Risk neutral and risk averse Stochastic Dual Dynamic Programming method

Alexander Shapiro and Wajdi Tekaya
School of Industrial and Systems Engineering,
Georgia Institute of Technology,
Atlanta, Georgia 30332-0205, USA

Joari Paulo da Costa and Murilo Pereira Soares
ONS - Operador Nacional do Sistema Elétrico
Rua da Quitanda, 196, Centro
Rio de Janeiro, RJ, 20091-005, Brasil

Abstract. In this paper we discuss risk neutral and risk averse approaches to multistage (linear) stochastic programming problems based on the Stochastic Dual Dynamic Programming (SDDP) method. We give a general description of the algorithm and present computational studies related to planning of the Brazilian interconnected power system.

Key Words: multistage stochastic programming, dynamic equations, Stochastic Dual Dynamic Programming, sample average approximation, risk averse, average value-at-risk, case studies

1 Introduction

In this paper we discuss risk neutral and risk averse approaches to multistage (linear) stochastic programming problems based on the Stochastic Dual Dynamic Programming (SDDP) method. We give a general description of the algorithm and present computational studies related to operation planning of the Brazilian interconnected power system. The Brazilian interconnected power system is a large scale system planned and constructed considering the integrated utilization of the generation and transmission resources of all agents and the use of inter-regional energy interchanges, in order to achieve cost reduction and reliability in power supply.

The power generation facilities as of December 2010 are composed of more that 200 power plants with installed capacity greater than 30 MW, owned by 108 public and private companies, called "Agents" – 57 Agents own 141 hydro power plants located in 14 large basins, 69 with large reservoirs (monthly regulation or above), 68 run-of-river plants and four pumping stations. Considering only the National Interconnected System (SIN), without captive self-producers, the installed capacity reaches 171.1 GW in 2020, with increment of 61.6 GW over the 2010 installed capacity of 109.6 GW. Hydropower accounts for 56% of the expansion (34.8 GW), while biomass and wind power account for 25% of the expansion (15.4 GW).

The main transmission grid is operated and expanded in order to achieve safety of supply and system optimization. The inter-regional and inter-basin transmission links allow interchanges of large blocks of energy between areas making it possible to take advantage of the hydrological diversity between river basins. The main transmission grid system has 100.7 $\times 10^3$km of lines above 230 kV, owned by 66 Agents, and is planned to reach 142.2 $\times 10^3$km in 2020, an increment of 41.2%, mainly with the interconnection of the projects in the Amazonian region.

It is worth noting that the installed capacity of hydro plants corresponds to 79.1% of the December 2010 total installed capacity, but its relative position should diminish to 71.0% in 2020. Nevertheless, considering the Domestic Electricity Supply, the hydro power supremacy will continue in 2020, standing for 73.4% of the total, as compared to 80.6% in Dec., 2010 (including imports). This relative reduction is offset by a strong penetration of biomass and wind generation. In this context, renewable sources maintain a high participation in Electricity Supply Matrix (87,7%) (see [2]).

1.1 Operation planning problem

The purpose of hydrothermal system operation planning is to define an operation strategy which, for each stage of the planning period, given the system state at the beginning of the stage, produces generation targets for each plant. The usual objective is to minimize the expected value of the total cost along the planning period, so as to meet requirements on the continuity of energy supply subject to feasibility constraints. The operation costs comprise fuel costs, purchases from neighboring systems and penalties for failure in load supply. This is referred to as the risk neutral approach: the total cost is optimized on average, and for a particular realization of the random data process the costs could be much higher than their average values. Risk averse approaches, on the other hand, aims at finding a compromise between minimizing the average cost and trying to control the upper limit of the costs for some possible realizations of the data set at every stage of the process. The risk averse approach will be discussed later in this article.

The hydrothermal operating planning can be seen as a decision problem under uncertainty because of unknown variables such as future inflows, demand, fuel costs and equipment availability. The existence of large multi-year regulating reservoirs makes the operation planning a multistage optimization problem; in the Brazilian case it is usual to consider a planning horizon of 5 years on

monthly basis. The existence of multiple interconnected hydro plants and transmission constraints characterizes the problem as large scale. Moreover, because the worth of energy generated in a hydro plant cannot be measured directly as a function of the plant state alone but rather in terms of expected fuel savings from avoided thermal generation, the objective function is also non-separable, [4].

In summary, the Brazilian hydro power operation planning problem is a multistage (60 stages), large scale (more than 200 power plants, of which 141 are hydro plants), nonlinear and nonseparable stochastic optimization problem. This setting far exceeds the computer capacity to solve it with adequate accuracy in reasonable time frame. The standard approach to solve this problem is to resort to a chain of models considering long, mid and short term planning horizon in order to be able to tackle the problem in a reasonable computational time. For the long-term problem, it is usual to consider an approximate representation of the system, the so-called aggregate system model, a composite representation of a multireservoir hydroelectric power system, proposed by Arvaniditis and Rosing [1], that aggregates all hydro power plants belonging to a homogeneous hydrological region in a single equivalent energy reservoir, and solve the resulting much smaller problem. The major components of the aggregate system model are: the equivalent energy reservoir model and the total energy inflow (controllable and uncontrollable), see Figure 1. The energy storage capacity of the equivalent energy reservoir can be estimated as the energy that can be produced by the depletion of the reservoirs of a system, provided a simplified operating rule that approximates the actual depletion policy.

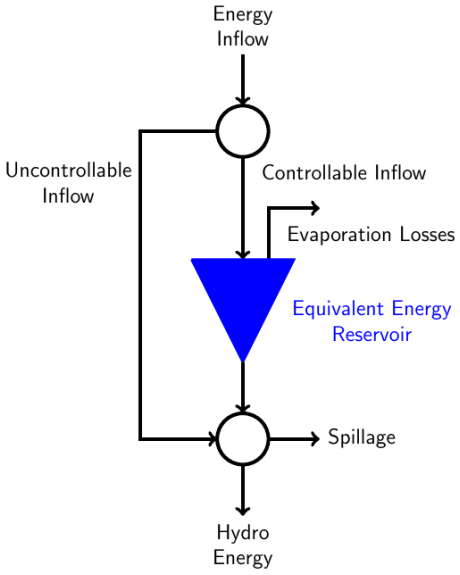

Figure 1: Aggregate system model

For the Brazilian interconnected power system it is usual to consider four energy equivalent reservoirs, one in each one of the four interconnected main regions, SE, S, N and NE. For this simplified problem, one can use the SDDP approach and obtain the cost-to-go functions for each of the stages of the planning period. The resulting policy obtained with the aggregate representation can then be further be refined, so as to provide decisions for each of the hydro and thermal power plants. This can be done by solving the mid-term problem, considering a planning horizon up to a few months and individual representation of the hydro plants with boundary conditions (the expected cost-to-go functions) given by the solution of the long-term problem. This is the approach

nowadays used for solving the long and mid term hydrothermal power planning in the Brazilian interconnected power system.

1.2 Article structure

This article is organized as follows. In the Introduction we gave a brief introduction of the hydrothermal operation planning problem with emphasis on the Brazilian power system case. In the next section we discuss general methodological aspects of the stochastic programming approach with focus on the Stochastic Dual Dynamic Programming (SDDP) algorithm (introduced by Pereira and Pinto [4],[5]). The performance of the SDDP algorithm is illustrated with the aid of some numerical experiments. Computational aspects of the algorithm convergence and solution stability are presented and discussed.

Up to this point, the standard risk neutral approach was implemented for planning of the Brazilian power system. The energy rationing that took place in Brazil in the period 2001/2002 raised the question of whether a policy that is based on a criterion of minimizing the expected cost is a valid one when it comes to meet the day-to-day supply requirements. As a consequence, a shift towards a risk averse criterion is underway, so as to enforce the continuity of load supply. In this way, a proposal of including the risk averse approach within the framework of SDDP algorithm is presented and discussed.

2 Mathematical formulation and modeling

Mathematical algorithms compose the core of the Energy Operation Planning Support System. The objective is to compute an operation strategy which controls the operation costs, over a planning period of time, in a reasonably optimal way. This leads to formulation of (linear) large scale multistage stochastic programming problems which in a generic form (nested formulation) can be written as

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

Components of vectors c_t, b_t and matrices A_t, B_t are modeled as random variables forming the stochastic data process $\xi_t = (c_t, A_t, B_t, b_t)$, $t = 2, ..., T$, with $\xi_1 = (c_1, A_1, b_1)$ being deterministic (not random). By $\xi_{[t]} = (\xi_1, ..., \xi_t)$ we denote history of the data process up to time t.

In order to set the hydrothermal operating planning problem within this framework one can proceed as follows. Considering the aggregate representation of the hydroplants, the energy conservation equation for each equivalent energy reservoir n can be written as

$$\mathrm{SE}_{t,n} = \mathrm{SE}_{t-1,n} + \mathrm{CE}_{t,n} - \mathrm{GH}_{t,n} - \mathrm{SP}_{t,n}. \tag{2.2}$$

That is, the stored energy (SE) at the end of each stage (start of the next stage) is equal to the initial stored energy plus controllable energy inflow (CE) minus total hydro generated energy (GH) and losses (SP) due to spillage, evaporation, etc.

At each stage, the net subsystem load L, given by the remaining load after discounting the uncontrolled energy inflow from the total load, has to be met by the total hydro, the sum of all thermal generation belonging to system n, given by the set NT_n, and the net interconnection energy

flow (NF) to each subsystem. In other words, the energy balance equation for system n is

$$\mathrm{GH}_{t,n} + \sum_{j \in \mathrm{NT}_n} \mathrm{GT}_{t,j} + \mathrm{NF}_{t,n} = L_{t,n}. \tag{2.3}$$

Note that this equation is always feasible (i.e., the problem has complete recourse) due to the inclusion of a dummy thermal plant with generation capacity equal to the demand and operation cost that reflects the social cost of not meeting the energy demand (deficit cost).

Constraints $B_t x_{t-1} + A_t x_t = b_t$ are obtained writing

$$x_t = (\mathrm{SE}, \mathrm{GH}, \mathrm{GT}, \mathrm{SP}, \mathrm{NF})_t^\top, \;\; b_t = (\mathrm{CE}, \mathrm{L})_t^\top, \;\; c_t = (0, 0, \mathrm{CT}, 0, 0)_t^\top,$$

$$A_t = \begin{pmatrix} I & I & \Delta & 0 & 0 \\ 0 & I & I & 0 & I \end{pmatrix}, \;\; B_t = \begin{pmatrix} -I & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

where $\Delta = \{\delta_{n,j} = 1 \text{ for all } j \in \mathrm{NT}_n \text{ and zero else}\}$, I and 0 are identity and null matrices, respectively, of appropriate dimensions and the components of CT are the unit operation cost of each thermal plant and penalty for failure in load supply. Note that hydroelectric generation costs are assumed to be zero. Physical constraints on variables like limits on the capacity of the equivalent reservoir, hydro and thermal generation, transmission capacity and so on are taken into account with constraints on x_t. More details can be found in [11] and [4].

It is often assumed in numerical approaches to solving multistage problems of the form (2.1) that the number of realizations (scenarios) of the data process is *finite*, and this assumption is essential in the implementations and analysis of the applied algorithms. In many applications, however, this assumption is quite unrealistic. In forecasting models (such as ARIMA) the errors are typically modeled as having continuous (say normal or log-normal) distributions. So one of the relevant questions is what is the meaning of the introduced discretizations of the corresponding stochastic process. We do not make the assumption of finite number of scenarios, instead the following assumptions will be made. These assumptions (below) are satisfied in the applications relevant for the Brazilian power system generation.

We make the basic assumption that the random data process is *stagewise independent*, i.e., random vector ξ_{t+1} is independent of $\xi_{[t]} = (\xi_1, ..., \xi_t)$ for $t = 1, ..., T-1$. In some cases across stages dependence can be dealt with by adding state variables to the model. In particular, the following construction is relevant for the considered applications. Suppose that only the right hand side vectors b_t are across stage dependent, while other parameters of the problem form a stagewise independent process (in particular could be deterministic). We are interested in cases where, for physical reasons, components of vectors b_t cannot be negative. Suppose that random vectors b_t follow p-th order autoregressive process with multiplicative error terms:

$$b_t = \varepsilon_t \bullet (\mu + \phi_1 b_{t-1} + ... + \phi_p b_{t-p}), \;\; t = 2, ..., T, \tag{2.4}$$

with vector μ and matrices $\phi_1, ..., \phi_p$ are estimated from the data. Here $\varepsilon_2, ..., \varepsilon_T$ are independent of each other error vectors such that with probability one their components are nonnegative and have expected value one, and $a \bullet b$ denotes the term by term (Hadamard) product of two vectors. The *multiplicative* error term model is considered to ensure that realizations of the random process b_t have nonnegative values.

The autoregressive process (2.4) can be formulated as a first order autoregressive process

$$\begin{bmatrix} b_t \\ b_{t-1} \\ b_{t-2} \\ \cdots \\ b_{t-p+1} \end{bmatrix} = \begin{bmatrix} \varepsilon_t \\ \mathbf{1} \\ \mathbf{1} \\ \cdots \\ \mathbf{1} \end{bmatrix} \bullet \left(\begin{bmatrix} \mu \\ 0 \\ 0 \\ \cdots \\ 0 \end{bmatrix} + \begin{bmatrix} \phi_1 & \phi_2 & \cdots & \phi_{p-1} & \phi_p \\ I & 0 & \cdots & 0 & 0 \\ 0 & I & \cdots & 0 & 0 \\ & & \cdots & & \\ 0 & 0 & \cdots & I & 0 \end{bmatrix} \begin{bmatrix} b_{t-1} \\ b_{t-2} \\ b_{t-3} \\ \cdots \\ b_{t-p} \end{bmatrix} \right), \tag{2.5}$$

where $\mathbf{1}$ is vector of ones and I is the identity matrix of an appropriate dimension. Denote by z_t the column vector in the left hand side of (2.5), and by ϵ_t, M and Φ the respective terms in the right hand side of (2.5). Then (2.5) can be written as

$$z_t = \epsilon_t \bullet (M + \Phi z_{t-1}),\ t = 2, ..., T. \tag{2.6}$$

Consequently the feasibility equations of problem (2.1) can be written as

$$z_t - \epsilon_t \bullet \Phi z_{t-1} = \epsilon_t \bullet M,\ B_t x_{t-1} + A_t x_t = b_t,\ x_t \geq 0,\ \ t = 2, ..., T. \tag{2.7}$$

Therefore by replacing x_t with (x_t, z_t), and considering the corresponding data process ξ_t formed by random elements of c_t, A_t, B_t and error vectors ε_t, $t = 2, ..., T$, we transform the problem to the stagewise independent case. The obtained problem is still linear with respect to the decision variables x_t and z_t.

We consider the following approach to solving the multistage problem (2.1). First, a (finite) scenario tree is generated by randomly sampling from the original distribution and then the constructed problem is solved by the *Stochastic Dual Dynamic Programming* (SDDP) algorithm. There are three levels of approximations in that setting. The first level is modelling. The inflows are viewed as seasonal time series and modelled as a periodic auto-regressive (PAR(p)) process. Any such modelling involves inaccuracies - autoregressive parameters should be estimated, errors distributions are not precise, etc. We will refer to an optimization problem based on a current time series model as the *"true"* problem.

The "true" model involves data process ξ_t, $t = 1, ..., T$, having *continuous* distributions. Since the corresponding expectations (multidimensional integrals) cannot be computed in a closed form, one needs to make a discretization of the data process ξ_t. So a sample $\tilde{\xi}_t^1, ..., \tilde{\xi}_t^{N_t}$, of size N_t, from the distribution of the random vector ξ_t, $t = 2, ..., T$, is generated. These samples generate a scenario tree with the total number of scenarios $N = \prod_{t=2}^{T} N_t$, each with equal probability $1/N$. Consequently the true problem is approximated by the so-called *Sample Average Approximation* (SAA) problem associated with this scenario tree. This corresponds to the second level of approximation in the current system. Note that in such sampling approach the stagewise independence of the data process is preserved in the constructed SAA problem.

If we measure computational complexity, of the true problem, in terms of the number of scenarios required to approximate true distribution of the random data process with a reasonable accuracy, the conclusion is rather pessimistic. In order for the optimal value and solutions of the SAA problem to converge to their true counterparts all sample sizes $N_2, ..., N_T$ should tend to infinity. Furthermore, available estimates of the sample sizes N_t required for a first stage solution of the SAA problem to be ε-optimal for the true problem, with a given confidence (probability), sums up to a total number of scenarios N which grows as $O(\varepsilon^{-2(T-1)})$ with decrease of the error level $\varepsilon > 0$ (cf., [8, section 5.8.2]). This indicates that from the point of view of the number of scenarios, complexity of multistage programming problems grows exponentially with increase of the number of stages. In other words even with a moderate number of scenarios per stage, say each $N_t = 100$,

the total number of scenarios N quickly becomes astronomically large with increase of the number of stages. Therefore a constructed SAA problem can be solved only approximately.

The SDDP method suggests a computationally tractable approach to solving SAA, and hence the "true", problems, and can be viewed as the third level of approximation in the current system. A theoretical analysis (cf., [10, section 5.2.2]) and numerical experiments indicate that the SDDP method can be a reasonable approach to solving multistage stochastic programming problems when the number of state variables is small even if the number of stages is relatively large.

3 Generic description of the SDDP algorithm

In this section we give a general description of the SDDP algorithm applied to the SAA problem. Suppose that N_t, $t = 2, ..., T$, points generated at every stage of the process. Let $\xi_t^j = (c_{tj}, A_{tj}, B_{tj}, b_{tj})$, $j = 1, .., N_t$, $t = 2, ..., T$, be the generated points. As it was already mentioned the total number of scenarios of the SAA problem is $N = \prod_{t=1}^{T} N_t$ and can be very large.

3.1 Backward step of the SDDP algorithm

Let $\bar{x}_t$, $t = 1, ..., T-1$, be trial points (we can, and eventually will, use more than one trial point at every stage of the backward step, an extension to that will be straightforward), and $\mathcal{Q}_t(x_{t-1})$ be the (expected value) cost-to-go functions of dynamic programming equations associated with the multistage problem (2.1) (see, e.g., [8, section 3.1.2]). Note that because of the stagewise independence assumption, the cost-to-go functions $\mathcal{Q}_t(x_{t-1})$ do not depend on the data process. Furthermore, let $\mathfrak{Q}_t(\cdot)$ be a current approximation of $\mathcal{Q}_t(\cdot)$ given by the maximum of a collection of *cutting* planes

$$\mathfrak{Q}_t(x_{t-1}) = \max_{k \in \mathfrak{I}_t} \left\{ \alpha_{tk} + \beta_{tk}^{\mathsf{T}} x_{t-1} \right\}, \ t = 1, ..., T-1. \tag{3.1}$$

At stage $t = T$ we solve N_T problems

$$\operatorname*{Min}_{x_T \in \mathbb{R}^{n_T}} c_{Tj}^{\mathsf{T}} x_T \ \text{ s.t. } B_{Tj}\bar{x}_{T-1} + A_{Tj}x_T = b_{Tj}, \ x_T \geq 0, \ j = 1, ..., N_T. \tag{3.2}$$

Recall that $Q_{Tj}(\bar{x}_{T-1})$ is equal to the optimal value of problem (3.2) and that subgradients of $Q_{Tj}(\cdot)$ at $\bar{x}_{T-1}$ are given by $-B_{Tj}^{\mathsf{T}}\pi_{Tj}$, where π_{Tj} is a solution of the dual of (3.2). Therefore for the cost-to-go function $\mathcal{Q}_T(x_{T-1})$ we can compute its value and a subgradient at the point $\bar{x}_{T-1}$ by averaging the optimal values of (3.2) and the corresponding subgradients. Consequently we can construct a supporting plane to $\mathcal{Q}_T(\cdot)$ and add it to collection of supporting planes of $\mathfrak{Q}_T(\cdot)$. Note that if we have several trial points at stage $T-1$, then this procedure should be repeated for each trial point and we add each constructed supporting plane.

Now going one stage back let us recall that $Q_{T-1,j}(\bar{x}_{T-2})$ is equal to the optimal value of problem

$$\operatorname*{Min}_{x_{T-1} \in \mathbb{R}^{n_{T-1}}} c_{T-1,j}^{\mathsf{T}} x_{T-1} + \mathcal{Q}_T(x_{T-1}) \ \text{ s.t. } B_{T-1,j}\bar{x}_{T-2} + A_{T-1,j}x_{T-1} = b_{T-1,j}, \ x_{T-1} \geq 0. \tag{3.3}$$

However, function $\mathcal{Q}_T(\cdot)$ is not available. Therefore we replace it by $\mathfrak{Q}_T(\cdot)$ and hence consider problem

$$\operatorname*{Min}_{x_{T-1} \in \mathbb{R}^{n_{T-1}}} c_{T-1,j}^{\mathsf{T}} x_{T-1} + \mathfrak{Q}_T(x_{T-1}) \ \text{ s.t. } B_{T-1,j}\bar{x}_{T-2} + A_{T-1,j}x_{T-1} = b_{T-1,j}, \ x_{T-1} \geq 0. \tag{3.4}$$

Recall that $\mathfrak{Q}_T(\cdot)$ is given by maximum of affine functions (see (3.1)). Therefore we can write problem (3.4) in the form

$$\begin{array}{ll} \underset{x_{T-1}\in\mathbb{R}^{n_{T-1}},\theta\in\mathbb{R}}{\text{Min}} & c_{T-1,j}^{\mathsf{T}}x_{T-1}+\theta \\ \text{s.t.} & B_{T-1,j}\bar{x}_{T-2}+A_{T-1,j}x_{T-1}=b_{T-1,j},\ x_{T-1}\geq 0 \\ & \theta\geq\alpha_{Tk}+\beta_{Tk}^{\mathsf{T}}x_{T-1},\ k\in\mathfrak{I}_T. \end{array} \tag{3.5}$$

Consider the optimal value, denoted $\underline{Q}_{T-1,j}(\bar{x}_{T-2})$, of problem (3.5), and let $\pi_{T-1,j}$ be the partial vector of an optimal solution of the dual of problem (3.5) corresponding to the constraint $B_{T-1,j}\bar{x}_{T-2}+A_{T-1,j}x_{T-1}=b_{T-1,j}$, and let

$$\underline{\mathcal{Q}}_{T-1}(\bar{x}_{T-2}):=\frac{1}{N_{T-1}}\sum_{j=1}^{N_{T-1}}\underline{Q}_{T-1,j}(\bar{x}_{T-2})$$

and

$$g_{T-1}=-\frac{1}{N_{T-1}}\sum_{j=1}^{N_{T-1}}B_{T-1,j}^{\mathsf{T}}\pi_{T-1,j}.$$

Consequently we add the corresponding affine function to collection of $\mathfrak{Q}_{T-1}(\cdot)$. And so on going backward in t.

3.2 Forward step of the SDDP algorithm

The computed approximations $\mathfrak{Q}_2(\cdot),...,\mathfrak{Q}_T(\cdot)$ (with $\mathfrak{Q}_{T+1}(\cdot)\equiv 0$ by definition) and a feasible first stage solution $\bar{x}_1$ can be used for constructing an implementable policy as follows. For a realization $\xi_t=(c_t,A_t,B_t,b_t)$, $t=2,...,T$, of the data process, decisions $\bar{x}_t$, $t=1,...,T$, are computed recursively going forward with $\bar{x}_1$ being the chosen feasible solution of the first stage problem, and $\bar{x}_t$ being an optimal solution of

$$\underset{x_t}{\text{Min}}\, c_t^{\mathsf{T}}x_t+\mathfrak{Q}_{t+1}(x_t)\ \text{ s.t. } A_tx_t=b_t-B_t\bar{x}_{t-1},\ x_t\geq 0, \tag{3.6}$$

for $t=2,...,T$. These optimal solutions can be used as trial decisions in the backward step of the algorithm. Note that $\bar{x}_t$ is a function of $\bar{x}_{t-1}$ and ξ_t, i.e., $\bar{x}_t$ is a function of $\xi_{[t]}=(\xi_1,...,\xi_t)$, for $t=2,...,T$. That is, policy $\bar{x}_t=\bar{x}_t(\xi_{[t]})$ is nonanticipative and by the construction satisfies the feasibility constraints for every realization of the data process. Thus this policy is implementable and feasible. If we restrict the data process to the generated sample, i.e., we consider only realizations $\xi_2,...,\xi_T$ of the data process drawn from scenarios of the SAA problem, then $\bar{x}_t=\bar{x}_t(\xi_{[t]})$ becomes an implementable and feasible policy for the corresponding SAA problem. On the other hand, if we draw samples from the true (original) distribution, this becomes an implementable and feasible policy for the true problem.

Since the policy $\bar{x}_t=\bar{x}_t(\xi_{[t]})$ is feasible, the expectation

$$\mathbb{E}\left[\sum_{t=1}^{T}c_t^{\mathsf{T}}\bar{x}_t(\xi_{[t]})\right] \tag{3.7}$$

gives an upper bound for the optimal value of the corresponding multistage problem. That is, if we take this expectation over the true probability distribution of the random data process, then the above expectation (3.7) gives an upper bound for the optimal value of the true problem. On

the other hand, if we restrict the data process to scenarios of the SAA problem, each with equal probability $1/N$, then the expectation (3.7) gives an upper bound for the optimal value of the SAA problem conditional on the sample used in construction of the SAA problem. Of course, if the constructed policy is optimal, then the expectation (3.7) is equal to the optimal value of the corresponding problem.

The *forward* step of the SDDP algorithm consists in generating M random realizations (scenarios) of the data process and computing the respective optimal values

$$\vartheta_j := \sum_{t=1}^{T} c_{tj}^{\mathsf{T}} \bar{x}_{tj}, \quad j = 1, ..., M. \tag{3.8}$$

That is, ϑ_j is the value of the corresponding policy for the realization $\xi_1, \xi_2^j, ..., \xi_T^j$ of the data process. As such, ϑ_j is an unbiased estimate of expected value of that policy, i.e., $\mathbb{E}[\vartheta_j] = \mathbb{E}\left[\sum_{t=1}^{T} c_t^{\mathsf{T}} \bar{x}_t(\xi_{[t]})\right]$. The forward step has two functions. First, some (all) of computed solutions $\bar{x}_{tj}$ can be used as trial points in the next iteration of the backward step of the algorithm. Second, these solutions can be employed for constructing a statistical upper bound for the optimal value of the corresponding multistage program (true or SAA depending on from what distribution the sample scenarios were generated).

Let

$$\tilde{\vartheta}_M := \frac{1}{M} \sum_{j=1}^{M} \vartheta_j \ \text{ and } \ \tilde{\sigma}_M^2 = \frac{1}{M-1} \sum_{j=1}^{M} (\vartheta_j - \tilde{\vartheta}_M)^2$$

be the respective sample mean and sample variance of the computed values ϑ_j. Since ϑ_j is an unbiased estimate of the expected value of the constructed policy, we have that $\tilde{\vartheta}_M$ is also an unbiased estimate of the expected value of that policy. By invoking the Central Limit Theorem we can say that $\tilde{\vartheta}_M$ has an approximately normal distribution provided that M is reasonably large. This leads to the following (approximate) $(1-\alpha)$-confidence upper bound for the value of that policy

$$\mathfrak{u}_{\alpha,M} := \tilde{\vartheta}_M + z_\alpha \frac{\tilde{\sigma}_M}{\sqrt{M}}. \tag{3.9}$$

Here $1-\alpha \in (0,1)$ is a chosen confidence level and $z_\alpha = \Phi^{-1}(1-\alpha)$, where $\Phi(\cdot)$ is the cdf of standard normal distribution. That is, with probability approximately $1-\alpha$ the expected value of the constructed policy is less than the upper bound $\mathfrak{u}_{\alpha,M}$. Since the expected value (3.7) of the constructed policy is bigger than or equal to the optimal value of the considered multistage problem, we have that $\mathfrak{u}_{\alpha,M}$ also gives an upper bound for the optimal value of the multistage problem with confidence at least $1-\alpha$. Note that the upper bound $\mathfrak{u}_{\alpha,M}$ can be used for the SAA or the true problem depending on from what distribution the sampled scenarios were generated.

Since $\mathfrak{Q}_t(\cdot)$ is the maximum of cutting planes of the cost-to-go function $\mathcal{Q}_t(\cdot)$ we have that

$$\mathcal{Q}_t(\cdot) \geq \mathfrak{Q}_t(\cdot), \ t = 2, ..., T. \tag{3.10}$$

Therefore the optimal value of the problem computed at a backward step of the algorithm, gives a lower bound for the optimal value of the considered SAA problem. This lower bound is deterministic (i.e., is not based on sampling) if applied to the corresponding SAA problem. As far as the true problem is concerned, recall that the optimal value of the true problem is greater than or equal to the expectation of the optimal value of the SAA problem. Therefore on average this is also a lower bound for the optimal value of the true problem. On the other hand, the upper bound $\mathfrak{u}_{\alpha,M}$ is a function of generated scenarios and thus is stochastic even for considered (fixed) SAA problem.

This upper bound may vary for different sets of random samples, in particular from one iteration to the next of the forward step of the algorithm.

The SDDP algorithm within the risk neutral framework is summarized in Algorithm 1.

Algorithm 1 Risk neutral SDDP

Require: $\{\mathfrak{Q}_t^0\}_{t=2,\ldots,T+1}$(Init. Lower approx.) and $\epsilon > 0$(accuracy)

Initialize: $i \leftarrow 0$, $\bar{z} = \infty$ (Upper bound), $\underline{z} = -\infty$ (Lower bound)
while $\bar{z} - \underline{z} > \epsilon$ **do**
  Sample M scenarios: $\left\{\{c_{tk}, A_{tk}, B_{tk}, b_{tk}\}_{2\le t\le T}\right\}_{1\le k\le M}$
  (Forward step)
  for $k = 1 \to M$ **do**
    for $t = 1 \to T$ **do**
      $\bar{x}_t^k \leftarrow \arg\min_{x_t\in\mathbb{R}^{n_t}} \left\{ \begin{array}{l} c_{tk}^\top x_t + \mathfrak{Q}_{t+1}^i(x_t) : \\ B_{tk}x_{t-1} + A_{tk}x_t = b_{tk}, x_t \ge 0 \end{array} \right\}$
    end for
    $\vartheta_k \leftarrow \sum_{t=1}^T c_{tk}^\mathsf{T}\bar{x}_t^k$
  end for
  (Upper bound update)
  $\overline{z} \leftarrow \tilde{\vartheta}_M + z_{\alpha/2}\frac{\tilde{\sigma}_M}{\sqrt{M}}$
  (Backward step)
  for $k = 1 \to M$ **do**
    for $t = T \to 2$ **do**
      for $j = 1 \to N_t$ **do**
        $\left[\widetilde{Q}_{tj}(\bar{x}_{t-1}^k), \widetilde{\pi}_{tj}^k\right] \leftarrow \min_{x_t\in\mathbb{R}^{n_t}} \left\{ \begin{array}{l} c_{tj}^\top x_t + \mathfrak{Q}_{t+1}^i(x_t) : \\ B_{tj}\bar{x}_{t-1} + A_{tj}x_t = b_{tj}, x_t \ge 0 \end{array} \right\}$
      end for
      $\widetilde{\mathcal{Q}}_t(\bar{x}_{t-1}^k) := \frac{1}{N_t}\sum_{j=1}^{N_t}\widetilde{Q}_{t,j}(\bar{x}_{t-1}^k)$; $\tilde{g}_t^k := -\frac{1}{N_t}\sum_{j=1}^{N_t}\tilde{\pi}_{tj}^k\widetilde{B}_{t,j}$
      $\mathfrak{Q}_t^{i+1} \leftarrow \{x_{t-1} \in \mathfrak{Q}_t^i : -\tilde{g}_t^k x_{t-1} \ge \widetilde{\mathcal{Q}}_t(\bar{x}_{t-1}^k) - \tilde{g}_t^k\bar{x}_{t-1}^k\}$
    end for
  end for
  (Lower bound update)
  $\underline{z} \leftarrow \min_{x_1\in\mathbb{R}^{n_1}}\left\{c_1^\top x_1 + \mathfrak{Q}_2(x_1) : A_1x_1 = b_1, x_1 \ge 0\right\}$
  $i \leftarrow i + 1$
end while

4 Adaptive risk averse approach

In formulation (2.1) the expected value $\mathbb{E}\left[\sum_{t=1}^T c_t^\mathsf{T}x_t\right]$ of the total cost is minimized subject to the feasibility constraints. That is, the total cost is optimized (minimized) *on average*. Since the costs are functions of the random data process, they are random and hence are subject to random

perturbations. For a particular realization of the data process these costs could be much bigger than their average (i.e., expectation) values. We will refer to the formulation (2.1) as *risk neutral* as opposed to *risk averse* approaches which we will discuss below.

The goal of a risk averse approach is to avoid large values of the costs for some possible realizations of the data process at *every* stage of the considered time horizon. One such approach will be to maintain constraints $c_t^{\mathsf{T}} x_t \leq \theta_t$, $t = 1, ..., T$, for chosen upper levels θ_t and *all* possible realizations of the data process. However, trying to enforce these upper limits under any circumstances could be unrealistic and infeasible. One may try to relax these constraints by enforcing them with a high (close to one) probability. However, introducing such so-called *chance constraints* can still result in infeasibility and moreover is very difficult to handle numerically. So we consider here penalization approaches. That is, at every stage the cost is penalized while exceeding a specified upper limit.

In a simple form this leads to a risk averse formulation where costs $c_t^{\mathsf{T}} x_t$ are penalized by $\Phi_t[c_t^{\mathsf{T}} x_t - \theta_t]_+$, with θ_t and $\Phi_t \geq 0$, $t = 2, ..., T$, being chosen constants. That is, the costs $c_t^{\mathsf{T}} x_t$ are replaced by functions $f_t(x_t) = c_t^{\mathsf{T}} x_t + \Phi_t[c_t^{\mathsf{T}} x_t - \theta_t]_+$ in the objective of the problem (2.1). The additional penalty terms represent the penalty for exceeding the upper limits θ_t. An immediate question is how to choose constants θ_t and Φ_t. It could be noted that in that approach the upper limits θ_t are fixed (chosen a priori) and not adapted to a current realization of the random process. Let us observe that optimal solutions of the corresponding risk averse problem will be not changed if the penalty term at t-th stage is changed to $\theta_t + \Phi_t[c_t^{\mathsf{T}} x_t - \theta_t]_+$ by adding the constant θ_t. Now if we adapt the upper limits θ_t to a realization of the data process by taking these upper limits to be $(1 - \alpha_t)$-quantiles of $c^{\mathsf{T}} x_t$ *conditional* on observed history $\xi_{[t-1]} = (\xi_1, ..., \xi_{t-1})$ of the data process, we end up with penalty terms given by $\mathsf{AV@R}_{\alpha_t}$ with $\alpha_t = 1/\Phi_t$. Recall that the *Average Value-at-Risk* of a random variable[1] Z is defined as

$$\mathsf{AV@R}_\alpha[Z] = \mathsf{V@R}_\alpha(Z) + \alpha^{-1}\mathbb{E}\left[Z - \mathsf{V@R}_\alpha(Z)\right]_+, \tag{4.1}$$

with $\mathsf{V@R}_\alpha(Z)$ being the (say left side) $(1 - \alpha)$-quantile of the distribution of Z, i.e., $\mathsf{V@R}_\alpha(Z) = F^{-1}(1 - \alpha)$, where $F(\cdot)$ is the cumulative distribution function (cdf) of the random variable Z.

This leads to the following nested risk averse formulation of the corresponding multistage problem (cf., [7])

$$\underset{\substack{A_1 x_1 = b_1 \\ x_1 \geq 0}}{\mathrm{Min}} c_1^{\mathsf{T}} x_1 + \rho_{2|\xi_1}\left[\min_{\substack{B_2 x_1 + A_2 x_2 = b_2 \\ x_2 \geq 0}} c_2^{\mathsf{T}} x_2 + \cdots + \rho_{T|\xi_{[T-1]}}\Big[\min_{\substack{B_T x_{T-1} + A_T x_T = b_T \\ x_T \geq 0}} c_T^{\mathsf{T}} x_T\Big]\right]. \tag{4.2}$$

Here $\xi_2, ..., \xi_T$ is the random process (formed from the random elements of the data c_t, A_t, B_t, b_t), $\mathbb{E}\left[Z|\xi_{[t-1]}\right]$ denotes the conditional expectation of Z given $\xi_{[t-1]}$, $\mathsf{AV@R}_{\alpha_t}\left[Z|\xi_{[t-1]}\right]$ is the conditional analogue of $\mathsf{AV@R}_{\alpha_t}[Z]$ given $\xi_{[t-1]}$, and

$$\rho_{t|\xi_{[t-1]}}[Z] = (1 - \lambda_t)\mathbb{E}\left[Z|\xi_{[t-1]}\right] + \lambda_t \mathsf{AV@R}_{\alpha_t}\left[Z|\xi_{[t-1]}\right], \tag{4.3}$$

with $\lambda_t \in [0, 1]$ and $\alpha_t \in (0, 1)$ being chosen parameters. In formulation (4.2) the penalty terms $\alpha_t^{-1}\left[c_t^{\mathsf{T}} x_t - \mathsf{V@R}_\alpha(c_t^{\mathsf{T}} x_t)\right]_+$ are conditional, i.e., are adapted to the random process by the optimization procedure. Therefore we refer to the risk averse formulation (4.2) as *adaptive*.

It is also possible to give the following interpretation of the adaptive risk averse formulation (4.2). It is clear from the definition (4.1) that $\mathsf{AV@R}_\alpha[Z] \geq \mathsf{V@R}_\alpha(Z)$. Therefore $\rho_{t|\xi_{[t-1]}}[Z] \geq \varrho_{t|\xi_{[t-1]}}[Z]$, where

$$\varrho_{t|\xi_{[t-1]}}[Z] = (1 - \lambda_t)\mathbb{E}\left[Z|\xi_{[t-1]}\right] + \lambda_t \mathsf{V@R}_{\alpha_t}\left[Z|\xi_{[t-1]}\right]. \tag{4.4}$$

[1] In some publications the Average Value-at-Risk is called the Conditional Value-at-Risk and denoted $\mathsf{CV@R}_\alpha$. Since we deal here with conditional $\mathsf{AV@R}_\alpha$, it will be awkward to call it conditional $\mathsf{CV@R}_\alpha$.

If we replace $\rho_{t|\xi_{[t-1]}}[Z]$ in the risk averse formulation (4.2) by $\varrho_{t|\xi_{[t-1]}}[Z]$, we will be minimizing the weighted average of means and $(1-\alpha)$-quantiles, which will be a natural way of dealing with the involved risk. Unfortunately such formulation will lead to a nonconvex and computationally intractable problem. This is one of the main reasons for using $\mathsf{AV@R}_\alpha$ instead of $\mathsf{V@R}_\alpha$ in the corresponding risk averse formulations. It is possible to show that in a certain sense $\mathsf{AV@R}_\alpha(\cdot)$ gives a best possible upper convex bound for $\mathsf{V@R}_\alpha(\cdot)$, [3].

It also could be mentioned that the (conditional) risk measures of the form (4.4) are not the only ones possible. We will discuss below some other examples of risk measures.

4.1 Risk averse SDDP method

With a relatively simple additional effort the SDDP algorithm can be applied to risk averse problems of the form (4.2). Let ρ_t, $t = 2, ..., T$, be a sequence of chosen (law invariant coherent) risk measures and $\rho_{t|\xi_{[t-1]}}$ be their conditional analogues (see, e.g., [8, Chapter 6] for a discussion of optimization problems involving coherent risk measures). Specifically, apart from risk measures of the form (4.3), we also consider mean-upper-semideviation risk measures of order p:

$$\rho_t(Z) = \mathbb{E}[Z] + \kappa_t \left(\mathbb{E}\big\{ \big[Z - \mathbb{E}[Z]\big]_+^p \big\} \right)^{1/p}, \tag{4.5}$$

where $p \in [1, \infty)$ and $\kappa_t \in [0, 1]$. In particular, for $p = 1$ this becomes

$$\rho_t(Z) = \mathbb{E}[Z] + \kappa_t \mathbb{E}\big[Z - \mathbb{E}[Z]\big]_+ = \mathbb{E}[Z] + \tfrac{1}{2}\kappa_t \mathbb{E}\big|Z - \mathbb{E}[Z]\big|. \tag{4.6}$$

Assuming that the stagewise independence condition holds, the corresponding dynamic programming equations are

$$Q_t\left(x_{t-1}, \xi_t\right) = \inf_{x_t \in \mathbb{R}^{n_t}} \big\{ c_t^{\mathsf{T}} x_t + \mathcal{Q}_{t+1}(x_t) : B_t x_{t-1} + A_t x_t = b_t,\ x_t \geq 0 \big\}, \tag{4.7}$$

with

$$\mathcal{Q}_{t+1}\left(x_t\right) := \rho_{t+1}\left[Q_{t+1}\left(x_t, \xi_{t+1}\right)\right]. \tag{4.8}$$

At the first stage problem

$$\operatorname*{Min}_{x_1 \in \mathbb{R}^{n_1}} c_1^{\mathsf{T}} x_1 + \mathcal{Q}_2(x_1) \ \text{ s.t. } A_1 x_1 = b_1,\ x_1 \geq 0, \tag{4.9}$$

should be solved. Note that because of the stagewise independence, the cost-to-go functions $\mathcal{Q}_{t+1}\big(x_t\big)$ do not depend on the data process. Note also that since the considered risk measures are convex and monotone, the cost-to-go functions $\mathcal{Q}_{t+1}\big(x_t\big)$ are convex (cf., [8, section 6.7.3]).

4.1.1 Backward step for mean-upper-semideviation risk measures

The corresponding SAA problem is obtained by replacing the expectations with their sample average estimates. For risk measures of the form (4.5), the dynamic programming equations of the SAA problem take the form

$$Q_{tj}\big(x_{t-1}\big) = \inf_{x_t \in \mathbb{R}^{n_t}} \big\{ c_{tj}^{\mathsf{T}} x_t + \mathcal{Q}_{t+1}\left(x_t\right) : B_{tj} x_{t-1} + A_{tj} x_t = b_{tj},\ x_t \geq 0 \big\}, \tag{4.10}$$

for $j = 1, ..., N_{t-1}$, with

$$\mathcal{Q}_{t+1}\big(x_t\big) = \widehat{Q}_{t+1}\big(x_t\big) + \kappa_t \left(\frac{1}{N_t} \sum_{j=1}^{N_t} \Big[Q_{t+1,j}\big(x_t\big) - \widehat{Q}_{t+1}\big(x_t\big) \Big]_+^p \right)^{1/p}, \tag{4.11}$$

$t = T, ..., 2$ and $Q_{T+1}(\cdot) \equiv 0$, where

$$\widehat{Q}_{t+1}(x_t) = \frac{1}{N_t} \sum_{j=1}^{N_t} Q_{t+1,j}(x_t).$$

The optimal value of the SAA problem is given by the optimal value of the first stage problem

$$\operatorname*{Min}_{x_1 \in \mathbb{R}^{n_1}} c_1^{\mathsf{T}} x_1 + \mathcal{Q}_2(x_1) \ \text{ s.t. } A_1 x_1 = b_1, \ x_1 \geq 0. \tag{4.12}$$

In order to apply the backward step of the SDDP algorithm we need to know how to compute subgradients of the right hand side of (4.11). Let us consider first the case of $p = 1$. Then (4.11) becomes

$$\mathcal{Q}_{t+1}(x_t) = \widehat{Q}_{t+1}(x_t) + \frac{\kappa_t}{N_t} \sum_{j=1}^{N_t} \left[Q_{t+1,j}(x_t) - \widehat{Q}_{t+1}(x_t) \right]_+ . \tag{4.13}$$

Let $\gamma_{t+1,j}$ be a subgradient of $Q_{t+1,j}(x_t)$, $j = 1, ..., N_t$, at the considered point x_t. In principle it could happen that $Q_{t+1,j}(\,\cdot\,)$ is not differentiable at x_t, in which case it will have more than one subgradient at that point. Fortunately we need just one (any one) of its subgradients.

Then the corresponding subgradient of $\widehat{Q}_{t+1}(x_t)$ is

$$\widehat{\gamma}_{t+1} = \frac{1}{N_t} \sum_{j=1}^{N_t} \gamma_{t+1,j}, \tag{4.14}$$

and the subgradient of $\left[Q_{t+1,j}(x_t) - \widehat{Q}_{t+1}(x_t) \right]_+$ is

$$\nu_{t+1,j} = \begin{cases} 0 & \text{if} \quad Q_{t+1,j}(x_t) - \widehat{Q}_{t+1}(x_t) < 0, \\ \gamma_{t+1,j} - \widehat{\gamma}_{t+1} & \text{if} \quad Q_{t+1,j}(x_t) - \widehat{Q}_{t+1}(x_t) > 0, \end{cases} \tag{4.15}$$

and hence the subgradient of $\mathcal{Q}_{t+1}(x_t)$ is

$$g_{t+1} = \widehat{\gamma}_{t+1} + \frac{\kappa_t}{N_t} \sum_{j=1}^{N_t} \nu_{t+1,j}. \tag{4.16}$$

In the backward step of the SDDP algorithm the above formulas are applied to the piecewise linear lower approximations $\mathfrak{Q}_{t+1}(\cdot)$ exactly in the same way as in the risk neutral case (discussed in section 3).

Let us consider now the case of $p > 1$. Note that then the cost-to-go functions of the SAA problem are no longer piecewise linear. Nevertheless the lower approximations $\mathfrak{Q}_{t+1}(\cdot)$ are still constructed by using cutting planes and are convex piecewise linear. Similar to (4.16) the corresponding subgradient of $\mathcal{Q}_{t+1}(x_t)$ is (by the chain rule)

$$g_{t+1} = \widehat{\gamma}_{t+1} + p^{-1} \kappa_t q^{p^{-1}-1} \frac{1}{N_t} \sum_{j=1}^{N_t} \eta_{t+1,j}, \tag{4.17}$$

where $q = \frac{1}{N_t} \sum_{j=1}^{N_t} \left[Q_{t+1,j}(x_t) - \widehat{Q}_{t+1}(x_t) \right]_+^p$ and

$$\eta_{t+1,j} = \begin{cases} 0 & \text{if} \quad Q_{t+1,j}(x_t) - \widehat{Q}_{t+1}(x_t) < 0, \\ p[Q_{t+1,j}(x_t) - \widehat{Q}_{t+1}(x_t)]^{p-1}(\gamma_{t+1,j} - \widehat{\gamma}_{t+1}) & \text{if} \quad Q_{t+1,j}(x_t) - \widehat{Q}_{t+1}(x_t) > 0. \end{cases} \tag{4.18}$$

4.1.2 Backward step for mean-AV@R risk measures

Let us consider risk measures of the form (4.3), i.e.,

$$\rho_t(Z) = (1-\lambda_t)\mathbb{E}[Z] + \lambda_t \mathsf{AV@R}_{\alpha_t}[Z]. \tag{4.19}$$

The cost-to-go functions of the corresponding SAA problem are

$$\mathcal{Q}_{t+1}\big(x_t\big) = (1-\lambda_t)\widehat{Q}_{t+1}\big(x_t\big) + \lambda_t \left(Q_{t+1,\iota}\big(x_t\big) + \frac{1}{\alpha_t N_t}\sum_{j=1}^{N_t} \big[Q_{t+1,j}\big(x_t\big) - Q_{t+1,\iota}\big(x_t\big)\big]_+ \right), \tag{4.20}$$

where $\iota \in \{1,...,N_t\}$ corresponds to the $(1-\alpha_t)$ sample quantile, i.e., numbers $Q_{t+1,j}\big(x_t\big)$, $j = 1,...,N_t$, are arranged in the increasing order $Q_{t+1,\pi(1)}\big(x_t\big) \leq ... \leq Q_{t+1,\pi(N_t)}\big(x_t\big)$ and $\iota = \hat{j}$ such that $\pi(\hat{j})$ is the smallest integer such that $\pi(\hat{j}) \geq (1-\alpha_t)N_t$. Note that if $(1-\alpha_t)N_t$ is not an integer, then ι remains the same for small perturbations of x_t.

The corresponding subgradient of $\mathcal{Q}_{t+1}\big(x_t\big)$ is

$$g_{t+1} = (1-\lambda_t)\widehat{\gamma}_{t+1} + \lambda_t \left(\gamma_{t+1,\iota} + \frac{1}{\alpha_t N_t}\sum_{j=1}^{N_t} \zeta_{t+1,j} \right), \tag{4.21}$$

where

$$\zeta_{t+1,j} = \begin{cases} 0 & \text{if} \quad Q_{t+1,j}\big(x_t\big) - Q_{t+1,\iota}\big(x_t\big) < 0, \\ \gamma_{t+1,j} - \gamma_{t+1,\iota} & \text{if} \quad Q_{t+1,j}\big(x_t\big) - Q_{t+1,\iota}\big(x_t\big) \geq 0. \end{cases} \tag{4.22}$$

The above approach is simpler than the one suggested in [9], and seems to be working as well.

4.1.3 Forward step

The constructed lower approximations $\mathfrak{Q}_t(\cdot)$ of the cost-to-go functions define a feasible policy and hence can be used in the forward step procedure in the same way as it was discussed in section 3.2. That is, for a given scenario (sample path), starting with a feasible first stage solution $\bar{x}_1$, decisions $\bar{x}_t$, $t = 2,...,T$, are computed recursively going forward with $\bar{x}_t$ being an optimal solution of

$$\operatorname*{Min}_{x_t} c_t^{\mathsf{T}} x_t + \mathfrak{Q}_{t+1}(x_t) \;\; \text{s.t. } A_t x_t = b_t - B_t \bar{x}_{t-1},\; x_t \geq 0, \tag{4.23}$$

for $t = 2,...,T$. These optimal solutions can be used as trial decisions in the backward step of the algorithm.

Unfortunately there is no easy way to evaluate value

$$c_1^{\mathsf{T}}\bar{x}_1 + \rho_{2|\xi_1}\left[c_2^{\mathsf{T}}\bar{x}_2(\xi_{[2]}) + \cdots + \rho_{T|\xi_{[T-1]}}\left(c_T^{\mathsf{T}}\bar{x}_T(\xi_{[T]})\right)\right]$$

of the obtained policy, and hence to construct an upper bound for the optimal value of the corresponding risk-averse problem (4.2). Therefore a stopping criterion based on stabilization of the lower bound was used in numerical experiments. Of course, the expected value (3.7) of the constructed policy can be estimated in the same way as in the risk neutral case by the averaging procedure.

	% of total load curtailment	Cost
1	0 – 5	1142.80
2	5 – 10	2465.40
3	10 – 20	5152.46
4	20 – 100	5845.54

Table 1: Deficit costs and depths

		to				
		SE	S	NE	N	IM
from	SE	–	7379	1000	0	4000
	S	5625	–	0	0	0
	NE	600	0	–	0	2236
	N	0	0	0	–	–
	IM	3154	0	3951	3053	–

Table 2: Interconnection limits between systems

5 Case study description

The numerical experiments were carried out considering instances of multi-stage linear stochastic problems based on an aggregate representation of the Brazilian Interconnected Power System long-term operation planning problem, as of January 2010, which can be represented by a graph with four generation nodes - comprising sub-systems Southeast (SE), South (S), Northeast (NE) and North (N) – and one (Imperatriz, IM) transshipment node (see Figure 2).

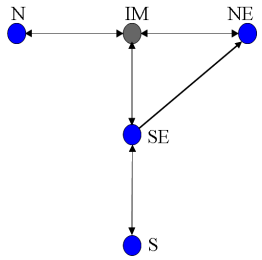

Figure 2: Case-study interconnected power system

The load of each area must be supplied by local hydro and thermal plants or by power flows among the interconnected areas. A slack thermal generator of high cost that increases with the amount of load curtailment accounts for load shortage at each area (Table 1). Interconnection limits between areas may differ depending of the flow direction, see Table 2. The energy balance equation for each sub-system has to be satisfied for each stage and scenario. There are bounds on stored and generated energy for each sub-system aggregate reservoir and on thermal generations.

As mentioned in the Introduction, the long-term planning horizon for the Brazilian case comprises 60 months, due to the existence of multi-year regulation capacity of some large reservoirs. In order to obtain a reasonable cost-to-go function that represents the continuity of the energy supply after these firsts 60 stages, a common practice is to add 60 more stages to the problem and consider a zero cost-to-go function at the end of the 120^{th} stage. Hence, the objective function of the planning problem is to minimize the expected cost of the operation along the 120 months

planning horizon, while supplying the area loads and obeying technical constraints. In the case of risk neutral approach, the objective function is the minimization of the expected value along the planning horizon of thermal generation costs plus a penalty term that reflects energy shortage.

The case's general data, such as hydro and thermal plants data and interconnections capacities were taken as static values through time. The demand for each system and the energy inflows in each reservoir were taken as time varying.

5.1 Time series model for the inflows

5.1.1 The historical data

The historical data is composed of 79 observations of the natural monthly energy inflow (from year 1931 to 2009) for each of the four systems. Let X_t, $t = 1, \ldots, 948$ denote a time series of monthly inflows for one of the regions. Figure 3 shows the histograms for the historical observations for each of the 4 systems in the aggregated model. We can see that X_t is highly skewed to the right for each of the 4 systems. This observation motivates considering $Y_t = \log(X_t)$ for analysis.

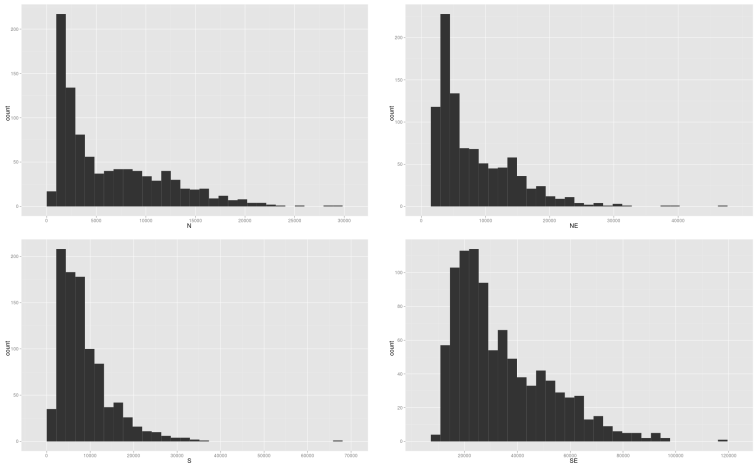

Figure 3: Histograms of the historical observations

Figure 4 shows the histograms for Y_t - the logarithm of historical observations for each of the 4 systems. After taking the logarithm, the histograms become more symmetric.

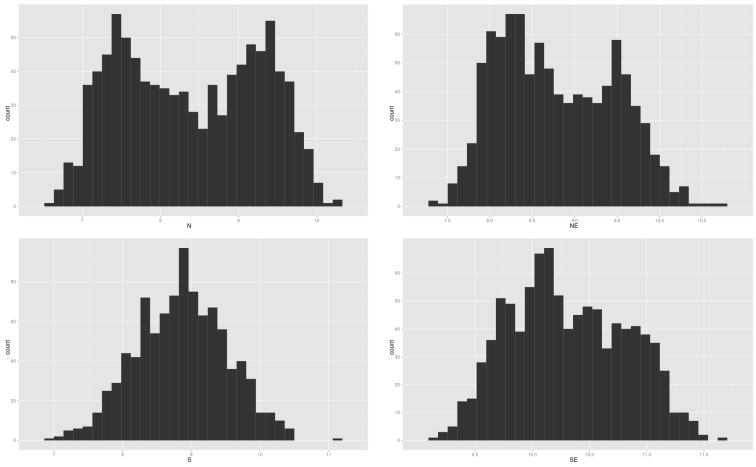

Figure 4: Histograms of the log-observations

Figure 5 shows monthly box plots of regions inflows. It could be seen that inflows of the N, NE and SE systems have a clear seasonal behavior, while for the S system it is not obvious.

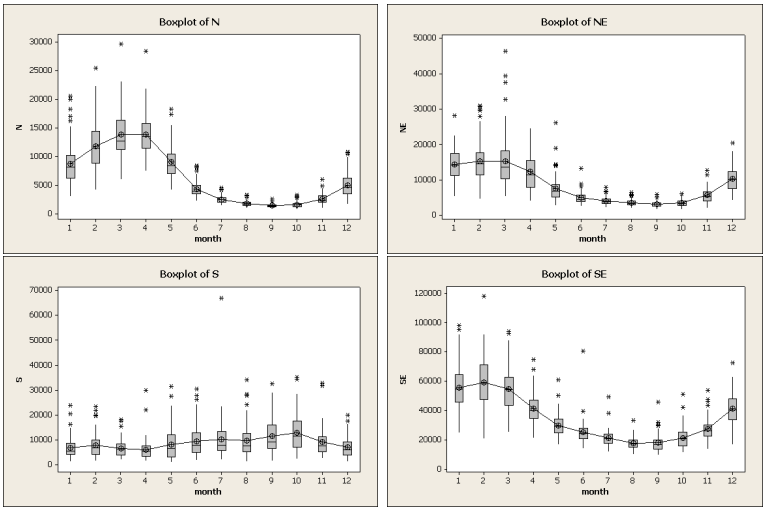

Figure 5: Box plot of the inflows for each system

5.1.2 Time series analysis of SE

As an example we give below analysis of the time series X_t of the SE data points. Analysis of the other 3 regions were carried out in a similar way. Figure 6 shows box plots of monthly inflows of the log-observations $Y_t = \log(X_t)$ of SE inflows. One can clearly note the seasonal behavior of the series, suggesting that a periodic monthly model could be a reasonable framework for this series.

Let $\hat{\mu}_t = \hat{\mu}_{t+12}$ be the monthly averages of Y_t and $Z_t = Y_t - \hat{\mu}_t$ be the corresponding residuals. Figure 7 shows the partial autocorrelation of the Z_t time series. High value at lag 1 and insignificant

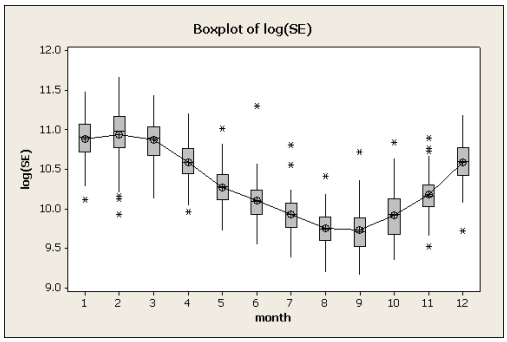

Figure 6: Box plot of the log-observations of SE inflows

values for larger lags suggest the first order $AR(1)$ autoregressive time series model for Z_t:

$$Z_t = \alpha + \phi Z_{t-1} + \epsilon_t. \tag{5.1}$$

For adjusted model the estimate for the constant term α resulted highly insignificant and could be removed from the model. This is not surprising since values Z_t by themselves are already residuals. Trying second order $AR(2)$ model for Z_t didn't give a significant improvement of the fit.

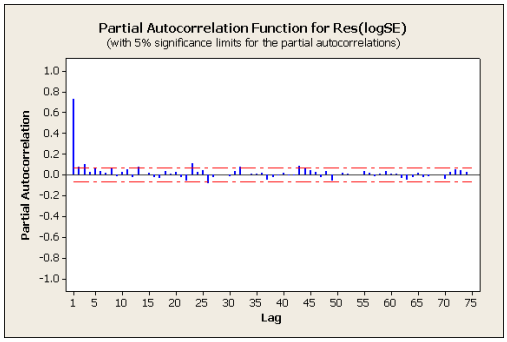

Figure 7: Partial autocorrelation of the residuals of the log-observations of SE inflows

Similar results were obtained for the other three subsystems. Therefore, we consider AR(1) model for all subsystems in the subsequent analysis.

5.1.3 Model description

The analysis of section 5.1.2 suggests the following model for the time series Y_t for a given month

$$Y_t - \hat{\mu}_t = \phi(Y_{t-1} - \hat{\mu}_{t-1}) + \epsilon_t, \tag{5.2}$$

where ϵ_t is iid sequence having normal distribution $N(0, \sigma^2)$. For the original times series X_t this gives

$$X_t = e^{\epsilon_t} e^{\hat{\mu}_t - \phi\hat{\mu}_{t-1}} X_{t-1}^{\phi}. \tag{5.3}$$

Unfortunately this model is not linear in X_t and would result in a nonlinear multistage program. Therefore we proceed by using the following (first order) approximation of the function $y = x^{\phi}$ at $e^{\hat{\mu}_{t-1}}$

$$x^{\phi} \approx (e^{\hat{\mu}_{t-1}})^{\phi} + \phi(e^{\hat{\mu}_{t-1}})^{\phi-1}(x - e^{\hat{\mu}_{t-1}}),$$

which leads to the following approximation of the model (5.3)

$$X_t = e^{\epsilon_t}\left[e^{\hat{\mu}_t} + \phi e^{\hat{\mu}_t - \hat{\mu}_{t-1}}(X_{t-1} - e^{\hat{\mu}_{t-1}})\right]. \tag{5.4}$$

We allow, further, the constant ϕ to depend on the month, and hence to consider the following time series model

$$X_t = e^{\epsilon_t}\left[e^{\hat{\mu}_t} + \gamma_t e^{\hat{\mu}_t - \hat{\mu}_{t-1}}(X_{t-1} - e^{\hat{\mu}_{t-1}})\right] \tag{5.5}$$

with $\gamma_t = \gamma_{t+12}$.

We estimate the parameters of model (5.5) directly from the data.

- Denote by $R_t = \frac{X_t - e^{\hat{\mu}_t}}{e^{\hat{\mu}_t}}$. If the error term ϵ_t is set to zero, i.e., the multiplicative error term e^{ϵ_t} is set to one, (5.5) can be written as:

  $$R_t = \gamma_t R_{t-1} \tag{5.6}$$

  For each month, we perform a least square fit to the R_t sequence to obtain the monthly values for γ_t, assuming that $\gamma_t = \gamma_{t+12}$.

- The errors ϵ_t are modelled as a component of multivariate normal distribution $N(0, \widehat{\Sigma}_t)$, where $\widehat{\Sigma}_t$ is the sample covariance matrix for

  $$\log\left(\frac{X_t}{[e^{\hat{\mu}_t} + \gamma_t e^{\hat{\mu}_t - \hat{\mu}_{t-1}}(X_{t-1} - e^{\hat{\mu}_{t-1}})]}\right)$$

  on a *monthly basis*, i.e., $\widehat{\Sigma}_{t+12} = \widehat{\Sigma}_t$.

5.1.4 Model validation

In this experiment we compute the one step ahead predictions and compare the obtained results against real realizations for the past 2 years of the available historical data. At each time step we compute the predicted value $\widehat{X}_t$ using the corresponding realization X_{t-1}, from the historical values, by employing the constructed model

$$\widehat{X}_t = e^{\hat{\mu}_t} + \gamma_t e^{\hat{\mu}_t - \hat{\mu}_{t-1}}(X_{t-1} - e^{\hat{\mu}_{t-1}}). \tag{5.7}$$

Consequently we compare the predicted value $\widehat{X}_t$ with the real realization X_t. Figure 8 shows the obtained results for the past 2 years of the available historical data. The model exhibits a good fit for systems (N), (NE) and (SE). For system (S) the model gives sometimes lower peaks than the real realizations.

We generate 200 scenarios using the model (5.5) described in section 5.1.3. Each scenario contains 948 points (similar length as the historical data). For each scenario, the mean and standard deviation is computed for each one of the 4 systems starting from the observation 121. We start from observation 121 to eliminate the initial value effect. The obtained values are then compared to the historical data values. Figures 9 and 10 show the box plots for the obtained means against the historical data realizations (red segments).

The average values, obtained for all systems, are relatively close to the historical values in most of the cases.

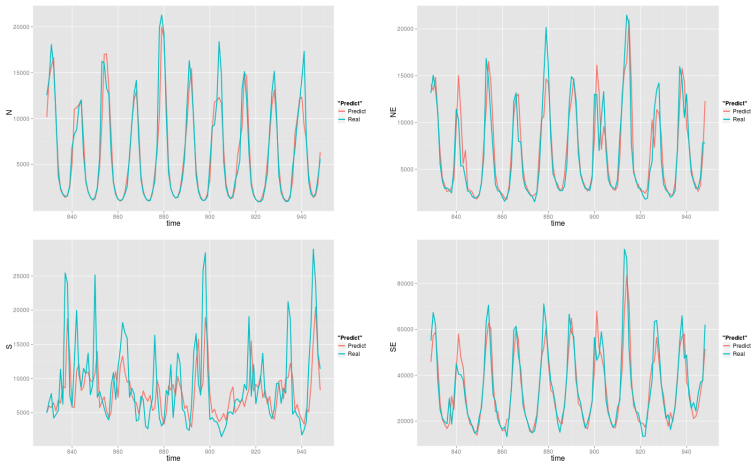

Figure 8: One step ahead prediction

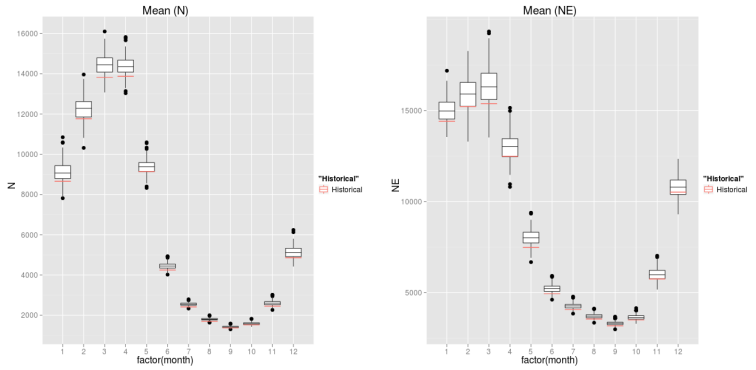

Figure 9: Monthly means boxplot: (N) and (NE)

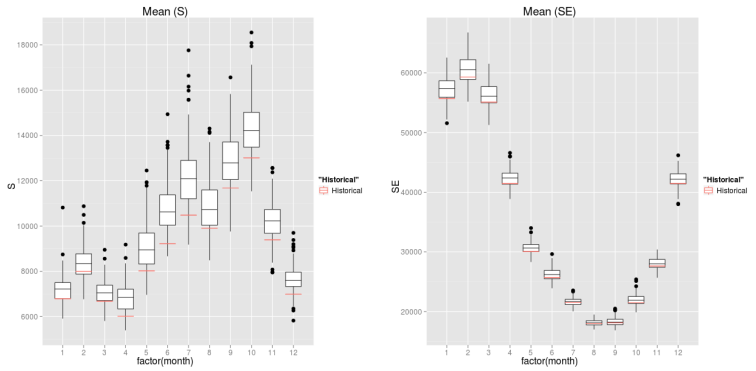

Figure 10: Monthly means boxplot: (S) and (SE)

Figures 11 and 12 illustrate the box plots for standard deviations against the historical data realizations (red segments). The suggested model also gives a reasonably good fit to variability

when compared to the historical data values.

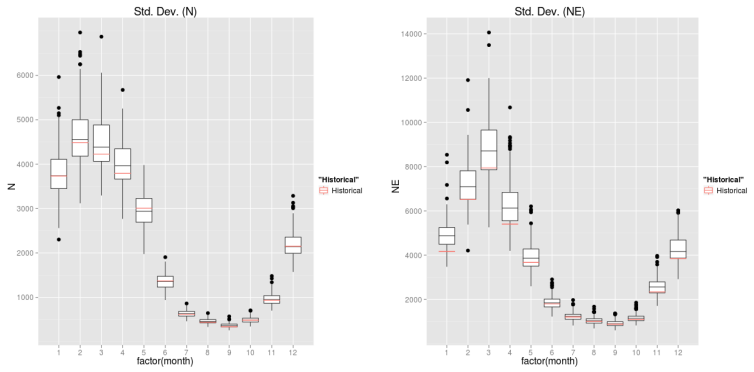

Figure 11: Monthly standard deviations boxplot: (N) and (NE)

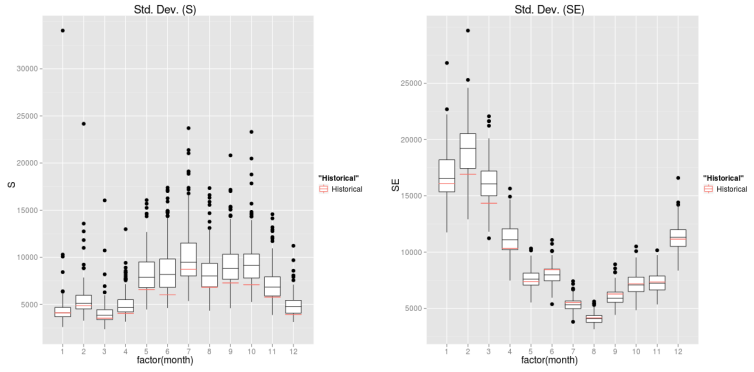

Figure 12: Monthly standard deviations boxplot: (S) and (SE)

6 Computational experiments

The numerical experiments are performed on an aggregated representation of the Brazilian Interconnected Power System operation planning problem with historical data as of January 2011. The study horizon is of 60 stages and the total number of considered stages is 120. We use the high demand profile setting described in [13]. We implement two versions of the risk averse SDDP algorithm, one with the mean-AV@R and one with the mean-upper-semideviation risk measures both applied to solve the problem with the statical model suggested in section 5 (with 8 state variables at each stage).

The SAA tree, generated in both cases, has 100 realizations per every stage with the total number of scenarios $1 \times 100 \times \cdots \times 100 = 100^{119}$. In the following experiments we run the SDDP algorithm with 1 trial solution per iteration.The individual stage costs and policy value are evaluated using 3000 randomly generated scenarios. Both implementations were written in C++ and using Gurobi 4.6. Detailed description of the algorithms can be found in [13]. The codes were run on 1 core of (2 quad-core Intel E5520 Xeons 2.26GHz, and 24GB RAM) machine. Dual simplex was used as a default method for the LP solver.

In section 6.1 the results for the SDDP algorithm with the time series model (5.5) are discussed within the risk neutral framework. The following section 6.2 discusses the computational experiments for the risk averse approach with mean-AV@R and mean-upper-semideviation risk measures. Finally, we conclude this part by discussing variability of the SAA problems and sensitivity to the initial conditions in sections 6.3 and 6.4.

6.1 Risk neutral approach results

In this section we investigate some computational issues related to the risk neutral SDDP applied to the problem with time series model (5.5). Figure 13 shows the bounds for risk neutral SDDP for more than 7000 iterations. In the legend, we have the following notation:

- LB: the lower bound (i.e., the first stage optimal value)
- UB: the upper end of the 95% confidence interval of the upper bound computed approximately using as observations the past 100 forward step realizations
- UBMA: moving average of UB using the past 100 values

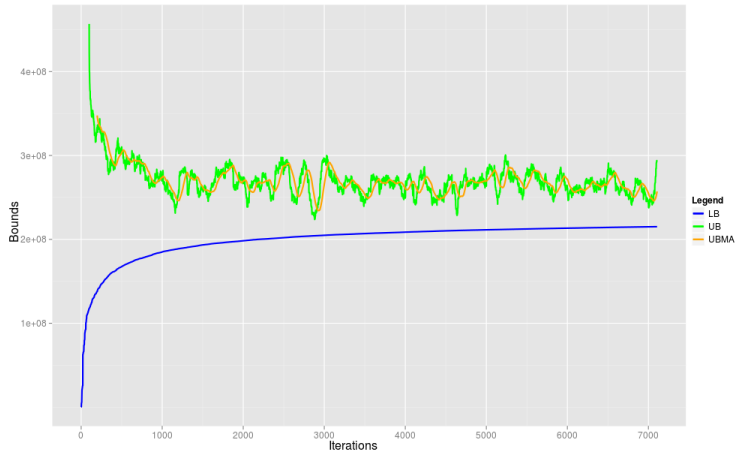

Figure 13: Bounds for Risk Neutral SDDP with time series model

Figure 13 illustrates typical behavior of the SDDP bounds - fast increase in the lower bound for the first iterations and then a slow increase in later iterations. The upper bound exhibits some variability along with a decreasing trend. We should notice the relatively slow convergence.

Figure 14 show the evolution of the approximate gap over iterations. The blue line provides a smoothing of the observations to get the approximate trend.

We consider an approximate gap defined by $\frac{UBMA-LB}{LB}$. This is just an approximation of the real gap defined as the difference between the upper end of 95% confidence interval and the lower bound. Due to the significant computational effort to evaluate adequately this gap, we approximate the observations by taking the past 100 forward step realizations. We perform at iteration 2000 and iteration 3000 a proper forward step with 3000 scenarios (see Table 3 for the details). We can see in Figure 14 the fast decrease in the first 1000 iterations and then a relatively slow decay for later ones.

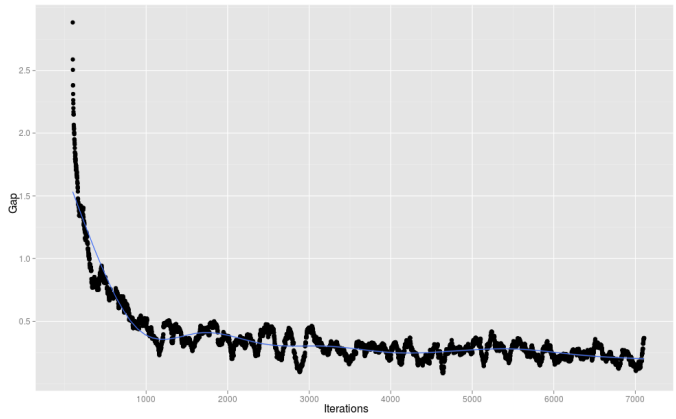

Figure 14: Approximate gap for Risk Neutral SDDP

Table 3 shows the lower bound, upper bound 95% confidence interval (mean and upper end), the CPU time along with the gap (i.e., $\frac{UB_{upper}-LB}{LB}$) at iteration 2000 and 3000. The confidence interval for the upper bound was computed using 3000 randomly generated scenarios.

Iteration	CPU time (sec)	LB	UB average	UB upper end	Gap
2000	58,841.3	198,255,341.2	246,090,917.6	250,193,144.8	26.19%
3000	136,740.7	204,835,524.2	246,164,363.1	250,136,700.6	22.11%

Table 3: Total CPU time, bounds status and gap at iteration 2000 and 3000

The approximate gap (of 35.84% and 29.24% at iterations 2000 and 3000, respectively) gives a slightly higher value than the accurate gap. Also, to reach a gap of 22.11%, which is quite large, 3000 iterations were needed. The experiment took 136,740.7 seconds (i.e., approximately 38 hours). Figure 15 shows the individual stage costs at iteration 2000 and iteration 3000 for the risk neutral SDDP.

Some differences are perceivable between the individual stage costs distributions at iteration 2000 and iteration 3000. These differences are noticeable in the 95% quantile. However, these differences don't have a dramatic effect on the general shape of the distribution. This observation is expected since running further the algorithm will provide a more accurate representation of the optimal solution. For these research purposes, it seems reasonable to stop the algorithm at a computationally acceptable running time without any significant impact on the general conclusions.

Figure 16 shows the CPU time per iteration for the risk neutral SDDP for more than 7000 iterations. We can see the linear trend of the CPU time per iteration. The discrepancy occurring at some iterations is most likely due to the shared resources feature of computing environment.

6.2 Risk averse approach results

6.2.1 Mean-AV@R risk measures

In this section, we investigate some computational issues related to the mean-AV@R risk averse SDDP applied to the operation planning problem with the time series model suggested in section 5.1. Figure 17 shows the total policy value for the first 60 stages at iteration 3000 for $\alpha \in \{0.05, 0.1\}$

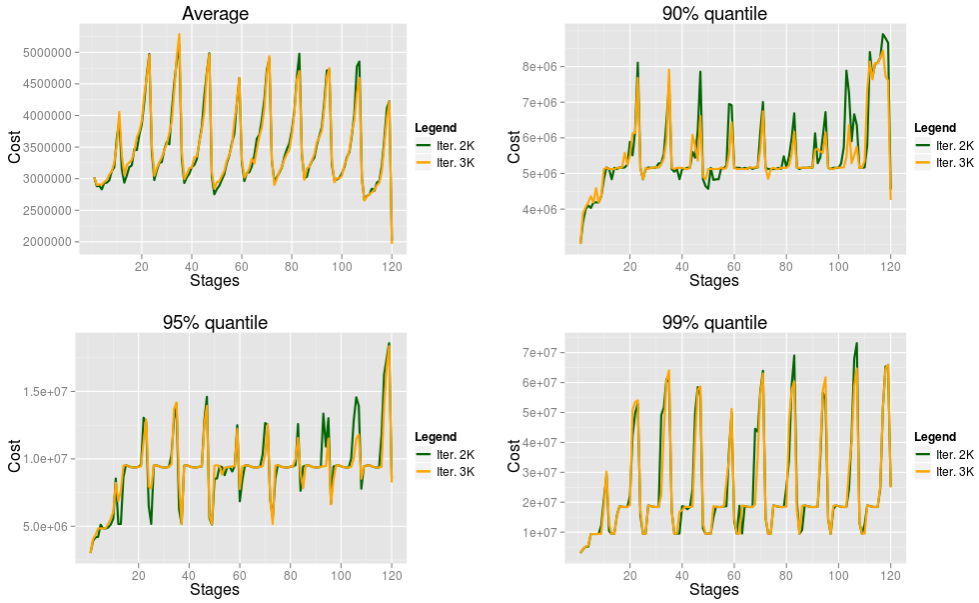

Figure 15: Individual stage costs at iteration 2000 and 3000

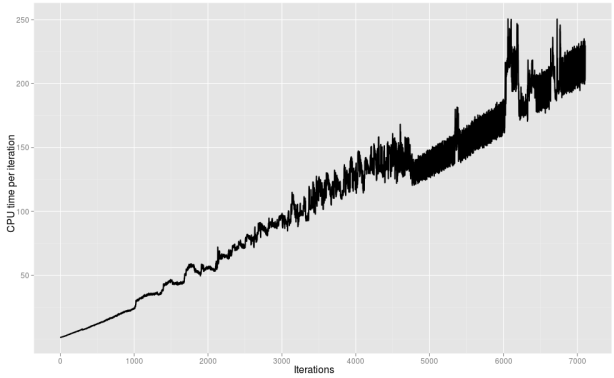

Figure 16: CPU time per iteration

and $\lambda \in \{0, 0.05, 0.1, \ldots, 0.35\}$. The dotted line corresponds to $\alpha = 0.1$ and the continuous line corresponds to $\alpha = 0.05$.

Practically there is no significant difference between the 60 stages total cost for $\alpha = 0.05$ and for $\alpha = 0.1$ when $0 \leq \lambda \leq 0.3$. For $\lambda = 0.35$, the total 60 stages policy value is lower for $\alpha = 0.1$. Furthermore, we can see that as λ increases (i.e. more importance is given to the high quantile minimization) the average policy value increases and an improvement in some quantiles is observed for $0 < \lambda \leq 0.30$. The "price of risk aversion" is what we lose on average compared to the risk neutral case (i.e. $\lambda = 0$). It is the price "paid" for some protection against extreme values.

Figure 18 shows the individual stage costs at iteration 3000 for $\lambda = 0.15$. In this figure, we compare the individual stage costs for the risk neutral case, $\alpha = 0.05$ and $\alpha = 0.1$. When we compare the risk averse approach and the risk neutral approach, we can see the significant reduction in the 99% quantile and the loss in the average policy value that occurs mostly in the first stages.

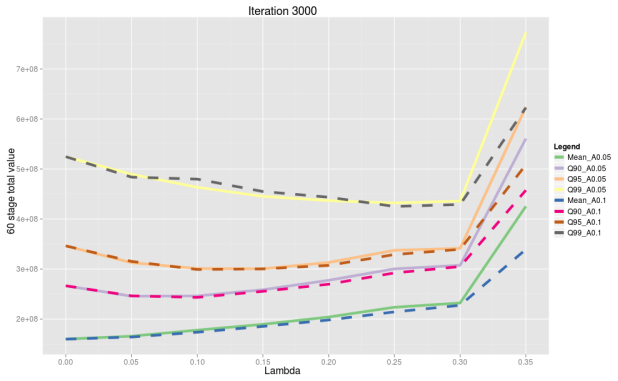

Figure 17: Total policy value for 60 stages for $\alpha \in \{0.05, 0.1\}$ as function of λ

Most of the reduction of the 90% and 95% quantiles happens in the last 15 stages. Furthermore, there is no significant difference between the individual stage costs for $\alpha = 0.05$ and $\alpha = 0.1$.

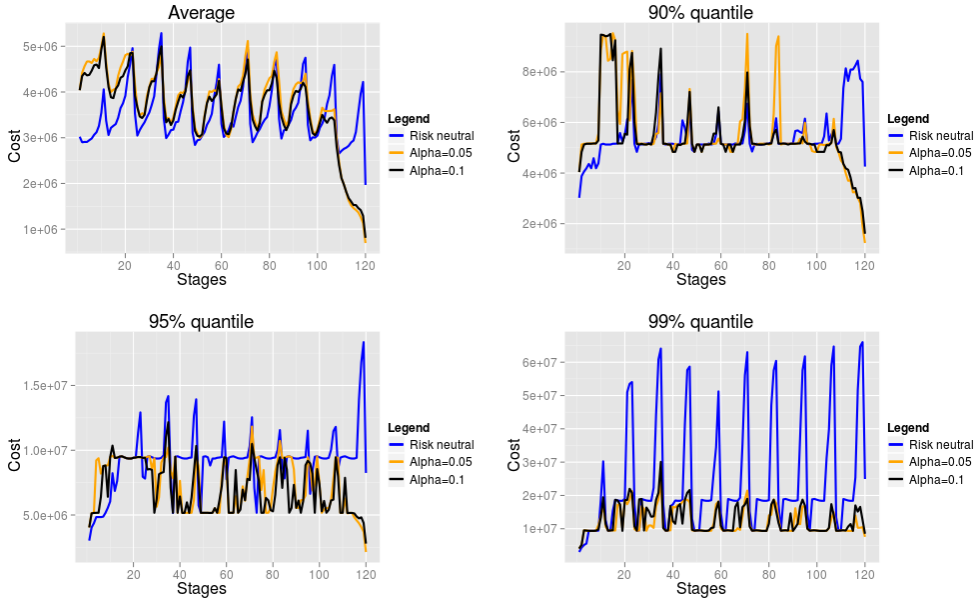

Figure 18: Individual stage costs for $\lambda = 0.15$ and $\alpha \in \{0.05, 0.1\}$

As an example of the impact of risk averse approach in the decision variables, Figure 19 shows the evolution of SE system stored volumes along the planning period, for risk neutral and risk averse for $\lambda = 0.15$ and $\alpha = 0.05$, where one can observe higher values of the stored volumes with the risk averse approach, as is expected. The availability of higher stored volumes makes it possible, in case of droughts, to be able to avoid large deficits and costs spikes as shown in Figure 18.

Among the interesting questions that we can ask is: how much does the risk averse approach costs in terms of CPU time compared to the risk neutral approach? Figure 20 shows the CPU time per iteration for the risk neutral and the AV@R risk averse approach with $\lambda = 0.15$ for the first 3000 iterations. Practically there is no loss in CPU time per iteration when compared to the risk

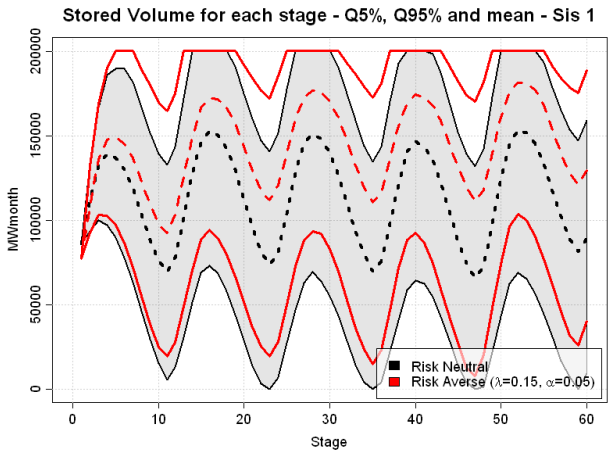

Figure 19: SE System quantiles of stored volumes for each stage

neutral approach.

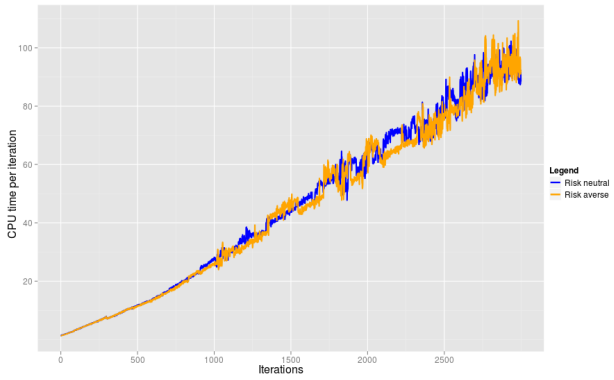

Figure 20: CPU time per iteration for risk neutral and risk averse approaches

6.2.2 Mean-upper-semideviation risk measures

In this section, we investigate some computational issues related to the mean-upper-semideviation risk averse SDDP discussed in section 4.1.

Figure 21 shows the mean, 90%, 95% and 99% quantiles of the total cost for the first 60 stages for different values of $p \in \{1, 2, 3\}$. Notice that a similar general behavior occurs as in the mean-AV@R risk averse approach: an increase of the average value against a decrease of the high quantiles on a specific range of values. For $p = 1$ and constant $\kappa_t = c$, we observe a continuous decrease of the 99% quantile for all values of $\kappa \in \{0.0, \ldots, 0.9\}$. Also, notice the slow increase of the average policy value. For $p = 2$, less reduction of the 99% quantile happens compared to the case of $p = 1$. However, the best reduction in 95% and 90% quantiles moves toward lower penalty value along with an increase of the quantiles values for the higher penalization. For $p = 3$, we can observe a significant increase of the total policy quantiles for the high values of c and a similar trend of moving the best quantiles reduction toward the low penalty values.

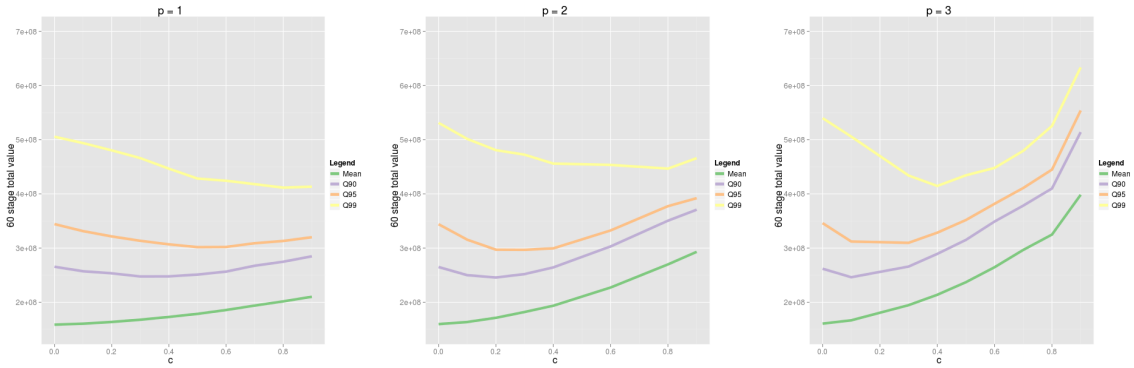

Figure 21: 60 stages total costs for $p \in \{1, 2, 3\}$

Figure 22 shows the individual stage costs plots at iteration 3000 for risk neutral case and the risk averse case for $\{p = 2, c = 0.4\}$ and $\{p = 3, c = 0.4\}$. Similar observations hold as the mean-AV@R risk averse approach: a loss on average occurring in the first stages and a significant impact on the 99% quantile.

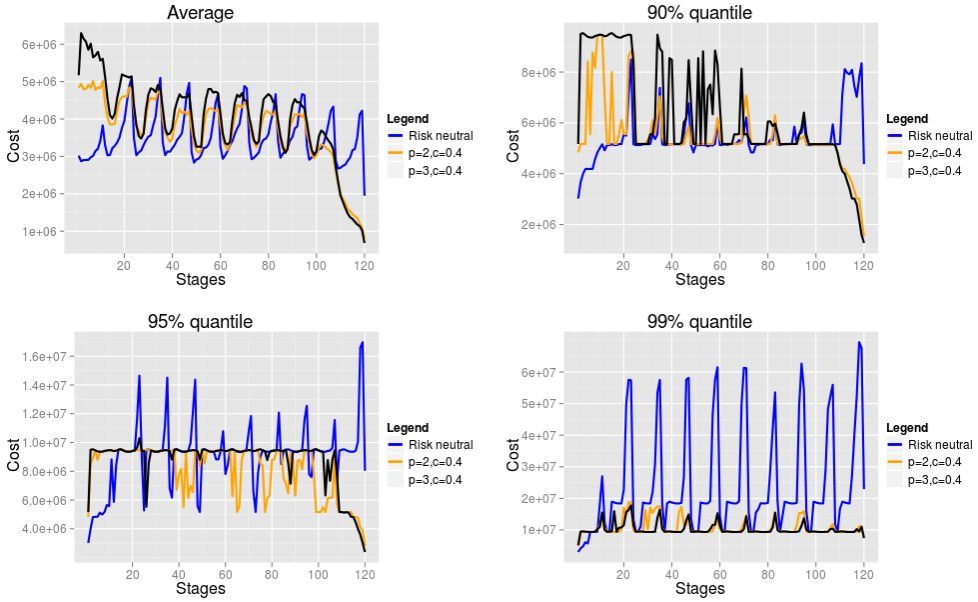

Figure 22: Individual stage costs at iteration 3000 for $c = 0.4$ and $p \in \{2, 3\}$

Figure 23 shows the CPU time per iteration for the mean-AV@R ($\lambda = 0.15$) and mean-upper-semideviation ($c = 0.4, p = 2$) risk measures. There is no significant difference between both of the approaches.

6.3 Variability of SAA problems

In this section we discuss variability of the bounds of optimal values of the SAA problems. Recall that an SAA problem is based on a randomly generated sample, and as such is subject to random

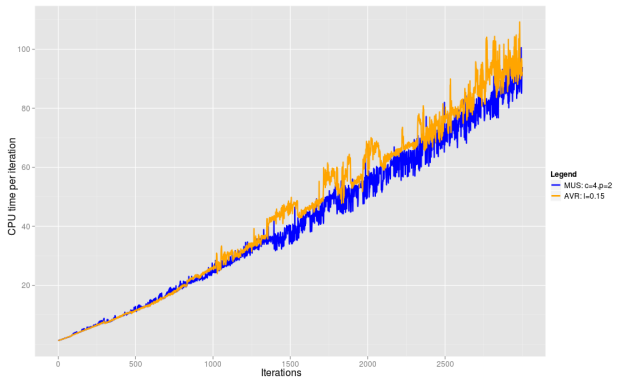

Figure 23: CPU time per iteration: AV@R vs. mean-upper-semideviation

perturbations, and in itself is an approximation of the "true" problem. As it was discussed in [12], for a simple model of the inflows (referred to as the *independent model*) there was little variability in the lower bound - it was in the order of 0.8% of the average lower bound over a sample of 20 scenario trees.

In this experiment we generate ten SAA problems, using time series model (5.5), each one having $1 \times 100 \times \cdots \times 100 = 100^{119}$ scenarios. Then we run 3000 iterations of the SDDP algorithm for the risk neutral approach. At the last iteration, we perform a forward step to evaluate the obtained policy with 5000 scenarios. Figure 24 shows the lower bound evolution for each of the considered SAA problems for 3000 iterations. We can see an increasing range of the lower bound value with the increase of the number of iterations in the beginning, followed by a stabilization of the range in later iterations.

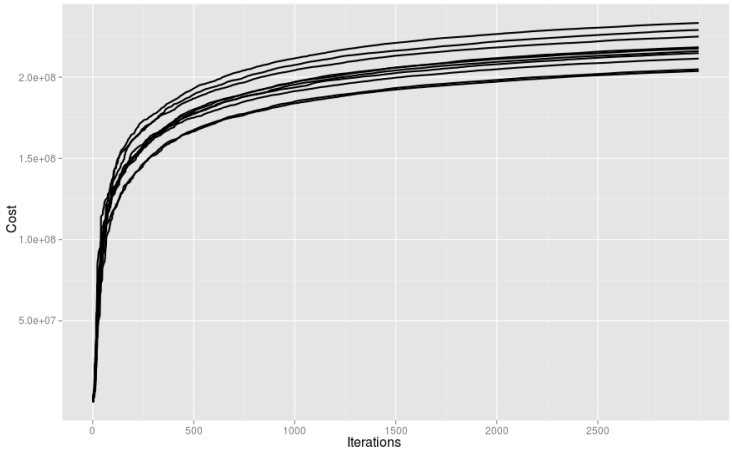

Figure 24: Lower bound variability for risk neutral SDDP

Table 4 shows the 95% confidence interval for the lower bound, average policy value and upper bound at iteration 3000 over a sample of 10 SAA problems. Each of the observations was computed using 5000 scenarios. The last 2 columns of the table shows the range divided by the average of the lower bound where the range is the difference between the maximum and minimum observation and the variability. The variability is defined as the standard deviation divided by the average value of

the lower bound.

	95% C.I. left	Average	95% C.I. right	Range/(Avg. LB)	Variability
Lower bound	211,456,233.9	217,439,241.1	223,422,248.4	13.61%	4.43%
Average policy	253,239,278.6	260,672,455.4	268,105,632.1	17.04%	5.51%
Upper bound	256,466,026.7	264,006,684.4	271,547,342.1	17.27%	5.59%

Table 4: Total CPU time and gap at iteration 2000

The lower bound variability defined as the ratio of the standard deviation to the average value of the time series model is equal to 4.43% compared to 0.8% of the simple statistical model. There is an increase of the variability of the observations but it remains within the acceptable range.

Figure 25 shows the lower bound evolution for the mean-AV@R risk averse SDDP with $\lambda = 0.15$ and $\alpha = 0.05$ for 3000 iterations. Similar to the risk neutral case, we can see an increasing range of the lower bound value with the increase of the number of iterations in the beginning, followed by a stabilization of the range in later iterations. Furthermore, some fluctuations in the bounds occur.

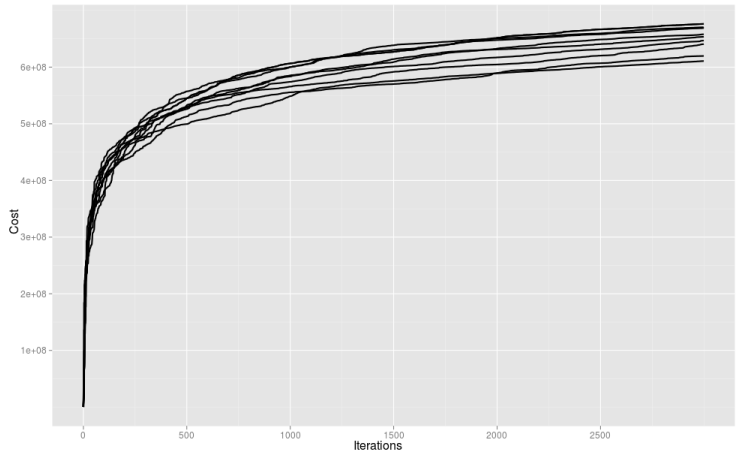

Figure 25: Lower bound variability for risk averse SDDP ($\lambda = 0.15$)

Table 5 shows the 95% confidence interval for the lower bound and average policy value at iteration 3000 over a sample of ten SAA problems. Each of the observations was computed using 3000 scenarios. The last column of the table shows the range divided by the average of the lower bound.

	95% C.I. left	Average	95% C.I. right	Range/(Avg. LB)	Variability
Lower bound	637,924,730.9	652,146,562.4	666,368,394.0	10.03%	3.51%
Average policy	284,392,636.5	292,042,297.6	299,691,958.7	5.82 %	1.89%

Table 5: Total CPU time and gap at iteration 2000

The risk averse approach shows lower variability of the lower bound and the average policy value.

6.4 Sensitivity to the initial conditions

The purpose of this experiment is to investigate the impact of changing initial conditions of the simulated time series (with the multiplicative error terms) on the distribution of the individual stage costs. In the above experiments, we considered the initial volumes and the reservoirs level based on January 2011 data. Table 6 shows the numerical values for these parameters.

	SE	S	NE	N
Inflows	53,442.7	6,029.9	18,154.9	5,514.4
Reservoir level	59,419.3	5,874.9	12,859.2	5,271.5
Total	112,862.0	11,904.8	31,014.1	10,785.9
% of Maximum capacity	56.23 %	60.68%	59.86%	84.63%

Table 6: Initial conditions

Table 7 shows the maximum storage value for each system.

	SE	S	NE	N
Maximum storage	200,717.6	19,617.2	51,806.1	12,744.9

Table 7: Maximum stored volume

We consider the following 2 initial levels: 25% and 75% of the maximum storage capacity in each system. Figure 26 shows the individual stage costs for the risk neutral approach in two cases: all the reservoirs start at 25% or at 75% of the maximum capacity. The yellow curve denotes the 75% initial reservoir level and the dark green denotes the 25% initial level.

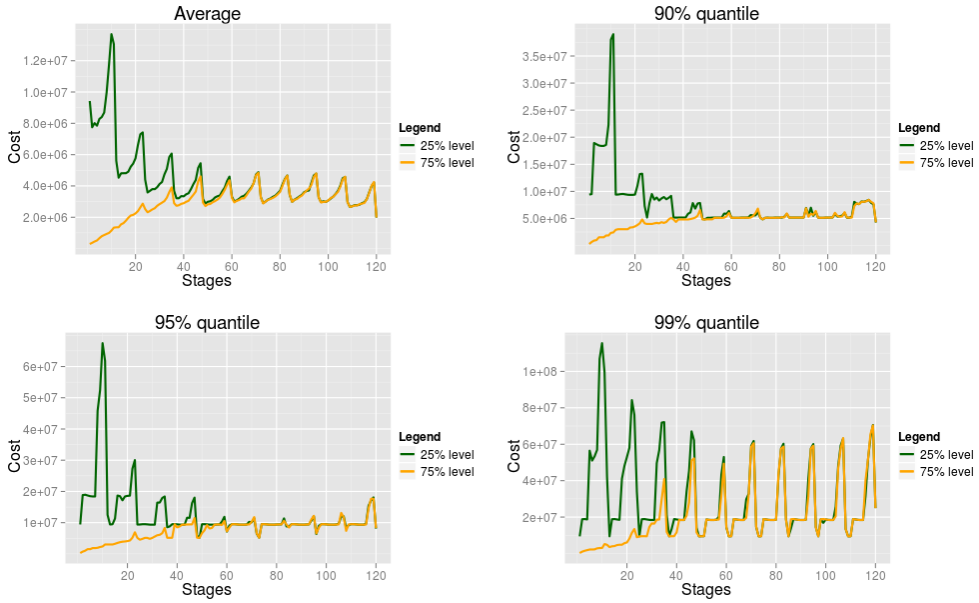

Figure 26: Sensitivity to initial conditions (Risk neutral)

When we start with 25% of the maximum capacity in all reservoirs, high costs occur for the first stages which reflects the recourse to the expensive thermal generation to satisfy the demand.

Similarly, when we start with 75% of the maximum capacity in all reservoirs low costs occur in the first stages. Interestingly in both cases the costs become identical starting from 60^{th} stage.

Similarly, Figure 27 shows the individual stage costs for the mean-AV@R risk averse approach with $\lambda = 0.15$ in two cases: all the reservoirs start at 25% or at 75% of the maximum capacity.

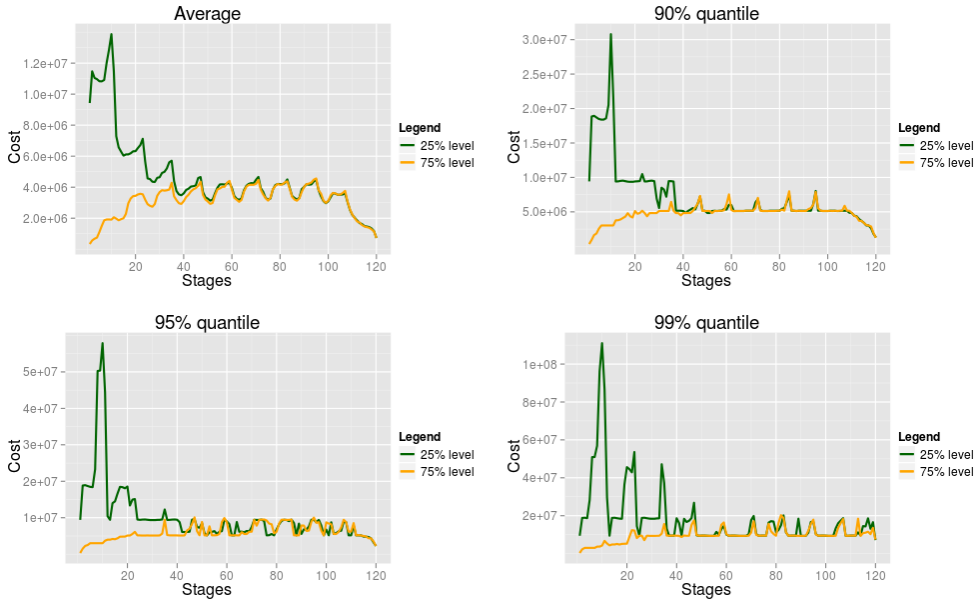

Figure 27: Sensitivity to initial conditions (Risk averse: $\lambda = 0.15$)

Similar observations as the risk neutral case hold for the risk averse approach: high costs occur for the first stages when we start with 25% levels and low costs occur in the first stages when we start with 75% levels. The two costs have similar distributions at later stages. However, one key difference with the risk neutral case is the values attained in the peaks are lower (see Figure 26).

7 Conclusions

The problem formulation and modelling related issues were discussed in sections 1 and 2. A generic description of the SDDP algorithm is presented in section 3. Adaptive risk averse approach was discussed in section 4 under two different perspectives: one through the mean-AV@R and the other using the mean-upper-semideviation risk measures.

The case study along a time series model for the inflows was presented and discussed in section 5. The computational results for the hydrothermal system operation planning problem of the Brazilian interconnected power system was presented in section 6. We discussed in 6.1 some computational aspects of the risk neutral SDDP applied to solve the problem with the time series model suggested in 5.1. We observed a slow convergence of the algorithm along a linear trend of the CPU time per iteration. Section 6.2 summarizes the results for the mean-AV@R and mean-upper-semideviation risk measures. We have seen that the risk averse approach ensures a reduction in the high quantile values of the individual stage costs. This protection comes with an increase of the average policy value - the price of risk aversion. Furthermore, both of the risk averse approaches come with practically no extra computational effort. In section 6.3 we have seen that there was no significant

variability of the SAA problems. Finally, we investigated in section 6 the impact of the initial conditions on the individual stage costs.

References

[1] N. V. Arvaniditis and J. Rosing, *Composite representation of a multireservoir hydroelectric power system*, IEEE Transactions on Power Apparatus and Systems, 89 (1970), no. 2, 319–326.

[2] Ministry of Mines and Energy MME, *Electricity in the Brazilian Energy Plan – PDE 2020*, July 2011, Secretariat of Energy Planning and Development.

[3] Nemirovski, A. and Shapiro, A. Convex approximations of chance constrained programs. *SIAM Journal on Optimization*, 17 (2006), 969–996.

[4] M. V. F. Pereira and L. M. V. G Pinto, *Stochastic optimization of a multireservoir hydroelectric system—a decomposition approach*, Water Resources Research, 21 (1985), no. 6, 779–792.

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

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

[7] Ruszczyński, A. and Shapiro, A. Conditional risk mappings. *Mathematics of Operations Research*, 31 (2006), 544–561.

[8] Shapiro, A., Dentcheva, D. and Ruszczyński, A., *Lectures on Stochastic Programming: Modeling and Theory*, SIAM, Philadelphia, 2009.

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

[10] Shapiro, A., *Topics in Stochastic Programming*, CORE Lecture Series,Universite Catholique de Louvain, 2011.

[11] L. A. Terry, M. V. F. Pereira, T. A. Araripe Neto, L. E. C. A. Pinto, and P. R. H. Sales, *Coordinating the energy generation of the Brazilian national hydrothermal electrical generating system.*, INTERFACES **16** (1986), no. 1, 16–38.

[12] Shapiro A., Tekaya W., Da Costa J.P. and Soares M.P., Report for technical cooperation between Georgia Institute of Technology and ONS - Operador Nacional do Sistema Eletrico - Phase 1, Technical Report, 2011.

[13] Shapiro, A. and Tekaya, W., Report for technical cooperation between Georgia Institute of Technology and ONS - Operador Nacional do Sistema Eletrico - Risk Averse Approach, Technical Report, 2011.