An effective streamflow process model for optimal reservoir operation using stochastic dual dynamic programming

L. Raso, Pierre-Olivier Malaterre, Jean-Claude Bader

Author-produced version of the article published in Journal of Irrigation and Drainage Engineering-ASCE, 2017, N°143(4)
The original publication is available at http://ascelibrary.org
Doi: 10.1061/(ASCE)WR.1943-5452.00007466

An effective streamflow process model for optimal reservoir operation using Stochastic Dual Dynamic Programming

Luciano Raso[1], Pierre-Olivier Malaterre[1], and Jean-Claude Bader[2]

Corresponding author: Luciano Raso, IRSTEA, Montpellier, France. (luciano.raso@irstea.fr)

[1]IRSTEA, UMR-GEAU, Montpellier,

France.

[2]IRD, UMR-GEAU, Montpellier, France.

Author-produced version of the article published in Journal of Irrigation and Drainage Engineering-ASCE, 2017, N°143(4)
The original publication is available at http://ascelibrary.org
Doi: 10.1061/(ASCE)WR.1943-5452.00007466

Abstract. We present an innovative streamflow process model to be used for reservoir operational rule design in Stochastic Dual Dynamic Programming (SDDP). Model features, which can be applied independently, are: i) a non-linear multiplicative process model for the forward phase, which produces positive streamflow values only, and its linearized version for the backward phase, and ii) a non-uniform time-step, which divides the hydrological period in time-steps of different length in order to have a process with approximately constant variance. Model identification is straightforward as for additive periodic autoregressive model generally used in SDDP. We applied this model on the Senegal River for the optimal operation of Manantali reservoir, and evaluated the proposed solutions against streamflow process model currently used in the water management literature.

Author-produced version of the article published in Journal of Irrigation and Drainage Engineering-ASCE, 2017, N°143(4)
The original publication is available at http://ascelibrary.org
Doi: 10.1061/(ASCE)WR.1943-5452.00007466

1. Introduction

In reservoir operation, present benefits must be balanced with future, uncertain ones [*Soncini-Sessa et al.*, 2007; *Castelletti et al.*, 2008]. After each release decision, new information becomes available and partially reduces uncertainty. Optimal reservoir operation can be framed as a Multistage Stochastic Programming problem [*Birge and Louveaux*, 1997; *Shapiro and Andrzej*, 2003], which, for long horizon, is conveniently solved by Stochastic Dynamic Programming [*Bellman and Dreyfus*, 1966]. SDP, notwithstanding its elegance and potential, is affected by the so-called "curse of dimensionality", limiting its application to systems made of few variables [*Stedinger et al.*, 1984; *Trezos and Yeh*, 1987]. In the literature, some alternatives try to circumvent these limitations, for example by defining an optimal trajectory [*Turgeon*, 1980] or fixing the release policy family and find the parameters by evolutionary algorithm [*Nicklow et al.*, 2009; *Reed et al.*, 2013]. These solutions, even if advantageous for some aspects, have rarely been tested over large systems, i.e. made of a large number of reservoirs. Stochastic Dual Dynamic Programming [*Pereira and Pinto*, 1991] (SDDP) is an approximation of SDP that largely attenuates the curse of dimensionality. SDDP, however, requires the optimization problem to be modeled as linear, since problem linearity ensures cost-to-go function convexity.

SDDP requires identifying a linear stochastic inflow model which reproduces the streamflow process and its uncertainty. Streamflow process model identification is a critical step in dynamic programming problem setting, sometimes referred to as "curse of modelling" [*Tsitsiklis and Van Roy*, 1996; *Bertsekas and Tsitsiklis*, 1995], to stress that model identification can be problematic, hence a limitation for the methodology. SDDP applications

Author-produced version of the article published in Journal of Irrigation and Drainage Engineering-ASCE, 2017, N°143(4)
The original publication is available at http://ascelibrary.org
Doi: 10.1061/(ASCE)WR.1943-5452.00007466

generally use a standardized Periodic Autoregressive (PAR) model of lag 1 [*Tilmant et al.*, 2008; *Tilmant and Kelman*, 2007; *Tilmant et al.*, 2007, 2010, 2012; *Tilmant and Kinzelbach*, 2012; *Tilmant et al.*, 2009; *Goor et al.*, 2010; *Arjoon et al.*, 2014; *Marques and Tilmant*, 2013; *Gjelsvik et al.*, 2010], also known as Thomas-Fiering model [*Loucks*, 1992]. The drawback of this additive model is the non-neglegible probability of negative discharge values. This is to be avoided, because negative discharges have no physical sense. Existing solutions dealing with negative discharges [*Stedinger and Taylor*, 1982; *Pereira et al.*, 1984; *Bezerra et al.*, 2012] use non-linear transformations that make these models not usable in SDDP.

The monthly time-step preserves the process as markov, but it risks to underestimate the system adaptivity to changing conditions. Such a large time-step, in fact, may not take into account the adaptivity at a smaller time-step, and it can be a limitation to the analysis of system response to fast processes, such as flood, resulting in an underestimation of system capacity to react to this type of events.

Short-term system adaptation can be taken into account by a time decomposition approach [*Karamouz et al.*, 2003]. In time decomposition, long-term policies are refined by optimizations at shorter-term windows, using results from the long-term optimization as boundary conditions. Even if time decomposition increases the accuracy of performance estimation, short-term optimization can use more information than the long-term optimization supposes, leading to an underestimation of the performance value in the long term planning [*Weijs*, 2011]. Therefore, aggregating discharges at large time steps is an approximation with negative impact on performance, and time-step length must be a trade-off between i) calculation time and capacity to represent the process as markov, and

Author-produced version of the article published in Journal of Irrigation and Drainage Engineering-ASCE, 2017, N°143(4)
The original publication is available at http://ascelibrary.org
Doi: 10.1061/(ASCE)WR.1943-5452.00007466

ii) accurate representation of the relevant processes. The former requires a long time-step, the latter a short one.

Time-step length depends on both the system characteristics and the hydrological process that we intend to model. For some hydrological systems, variability is not uniform along the year, depending on local climate. For example, where a rainy season is separated from a dry one, the latter has generally less variability than the first, for drought is a relatively slow process compared to flood.

In this paper we present an innovative stochastic streamflow model, to be used within SDDP, that avoids some important limitations of existing models. This paper is structured as in the following. In Section 2 we introduce the methodology, from the original optimal control problem, until the SDDP as a way to solve a Multistage Stochastic Programming problem; Section 2.1 presents a procedure to estimate a streamflow model with a multiplicative stochastic component, and its linearized version. This model guarantees a negligible probability of negative discharge values, and its identification is straightforward. Section 2.2 exposes a procedure to identify non-uniform time-step lengths, to take advantage of this hydrological variability to have better distributed decision instants. The proposed features are independent from each other, then each of them can be applied separately. In Section 3 we test the proposed solutions for modeling the streamflow process on the Senegal river, West Africa. In Section 4 we draw the conclusions.

2. Methodology

Consider a water system composed of N_{res} reservoirs that is operated by N_{dec} discharge decisions. Discharge decisions are diversions from rivers and releases from reservoirs. A reservoir may have multiple releases (by different structures or for different users). The system is influenced by N_{scen} scenarios, such as future inflows. Equations (1) define the control problem.

$$\text{Find } \pi_t, \forall t \in \{1, \dots, H\}: \tag{1a}$$

$$\max_{\pi_t} \sum_{t=1}^{H} \mathbb{E}_{\mathbf{q}_t}\left[g_t(\mathbf{v}_t, \mathbf{r}_t, \mathbf{q}_t)\right] \tag{1b}$$

Subject to:

$$\mathbf{v}_t = \mathbf{v}_{t-1} + \Delta t \cdot (I \cdot [\mathbf{r}_t, \mathbf{q}_t] - O \cdot [\mathbf{r}_t, \mathbf{q}_t]) \tag{1c}$$

$$\mathbf{c}_t(\mathbf{v}_t, \mathbf{r}_t, \mathbf{q}_t) \leq 0 \tag{1d}$$

$$\mathbf{q}_t \sim f_{\mathbf{Q}_t} \tag{1e}$$

$$\tag{1f}$$

In Problem (1), vectors $\mathbf{v}_t \in \mathbb{R}^{N_{\text{res}}}, \mathbf{r}_t \in \mathbb{R}^{N_{\text{dec}}}, \mathbf{q}_t \in \mathbb{R}^{N_{\text{scen}}}$ represent reservoir volumes, discharge decisions, and scenarios, at instant t for stocks and in the period $[t - \Delta t, t)$ for flows; π_t is the optimal release rule, that suggests the optimal release decision $\mathbf{r}_t$ in function of the occurring scenario, i.e. the realization of $\mathbf{q}_t$. In Expression (1b), $g_t(\cdot)$ is a $\mathbb{R}^N$ to $\mathbb{R}$ function, representing the system objective at t, where $N = N_{\text{res}} + N_{\text{dec}} + N_{\text{scen}}$. Equation (1c) is the continuity equation, represented by the reservoirs mass balance. In Equation (1c), Δt is time-step length, I and O are the input and output matrix, of dimension $N_{\text{res}} \times (N_{\text{dec}} + N_{\text{scen}})$, associating at each inflow and discharge decision to its reservoir. $O(i, j)$ and $I(i, j)$ is 1 if the i variable is input or output of reservoir j, 0 elsewhere. Scenarios q_t, in Expressions (1e), are either stochastic or deterministic scenarios. Deterministic scenarios are a vector of values, stochastic scenarios are random variables distributed as $f_{\mathbf{Q}_t}(\mathbf{q}_t)$. Future inflows to the reservoirs are described as stochastic scenarios, while other

Author-produced version of the article published in Journal of Irrigation and Drainage Engineering-ASCE, 2017, N°143(4)
The original publication is available at http://ascelibrary.org
Doi: 10.1061/(ASCE)WR.1943-5452.00007466

variables, such as evaporation, for which uncertainty can be neglected, are considered as deterministic scenarios.

$\mathbf{c}_t$, in Inequality (1d), defines other constraints that apply to the system, such as physical constraints, or other legal or environmental requirements treated as constraints. For example, discharge decisions can have a physical upper limit and be limited by water availability.

Problem (1) is to be solved for an optimization horizon of H time-steps, from $t = 1$, where the initial condition, $\mathbf{v}_0$, is given. Decisions and realizations of stochastic variables come in recursive mode, therefore, at each decision step, release can be adjusted thanks to the new information on the occurring scenario. In this case the optimization problem is set as Multistage Stochastic Programming [*Shapiro and Andrzej*, 2003], as in Expression (2):

$$\max_{\mathbf{r}_1} g_1(\mathbf{v}_1, \mathbf{r}_1, \mathbf{q}_1) + \mathop{\mathbb{E}}_{q_{2,i}} \Bigg[\max_{\mathbf{r}_{2,i}} g_2(\mathbf{v}_{2,i}, \mathbf{r}_{2,i}, \mathbf{q}_{2,i}) + \mathop{\mathbb{E}}_{q_{3,i}} \Bigg[\dots \\ \dots + \mathop{\mathbb{E}}_{q_{H,i}} \Bigg[\max_{\mathbf{r}_{H,i}} g(\mathbf{v}_{H,i}, \mathbf{r}_{H,i}, \mathbf{q}_{H,i}) \Bigg] \dots \Bigg] \Bigg] \tag{2}$$

under conditions given by Equations (1c,**??**,1d), and with an initial condition $\mathbf{v}_0$. In some case, a condition on final time-step $\mathbf{c}_H(\mathbf{v}_H) \leq \mathbf{b}_H$ may be present, in the form of hard or soft constraint [*van Overloop et al.*, 2008].

By solving Problem (2), the optimization procedure finds an optimal release rule for a future horizon H. In expression (2), the release rule $\{\pi_t\}_{t=1}^{H}$ is a decision tree, $\mathbf{r}_{t,i} \forall t, \forall i$, made of multiple bifurcations at each time-steps, $t \in [1 : H]$, representing the optimal decision strategy adapted to the i realisation of the stochastic variable, $\mathbf{q}_{t,i}$. The main drawback of Multistage Stochastic Programming is its computational complexity, which increases exponentially with H. Multistage Stochastic Programming can be applied when

Author-produced version of the article published in Journal of Irrigation and Drainage Engineering-ASCE, 2017, N°143(4)
The original publication is available at http://ascelibrary.org
Doi: 10.1061/(ASCE)WR.1943-5452.00007466

H is small, for example in short term management [*Raso et al.*, 2014]. In same case the problem has been aggregated and reduced to seasonal decisions [*Seifi and Hipel*, 2001]. For long-term optimization, however, MSP can be considered as a theoretical, rather than practical, method [*Mayne et al.*, 2000].

Stochastic Dynamic Programming (SDP) decomposes the MSP problem in step-by-step optimal decision problems. Then, the optimization problem in Expression (2) can be written as Bellman Chain, as in Equation (3). SDP can solve problems with a much longer horizon, because problem complexity increases only linearly with H.

$$F_t(\mathbf{v}_t, \mathbf{q}_t) = \max_{\mathbf{r}_t} g_t(\mathbf{v}_t, \mathbf{r}_t, \mathbf{q}_t) + \mathop{\mathbb{E}}_{\mathbf{q}_{t+1}} \left[F_{t+1}(\mathbf{v}_{t+1}, \mathbf{q}_{t+1}) \right] \tag{3}$$

In Equation (3), F_t is the cost-to-go function, the average cost for leaving the system in the state $[\mathbf{v}_t, \mathbf{q}_t]$, which is the compromise of present and future benefits, $g_t(\cdot)$ and F_{t+1}. In SDP, Equation (3) is the release rule π_t, which maps the system state to the optimal release decisions.

Equation (3) can be solved backwards, from $t = H$ to the initial time-step. Condition (1e) is substituted by Condition (4). The probability transition from $\mathbf{q}_{t-1}$ to $\mathbf{q}_t$, together with the continuity Equation (1c), makes up the transition equations, which describes the system dynamic from one state to the next one.

SDP requires the stochastic transition to be expressed as a Markov process [*Rabiner and Juang*, 1986], i.e. the probability of each event, $f(\mathbf{q}_t)$, depends only on the state at previous instant, $\mathbf{q}_{t-1}$, as in Equation (4).

$$\mathbf{q}_t \sim f_{Q_t}(\mathbf{q}_t | \mathbf{q}_{t-1}) \tag{4}$$

Author-produced version of the article published in Journal of Irrigation and Drainage Engineering-ASCE, 2017, N°143(4)
The original publication is available at http://ascelibrary.org
Doi: 10.1061/(ASCE)WR.1943-5452.00007466

When an autoregressive lag 1 process is not sufficient, $\mathbf{q}_{t-1}$ can be enlarged to contain all the informative variables [*Turgeon*, 1980]. Then, the process 'memory" $\mathbf{q}_{t-1}$ includes $[\mathbf{q}_{t-1}, \ldots, \mathbf{q}_{t-p}, \mathbf{e}_{t-1}, \ldots, \mathbf{e}_{t-q}]$, where $\mathbf{e}_{t-j}$ is the difference between the observed value and the expected value of the model forecast at $\mathbf{q}_{t-j}$. In SDP, this state variables augmentation allows to represent the stochastic process by an Periodic ARMA(p, q) model. However, hydrological processes representation with more than one variable has rarely been applied, and SDP applications are often limited to strategic reservoir operation. In fact, SDP suffers from the so-called "curse of dimensionality,, i.e. complexity increases exponentially with the number of system variables.

Stochastic Dual Dynamic Programming (SDDP) [*Shapiro*, 2011] is an approximation of the original SDP problem. SDDP attenuates the curse of dimensionality, and can be applied to larger systems. SDDP does not require variables discretization, but the time-step optimization problem must be linear: transition equations (1c, 4) and objective (1b) must be linear, and equations defining the constraints space (1d) must be affine. SDDP is solved iteratively forward and backward. In the backward stage, at each $t = H, \ldots, 2$, the optimization finds the minimum average cost to pass from $[\mathbf{v}_{t-1}, \mathbf{q}_{t-1}]$ to $[\mathbf{v}_t, \mathbf{q}_t]$, adding an extra cut $l_k(\mathbf{v}_t, \mathbf{q}_t)$ to the approximation of the cost-to-go function $\mathcal{F}_t(\mathbf{v}_t, \mathbf{q}_t)$, such that $\mathcal{F}_t(\cdot) := \max\{\mathcal{F}_t(\cdot), l_k(\cdot)\}$. In the forward stage, the approximate problem is solved from $t = 1$ to H to find the optimal trajectories to be used in the next backward phase. By successive iterations, $\mathcal{F}_t$ converges to the real cost-to-go function F_t, as demonstrated, under mild conditions, by *Philpott and Guan* [2008] and *Linowsky and Philpott* [2005].

2.1. Multiplicative error model identification

Author-produced version of the article published in Journal of Irrigation and Drainage Engineering-ASCE, 2017, N°143(4)
The original publication is available at http://ascelibrary.org
Doi: 10.1061/(ASCE)WR.1943-5452.00007466

In this section we propose a multiplicative error model to deal with the problem of generating negative discharges.

In SDDP, the stochastic hydrological model must be linear, but no conditions are imposed on distribution $f_t(\mathbf{q}_t)$. There is not specific reason to prefer separate additive terms for the predictive and the uncertain parts of the model [*Koutsoyiannis*, 2009]. In the following, we show a procedure to estimate a linear model in $\mathbf{q}_t$, with a multiplicative stochastic component.

We start from a multivariate signal of observed discharge, which we want to reproduce, $\mathbf{q}_t$, made of T samples for N_y years of data, where T is the number of time-steps per year. The original signal is transformed into $\mathbf{y}_t$ according to Equation (5).

$$\mathbf{y}_t = \log(\mathbf{q}_t) - \log(\overline{\mathbf{q}}_\tau) \tag{5}$$

In Equation (5), $\tau = 1, \ldots, T$ is the periodic time index. Equation (6) defines $\overline{\mathbf{q}}_\tau$, i.e. the periodic geometric average of $\mathbf{q}_t$ for all $\mathbf{t} = \tau$, where $\mathbf{t} = (t-1) \bmod T + 1$, and mod is the modulus (or remainder) operator.

$$\overline{\mathbf{q}}_\tau = \left(\prod_{\mathbf{t}}^{N_y} \mathbf{q}_t\right)^{1/N_y} = \exp\left(\frac{1}{N_y}\sum_{\mathbf{t}}^{N_y} \log \mathbf{q}_t\right), \; \forall \mathbf{t} = \tau \tag{6}$$

Logarithm smooths extremes and long tails, making model identification easier. In fact, logarithm is a particular case of the Box-Cox transformation, suggested in *Box et al.* [1970] to deal with non-normal residuals. We use the $\mathbf{y}_t$ signal to identify an ARMA model on the N_{stoch} dimensional signal. Contemporaneous ARMA (CARMA) [*Camacho et al.*, 1987] are an effective sub-class of multivariate ARMA model that proved effective in hydrological applications. A CARMA model is identified as a set of N_{stoch} univariate

Author-produced version of the article published in Journal of Irrigation and Drainage Engineering-ASCE, 2017, N°143(4)
The original publication is available at http://ascelibrary.org
Doi: 10.1061/(ASCE)WR.1943-5452.00007466

ARMA models, with correlated residuals. For simplicity of explanation and notation, we describe the model identification procedure for an univariate ARMA model, recalling that possible correlation among sub-catchments can be included in the residuals covariance matrix [*Salas et al.*, 1985].

In hydrological systems, parameters are generally time-variant [*Salas*, 1980; *Hipel and McLeod*, 1994], because of climatic effect and because different hydrological processes are dominant at different periods in the years. Time-variant parameters can effectively dealt with Periodic ARMA model (PARMA). Equation (7) defines the $PARMA_\tau(p,q)$ model.

$$y_t = \sum_{i=1}^{p} \phi_{\tau,i} y_{t-i} + \sum_{j=1}^{q} \psi_{\tau,i} \varepsilon_{t-i} + \varepsilon_t \tag{7}$$

In Equation (7), ε_t is the stochastic process, extracted from $\mathcal{N}(0, \sigma_\tau^2)$, independent on previous ε_{t-i}, where $\varepsilon_{t-i} = y_{t-i} - \hat{y}_{t-i}$; y_t is observed value, and $\hat{y}_t$ is expected value of Model (7), i.e. when $\varepsilon_t = 0$. Identifying a $PARMA_\tau(p,q)$ model is defining parameters $\phi_{\tau,i}, \psi_{\tau,i}, \sigma_\tau$ of Equation (7) $\forall i, \forall j, \forall \tau$.

Equation (7) can be written as transition from q_{t-i} to q_t. By inversion of Equation (5) and some rearrangement, one can obtain Equation (8).

$$q_t = \alpha_\tau \cdot \prod_{i=1}^{p} {q_{t-i}}^{\phi_{\tau,i}} \cdot \prod_{j=1}^{q} {\xi_{t-i}}^{\psi_{\tau,j}} \cdot \xi_t \tag{8}$$

where $\xi_t \sim \ln \mathcal{N}(0, \sigma_\tau)$, $\alpha_\tau = \overline{q}_\tau / \prod_{i=1}^{p} \overline{q}_{\tau-i}^{\phi_{\tau,i}}$, with $\overline{q}_\tau$ defined as in Equation (6), and $\xi_{t-j} = \frac{q_{t-j}}{\hat{q}_{t-j}}$.

If in model (7) residuals are normal and additive, in model (8) residuals become log-normal and multiplicative. For positive initial condition, $q_0 > 0$, multiplicative random process ensures non-negative values of inflow process, offering a better representation of the hydrological process.

In Model (8), dependencies on q_{t-i} are non-linear. Model (8) can be directly employed in the forward phase of SDDP, also using a less parsimonious model than in backwards, which can produce, in some cases, considerably better results [*Bartolini and Salas*, 1993]. SDDP forward phase requires in fact a streamflow time series, regardless of the method to generate it. In backward phase, instead, SDDP must solve a linear optimization, therefore transition from q_{t-1} to q_t must be linear. Model (8), to be applied in backward phase of SDDP, must be written as in Equation (9), which is Model (8) linearized by taylor expansion on the argument median, i.e. $q_{t-i} = \overline{q}_{\tau-i}$ and $\xi_{\tau-j} = 1$, $\forall\tau, \forall i, \forall j$.

$$q_t = \left[\sum_{i=1}^{p} \rho_{\tau,i} q_{t-i} + \sum_{j=1}^{q} \omega_{\tau,i} \xi_{t-i} + \kappa_\tau\right] \cdot \xi_\tau \tag{9}$$

where parameters are defined in Equations (10) and derived in Appendix A.

$$\rho_{\tau,i} = \phi_{\tau,i} \cdot \frac{\overline{q}_\tau}{\overline{q}_{\tau-i}} \tag{10a}$$

$$\omega_{\tau,i} = \psi_{\tau,i} \cdot \overline{q}_\tau \tag{10b}$$

$$\kappa_\tau = \overline{q}_\tau \cdot \left(1 - \sum_{i=1}^{p} \phi_{\tau,i} - \sum_{j=1}^{q} \psi_{\tau,i}\right) \tag{10c}$$

Linearisation introduces an approximation error that must be quantified. Equation (11) defines e_t, the error due to linearisation.

$$e_t = q_t^{\mathrm{lin}} - q_t^{\mathrm{nl}} \tag{11}$$

where q_t^{lin} and q_t^{nl} are the output of Model (9) and Model (8) for $\xi_t = 1$. Considering, for simplicity of notation, an univariate Multiplicative Periodic Autoregressive Model,

Author-produced version of the article published in Journal of Irrigation and Drainage Engineering-ASCE, 2017, N°143(4)
The original publication is available at http://ascelibrary.org
Doi: 10.1061/(ASCE)WR.1943-5452.00007466

$PAR\tau(1)$, then e_t can be written as function of q_{t-1} only. Knowing the distribution $f_\tau(q_{t-1})$ we can estimate that of e_t. Specifically, we are interested in the average and extreme quantiles of e_t, defined as in Equations (12).

$$\mathbb{E}(e_\tau) = \int_0^{+\infty} e_t(q_{t-1}) \cdot f_\tau(q_{t-1}) \cdot dq_{t-1} \tag{12a}$$

$$q_p(e_\tau) = e_t\left(F^{-1}_{q_{\tau-1}}(p)\right) \tag{12b}$$

Where $q_p(e_\tau)$, is the p quantile of e_τ, and $F_{q_{\tau-1}}(\cdot)$ is the Cumulative Density Function of $q_{\tau-1}$. Distribution of $q_{\tau-1}$ can be obtained from $f(y_{\tau-1})$ through Equation (5). From *Vecchia* [1985] *Bartolini et al.* [1988], we know that $f(y_{\tau-1})$ is normal with known average and variance. Therefore $f(q_{\tau-1})$ is lognormal, with parameters derived from $f(y_{\tau-1})$ average and standard deviation.

$\mathbb{E}(e_\tau)$, from Equation (12a), can be used to correct results of linear Model (9) by shifting the output value in order to have zero bias. In this case, $-\mathbb{E}(e_\tau)$ is to be added in Equation (10c).

2.2. Non-uniform time-step length

In this section we propose a non-uniform discharge aggregation, which modulates the time-step length to have a fine discharge representation only when needed.

Hydrological models in SDDP have so far always used a fixed time-step and periodically variable parameters. Predictive uncertainty, changing along the period, is included in the process model by considering heteroscedastic residuals, i.e. residuals with different variability. We consider here, instead, a variable time-step that divides the hydrological

Author-produced version of the article published in Journal of Irrigation and Drainage Engineering-ASCE, 2017, N°143(4)
The original publication is available at http://ascelibrary.org
Doi: 10.1061/(ASCE)WR.1943-5452.00007466

period in time-steps of different length in order to maintain an approximately homogeneous variance.

In the following, we describe the proposed procedure to select the non-uniform time-step length, $\Delta k(\tau)$, to aggregate data from the finest time-step, $\tau_d = 1, \dots, T_d$ to the desired level of aggregation, $\tau = 1, \dots, T$, such that the aggregated time-step is $\Delta\tau = \Delta\tau_d \cdot \Delta k$.

Starting from the data at the finest time aggregation of period T_d and time index $\tau_d = 1, \dots, T_d$, we identify a $\mathrm{PAR}_{T_d}(1)$ model on q_t, having parameters $\phi_{\tau_d}, \sigma^2_{\tau_d}, \forall \tau_d \in \{1, \dots, T_d\}$, defined as in Equation (7). Even if such simple model may be not accurate enough for prediction purposes, it is generally sufficient to catch the dominant dynamics.

Time-steps of homogeneous variability, $\Delta k(\tau)$, must be such that $VAR(q_{\tau 0+\Delta k(\tau)}|q_{\tau 0})$ is approximately homogeneous for all τ, and $\sum \Delta k(\tau) = T_d$. Equation (13) defines variance of $q_{\tau 0+\Delta k(\tau)}$ conditional to q_{τ_0} as function of time-step residual variance, $\sigma^2_{\tau_d}$.

$$VAR(q_{\tau 0+\Delta k(\tau)}|q_{\tau 0}) = \sum_{\tau_d=\tau_0+1}^{\tau_0+\Delta k(\tau)} \prod_{i=\tau_d+1}^{\tau_0+\Delta k(\tau)} \phi_i^2 \cdot \sigma^2_{\tau_d} \tag{13}$$

where, by convention, $\phi_i^2 = 1$ if $i > \tau_0 + \Delta k(\tau)$.

In Equation (13), ϕ_{τ_d} is generally close to one, especially for τ_d where variability is small, and time-steps Δk can be larger. Considering $\prod \phi^2_{\tau_d}$ in Equation (13) as equal to one, $VAR(q_{\tau 0+\Delta k_i}|q_{\tau 0})$ can be written as proportional to the sum of variances only, which we take as indicator of variability. This allows us to define a cumulated variability that depends on τ_d only.

The residual variance is used to define a Cumulative Variability in function of τ, as in Equation (14).

Author-produced version of the article published in Journal of Irrigation and Drainage Engineering-ASCE, 2017, N°143(4)
The original publication is available at http://ascelibrary.org
Doi: 10.1061/(ASCE)WR.1943-5452.00007466

$$CV(\tau_d) = \frac{\sum_{t=1}^{\tau_d} \sigma_{\tau_d}^2}{\sum_{t=1}^{T_d} \sigma_{\tau_d}^2} \quad (14)$$

In Equation (14), the numerator is the cumulative variance until τ_d, the denominator is the cumulated variability for the entire period, T_d, to standardize CV between zero and one. The non-uniform time-step $\Delta k(\tau)$ is then chosen by splitting the hydrological period in T time-steps having approximately homogenous variability, as in Equation (15).

$$\Delta k(\tau) = \left\lfloor CV^{-1}\left(\frac{\tau}{T}\right) - CV^{-1}\left(\frac{\tau-1}{T}\right)\right\rceil;\ \forall\tau = \{1, \cdots, T\} \quad (15)$$

where operator $\lfloor\cdot\rceil$ returns the nearest integer. By convention, $CV^{-1}(0) = 0$.

The discharge signal is aggregated using the variable time-step, and a model is identified on the aggregated signal according to procedure in Section 2.1, or others.

3. Application to the Senegal river

The Senegal River, West Africa, is a 1790 km long river. Its drainage basin extension is 270.000 km^2, over Guinea, Mali, Senegal and Mauritania. The river inflow is extremely variable, following the tropical raining seasonality with a marked difference between the dry season, in January-June, and the raining one, in July-October, when most of the water falls in the upper part of the basin [*Albergel et al.*, 1997]. Figure 2 displays the discharge at Soukoutali, inflow to Manantali for 64 years.

Manantali is an annual reservoir for hydropower, located in Mali and controlling about 50% of the total water flow of the Senegal River. Manantali was completed in 1987 and started to produce electricity in 2002. Its benefits are shared among Mali, Mauritania, and Senegal. These three countries participate with Guinea to the Organization for the

Author-produced version of the article published in Journal of Irrigation and Drainage Engineering-ASCE, 2017, N°143(4)
The original publication is available at http://ascelibrary.org
Doi: 10.1061/(ASCE)WR.1943-5452.00007466

Valorization of Senegal River (Organisation pour la Mise en Valeur du fleuve Senegal, OMVS). Manantali reservoir volume is $12 \times 10^9 m^3$, its installed capacity is 205 MW, the average inflow is 270 m^3/s, the average residence time is about one year.

The OMVS have gradually embraced the Integrated Water Resources Management paradigm, in which water allocation decisions are based on economic, social, technical and political factors, in accordance with stakeholders' interests. This led the OMVS to prepare a reservoir management optimization program [*Fraval et al.*, 2002]. Manantali operation was originally designed to satisfy different rival uses: energy production and low-flow augmentation on one hand; flood support on the other. Operating the reservoir for hydroelectric production reduces the annual streamflow variability, with negative effects on the ecological equlilbrium and on some traditional activities in the valley, but with positive effects on water availability for irrigation and navigation, which will become a more important objective in the near future [*Bader et al.*, 2003]. This analysis, however, focuses on methodological aspects, so we consider energy production objective only.

In the following, at section 3.1, we define the non-uniform time-step, then we identify a streamflow process model as defined in Equations (8) and (9), and a classic Thomas-Fiering Model, for comparison. At section 3.2, we test the added value of non-uniform time-step in terms of reservoir operation.

3.1. Streamflow process model identification

The hydrological process on the Senegal river is characterised by a strong periodical component. Following the procedure described in section 2.2, we select $\Delta k(\tau)$ using $T = 12$ time-steps, for comparison with monthly time-step. Figure 3 shows the $CV(\tau_d)$ function, defined in Equation (14), and the $\Delta k(\tau)$ for $T = 12$, for the entire periodic year,

Author-produced version of the article published in Journal of Irrigation and Drainage Engineering-ASCE, 2017, N°143(4)
The original publication is available at http://ascelibrary.org
Doi: 10.1061/(ASCE)WR.1943-5452.00007466

on plot (a), and a zoom on the period of high variability, on plot (b), for τ_d between 135 (May 15) and 285 (October 12).

The non-uniform time-step defined by $\Delta k(\tau)$ is used to aggregate the daily inflow signal. The same inflow is also aggregated at the monthly time-step to be compared to the non-uniform aggregation. Figure 4 shows the aggregated observed inflow with non-uniform time-step and monthly time-step.

Non-uniform aggregation allocates only two time-steps for the entire recession curve, going from October to May,using the ten remaining time-steps for the raising part of the hydrography, against the seven and five time-steps used by monthly aggregation. A finer time-step during the raising part of the hydrograph should allow to better adapt to the incoming information on the inflow value.

Figure 5 shows the autocorrelation lag 1 on the de-trended logarithmic of discharge signal, equivalent to ϕ_τ of Model (7) for a PAR(1), being $\phi_\tau = \frac{COV(y_t,y_{t-1})}{VAR(y_{t-1})}$ [*Box et al.*, 1970]. In monthly aggregation, ϕ_τ stays close to 1 for different months during the dry period, reaching 0.5 at $\tau = 8$. The non-uniform aggregation autocorrelation is more regular: most of the cases lay between 0.6 and 0.8. This results suggest that the non-uniform aggregation can provide a more effective distribution of decision instants. This is further investigated in section 3.2.

From the non-uniform aggregated inflow signals, following the procedure described in 2.1, we identify a multiplicative model as in Equations (9) and (8) and an additive Thomas-Fiering model. We refer to these models as multiplicative and additive model. A $PAR(1)$ represents the process sufficiently well. The residuals autocorrelation is approximately zero for both the multiplicative and the additive model, being within the 95% confidence

Author-produced version of the article published in Journal of Irrigation and Drainage Engineering-ASCE, 2017, N°143(4)
The original publication is available at http://ascelibrary.org
Doi: 10.1061/(ASCE)WR.1943-5452.00007466

band, for most of the lags larger than zero. This confirms the validity of the non-uniform aggregation, for it preserves the process as markov.

We test whether the residual sampling distribution adhere to the prescribed one, log-normal for the multiplicative case, normal for the additive. Both residuals must be standardised by the time variant variance, σ_τ. For the multiplicative residuals, we test their normality on $\log(\Xi_t)$. Figure 6 presents the normality plot for both the logarithm of residuals for the multiplicative model and the residuals for the additive model, showing that the multiplicative model follows the prescribed distribution, whereas the additive model suffers from a fat tail on higher values. A Kolmogorov-Smirnov test [*Dekking*, 2005] on the standardized logarithms of residuals does not reject the hypothesis of a normal distribution, with a p-value of 0.07. The same test on the additive model standardized residual gives a p-value of 10^{-6}, leading to rejecting the assumption of residuals being normally distributed with sufficient confidence. The positive results on the (ξ_t) distribution correctness gives us further confidence in the validity of the multiplicative model.

Figure 7 shows the observed discharge data and their match with the 95% confidence bands for both the multiplicative and the additive model. The confidence bands are $\overline{q}_\tau \cdot \exp\big(\pm 2 \cdot VAR(y_t)\big)$ for the multiplicative model (bold continuous lines) and $\mathbb{E}(q_t) \pm 2 \cdot VAR(q_t)$ for the additive model (bold dashed lines). The signal variance $VAR(q_t)$ or $VAR(y_t)$, is derived from $\{\sigma_\tau\}_{\tau=1}^{T}$ as described in *Bartolini et al.* [1988].From Figure 7 we see how the additive model has a non-negligible probability of producing negative inflow. Moreover, lower confidence band for the additive model is, for some time-steps, much lower than the lowest observed discharge. whereas the multiplicative model follows closely the observed signal variance for the entire period.

Author-produced version of the article published in Journal of Irrigation and Drainage Engineering-ASCE, 2017, N°143(4)
The original publication is available at http://ascelibrary.org
Doi: 10.1061/(ASCE)WR.1943-5452.00007466

We quantify the approximation due to linearization using Equations (12) for error average, and 5% and 95% quantile of $q_{\tau-1}$, examining both the average for all τ and the largest value in τ. The average error is 0.6 m^3/s, or 0.3% of median discharge, which we consider a relatively small value. Even for the largest error average, at $\tau = 7$, is 23 m^3/s, which is less than 3% of median discharge. Error at 5% quantile is, on average, 4% of median discharge, with a peak of 7% at $\tau = 4$ (15 m^3/s in absolute value). Error at 95% quantile is 11% of median discharge, with a maximum absolute value of 88 m^3/s at $\tau = 7$, which is 11% of median discharge. As a summary, we can state that error due to linearization is adequatly small; its average is generally negligible, growing to about 5% and 10% of median discharge at 5% and 95% quantiles of $q_{\tau-1}$.

Figure 8 shows the detail of the error due to linearization for $\tau = 7$. Plot (a) compares output of nonlinear and linearized multiplicative model, i.e. model (8) and model (9) for $\xi_t = 1$, in function of q_{t-1}. The linear model is tangent to the nonlinear one at the linearisation point, i.e. at $\bar{q}_{\tau-1}$. Going further from the median the models diverge, even if the difference stays small. Plot (b) presents the error magnitude in function of q_{t-1} next to its probability of occurrence, $f(q_{\tau-1})$.

On a multiplicative $PAR_\tau(1)$, $\phi_\tau < 1$ for most of τ, because stationarity condition requires that $\prod_{\tau=1}^{T} \phi_\tau < 1$. This implies that model (8) is a concave function in q_{t-1} and $e_t > 0, \forall q_{t-1}$, i.e. Model (9) systematically overestimate its nonlinear version. A large number of time-steps per period results in a shorter time-steps length. As a consequence, ϕ_τ values will be closer to one, for ϕ_τ is the autocorrelation lag 1 in a AR model [*Box et al.*, 1970], resulting in a smaller error due to linearization.

3.2. Effects of non-uniform aggregation on reservoir operation

We compare the system performances and behaviour of reservoir operation between monthly and non-uniform discharge aggregation to assess the advantage of the latter.

The system model is made of a reservoir and a hydrological component, as in Figure 9. Streamflow at Soukoutali, output of the hydrological model, is the input to the reservoir. The objective is energy production. Each element of the optimization problem is detailed in the following.

The reservoir is modelled as in Equation (1c). The reservoir input is the discharge from the hydrological component q_t. Reservoir outputs are: discharge through turbines $r_{\text{turb},t}$, and discharge through spillways, $r_{\text{spill},t}$. Evaporation is not considered in this analysis.

Inequalities (16) are constraints on discharge decisions and reservoir volume.

$$v_{\min} \leq v_t \leq v_{\max} \tag{16a}$$

$$0 \leq r_{\text{turb},t} \leq r_{\text{turb,max}} \tag{16b}$$

$$r_{\text{spill},t} \geq m_{\text{safety}} \cdot (v_t - v_{\text{safety}}) \tag{16c}$$

Inequalities (16a) and (16b) are physical constraints, derived from the system characteristics. Inequality (16c) is a legal condition that forces to draw down the reservoir when its volume exceeds the safety threshold, v_{safety}. Constraint on release through spillages is large enough for never being active during the simulation period, as verified a posteriori, and therefore it is not included. Inequalities (16b) and (16c) are implemented as hard constraint, inequality (16a) is implemented as soft constraint to avoid non-feasibility.

The system objective is the yearly average energy production $J_E = 1/N_{years} \sum_{t=1}^{T \times N_{years}} E_t$, composed of the sum of the daily energy production, E_t, as defined in Equation (17), for the entire simulation horizon $T \times N_{years}$, where $N_{years} = 43$, from 1970 to 2012.

Author-produced version of the article published in Journal of Irrigation and Drainage Engineering-ASCE, 2017, N°143(4)
The original publication is available at http://ascelibrary.org
Doi: 10.1061/(ASCE)WR.1943-5452.00007466

$$E_t = \eta \cdot \Delta k(\tau) \cdot \Delta h_t \cdot r_{\mathrm{turb},t} \tag{17}$$

In Equation (17), Δh_t is the hydraulic head $[m]$, $r_{\mathrm{turb},t}$ is discharge trough turbines $[m^3/s]$. η is a multiplicative factor, such that $\eta = \rho \cdot g \cdot \widehat{\eta}(r_t, \Delta h) \cdot 24 \cdot 10^{-6}$, where ρ is the water density, 1000 $[kg/m3]$, g is the gravity acceleration, 9.8 $[m/s^2]$, $\widehat{\eta}$ is the efficiency coefficient, considered equal to 0.9, 24 and 10^{-6} are unit transformation coefficients, $[h/d]$ and $[MW/W]$. E_t is expressed in $[MWh]$.

Equation (17), to be employed in linear optimization within SDDP, is approximated by expressing it as linear function of r_t and v_t, linearized at an operational point, under the hypothesis of cylindric reservoir. Equation (18) defines the operational time-step objective indicator used in SDDP, being the weighted sum of releases and volume.

$$\begin{aligned} E_t^{\mathrm{op}} &= E_{0,t} + \\ + \eta \cdot \Delta k(\tau) \cdot \Bigg\{ &+ \left[\frac{r_{\mathrm{turb},0}}{A_0}\right] \cdot v_t \\ &+ [h_0 - h_0^v - m_v(1 + r_{\mathrm{turb},0})] \cdot r_{\mathrm{turb},t} \\ &+ [-m_v] \cdot r_{\mathrm{spill},t} \Bigg\} \end{aligned} \tag{18}$$

In Equation (18), $E_{0,t} = \eta \cdot \Delta k(\tau) \cdot \Big\{ [h_0 - h_0^v] \cdot r_{\mathrm{turb},0} - \left[\frac{r_{\mathrm{turb},0}}{A_0}\right] \cdot v_0 - [h_0 - h_0^v - m_v(1 + r_{\mathrm{turb},0})] \cdot r_{\mathrm{turb},0} + m_v \cdot r_{\mathrm{spill},0} \Big\}$. Parameters of Equation (18) and their derivation are described in Appendix B.

Equation (17) is linearized at a normal operational point, that is the reservoir state at which the reservoir is mostly operated, either historically observed or deduced from system characteristics. We infer the operational point from the reservoir characteristics,

Author-produced version of the article published in Journal of Irrigation and Drainage Engineering-ASCE, 2017, N°143(4)
The original publication is available at http://ascelibrary.org
Doi: 10.1061/(ASCE)WR.1943-5452.00007466

considering $v_0 = v_{\text{safety}}$, i.e. the safety limit, $r_{turb,0} = \mathbb{E}(q_t)$, i.e. the average inflow to the reservoir, and $r_{spill,0} = 0\ m^3/s$, which considers no release through the spillages.

The optimization is performed using 25 extractions for the forward phase and 25 for the backward one until convergence, attained at an accuracy level of 2×10^7. This accuracy level lays within the $\pm 2\sigma$ of forward simulation results. We can consider with sufficient confidence that the algorithm has converged to the optimum.

Performance results show a moderate improvement for non-uniform aggregation. Optimal solution gives J_E equal to 930 GWh/year for the monthly aggregation, and 945 GWh/year in the non-uniform case, equivalent to an improvement of one week of average energy production.

We analyse the reservoir operation behaviour. Figure 10 shows the reservoir volume and the release trough turbines for the monthly and the non-uniform aggregation in response to the 2005 inflow scenario. Inflow peak is larger than maximum discharge trough turbines; therefore, to avoid spillage, the reservoir must be drawn down before the high flow period, in order to create a buffer that stores part of the incoming water. SDDP optimal operating rules are the results of an optimal compromise between the objectives of keeping a high water level and avoiding spillages.

Figure 10 plot (b) shows how the reservoir operation using non-uniform steps adjusts decisions at higher frequency during high uncertainty periods, adapting more rapidly to the new observed discharge. Thanks to this rapid adaptation during the high discharge period, reservoir operation using non-uniform aggregation can draw down the reservoir less, as shown in Figure 10 plot (a).

Author-produced version of the article published in Journal of Irrigation and Drainage Engineering-ASCE, 2017, N°143(4)
The original publication is available at http://ascelibrary.org
Doi: 10.1061/(ASCE)WR.1943-5452.00007466

4. Conclusions and Discussion

SDDP, to be employed for reservoir operational rules design, requires the identification of a linear streamflow process model. Presently, models from the literature use almost always a periodic autoregressive model with monthly time-steps. In this study we proposed an innovative streamflow process model to be used in SDDP. Model features are i) a log-normal multiplicative stochastic component, which guarantees positive discharge values, and ii) non-uniform time-steps, which makes the process approximately homoscedastic i.e. having constant variability. The multiplicative non-linear model can be employed in the SDDP forward phase directly, whereas a linearized version must be used in the backward phase. We showed how to identify the multiplicative streamflow process model and its linearized version, and how to derive the non-uniform time-steps lengths from discharge data. Model identification for the proposed model is not more complex than for classic periodic autoregressive models with monthly time-step. The proposed features are independent from each other, then each of them can be applied separately. This work address specific problems encountered in SDDP, but some results may have a broader (potential) validity in time series modeling for synthetic streamflow sequences generation.

We applied the model to the streamflow process at Soukoutali, on the Senegal River, for the operational rules design of a single reservoir system. Model identification would not be different if the reservoirs were many. The proposed multiplicative model offers a better representation of the streamflow process both in the forward phase, where it correctly represents the streamflow dynamics and the discharge distribution, and in the backward phase, where it correctly represents the residual distribution, avoiding the fat-tail phenomenon, otherwise present in the classic Thomas-Fiering model. The model

Author-produced version of the article published in Journal of Irrigation and Drainage Engineering-ASCE, 2017, N°143(4)
The original publication is available at http://ascelibrary.org
Doi: 10.1061/(ASCE)WR.1943-5452.00007466

using non-uniform time-steps has a relatively homogeneous variance. This brings in a practical advantage: the non-uniform time-steps follow closer the changing hydrological variability along the year, adapting the decision more frequently during high variability periods, resulting in enhanced system performance evaluation. If time-step aggregation is sufficiently fine, a non-uniform aggregation may even make time decomposition needless.

Model linearization, used in the backward phase, introduces an error. The analysis on error due to linearisation show that the error average is negligible, growing to about 5%-10% of median discharge at 5%-95% quantiles of $q_{\tau-1}$, which we consider satisfactory; this depends, however, on the specific test-case and, a priori, we cannot exclude it to be a limitation for this model.

Appendix A: Linear model parameters derivation

Parameters $\rho_{\tau,i}, \omega_{\tau,i}, \kappa_\tau$ are derived by linerization of Model (8) on the median of its deterministic inputs, at $q_{t-i} = \bar{q}_{\tau-i}$, with $\bar{q}_\tau$ defined in Equation (6), and $\xi_{t-j} = 1$.

Equation A1 is model (8) for $\xi = 1$ written as Taylor expansion on its deterministic inputs.

$$\begin{aligned} q_t^{\text{nl}} \approx & \sum_{i=1}^{p} \frac{\partial q_t^{\text{nl}}}{\partial q_{t-i}} \cdot (q_{t-i}^{\text{nl}} - \bar{q}_{t-i}) + \\ & \sum_{j=1}^{q} \frac{\partial q_t^{\text{nl}}}{\partial \xi_{t-j}} (\xi_{t-j} - \bar{\xi}_{t-j}) + \\ & q_t^{\text{nl}}(q_{t-i}, \xi_{t-j}) \end{aligned} \tag{A1}$$

Equations (A2) are the derivatives of q_t on inputs q_{t-i} and ξ_{t-j}.

$$\frac{\partial q_t^{\text{nl}}}{\partial q_{t-i}} = \alpha_\tau \cdot \left[\phi_{\tau,i} \cdot {q_{t-i}}^{(\phi_{\tau,i}-1)} \cdot \prod_{k\in\{1,\dots,p\}\setminus i} {q_{t-k}}^{\phi_{\tau,k}} \cdot \prod_{j=1}^{q} {\xi_{t-j}}^{\psi_{\tau,i}} \right] \tag{A2a}$$

$$\frac{\partial q_t^{\text{nl}}}{\partial \xi_{t-j}} = \alpha_\tau \cdot \left[\psi \cdot {\xi_{t-j}}^{(\psi_{\tau,j}-1)} \cdot \prod_{i}^{p} {q_{t-i}}^{\phi_{\tau,i}} \cdot \prod_{k\in\{1,\dots,q\}\setminus j} {\xi_{t-k}}^{\psi_{\tau,k}} \right] \tag{A2b}$$

Separating members of Equation (A1) by their inputs, and considering that $\alpha_\tau = \bar{q}_\tau / \prod_{i=1}^{p} \bar{q}_{\tau-i}^{\phi_{\tau,i}}$ we derive, in Equations (A3), parameters (10) of Model (9).

$$\begin{aligned} \rho_{\tau,i} =& \frac{\partial q_t^{\text{nl}}}{\partial q_{t-i}}(\overline{q}_{t-i}, \overline{\xi}_{t-j}) = \\ =& \frac{\bar{q}_\tau}{\prod_{i=1}^{p} \bar{q}_{\tau-i}^{\phi_{\tau,i}}} \cdot \left[\phi_{\tau,i} \cdot \frac{\bar{q}_{t-i}^{\phi_{\tau,i}}}{\bar{q}_{t-i}} \cdot \prod_{k\in\{1,\dots,p\}\setminus i} \bar{q}_{t-k}^{\phi_{\tau,k}} \right] = \\ =& \phi_{\tau,i} \cdot \frac{\overline{q}_\tau}{\overline{q}_{\tau-i}} \end{aligned} \tag{A3a}$$

$$\begin{aligned} \omega_{\tau,j} =& \frac{\partial q_t^{\text{nl}}}{\partial \xi_{t-j}}(\overline{q}_{t-i}, \overline{\xi}_{t-j}) = \\ =& \frac{\bar{q}_\tau}{\prod_{i=1}^{p} \bar{q}_{\tau-i}^{\phi_{\tau,i}}} \cdot \left[\psi \cdot \prod_{i}^{p} {\overline{q}_{t-i}}^{\phi_{\tau,i}} \right] = \\ =& \psi_{\tau,i} \cdot \overline{q}_\tau \end{aligned} \tag{A3b}$$

$$\begin{aligned} \kappa_\tau =& q_t^{\text{nl}}(\overline{q}_{t-i}, \overline{\xi}_{t-j}) - \sum_{i=1}^{p} \frac{\partial q_t^{\text{nl}}}{\partial q_{t-i}}(\overline{q}_{t-i}, \overline{\xi}_{t-j}) \cdot \overline{q}_{t-i} - \sum_{j=1}^{q} \frac{\partial q_t^{\text{nl}}}{\partial \xi_{t-j}}(\overline{q}_{t-i}, \overline{\xi}_{t-j}) \cdot \overline{\xi}_{t-j} = \\ =& \overline{q}_\tau \cdot \left(1 - \sum_{i=1}^{p} \phi_{\tau,i} - \sum_{j=1}^{q} \psi_{\tau,i} \right) \end{aligned} \tag{A3c}$$

Appendix B: Energy objective linearization

Energy function, from Equation 17, is written in Equation (B1) as function of problem variables.

Author-produced version of the article published in Journal of Irrigation and Drainage Engineering-ASCE, 2017, N°143(4)
The original publication is available at http://ascelibrary.org
Doi: 10.1061/(ASCE)WR.1943-5452.00007466

$$E_t = \eta \cdot \Delta k(\tau) \cdot [R_1(v_t) - R_2(r_{\text{turb},t} + r_{\text{spill},t})] \cdot r_{\text{turb},t} \tag{B1}$$

In Equation (B1), the hydraulic head Δh, is written as function of problem variables; $h_t = R_1(v_t)$ is the stage-storage curve and $h_t^v = R_2(r_{\text{turb},t} + r_{\text{spill},t})$ the rating-curve, where h_t is the water level in the reservoir and h_t^v the tailwater elevation, downstream of the reservoir.

Equations (B2) are the partial derivatives of Equation (B1).

$$\begin{aligned}
\frac{\partial E_t}{\partial v_t} &= \eta \cdot \Delta k(\tau) \cdot {R_1}'(v_t) \cdot r_{\text{turb},t} \\
\frac{\partial E_t}{\partial r_{\text{turb},t}} &= \eta \cdot \Delta k(\tau) \cdot \Big[R_1(v_t) - \\
&\quad {R_2}'(r_{\text{turb},t} + r_{\text{spill},t}) \cdot r_{\text{turb},t} - R_2(r_{\text{turb},t} + r_{\text{spill},t}) \Big] \\
\frac{\partial E_t}{\partial r_{\text{spill},t}} &= -\eta \cdot \Delta k(\tau) \cdot {R_2}'(r_{\text{turb},t} + r_{\text{spill},t})
\end{aligned} \tag{B2}$$

In Equations (B3), we consider a cylindrical reservoir in proximity of the operational reservoir water level, h_0, and a linear rating curve in proximity of discharges $r_{\text{turb},0} + r_{\text{spill},0}$.

$$R_1(v_t) \approx h_0 + \frac{1}{A_0} \cdot (v_t - v_0) \tag{B3a}$$

$$R_2(r_{\text{turb},t} + r_{\text{spill},t}) \approx h_0^v + m_v(r_{\text{turb},t} - r_{\text{turb},0} + r_{\text{spill},t} - r_{\text{spill},0}) \tag{B3b}$$

where A_0 is the reservoir surface corresponding to h_0, $h_0^v = R_2(r_{\text{turb},0} + r_{\text{spill},0})$, and $m_v = {R_2}'(r_{\text{turb},0} + r_{\text{spill},0})$.

Considering Equations (B2) and (B3), we get, in Equation (B4), the linear approximation of Equation (B1).

Author-produced version of the article published in Journal of Irrigation and Drainage Engineering-ASCE, 2017, N°143(4)
The original publication is available at http://ascelibrary.org
Doi: 10.1061/(ASCE)WR.1943-5452.00007466

$$\begin{aligned} E_t \approx & E_t(v_0, r_{\text{turb},0}, r_{\text{spill},0}) + \\ & + \frac{\partial E_t}{\partial v_t} \cdot (v_t - v_0) + \\ & + \frac{\partial E_t}{\partial r_{\text{turb},t}} \cdot (r_{\text{turb},t} - r_{\text{turb},0}) + \\ & + \frac{\partial E_t}{\partial r_{\text{spill},t}} (r_{\text{spill},t} - r_{\text{spill},0}) = \end{aligned} \tag{B4a}$$

$$\begin{aligned} = \eta \cdot \Delta k(\tau) \cdot \Big\{ & [h_0 - h_0^v] \cdot r_{\text{turb},0} + \\ & + \left[\frac{r_{\text{turb},0}}{A_0}\right] \cdot (v_t - v_0) + \\ & + [h_0 - h_0^v - m_v(1 + r_{\text{turb},0})] \, (r_{\text{turb},t} - r_{\text{turb},0}) \\ & + [-m_v](r_{\text{spill},t} - r_{\text{spill},0}) \Big\} \end{aligned} \tag{B4b}$$

Equation (B4b) gives the weights for $v_t, r_{\text{turb},t}$, and $r_{\text{spill},t}$ that maximise energy, as in Equation (18). For constant tailwater elevation, then $m_v = 0$, and Equations (B4b) and (18) can be simplified.

Appendix C: List of main variables

We use here the classic convention of representing vectors in bold.

$\mathbf{v}_t$ Reservoir volumes $[m^3]$

$\mathbf{r}_t$ Discharge decision $[m^3/s]$

$\mathbf{q}_t$ (Flow) scenarios $[m^3/s]$

N_{res} Number of reservoirs $[-]$

N_{dec} Number of discharge decisions $[-]$

N_{scen} Number of scenarios $[-]$

I Input matrix $[-]$

O Output matrix $[-]$

$\mathbf{c}_t(\cdot)$ Inequality constraints $[-]$

$g(\cdot)$ Time-step objective function $[-]$

$F(\cdot)$ Cost-to-go function $[-]$

$\mathcal{F}(\cdot)$ cost-to-go function approximation by Bender's cuts $[-]$

$\mathbb{E}$ Expected value $[-]$

$f_{\mathbf{Q}_t}(\mathbf{q}_t)$; Probability density function of q_t $[-]$

$\alpha_{\tau,i}, \phi_{\tau,i}, \psi_{\tau,i}, \sigma^2_\tau$ Periodic ARMA parameters

$\rho_{\tau,i}, \omega_{\tau,i}, \kappa_\tau$ Linearized multiplicative model parameters

$\bar{q}_\tau$ Climatic average of Q_τ m^3/s

t Time-step index $[-]$

τ Periodic time-step index $[-]$

Δt Daily time-step length $[86400s]$

Δk Number of daily time-steps $[days]$

T Period length $[-]$

H Optimization horizon $[-]$

e_t Error due to model linearization $[m^3/s]$

CV Cumulative variance $[-]$

$v_{\min}$ Minimum reservoir volume $[3.9 \times 10^9 m^3]$

$v_{\max}$ Maximum reservoir volume $[1.5 \times 10^{10} m^3]$

v_{safety} Reservoir volume safety limit $[1.18 \times 10^{10} m^3]$

$r_{\text{spill,max}}$ Maximum discharge through turbines $[500 m^3/s]$

J_E Annual average energy production $[GWh]$

E Daily energy production $[MWh]$

E^{op} Linearized energy production objective $[MWh]$

v_0 Operational volume $[1.18 \times 10^{10} m^3]$

$r_{\mathrm{turb},0}$ Operational release through turbines $[270 m^3/s]$

$r_{\mathrm{spill},0}$ Operational release through spillages $[0 m^3/s]$

Acknowledgments. Luciano Raso's research is funded by the AXA Research Fund.

References

Albergel, J., J.-C. Bader, and J.-P. Lamagat, Flood and drought: application to the senegal river management, *IAHS Publications-Series of Proceedings and Reports-Intern Assoc Hydrological Sciences*, *240*, 509–518, 1997.

Arjoon, D., Y. Mohamed, Q. Goor, and A. Tilmant, Hydro-economic risk assessment in the eastern nile river basin, *Water Resources and Economics*, *8*, 16–31, 2014.

Bader, J.-C., J.-P. Lamagat, and N. Guiguen, Gestion du barrage de manantali sur le fleuve sénégal: analyse quantitative d'un conflit d'objectifs, *Hydrological sciences journal*, *48*(4), 525–538, 2003.

Bartolini, P., and J. D. Salas, Modeling of streamflow processes at different time scales, *Water resources research*, *29*(8), 2573–2587, 1993.

Bartolini, P., J. D. Salas, and J. Obeysekera, Multivariate periodic arma (1, 1) processes, *Water Resources Research*, *24*(8), 1237–1246, 1988.

Bellman, R. E., and S. E. Dreyfus, *Applied dynamic programming*, vol. 7962, Princeton University Press, 1966.

Bertsekas, D. P., and J. N. Tsitsiklis, Neuro-dynamic programming: an overview, in *Decision and Control, 1995., Proceedings of the 34th IEEE Conference on*, vol. 1, pp. 560–564, IEEE, 1995.

Bezerra, B., Á. Veiga, L. A. Barroso, and M. Pereira, Assessment of parameter uncertainty in autoregressive streamflow models for stochastic long-term hydrothermal scheduling, in *Power and Energy Society General Meeting, 2012 IEEE*, pp. 1–8, IEEE, 2012.

Birge, J. R., and F. Louveaux, *Introduction to stochastic programming*, Springer Verlag, 1997.

Box, G., G. Jenkins, and G. Reinsel, *Time series analysis*, Holden-day San Francisco, 1970.

Camacho, F., A. McLeod, and K. Hipel, Multivariate contemporaneous arma model with hydrological applications, *Stochastic Hydrology and Hydraulics*, *1*(2), 141–154, 1987.

Castelletti, A., F. Pianosi, and R. Soncini-Sessa, Water reservoir control under economic, social and environmental constraints, *Automatica*, *44*(6), 1595–1607, doi: 10.1016/j.automatica.2008.03.003, 2008.

Dekking, M., *A modern introduction to probability and statistics: understanding why and how*, Springer Verlag, 2005.

Fraval, P., J. Bader, L. Mané, H. David-Benz, J. Lamagat, and O. D. Diagne, The quest for integrated and sustainable water management in the senegal river valley, in *5th Inter-Regional Conference on Environment and Water ENVIROWATER*, pp. 5–8, 2002.

Gjelsvik, A., B. Mo, and A. Haugstad, Long-and medium-term operations planning and stochastic modelling in hydro-dominated power systems based on stochastic dual dynamic programming, in *Handbook of Power Systems I*, pp. 33–55, Springer, 2010.

Goor, Q., R. Kelman, and A. Tilmant, Optimal multipurpose-multireservoir operation model with variable productivity of hydropower plants, *Journal of Water Resources Planning and Management*, *137*(3), 258–267, 2010.

Hipel, K. W., and A. I. McLeod, *Time series modelling of water resources and environmental systems*, Elsevier, 1994.

Karamouz, M., F. Szidarovszky, and B. Zahraie, *Water resources systems analysis*, CRC Press, 2003.

Koutsoyiannis, D., A random walk on water, *Hydrology and Earth System Sciences Discussions*, *6*, 6611–6658, 2009.

Linowsky, K., and A. B. Philpott, On the convergence of sampling-based decomposition algorithms for multistage stochastic programs, *Journal of optimization theory and applications*, *125*(2), 349–366, 2005.

Loucks, D. P., Water resource systems models: their role in planning, *Journal of Water Resources Planning and Management*, *118*(3), 214–223, 1992.

Marques, G. F., and A. Tilmant, The economic value of coordination in large-scale multireservoir systems: The parana river case, *Water Resources Research*, *49*(11), 7546–7557, 2013.

Mayne, D. Q., J. B. Rawlings, C. V. Rao, and P. O. M. Scokaert, Constrained model predictive control: Stability and optimality, *AUTOMATICA-OXFORD-*, *36*, 789–814, doi:10.1016/S0005-1098(99)00214-9, 2000.

Nicklow, J., et al., State of the art for genetic algorithms and beyond in water resources planning and management, *Journal of Water Resources Planning and Management*, *136*(4), 412–432, 2009.

Pereira, M., and L. M. Pinto, Multi-stage stochastic optimization applied to energy planning, *Mathematical Programming*, *52*(1), 359–375, doi:10.1007/BF01582895, 1991.

Pereira, M., G. Oliveira, C. Costa, and J. Kelman, Stochastic streamflow models for hydroelectric systems, *Water Resources Research*, *20*(3), 379–390, 1984.

Philpott, A. B., and Z. Guan, On the convergence of stochastic dual dynamic programming and related methods, *Operations Research Letters*, *36*(4), 450–455, 2008.

Rabiner, L. R., and B.-H. Juang, An introduction to hidden markov models, *ASSP Magazine, IEEE*, *3*(1), 4–16, 1986.

Raso, L., D. Schwanenberg, N. van de Giesen, and P. van Overloop, Short-term optimal operation of water systems using ensemble forecasts, *Advances in Water Resources*, *71*, 200–208, 2014.

Reed, P. M., D. Hadka, J. D. Herman, J. R. Kasprzyk, and J. B. Kollat, Evolutionary multiobjective optimization in water resources: The past, present, and future, *Advances in Water Resources*, *51*, 438–456, 2013.

Salas, J. D., *Applied modeling of hydrologic time series*, Water Resources Publication, 1980.

Salas, J. D., G. Q. Tabios, and P. Bartolini, Approaches to multivariate modeling of water resources time series1, *JAWRA Journal of the American Water Resources Association*, *21*(4), 683–708, 1985.

Seifi, A., and K. W. Hipel, Interior-point method for reservoir operation with stochastic inflows, *Journal of water resources planning and management*, *127*(1), 48–57, 2001.

Shapiro, A., Analysis of stochastic dual dynamic programming method, *European Journal of Operational Research*, *209*(1), 63–72, 2011.

Shapiro, A., and P. R. Andrzej, *Stochastic programming*, Elsevier, 2003.

Soncini-Sessa, R., A. Castelletti, and E. Weber, *Integrated and participatory water resources management*, Elsevier Science, 2007.

Stedinger, J. R., and M. R. Taylor, Synthetic streamflow generation: 1. model verification and validation, *Water resources research*, *18*(4), 909–918, 1982.

Stedinger, J. R., B. F. Sule, and D. P. Loucks, Stochastic dynamic programming models for reservoir operation optimization, *Water Resources Research*, *20*(11), 1499–1505, doi:10.1029/WR020i011p01499, 1984.

Tilmant, A., and R. Kelman, A stochastic approach to analyze trade-offs and risks associated with large-scale water resources systems, *Water resources research*, *43*(6), 2007.

Tilmant, A., and W. Kinzelbach, The cost of noncooperation in international river basins, *Water Resources Research*, *48*(1), 2012.

Tilmant, A., J. Lettany, and R. Kelman, Hydrological risk assessment in the euphrates-tigris river basin: A stochastic dual dynamic programming approach, *Water International*, *32*(2), 294–309, 2007.

Tilmant, A., D. Pinte, and Q. Goor, Assessing marginal water values in multipurpose multireservoir systems via stochastic programming, *Water Resources Research*, *44*(12), W12,431, doi:10.1029/2008WR007024, 2008.

Tilmant, A., Q. Goor, and D. Pinte, Agricultural-to-hydropower water transfers: sharing water and benefits in hydropower-irrigation systems, *Hydrology and Earth System Sciences*, *13*(7), 1091–1101, 2009.

Tilmant, A., L. Beevers, and B. Muyunda, Restoring a flow regime through the coordinated operation of a multireservoir system: The case of the zambezi river basin, *Water*

Resources Research, *46*(7), 2010.

Tilmant, A., W. Kinzelbach, D. Juizo, L. Beevers, D. Senn, and C. Casarotto, Economic valuation of benefits and costs associated with the coordinated development and management of the zambezi river basin, *Water Policy*, *14*(3), 490, 2012.

Trezos, T., and W. W.-G. Yeh, Use of stochastic dynamic programming for reservoir management, *Water Resources Research*, *23*(6), 983–996, doi:10.1029/WR023i006p00983, 1987.

Tsitsiklis, J. N., and B. Van Roy, Feature-based methods for large scale dynamic programming, *Machine Learning*, *22*(1-3), 59–94, 1996.

Turgeon, A., Optimal operation of multireservoir power systems with stochastic inflows, *Water Resources Research*, *16*(2), 275–283, doi:10.1029/WR016i002p00275, 1980.

van Overloop, P. J., S. Weijs, and S. Dijkstra, Multiple model predictive control on a drainage canal system, *Control Engineering Practice*, *16*(5), 531–540, doi: 10.1016/j.conengprac.2007.06.002, 2008.

Vecchia, A., Periodic autoregressive-moving average (parma) modeling with applications to water resources1, *JAWRA Journal of the American Water Resources Association*, *21*(5), 721–730, 1985.

Weijs, S. V., Information theory for risk-based water system operation, Ph.D. thesis, Delft University of Technology, 2011.

Author-produced version of the article published in Journal of Irrigation and Drainage Engineering-ASCE, 2017, N°143(4)
The original publication is available at http://ascelibrary.org
Doi: 10.1061/(ASCE)WR.1943-5452.00007466

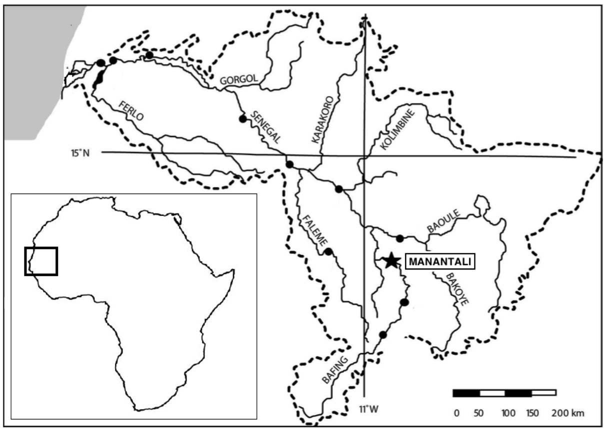


Figure 1. Map of the Senegal Basin.

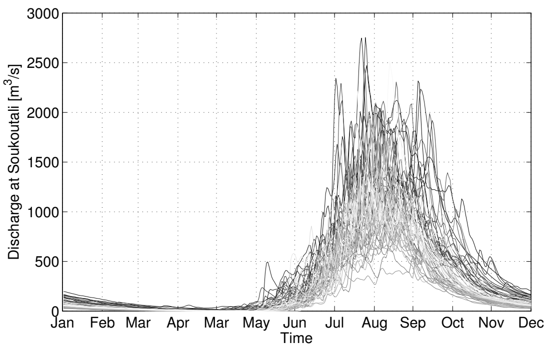
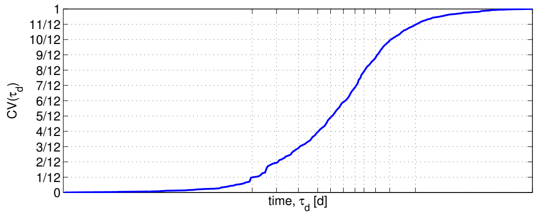
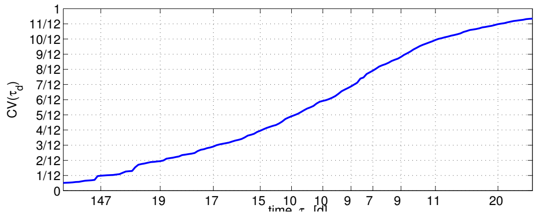


Figure 2. Inflow at Soukoutali from 1 January 1950 to 31 December 2013, daily time-step.

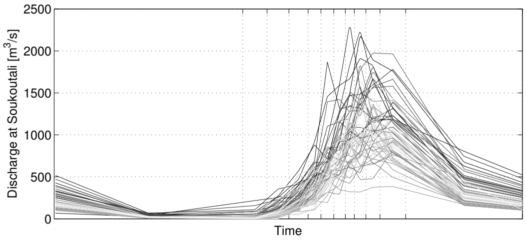


(a) τ_d between 1 and 365

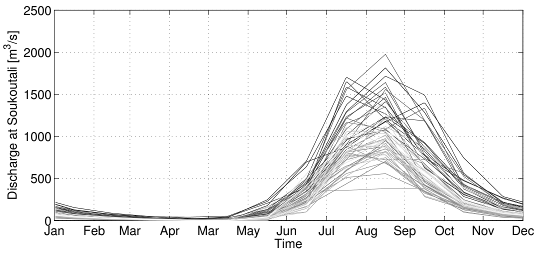


(b) τ_d between 135 and 285

Figure 3. Cumulative Variance on τ_d.



(a) Non-uniform aggregation



(b) Monthly aggregation

Figure 4. Inflow at Soukoutali from 1 January 1950 to 31 December 2013, aggregated observed discharge.

Author-produced version of the article published in Journal of Irrigation and Drainage Engineering-ASCE, 2017, N°143(4)
The original publication is available at http://ascelibrary.org
Doi: 10.1061/(ASCE)WR.1943-5452.00007466

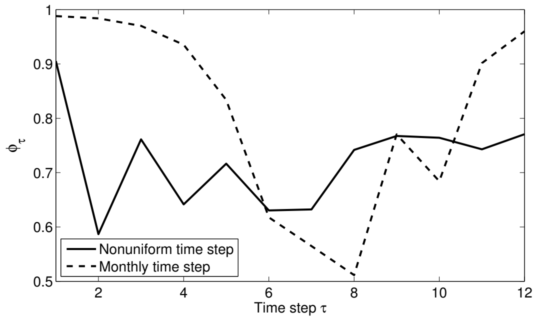


Figure 5. ϕ_τ for non-uniform and monthly time-step

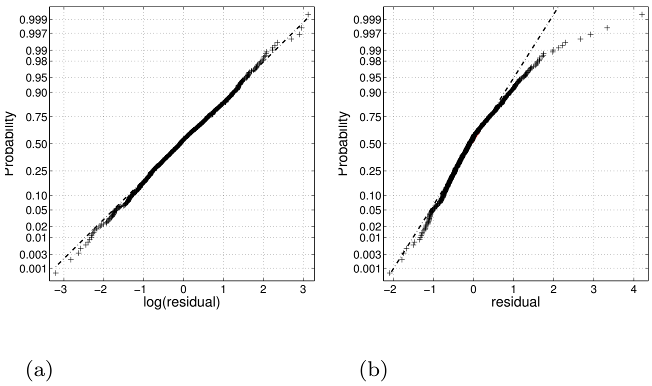


Figure 6. Normality plot for logarithms of model residuals of Model (9), Multiplicative Model (a), and residuals of Thomas-Fiering, Additive Model (b).

Author-produced version of the article published in Journal of Irrigation and Drainage Engineering-ASCE, 2017, N°143(4)
The original publication is available at http://ascelibrary.org
Doi: 10.1061/(ASCE)WR.1943-5452.00007466

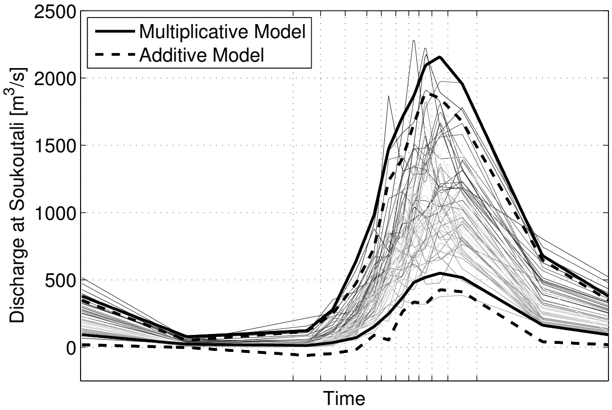
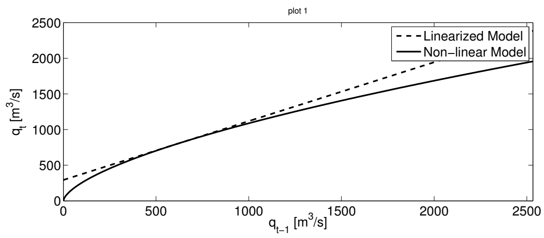
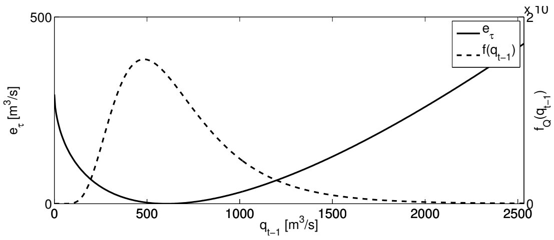


Figure 7. Observed discharge data and 95% confidence bands for Model (8), or multiplicative model (continuous bold lines) and Thomas-Fiering, or additive model (dashed bold lines).



(a) Linear vs non-linear model.



(b) error due to linearisation vs q_{t-1} distribution.

Figure 8. Error due to linearisation at $\tau = 7$.

Author-produced version of the article published in Journal of Irrigation and Drainage Engineering-ASCE, 2017, N°143(4)
The original publication is available at http://ascelibrary.org
Doi: 10.1061/(ASCE)WR.1943-5452.00007466

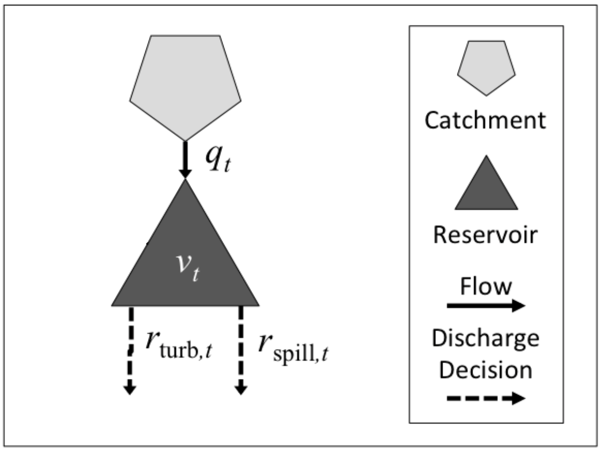
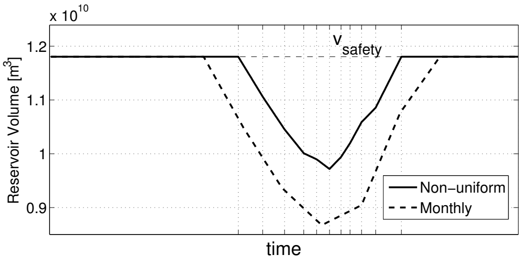
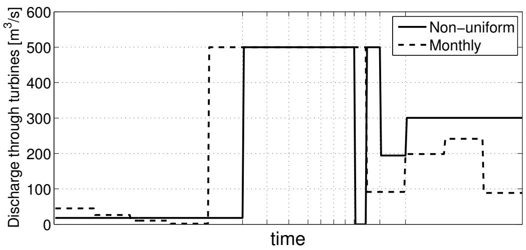


Figure 9. Schema of the Senegal River System at Manantali, including: Manantali reservoir volume, v_t, inflows at Soukoutali, q_t, and discharge decisions, $\mathbf{r}_t$.



(a)



(b)

Figure 10. Reservoir volume (a) and discharge trough turbines (b) for monthly (dashed line) and non-uniform (continuous line) aggregation, year 2005.