Modeling Time-dependent Randomness in Stochastic Dual Dynamic Programming

Nils Löhndorf

Department of Information Systems and Operations

Vienna University of Economics and Business

Vienna, Austria

nils.loehndorf@wu.ac.at

Alexander Shapiro*

School of Industrial and Systems Engineering

Georgia Institute of Technology

Atlanta, GA 30332-0205, USA

ashapiro@isye.gatech.edu

Abstract

We consider the multistage stochastic programming problem where uncertainty enters the right-hand sides of the problem. Stochastic Dual Dynamic Programming (SDDP) is a popular method to solve such problems under the assumption that the random data process is stagewise independent. There exist two approaches to incorporate dependence into SDDP. One approach is to model the data process as an autoregressive time series and to reformulate the problem in stagewise independent terms by adding state variables to the model (TS-SDDP). The other approach is to use Markov Chain discretization of the random data process (MC-SDDP). While MC-SDDP can handle any Markovian data process, some advantages of statistical analysis of the policy under the true process are lost.

In this work, we compare both approaches based on a computational study using the long-term operational planning problem of the Brazilian interconnected power systems. We find that the optimality bounds computed by the MC-SDDP method close faster than its TS-SDDP counterpart, and the MC-SDDP policy dominates the TS-SDDP policy. When implementing the optimized policies on real data, we observe that not only the method but also the quality of the stochastic model has an impact on policy performance and that using an AV@R formulation is effective in making the policy robust against a misspecified stochastic model.

Keywords: Stochastic programming, dynamic programming, Markov decision process, coherent risk measures

*Research of this author was partly supported by NSF grant 1633196 and DARPA EQUiPS program, grant SNL 014150709.

1 Introduction

A scalable way of solving multistage stochastic decision problems is based on approximate dynamic programming. In particular, the stochastic dual dynamic programming (SDDP) method based on Pereira and Pinto's seminal work [18] became popular in many applications. In SDDP it is crucial to assume randomness of the data process to be stagewise independent, because in this case the respective cost-to-go (value) functions of the dynamic programming equations are independent of the data process. However, in many applications the data process has a Markovian type dependence structure. To deal with this problem two approaches have been suggested in the literature.

One approach is to model the data process as an autoregressive process and to add the time series transition equations as additional equality constraints into the optimization problem, so that realizations of the data process are treated as decision variables [17, 25]. This allows to reformulate the problem in terms of the stagewise independent errors of the considered time series, at the price of increasing the number of state variables. We refer to this approach as the TS-SDDP method.

An alternative approach is based on Markov Chain discretization of the Markovian data process using optimal quantization [6, 9, 21]. We refer to this approach as the MC-SDDP method (this approach is referred to as Approximate Dual Dynamic Programming (ADDP) in [9, 11]).

Both approaches have advantages and disadvantages. While TS-SDDP is restricted to the linear case and right-hand side uncertainty, MC-SDDP allows modeling any parameter as Markovian data process. MC-SDDP therefore allows usage of a much broader range of stochastic models which may better represent the true process. Yet MC-SDDP makes no convergence guarantees with respect to the true process, whereas the properties of the TS-SDDP bounds are well-established.

Whichever approach provides the better policy in practice has not yet been studied. We intend to close this gap and compare both approaches theoretically as well as numerically. In Section 2 we introduce the two approaches and discuss constructions of lower bounds for the optimal objective value. In Section 3 we discuss how a continuous data process can be discretized to a lattice using optimal quantization to provide a time-nonhomogeneous Markov chain. Starting from Section 4, we test the performance of both approaches based on the long-term operational planning problem of the Brazilian interconnected power system. The problem is widely used in the literature on SDDP due to its practical importance for providing generation targets of power plants in the Brazilian power system [13].

In previous work on the performance of SDDP [25], it was found that the TS-SDDP

approach cannot close the gap between lower and upper bounds when time-dependent randomness is represented by an appropriate autoregressive model. In Section 5, we demonstrate that the MC-SDDP method computes tighter lower bounds for this problem than its TS-SDDP counterpart, and provides slightly better policies. Moreover, MC-SDDP enables the use of inflow models that show better performance when implementing the decision policy on real data.

2 Modeling Time-dependent Randomness

Consider the following risk averse multistage linear stochastic programming problem

$$\min_{\substack{A_1x_1=b_1\\ x_1\geq 0}} c_1^\mathsf{T}x_1 + \rho_{2|\xi_1}\left[\min_{\substack{B_2x_1+A_2x_2=b_2\\ x_2\geq 0}} c_2^\mathsf{T}x_2 + \cdots + \rho_{T|\xi_{[T-1]}}\left[\min_{\substack{B_Tx_{T-1}+A_Tx_T=b_T\\ x_T\geq 0}} c_T^\mathsf{T}x_T\right]\right]. \tag{2.1}$$

Components of vectors c_t, b_t and matrices A_t, B_t are modeled as random variables forming the stochastic data process[1] $\xi_t = (c_t, A_t, B_t, b_t)$, $t = 2, ..., T$, with $\xi_1 = (c_1, A_1, b_1)$ being deterministic (not random), $\xi_{[t]} = (\xi_1, ..., \xi_t)$ denotes history of the data process up to time t, and $\rho_{t|\xi_{[t-1]}}$ are conditional risk measures (see, e.g., [26]). In particular if risk measures $\rho_{t|\xi_{[t-1]}}$ are given as conditional expectations $\mathbb{E}_{|\xi_{[t-1]}}$, problem (2.1) becomes a risk neutral multistage stochastic programming problem. It is said that data process ξ_t as *stagewise independent* if random vector ξ_{t+1} is independent of $\xi_{[t]}$, $t = 1, ..., T-1$.

2.1 Static Case

In this section we consider two stage problems. We discuss construction of lower bounds for the optimal value of the respective stochastic problem.

We start with the risk neutral problem of the form

$$\min_{x\in\mathcal{X}} \mathbb{E}[F(x,\xi)], \tag{2.1}$$

where $\mathcal{X} \subset \mathbb{R}^n$ is a nonempty set and $F : \mathcal{X} \times \mathbb{R}^d \to \mathbb{R}$. In the two stage setting the function $F(x, \xi)$ can be viewed as the cost of the first stage plus the cost of the second stage problem, $\mathcal{X}$ as a feasible set of the first stage and ξ is composed from random elements of (c_2, A_2, B_2, b_2). The expectation in (2.1) is taken with respect to random vector ξ whose

[1] Of course, not all elements of the data vectors ξ_t should be random. For example, we can model only the right-hand side vectors b_t as random while all other elements of ξ_t being fixed (known). We use the same notation ξ_t for random vector and its particular realization; meaning of that will be clear from the context.

probability distribution (probability measure) P is supported on set $\Xi \subset \mathbb{R}^d$. The set Ξ equipped with its Borel sigma algebra $\mathcal{F}$ and measure P becomes a probability space. We also use notation $F_x(\xi) := F(x,\xi)$ when considering $F(x,\xi)$ as a function of ξ for fixed $x \in \mathcal{X}$. It is assumed that the expectation $\mathbb{E}[F(x,\xi)]$ is well defined and finite valued for every $x \in \mathcal{X}$. In particular, this means that for every $x \in \mathcal{X}$ the second stage problem is feasible w.p.1., i.e., we assume relatively complete recourse.

Let $\Xi = \cup_{i=1}^K \Xi_i$ be a disjoint partition of the set Ξ into measurable subsets $\Xi_i \subseteq \Xi$, $i = 1, ..., K$. Assuming that probabilities $\pi_i := P(\xi \in \Xi_i)$, $i = 1, ..., K$, are positive, conditional expectations are given by

$$\mu_i := \mathbb{E}[\xi | \xi \in \Xi_i] = \frac{1}{\pi_i} \int_{\Xi_i} \xi dP(\xi). \tag{2.2}$$

Note that $\sum_{i=1}^K \pi_i \mu_i = \mathbb{E}[\xi]$. Suppose that for every $x \in \mathcal{X}$ the function $F(x,\cdot)$ is convex[2]. Then by Jensen's inequality we have for every $x \in \mathcal{X}$,

$$\mathbb{E}[F_x(\xi)] = \sum_{i=1}^K \pi_i \mathbb{E}[F_x(\xi) | \xi \in \Xi_i] \geq \sum_{i=1}^K \pi_i F_x(\mu_i). \tag{2.3}$$

Hence, the optimal value of the following problem gives a lower bound for the optimal value of problem (2.1)

$$\min_{x \in \mathcal{X}} \sum_{i=1}^K \pi_i F(x, \mu_i). \tag{2.4}$$

This lower bound is well known in stochastic programming (e.g., [5, Section 4.3] and references therein).

Remark 2.1 A natural question is how well problem (2.4) approximates the original problem (2.1), i.e., how fine should the partition be in order for the difference between the optimal values of these problems to be sufficiently small. Let $g_i \in \partial F_x(\mu_i)$ be a subgradient of $F_x(\cdot)$ at μ_i, $i = 1, ..., K$. Then we have that $F_x(\mu_i) = \mathbb{E}\left[F_x(\mu_i) + g_i^{\mathsf{T}}(\xi - \mu_i) \middle| \xi \in \Xi_i\right]$, and hence

$$\mathbb{E}[F_x(\xi)] - \sum_{i=1}^K \pi_i F(x, \mu_i) = \sum_{i=1}^K \pi_i \mathbb{E}\left[F_x(\xi) - F_x(\mu_i) - g_i^{\mathsf{T}}(\xi - \mu_i) \middle| \xi \in \Xi_i\right]. \tag{2.5}$$

It follows that if for every $i = 1, ..., K$ and all $\xi \in \Xi_i$ the difference between $F_x(\xi)$ and the cutting plane $F_x(\mu_i) + g_i^{\mathsf{T}}(\xi - \mu_i)$ is smaller than an accuracy constant $\varepsilon > 0$, then the

[2]In the two stage setting, if the recourse is fixed, i.e., the matrix A_2 is fixed (not random), then the optimal value of the second stage is convex in (B_2, b_2) and concave in c_2.

left-hand side of (2.5) is also less than ε. This shows that, for sufficiently fine partition of the set Ξ, the difference between optimal values of problems (2.1) and (2.4) can be made arbitrary small. However, for a general convex function of d variables the number of cutting planes needed to uniformly approximate this function grows exponentially with dimension d.

2.2 Risk-averse Case

The Jensen's inequality lower bound has a natural extension for risk averse problems of the form

$$\min_{x\in\mathcal{X}} \rho[F_x(\xi)], \tag{2.6}$$

with ρ being a coherent risk measure. That is, consider the linear space $\mathcal{Z} := L_p(\Xi, \mathcal{F}, P)$ of random variables $Z : \mathcal{Z} \to \mathbb{R}$ having finite p-th order moments, $p \in [1, \infty)$. The dual of space $\mathcal{Z}$ is the space $\mathcal{Z}^* = L_q(\Xi, \mathcal{F}, P)$, where $q \in (1, \infty]$ is such that $1/p + 1/q = 1$. A functional $\rho : \mathcal{Z} \to \mathbb{R}$ is called *coherent risk measure* if it satisfies the axioms of convexity, monotonicity, translation equivariance and positive homogeneity (cf. [1]). Every real valued coherent risk measure has the following dual representation

$$\rho(Z) = \sup_{Q\in\mathfrak{M}} \mathbb{E}_Q[Z], \tag{2.7}$$

where $\mathfrak{M}$ is a set of probability measures Q is absolutely continuous with respect to the reference measure P and such that the corresponding density $\zeta = dQ/dP$ belongs to the space $\mathcal{Z}^*$ (cf. [22]).

Assuming that random variable $F_x(\xi)$ belongs to the space $\mathcal{Z}$, and using Jensen's inequality for every $Q \in \mathfrak{M}$ and convex $F_x(\cdot)$, we can write

$$\rho[F_x(\xi)] \geq \mathbb{E}_Q[F_x(\xi)] \geq \sum_{i=1}^{K} \pi_i^Q F(x, \mu_i^Q),$$

where $\pi_i^Q := Q(\xi \in \Xi_i)$ and $\mu_i^Q := \mathbb{E}_Q[\xi|\xi \in \Xi_i]$, $i = 1, ..., K$. Therefore, for the risk averse problem (2.6), a lower bound is given by the optimal value of the following minimax problem

$$\min_{x\in\mathcal{X}} \max_{Q\in\mathfrak{M}} \sum_{i=1}^{K} \pi_i^Q F(x, \mu_i^Q). \tag{2.8}$$

Formula (2.8), although general, could be difficult to use in numerical implementations. In some cases it is possible to construct lower bounds in a more direct way. For example consider

the following coherent risk measure

$$\rho(Z) := (1-\lambda)\mathbb{E}[Z] + \lambda \mathsf{AV@R}_\alpha(Z), \tag{2.9}$$

where $\mathsf{AV@R}_\alpha$ is the so-called Average Value-at-Risk (also called Conditional Value-at-Risk, Expected Shortfall and Expected Tail Loss)

$$\mathsf{AV@R}_\alpha(Z) := \inf_{u\in\mathbb{R}} \left\{ u + \alpha^{-1}\mathbb{E}[Z-u]_+ \right\} \tag{2.10}$$

and the expectation is taken with respect to the reference probability measure P. In that case we can write problem (2.6) in the following equivalent form

$$\min_{y\in\mathcal{Y}} \mathbb{E}[G_y(\xi)], \tag{2.11}$$

where $y := (x,u)$, $\mathcal{Y} := \mathcal{X}\times\mathbb{R}$ and

$$G_y(\xi) := (1-\lambda)F_x(\xi) + \lambda u + \lambda\alpha^{-1}[F_x(\xi)-u]_+. \tag{2.12}$$

Note that if $F_x(\cdot)$ is convex, then $G_x(\cdot)$ is also convex. Hence, the optimal value of the following problem gives a lower bound for the optimal value of problem (2.6)

$$\min_{y\in\mathcal{Y}} \sum_{i=1}^K \pi_i G_y(\mu_i). \tag{2.13}$$

2.3 Multistage Dynamic Case

Consider the multistage stochastic program (2.1). Let us start with the risk neutral case, that is suppose that $\rho_{t|\xi_{[t-1]}} := \mathbb{E}_{|\xi_{[t-1]}}$, and hence the corresponding multistage problem can be written as

$$\underset{\substack{A_1x_1=b_1\\ x_1\geq 0}}{\mathrm{Min}} c_1^\mathsf{T}x_1 + \mathbb{E}\left[\min_{\substack{B_2x_1+A_2x_2=b_2\\ x_2\geq 0}} c_2^\mathsf{T}x_2 + \mathbb{E}\Big[\cdots + \mathbb{E}\left[\min_{\substack{B_Tx_{T-1}+A_Tx_T=b_T\\ x_T\geq 0}} c_T^\mathsf{T}x_T \right] \Big] \right]. \tag{2.1}$$

Suppose further that the data process is stagewise independent. Then the dynamic programming equations for problem (2.1) can be written as

$$Q_t(x_{t-1},\xi_t) = \min\left\{ c_t^\mathsf{T}x_t + \mathcal{Q}_{t+1}(x_t) : B_tx_{t-1} + A_tx_t = b_t,\ x_t\geq 0 \right\}, \tag{2.2}$$

$$\mathcal{Q}_{t+1}(x_t) := \mathbb{E}\left[Q_{t+1}(x_t,\xi_{t+1}) \right], \tag{2.3}$$

$t = 2, ..., T$, with term $\mathcal{Q}_{T+1}$ omitted (cf., [26, Section 3.1.1]). Suppose that the cost-to-go functions $Q_t(x_{t-1}, \xi_t)$ are convex in ξ_t. This holds if only b_t and B_t are random. Let support Ξ_t of random vector ξ_t be partitioned,

$$\Xi_t = \cup_{i=1}^{K_t} \Xi_{ti}, \ t = 2, ..., T. \tag{2.4}$$

Consider probabilities $\pi_{ti} := P(\xi_t \in \Xi_{ti})$ and conditional expectations $\mu_{ti} := \mathbb{E}[\xi_t | \xi_t \in \Xi_{ti}]$. Then by Jensen's inequality we have

$$\mathcal{Q}_{t+1}(x_t) \geq \sum_{i=1}^{K_t} \pi_{ti} Q_{t+1}(x_t, \mu_{t+1,i}). \tag{2.5}$$

Hence, similar to the static case by discretizing the distribution of ξ_t, $t = 2, ..., T$, using points μ_{ti} and assigned probabilities π_{ti}, $i = 1, ..., K_t$, we obtain a multistage problem whose optimal value gives a lower bound for the optimal value of the original problem (2.1) (cf. [5, Section 11.1]).

Without the stagewise independence condition, the situation is considerably more involved. Then the expected value cost-to-go functions also depend on realizations of the data process. In order to deal with this in a reasonable way, we assume that the data process is Markovian, i.e., conditional distribution of ξ_{t+1} given $\xi_{[t]}$ is the same as conditional distribution of ξ_{t+1} given ξ_t , $t = 1, ..., T-1$. Then the corresponding dynamic programming equations become (cf., [26, Section 3.1.1])

$$Q_t(x_{t-1}, \xi_t) = \min \left\{ c_t^\mathsf{T} x_t + \mathcal{Q}_{t+1}(x_t, \xi_t) : B_t x_{t-1} + A_t x_t = b_t, \ x_t \geq 0 \right\}, \tag{2.6}$$

$$\mathcal{Q}_{t+1}(x_t, \xi_t) := \mathbb{E}\left[Q_{t+1}(x_t, \xi_{t+1}) \big| \xi_t \right]. \tag{2.7}$$

That is the cost-to-go functions $Q_t(x_{t-1}, \xi_t)$ and the expected value cost-to-go functions $\mathcal{Q}_{t+1}(x_t, \xi_t)$ depend only on ξ_t rather than the whole history of the data process.

One way of dealing with a Markovian data process is to model the state transition function by including additional equality constraints into the optimization problem (cf. [17]). Suppose that only the right-hand sides b_t are random, i.e. $\xi_t = b_t$, and the process b_t is modelled as a first order autoregressive process. That is $b_t = \mu + \Phi b_{t-1} + \varepsilon_t$ with the error process ε_t being stagewise independent. Then

$$\begin{aligned} \mathcal{Q}_t(x_t, \xi_t) &= \mathbb{E}[Q_{t+1}(x_t, \xi_{t+1}) | \xi_t] \\ &= \mathbb{E}[Q_{t+1}(x_t, \mu + \Phi \xi_t + \varepsilon_{t+1}) | \xi_t] \\ &= \mathbb{E}[Q_{t+1}(x_t, \mu + \Phi \xi_t + \varepsilon_{t+1})], \end{aligned} \tag{2.8}$$

where the last equality holds since ξ_t is modeled as an autoregressive process and hence is a function of $\varepsilon_t, \varepsilon_{t-1}, ...$, and thus is independent of ε_{t+1}. The last expectation in (2.8) is taken with respect to ε_{t+1}. In the considered case $\mathcal{Q}_{t+1}(x_t, b_t)$ is convex in (x_t, b_t). This can be proved by induction in $t = T-1, ..., 1$. Let us refer to the above approach as the Time Series (TS) approach.

Again consider partitions (2.4) of the supports of random vectors ξ_t. This leads to an approximation of the data process by a Markov chain with transition probabilities

$$\pi_{tij} := P(\xi_{t+1} \in \Xi_{t+1,j} | \xi_t \in \Xi_{ti}) \tag{2.9}$$

of moving from $\xi_t = \mu_{ti}$ to $\xi_{t+1} = \mu_{t+1,j}$. This approximation can be viewed as a recombined scenario tree or *lattice*. That is, at the first stage, the lattice has got one (root) node. There are K_2 nodes at the second stage, corresponding to points μ_{2j}, $j = 1, ..., K_2$, with probability of moving from the root node to node μ_{2j} given by π_{12j} at stage two. There are $K_2 K_3$ nodes at the third stage with the conditional probability of moving from node μ_{2i} to the respective node at stage 3 given by π_{2ij}, and so on for later stages. The total number of scenarios in that tree is $\prod_{t=2}^{T} K_t$. Although the total number of scenarios grows exponentially in the number of stages, the corresponding dynamic programming equations can still be solved in certain situations by approximate dynamic programming techniques. Let us refer to this as the Markov Chain (MC) approach.

Both approaches have advantages and disadvantages. One of the advantages of the TS approach is that the corresponding multistage problem can be formulated in terms of the stagewise independent error process ε_t while treating b_t as state variables. Consequently, the lower bound of the form (2.5) can be constructed by discretizing the error process ε_t. The price of such reformulation is an increase of the number of state variables and hence computational complexity of the problem. Also the TS approach is appropriate only when uncertainty in the data process is in the right-hand side, b_t. Otherwise, adding the corresponding state variables destroys convexity of the cost-to-go functions.

For the MC approach it is not clear how to construct a lower bound similar to (2.5). Indeed, by Jensen's inequality we have

$$\mathcal{Q}_{t+1}(x_t, \xi_t | \xi_t \in \Xi_{ti}) \geq \sum_{j=1}^{K_{t+1}} \pi_{tij} Q_{t+1}(x_t, \mu_{t+1,j}(\xi_t)), \tag{2.10}$$

where

$$\mu_{t+1,j}(\xi_t) := \mathbb{E}\left[\xi_{t+1} | \xi_{t+1} \in \Xi_{t+1,j};\, \xi_t\right]. \tag{2.11}$$

The problem is that here the points $\mu_{t+1,j}(\xi_t)$ also depend on ξ_t.

Consider now the risk averse problem (2.1). Suppose that risk measure $\rho_{t|\xi_{[t-1]}}$, $t = 2, ..., T$, is the conditional analogue of the risk measure (compare with (2.9))

$$\rho_t(Z) := (1 - \lambda_t)\mathbb{E}[Z] + \lambda_t \mathsf{AV@R}_{\alpha_t}(Z). \tag{2.12}$$

Assuming that the data process is stagewise independent the corresponding dynamic programming equations are similar to (2.2)–(2.3) with (2.3) replaced by

$$\mathcal{Q}_{t+1}(x_t) := \rho_t \left[Q_{t+1}(x_t, \xi_{t+1}) \right]. \tag{2.13}$$

Similar to (2.11)–(2.12) it is possible to reformulate this dynamic equations in a form of expectations by adding state variables u_t corresponding to the variational representation (2.10) of $\mathsf{AV@R}$ (cf., [24]). Consequently for the TS approach it is possible to construct lower bound similar to (2.5).

3 Markov Chain Stochastic Dual Dynamic Programming

When the TS approach is used, computation of the lower bound proceeds in the usual way (cf., [25]). For a description of the MC-SDDP method we can refer, e.g., to [26, pp. 258 - 259]. Note that by solving the problem for the constructed Markov Chain we generate a policy only for the set of scenarios generated by the considered Markov Chain. In order to extend such policy to all possible scenarios (sample paths), the usual approach is to find a scenario from the set of MC scenarios in some sense closest to a considered sample path of the continuous process and to use the respective approximations of the cost-to-go functions for evaluating policy decisions for the considered sample path (cf., [6]). Note also that the lower bound computed by the MC approach is valid for the MC discretized problem rather than the original continuous data problem.

Discretization via Lattice Quantization Let us describe a method that constructs a discretization of the continuous data process by separating the continuous data process into a finite number of disjoint partitions. Since partitions are found using optimal quantization and the set of quantizers forms a discrete probability lattice, we refer to this method as lattice quantization.

We are interested in creating a lattice with minimal discretization error. We can infer

from [4] that this error depends on the stagewise Wasserstein distance between the data process ξ_t and its discretization. More specifically, at each stage $t = 2, \ldots, T$, we find partition means (quantizers) μ_{ti}, $i = 1, ..., K_t$, which are solutions of the following problem

$$\min_{\mu_{t1},...,\mu_{tK_t}} \int \min_{i=1,...,K_t} \|\xi_t - \mu_{ti}\|^2 dP(\xi_t), \tag{3.14}$$

where $\|\cdot\|$ denotes the usual Euclidean norm. Given a solution $\mu_{t1}, ..., \mu_{tK_t}$ of (3.14), the space is partitioned into subsets Γ_{ti} consisting of points which are closest to μ_{ti}. The corresponding sets

$$\Gamma_{ti} := \{\xi_t : \|\xi_t - \mu_{ti}\| \leq \|\xi_t - \mu_{tj}\|,\ j = 1, ..., K_t,\ j \neq i\}, \tag{3.15}$$

represent the Voronoi partition associated with respective means μ_{ti}, $i = 1, ..., K_t$, which serve as the nodes of the lattice in t. Note the following relation between optimal partition and the center points μ_{ti}:

$$\mu_{ti} = \frac{\int_{\Gamma_{ti}} \xi_t dP(\xi_t)}{P(\Gamma_{ti})}.$$

That is, $\mu_{ti} = \mathbb{E}\left[\xi_t | \xi_t \in \Gamma_{ti}\right]$.

Unfortunately, finding the optimal center points is an $\mathcal{NP}$-hard problem [2]. In [4], faced with a similar problem, the authors propose a method based on stochastic approximation which we will adopt here. The stochastic approximation method draws S random sequences $(\hat{\xi}_t^s)_{t=1}^T$, $s = 1, ..., S$, from the continuous data process. With $(\beta_s)_{s=1}^S$ a sequence of stepsizes with $0 \leq \beta_s \leq 1$, $s = 1, ..., S$, means are updated recursively using

$$\mu_{ti}^s := \begin{cases} \mu_{ti}^{s-1} + \beta_k \left(\hat{\xi}_t^s - \mu_{ti}^{s-1}\right) & \text{if } i = \operatorname{argmin}\left\{||\hat{\xi}_t^s - \mu_{tk}^{s-1}||_2^2,\ k = 1, ..., K_t\right\}, \\ \mu_{ti}^{s-1} & \text{otherwise,} \end{cases} \tag{3.16}$$

with $\mu_{0i}^{s-1} \equiv 0$, for $i = 1, ..., K_t$, $t = 2, \ldots, T$, $s = 1, ..., S$. In [4] the authors show that if the sequence $(\beta_s)_{s=1}^S$ satisfies $\sum_{s=1}^\infty \beta_s = \infty$ and $\sum_{s=1}^\infty \beta_s^2 < \infty$, then the resulting means are local minimizers of (3.14).

To estimate the transition probabilities between partitions at subsequent stages, define the nodes of the lattice as $\mu_{ti} := \mu_{ti}^S$ and denote π_t, $t = 1, ..., T-1$, as the $|K_t| \times |K_{t+1}|$ transition matrix between layers t and $t+1$ with elements π_{tij}, where each π_{tij} defines the (conditional) probability of a state transition from μ_{ti} to $\mu_{t+1,j}$. The transition probabilities (2.9) can be estimated by counting the sample transition between successive partitions,

$$\hat{\pi}_{tij} = \frac{\sum_{s=1}^S \mathbf{1}_{\Gamma_{ti}}(\hat{\xi}_t^s)\mathbf{1}_{\Gamma_{tj}}(\hat{\xi}_{t+1}^s)}{\sum_{s=1}^S \mathbf{1}_{\Gamma_{ti}}(\hat{\xi}_t^s)},\ i = 1, ..., K_t,\ j = 1, ..., K_{t+1},\ t = 1, ..., T-1, \tag{3.17}$$

with $\mathbf{1}_A$ denoting the indicator function of the set A.

Stochastic Dual Dynamic Programming The SDDP method finds an outer approximation of the cost-to-go functions based on cutting planes which provides a lower bound of the optimal objective value of the multistage stochastic programming problem. In the MC-SDDP approach this is applied to the Markov Chain discretized problem with the respective lower bound valid for *discretized* MC problem. As it was already mentioned, a solution of the MC discretization provides a policy for the set of scenarios given by the constructed lattice with nodes μ_{ti}, $i = 1, ..., K_t$, $t = 1, ..., T-1$. In order to extend this policy to the original continuous data process we proceed in the following way. Consider a scenario (sample path) ξ_t, $t = 1, ..., T$, drawn from the continuous data process. For each $t = 1, ..., T$, find a point $\tilde{\mu}_{ti}$ closest to ξ_t among the points μ_{ti}, $i = 1, ..., K_t$. Consequently, for the considered scenario, use the constructed approximations of the cost-to-go functions $\mathcal{Q}_{t+1}(\cdot, \mu_{ti})$ for evaluating policy value in the forward step of the algorithm (cf., [6]).

4 Case Study: the Brazilian Interconnected Power System

Our numerical analysis is based on a multistage stochastic program that is widely used in the literature, the longterm planning problem of the Brazilian interconnected power system, which, due to its impact on the Brazilian spot price, presents an important problem in energy resource management.

In 2010, the Brazilian interconnected power system contained more than 200 power plants with a total installed generation capacity of 109.6 gigawatt of which 79.1% were hydropower. Long-term operational planning of this system is important to make efficient use of hydropower and take advantage of the hydrological diversity between different geographical regions.

4.1 The Long-term Operational Planning Problem

The objective of the long-term operational planning problem is to define an operational policy which provides generation targets for each plant to supply the load at minimal system cost constrained by generation capacity limits as well as transmission constraints. The existence of multiple multi-season reservoirs and the variability of natural inflows makes the problem a multistage optimization problem under uncertainty.

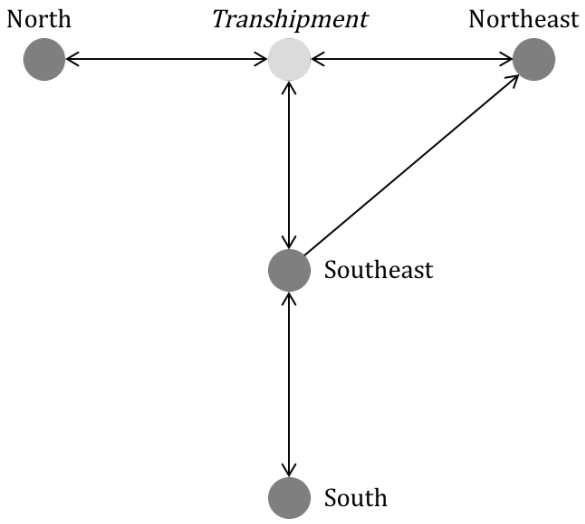

Figure 1: Interconnected power system with energy equivalent reservoirs.

The typical planning horizon of the long-term operational planning problem is 60 months due to some large reservoirs exhibiting multi-season regulation capacity. Since an individual representation of all 141 hydropower plants as well as multivariate randomness of all of their natural inflows results in a high-dimensional dynamic program, the problem is typically separated into long-term, medium-term, and short-term models. For the long-term model, hydropower plants that belong to a homogenous hydrological region are typically aggregated into four equivalent energy reservoirs [3]. The aggregate cost-to-go functions then enter a disaggregate medium-term model with a much shorter planning horizon which then again provides boundary conditions for a short-term model.

Figure 1 shows a graph with five nodes, where four nodes are associated with an energy equivalent reservoir, Southeast (SE), South(S), Northeast (NE), North (N), as well as one transshipment node (T). The load at each node must be supplied by hydropower or thermal power plants at that node or by power flow from another node. Available power flow capacity is limited and represented by an arc in the graph. Load shortage at one node results in load curtailment which is penalized by a convex cost function.

A more detailed description of the long-term operational planning problem as well as system parameters is given in Section 5 of [25].

4.2 Stochastic Inflow Models

Historical inflow data has been provided by the Brazilian system operator ONS (Operador Nacional do Sistema Eltrico) for the years 1931 to 2008. In [25], the authors find that a first-order periodic autoregressive model of the log-inflows provides a good description of

the data. This geometric periodic autoregressive (GPAR) model is defined as the first order autoregressive (periodical) process after making the transformation $\ln b_t$ (the logarithm of vector b_t is taken componentwise). That is

$$\ln b_t = \mu_t + \Phi_t \ln b_{t-1} + \varepsilon_t,\ \varepsilon_t \sim \mathcal{N}(0, \Sigma_t),\ t = 2, ..., T, \tag{4.18}$$

with vectors μ_t and matrices Φ_t, Σ_t having periodical behavior of 12 months. Here, the error terms (random vectors), ε_t, are assumed to form a stagewise independent process. The GPAR model not only deals with negative inflows but also reduces symmetry of the error distribution.

Unfortunately, the GPAR model destroys convexity of the cost-to-go functions with respect to $\xi_t = b_t$ and hence cannot be used with TS-SDDP method. On the other hand for the MC-SDDP method convexity with respect to ξ_t is not essential, and the model can be used.

To linearize, the log-linear autoregressive model, the authors in [25] propose a first-order approximation, which leads to

$$b_t = e^{\varepsilon_t} \circ (a_t + \Phi_t b_{t-1}),\ \varepsilon_t \sim \mathcal{N}(0, \Sigma_t),\ t = 2, ..., T, \tag{4.19}$$

with vector a_t and matrices Φ_t and Σ_t having periodical behavior of 12 months. The exponents e^{ε_t} are taken componentwise and $a \circ b$ denotes the componentwise (Hadamard) product of vectors a and b. Let us refer to this inflow model as the Shapiro-Tekaya-Costa-Soares (STCS) model.

Both models were fitted and tested on 78 years of data of historical inflows into the four energy equivalent reservoirs of the Brazilian interconnected power system. As both models exhibit approximately the same model fit measured by R^2 (STCS: 78.5%, GPAR: 78.2%)[3], let us compare them based on the inflow scenarios they generate.

Figure 2 shows multiple fan charts to visualize the percentiles of the unconditional distribution for each month based on 1000 simulated sample paths. Fan charts of the historical data with 78 sample paths are shown on the left. The differences between simulated data and original data are subtle. The STCS model seems to exhibit slightly heavier tails than the data, e.g., inflows above 30'000 megawatt months into the South EER between May and November occur more often than in the data. The GPAR model does not exhibit any apparent deviations. Nevertheless, both models produce inflow scenarios with approximately the same characteristics as the historical data.

[3]To obtain a joint R^2, residual sum of squares (RSS) and total sum of squares (TSS) are taken jointly over all four inflow dimensions, whereby TSS is computed using a separate mean for each inflow dimension.

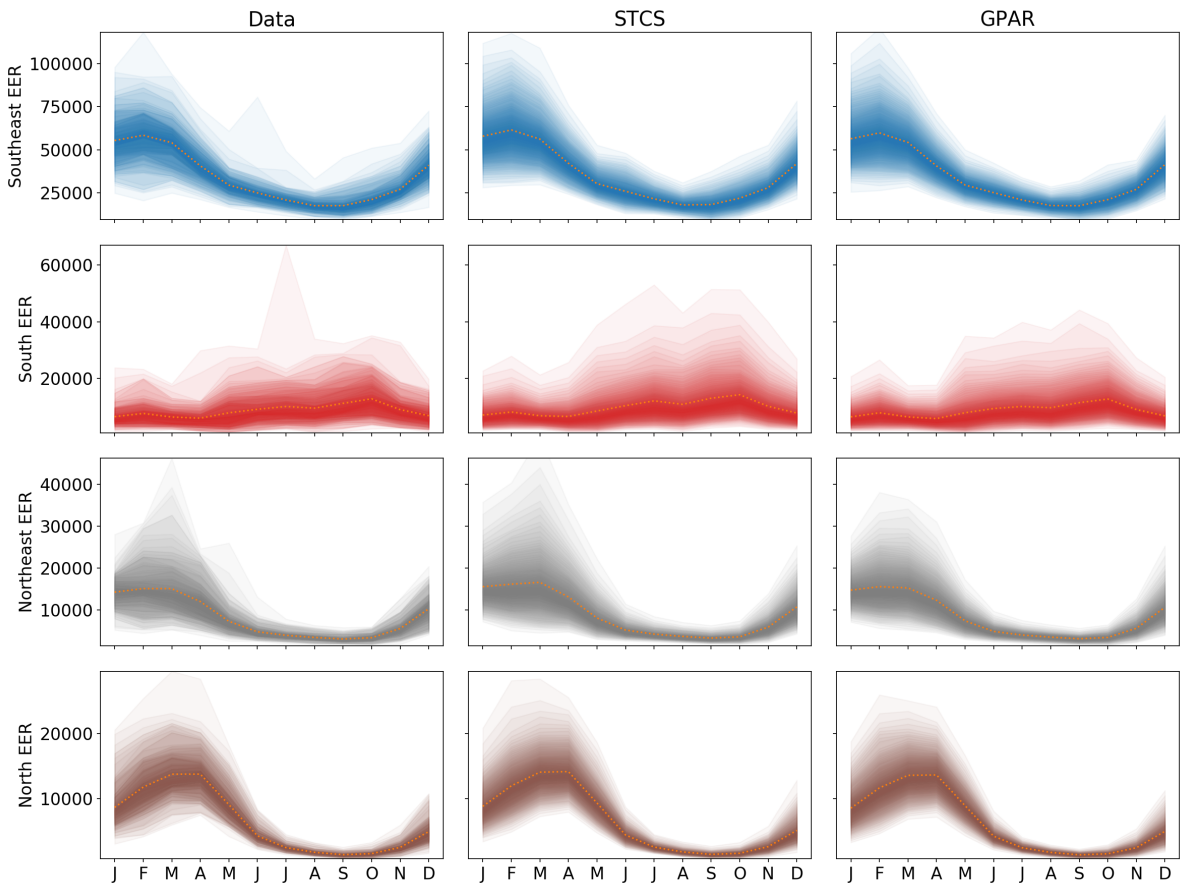

Figure 2: Monthly inflow percentiles with different stochastic inflow models.

Figure 3 shows Q-Q plots comparing the distribution of historical annual mean inflows against simulated annual mean inflows for both inflow models. The STCS model exhibits the largest deviation from the diagonal. Most notably, the means of the Southeast and the South EER are greater than their historical means, which implies that the STCS model overestimates the longterm inflow into these reservoirs. The GPAR model deviates less, although the upper quantiles of the distribution of annual means of the Northeast EER appear to be a little lower than in the data.

5 Computational Experiments

5.1 Implementation

For our computational experiments we used the same parameters as in [25] as well as the same computational setup. All models were formulated and solved in Python using the QUASAR stochastic optimization software [10].

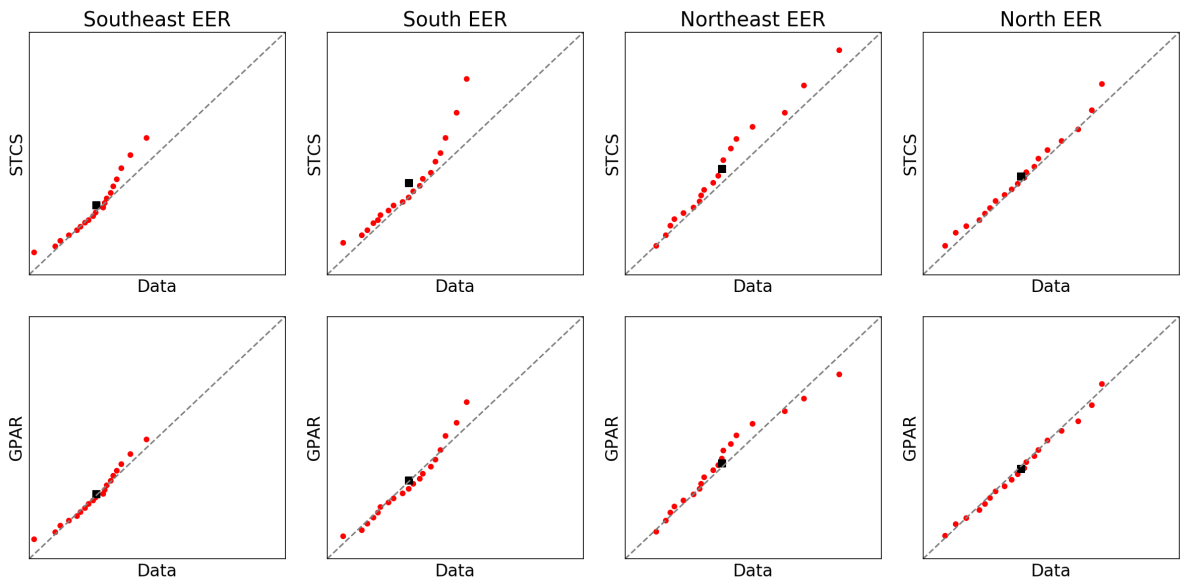

Figure 3: Q-Q-plots of annual mean inflows with different stochastic inflow models.

Unless stated otherwise, we used 100 different random realizations ($K_t = K = 100,\ t = 1, ..., T$, center points) per stage, whereby MC lattices were trained with 100k stochastic inflow scenarios. Convergence of TS-SDDP and MC-SDDP was tracked over 1000 iterations. To allow for pairwise comparison, both methods received the same sequence of inflow scenarios during the forward passes. As in [25], the optimization problem is solved with a planning horizon of 120 months.

5.2 Simulation Analysis

In our first analysis, we study the gap between lower and upper bound using TS-SDDP and MC-SDDP, respectively. Although MC-SDDP does not provide a lower bound for the STCS inflow model, we can study the policy performance. The convergence of the lower and upper bounds of MC-SDDP and TS-SDDP for five independent optimization runs is shown in Figure 4. For better visual comparison, the estimates of the upper bound were smoothed using a moving average with a right-centered window of size 100. Additionally, we aggregated the bounds for all five runs in the bottom right subfigure. Here, to smooth the upper bounds, we fitted a convex, monotone decreasing function by solving the following quadratic program,

$$\min_{\hat{y}_i \in \mathbb{R}, i=1,\dots,N} \left\{ \sum_{i=1}^{N} (y_i - \hat{y}_i)^2 : \hat{y}_i - \hat{y}_{i-1} \geq \hat{y}_{i-1} - \hat{y}_{i-2},\ \hat{y}_i \leq \hat{y}_{i-1} \right\},$$

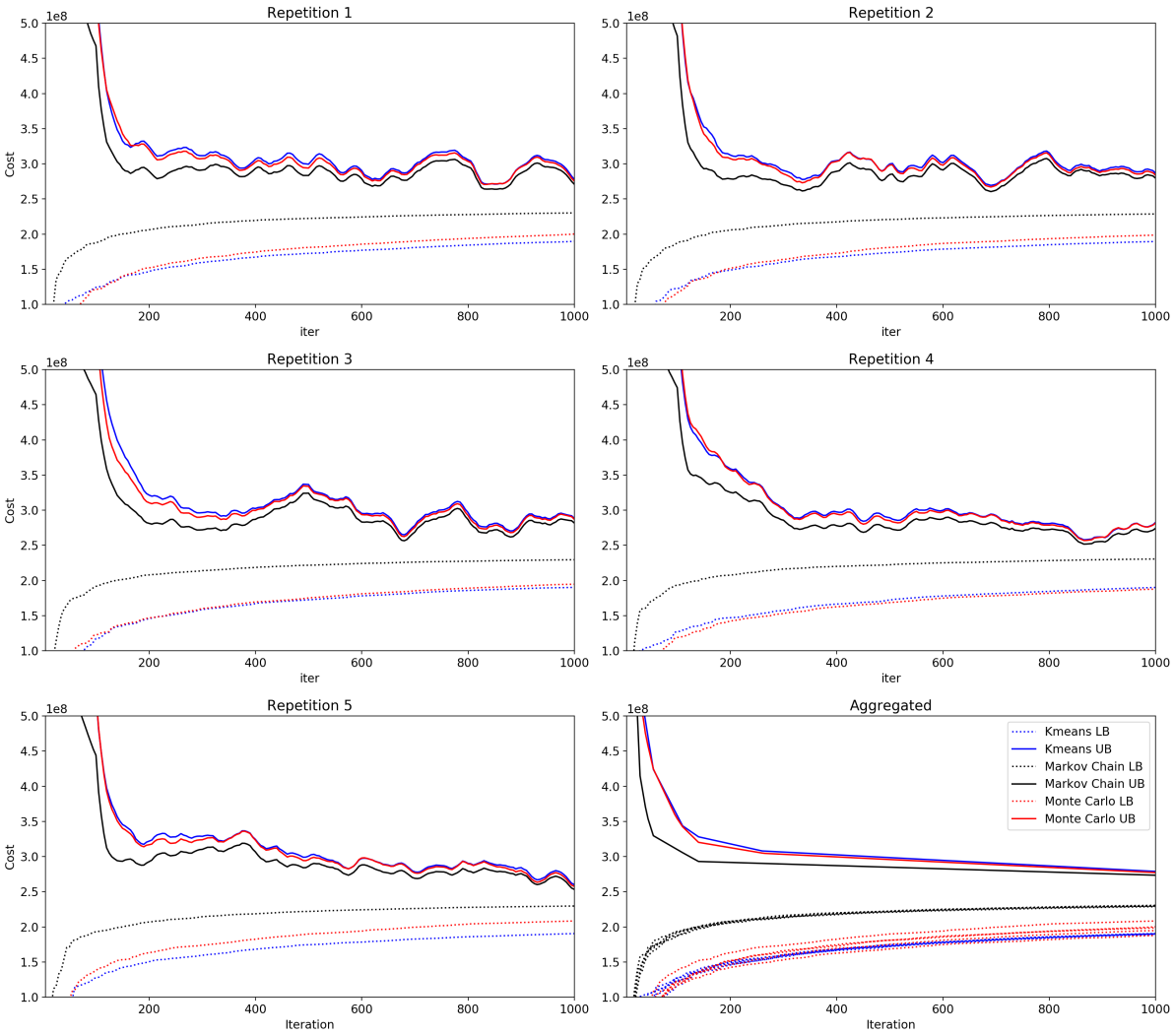

Figure 4: TS-SDDP and MC-SDDP bounds for STCS inflow model.

with y_i as random realization of the upper bound at iteration i.

It can be seen that, in the considered case, MC-SDDP consistently provides a policy that is slightly better than TS-SDDP, and that its lower bound converges faster to the upper bound. Although the Markov chain lower bound is not theoretically valid for the "true" problem, it is closer to the upper bound than the Monte Carlo lower bound. Using TS-SDDP with K-means to partition the error distribution is only marginally different from using a Monte Carlo sample, although we can see from the bottom right subfigure that the Monte Carlo lower bound exhibits significant variability.

We conjecture that the reason for the better performance of the MC-SDDP over TS-SDDP lies in the way in which the state space is being explored by either algorithm. TS-SDDP introduces additional equality constraints into the model to ensure stagewise independence,

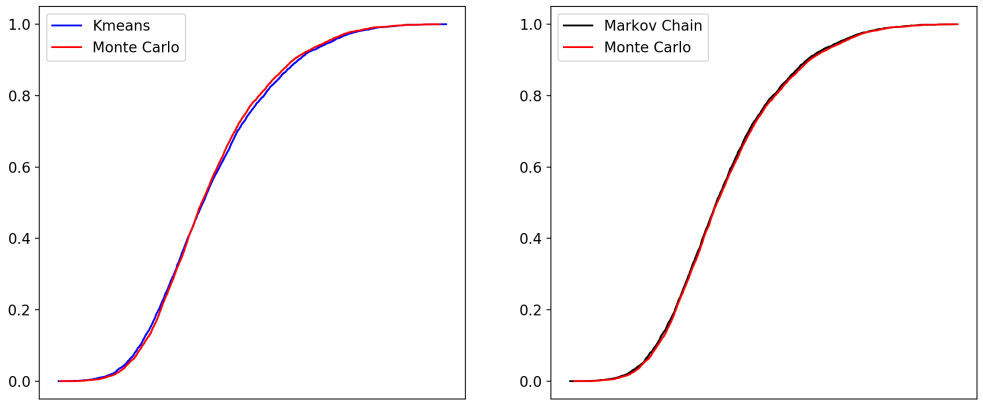

Figure 5: Cumulative cost distribution of the optimal policy after 1000 iterations.

which adds dimensions to the cost-to-go functions that are not under the control of the optimal policy, but are selected randomly during forward simulation of the decision process. Since the approximate cost-to-go functions always underestimates the true cost-to-go functions, sample decisions during forward passes are always explorative. Reservoir states can thus be explored efficiently by adding information about the cost-to-go in regions of the state space with the highest prospect of a cost decrease. Inflow states however cannot be explored but are sampled at random. Random sampling of inflow states might therefore not lead into inflow states with the largest information gain. Although the MC-SDDP does not resolve this problem, generating an MC lattice of inflow states in advance at least ensures that a diverse set of inflow states is selected beforehand. The numerical results indicate that this strategy leads to faster convergence and a slightly better policy in a shorter amount of time.

In addition to the sample average, let us study the cost distribution obtained with each approach. Figure 5 shows the cumulative log-cost distribution of the optimal policy after 1000 iterations for the same random realizations that were used during the forward passes for all five repetitions. Although Kmeans and Monte Carlo sampling produce nearly the same sample averages, we see that Kmeans leads to a higher probability of high cost scenarios. The cost distribution obtained using MC-SDDP by contrast dominates the cost distribution obtained using TS-SDDP with Monte Carlo sampling.

For our next analysis, we were interested at the convergence behavior of the MC-SDDP under the GPAR model, when varying the number of nodes at each stage from 10 to 1000. The convergence of the lower and upper bounds is shown in Figure 6. The graphs illustrate that a larger number of nodes leads to a better policy as well as a lower bound that is closer to the true upper bound, because it leads to a fine partitioning of the outcome space, which

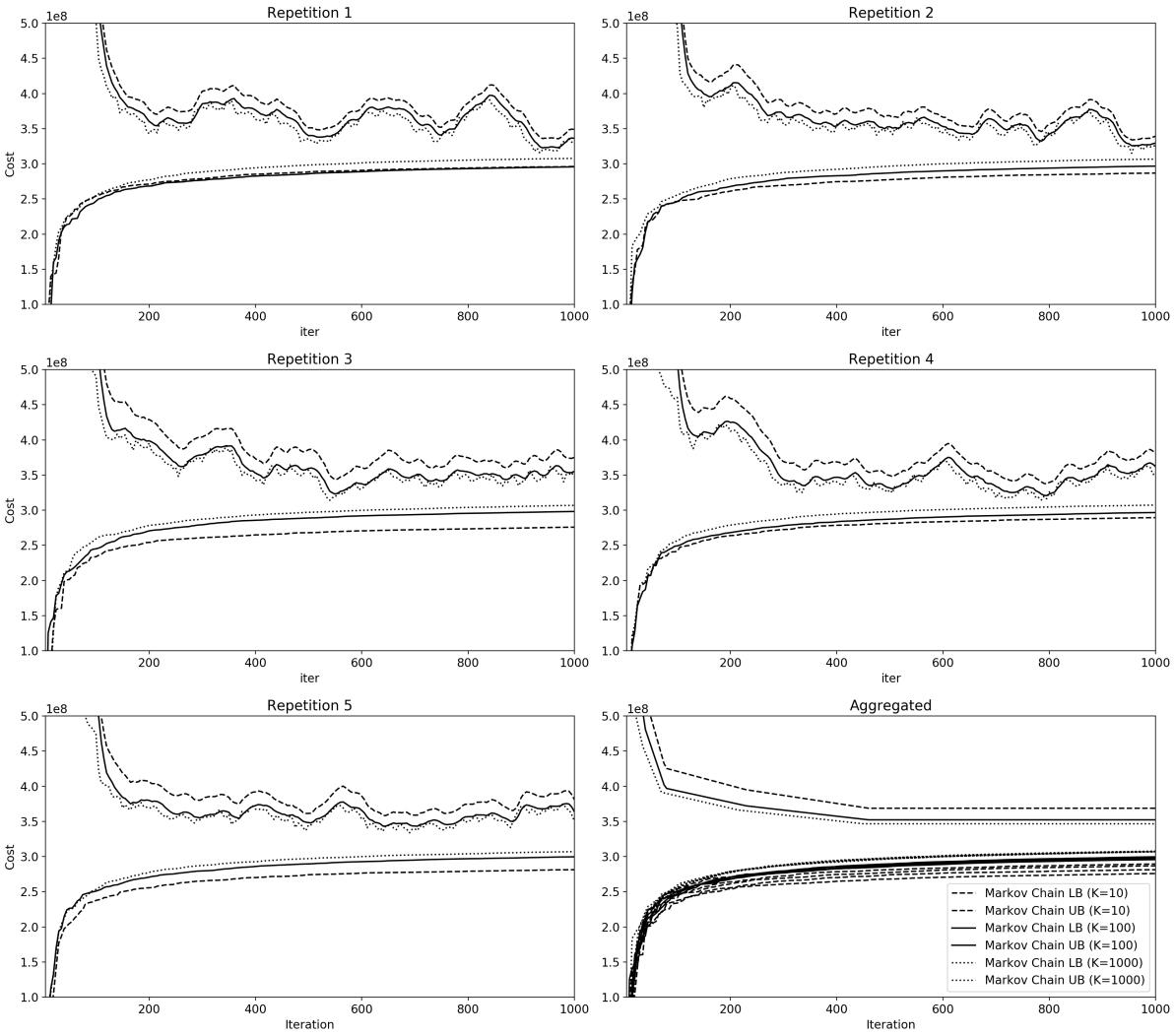

Figure 6: MC-SDDP convergence for the GPAR inflow model with varying number of nodes in the Markov chain.

is supported by Remark 2.1. Nevertheless, as in Figure 4, the gap between lower and upper bound does not close within 1000 iterations.

The convergence of the lower bounds of MC-SDDP and TS-SDDP for an AV@R with $\alpha = 0.1$ and $\lambda = 0.25$ is shown in Figure 7. Interestingly, the Monte Carlo version of TS-SDDP now provides a higher lower bound than MC-SDDP for the same number of random outcomes per stage. Only after increasing the number of partitions by a factor of 10 to $K = 1000$ nodes per stage, the Markov chain lower bound is again higher than the Monte Carlo lower bound.

Also, the relative increase in the lower bound when increasing the number of lattice nodes per stage is much greater with AV@R than without, so is the variability of the lower

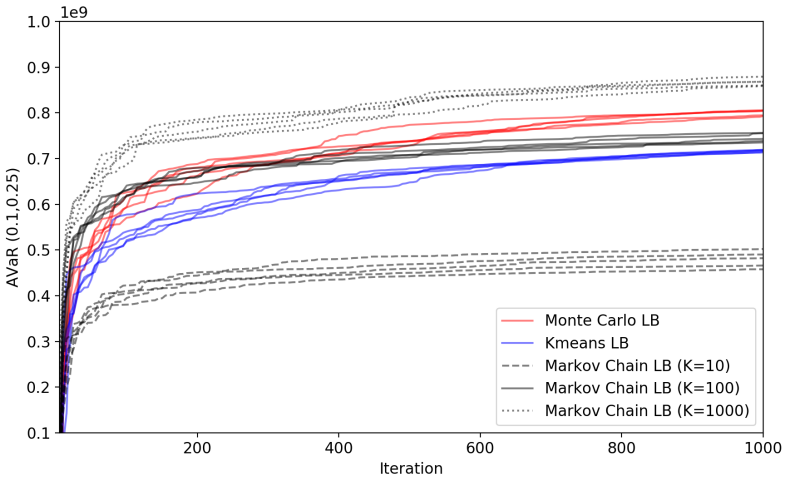

Figure 7: TS-SDDP and MC-SDDP lower bounds using AV@R in the objective.

bound. In lack of a valid upper bound, effect on the quality of the simulated policy however is unknown and subject to future research.

5.3 Analysis with Historical Data

In addition to numerical benchmarks based on Monte Carlo simulation, we were also interested in the quality of the policies when implemented on the historical inflow data that was used to fit the stochastic inflow models.

For this we separated the 78 years of inflow data into 68 overlapping 120-month horizons and implemented the approximate optimal policy after 1000 iterations on the resulting 68 sample paths. To avoid bias, we only report the cost of the intermediate 12 months, since we can expect initialization bias as well as end-of-horizon bias. The total cost accrued during the respective year then serves an estimate of the steady-state cost. We will refer to the corresponding distribution as the empirical cost distribution.

Figures 8 and 9 show the empirical cost distribution for different solution approaches and inflow models. While Figure 8 shows empirical cost for 68 historical time series averaged over five repetitions, Figure 9 shows empirical cost for five repetitions but averaged over 68 historical time series. We can make the following observations.

- MC-SDDP leads to significantly lower cost than TS-SDDP with STCS inflows ($p < 0.001$, $t(\text{n}=68) = 4.15$).
- MC-SDDP with GPAR inflows leads to significantly lower cost than TS-SDDP with STCS inflows ($p < 0.01$, $t(\text{n}=68) = 3.34$).

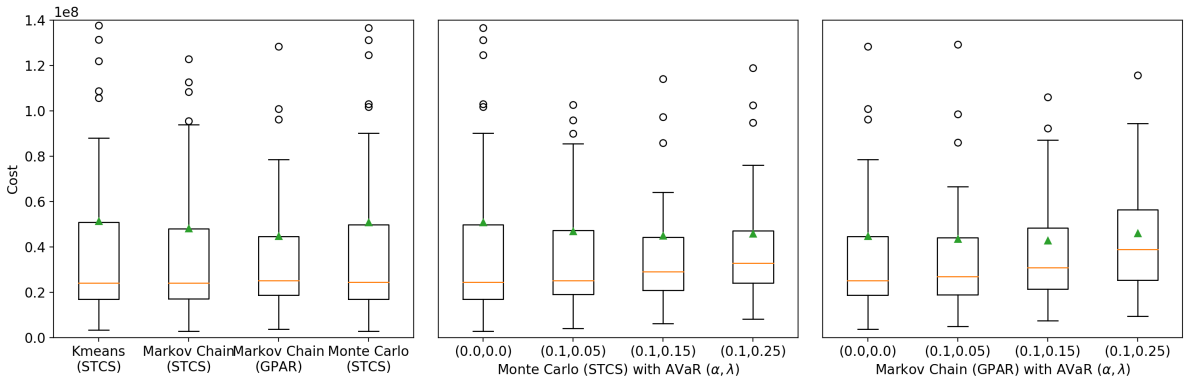

Figure 8: Empirical distribution of steady-state cost for different solution methods and inflow models (left), for TS-SDDP with STCS inflows for different AV@R (center), for MC-SDDP with GPAR inflows (right)

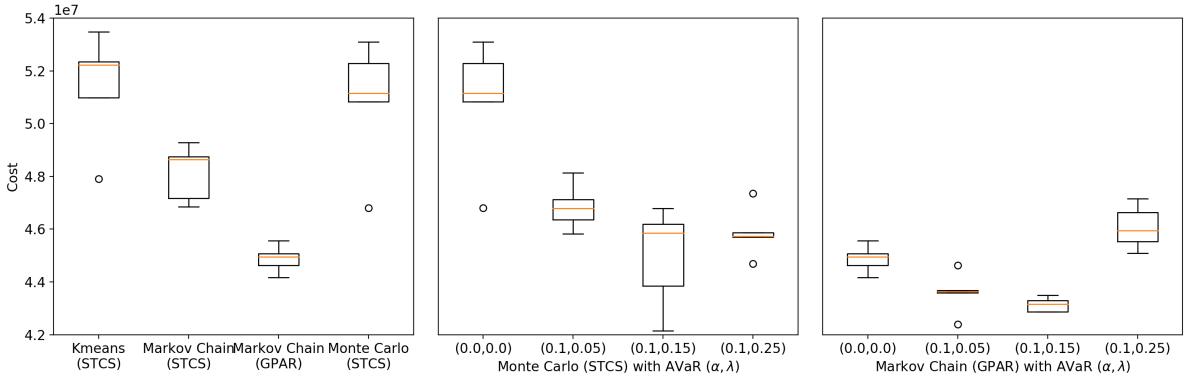

Figure 9: Average empirical cost over five repetitions for different solution methods and inflow models (left), for TS-SDDP with STCS inflows for different AV@R (center), for MC-SDDP with GPAR inflows (right)

- The decrease in cost for using an AV@R with $\alpha = 0.1$ and $\lambda = 0.05$ is already significant for TS-SDDP with STCS inflows ($p < 0.05$, $t(\text{n=68}) = 2.56$).
- The decrease in cost for using an AV@R with $\alpha = 0.1$ and $\lambda = 0.15$ is not significant for MC-SDDP with GPAR inflows ($p = 0.52$, $t(\text{n=68}) = 0.64$).
- The difference in cost between MC-SDDP with GPAR inflows but without AV@R and TS-SDDP with STCS inflows with AV@R with $\alpha = 0.1$ and $\lambda = 0.15$ are not significant ($p = 0.93$, $t(\text{n=68}) = -0.09$).

We interpret these results as follows. The optimal policy with the GPAR model results in better decisions than the optimal policy with the STCS model. In contrast to the GPAR policy, the STCS policy can be significantly improved by using an AV@R formulation. As

already shown in Figures 2 and 3, the STCS model overestimates the probability of high inflow outcomes, which may be the reason for the worse quality of the resulting optimal policy when implemented on real data. Using an AV@R has robustified the STCS policy against the misspecified inflow model and, given the right parameters, leads to a policy that is not significantly different from the optimal GPAR policy.

To strengthen this result, let us take a look at the implemented decision under the optimal policy obtained using TS-SDDP with the STCS model and MC-SDDP with the GPAR model, with and without using an AV@R in the objective.

Figure 10 shows reservoir content curves for the Southeast energy equivalent reservoir for each policy. The top subfigure shows the reservoir content over 68 years under each policy without using an AV@R formulation. The bottom panel also shows the reservoir content, but this time TS-SDDP used an AV@R formulation with $\alpha = 0.1$ and $\lambda = 0.05$. To better visualize the difference between the two policies, the middle two panels additionally show the monthly average reservoir content before and after introducing an AV@R.

Clearly, usage of AV@R has increased the amount of energy held in the reservoir. Both policies now produce nearly the same reservoir content curves, which demonstrates that introducing an AV@R not only minimizes risk, but can result in a cost decrease if the stochastic model does not correctly describe the true process that underlies the variability of the data process.

6 Conclusion

In this paper, we compare two different approaches to model time-dependent randomness in stochastic dual dynamic programming: introducing additional variables and constraints into the problem formulation to model right-hand side uncertainty as a linear time series model (TS-SDDP) versus discretizing the data process to a Markov chain (MC-SDDP).

We compare both approaches using case study data from the Brazilian interconnected power system, where randomness enters the problem through uncertainty in hydrological inflows into a number of interconnected reservoirs. We study SDDP bounds as well as policy quality using simulated and empirical data. For this, we fit a geometric periodic autoregressive (GPAR) model to historical inflow data and compare this to a linearized version from the literature (STCS) that meets the requirements of TS-SDDP.

We find that MC-SDDP produces tighter lower bounds as well as consistently better policies for the STCS inflow model. The gap is largest in the early iterations, but vanishes as the number of iterations increases.

We also compare solution quality using the original data and find that MC-SDDP leads

to significantly lower cost than TS-SDDP, irrespective of the inflow model. We also find that the GPAR model yields significantly lower cost than the STCS model, which roots back to the STCS model overestimating high inflow outcomes. The effect can be mitigated by using an AV@R formulation, which appears to robustify the solution under the STCS model. This finding provides empirical support for the ability of the AV@R to provide distributional robustness.

Of course, these are empirical observations for the studied case where data vectors are low-dimensional. Also the MC-SDDP lower bounds are valid for the Markov Chain discretization and are not guaranteed for the "true" continuous model. Future work should therefore study the performance of TS-SDDP and MC-SDDP for high-dimensional data processes with no or only low dependence between components of the data vectors.

Acknowledgements

The authors would like to thank Joari Paulo da Costa and Filipe Goulart Cabral from ONS for preparing the case study data that was used for the numerical experiments as well as for their valuable input and discussion.

References

[1] Artzner, P., Delbaen, F., Eber, J.-M., Heath, D., Coherent measures of risk, *Mathematical Finance* 9, 203–228, 1999.

[2] Aloise, D., Deshpande, A., Hansen, P., Popat, P., NP-hardness of Euclidean sum-of-squares clustering, *Machine Learning* 75:2, 245–248, 2009.

[3] Arvaniditis, N.V. , Rosing, J., Composite representation of a multireservoir hydroelectric power system, *IEEE Transactions on Power Apparatus and Systems* 89:2, 319-326, 1970

[4] Bally, V., Pagès, G., A quantization algorithm for solving multidimensional discrete-time optimal stopping problems, *Bernoulli* 6:9, 1003–1049, 2003.

[5] Birge, J.R., Louveaux, F.V., *Introduction to Stochastic Programming*, Springer, New York, 1997.

[6] Bonnans, J.F., Cen, Z., Christel, T., Energy contracts management by stochastic programming techniques, *Annals of Operations Research* 200, 199–222, 2012.

[7] Girardeau, P., Leclere, V., Philpott, A.B., On the convergence of decomposition methods for multistage stochastic convex programs, *Mathematics of Operations Research* 40:1, 130–145, 2014.

[8] Infanger, G., Morton, D.P., Cut sharing for multistage stochastic linear programs with interstage dependency, *Mathematical Programming* 75:2, 241–256, 1996.

[9] Löhndorf, N., Wozabal, D., Minner, S., Optimizing trading decisions of hydro storage plants using approximate dual dynamic programming, *Operations Research* 61:4, 810–823, 2013.

[10] Löhndorf, N., *QUASAR Optimization Software 2.1*, available at http://quantego.com, 2017.

[11] Löhndorf, N., Wozabal, D., Indifference pricing of natural gas storage contracts, *Working Paper*, 2017.

[12] Lund, R., Shao, Q., Basawa, I., Parsimonious perdiodic time series modeling, *Australian and New Zealand Journal of Statistics* 48, 33-47, 2006.

[13] Maceira, M. E. P., Duarte, V. S., Penna, D. D. J., Moraes, L. A. M., Melo, A. C. G., Ten years of application of stochastic dual dynamic programming in official and agent studies in brazil-description of the newave program, *16th PSCC-Power Syst Comp Conf*, 2008.

[14] Maceira, M.E.P., Damazio, J.M., Use of par(p) model in the stochastic dual dynamic programming optimization scheme used in the operation planning of the brazilian hydropower system, *Probability in the Engineering and Informational Sciences* 20, 143-156, 2006.

[15] Pagès, G., Printems, J., Optimal quantization for finance: from random vectors to stochastic processes, *Handbook of Numerical Analysis* 15, 595–648, 2008.

[16] Penna, D.D.J., Maceira, M.E.P, Damzio, J.M., Selective sampling applied to long-term hydrothermal generation planning, *17th PSCC-Power Syst Comp Conf*, 2011.

[17] Pereira, M. V. F., Optimal stochastic operations scheduling of large hydroelectric systems, *International Journal of Electrical Power and Energy Systems* 11:3, 161–169, 1989.

[18] Pereira, M.V.F., Pinto, L.M.V.G., Multi-stage stochastic optimization applied to energy planning, *Mathematical Programming* 52, 359–375, 1991.

[19] Pflug, G. C., Pichler, A., From empirical observations to tree models for stochastic optimization: convergence properties, *SIAM Journal on Optimization* 26:3, 1715–1740, 2016.

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

[21] Philpott, A.B. and de Matos, V.L., Dynamic sampling algorithms for multi-stage stochastic programs with risk aversion, *European Journal of Operational Research*, 218:2, 470–483, 2012.

[22] Ruszczyński, A., Shapiro, A., Optimization of convex risk functions, *Mathematics of Operations Research* 31, 433–452, 2006.

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

[24] Shapiro, A., Analysis of stochastic dual dynamic programming method, *European Journal of Operational Research* 209:1, 63–72, 2011.

[25] Shapiro, A., Tekaya, W., da Costa, J.P., Soares, M.P, Risk neutral and risk averse stochastic dual dynamic programming method, *European Journal of Operational Research* 224:2, 375–391, 2013.

[26] Shapiro, A., Dentcheva, D., Ruszczyński, A., *Lectures on Stochastic Programming: Modeling and Theory*, 2nd ed, SIAM, Philadelphia, 2014.

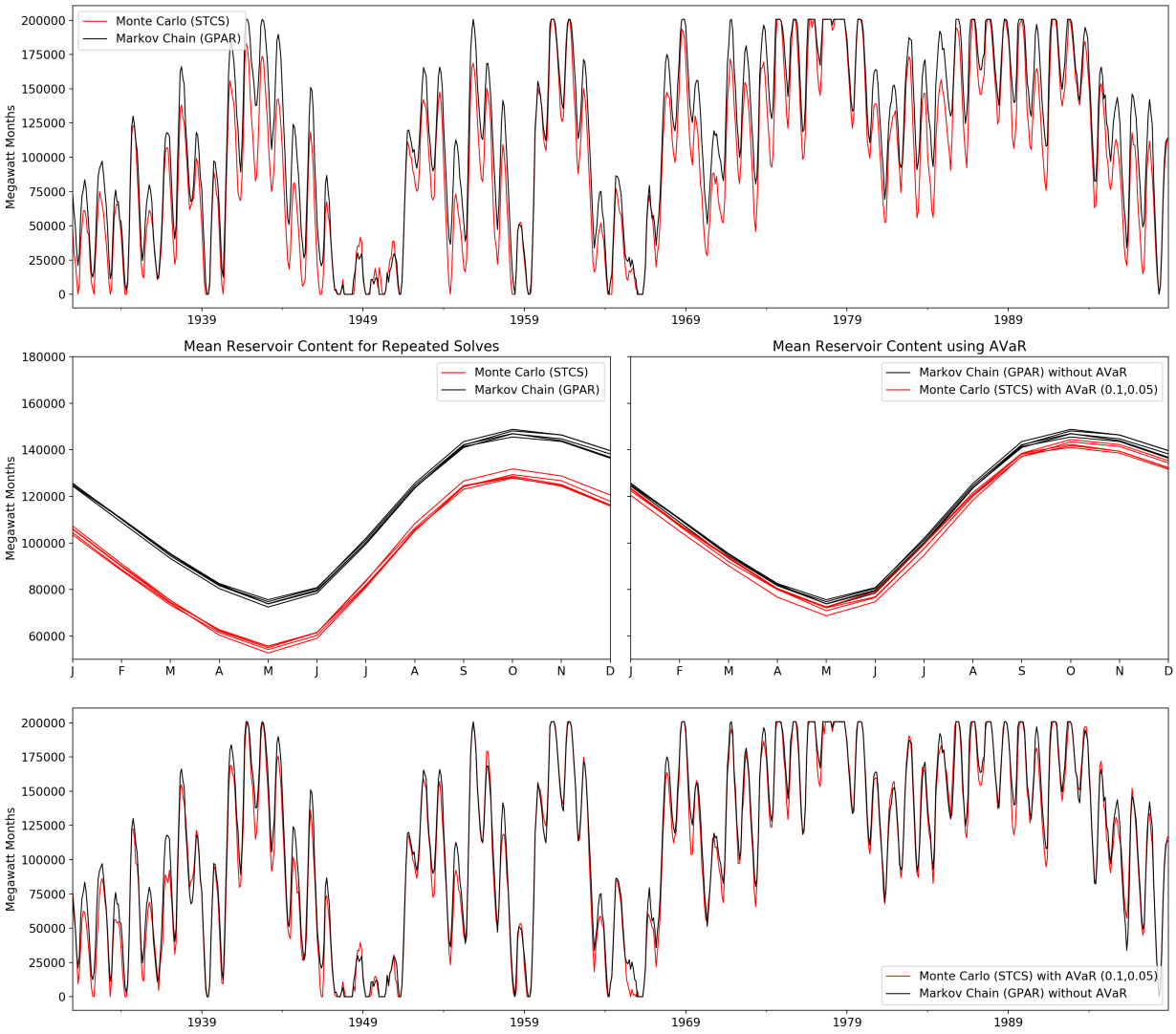

Figure 10: Southeast reservoir contents over 68 years from TS-SDDP (STCS) and MC-SDDP (GPAR) with and without AV@R