Markov Decision Processes: Models, Methods, Directions and Open Problems

Bernard F. Lamond & Abdeslem Boukhtouta
Département Opérations et systèmes de décision
Faculté des sciences de l'administration
Université Laval
Québec (Qc)
Canada G1K 7P4

December 15, 1999

Chapter 1

Water reservoir applications of Markov decision processes

Abstract

Decision problems in water resources management are usually stochastic, dynamic and multidimensional. Water inflows into reservoirs exhibit complex behaviors, with high variability and often serial and spatial correlations, but the statistical analysis of historical time series suggests the natural inflows can be modeled adequately using a Markovian stochastic process. Moreover, the transition equations of mass conservation for the reservoir storages are akin to those found in inventory theory. Therefore, MDP models have been used since the late forties for the planning and operation of reservoir systems. However, the "curse of dimensionality" has been a major obstacle to the numerical solution of MDP models for systems with several reservoirs. Also, the use of optimization models for the operation of multipurpose reservoir systems does not seem widespread. Instead, dam operators often rely on operating rules obtained by simulation models. We feel many research opportunities exist both in the enhancement of computational methods and in the modeling of reservoir applications.

1.1 Introduction

Dams and reservoirs have long been used for storing surplus water during rainy seasons to provide irrigation and drinking water during dry periods. They prevent flooding during periods of thaw or unusually high rainfall. They also serve to regulate flow and depth of water in lakes and rivers for navigational purposes, and to move ships up and down locks as in the Panama canal and the St-Lawrence seaway. Throughout the twentieth century, hydroelectric production has become a major economic benefit of dams, reservoirs and water resources.

The sequential nature of the reservoir management decisions, together with the inherent randomness of natural water inflows, explains the frequent modeling of reservoir management problems as Markov decision processes (MDP), and their optimization by stochastic dynamic programming (SDP). The first discussion of reservoir management in this framework is usually credited to Pierre Massé masse:a [28] in 1946. Optimization results for the hydroelectric production of a single reservoir were published a decade later, with the numerical computation of an optimal policy little:a [24] and the analytic structure of optimal policies for hydrothermal systems gessford:karlin [12]. These results paralleled similar developments that occured in inventory theory at the same epoch. There is an extensive literature on models and methods for reservoir optimization. Surveys can be found in lamond:boukhtouta,yak:a,yeh:a [20, 52, 53].

Nonetheless, large reservoir systems have been in operation for decades before optimization models were developed. Reservoir operators have thus relied on rule curves and other agreed upon operating rules, as well as their own judgement and experience in making reservoir release decisions loucks:sigvaldason [26]. While optimization models are now often used in practice for planning purposes, their use in real-time multiple-reservoir operation is not so widespread. According to oliveira:loucks [31],

> "The need for comprehensive negotiations and subsequent agreements on how to operate a reservoir system seems to be a main reason why most reservoir systems are still managed based on fixed predefined rules. [...] Optimization models can help define these predefined rules, rules that satisfy various constraints on system operation while minimizing future spills or maximizing energy production or minimizing expected future undesired deviations from various water release, storage volume and/or energy production targets."

The "optimization models" referred to in the above citation are usually based on linear (LP) or nonlinear programming (NLP), with the random variables of future inflows replaced by their most recent forecasts. These (deterministic) models must be solved every period with updated forecasts and their solutions provide an *open loop control*. By contrast, an optimal policy of an MDP gives a *closed loop control*, or *feedback solution*, which is more in the form of traditional operating rules.

On the other hand, for reservoir systems whose main purpose is hydroelectric generation, the use of solutions from optimization models is widespread. In [44, 46], for instance, MDP models are presented for the long term planning of the aggregated system, to obtain optimal policies for monthly release and storage targets. Then a hierarchy of deterministic models [13] are used for medium term (NLP) and short term scheduling (LP). A comparison of optimal MDP solutions with traditional rule curve solutions was made in [44] for the Brazilian system, where optimal MDP solutions were shown to have the same reliability than rule curve solutions, but with significantly increased profits.

Stochastic optimization models of hydroelectric production are usually needed when the planning horizon has a length of one or several years, with a time step of one month or longer. The long term scheduling of hydroelectric production is mainly concerned with the larger (annual or multiannual) reservoirs managed by a utility. Typical problems consider twenty or more such reservoirs and are therefore multidimensional. In general, an optimal decision rule for a given period would thus consist of twenty functions of at least twenty variables, each function giving a release target for one reservoir depending on the stored volumes at every reservoir in the system. Such functions are obviously impossible to tabulate numerically (curse of dimensionality). Hence research in this area has attempted to develop (1) aggregation-disaggregation methods, (2) numerical approximation and optimization methods, and (3) analytical solutions [20].

Research also addresses important modeling issues in statistical hydrology (stochastic processes of natural inflows with adequate representation of serial and spatial correlations) and energy economics (such as evaluating marginal production costs in the context of deregulated markets).

The sequel is organized as follows. A brief review of the basic reservoir management concepts and traditional operating policies, are given in §1.2. A survey of several models describing the stochastic processes of natural inflows is presented in §1.3 and a dynamic programming optimization model for a multireservoir hydroelectric system is given in §1.4. Different applications of the MDP models are presented in §1.5. A survey of recent research on MDP solution methods is presented in §1.6. Finally, §1.7 contains directions and open problems, in water reservoir applications of MDP, and some concluding remarks.

1.2 Reservoir management concepts

sec:rmc

More than 5000 years ago the Egyptians measured fluctuations of the Nile river and built dams and other hydraulic structures to divert water to agricultural fields. Since then, practical water knowledge has proliferated among dam operators, farmers and other users. But the concept of a water cycle (hydrological cycle) became firmly established in the scientific literature only in the seventeenth century shaw [37]. The *hydrological cycle* is the continuous circulation of water from the sea to the atmosphere, then to the land and back again to the sea. The water exchanges involved at the various stages of the cycle are evaporation, water-vapor transport, condensation, precipitation, and runoff. *Runoff* from land surface is the residual water of the hydrological cycle, which has not been evaporated by plants and has not infiltrated the ground surface, so it is available for use. The collection of land whose surface waters drain into a river valley forms the *hydrographic basin* of that river.

A *dam* is a barrier built across a watercourse for impounding water. By erecting dams, humans can obstruct and control the flow of water in a basin. A *reservoir* is a (possibly artificial) lake, usually the result of a dam, where water is collected and stored in quantity for use. Reservoirs must occupy the best available sites in the hydrographic basin because their development requires unique geological, hydrological, topographical and geographical characteristics. *Controlled inflows* into a reservoir include all releases from adjacent upstream reservoirs on the same river or its tributaries. Uncontrolled or *natural* inflows include all other inflows from surface runoffs, streams and undammed rivers. Water may flow out of a reservoir through various outlets such as derivations (to draw water for irrigation or other consumption), spillways (for flood protection) and penstocks (to produce electricity). Also, there may be water losses due to evaporation and seepage into the ground.

One of the most important uses of reservoirs is to produce electricity. In this case a hydro plant is provided near the reservoir. The quantity of energy produced by a hydro plant depends both on the flow through the turbines and the water head. The *water head* is the difference between forebay elevation and tailwater elevation, which are the reservoir levels respectively in front of the intake and at the exit of the draft tube.

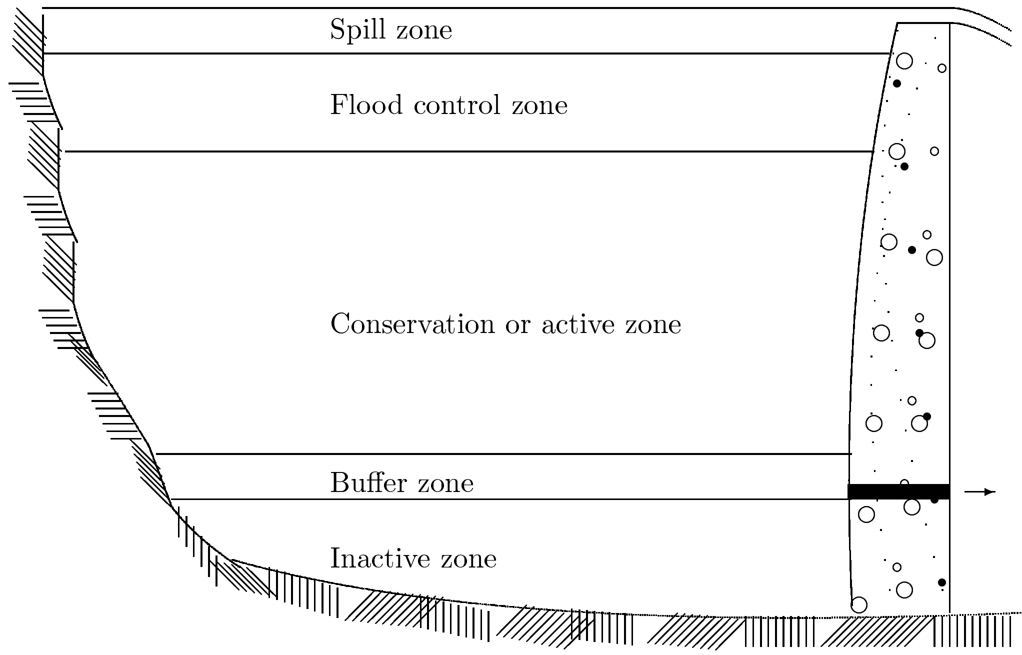

Figure 1.1: Storage allocation zones within a reservoir RZ

Reservoir systems are also used for a variety of other purposes such as flow control, depth regulation, flood control, water storage for irrigation or supply of drinking water, recreation, navigation, and fish and wildlife enhancement. The management of a multi-purpose reservoir system is a complicated process that must comply with public laws and regulations (public safety, environmental protection, and so forth). It usually attempts to find an effective compromise between the conflicting needs of different uses. For example, flood control requires depletion of reservoirs in advance of floods, and the maximum volume of unused storage has to be maintained until all danger of flooding is past. But energy needs require a full reservoir to allow greater turbine efficiencies for power generation. Moreover, recreational uses require a full pool during the vacation season, which coincides with the need to lower the pool to supply irrigation. Reservoir operators thus rely on *reservoir operating policies* to make release decisions that satisfy all conditions and allocate water equitably between the different uses.

Rule curve and storage allocation zones have been the first operating policies used to manage multi-purpose reservoirs. Operating policies associated to *rule curves* define the ideal storage pool level and discharges at different times of the year for each reservoir. The rule curve is based on historical operating practice. In operating policies based on multiple *zones*, the total storage volume of the reservoir is divided into several zones as in Figure 1.1, based on the placement of outlet structures and operational assignments. The *inactive zone* or *dead storage zone* represents the lower part of the reservoir that is not normally used. The *buffer zone* is above the inactive zone. Only essential needs are satisfied when the storage volumes are within this zone, usually as a result of a dry period. The *conservation* or *active zone* represents the volume of water that can be used to satisfy various beneficial uses including recreational and environmental needs. The *flood control zone* is above the conservation zone and it is reserved for flood detention especially during periods of abnormally high runoff. The *spill zone* is the upper portion of the pool, in which the downstream flows are at or near their maximum. See Loucks and Sigvaldason [26] for details about traditional operating policies.

The management and planning of multireservoir systems are often supported by a hierarchy of mathematical models. Stochastic optimization models usually lie in the top layer of this hierarchy, in which long term planning is performed. The planning horizon in these models exceeds one year, and reaches typically five to twenty years or more. Although reservoir operation happens continuously, the long term planning exercise normally separates the planning horizon into a number of time intervals, or periods, with a fixed time step of one month to one year. The purpose of the planning exercise is to assign storage and production targets in every period. These targets, in turn, are passed on to the next layer (medium term scheduling) in which a deterministic optimization model is applied to a shorter horizon with a smaller time step [13]. Discrete time MDP models are well suited for the long term planning of reservoir systems, especially for hydro-power production, because of the Markovian nature of the stochastic processes governing the natural water inflows.

1.3 Stochastic processes of natural inflows

sec:spni

The natural phenomena governing rainfall, runoff, river flows, and flood and drought characteristics are complex and largely unpredictable. The design and operation of dams and reservoirs must therefore take into account the high level of randomness present in these physical processes. The statistical properties of hydrological phenomena are usually obtained by analysing the time series based on historical records of river flows. In particular, largest annual flood intensities in successive years have been found to be independent random variables, while total annual flows tend to be autocorrelated and are often modeled as autoregressive or moving average processes [27].

Let the random variable D_{it} be the volume of natural inflows received in the i-th reservoir during period t. We denote by D_t the column vector of natural inflows in period t for all sites. The sequence $\{D_{it},\ t \in \mathbb{Z}\}$ is the stochastic process of natural inflows at reservoir i. Normally, the time index $t = 1$ corresponds to the first period of the planning horizon, and past periods have a negative or zero time index. Similarly, the sequence $\{D_t,\ t \in \mathbb{Z}\}$ is the (vector) stochastic process of natural inflows at all sites in the system.

Statistical hydrology

Statistical inflow data are usually available from historical records for a finite number of past periods, and most modelers also make a stationarity assumption about the stochastic process of natural inflows. Then various properties of the stochastic process can be inferred from the historical time series using appropriate statistical methods, such as the Box-Jenkins methods [4, 14, 30] for time series analysis. For example, McCleod et al. [30] analyzed the time series of average annual river flows from 1860 to 1957 for the Saint-Lawrence at Ogdensburg, New York. They obtained the following autoregressive model of order 3 (denoted AR(3)):

$$Z_t = 0.6219 Z_{t-1} + 0.1771 Z_{t-3} + E_t,$$

where $Z_t = D_t - \overline{D}$ are the centered inflows ($\overline{D}$ is the historic mean) and E_t are i.i.d.[1] Gaussian errors (innovations) with mean zero and coefficient of variation equal to six percent. Reservoir inflows can also be modeled as moving average (MA) processes or, more generally, as autoregressive moving average (ARMA) processes.

Multivariate models have also been proposed for the statistical analysis of natural inflows into multiple reservoir systems. For example, Salas et al. [35] examined the bivariate time series of annual river flows from 1932 to 1963 for the Skykomish and Green rivers located in the state of Washigton (USA). They obtained the following autoregressive process:

$$\begin{pmatrix} X_{1t} \\ X_{2t} \end{pmatrix} = \begin{pmatrix} 0.407 & 0 \\ 0 & 0.345 \end{pmatrix} \begin{pmatrix} X_{1,t-1} \\ X_{2,t-1} \end{pmatrix} + \begin{pmatrix} E_{1t} \\ E_{2t} \end{pmatrix},$$

[1] independent, identically distributed

where X_{it}, the standardized flow into site i in period t, is given by

$$X_{it} = \frac{D_{it} - \mu_i}{\sigma_i},$$

with μ_i the historical mean of the flows D_{it} at site i, and σ_i their standard deviation. The residuals E_{1t} and E_{2t} are independent of previous periods and follow a multivariate normal distribution with mean zero and variance-covariance matrix

$$\Sigma = \begin{pmatrix} 0.837 & 0.823 \\ 0.823 & 0.881 \end{pmatrix}.$$

Similarly, the annual flows of the Wolfe and Fox rivers, located in the state of Wisconsin (USA), were modeled in [5] using records from 1899 to 1965. After a logarithmic transformation of data, the following moving average process was obtained:

$$\begin{pmatrix} Z_{1t} \\ Z_{2t} \end{pmatrix} = E_t + \begin{pmatrix} -0.626 & 0 \\ 0 & -0.543 \end{pmatrix} E_{t-1},$$

where $Z_{it} = \log(D_{it}) - \mu_i$ with μ_i the mean of $\log(Z_{it})$, and the vector E_t of innovations is i.i.d., having a multivariate normal distribution with mean zero and variance-covariance matrix

$$\Sigma = \begin{pmatrix} 0.0552 & 0.0502 \\ 0.0502 & 0.075 \end{pmatrix}.$$

In the above models, the matrices of autoregressive and moving-average coefficients are diagonal. Such models are said to be *contemporaneous* because the correlations between flows at different sites in a given period are accounted for in the variance-covariance matrix, while serial correlations between flows at different sites are neglected. Contemporaneous ARMA models are much easier to estimate than the general multivariate ARMA models, and they were found to reproduce well the main statistical characteristics of the time series analyzed. Moreover, their selection is often justified a priori by physical considerations [5, 35, 40].

Stedinger et al. [40] examined different ways of fitting the following ARMA(1,1) model to bivariate hydrologic time series by building upon simple univariate procedures:

$$Z_t = \Phi Z_{t-1} - \Theta E_{t-1} + E_t,$$

where $Z_t = D_t - \mu$ is the 2×1 vector of centered flows in period t, Φ and Θ are the 2×2 coefficient matrices and E_t is the 2×1 vector of time-independent normally distributed random fluctuations. Three models were compared: a contemporaneous ARMA(1,1) model where the matrices Φ and Θ are diagonal, a univariate ARMA(1,1) model of aggregate flows on both rivers with a simple disaggregation procedure described in [41], and a non-diagonal ARMA(1,1) model. The study concluded that the first two models performed as well as the multivariate nondiagonal ARMA(1,1) model.

Aggregation/disaggregation models for multivariate water resources time series were proposed in [41, 49]. These models reproduce (relevant) appropriate statistics at several time intervals. Aggregation/disaggregation can be used to obtain flows at different sites and also to model seasonality. Seasonal inflow models can also be modeled as periodic ARMA processes but their estimation is sometimes difficult especially in the multivariate case [35].

Sophisticated ARMA or aggregation/disaggregation models have been used mostly for forecasting purposes or for synthetic streamflow generation. Inflow forecats are then used in deterministic optimization models to obtain open loop solutions. On the other hand, synthetic inflow sequences are generated at random for use in Monte Carlo simulation studies to evaluate the design of a reservoir system, or to compare various operating policies.

A review of stochastic models in hydrology is given in [54]. However, the stochastic structure of natural inflow phenomena is not fully understood and is the object of intense research. For instance, recent studies indicate that regional average streamflow statistics contain more information about the variability and persistence of streamflow at a particular site than does the individual streamflow record at this site. See [50].

Discretized inflow models

Most stochastic optimization models for reservoir operations use discrete random variables to model the natural inflows. This assumption is fundamental when a discrete dynamic programming model is used in which the reservoir levels and discharges are discretized. However, in more realistic models with continuous state and action variables, computation of the expected value of a state-action pair often requires the discretization of the natural inflow distribution, which corresponds to a quadrature rule for numerical integration [9, 18]. A numerical study of discretization error was also presented in [21]. Exceptions to this rule are the linear quadratic Gaussian controller [3] and the myopic-affine dynamic model [23], in which inflow discretization is avoided. Here, we review briefly some discrete inflow models used in reservoir applications of MDP.

In the long-term reservoir scheduling model of [45], the year is broken down in two periods of six months, representing the winter and summer seasons, respectively. The random variable D_t expresses the inflows in terms of potential energy added into the aggregate reservoir system during period t. The energy inflows in period t are assumed correlated with the previous summer's inflows through a seasonal autoregressive model with lag 1 dependence if t is the winter season, and lag 2 dependence if t is the summer season:

$$D_t = \begin{cases} \alpha_{1,w} + \alpha_{2,w} D_{t-1} + \alpha_{3,w}\xi_t, & \text{in winter,} \\ \alpha_{1,s} + \alpha_{2,s} D_{t-2} + \alpha_{3,s}\xi_t, & \text{in summer,} \end{cases}$$

where the coefficients $\alpha_{i,w}$ (for winter season) and $\alpha_{i,s}$ (for summer season), $i = 1, 2, 3$, are constants and the ξ_t are i.i.d. standard normal. Moreover, the random variable D_t is discretized into eleven levels.

In the aggregate reservoir model of [44], the inflows are also expressed in terms of energy. The seasonal autoregressive lag-one model, used to represent the stochastic inflow process, has the form

$$\frac{(D_t - \mu_t)}{\sigma_t} = \rho_t \frac{(D_{t-1} - \mu_{t-1})}{\sigma_{t-1}} + (1 - \rho_t^2)^{0.5} U_t$$

where D_t is the energy inflow during period t (a month), with mean μ_t and standard deviation σ_t, ρ_t is the correlation coefficient between inflows in stage t and stage $t-1$, and U_t is a 3-parameter (mean, standard deviation, and shift) lognormal random variable. The energy inflows are discretized into ten levels.

A different approach, called *sampling* stochastic dynamic programming, was presented in [17]. With this approach, the serial and spatial structures of the streamflow process are captured by using a large number of randomly generated 12 month streamflow sequences. A conditional distribution, developed using a historical time series of streamflow forecasts, is then assigned to various streamflow scenarios given a streamflow forecast. The approach was applied to the North Fork Feather River hydroelectric system composed of nine reservoirs and located in California. Similarly, the DP models developed in [39] derived a reservoir release policy by using the best forecast of the current period's inflow as hydrologic state variable instead of the previous period's inflow. The potentiel advantage of the proposed approach was illustrated using the Nile river basin as a case study.

In [42], the performance of operating policies derived using stochastic DP models with different sets of hydrologic state variables were compared.

Different choices for hydrologic state variables were considered: current period flow, previous period flow, and current period or seasonal flow forecasts. The stochastic process representing the hydrologic variables was described by a month-to-month Markov chain and was discretized to allow the calculation of transition probability matrices for the hydrologic state variables. The authors found that for a benefit function stressing energy maximization, all policies did nearly as well. However, for benefit functions involving large water and firm power targets and severe penalties for shortages, the policies that employed more complete hydrologic information performed significantly better.

1.4 Dynamic programming model

sec:dpm

We present general MDP model for the long term optimization of a multireservoir system composed of m reservoirs. The planning horizon of the study is divided into T periods $t = 1, \ldots, T$. Dynamic programming is an attractive tool for modeling such problems because the stochastic nature of inflows and the nonlinear functions associated with energy generation can be modeled explicitly. Definitions of state variables, decision variables, constraints, and objective function for the general multi-reservoir system, whose main purpose is hydro-power production, are provided below.

State variables

Including hydrologic variables in the state vector allows consideration of the serial correlation of natural inflows. Additional hydrologic state variables can be added to model spatial correlation for multireservoir systems. However, incorporating more hydrologic information in the model improves reservoir operation, but it adds dimensionality into the model.

The state vector at period t, considered in the model, includes the volume of water in storage in each reservoir and also some hydrologic information describing recent hydrological activity (hydrologic variables). The volume of water in storage in reservoir i at period t is denoted as S_{it}. We suppose that the volume of water inflow D_{it} at site i in period t is a function of the inflows in the K previous periods. Hence, the state of the system X_t at period t is represented by $K+1$ column m-vectors (m sites) and may be written as $X_t = (D_{t-K}, \ldots, D_{t-1}, S_t)^T$. The superscript T represents the vector transpose. It is convenient to assume that the natural inflows D_{it} are collected, at site i, at a constant flow rate over the duration of period t.

Decision variables

The decision to be taken at each period represents the release through the turbines Z_t and the spill Y_t. We assume that the decision depends on both the state X_t and the inflow D_t (see e.g. [45]). In practice, the release Z_t and the spill Y_t represent the mean flow during the period, and the model considers they will flow at a constant rate over the period. Thus, the action, unknown at the beginning of the period, is a random variable with a well known distribution conditioned on X_t.

The action at each period, denoted as A_t, is represented by the pair of m-vectors (Z_t, Y_t). The ith entry Z_{it} of the vector Z_t represents the release at reservoir i during period t, used to produce energy, and the ith entry Y_{it} of the vector Y_t represents the volume of water spilled (without energy production) at reservoir i during the period t. Usually, in reservoir applications, the state X_t and the action A_t are assumed to be random vectors resulting from random variables with continuous joint distributions, except possibly for a small number of mass points (e.g., droughts, storage capacities, etc.). However, to simplify the analysis and for the purpose of numerical computations, many studies assume that X_t and A_t have finite, discrete distributions instead. We denote the realisations of X_t and A_t respectively by x_t and a_t.

Transition equations

The first transition equation describes the dynamic behavior of the system. It is the usual water conservation equation

$$S_{t+1} = S_t - BZ_t - CY_t + D_t, \tag{1.1}$$

eq:balance

where B and C are the $m \times m$ connectivity matrices (or network incidence matrices), used to allocate releases and spills from upstream reservoirs. For systems in which the spilled water is routed on the same river as the turbine releases, we have $C = B$. On other systems, the spilled water is expelled from the system, so that $C = I$. We note that evaporation and leakage are neglected in this model. The second transition equation represents the evolution of the stochastic process of the natural inflows. Models used to represent such processes were discussed in the previous section. We can assume, for example, a multivariate autoregressive process of order K, denoted as AR(K), with transition equation

$$D_t = \sum_{\ell=1}^{K} \Phi_t^{\ell} D_{t-\ell} + E_t, \tag{1.2}$$

eq:autoreg

where Φ_t^{ℓ} is an $m \times m$ matrix of "lag ℓ" autoregressive coefficients and E_t is a random vector of innovations, assumed independent of all prior states and actions and with a known, arbitrary joint distribution. Verification of the Markovian property is immediate by inspection of (eq:balance 1.1) and (eq:autoreg 1.2).

Objective and economic structure

The total amount E_t of energy produced by the whole system in period t can be represented as

$$E_t(S_t, Z_t) = \sum_{i=1}^{m} E_{it}(S_{it}, Z_{it}).$$

The quantity of electricity E_{it} produced at the hydro-plant i in period t is a nonlinear function of the volume of water in storage S_{it} and the volume released through the turbines Z_{it}. For instance, this function can be expressed as in soares:carneiro [38] by

$$E_{it}(S_{it}, Z_{it}) = K[H_i^{up}(S_{it}) - H_i^{dw}(Z_{it})]Z_{it}$$

where K is a constant and H^{up} and H^{dw} are the upstream and downstream reservoir level functions. The difference $H_i^{up} - H_i^{dw}$ is the water head factor, which greatly influences the efficiency of electric generation turbines.

The objective is to determine a policy that will maximize the expected discounted rewards of system operation over a finite horizon. The benefit function is expressed as a nonlinear no concave function of the system states and decisions variables. The system operating function (or one-step reward), in period t, can be considered as

$$R_t = g_t\left(E_t(S_t, Z_t)\right) + h_t(S_t)$$

where the function $g_t\left(E_t(S_t, Z_t)\right)$ is the benefit of the hydropower system in period t and the function $h_t(S_t)$ represent the sum of revenues from other uses of the water in the reservoirs. For many utilities, the function g_t is a concave piecewise linear function of the energy produced $E_t(S_t, Z_t)$.

Inequality constraints

Inequality constraints account for physical limitations on the states and decision variables. The set $\mathbb{A}_t(x_t, d_t)$ of allowable actions in period t comprises all pairs of vectors z_t and y_t satisfying these inequalities as well as the balance equation (eq:balance 1.1). There are lower and upper limits on the allowable storage volumes during the period. These limits may include legal restrictions (such as navigational safety, flood control or scheduled maintenance) as well as reservoir capacity constraints. They are represented by constraints on the stored volumes at epoch $t + 1$:

$$\underline{s}_{i,t+1} \leq s_{i,t+1} \leq \overline{s}_{i,t+1}. \tag{1.3}$$ eq:storage:a

There are lower an upper limits on the allowable volumes that can flow through the turbine during the period. These limits incorporate nonnegativity, legal requirements and turbine capacities:

$$\underline{z}_{it} \leq z_{it} \leq \overline{z}_{it}. \tag{1.4}$$ eq:turbine

Finally, there are nonnegativity and upper bound constraints on the spill variables,

$$0 \leq y_{it} \leq \overline{y}_i, \tag{1.5}$$ eq:spill

where $\overline{y}_i$ is the spillway capacity of the site i.

The case where there is no feasible solutions satisfying the previous constraints can be considered by adding penalty functions in the objective to control release water shortage and flood. See terry:etal [44], for objective function with penalities for failure in load supply.

Optimization

The expected future consequence of choosing an action a_t is given by the cost to go function that can be written as

$$V_t(x_t) = \mathbb{E}[V_t(x_t; D_t) \mid D_{t-k} = d_{t-k}, \ldots, D_{t-1} = d_{t-1}], \tag{1.6}$$ eq:future

where $\mathbb{E}$ is the expectation and $V_t(x_t; d_t)$ is the dynamic programming optimal value function, representing the expected operating reward from stage t to the end of the planning horizon. This function is the recursive DP equation that can be written, for period $t = T - 1, \ldots, 1$, as

$$V_t(x_t; d_t) = \max_{a_t \in \mathbb{A}_t(x_t, d_t)} [r_t(x_t, a_t) + \beta V_{t+1}(x_{t+1})], \tag{1.7}$$ eq:bellman

where $r_t(x_t, a_t)$ is the immediate reward.

The functions $V_t(x_t)$ and $V_t(x_t; d_t)$ are obtained by backward induction, starting with $t = T$, using (1.6) and (1.7). At each epoch t, we need to solve Bellman's equation (1.7) for all possible states $x_t \in \mathbb{X}$.

The expected terminal reward function, is assumed given and depends only on the state at the final stage T. It can be written as

$$V_T(x_T) = \mathbb{E}\left[f(x_T) \mid X_T = x_T\right],$$

where $f(x_T)$ is the terminal reward function.

The cost to go function generally cannot be calculated exactly even if all the elements in the state vector are continuous. An approximate solution of V_t can be obtained after the discretization of D_t and X_t. So, the continuous state space is replaced by a grid and the value of the cost to go function is calculated at the grid points. See [9, 18, 19, 22].

1.5 Applications

The model of [44], adopted officially in 1979 to manage the Brazilian national electrical generating system, determines the optimal hydro and thermal generations in the system. In addition, this model is used to calculate the expected incremental costs of producing the thermal generation. These costs represent the increase of the expected future operation cost if hydro generation is increased by one Mwh. They are used to make decisions about selling or purchasing energy. The model belongs to a chain of generation expansion planning models used to establish when and where to build the new plants. It has also been used, to determine the reliability indices of trial expansion plans for 10 to 30 years.

The PERESE model for the long-term scheduling of reservoirs, developed especially for the Hydro-Québec system, is presented in [45]. The model is useful to study different generation expansion scenarios in order to determine if the demand can be satisfied, for each scenario, with the desired reliability and to calculate the expected benefits and costs. The model is also useful to decide about construction of the generation expansion plan and to determine guidelines for middle-term horizon studies. It can also be used to determine if additional firm energy can be sold on the spot markets.

The optimal operating policy developed in [46] gives not only the hydroelectric energy to produce in a month but also the expected marginal cost of the hydroelectric energy produced. This last information is very important in the current context of deregulated markets. To deal in this deregulated market, certain hydroelectric producers purchase energy in spot markets when the price is low. They store it in their reservoirs (by letting them fill with water) and sell it in the spot markets later when the prices are higher. To realize a profit it is important for the producers to determine the exact marginal production cost. The model can also be used to determine the volume of energy that can be sold in the spot markets without endangering the reliability of the system, and to study different scenarios of electrical energy demand and system expansion.

A model of the Shasta-Trinity subsystem, which is a part of the Central Valley Project located in Northern California and operated by the United States Bureau of Reclamation, is formulated in [42]. The benefit function of the model seeks to maximize energy production and to meet reliably water and firm power targets. In addition, the Shasta-Trinity subsystem must produce a part of the energy that the Central Valley Project contracted to provide. Other applications on this system are reported in [43] using other approaches with the objective of meeting water and energy targets.

The control method presented in [11] can be used for real time operation of a reservoir system as well as for developing policy-making guidelines. It was implemented for the High Aswan dam on Lake Nasser in Egypt. The application seeks to determine the optimal release sequence over a 36 months horizon. In the case when storage exceeds the reservoir capacity, the spilled water is diverted to a depression area where it evaporates. The objective is to maximise the expected energy generation while satisfying downstream water supply requirements, and taking into account the monthly evaporation rates. In spite of the fact that the High Aswan dam suffers heavy evaporation losses, its storage capacity is adequate for current water supply purposes. The tradeoff sought in the Aswan dam application is to maintain the lowest reservoir elevation necessary for meeting the water supply requirements at the expense of energy losses due to the lower hydraulic head. This policy tends to minimize water losses due to evaporation, since evaporation depends on the reservoir's volume and surface. This policy has practical interest since water rather than hydropower availability is the limiting factor in the development of the Egyptian economy. Another benefit of operating the High Aswan dam at low reservoir levels is safety in case of seismic activity. The High Aswan Dam system has also been modeled in [39] to define a reservoir release policy and to calculate the expected benefits from future operations. The application used alternative formulations of SDP models and numerical examples on the High Aswam dam system to demonstrate that the approach used can identify more efficient reservoir operations policies.

A real time optimal control approach was applied in mclaughlin:a [29], to a system of hydropower reservoirs in the Caroni river basin in Venezuela. The system is composed of one very large and one moderately sized reservoir. The objective of the application is to track an optimal trajectory that provides a reliable power output to satisfy contractual obligations. Inflows records, collected since 1960, allow consideration of the hydrologic seasonality inputs for the application. The results of the Caroni system application illustrate that there is a trade-off between operation strategies which sacrifice hydrologic complexity in exchange for a more nearly optimal solution and those which sacrifice theorical optimality in exchange for more accurate hydrologic predictions. For this specific application, the authors believe that improvements in the model that predict reservoir inflows have more impact on the performance of the system than improvements in the optimization algorithm. The method used in this application is expected by its authors to work best when used with large reservoirs.

The control approach of fletcher:a [8] was applied to the Great Lakes levels regulation problem. The Great Lakes, forming a chain of natural reservoirs situated between USA and Canada, consist of Lakes Superior, Michigan, Huron, Erie, and Ontario. The larger lakes, Superior, Michigan and Huron are modeled in the application as infinite capacity reservoirs with no upper or lower bounds. The management of the Great Lakes system should not cause any disagreement between its five major groups of interest: commercial navigation, riparian or shore property, hydro-power, recreational boating, and environment. Thus, to provide a Great Lakes levels regulation plan, the application is performed to minimize the variations of lakes levels over the planning horizon, knowing the storage and release targets, prescribed over a twelve month time horizon. The application can also provide the estimation of reliability and failure probability. Such informations are useful to determine the risks associated with investment decisions involving lake resources and they are useful also in the context of reservoir operation and design.

1.6 Computational issues

sec:ci

The optimization of operating policies for a multireservoir system is a stochastic nonlinear dynamic programming problem. Its main drawback is the need to discretize the state variables. This may lead to a very large number of combinations even with a small number of variables. For small systems the optimization problem can be solved by classical discrete dynamic programming, but for large systems the usefulness of this approach is limited, because computational complexity increases exponentially with the number of reservoirs in the system. This is the well known *curse of dimensionality* (see, e.g., [20] for more details).

Early water reservoir applications of MDP date back more than fifty years. They were concentrated on single reservoir and on dynamic programming algorithms for solving these problems. In the sixties and seventies, models with two reservoirs were solved by SDP [52, 53]. Then models with three or four reservoirs were solved in the eighties [9]. Researchers have thus had to resort to other methodologies by combining SDP with other sophisticated approaches to extend the studies to models with more than four reservoirs.

Multistage stochastic programming has been intensively used recently for water reservoir applications. The notion of inflow scenarios is normally used in algorithms based on this approach. In this field, Benders decomposition type models seem to perform well for the linear and piecewise linear problems. The approach of Pereira and Pinto [32, 33], called SDDP (Stochastic Dual Dynamic Programming), is based on Benders decomposition and it uses duality theory to approximate the value function by a piecewise linear function. A comparaison of the Benders and dynamic programming approaches is given in [2]. An approach using duality theory and parametric linear programming, but not Benders decomposition, was developed in [34]. This approach was applied to the stochastic problem of scheduling hydro and thermal power generation of a system with two reservoirs. The procedure presented in [15] is based on the same concepts as Pereira and Pinto's algorithm and it was used for scheduling a large hydroelectric generation problem with a 24 month planning horizon. Another algorithm, based on Benders decomposition, but for nonlinear convex multistage stochastic programming problem was developed in [36]. This algorithm is specialized for a hydropower scheduling application and was applied to a problem with 176 powerhouses on 94 reservoirs with 174 additional controlled water spillways over a 24 month planning horizon. The weakness of these approaches is that the head effects were neglected.

Other multi-reservoir methods with continuous state and action variables have been developed using an approximation of the expected value function. In the parameter iteration method, developed in [10], the value function (or cost to go function) is approximated at each stage, of the dynamic program, by a simple functional form with a small number of parameters. The control is a function of the state characterized by a set of parameters that are improved at each iteration by least squares minimization. In their gradient dynamic programming method, Foufoula-Georgiou and Kitanidis [9] use Hermite interpolation to approximate the cost to go function and they reduce dimensionality by using a coarser state discretization. A similar approach, but using tensor-product cubic splines, has been developed in [16], with a parallel computing extension in [7]. The approach was further extended in [6] to higher dimensions using multivariate adaptive regression splines on orthogonal arrays.

Another way to avoid the curse of dimensionality consists of formulating an MDP problem with a small number of discrete state and action variables by using composite reservoirs that can be solved by discrete DP. Models in which the storage capacities of the many reservoirs in the system are aggregated into a single composite reservoir of potentiel energy are proposed in [44, 45]. The mains drawbacks of such single composite reservoir models is that it is difficult to derive optimal operating policies for individual reservoirs in the system, from the aggregate policy. To address this difficulty, an approach for the optimal operation was proposed in [47] for a system composed of reservoirs in parallel and in [48] in the case where reservoirs are in series. The two approaches consist of using aggregation and decomposition to break up the original model into subproblems of two reservoirs. In the case of a general arborescent multireservoir system, an aggregate stochastic dynamic programming model for determining an operating policy was proposed in [1].

Another set of models, belonging to the optimal control approach, and involving no discretization of the system variables, have been used in water reservoirs applications. In these models, based on the linear quadratic Gaussian controller [3], the optimal solution of the cost to go function is deduced analytically, with parameters that can be determined numerically by solution of a "static" optimization problem. An extended linear quadratic Gaussien (ELQG) control technique was applied to a three reservoir system in [11]. In this approach the state variables are replaced by the mean and variance of the storages and an assumption on the probability distribution of the storage state variables is required. The ELQG technique is an iterative refinement of the linear-quadratic control approach, used in [25, 51] in the context of operation of a multireservoir system for flood control. A similar approach was used in [29] for the real-time control of a hydropower system with two reservoirs. The main drawback of the linear quadratic Gaussian controller is that it is limited to unbounded reservoir systems. The stochastic control approach presented recently in [8] remedy to this problem by considering the storage bounds explicitly in the expressions for the reservoir systems dynamics by modifying the dynamic equation. As in ELQG, the state variables are the mean and variance of the storages but no assumption on the probability distribution of the storage state variables is required. The limitation of this method lies in the fact that Taylor series approximations are used in the derivation.

1.7 Research perspectives

sec:rp

Rising demands for water for different uses are forcing stiff competition over allocation of scare water resources among different users. Decision makers face the challenge of having to arbitrate between two conflicting objectives: on the one hand to manage and conserve water supplies, and on the other hand, to satisfy all needs in face of growing demand from population growth and industries. There is a clear need for more research in the development and refinement of models and methods to help deriving efficient operating policies for reservoir systems, and we feel that the potential of the MDP tools has not yet been fully exploited.

Recent technological developments in digital computing have permitted an evolution in the size of models that can be solved by discrete DP, spline approximations, and stochastic programming. But despite this progress, more research is still needed to solve large stochastic reservoir systems in a precise, detailed manner, under more realistic assumptions. Methods still have to be developed that consider a realistic representation of the natural nonlinearities of hydroelectric generation, by taking head effect into account, and reservoir operations. These methods should also consider a more detailed description of the physical system and uncertainty on inflows, demands, and price of energy and fuel.

Bibliography

archibald:etal [1] Archibald, T. W. A and K. I. M. McKinnon, K.I.M and Thomas, L.C. (1997). An aggregate stochastic dynamic programming model of multiple reservoir systems. *Water Resources Research* **33**, pp. 333–340.

archibald:buchanan:mckinnon [2] Archibald, T.W. and Buchanan, C.S. and McKinnon K.I.M and Thomas, L.C., (1996) Nested Benders decomposition and dynamic programming for reservoir optimization. Working paper, Department of Business Studies, The University of Edinburgh, Edinburgh, UK.

bertsekas:a [3] Bertsekas, D.P. (1976). *Dynamic Programming and Stochastic Control*. Academic Press, New York.

box:jenkins [4] Box, G.E.P. and Jenkins, G.M. (1976). *Time Series Analysis: Forecasting and Control*. Holden-Day, San Francisco.

camacho:a [5] Camacho, F. and McLeod, A.I. and Hipel, W.H. (1985). Contemporaneous autoregressive-moving average (CARMA) modeling in water resources. *Water Resources Research* **21**, pp. 709–720.

chen:etal [6] Chen, V.C.P., Ruppert, D. and Shoemaker, C.A. (1999). Applying experimental design and regression splines to high-dimensional continuous-state stochastic dynamic programming. *Operations Research* **47**, pp. 38–53.

eschenbach:etal [7] Eschenbach, E.A., Shoemaker, C.A. and Caffey, H.M. (1995). Parallel algorithms for stochastic dynamic programming with continuous state and control variables. *ORSA J. on Computing* **7**, pp. 386–401.

fletcher:a [8] Fletcher, S.G. and Ponnambalam, K. (1998). A Constrained state formulation for the stochastic control of multireservoir systems. *Water Resources Research* **34**, pp. 257–270.

foufoula:a [9] Foufoula-Georgiou, E. and Kitanidis, P.K., (1988). Gradient dynamic programming for stochastic optimal control of multidimensional water resources systems. *Water Resources Research* **24**, pp. 1345–1359.

gal:a [10] Gal, S. (1979). Optimal management of a multireservoir water supply system. *Water Resources Research* **15**, pp. 737–749.

georgakakos:b [11] Georgakakos, A.P., (1989). Extended linear quadratic Gaussian control for the real time operation of reservoir systems. In *Dynamic Programming for Optimal Water Resources Systems Analysis*, Prentice Hall, Esogbue Editor, pp. 329–360.

gessford:karlin [12] Gessford, J. and Karlin, S. (1958). Optimal policy for hydroelectric operations. In *Studies on the Mathematical Theory of Inventory and Production*, Stanford University Press, Stanford, Calif., pp. 179–200.

hanscom:etal [13] Hanscom, M. A., Lafond, L., Lasdon, L. and Pronovost, G. (1980). Modelling and resolution of the medium term energy generation planning problem for a large hydro-electric system. *Management Science* **26**, pp. 659–668.

hipel:etal [14] Hipel, K.W. and McLeod, A.I. and Lennox, W.C. (1977). Advances in Box-Jenkins modeling 1: model construction. *Water Resources Research* **13**, pp. 567–575.

jacobs:al [15] Jacobs, J.J., Freeman, G., Grygier, J., Morton, D., Schultz, G.L., Staschus, K., and Stedinger, J.R. (1995). Socrates: a system for scheduling hydroelectric generation under uncertainty. *Annals of Operations Reseach* **59**, pp. 99–133.

johnson:etal:a [16] Johnson, S.A. and Stedinger J.R. and Shoemaker C.A., Li, Y. and Tejada-Guibert, J.A. (1993). Numerical solution of continuous-state dynamic programs using linear and spline interpolation. *Operations reseach* **41** pp. 484–500.

kelman:a [17] Kelman, J. and Stedinger, J.R. and Cooper, L.A. and Hsu, E. and Tuan, S. (1990). Sampling stochastic dynamic programming applied to reservoir operation. *Water Resources Research* **26**, pp. 447–454.

kitanidis:a [18] Kitanidis, P.K. and Foufoula-Georgiou, E. (1987). Error analysis of conventional discrete and gradient dynamic programming. *Water Resources Research* **23**, pp. 845–856.

lamond [19] Lamond, B. F. (1997) Stochastic reservoir optimization using piecewise polynomial approximations. Working paper, Faculté des sciences de l'administration, Université Laval, Québec, Canada.

lamond:boukhtouta [20] Lamond, B. F. and Boukhtouta, A. (1996). Optimizing long-term hydro-power production using Markov decision processes. *Int. Trans. Opl. Res.* **3**, pp. 223–241.

lamond:bachar [21] Lamond, B. F. and Bachar, B. (1998). Une étude numérique de la discrétisation des apports aléatoires pour un réservoir non saisonnier. *INFOR* **36**, pp. 247–260.

lamond:lang [22] Lamond, B. F. and Lang, P. (1996). Lower bounding aggregation and direct computation for an infinite horizon one-reservoir model. *European Journal of Operational Research* **95**, pp. 404–410.

lamond:sobel [23] Lamond, B.F. and Sobel, M.J. (1995). Exact and approximate solutions of affine reservoir models. *Operations Research* **43**, pp. 771–780.

little:a [24] Little, J. D. C. (1955). The use of storage water in a hydroelectric system. *Operations Research* **3**, pp. 187–197.

loaiciga:a [25] Loaiciga, H.A. and Mariño, M.A. (1985) An approach to parameter estimation and stochastic control in water resources with an application to reservoir operation. *Water Resources Research* **21**, pp. 1575–1584.

loucks:sigvaldason [26] Loucks, D. P. and Sigvaldason, O. T. (1982). Multiple-reservoir operation in North America. In Z. Kaczmarek and J. Kindler (Eds), *The Operation of Multiple Reservoir Systems*, pp. 2–103. IIASA, Laxenburg, Austria.

maidment [27] Maidment, D. R. (1993). *Handbook of Hydrology.* McGraw-Hill, New-York.

masse:a [28] Massé, P. B. D. (1946). *Les Réserves et la Régulation de l'Avenir dans la Vie Économique.* Hermann, Paris.

mclaughlin:a [29] McLaughlin, D. and Velasco, H.L. (1990). Real-time control of a system of large hydropower reservoirs. *Water Resources Research* **26**, pp. 623–635.

mcleod:etal [30] McLeod, A.I. and Hipel, K.W and Lennox, W.C. (1977). Advances in Box-Jenkins modeling 2: Applications. *Water Resources Research* **13**, pp. 577–586.

oliveira:loucks [31] Oliveira, R. and Loucks, D. P. (1997). Operating rules for multireservoir systems. *Water Resources Research* **33**, pp. 839–852.

pereira:b [32] Pereira, M.V.F. and Pinto, L.M.V.G. (1985). Stochastic optimization of a multireservoir hydroelectric system: A decomposition approach. *Water Resources Research* **21**, pp. 779–792.

pereira:d [33] Pereira, M.V.F. and Pinto, L.M.V.G. (1991). Multi-stage stochastic optimization applied to energy planning. *Mathematical Programming* **52**, pp. 359–375.

read:a [34] Read, E. G. (1989). A dual approach to stochastic programming for reservoir release scheduling. In *Dynamic Programming for Optimal Water Resources Systems Analysis*, Prentice Hall, Esogbue Editor, pp. 361–372.

salas:a [35] Salas, J.D. and Tabios III, G.Q. and Bartolini, P. (1985). Approaches to multivariate modeling of water resources time series. *Water Resources Bulletin* **21**, pp. 683–708.

salinger:a [36] Salinger, D.H. (1997). *A Splitting Algorithm for Multistage Stochastic Programming with Application to Hydropower Scheduling*. Ph.D. thesis, University of Washington.

shaw [37] Shaw, E. M. (1988). *Hydrology in Practice*. Van Nostrand Reinhold, England.

soares:carneiro [38] Soares, S. and Carneiro, A.A.F.M. (1990). Optimal operation of reservoirs for electric generation. *Proceedings of the IEEE Power Engineering Society 1990 Summer Meeting*, Minneapolis, Minnesota.

stedinger:b [39] Stedinger, J.R. and Sule, B.F. and Loucks, D.P. (1984). Stochastic dynamic programming models for reservoir operation optimization. *Water Resources Research* **20**, pp. 1499–1505.

stedinger:c [40] Stedinger, J.R. and Lettenmaier, D.P. and Vogel, R.M. (1985). Multisite ARMA(1,1) and disaggregation models for annual streamflow generation. *Water Resources Research* **21**, pp. 497–509.

stedinger:d [41] Stedinger, J.R. and Vogel, R. M. (1984). Disaggregation procedures for generating serially correlated flow vectors. *Water Resources Research* **20**, pp. 47–56.

tejada:a [42] Tejada-Guibert, J.A., Johnson, S.A. and Stedinger, J.R. (1995). The value of hydrologic information in stochastic dynamic programming models of a multireservoir system. *Water Resources Research* **31**, pp. 2571–2579.

tejada:b [43] Tejada-Guibert, J.A. and Johnson, S.A. and Stedinger, J.R. (1993). Comparaison of two approaches for implementing multireservoir operating policies derived using stochastic dynamic programming. *Water Resources Research* **29**, pp. 3969–3980.

terry:etal [44] Terry, L. A., Pereira, M. V. F., Araripe Neto, T. A., Silva, L. F. C. A. and Sales, P. R. H. (1986). Coordinating the energy generation of the Brazilian national hydrothermal electrical generating system. *Interfaces* **16**, pp. 16–38.

turgeon:d [45] Turgeon, A. (1992). PERESE: The New Hydro-Québec optimization model for the long-term scheduling of reservoirs. In *Hydropower '92: Proceedings of the 2nd International Conference on Hydropower*, Held in Lillehamer, Norway Ed. Broch, E. and Lysne, D.K., Balkema publisher, Rotterdam, pp. 603–612.

turgeon:charbonneau [46] Turgeon, A. and Charbonneau, R. (1998). An aggregation - disaggregation approach to long-term reservoir management. *Water Resources Research* **34**, pp. 3585–3594.

turgeon:a [47] Turgeon, A. (1980). Optimal operation of multireservoir power systems with stochastic inflows. *Water Resources Research* **16**, pp. 275–283.

turgeon:c [48] Turgeon, A. (1981) A decomposition method for the long-term scheduling of reservoirs in series. *Water Resources Research* **17**, pp. 1565–1570.

valencia:schaake [49] Valencia, R.D. and Schaake, J.C. (1973). Disaggregation processes in stochastic hydrology. *Water Resources Research* **9**, pp. 580–585.

vogel:etal [50] Vogel, R.M., Tsai, Y. and Limbrunner, J.F. (1998). The regional persistence and variability of annual steamflow in the United States. *Water Resources Research* **34**, pp. 3445–3459.

wasimi:a [51] Wasimi, S.A. and Kitanidis, P.K. (1983). Real-time forecasting and daily operation of a multireservoir system during floods by linear quadratic Gaussian control. *Water Resources Research* **19**, pp. 1511–1522.

yak:a [52] Yakowitz, S. (1982). Dynamic programming applications in water resources. *Water Resources Research* **18**, pp. 7673–7696.

yeh:a [53] Yeh, W. W.-G. (1985). Reservoir management and operations models: a state-of-the-art review. *Water Resources Research* **21**, pp. 1797–1818.

yevjevich:a [54] Yevjevich, V. (1987). Stochastic models in hydrology. *Stochastic Hydrological Hydraulic* **1**, pp. 17–36.