and S2 in the supporting information) were obtained for every month of the year ($t = 1, \ldots, 12$):

- LDR parameters (k_t);
- First ($\mathbb{E}(S_t)$) and second moments ($\mathbb{E}(S_t^2)$) as well as variance ($Var(S_t)$) of storage;
- Probabilities of containment ($\mathbb{P}_t^{\text{con}}$), deficit ($\mathbb{P}_t^{\text{def}}$), and spill ($\mathbb{P}_t^{\text{sp}}$);
- First ($\mathbb{E}(\delta_t)$) and second moments ($\mathbb{E}(\delta_t^2)$) as well as variance ($Var(\delta_t)$) of deficit;
- First ($\mathbb{E}(Sp_t)$) and second moments ($\mathbb{E}(Sp_t^2)$) as well as variance ($Var(Sp_t)$) of spill.

Next, same statistics (M_1 and M_2 stand for first and second moments, respectively) were calculated using the optimal values of k_t by conducting a simulation model under the generated 1,000-year inflow scenario. Therefore, the FP model results were validated if they were close to those obtained by the long-period simulation in terms of the objective function value and the storage⁄deficit⁄spill moments as well as probabilities of containment⁄deficit⁄spill. Figure 3 compares the FP model optimization and simulation results when the FP optimal policies derived under Gaussian inflow assumption are simulated against a 1,000-year independent non-Gaussian inflow series. The agreement is very good, and the difference between optimization and simulation objective function values is just 0.32%. The only major issue was an underestimation of the moment of spill for the month of March (optimization provided $\mathbb{E}(Sp_t) = 0.0081$ against the simulated $M_1(Sp_t) = 0.0401$, as displayed in Tables S2 and S4 in the supporting information, respectively).

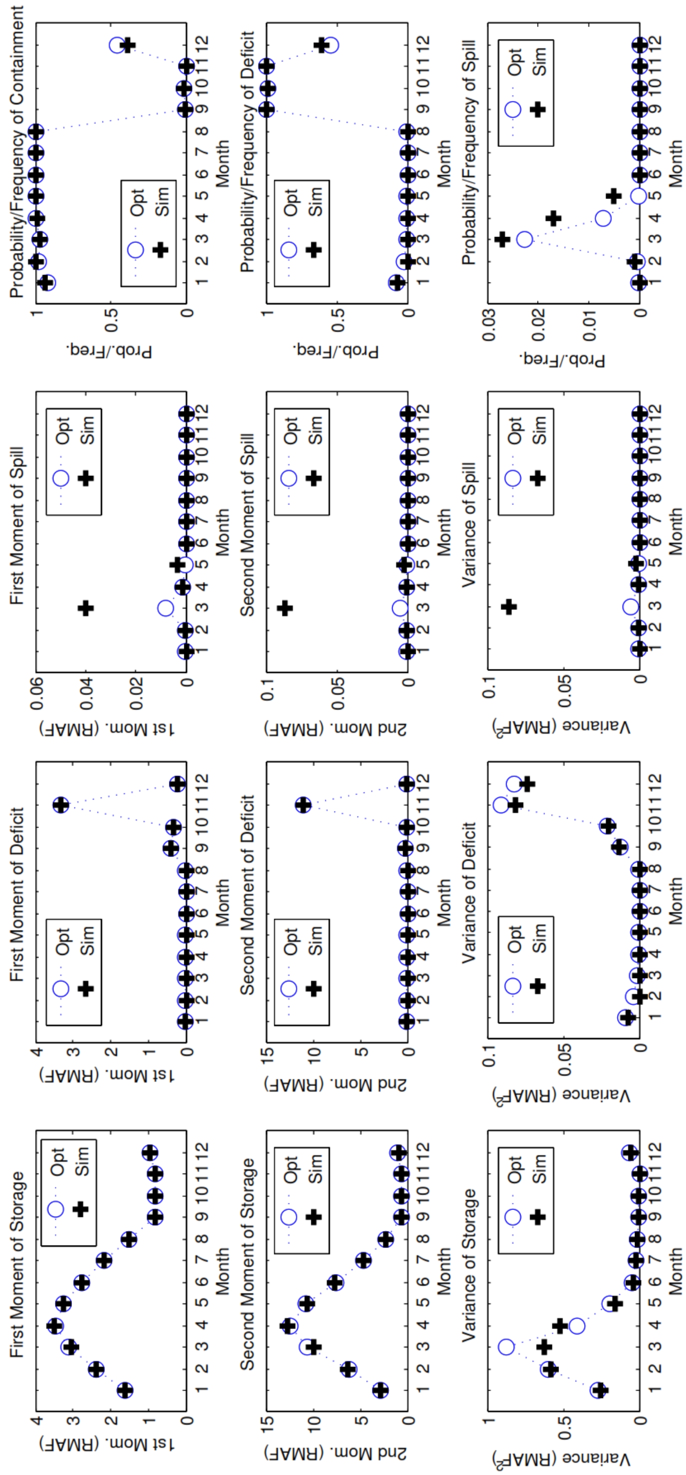

Figure 3. Comparison of FP optimization and simulation results when the optimal policies derived under Gaussian inflow assumption are simulated against a long-term non-Gaussian inflow series

Therefore, normality assumption has not been a restriction in the FP model for the case studied. Moreover, FP model provides accurate estimations of random variables up to second moments and also accurate estimations of probabilities of important storage states. However, to further investigate the issue and to quantify the impact of both normality and independence of inflows assumptions on the optimal polices derived by the FP model, we subsequently compare the results of FP, SDP, and TSP2 methods against different inflow scenarios. These scenarios include non-normal, serially-correlated historical inflow times series having lag-1 serial correlation coefficients as reported in Table 2. Among different TSP models, TSP2 is used here because it is the implicit stochastic optimization counterpart model of the FP model as both of them consider non-zero second moments of storages and employ S-type operating policy.

Table 2: Serial correlation coefficients of historical time series

Month	1	2	3	4	5	6	7	8	9	10	11	12
Corr. Coef.	0.54	0.76	0.65	0.82	0.93	0.97	0.98	0.95	0.84	0.73	0.66	-0.04

Different optimization and simulation experiments are conducted including 1) simulating the derived-by-FP policies against a) a 85-year historical inflow time series where inflows are neither Gaussian nor independent, b) a 85-year Gaussian independent synthetic inflow time series, and c) a 1,000-year Gaussian independent synthetic inflow time series, 2) running the TSP2 model using the 85-year Non-Gaussian correlated historical inflows, then simulating the resulting policies against the three inflow scenarios a-c, and 3) running the TSP2 model using the 85-year Gaussian independent synthetic inflow series, then simulating the resulting policies against the three mentioned a-c inflow scenarios. Additionally, SDP transitional probabilities are determined from the 85-year historical series (scenarios a) with $nclass = 7$, and its policies are simulated against scenarios a-c. Table 3 presents the results obtained using MATLAB ® in a Windows 10 Intel5 laptop:

Table 3: Analysis of the role of normality/non-normality and independence/dependence of random inflow series

Model	Obj. func. in optimization	simulation with 85-year historical non-Gaussian correlated inflows	simulation with 85-year uncorrelated Gaussian synthetic inflows	simulation with 1,000-year uncorrelated Gaussian synthetic inflows
FP	27.08	28.04	28.27	27.30
CPU seconds	1.44	1.46	---	---
SDP	27.49	27.84	28.22	27.29
CPU seconds	4.01	1.46	---	---
TSP2-Hist	27.92	27.92	28.43	27.50
CPU seconds	~6000	---	---	---
TSP2-Gauss	27.88	27.77	27.88	27.12

TSP2-Hist uses the 85-year historical monthly inflows, whereas TSP2-Gauss works with Gaussian independent synthetic inflow time series having the same length and same first and second moments as those of the historical time series. Therefore, in above results, 28.43 is about simulating optimal policies obtained from correlated historical inflows (85 years) against independent Gaussian inflows of the same size (85 years), and 27.77 is about simulating the policies obtained from 85-year Gaussian inflow series against 85-year correlated historical inflows. Additionally, 27.5038 is for simulation the policies obtained from 85-year historical inflows against a 1,000-year Gaussian independent series, whereas 27.1241 is about simulating the optimal policies obtained from 85-year Gaussian independent flows against a 1,000-year Gaussian independent inflow time series. We also mentioned that SDP policies have been derived using serially correlated non-Gaussian historical inflows, and they are then simulated against three different scenarios of correlated and non-correlated inflows.

The results presented in Table 3 demonstrate that the assumption of normality and independence for inflows do not have significant impacts on the optimal policies derived by the proposed FP and SDP models as the objective function values resulted from optimization and simulation under the examined scenarios are close and their differences are between 1-4%. Even if we cannot generalize such an outcome to all other case studies, we believe the same situation would be the case for long-term reservoir operation problems according to previous experiences such as Zhang and Ponnambalam (2005). A same analysis and examination can be carried out for a multireservoir system with respect to the impact of cross correlations of inflows, where the FP model has a significant advantage over other techniques such as SDP in dealing with the curse of dimensionality problem.

4.3. Final Remarks and Discussion

In terms of implementation, the FP method here needs only the LDR and equation (26) to be minimized as an unconstrained objective function while calculating moments in equations (11) and (12) and if necessary, the various probabilities can be calculated as well using equations in Appendix B. This extends to multireservoir systems in a similar form as in Mahootchi et al. (2010) but using an appropriate extension of equation (26). Hardly any stochastic method can be as simple as this method. Analyses and results presented in Sections 4.1 and 4.2 revealed that the FP method even under simplifying assumptions of linear decision rules and the non-correlated inflows still performed well for the case studied. We also elaborated on how the proposed FP method can deal with other objective functions accounting for storage-dependent purposes such as recreation and hydropower operations. However, we provide a brief discussion here on the applicability of the results for other problems including multireservoir systems.

Although more investigation is required regarding simple linear release rules assumption for large reservoir systems that carry storage crossing years under different inflow and demand variability and correlation conditions, we think the reason the simple linear policies works well (like in this case study where there are inter-annual storage happens) is that all future statistics are used when deriving the parameters k_t. Of course, if the inflow data is not stationary these parameters not varying over different years won't work, but that is a completely different problem which should be studied separately. Note that for this case studied here, inflows were highly correlated and demands are too as they were considered the same value for all periods.

While expanding the nonlinear quadratic objective function, we also assumed the beginning-of-month storage and deficit/spill in that month are independent which is another limitation that need to be considered further in the future. Our simulation experiments showed the validity of this assumption for most but not all months. The limitations of such investigations have been studied in Fletcher & Ponnambalam (2008) for systems having high probabilities of spill/deficit compared to probability of containment, i.e. systems with small storage capacity and high inflow variability that frequently become full and empty, or systems staying at full or empty storage state for long sustained periods. For example, they also considered correlation of inflow noise with beginning-of-period storage as a variable whose result was available from optimization. The simulation results compared well with the FP model results for this correlation. Additionally, in the problem studied here the probability of deficit has been equal to one for three months and some few months with nonzero probability of spill, so the bounds have been hit in some months even for this relatively large reservoir, and the FP method accounting explicitly for the probabilities of deficits and spills has performed well in terms of the match between optimization and simulation results for problems where the bounds have been hit frequently.

As a summary, the FP model 1) accounts for stochasticity of independent, Gaussian and non-Gaussian inflows explicitly, 2) it has no dimensionality problem and 3) it can address the nonlinear objective functions no worse than most other optimization methods that use only up to a second order approximation. These advantages are important considering that there is still no explicit stochastic optimization method capable of addressing all aspects of nonlinearity, stochasticity and dimensionality challenges perfectly at such rapid speed as this method. While FP method can be used to solve systems with hundreds of reservoirs (especially for the long term operations), other methods will be impossible to apply without significant approximations. The tradeoffs between approximations in such methods and the simpler linear decision rule used in multireservoirs and certain independent assumptions in FP method are yet to be studied.

5 Conclusions

This paper proposes novel extensions to the FP explicit stochastic optimization method applied to the operation of a water supply reservoir. The main conclusions and contributions are:

1) When the FP approach was introduced by Fletcher and Ponnambalam (1996), Taylor series approximations were used for the derivation of the first and second storage moments and the final optimization model had to include also the moments as decision variables. These typically led to an optimization problem with 36 decision variables, 12 equality constraints, 12 inequality constraints and 24 bound constraints, which has been applied in all applications of the FP method. The new implementation in this paper considerably simplified the original highly-constrained nonlinear optimization problem to a completely unconstrained, vectorized and faster 12-variable (linear decision rule parameters) optimization model that is able to explicitly determine first and second statistical moments of storage, deficit, and spill as well as probabilities of deficit, containment, and spill. Also, it is easy to see that this provides at least a 27 times speedup. In addition to this, the computational efficiency also increases significantly for using unconstrained instead of constrained optimization. The significance of the proposed modifications was investigated through the application of the new procedure to the monthly operation of the Sobradinho reservoir, Brazil.

2) New expressions were proposed for first and second moments of deficit and spill terms. These expressions together with already-derived second moments of storage were then incorporated in the FP model's nonlinear objective function and provided new information that considerably improved the model's ability to estimate the

expected value of the sum of squared deviations between releases and demands. The results obtained by the new FP formulation showed agreement with those obtained by simulating the reservoir operation over a long period using the derived-by-FP release policies.

3) We also conducted detailed analyses to assess the role of simple linear decision rules (LDR) and Gaussian independent inflows assumptions employed in the FP method. The FP method's results revealed that the derived-by-FP policies based on LDR performed quite satisfactorily compared to SDP, TSP, and ISP methods, benefiting from more sophisticated operation policies, even when the derived policies were simulated against non-Gaussian correlated inflows.

Together with the non-requirement for discretization of storage and inflow state variables, these characteristics can be of great advantage when compared to other strategies based on for example SDP, and are especially valuable to the design and operation of multireservoir systems. The application of the newly proposed extensions to the FP method to multireservoir systems under different inflow and demand variability and correlation conditions should be studied in future.

Acknowledgments and Data

The data is available for anyone from: https://doi.org/10.5683/SP2/SBQFWO

The programs in MATLAB ® associated with this study will be available in the same URL after the paper is accepted.

The author SJM thanks Amirkabir University of Technology for partial support during his sabbatical research period at the University of Waterloo. The author KP thanks the Natural Sciences and Engineering Research Council of Canada (NSERC) for their Discovery and CRD grants. He acknowledges also the Ontario Power Generation (OPG) grant on Dam Safety. We also thank Ifeanyi E. Okwuchi for his help in TSP1 model implementation.

Authors declare no conflicts of interests.

A: Derivation of the New Expressions for the Moments of Deficit and Spill

A.1 First Moment of Deficit

Taking expectation of equation (17) gives

$$\begin{aligned}\mathbb{E}(\delta_t) &= \mathbb{E}\left[S_t^{min} - (\bar{I}_t - k_t) - \eta_t\right] \cdot \mathbb{1}_{[-\infty, S_t^{min}]}(\hat{S}_t)) \\ &= \left[S_t^{min} - (\bar{I}_t - k_t)\right] \cdot \mathbb{E}(\mathbb{1}_{[-\infty, S_t^{min}]}(\hat{S}_t)) - \mathbb{E}(\eta_t \cdot \mathbb{1}_{[-\infty, S_t^{min}]}(\hat{S}_t))\end{aligned} \tag{A1}$$

The expected value of the indicator function of a random variable over any region is the probability of that random variable occurring within that same region. Thus, the first expectation in equation (34) represents the probability of deficit $\mathbb{P}_t^{\mathrm{def}}$ (i.e., projected storage below the minimum) and can be calculated as

$$\begin{aligned}\mathbb{E}\left(\mathbb{1}_{[-\infty, S_t^{min}]}(\hat{S}_t)\right) &= \mathbb{P}_t^{\mathrm{def}} = \Pr\left(\hat{S}_t < S_t^{min}\right)\\ &= \Pr\left(\bar{I}_t + \eta_t - k_t < S_t^{min}\right)\\ &= \Pr\left(\eta_t < S_t^{min} - (\bar{I}_t - k_t)\right) = \int_{-\infty}^{S_t^{min}-(\bar{I}_t-k_t)} f_{\eta_t}(\eta_t)\,d\eta_t\end{aligned} \tag{A2}$$

in which Pr( ) denotes probability. The second term on the right-hand side of equation (A1) represents the expectation of a function $g(\eta_t) = \eta_t \,.\, \mathbb{1}_{[S_t^{min}, S_t^{max}]}(\hat{S}_t)$ of the random variable η_t. Given the expectation property $\mathbb{E}(g(X)) = \int_{-\infty}^{+\infty} g(x).f(x)d(x)$ in which *X* is a random variable and *f*(*x*) is its probability density function, then

$$\mathbb{E}\left(\eta_t \cdot \mathbb{1}_{[-\infty, S_t^{min}]}(\hat{S}_t)\right) = \int_{-\infty}^{+\infty}\left[\eta_t \cdot \mathbb{1}_{[-\infty, S_t^{min}]}(\hat{S}_t)\right] f(\eta_t)\,d\eta_t \tag{A3}$$

This integral can be separated into two parts corresponding to intervals $(-\infty, S_t^{min} - (\bar{I}_t - k_t)]$ and $(S_t^{min} - (\bar{I}_t - k_t), +\infty)$ and finally be expressed only for the limits where the indicator function is the unity (first interval) as

$$\mathbb{E}\left(\eta_t \cdot \mathbb{1}_{[-\infty, S_t^{min}]}(\hat{S}_t)\right) = \int_{-\infty}^{S_t^{min}-(\bar{I}_t-k_t)} \eta_t f_{\eta_t}(\eta_t)\,d\eta_t \tag{A4}$$

Thus, equation (A1) for the first moment of deficit finally becomes

$$\mathbb{E}(\delta_t) = \left[S_t^{min} - (\bar{I}_t - k_t)\right] \cdot \int_{-\infty}^{S_t^{min}-(\bar{I}_t-k_t)} f_{\eta_t}(\eta_t)\,d\eta_t - \int_{-\infty}^{S_t^{min}-(\bar{I}_t-k_t)} \eta_t f_{\eta_t}(\eta_t)\,d\eta_t \tag{A5}$$

A.2 Second Moment of Deficit

Taking expectation of equation (19) gives

$$\begin{aligned}\mathbb{E}(\delta_t^2) = \left[S_t^{min} - (\bar{I}_t - k_t)\right]^2 \cdot \mathbb{E}\left(\mathbb{1}_{[-\infty, S_t^{min}]}(\hat{S}_t)\right) - 2\left[S_t^{min} - (\bar{I}_t - k_t)\right]\\ \cdot\, \mathbb{E}\left(\eta_t \cdot \mathbb{1}_{[-\infty, S_t^{min}]}(\hat{S}_t)\right) + \mathbb{E}\left(\eta_t^2 \cdot \mathbb{1}_{[-\infty, S_t^{min}]}(\hat{S}_t)\right)\end{aligned} \tag{A6}$$

Using the same principle applied in equation (A3) for the second and third terms and substituting equation (A2) yields the expression for the second moment of deficit:

$$\mathbb{E}(\delta_t^2) = \left[S_t^{min} - (\bar{I}_t - k_t)\right]^2 \cdot \int_{-\infty}^{S_t^{min}-(\bar{I}_t-k_t)} f_{\eta_t}(\eta_t)d\eta_t - 2\left[S_t^{min} - (\bar{I}_t - k_t)\right] \cdot \int_{-\infty}^{S_t^{min}-(\bar{I}_t-k_t)} \eta_t f_{\eta_t}(\eta_t)d\eta_t + \int_{-\infty}^{S_t^{min}-(\bar{I}_t-k_t)} \eta_t^2 f_{\eta_t}(\eta_t)d\eta_t \tag{A7}$$

A.3 First Moment of Spill

Taking expectation of equation (16) gives

$$\begin{aligned}\mathbb{E}(Sp_t) &= \mathbb{E}\left(\left[(\bar{I}_t - k_t) - S_t^{max} + \eta_t\right] \cdot \mathbb{1}_{[S_t^{max}.+\infty]}(\hat{S}_t)\right) \\ &= \left[(\bar{I}_t - k_t) - S_t^{max}\right] \cdot \mathbb{E}\left(\mathbb{1}_{[S_t^{max}.+\infty]}(\hat{S}_t)\right) + \mathbb{E}\left(\eta_t \cdot \mathbb{1}_{[S_t^{max}.+\infty]}(\hat{S}_t)\right)\end{aligned} \tag{A8}$$

The first expectation in equation (A8) represents the probability of spill $\mathbb{p}_t^{\mathrm{sp}}$ (i.e., projected storage above maximum) and can be calculated as

$$\begin{aligned}\mathbb{E}\left(\mathbb{1}_{[S_t^{max}.+\infty]}(\hat{S}_t)\right) &= \mathbb{p}_t^{\mathrm{sp}} \\ &= \Pr\left(\hat{S}_t > S_t^{max}\right) = \\ &= \Pr\left(\bar{I}_t + \eta_t - k_t > S_t^{max}\right) \\ &= \Pr\left(\eta_t > S_t^{max} - (\bar{I}_t - k_t)\right) = \int_{S_t^{max}-(\bar{I}_t-k_t)}^{+\infty} f_{\eta_t}(\eta_t)d\eta_t\end{aligned} \tag{A9}$$

Using the same principle applied in equation (A3) for the second expectation in (A8) and substituting equation (A9) yields the expression for the first moment of spill:

$$\mathbb{E}(Sp_t) = \left[(\bar{I}_t - k_t) - S_t^{max}\right] \cdot \int_{S_t^{max}-(\bar{I}_t-k_t)}^{+\infty} f_{\eta_t}(\eta_t)d\eta_t + \int_{S_t^{max}-(\bar{I}_t-k_t)}^{+\infty} \eta_t f_{\eta_t}(\eta_t)d\eta_t \tag{A10}$$

A.4 Second Moment of Spill

Taking expectation of equation (18) gives

$$\mathbb{E}(Sp_t^2) = [(\bar{I}_t - k_t) - S_t^{max}]^2 \cdot \mathbb{E}\left(\mathbb{1}_{[S_t^{max}.+\infty]}(\widehat{S}_t)\right) + 2\,[(\bar{I}_t - k_t) - S_t^{max}] \cdot \mathbb{E}\left(\eta_t \cdot \mathbb{1}_{[S_t^{max}.+\infty]}(\widehat{S}_t)\right)$$
$$+\ \mathbb{E}\left(\eta_t^2 \cdot \mathbb{1}_{[S_t^{max}.+\infty]}(\widehat{S}_t)\right) \tag{A11}$$

Using the same principle applied in equation (A3) for the second and third terms and substituting equation (A9) yields the expression for the second moment of spill:

$$\mathbb{E}(Sp_t^2) = [(\bar{I}_t - k_t) - S_t^{max}]^2 \cdot \int_{S_t^{max}-(\bar{I}_t-k_t)}^{+\infty} f_{\eta_t}(\eta_t)d\eta_t + 2\,[(\bar{I}_t - k_t) - S_t^{max}]$$
$$\cdot \int_{S_t^{max}-(\bar{I}_t-k_t)}^{+\infty} \eta_t f_{\eta_t}(\eta_t)d\eta_t + \int_{S_t^{max}-(\bar{I}_t-k_t)}^{+\infty} \eta_t^2 f_{\eta_t}(\eta_t)d\eta_t \tag{A12}$$

Similar to equations (A2) and (A9), the probability of containment $\mathbb{P}_t^{\mathrm{con}}$ can be expressed as

$$\mathbb{E}\left(\mathbb{1}_{[S_t^{min}.S_t^{max}]}(\hat{S}_t)\right) = \mathbb{P}_t^{\mathrm{con}} = \Pr\left(S_t^{min} \le \hat{S}_t \le S_t^{max}\right)$$
$$= \Pr\left(S_t^{min} \le \bar{I}_t + \eta_t - k_t \le S_t^{max}\right)$$
$$= \Pr\left(S_t^{min} - (\bar{I}_t - k_t) \le \eta_t \le S_t^{max} - (\bar{I}_t - k_t)\right) = \int_{S_t^{min}-(\bar{I}_t-k_t)}^{S_t^{max}-(\bar{I}_t-k_t)} f_{\eta_t}(\eta_t)d\eta_t \tag{A13}$$

B: Expressions Assuming Gaussian Inflows

The probability density function of a zero-mean random variable η following a Gaussian distribution of the form $N(0, \mathrm{Var}(\eta_t))$ is given by

$$f_{\eta_t}(\eta) = \frac{1}{\sqrt{2\pi\mathrm{Var}(\eta_t)}}\exp\left[\frac{\eta^2}{2\mathrm{Var}(\eta_t)}\right] \tag{B1}$$

Its correspondent cumulative distribution function (CDF) is

$$F_{\eta_t}(\eta) = \Pr(\eta_t \le \eta) = \int_{-\infty}^{\eta} f(t)dt = \frac{1}{2}\left[1 + \mathrm{erf}\left(\frac{\eta}{\sqrt{2\mathrm{Var}(\eta_t)}}\right)\right] \tag{B2}$$

in which $\mathrm{erf}(.)$ is the *error function* formulated as

$$\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2}\, dt \tag{B3}$$

With these, the solutions for the three types of integrals appearing in the expressions of moments of storage (equations (13) and (14)), deficit (equations (19) and (20)) and spill (equations (21) and (22)), as well as in the expressions for probabilities (equations (A3), (A9) and (A13)) are given as below, assuming generic lower L and upper U limits of integration:

$$\int_L^U f_{\eta_t}(\eta) d\eta = F_{\eta_t}(U) - F_{\eta_t}(L) = \frac{1}{2}\left[erf\left(\frac{U}{\sqrt{2Var(\eta_t)}}\right) - erf\left(\frac{L}{\sqrt{2Var(\eta_t)}}\right)\right] \tag{B4}$$

$$\begin{aligned}\int_L^U \eta f_{\eta_t}(\eta) d\eta_t &= \frac{1}{\sqrt{2\pi Var(\eta_t)}} \int_L^U \eta\, exp\left[\frac{\eta^2}{2Var(\eta_t)}\right] d\eta \\ &= -\sqrt{\frac{Var(\eta_t)}{2\pi}}\left[exp\left(\frac{-U^2}{2Var(\eta_t)}\right) - exp\left(\frac{-L^2}{2Var(\eta_t)}\right)\right]\end{aligned} \tag{B5}$$

$$\begin{aligned}\int_L^U \eta^2 f_{\eta_t}(\eta) d\eta_t &= \frac{1}{\sqrt{2\pi Var(\eta_t)}} \int_L^U \eta^2\, exp\left[\frac{\eta^2}{2Var(\eta_t)}\right] d\eta = \\ &= -\sqrt{\frac{Var(\eta_t)}{2\pi}}\left[U\, exp\left(\frac{-U^2}{2Var(\eta_t)}\right) - L\, exp\left(\frac{-L^2}{2Var(\eta_t)}\right)\right] \\ &+ \frac{Var(\eta_t)}{2}\left[erf\left(\frac{U}{\sqrt{2Var(\eta_t)}}\right) - erf\left(\frac{L}{\sqrt{2Var(\eta_t)}}\right)\right]\end{aligned} \tag{B6}$$

The limits of integration L and U can be changed accordingly in order to derive the final expressions. The expressions for the storage moments were already shown in equations (13) and (14). The final expressions for probabilities and moments of deficit and spill are displayed below, using LB and UB defined in section 2.1.

- Probability of containment:

$$\mathbb{P}_t^{con} = \frac{1}{2}[erf(UB) - erf(LB)] \tag{B7}$$

- Probability of deficit:

$$\mathbb{P}_t^{\mathrm{def}} = \frac{1}{2}[1 + \mathrm{erf}(\mathrm{LB})] \tag{B8}$$

- Probability of spill:

$$\mathbb{P}_t^{\mathrm{sp}} = \frac{1}{2}[1 - \mathrm{erf}(\mathrm{UB})] \tag{B9}$$

- First moment of deficit:

$$\mathbb{E}(\delta_t) = \left[S_t^{min} - (\bar{I}_t - k_t)\right] \mathbb{P}_t^{\mathrm{def}} + \sqrt{\frac{\mathrm{Var}(\eta_t)}{2\pi}} \exp(-\mathrm{LB}^2) \tag{B10}$$

- Second moment of deficit:

$$\mathbb{E}(\delta_t^2) = \left[S_t^{min} - (\bar{I}_t - k_t)\right]^2 \mathbb{P}_t^{\mathrm{def}} - 2\left[S_t^{min} - (\bar{I}_t - k_t)\right]\left[-\sqrt{\frac{\mathrm{Var}(\eta_t)}{2\pi}} \exp(-\mathrm{LB}^2)\right] - \sqrt{\frac{\mathrm{Var}(\eta_t)}{2\pi}}\left[S_t^{min} - (\bar{I}_t - k_t)\right] \exp(-\mathrm{LB}^2) + \frac{\mathrm{Var}(\eta_t)}{2}[1 + \mathrm{erf}(LB)] \tag{B11}$$

- First moment of spill:

$$\mathbb{E}(Sp_t) = [(\bar{I}_t - k_t) - S_t^{max}]\, \mathbb{P}_t^{\mathrm{sp}} + \sqrt{\frac{\mathrm{Var}(\eta_t)}{2\pi}} \exp(-\mathrm{UB}^2) \tag{B12}$$

- Second moment of spill:

$$\mathbb{E}(Sp_t^2) = [(\bar{I}_t - k_t) - S_t^{max}]^2\ \mathbb{P}_t^{\mathrm{sp}} + 2\,[(\bar{I}_t - k_t) - S_t^{max}]\left[\sqrt{\frac{\mathrm{Var}(\eta_t)}{2\pi}} \exp(-\mathrm{UB}^2)\right] + \sqrt{\frac{\mathrm{Var}(\eta_t)}{2\pi}}[S_t^{max} - (\bar{I}_t - k_t)] \exp(-\mathrm{UB}^2) + \frac{\mathrm{Var}(\eta_t)}{2}[1 - \mathrm{erf}(\mathrm{UB})] \tag{B13}$$

C: Vectorization

Let $\boldsymbol{k} = \{k_1, \cdots, k_{12}\}^{\intercal}$ be the vector formed by the twelve unknown LDR parameters. Similarly, we can define vectors for minimum and maximum storages as well as for monthly mean inflow and inflow variances, respectively:

$$\mathrm{S}^{min} = \left\{S_1^{min}. \cdots . S_{12}^{min}\right\} \tag{C1}$$

$$\mathrm{S}^{max} = \{S_1^{max}. \cdots . S_{12}^{max}\} \tag{C2}$$

$$\bar{\mathrm{I}} = \{\bar{I}_1. \cdots . \bar{I}_{12}\} \tag{C3}$$

$$\mathrm{Var}\eta = \{\mathrm{Var}(\eta_1). \cdots . \mathrm{Var}(\eta_{12})\} \tag{C4}$$

Corresponding vectorized versions of LB and UB may be written as

$$LB = \frac{S^{min} - (\bar{I} - k)}{\sqrt{2Var\eta}} \tag{C5}$$

$$UB = \frac{S^{max} - (\bar{I} - k)}{\sqrt{2Var\eta}} \tag{C6}$$

which, in turn, provide a means to write the vectorized expression for the first moment of storage (equation (13)):

$$\begin{aligned} \mathbb{E}1 = \frac{\bar{I} - k}{2}[erf(UB) - erf(LB)] - \sqrt{\frac{Var\eta}{2\pi}}[exp(-UB^2) - exp(-LB^2)] \\ + \frac{S^{min}}{2}[1 + erf(LB)] + \frac{S^{max}}{2}[1 - erf(UB)] \end{aligned} \tag{C7}$$

where $\mathrm{E1} = \{\mathbb{E}(S_{12}).\mathbb{E}(S_1). \cdots . \mathbb{E}(S_{11})\}^{\intercal}$ and all operations are conducted element-wise. Alternative vector expressions can be easily derived for second storage moment (E**2**) and moments of deficit ($\mathbb{E}\mathbf{1}_{\delta}$, $\mathrm{E}\mathbf{2}_{\delta}$) and spill ($\mathrm{E}\mathbf{1}_{Sp}$, $\mathrm{E}\mathbf{2}_{Sp}$). Defining two other vectors

$$\mathbb{E}1_0 = \{\mathbb{E}(S_{12}). \mathbb{E}(S_1). \cdots . \mathbb{E}(S_{11})\}^{\intercal} \tag{C8}$$

$$\mathbb{E}2_0 = \{\mathbb{E}(S_{12}^2). \mathbb{E}(S_1^2). \cdots . \mathbb{E}(S_{11}^2)\}^{\intercal} \tag{C9}$$

the vectorized version of the objective function (28) may be written as

$$\begin{aligned} Z = sum\big[\mathbb{E}2_0 + 2 \cdot (k - D) \cdot \mathbb{E}1_0 + (k - D)^2 + \mathbb{E}2_{\delta} - 2 \cdot \mathbb{E}1_0 \cdot \mathbb{E}1_{\delta} - 2 \cdot (k - D) \cdot \mathbb{E}1_{\delta} \\ + \mathbb{E}2_{Sp} + 2 \cdot \mathbb{E}1_0 \cdot \mathbb{E}1_{Sp} + 2 \cdot (k - D) \cdot \mathbb{E}1_{Sp}\big] \end{aligned} \tag{C10}$$

for demand vector $\mathrm{D} = \{D_1, \cdots, D_{12}\}^{\top}$ and operator sum [.] representing the sum of array elements. All these vectorized expressions are straightforwardly implemented in matrix programming environments such as MATLAB or Octave.

D: Evaluating the Role of the Second Moment of Storage

Omitting the deficit and surplus terms at this stage, by comparing the results of the FP method in which Z_1 (equation (25)) and Z_2 (equation (26)) are used as the objective function and simulating their policies, we can evaluate how important the role of incorporating the second moments of storage is.

For convenience, the implementations using Z_1 and Z_2 were named Model 1 and Model 2, respectively. Figures D1 (Model 1) and D2 (Model 2) show comparison of statistics obtained by optimization and simulation for both models. Note that in both optimization and simulation modes, the values of variables (inflow, storage, release, spill, deficit) in units of volume were scaled by the volume equivalent to the mean annual flow.

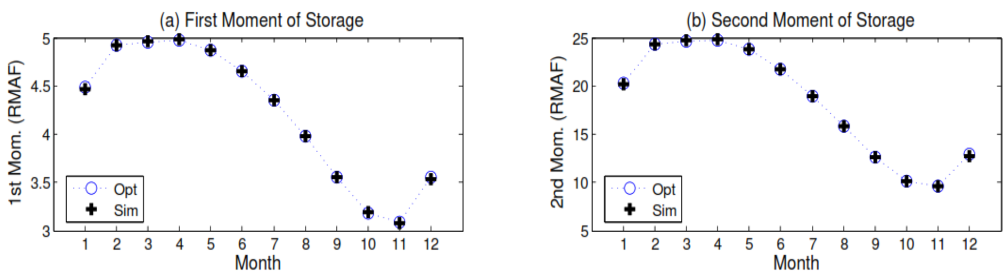

Figure D1. Comparison of (a) first and (b) second moments of storage found by found by the FP method for Model 1

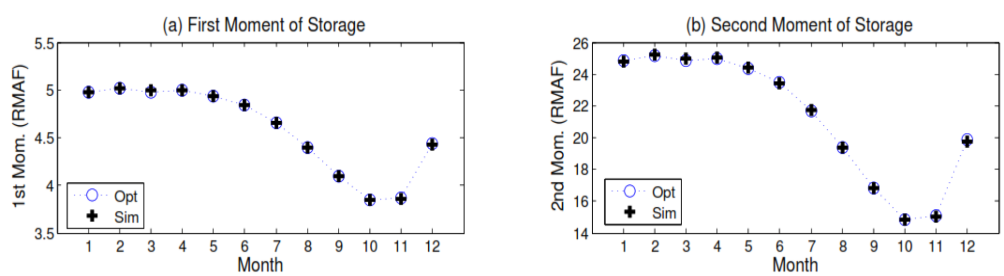

Figure D2. Comparison of (a) first and (b) second moments of storage found by the FP method for Model 2

As mentioned before, the simulation of the reservoir operation employed the LDR-guided policies derived from optimization (optimal k_t values) for 1,000 years from which the simulated first and second (sample) moments of storage were calculated for every month of the year. From the figures, acceptable match between simulated and optimization-based first and second moments are seen. However, to be more precise, the sums of squared errors between optimized and simulated first (SSE_1) and second moments (SSE_2) for both models were used for comparison. SSE_1 values were 0.00094 (0.00065) whereas SSE_2 was 0.060 (0.049) for Model 1 (Model 2). Therefore, the match between optimization and simulation-based raw moments of storage for Model 2 (with exact objective function) was better than that for Model 1 (with zeroth-order Taylor approximation of the objective function). It is also important to evaluate the performance of these models in terms of the most important optimality criterion, i.e., the objective function value. For Model 1, while the objective function value of the optimization model was almost zero ($Z_1^{opt} = 2.75 \times 10^{-7}$), the simulated objective function value was quite different ($Z_1^{sim} = 0.79$). However, for Model 2, not only the simulated objective function ($Z_2^{sim} = 0.61$) was about 23% better than that of Model 1, it also better matched the optimization objective function ($Z_2^{opt} = 0.70$).

E: Evaluating the Role of Deficits and Spills

Looking carefully at the most important set of equations (3), representing the dynamics of a nonlinear bounded system, one can notice that U_t is not the total release from the reservoir, but part of the release that makes the end-of-period storage volume contained. In all applications of the FP model so far only U_t has been used in the objective function meaning that the role of deficit and spill terms have not been included in the objective function evaluation of any candidate solution. However, we show here that consideration of deficit and spill terms is quite important when a nonlinear objective function like the one used in this study is being considered. The importance of the issue is because penalizing the objective function due to deficit or spill occurrences is all what the model's objective is about. To account for these terms, we derived new expressions for the first and second moments of deficit and spill and used them in the expected value of the objective function. We first analyze the role of incorporating the deficit term. Typically, spillway capacity is very large, so in cases where the downstream river's safe discharge is also large enough, we may not care about spill volumes to be penalized in the objective function.

Role of deficits

To evaluate how important the incorporation of the deficit term in the objective function is, two other FP formulations were compared, one that uses only U_t in the objective function (with consideration of the second moment of storage) (Model 2B), and another using the deficit term and new expressions for its first ($\mathbb{E}(\delta_t)$) and second $\mathbb{E}(\delta_t^2)$ moments added to the optimization model formulation (Model 3). However, in both cases the release made in the simulation model is the actual total release including U_t and δ_t. Therefore, the difference between simulated objective functions in Model 2B and Model 3 will be due to the impact of how the deficit term has been considered in the optimization model's formulation. Note that Model 2B is the same as Model 2 introduced in the previous section in optimization mode, and their difference is just in simulation mode. The deficit term is included in simulated releases in Model 2B whereas they are not in Model 2. For Model 3, the objective function is

$$Z_3 = \sum_{t=1}^{12} [\mathbb{E}(S_{t-1}^2) + 2(k_t - D_t).\mathbb{E}(S_{t-1}) + (k_t - D_t)^2 + \mathbb{E}(\delta_t^2) - 2\mathbb{E}(S_{t-1}) \cdot \mathbb{E}(\delta_t) - 2(k_t - D_t).\mathbb{E}(\delta_t)] \quad \text{(E1)}$$

The objective function values in optimization (simulation) for Model 2B and Model 3 were 0.7 (0.61) and 3.69 (3.52), respectively. We also tested the case when the target demand (80% of the mean annual flow) was doubled because the larger the demand, the more important the impact of incorporating deficit is expected to be. The objective function values in optimization (simulation) were 1.31 (9.63) and 7.15 (6.81) for Models 2B and 3, respectively. We observe that for the newly derived objective function expressions (Model 3), the objective function values in simulation and optimization matched better. However, there was a big gap between these values with the old expressions (Model 2B) where the optimization model always underestimated the real objective function value (simulated value). Another interesting point to know is what we would lose if we modeled the second moment of storage accurately, but still did not account for deficit (Model 2). The Model 3's objective function value (both simulation and optimization) as the correct value was about 3.62 (estimated by averaging optimization and simulation values), whereas it was underestimated as 0.70 by Model 2. Therefore, 3.62 - 0.70 = 2.92 is due to not accounting for the role of deficits in the optimization model formulation. On the other hand, the difference between the objective functions values of Model 3 and Model 1 is $3.62 - 2.75 \times 10^{-7} = 3.62$. Therefore, from the two sources of error associated with Model 1, (considering neither the second moments of storages nor first and second moments of deficits), 0.70⁄3.62 = 19% is because of not accounting for the second moments of storages and 2.92⁄3.62 = 81% is due to not modeling deficits appropriately.

Role of spills

A similar analysis was conducted for evaluating the role of incorporating spills by running two other types of models, one where the spill term is not accounted for in the optimization model formulation (Model 3B) versus another in which such term is included using the newly derived expressions for the first ($\mathbb{E}(Sp_t)$) and second ($\mathbb{E}(Sp_t^2)$) moments of spill (Model 4). Note that for both cases the surplus term is included in the simulation model while determining reservoir releases and evaluating the objective function value. Additionally, to be fair and to analyze only the effect of spills without having the results being affected by the influence of deficit, the deficit term is considered in both optimization and simulation for both Models 3B and 4. Model 3B is the same as Model 3 in optimization mode, and their difference is only in simulation mode. For Model 3B, spills are considered while simulating FP's optimal policies whereas they are not for Model 3. For Model 4, the objective function is equation (26), including all moments of storage, deficit, and spill in both optimization and simulation. To have the role of spills more sensed, experiments were carried out for inflow mean values equal to 2 times of the normal inflows. The objective function values in optimization (simulation) were 8.91 (26.04) and 21.30 (21.52) for Models 3B and 4, respectively. We see that Model 4 has improved the agreement between optimization and simulation significantly as the difference between optimization and simulation objective function values is around 192% for Model 3B, whereas it is only 1% for Model 4. See Figure E1 for a comparison of simulation-optimization results for 1000 years of simulated Gaussian inflows.

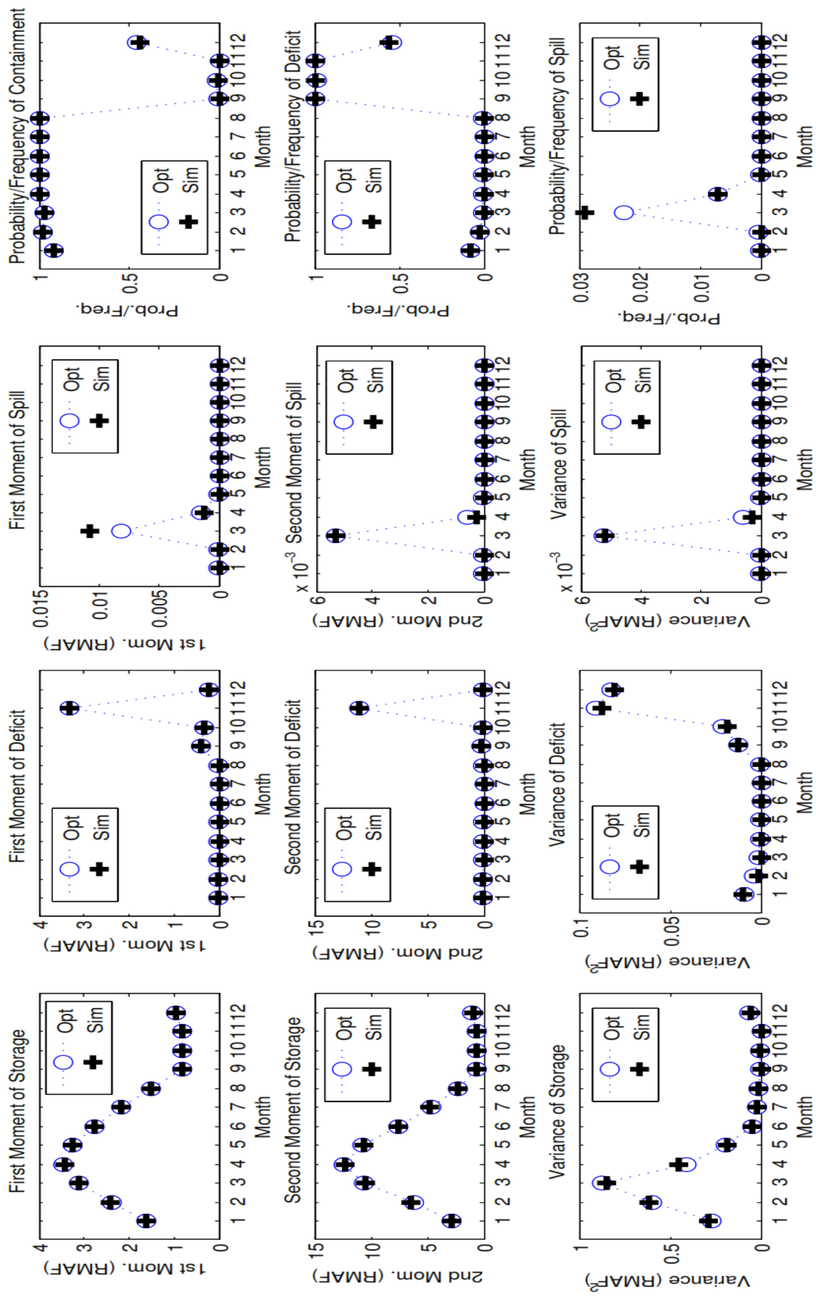

Figure E1. Results from the proposed formulation/implementation of the FP model applied to the Sobradinho reservoir with Gaussian inflows.

References

Adams, B. J., & Ponnambalam, K. (1994). An algorithm for determining closed-loop operations policies for multireservoir systems. In Hipel, K. W. & Fang, L., editors, *Stochastic and Statistical Methods in Hydrology and Environmental Engineering: Effective Environmental Management for Sustainable Development.* 389–396. https://link.springer.com/book/10.1007/978-94-017-3081-5

Agência Nacional de Águas (2015). *Conjuntura dos recursos hídricos no Brasil: regiões hidrográficas brasileiras – Edição Especial.* Brazilian National Water Agency, Brasília.

Alizadeh, H., Mousavi, S. J., & Ponnambalam, K. (2018). Copula-based chance-constrained hydro-economic optimization model for optimal design of reservoir-irrigation district systems under multiple interdependent sources of uncertainty. *Water Resources Research*, 54(8), 5763–5784. https://doi.org/10.1029/2017WR022105

Mousavi, S. J., Alizadeh, H., & Ponnambalam, K. (2014). Storage-yield analysis of surface water reservoirs: the role of reliability constraints and operating policies. *Stochastic Stochastic Environmental Research and Risk Assessment*, 28(8), 2051-2061. https://doi.org/10.1007/s00477-014-0910-7

Archibald, T. W., & Marshall, S. E. (2018). Review of mathematical programming applications in water resource management under uncertainty. *Environmental Modeling & Assessment*. 23, 753–777. https://doi.org/10.1007/s10666-018-9628-0

Cai, X. and M. Rosegrant (2004). Irrigation technology choices under hydrologic uncertainty- A case study from the Maipo River Basin, Chile, *Water Resources Research*, 40, W04103, doi:10.1029/2003WR002810.

Celeste, A. B., & Billib, M. (2009). Evaluation of stochastic reservoir operation optimization models. *Advances in Water Resources*, 32, 1429–1443. https://doi.org/10.1016/j.advwatres.2009.06.008

Fayaed, S. S., El-Shafie, A., & Jaafar, O. (2013). Reservoir-system simulation and optimization techniques. *Stochastic Environmental Research and Risk Assessment*, 27(7):1751–1772. https://doi.org/10.1007/s00477-013-0711-4

Fletcher, S. G., & Ponnambalam, K. (1996). A new formulation for the stochastic control of systems with bounded state variables: An application to a single reservoir system. *Stochastic Hydrology and Hydraulics*, 10(3),167–186. https://doi.org/10.1007/BF01581462

Fletcher, S. G., & Ponnambalam, K. (1998). A constrained state formulation for the stochastic control of multireservoir systems, *Water Res. Res*., 34(2), 257-270.

Fletcher, S. G., & Ponnambalam, K. (2008). Stochastic control of reservoir systems using indicator functions: New enhancements. *Water Resources Research*, 44(12). https://doi.org/10.1029/2006WR005269

Ganji, A., & Jowkarshorijeh, L. (2012). Advance first order second moment (AFOSM) method for single reservoir operation reliability analysis: a case study. *Stochastic Environmental Research and Risk Assessment*, 26(1), 33–42. https://doi.org/10.1007/s00477-011-0517-1

Jones, M. (2009). Kumaraswamy's distribution: A beta-type distribution with some tractability advantages. *Statistical Methodology*, 6(1), 70–81. https://doi.org/10.1016/j.stamet.2008.04.001

Joodavi, A., Zare, M., Ziaei, A. N., & Ferré, T. P. (2017). Groundwater management under uncertainty using a stochastic multi-cell model. *Journal of Hydrology*, 551, 265–277. https://doi.org/10.1016/j.jhydrol.2017.06.003

Labadie, J. W. (2004). Optimal operation of multireservoir systems: State-of-the-art review. *Journal of Water Resources Planning and Management*, *130*(2), 93–111. https://doi.org/10.1061/(ASCE)0733-9496(2004)130:2(93)

Lilliefors, H. W. (1967). On the Kolmogorov-Smirnov test for normality with mean and variance unknown. *Journal of the American Statistical Association*, 62(318), 399–402. doi: 10.1080/01621459.1967.10482916

Lopez, J. A., Ponnambalam, K., & Quintana, V. H. (2007). Generation and transmission expansion under risk using stochastic programming. *IEEE Transactions on Power Systems*, 22(3), 1369–1378. doi: 10.1109/TPWRS.2007.901741.

Loucks, D.P., Van Beek, E. (2017). *Water resource systems planning and management: An introduction to methods, models, and applications*. Springer. ISBN 978-3-319-44234-1

Mahootchi, M. , K. Ponnambalam, H. Tizhoosh, 2010. Comparison of risk-based optimization models for reservoir management, *Canadian J. of Civil Engineering*, 37, 112-124.

Mahootchi, M., Ponnambalam, K., & Tizhoosh, H. (2010). Operations optimization of multireservoir systems using storage moments equations. *Advances in Water Resources*, *33*(9), 1150–1163. https://doi.org/10.1016/j.advwatres.2010.07.004

Mahootchi, M., Ahmadi, T., & Ponnambalam, K. (2012). Introducing a new formulation for the warehouse inventory management systems: with two stochastic demand patterns. *International Journal of Industrial Engineering & Production Research*, *23*(4), 277–284.

Mahootchi, M., & Ponnambalam, K. (2013). Parambikulam-aliyar project operations optimization with reliability constraints. *Journal of Water Resources Planning and Management, 139*(4), 364–375. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000261

Mendes, L. A., de Barros, M. T. L., Zambon, R. C., & Yeh, W. W.-G. (2015). Trade-off analysis among multiple water uses in a hydropower system: Case of São Francisco River Basin, Brazil. *Journal of Water Resources Planning and Management*, *141*(10), 04015014. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000527

Mousavi, S. J., Karamouz, M., & Menhadj, M.B. (2004). Fuzzy-State Stochastic Dynamic Programming for Reservoir Operation. *Journal of Water Resources Planning and Management, 130*(6), 460-470. https://doi.org/10.1061/(ASCE)0733-9496(2004)130:6(460)

Mousavi, S. J., Ponnambalam, K., & Karray, F. (2007). Inferring operating rules for reservoir operations using fuzzy regression and ANFIS. *Fuzzy Sets and Systems*, *158*(10), 1064-1082. https://doi.org/10.1016/j.fss.2006.10.024

Nagy, I. V., Asante-Duah, K., & Zsuffa, I. (2002). *Hydrological dimensioning and operation of reservoirs: practical design concepts and principles*. Kluwer Academic Publishers, Dordrecht. ISBN 978-94-015-9894-1

Oliveira Dantas, C. E. (2005). Análise dos efeitos dinâmicos em reservatórios de grande extensão. Estudo de caso: reservatório de Sobradinho. Master’s thesis, Universidade Federal de Pernambuco, Recife, Brazil.

Pan, L., Housh, M., Liu, P., Cai, X., & Chen, X. (2015). Robust stochastic optimization for reservoir operation. *Water Resources Research*, 51(1), 409–429. https://doi.org/10.1002/2014WR015380

Ponnambalam, K. (1980). A generalized probability density function for double-bounded random processes. *Journal of Hydrology*, *46*(1), 79–88. https://doi.org/10.1016/0022-1694(80)90036-0

Ponnambalam, K., & Adams, B. J. (1996). Stochastic optimization of multireservoir systems using a heuristic algorithm: case study from India. *Water Resources Research*, 32(3), 733-741. https://doi.org/10.1029/95WR03484

Ponnambalam, K., & Adams, B. J. (1987). Experiences with integrated irrigation system optimization analysis. *IAHS Publication*, 169, 229–245.

Ponnambalam, K., Saad, Y. E., Mahootchi, M., & Heemink, A. W. (2010). Comparison of methods for battery capacity design in renewable energy systems for constant demand and uncertain supply. In *7th International Conference on the European Energy Market*, 1–5. doi: 10.1109/EEM.2010.5558745.

ReVelle, C., Joeres, E., & Kirby, W. (1969). The linear decision rule in reservoir management and design: 1, development of the stochastic model. *Water Resources Research*, 5(4), 767–777. https://doi.org/10.1029/WR005i004p00767

Saadat, M., & Asghari, K. (2017). Reliability improved stochastic dynamic programming for reservoir operation optimization. *Water Resources Management*, *31*(6), 1795–1807. https://doi.org/10.1007/s11269-017-1612-y

Séguin, S., Audet, S, & Côté, P (2017). Scenario-Tree Modeling for Stochastic Short-Term Hydropower Operations Planning, *J. Water Resour. Plann. Manage*., 2017, 143(12).

Svanidze, G. (1980). *Mathematical modeling of hydrologic series (for hydroelectric and water resources computations)*. Water Resources Publications, Fort Collins, Colorado. ISBN-10: 0918334322

Turgeon, A. (1981). A decomposition method for the long-term scheduling of reservoirs in series. *Water Resources Research*, 17(6), 1565–1570. https://doi.org/10.1029/WR017i006p01565

Vedula, S., & Mujumdar, P. P. (2005). Water Resources Systems: Modelling Techniques and Analysis. Tata McGraw-Hill Education (made freely available by authors since 2016).

Zhang, J. L, & Ponnambalam, K. (2006). Hydro energy management optimization in a deregulated electricity market, *Optimization and Engineering*, 7(1), 47-61.