Guidelines for choosing parameters λ and α for the risk averse approach

Final Report

Second activity of the technical cooperation between
Georgia Institute of Technology
and
ONS - Operador Nacional do Sistema Elétrico

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

Filipe Goulart Cabral and Joari Paulo da Costa
ONS - Operador Nacional do Sistema Elétrico
Rua Júlio do Carmo, 251 - Cidade Nova
20211-160 – Rio de Janeiro – RJ - Brasil

October, 2015

Contents

1 Introduction

The Electric System National Operator (ONS) of Brazil is a private non-profitable entity created on 26 August 1998. It is responsible for coordinating and controlling the operation of generation and transmission facilities in the National Interconnected Power System (NIPS) under supervision and regulation of the Electric Energy National Agency (ANEEL).

The Brazilian power system generation is hydro dominated (about 75% of the installed capacity) and characterized by large reservoirs presenting multi-year regulation capability, arranged in complex cascades over several river basins. The hydro plants use stored water in the reservoirs to produce energy in the future, replacing fuel costs from the thermal units. Since the water inflows depend on rainfalls, the amount of future inflows is uncertain and cannot be predicted with a high accuracy. Moreover, historical records indicate possibility of some very dry time periods which, of course, can put a severe burden on hydro power generation.

Mathematical algorithms compose the core of the Energy Operation Planning Support System. The former objective was to calculate an operation strategy which minimizes the expected value of the operation costs over a planning period of time. This lead to formulation of large scale multistage stochastic programming problems. *Stochastic Dual Dynamic Programming* (SDDP) method is the main algorithmic tool and it is currently employed to solve the involved stochastic optimization problems.

In the previous Technical Agreement signed between Georgia Tech and ONS, that continued from September 2010 until August 2012, several topics regarding the problem formulation methodology and the SDDP algorithm were investigated and certain developments were suggested. One of the main contributions of that project was the development of a risk averse approach based on the *Average Value-at-Risk* (AV@R) (also called *Conditional Value-at-Risk*) risk measure that complies with the SDDP methodology and results in a policy that is less sensitive to extreme scenarios with a reasonable increase on average costs when compared with the risk neutral (traditional) approach. This methodology is currently in use in the planning procedures of the NIPS. The goal of the current Technical Agreement report is to continue investigation of the involved modeling and computational issues.

This document reports the investigations carried out from February to September 2015 regarding guidelines for choosing parameters λ and α in the risk averse approach. Note that the optimal weight of the risk averse component, governed by λ, should not give rise to load curtailment when there is still enough stored energy in the system. The envisaged criterion consists of choosing a pair (λ^*, α^*) that minimizes the expected deficit cost. As this criterion does not result in a closed-form expression for the minimization problem, a grid search procedure was performed.

One of the difficulties of this search procedure is that, given a value for α, the corresponding optimal λ value which minimizes the expected deficit cost is subject to the inherent solution variability of the tree sampling and SDDP algorithm. We discuss briefly in Appendix A some aspects of discretization of continuous distributions of the "true" model that result in the scenarios tree employed in the SDDP algorithm. Two approaches, the usual Monte Carlo method and the Wasserstein distance (also known as K-means algorithm) were considered. We present the numerical studies performed to assess the variability of SDDP solutions with respect to the tree sampling and the tree branching (N_t). Solution variability control is important to ensure distinctions between policies with different (λ, α) parameters.

This report is organized as follows. In section 2 we review the operation planning, based on mean-AV@R problem formulation, and the solution approach currently used. In section 3 we present the mathematical framework that justifies the proposed methodology to choose the parameters

of the risk averse approach. In order to do so, we recall two equivalent formulations for the multistage stochastic programming problem and formalize the adopted procedure. The criterion for choosing (λ, α) is presented in section 4, and the performed numerical experiments are reported in section 5. In particular, in section 5.3 one can see that, for the presented case, it was always possible to find a value of (λ, α) that meets the criterion, and as reported in section 5.4 these values are not distinguishable from the point of view of total thermal generation. Additionally, section 5.5 indicates a preliminary result regarding the issue of when to revise the optimal value choice. Section 6 presents the suggested algorithm to choose the optimal value for (λ, α). Finally, section 7 is devoted to conclusion.

2 Problem general description

The goal of a risk averse approach currently in use for the planning operation problem is to avoid large values of the costs for some possible realizations of the data process at every stage of the considered time horizon. The nested risk averse formulation of the corresponding multistage stochastic programming problem is:

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

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), with $\xi_1 = (c_1, A_1, b_1)$ being deterministic (not random). By $\xi_{[t]} = (\xi_1, \ldots, \xi_n)$ we denote the history of the data process up to time t. The decision vectors x_t in (1) are functions of the data process. At time period t a decision $x_t = x_t(\xi_{[t]})$ can depend on the past observations, but not on the future values of the data process. A sequence $x_t(\xi_{[t]})$, $t = 1, \ldots, T$, of such decision vectors is called a *policy*, or *decision rule*. A policy $x_t(\cdot)$, $t = 1, \ldots, T$, is *feasible* if it satisfies all constraints of (1) with probability one. A policy $x_t^*(\cdot)$, $t = 1, \ldots, T$, is said to be *optimal* if it is feasible and attains the minimum value of (1).

We use the following (conditional) risk measures applied at every stage t of the process:

$$\rho_{t|\xi_{[t-1]}}[Z] = (1-\lambda)\mathbb{E}\left[Z|\xi_{[t-1]}\right] + \lambda \mathsf{AV@R}_\alpha\left[Z|\xi_{[t-1]}\right], \tag{2}$$

where

$\mathbb{E}\left[Z|\xi_{[t-1]}\right]$ denotes the conditional expectation of Z given $\xi_{[t-1]}$;
$\mathsf{AV@R}_\alpha\left[Z|\xi_{[t-1]}\right]$ is the conditional analogue of $\mathsf{AV@R}_\alpha[Z]$, given $\xi_{[t-1]}$;
$\mathsf{AV@R}_\alpha[Z]$ is the Average Value-at-Risk: $\mathsf{AV@R}_\alpha[Z] = \inf_{t\in\mathbb{R}}\left\{t + \alpha^{-1}\mathbb{E}[Z-t]_+\right\}$,
where $\lambda \in [0,1]$ and $\alpha \in (0,1)$ are chosen parameters.

Note that the optimal policy depends on parameters λ and α, i.e., $x_t^* = x_t^*(\xi_{[t]}, \lambda, \alpha)$, $t = 1, \ldots, T$, since the objective function depends on these parameters.

Regarding this formulation, there are two parameters which should be specified: λ and α. Recall that these parameters have different functions. The parameter $\lambda \in [0,1]$ represents a compromise between minimizing the total cost on average and risk aversion of high costs at every stage of the process. For small values of λ the risk measure behaves almost like the risk neutral case, that is, the resulting policy does not meet the requirement of smoothing high peaks of the operation cost due to the extreme scenarios. On the other hand, for large values of λ one becomes too much risk averse,

and the immediate consequence is to cut the load preemptively, which can lead to unnecessary impacts on the society. Parameter $\alpha \in (0,1)$ represents the α-percentile of extreme costs that are to be avoided. This parameter is responsible for smoothing out extreme costs at every stage of the process. Note that for $\alpha = 1$ the $\mathsf{AV@R}_1[\cdot]$ is the expectation functional and for α tending to zero the $\mathsf{AV@R}_\alpha[\cdot]$ tends to the max-risk functional.

Before we state a problem for choosing λ and α, we should mention how to solve problem (1) with fixed parameters. In the next section we review the SDDP algorithm approach to solving problem (1). We assume that the data process $\{\xi_t\}_{t=1}^T$ is *stagewise independent*, i.e., random vector ξ_{t+1} is independent of $\xi_{[t]}$, $t = 1, ..., T-1$.

2.1 Dynamic programming formulation

The dynamic programming formulation is useful to describe practical stochastic problems and to present some classes of algorithms for their solutions. Its general idea is to decompose problem (1) into smaller optimization subproblems connected by a recursive relation, where the base case ($t = T$) is a trivial problem and the last case ($t = 1$) is the problem to be solved.

Let us consider the stochastic programming problem (1) driven by the random data process $\xi_t = (c_t, A_t, B_t, b_t)$, $t = 1, ..., T$. In order to express the dynamic programming recursive relation, we define value (also called cost-to-go) function at stage $t = 2, \ldots, T$ as

$$\mathcal{Q}_t(x_{t-1}, \xi_{[t-1]}) = \rho_{t|\xi_{[t-1]}}\left[\min_{\substack{B_t x_{t-1} + A_t x_t = b_t \\ x_t \geq 0}} c_t^\mathsf{T} x_t + \rho_{t+1|\xi_{[t]}}\Big[\cdots + \rho_{T|\xi_{[T-1]}}\Big[\min_{\substack{B_T x_{T-1} + A_T x_T = b_T \\ x_T \geq 0}} c_T^\mathsf{T} x_T\Big]\Big]\right]. \tag{3}$$

Because of the stagewise independence assumption we have that if Z is a function of $(\xi_t, \ldots, \xi_T)$, then $\rho_{t|\xi_{[t-1]}}[Z] = \rho_t[Z]$. It follows that the cost-to-go functions do not depend on $\xi_{[t-1]}$ and can be written as

$$\mathcal{Q}_t(x_{t-1}) = \rho_t\left[\min_{\substack{B_t x_{t-1} + A_t x_t = b_t \\ x_t \geq 0}} c_t^\mathsf{T} x_t + \underbrace{\rho_{t+1}\Big[\cdots + \rho_T\Big[\min_{\substack{B_T x_{T-1} + A_T x_T = b_T \\ x_T \geq 0}} c_T^\mathsf{T} x_T\Big]\Big]}_{\mathcal{Q}_{t+1}(x_t)}\right], \quad t = 2, \ldots, T. \tag{4}$$

That is, we can write dynamic programming equations in the following recursive way. Set $\mathcal{Q}_{T+1}(\cdot)$ to be identically zero, and define going backward in time $t = T, \ldots, 2$,

$$Q_t(x_{t-1}, \xi_t) := \min_{\substack{B_t x_{t-1} + A_t x_t = b_t \\ x_t \geq 0}} c_t^\mathsf{T} x_t + \mathcal{Q}_{t+1}(x_t), \tag{5}$$

and

$$\mathcal{Q}_t(x_{t-1}) = \rho_t\left[Q_t(x_{t-1}, \xi_t)\right]. \tag{6}$$

Finally at the first stage the following problem should be solved

$$\min_{\substack{A_1 x_1 = b_1 \\ x_1 \geq 0}} c_1^\mathsf{T} x_1 + \mathcal{Q}_2(x_1). \tag{7}$$

For uniformity of notation, we set B_1 and x_0 as zero.

2.2 Solving the problem

One of the difficulties in solving dynamic programming equations (5)–(7) numerically is evaluation of the risk measure $\rho_t[\cdot]$. Recall that

$$\rho_t[Z] = (1-\lambda)\mathbb{E}[Z] + \lambda \mathsf{AV@R}_\alpha[Z]. \tag{8}$$

For continuous distributions evaluation of $\mathbb{E}[Z]$ and $\mathsf{AV@R}_\alpha[Z]$ involves computation of multivariate integrals. Typically these integrals cannot be written in a closed form and in high dimensional cases cannot be computed with high accuracy. In order to solve the problem numerically we use a discretization approach employing Monte Carlo sampling techniques.

We generate an iid sample of size N_t from the marginal distribution of random vector ξ_t, independently for each $t = 2, ..., T$. By assigning equal probabilities $1/N_t$ to each generated point, we create a discretization of the probability distribution of ξ_t. We denote by $\widehat{\xi}_t$ random vector having this discretized distribution. The obtained random process $\widehat{\xi}_2, \ldots, \widehat{\xi}_T$ can be represented by scenario tree with $N = \prod_{t=2}^{T}$ total number of scenarios. Note that in such construction the stagewise independence of the original process is preserved in the discretized process $\widehat{\xi}_2, \ldots, \widehat{\xi}_T$. In the numerical experiments we usually use the same number of generated samples at every stage time t.

We denote by $\widehat{\mathbb{E}}[\cdot]$ and $\widehat{\mathsf{AV@R}}_\alpha[\cdot]$ the expectation and the Average Value-at-Risk with respect to the process $\widehat{\xi}_2, \ldots, \widehat{\xi}_T$. Since the process $\widehat{\xi}_2, \ldots, \widehat{\xi}_T$ has a finite number of realizations, values $\widehat{\mathbb{E}}[\cdot]$ and $\widehat{\mathsf{AV@R}}_\alpha[\cdot]$ can be computed with a reasonable effort. Following eq.(8), we define the Monte Carlo estimate $\widehat{\rho}_t$ of the risk functional ρ_t as

$$\widehat{\rho}_t[Z] = (1-\lambda)\widehat{\mathbb{E}}[Z] + \lambda \widehat{\mathsf{AV@R}}_\alpha[Z]. \tag{9}$$

In that way we construct the so-called *Sample-Average Approximation* (SAA) of the "true" problem (1). The dynamic programming equations for the SAA problem can be written as (compare with (5)–(6))

$$\begin{aligned} \widehat{Q}_t(x_{t-1}, \widehat{\xi}_t) &= \min \quad c_t^{\mathsf{T}} x_t + \widehat{\mathcal{Q}}_{t+1}(x_t) \\ &\qquad \text{s.t.} \quad B_t x_{t-1} + A_t x_t = b_t,\ x_t \geq 0, \\ \widehat{\mathcal{Q}}_{t+1}(x_t) &= \begin{cases} \widehat{\rho}_{t+1}\left[\widehat{Q}_{t+1}(x_t, \widehat{\xi}_{t+1})\right] & , \quad t \in \{1, \ldots, T-1\}, \\ 0 & , \quad t = T, \end{cases} \end{aligned} \tag{10}$$

where $\widehat{\xi}_t = (c_t, A_t, B_t, b_t)$, for $t = 1, \ldots, T$. An important property of SAA approach is that solutions of (10) converge with probability 1 to a solution of the original (true) problem as the sample size increases. Since we can evaluate $\widehat{\rho}_{t+1}[\cdot]$, we focus on solving problem (10).

For solving problem (10) we use the *Stochastic Dual Dynamic Programming* (SDDP) algorithm. The SDDP algorithm belongs to the Nested Cutting Plane class of algorithms and is specialized for stagewise independent random processes. The main idea of such algorithms, based on cutting planes, is to approximate the objective function from *below* by a convex piecewise linear function and to refine this approximation iteratively. However, this approach is only possible if objective function and constraints of the corresponding optimization problem are *convex*. Indeed, the SAA approach (10) fulfills the convexity requirement on x variable for both cost-to-go and risk cost-to-go functions. We denote by $\mathfrak{Q}_{t+1}(\cdot)$ and $\underline{Q}_t(\cdot, \cdot)$ the lower approximations of $\widehat{\mathcal{Q}}_{t+1}(\cdot)$ and $\widehat{Q}_t(\cdot, \cdot)$

respectively, constructed with the SDDP algorithm. At each iteration, new cuts are added to the respective function improving its approximation:

$$\begin{array}{ccccc} \mathfrak{Q}_{t+1}^{\mathrm{old}}(x_t) & \leq & \mathfrak{Q}_{t+1}^{\mathrm{new}}(x_t) & \leq & \widehat{\mathcal{Q}}_{t+1}(x_t) \quad , \quad \forall x_t, \\ \underline{Q}_t^{\mathrm{old}}(x_{t-1}, \widehat{\xi}_t) & \leq & \underline{Q}_t^{\mathrm{new}}(x_{t-1}, \widehat{\xi}_t) & \leq & \widehat{Q}_t(x_{t-1}, \widehat{\xi}_t) \ , \quad \forall x_{t-1}, \ \forall \widehat{\xi}_t. \end{array}$$

A theoretical property of the SDDP algorithm is that it can compute an optimal solution of the SAA problem in a *finite* number of iterations. However, computational time for the final convergence can be extremely large. Therefore in practical applications the algorithm is stopped if there is no significant improvement in $\mathfrak{Q}_{t+1}(\cdot)$ over a given number of iterations.

In summary, for solving the true problem (1) with fixed λ and α, we approximate it by the corresponding SAA problem, and consequently solve the SAA problem by the SDDP algorithm. We will describe in the next section a procedure for choosing λ and α by using the solutions provided by this approach.

3 Choice of parameters in the risk averse approach

Ideally we would like to choose values of the parameters λ and α which are "best" in some sense.

3.1 Optimal policy estimation

The connection between solutions of the dynamic programming equations (5)–(6) and the nested formulation (1) is stated in the following theorem:

Theorem 3.1 *A policy $x_t^*(\xi_{[t]})$, $t = 1, \ldots, T$, is optimal if, and only if, for all $t = 1, \ldots, T$ and almost every realization of the data process ξ_t, it holds that*

$$x_t^*(\xi_{[t]}) \in \underset{x_t \in \mathbb{R}^{n_t}}{\arg\min} \left\{ c_t^\top x_t + \mathcal{Q}_{t+1}(x_t) : B_t x_{t-1}^*(\xi_{[t-1]}) + A_t x_t = b_t, \ x_t \geq 0 \right\}. \tag{11}$$

Because of (11), optimal policy is a function $x_t^* = x_t^*(x_{t-1}^*, \xi_t)$, $t = 1, ..., T$.

As pointed in section 2.2, we cannot solve the dynamic programming equations (5)–(6) directly. Instead, first we approximate problem (1) by the corresponding SAA problem (10). Similar to (11), an optimal policy $\widehat{x}_t(\xi_{[t]})$, $t = 1, \ldots, T$, of the SAA problem should satisfy the corresponding dynamic programming equations:

$$\widehat{x}_t(\xi_{[t]}) \in \underset{x_t \in \mathbb{R}^{n_t}}{\arg\min} \left\{ c_t^\top x_t + \widehat{\mathcal{Q}}_{t+1}(x_t) : B_t \widehat{x}_{t-1}(\xi_{[t-1]}) + A_t x_t = b_t, \ x_t \geq 0 \right\}, \tag{12}$$

for $t = 1, \ldots, T$. Note that the sampled process $\hat{\xi}_1, ..., \hat{\xi}_T$ of the SAA problem has a finite number of realizations (scenarios). Therefore the requirements for equations (12) to hold for *every* and *almost every* realization of the process $\hat{\xi}_1, ..., \hat{\xi}_T$ are the same. We emphasize that the SAA cost-to-go approximation $\widehat{\mathcal{Q}}_{t+1}(\cdot)$ also defines a feasible policy for the true problem when applying (12) to the true process $\xi_1, ..., \xi_T$. Additionally, it can be shown that as the sample sizes N_t, $t = 2, ..., T$, tend to infinity, the optimal value and optimal solutions of the SAA problem converge with probability one to their counterparts of the true problem.

As it was discussed in section 2.2, we solve the SAA problem by employing the SDDP algorithm. At the end of the iteration process we get an estimate the cost-to-go functions $\widehat{Q}_t(x_{t-1})$

and $\widehat{\mathcal{Q}}_{t+1}(x_t)$, which are denoted by $\underline{Q}_t(x_{t-1})$ and $\mathfrak{Q}_{t+1}(x_t)$, respectively. We define the *SDDP estimator* of the optimal policy as

$$\mathfrak{x}_t(\xi_{[t]}) \in \arg\min_{x_t \in \mathbb{R}^{n_t}} \left\{ c_t^\top x_t + \mathfrak{Q}_{t+1}(x_t) : B_t \mathfrak{x}_{t-1}(\xi_{[t-1]}) + A_t x_t = b_t,\ x_t \geq 0 \right\} \tag{13}$$

for $t = 1, \ldots, T$. Note that equation (13) defines a feasible policy for the true as well as for the SAA problems, depending from what distribution the process $\xi_t = (c_t, A_t, B_t, b_t)$ is generated.

Considering the concepts introduced thus far, we are in a position to state the criterion for choosing λ and α.

3.2 Criterion

Suppose that performance of a policy $x_t = x_t(\xi_{[t]})$, $t = 1, \ldots, T$, can be assessed by the following statistic

$$\mathbb{E}\left[F_1(x_1, \xi_1) + \cdots + F_T(x_T, \xi_T)\right], \tag{14}$$

where $F_t(x_t, \xi_t)$ is a *selected index* of decision x_t and observation ξ_t at stage t. For example, in the long term operation planning, if the selected index is the deficit (load curtailment) cost the resulting statistic (14) is the expected value of total deficit cost. Assuming that smaller value of statistic (14) the better, we can formulate the *ideal* criterion for choosing λ and α as follows. Consider an optimal solution $x_t^* = x^*(\xi_{[t]})$, $t = 1, ..., T$, of the "true" problem (1). Such an optimal solution depends on the parameters λ and α used in the respective risk measure (2). We write this explicitly by denoting this optimal solution as $x_t^*(\lambda, \alpha) = x_t^*(\xi_{[t]}, \lambda, \alpha)$, $t = 1, ..., T$, emphasizing that it is also a function of λ and α. Consequently we would like to choose $\lambda \in [0, 1]$ and $\alpha \in (0, 1)$ as optimal solutions of the following problem

$$\min_{\lambda, \alpha} \mathbb{E}\big[F_1(x_1^*(\lambda, \alpha), \xi_1) + \cdots + F_T(x_T^*(\lambda, \alpha), \xi_T)\big]. \tag{15}$$

In section 3.1, we emphasized that the optimal policy cannot be computed with a high accuracy. Therefore, we resort to the SAA optimal policy estimator (12) to formulate a criterion which approximates the ideal criterion (15). Considering this approach, an alternative objective function is the *plug-in* estimator:

$$\mathbb{E}_{|\widehat{\xi}}\left[F_1(\widehat{x}_1, \xi_1) + \cdots + F_T(\widehat{x}_T, \xi_T)\right], \tag{16}$$

where $\widehat{x}_t = \widehat{x}_t(\xi_{[t]}, \lambda, \alpha)$, for $t = 1, \ldots, T$. Note that the policies $\widehat{x}_t = \widehat{x}_t(\xi_{[t]}, \lambda, \alpha)$ are functions of the sampled process $\widehat{\xi}$ used in the corresponding SAA problem. The sampled process $\widehat{\xi}$ is generated by Monte Carlo sampling and in itself is random, and hence in that way the corresponding scenario tree is also random. Therefore the expectation (16) is conditional on the sampled process and hence is a random variable. As such it is a subject to variations inherited from the respective variations of sampled scenario trees. A less sensitive estimator of the objective function (15) is the expected value of (16) with respect to the probability distribution of sampled scenario trees:

$$\mathbb{E}\left[\mathbb{E}_{|\widehat{\xi}}\left[F_1(\widehat{x}_1, \xi_1) + \cdots + F_T(\widehat{x}_T, \xi_T)\right]\right]. \tag{17}$$

Note that the computational effort to evaluate (17) can be prohibitive. Moreover, it is reasonable to assume that both estimators (16) and (17) converge to the ideal objective function (15) as the sample size increases.

In summary, the plug-in estimator (16) is sensitive to variations of the sampled SAA trees. The expected plug-in estimator (17) is more robust, but requires computationally intensive calculations to estimate SAA policy for several probability trees. In the next subsections we will describe with more details how we solved approaches (16) and (17).

3.2.1 Single probability tree approach

Considering the plug-in estimator (16), we can approximate the ideal criterion (15) by the following problem:

$$\min_{\lambda,\alpha} \mathbb{E}\big[F_1(\widehat{x}_1(\lambda,\alpha),\xi_1) + \cdots + F_T(\widehat{x}_T(\lambda,\alpha),\xi_T)\big]. \tag{18}$$

As seen in section 3.1, the SDDP algorithm can be used to solve the SAA problem and therefore we approximate problem (18) by

$$\min_{\lambda,\alpha} \mathbb{E}\left[F_1(\mathfrak{x}_1(\lambda,\alpha),\xi_1) + \cdots + F_T(\mathfrak{x}_T(\lambda,\alpha),\xi_T)\right], \tag{19}$$

where $\mathfrak{x}_t(\lambda,\alpha)$, $t = 1,\ldots,T$, is a policy produced by the SDDP algorithm with given $\lambda \in [0,1]$ and $\alpha \in (0,1)$.

Finally, we estimate the objective function expectation of (19) by averaging the corresponding Monte Carlo samples:

$$\min_{\lambda,\alpha} \frac{1}{N_{\text{scen}}} \sum_{\omega=1}^{N_{\text{scen}}} F_1(\mathfrak{x}_1^{\omega}(\lambda,\alpha),\xi_1^{\omega}) + \cdots + F_T(\mathfrak{x}_T^{\omega}(\lambda,\alpha),\xi_T^{\omega}), \tag{20}$$

where ξ_t^{ω}, $t = 1,...,T$, is a sampled process from the true process, for $\omega = 1,\ldots,N_{\text{scen}}$, and $\mathfrak{x}_t^{\omega}(\lambda,\alpha) = \mathfrak{x}_t(\xi_{[t]}^{\omega},\lambda,\alpha)$, for $t = 1,\ldots,T$.

3.2.2 Several probability trees approach

In this section, we consider the case when one can calculate several SAA optimal policies $\widehat{x}_t = \widehat{x}_t(\xi_{[t]},\lambda,\alpha)$, for $t = 1,\ldots,T$, each one associated with a sampled SAA tree. Considering the expected plug-in estimator (17), we can approximate the ideal criterion (15) by the following problem:

$$\min_{\lambda,\alpha} \mathbb{E}\left[\mathbb{E}_{|\widehat{\xi}}\big[F_1(\widehat{x}_1(\lambda,\alpha),\xi_1) + \cdots + F_T(\widehat{x}_T(\lambda,\alpha),\xi_T)\big]\right]. \tag{21}$$

We use the SDDP algorithm to solve the SAA problem and therefore we approximate problem (21) by

$$\min_{\lambda,\alpha} \mathbb{E}\left[\mathbb{E}_{|\widehat{\xi}}\big[F_1(\mathfrak{x}_1(\lambda,\alpha),\xi_1) + \cdots + F_T(\mathfrak{x}_T(\lambda,\alpha),\xi_T)\big]\right], \tag{22}$$

where $\mathfrak{x}_t(\lambda,\alpha)$, $t = 1,\ldots,T$, is a policy produced by the SDDP algorithm with given $\lambda \in [0,1]$ and $\alpha \in (0,1)$. Consequently the outer expectation in the expected plug-in estimator (22) is estimated by the respective averaging. This leads to the following optimization problem, for $\vartheta = 1,\ldots,N_{\text{trial}}$ sampled SAA trees and $\omega = 1,\ldots,N_{\text{scen}}$ scenarios,

$$\min_{\lambda,\alpha} \frac{1}{N_{\text{trial}}N_{\text{scen}}} \sum_{\vartheta=1}^{N_{\text{trial}}} \sum_{\omega=1}^{N_{\text{scen}}} F_1(\mathfrak{x}_1^{\vartheta,\omega}(\lambda,\alpha),\xi_1^{\omega}) + \cdots + F_T(\mathfrak{x}_T^{\vartheta,\omega}(\lambda,\alpha),\xi_T^{\omega}). \tag{23}$$

4 Long term operation planning problem

The dynamic programming formulation for the long term operation planning problem can be written as:

$$\begin{aligned} Q_t([v_t, a_{[t-p,t-1]}], \eta_t) = \quad & \min \quad \sum_{k\in K}\left(\sum_{j\in T_k} c_j g_{tj} + \sum_{i\in U_k} \tilde{c}_i Def_{ti}\right) + \beta \mathcal{Q}_{t+1}([v_{t+1}, a_{[t-p+1,t]}]) \\ & s.t. \quad v_{t+1} = v_t + a_t + q_t + s_t \\ & \quad a_t = \text{diag}\,(\eta_t)(\Phi_{t,0} + \textstyle\sum_{\nu=1}^{p} \Phi_{t,\nu} a_{t-\nu}) \\ & \quad q_{tk} + \textstyle\sum_{j\in T_k} g_{tj} + \sum_{i\in U_k} Def_{ti} + \sum_{l\in\Omega_k}(f_{tlk} - f_{tkl}) = d_{tk}, \quad \forall k \in K \\ & \quad 0 \le v_{t+1} \le \overline{v}, \quad 0 \le q_t \le \overline{q}, \quad 0 \le s_t, \\ & \quad \underline{g} \le g_t \le \overline{g}, \quad 0 \le Def_t \le \overline{Def}, \quad \underline{f} \le f_t \le \overline{f} \end{aligned} \tag{24}$$

$$\mathcal{Q}_{t+1}([v_{t+1}, a_{[t-p+1,t]}]) = \begin{cases} \rho_{t+1}\left[Q_{t+1}([v_{t+1}, a_{[t-p+1,t]}], \eta_{t+1})\right] & , \quad t \in \{1, \ldots, T-1\} \\ 0 & , \quad t = T \end{cases}, \tag{25}$$

for all $t = 1, \ldots, T$.

The risk measure used considers a mean-AV@R convex combination, given by:

$$\rho_t[Z_t] = (1-\lambda)\mathbb{E}[Z_t] + \lambda \text{AV@R}_\alpha[Z_t]$$

with $\lambda \in [0, 1]$ and $\alpha \in [0, 1]$ being chosen parameters.

The *objective function*

$$\sum_{k\in K}\left(\sum_{j\in T_k} c_j g_{tj} + \sum_{i\in U_k} \tilde{c}_i Def_{ti}\right) + \beta \mathcal{Q}_{t+1}([v_{t+1}, a_{[t-p+1,t]}])$$

represents the sum of the total cost for thermal generation and deficit with $\mathcal{Q}_{t+1}([v_{t+1}, a_{[t-p+1,t]}])$, where:

- β discount factor;
- K subsystem set;
- T_k thermal set for subsystem k; and
- U_k deficit set for subsystem k.

The *energy balance* equation for each reservoir k is:

$$v_{t+1} = v_t + a_t + q_t + s_t$$

where:

- v_t stored energy in the reservoir at the beginning of stage t;
- a_t energy inflow during stage t;
- q_t generated energy during stage t; and
- s_t spilled energy during stage t.

The *time-series* model for the energy inflow is:

$$a_t = \text{diag}(\eta_t)\left(\Phi_{t,0} + \sum_{\nu=1}^{p} \Phi_{t,\nu} a_{t-\nu}\right)$$

where:

$a_{t-\nu}$	energy inflow during stage t;
$\Phi_{t,\nu}$	coefficient of PMAR vector time-series model; and
η_t	multiplicative noise of PMAR. (independent random variable)

The *load balance* equation, in MW month, for each subsystem k and stage t is:

$$q_{tk} + \sum_{j \in T_k} g_{tj} + \sum_{i \in U_k} Def_{ti} + \sum_{l \in \Omega_k} (f_{tlk} - f_{tkl}) = d_{tk}$$

where:

d_{tk}	load;
q_{tk}	hydro generation;
$\sum_{j \in T_k} g_{tj}$	thermal generation;
$\sum_{i \in U_k} Def_{ti}$	deficit generation;
$\sum_{l \in \Omega_k} (f_{tlk} - f_{tkl})$	net energy interchange;
f_{tlk}	energy flow from subsystem l to subsystem k; and
Ω_k	set of subsystems directly connected to subsystem k.

Bounds on variables are:

$0 \leq v_{t+1} \leq \overline{v}$	stored energy lower and upper bound;
$0 \leq q_t \leq \overline{q}$	generated energy lower and upper bound;
$0 \leq s_t$	non-negative spillage energy;
$\underline{g} \leq g_t \leq \overline{g}$	thermal generation lower and upper bound;
$0 \leq Def_t \leq \overline{Def}$	deficit energy lower and upper bound; and
$\underline{f} \leq f_t \leq \overline{f}$	energy flow lower and upper bound.

The main idea of the deficit is to penalize the load cut by a convex piecewise linear cost function which is dependent on the load cut depth. Regarding this approach, it is important to emphasize that the deficit variable with highest associated cost is unbounded above.

For each stage t the decision vector is $x_t = (v_{t+1}, q_t, s_t, g_t, Def_t, f_t, a_t)$. In the long term operation planning problem the only considered uncertainty is the independent multiplicative noise, that is, $\xi_t = \eta_t$, and the cost-to-go function depends only on $[v_t, a_{[t-p,t-1]}]$ of last p previous decision $x_{[t-p,t-1]}$. In this way, it is usual to write $Q_t([v_t, a_{[t-p,t-1]}], \eta_t)$ instead of $Q_t(x_{[t-p,t-1]}, \eta_t)$.

4.1 Criterion for choosing (λ, α)

Given a pair (λ, α), the current implementation of the risk averse SDDP method aims to minimize the average total cost with penalization of its high quantiles. Note that, for a given combination of the pair (λ, α), the performance of the resulting policy, as measured by considering additional criteria, may vary. For instance, one can consider several other variables of interest, apart from the total cost, to provide additional information regarding the performance of a proposed operation policy, such as deficit cost, thermal cost, deficit risk, etc. In this Report, the criterion we propose to choose among different pairs (λ, α) is to use the combination that results in the minimum expected total deficit cost.

For the "true" problem the criterion is given by:

$$\mathbb{E}\left[\sum_{k\in K}\sum_{i\in U_k}\tilde{c}_i\Big(Def^*_{1i}+\cdots+Def^*_{Ti}\Big)\right],$$

where $Def^*_{ti} = Def^*_{ti}(\xi_{[t]},\lambda,\alpha)$. Following the same reasoning as before, the SAA approximation is given by

$$\mathbb{E}\left[\mathbb{E}_{|\widehat{\xi}}\left[\sum_{k\in K}\sum_{i\in U_k}\tilde{c}_i\Big(\widehat{Def}_{1i}+\cdots+\widehat{Def}_{Ti}\Big)\right]\right],$$

where $\widehat{Def}_{ti} = \widehat{Def}_{ti}(\xi_{[t]},\lambda,\alpha)$. The corresponding SDDP approximation is

$$\mathbb{E}\left[\mathbb{E}_{|\widehat{\xi}}\left[\sum_{k\in K}\sum_{i\in U_k}\tilde{c}_i\Big(\mathfrak{Def}^{\vartheta,\omega}_{1i}+\cdots+\mathfrak{Def}^{\vartheta,\omega}_{Ti}\Big)\right]\right],$$

where $\mathfrak{Def}^{\vartheta,\omega}_{ti} = \mathfrak{Def}_{ti}(\xi^{\omega}_{[t]},\lambda,\alpha,\widehat{\xi}^{\vartheta}_{[T]})$. Finally, the expectation approximation is

$$\frac{1}{N_{\text{trial}}N_{\text{scen}}}\sum_{\vartheta=1}^{N_{\text{trial}}}\sum_{\omega=1}^{N_{\text{scen}}}\left[\sum_{k\in K}\sum_{i\in U_k}\tilde{c}_i\Big(\mathfrak{Def}^{\vartheta,\omega}_{1i}+\cdots+\mathfrak{Def}^{\vartheta,\omega}_{Ti}\Big)\right].$$

We denote, for further reference, the following estimate for each SAA tree

$$\mathfrak{CDef}^{\vartheta} = \frac{1}{N_{\text{scen}}}\sum_{\omega=1}^{N_{\text{scen}}}\left[\sum_{k\in K}\sum_{i\in U_k}\tilde{c}_i\Big(\mathfrak{Def}^{\vartheta,\omega}_{1i}+\cdots+\mathfrak{Def}^{\vartheta,\omega}_{Ti}\Big)\right]$$

for $\vartheta = 1,\ldots,N_{\text{trial}}$.

Similar reasoning follows if we substitute the deficit variable Def_{ti} by the thermal generation variable g_{ti}.

5 Numerical experiments

This section presents the numerical experiments that were performed in order to evaluate the proposed criterion for choosing the values of λ and α.

5.1 Case study

The numerical experiments described in this report 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 2015. This system 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 1).

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

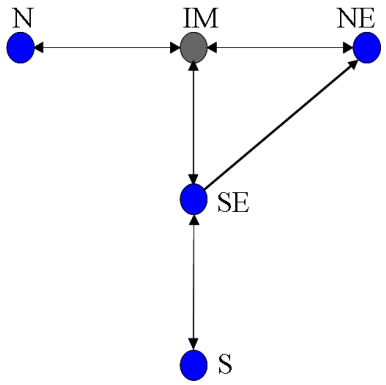


Figure 1: Aggregate representation of the Brazilian interconnected power system

amount of load curtailment accounts for load shortage at each area (Table 1). Interconnection limits between areas may differ depending on 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.

	% of total load curtailment	Cost
P1	0 – 5	1420.34
P2	5 – 10	3064.15
P3	10 – 20	6403.81
P4	20 – 100	7276.40

Table 1: Deficit costs and depths ($/MWh)

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 convex combination of the expectation and Average Value-at-Risk costs along the 120 months planning horizon, while supplying the area loads and obeying technical constraints. The total operation cost is the sum of thermal generating costs plus a penalty term that reflects energy shortage.

A scenario tree consisting of $1 \times N_2 \times N_3 \times \cdots \times N_{120}$ scenarios, for 120 stages, was constructed

		to				
		SE	S	NE	N	IM
from	SE	–	7500	1000	0	4000
	S	5470	–	0	0	0
	NE	600	0	–	0	3500
	N	0	0	0	–	∞
	IM	2940	0	3300	4407	–

Table 2: Interconnection limits between systems (MWave)

based on sampling of a periodic autoregressive multivariate statistical (PMAR) model with multiplicative error for the energy inflow record. In this (seasonal) model, the empirical distribution for each month and for every system is used to represent the corresponding noise distribution. The scenario tree is generated by sampling from these empirical distributions. The input data for this statistical model is based on 80 observations of the natural energy inflow (from year 1931 to 2010) for each of the considered 4 systems.

The study case general data, such as hydro and thermal plants data and interconnections capacities, were taken as static values throughout the planning horizon (120 months). The monthly seasonality of the demand is taken into account, but the annual value is constant. The energy inflows may vary along the stages.

5.2 Choice procedure

Closed form expressions for the minimum expected total deficit cost objective values, as functions of (λ, α), are not available. Therefore we resorted to computing objective values for a grid of λ and α values. We considered α values equal to 5%, 10%, 20% and 50%, and λ values ranging from 0 to 60%. Note that $\lambda = 0$ is equivalent to the risk neutral approach.

Let us also mention some involved computational issues. We are really interested in performance of different criteria for the "true" model with continuous distributions of the corresponding random vectors ξ_t. The discretization procedure is applied for computational purposes. The obtained so-called Sample Average Approximation (SAA) problems are functions of the constructed discretizations and as such are subject to variations. The Monte Carlo sampling approach allows to validate variability of the obtained optimal values for different sample sizes N_t. Previous numerical experiments [3] showed that for sample sizes $N_t = 100$, $t = 2, ..., T$, the computed optimal values became reasonably stable.

The $\mathsf{AV@R}_\alpha$ risk measure penalizes $N_t\alpha$ extreme scenarios, where N_t is the number of branching scenarios per stage. Therefore the number $N_t\alpha$ should not be "too small". For example for $N_t = 100$ and $\alpha = 0.1$ we have $N_t\alpha = 10$, which seems to be reasonable.

As mentioned in section 2.2, the main criterion proposed for choosing λ is the expected value of the total deficit cost. In order to select the pair (λ, α) that meet the criteria of minimizing the expected total deficit cost, the numerical experiments were performed considering:

- 20 policies defined with 20 sampled scenarios trees, each one generated from PMAR model by Monte Carlo sampling;
- for each scenario tree, the policy is defined considering as stopping criteria 3000 iterations, with one in-sample forward scenario for each iteration;
- for each of the 20 policies defined by 20 sampled scenario trees, 5000 index values were calculated considering a *fixed* out-of-sample set of 5000 energy inflow scenarios, each one with 120 stages (10 years), generated using the same PMAR model. Each index value was obtained adding the corresponding variable along the 120 stages.

5.3 (λ, α) choice

Figure 2 shows the deficit costs for the considered (λ, α) pairs. One can see that for every value of α it was possible to find a λ value that minimizes $\mathbb{E}[\mathfrak{CDef}]$. Moreover, although the optimal solution (λ, α) differs, the minimum $\mathbb{E}[\mathfrak{CDef}]$ value is approximately the same for all cases.

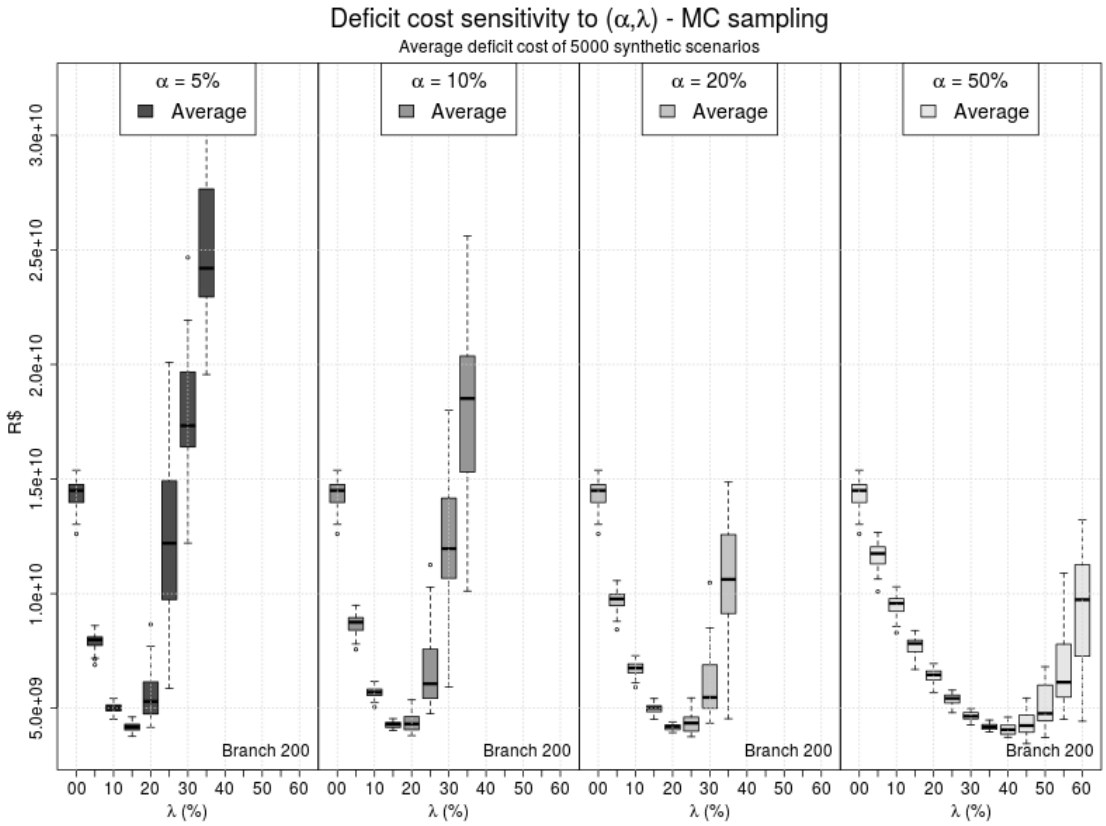
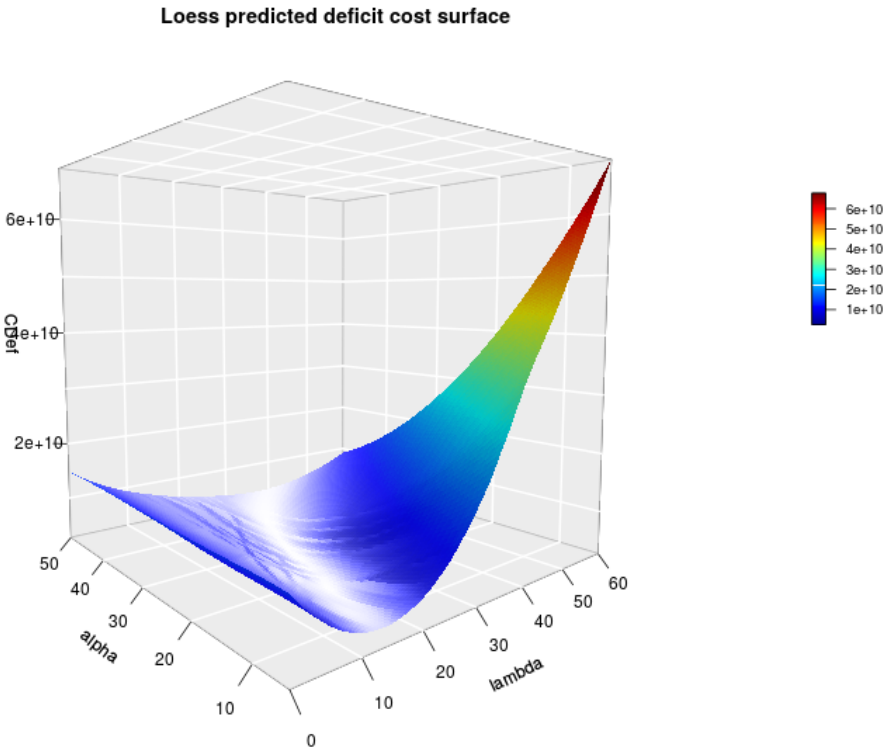
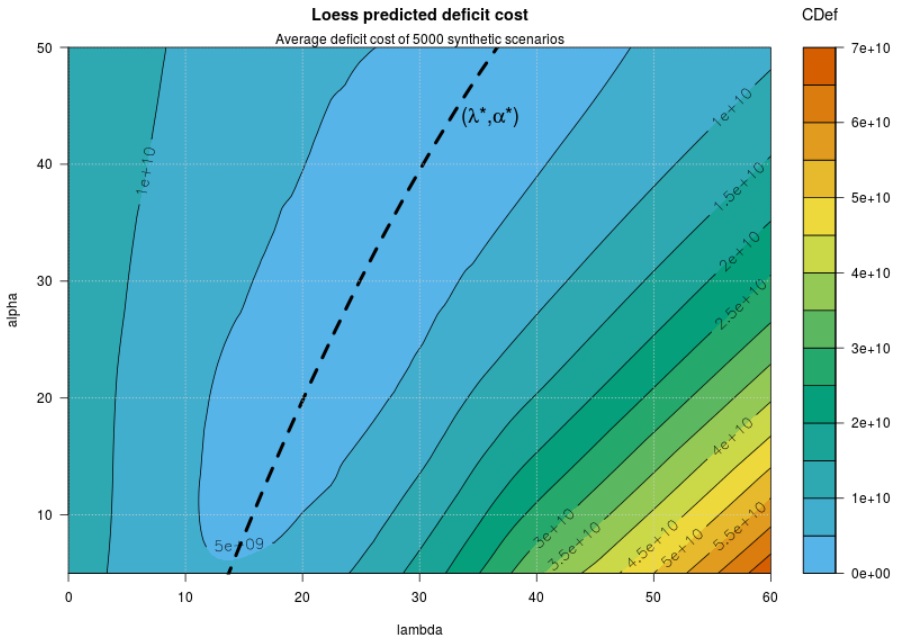


Figure 2: (λ, α) choice

Figure 3a shows a smooth surface adjusted to the experimental results using the non-parametric regression technique *loess* (locally weighted scatterplot smoothing). The corresponding adjusted contour curve, Figure 3b, suggests that we can always find the same minimum $\mathbb{E}[\mathfrak{CDef}]$ value for different combinations of λ and α.



(a) adjusted surface



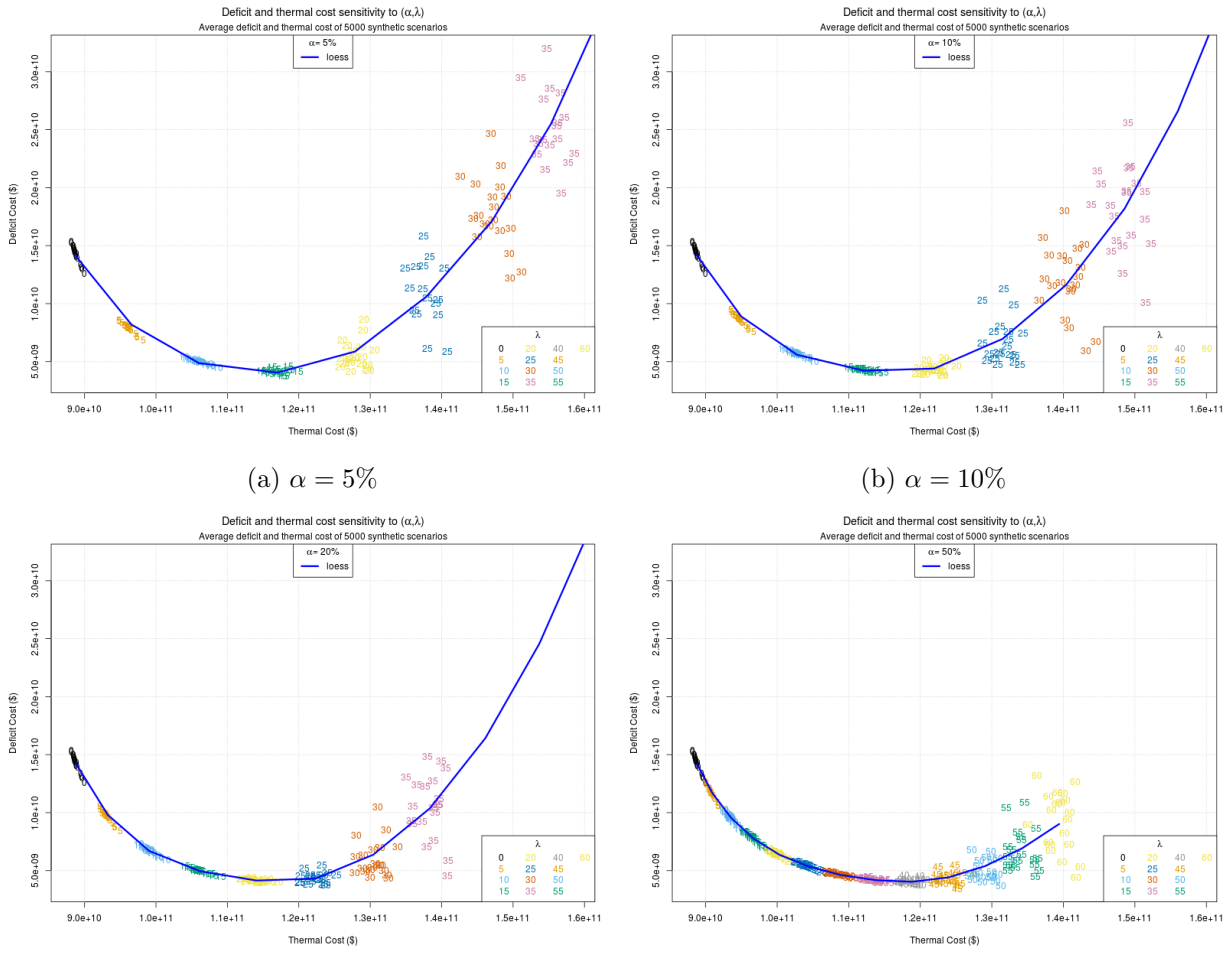
(b) adjusted contour curve

Figure 3: $\mathbb{E}[\mathfrak{CDef}] \times (\lambda, \alpha)$

5.4 Deficit and thermal costs

The results of section 5.3 lead to the conjecture that it is always possible to find a pair (λ, α) that meet the criterion of minimizing $\mathbb{E}[\mathfrak{CDef}]$. A natural question that arises is if one can propose an additional criterion to distinguish among these (λ, α) optimal pairs. A possible additional criterion is to minimize the total thermal costs for all (λ, α) minimum expected total deficit cost $\mathbb{E}[\mathfrak{CDef}]$. In order to assess these proposition, we plotted the average thermal generation costs against the average deficit cost along the planning horizon. Considering the smooth curve adjusted with loess to the data obtained for the α values 5%, 10%, 20% and 50%, Figure 4 shows that there is no clear distinction: the operating policies associated with (λ, α) corresponding to the minimum $\mathbb{E}[\mathfrak{CDef}]$ result in approximately the same values for the average thermal costs.

Additionally, these graphs show that the minimum $\mathbb{E}[\mathfrak{CDef}]$ sensitivity to the λ values diminishes as α increases. In this way, a cautious option is to select the greatest possible value for α.



(a) $\alpha = 5\%$

(b) $\alpha = 10\%$

(c) $\alpha = 20\%$

(d) $\alpha = 50\%$

Figure 4: Thermal and deficit operation cost ($)

5.5 Parameter revision

In this section we address the issue of when, if ever, the defined parameters (λ, α) have to be reassessed. In order to do so, we performed a simplified study regarding the sensitivity of the choice with respect to the relation supply/demand as a proxy to the variation of the power system configuration. The experiments considered that, for the same hydropower system, the demand would increase up to 110% of the designed system. Note that considering the demand equal to 110% is a significant departure from the demand for which the hydropower system was designed and implies significant differences in the corresponding operation policy.

Figure 5 shows that, given $\alpha = 5\%$:

- it was possible to identify a λ value that corresponds to the minimum expected value of the deficit cost;
- the $\mathbb{E}[\mathfrak{CDef}]$ differs according to the relation supply/demand, that is, the optimal pair (λ, α) depends on the system configuration;
- the proposed methodology can be applied for a rather large range of supply/demand relations.

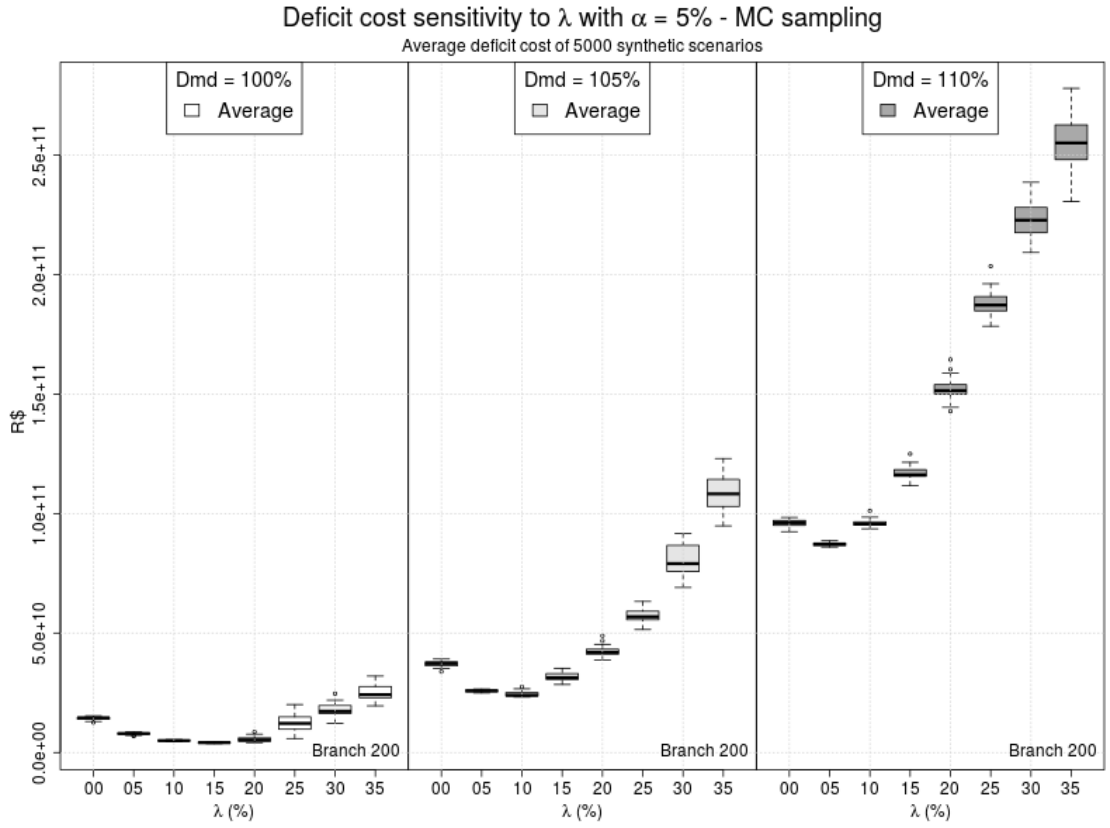


Figure 5: (λ, α) sensitivity to supply/demand relation

6 Algorithm

The reported numerical studies showed that:

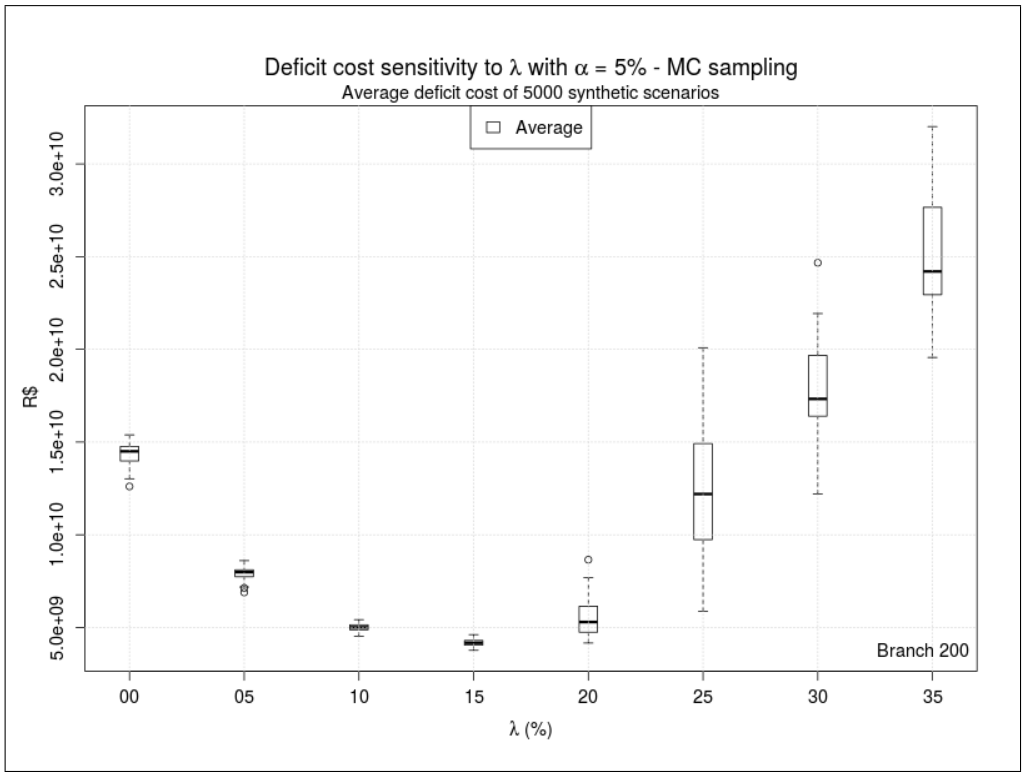
- for $\alpha \in (5\%, 50\%)$ there is a λ value for which the minimum expected value of the deficit cost are "equal";
- as the α value increases, the sensitivity of the estimation of minimum expected value of the deficit cost with λ diminishes;

In other words, the variability of the criterion with λ is smaller the greater the value of α is. Based in this reasoning, we propose to select the pair (λ, α) associated with the largest value for α that satisfies the min $\mathbb{E}[\mathfrak{CDef}]$ criterion. The following Algorithm 1 implements this procedure.

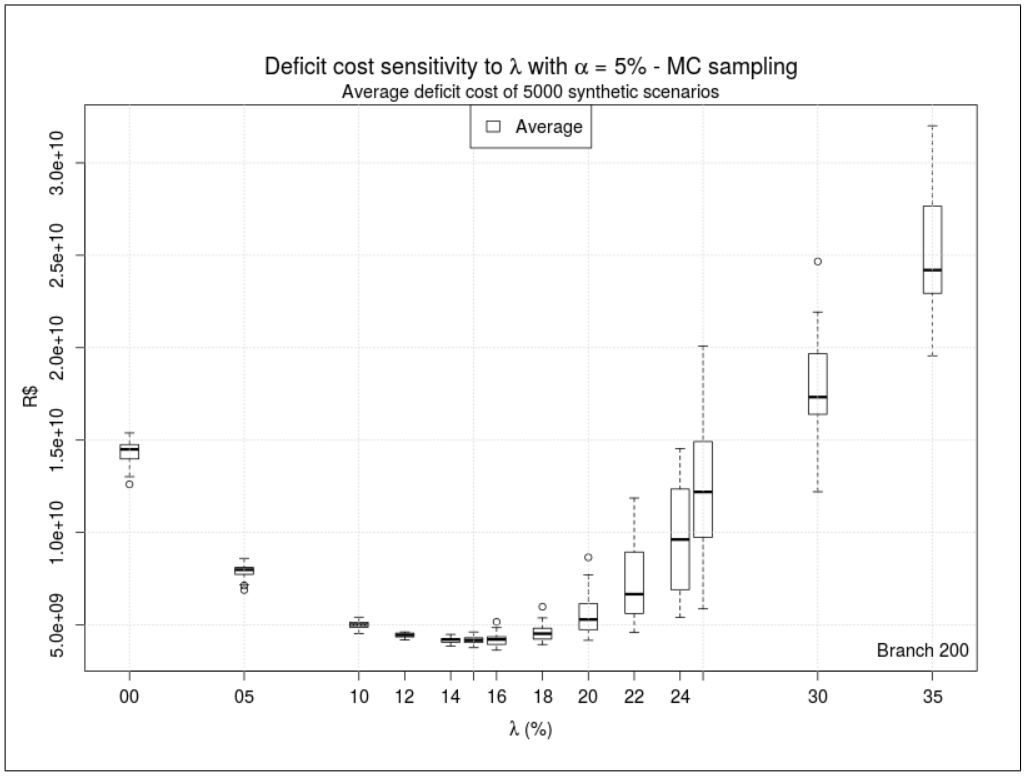
Algorithm 1: Algorithm to select (λ, α)

Initialization
  choose α_{ref} (=5%, e.g.)
  set $k \leftarrow 1$
Step 1 obtain reference value
  find $\lambda_{\text{ref}} \in [0, 1]$ that minimizes $\mathbb{E}[\mathfrak{CDef}]\big|_{\lambda, \alpha_{\text{ref}}}$
  define $V_{\text{ref}} \leftarrow \mathbb{E}[\mathfrak{CDef}]\big|_{\lambda_{\text{ref}},\, \alpha_{\text{ref}}}$
Step 2 find new λ and α
  increase α_k with $\alpha_k \in (0, 1)$
  find $\lambda_k \in [0, 1]$ that minimizes $\mathbb{E}[\mathfrak{CDef}]\big|_{\lambda, \alpha_k}$
  define $V_k \leftarrow \mathbb{E}[\mathfrak{CDef}]\big|_{\lambda_k, \alpha_k}$
Step 3 Convergence test
  if $V_k \leq V_{ref}$ **then**
    $k \leftarrow k + 1$
    go to **Step 2**
  else
    $(\lambda^*, \alpha^*) \leftarrow (\lambda_{k-1}, \alpha_{k-1})$
    return (λ^*, α^*)

One possible way to implement the proposed procedure is to start with a crude grid for λ and then refine the studies in the region of interest, as can be seen in Figure 6.



(a) Initial Monte Carlo sampling



(b) Refined Monte Carlo sampling

Figure 6: λ search procedure — $\mathbb{E}[\mathfrak{CDef}]$ ($)

7 Conclusion

This report aims at providing a methodology for defining the λ value to be used in mean-AV@R criterion of the operation planning policy procedure.

The envisaged criteria consists in finding the (λ, α) that minimizes the expected deficit cost. Given that there is no closed expression to evaluate the policies indexes as a function of (λ, α), a grid procedure was used. In order to discriminate indexes among different (λ, α)'s with SDDP algorithm it is necessary to have sufficient precision in the numerical experiments. In this report a value of branching 200 was used as it represents a good compromise between the necessary precision and associated computational effort. Appendix A details the numerical experiments that justify this option.

The results indicate that for α in the range from 5% to 50% it was possible to find a value for λ that minimizes the deficit cost average index, $\mathbb{E}[\mathfrak{CDef}]$. Moreover, for the given system configuration, it was possible to define a (λ, α) such that the $\mathbb{E}[\mathfrak{CDef}]$ is approximately the same. In order to get the less sensitivity of the proposed criterion with the (λ, α) pair, we suggest to choose the (λ, α) pair associated to the highest value for α.

On the other hand, the "optimal" (λ, α) pair is associated with the configuration of the system. Therefore, one must review its definition if the system configuration changes.

Acknowledgments

Authors would like to thank Sérgio Henrique Lisboa Mafra and Fabio Gibson Soares Santos for their support with the computational infra-structure which allowed the use of the Amazon Cloud Computing resources. The completion of the more than 2,200 numerical experiments was only possible due to their engagement in this research.

Appendix A Monte Carlo and K-means scenario tree generation

Recall that in the "true" model the involved random process $\xi_1, ..., \xi_T$, with ξ_1 being deterministic, is modeled as stagewise independent with d-dimensional random vectors ξ_t, $t = 1, ..., T$, having *continuous* distributions. In order to solve the corresponding multistage program numerically one has to discretize the continuous distribution of each ξ_t. In the Monte Carlo approach, a random sample of size N_t of ξ_t is generated. In an implicit way, this generates a scenarios tree with $N = \prod_{t=2}^{T} N_t$ number of scenarios.

Note that 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. This indicates that from the point of view of the number of scenarios, complexity of multistage stochastic programming problems grows exponentially with increase of the number of stages. In other words it is hopeless to try to solve multistage programs by enumerating all scenarios. The SDDP method approaches the problem in a different way by using an approximate dynamic programming techniques.

Yet an important question is how to discretize the continuous distribution of each random vector ξ_t. One can be tempted to resort to other approaches, apart from the Monte Carlo method, to discretization of the distribution of ξ_t. The aim of these approaches is to approximate the distribution of ξ_t in a certain uniform way. Since distribution of each ξ_t is approximated for different stages independently of each other, we drop the subscript t and consider a random vector ξ having probability distribution P supported on a set $\Xi \subset \mathbb{R}^d$.

Let us consider the Wasserstein distance approach (cf., Pflug and Pichler [1]). This approach aims at constructing finite approximation[1] $P_k = \sum_{i=1}^{k} p_i \delta(\mu_i)$ by solving the following optimization problem[2]

$$\min_{\mu_1,...,\mu_k \in \mathbb{R}^d} \int \min\left\{\|\xi - \mu_1\|^2, ..., \|\xi - \mu_k\|^2\right\} dP(\xi). \tag{26}$$

Given a solution $\mu_1^\star, ..., \mu_k^\star$ of problem (26), space $\mathbb{R}^d$ is partitioned into subsets Ξ_i consisting of those points of $\mathbb{R}^d$ which are closest to $\mu_i^\star$, $i = 1, ..., k$. Consequently to probabilities $p_i^\star$ are assigned values $P(\Xi_i)$, and $P_k^\star = \sum_{i=1}^{k} p_i^\star \delta(\mu_i^\star)$ is taken as the required discretization of P. An important property of problem (26), which is specific for Wasserstein distance of *order* $r = 2$, is the following relation between optimal partition and center points $\mu_i^\star$:

$$\mu_i^\star = \frac{\int_{\Xi_i} \xi dP(\xi)}{P(\Xi_i)}, \ i = 1, ..., k. \tag{27}$$

Note that problem (26) could have more than one optimal solution. Suppose for example that random vector ξ has standard normal distribution $N(0, I_d)$. Then for any orthogonal matrix[3] Q, the random vector $Q^\mathsf{T}\xi$ has also standard normal distribution[4] and the transformation $\mu_i \mapsto Q\mu_i$, $i = 1, ..., k$, will not change the objective value of problem (26). Indeed,

$$\|\xi - Q\mu_i\|^2 = (\xi - Q\mu_i)^\mathsf{T}(\xi - Q\mu_i) = (Q^\mathsf{T}\xi - \mu_i)^\mathsf{T}(Q^\mathsf{T}\xi - \mu_i) = \|Q^\mathsf{T}\xi - \mu_i\|^2$$

[1] By $\delta(\mu)$ we denote probability distribution having mass 1 at $\mu \in \mathbb{R}^d$. Unless stated otherwise, $\|x\| = \sqrt{x_1^2 + ... + x_k^2}$ denotes the Euclidean norm of vector $x \in \mathbb{R}^k$.

[2] Here we consider Wasserstein distance of order $r = 2$. This is motivated by applicability of the K-means algorithm discussed below.

[3] Matrix Q is orthogonal if $QQ^\mathsf{T} = I_d$, where I_d denotes $d \times d$ identity matrix.

[4] If $X \sim N(\mu, \Sigma)$ then, for any linear transformation T, it holds $TX \sim N(T\mu, T\Sigma T^\top)$.

has the same distribution as $\|\xi - \mu_i\|^2$.

In order to solve problem (26) we can apply the Sample Average Approximation (SAA) method. We can write problem (26) as

$$\min_{\mu} \mathbb{E}_P[F(\mu, \xi)], \tag{28}$$

where $\mu = (\mu_1, ..., \mu_k)$,

$$F(\mu, \xi) = \min\left\{\|\xi - \mu_1\|^2, ..., \|\xi - \mu_k\|^2\right\}$$

and $\mathbb{E}_P$ denotes the expectation with respect to probability distribution P. Let $\xi^1, ..., \xi^n$ be a random sample, of large size n, of the random vector ξ. The SAA method estimates the expectation $\mathbb{E}_P[F(\mu, \xi)]$ by the sample average $n^{-1}\sum_{j=1}^n F(\mu, \xi^j)$. That is, problem (26) is approximated by the SAA problem

$$\min_{\mu_1,...,\mu_k \in \mathbb{R}^d} \sum_{j=1}^n \min\left\{\|\xi^j - \mu_1\|^2, ..., \|\xi^j - \mu_k\|^2\right\}. \tag{29}$$

Given a solution $\mu_1^*, ..., \mu_k^*$ of problem (29), the set $S = \{\xi^1, ..., \xi^n\}$ is partitioned into subsets S_i consisting of those points of S which are closest to μ_i^*, $i = 1, ..., k$, and probabilities $p_i^\star$ are approximated by $p_i^* = |S_i|/n$.

It is known that an optimal solution of the SAA problem (29) converges w.p.1 as n tends to infinity to the set of optimal solutions of problem (26). In particular if problem (26) has unique optimal solution, then an optimal solution of the SAA problem (29) converges w.p.1 to this optimal solution (e.g., [2, section 5.1.1]). Also by the Law of Large Numbers the estimates $p_i^* = |S_i|/n$ converge w.p.1 to the respective probabilities $p_i^\star$.

To solve problem (29) the standard K-means algorithm, also referred to as Lloyd's algorithm, uses an iterative refinement technique. Given an initial set of k means $\mu_1, ..., \mu_k$, the algorithm proceeds by alternating between two steps:

Assignment step. The set S is partitioned into subsets S_i, $i = 1, ..., k$, according to the minimal distances to the points $\mu_1, ..., \mu_k$.

Update step. Update points μ_i by the respective averages $\frac{1}{|S_i|}\sum_{\xi^j \in S_i} \xi^j$, $i = 1, ..., k$.

This algorithm converges when the assignments no longer change.

Remark 1 Since the distribution of random vector ξ is supported on the set Ξ with probability one, the sample vectors ξ^j belong to Ξ, $j = 1, ..., n$. If the set Ξ is convex, then it follows that the averages $\frac{1}{|S_i|}\sum_{\xi^j \in S_i} \xi^j$ also belong to Ξ. Therefore in case the set Ξ is convex, the iteration points μ_i in the K-means algorithm stay in the set Ξ.

The problem (29) is not convex and is computationally difficult[5]. The above K-means algorithm may converge to a local, rather than global, solution of problem (29). If the original problem (26) has one optimal solution $\mu^\star = (\mu_1^\star, ..., \mu_k^\star)$, then as it was pointed above a *globally* optimal solution of (29) converges with probability one to $\mu^\star$ as n tends to infinity. However, it could happen that problem (26) has more than one optimal solution. In that case an optimal solution of (29) converges to the *set* of optimal solutions of (26) and may oscillate between different points of that optimal set.

[5]Partitioning of the set S into optimal subsets S_i is an NP-hard problem.

Remark 2 As far as this procedure is used for generating scenarios for the SDDP algorithm let us make the following observations. If the problem (26) has unique optimal solution, then for large n, say $n = 10^4$ and $k = 50$, variability of an optimal solution of problem (29) could be a result of computing a local rather than global solution of problem (29). If problem (26) has more than one optimal solution, then this also may effect variability of optimal solutions of problem (29). In any case increasing the sample size n, say from $n = 10^4$ to $n = 10^5$ or even $n = 10^6$, does not help in verification quality of the constructed discretization of the "true" problem. That is, for large n, variability of the obtained discrete approximation $P_k^* = \sum_{i=1}^k p_i^* \delta(\mu_i^*)$ could be a result of not computing a global solution of the problem (29) or because problem (26) has more than one optimal solution. *In other words increasing the sample size n, used in the above K-means algorithm, does not necessarily leads to better approximation of the corresponding distribution and therefore to better approximation for the required integral values.* This can be compared with the Monte Carlo (MC) sampling approach which allows estimation of variability of the corresponding sample average approximations, and hence gives an indication of the required sample sizes N_t.

Remark 3 Let us point out that in order to obtain an approximating measure, of the form $P_k^\star = \sum_{i=1}^k p_i^\star \delta(\mu_i^\star)$, with the distance (to the "true" measure P) no more than $\varepsilon > 0$, a total of $k = \mathcal{O}(\varepsilon^{-d})$ approximating points are necessary (cf., [1, p.147]). That is, the number k of points required to have an accurate approximation of the distribution P grows exponentially with the dimension d. However we do not need an accurate approximation of the continuous distribution P, but rather our goal is to obtain a good estimation of the integral values used in the respective objective functions. From this point of view we may compare accuracy of the K-means and the MC approaches as tools for constructing discretizations (scenarios) for the "true" problem with the *same number* of discretization points per stage. It could be noted that the MC method does not try to construct an accurate approximation of the continuous distribution P. Rather it is aimed at estimation of the corresponding integral values. Its complexity does not depend directly on the dimension d and is of order $\mathcal{O}(\varepsilon^{-2})$.

Another important advantage of the MC approach is that it allows to evaluate variability of the constructed approximations of the "true" problem by resampling of scenarios, i.e., by constructing several SAA problems and evaluating their variability. This is important in making a decision of how many discretization points per stage to use. On the other hand ideally the K-means reduction produces just one point (for given number k of discretization points) – coming from the optimal solution of problem (26). As it was pointed before, for large sample size n used in constructing the approximating problem (29), an observed variability of the discretization points produced by the K-means algorithm could be attributed to computing local optima rather than using different samples $\xi^1, ..., \xi^n$.

A.1 Monte Carlo and K-means policy evaluation

The purpose of this section is to examine the performance of both Monte Carlo sampling and K-means reduction approaches with respect to the variability of the results obtained. Monte Carlo sampling is the standard way to generate scenario trees and presents well known statistical properties. The main justifications for using the K-means reduction method are: this is the current approach used for generating scenarios tree in the official implementation used in Brazil and it is also a common technique proposed in the literature.

The numerical experiments setting is the same as the one described in section 5.3. All experiments reported in this appendix considered a fixed value of 15% for λ, 5% for α parameter, and

branching values N_t =25, 50, 75, 100, 125, 150, 175, 200, 225 for all stages t in $\{2, \ldots, 120\}$. As usual, the first stage is assumed known.

For each one of the 20 different scenario trees defined by the same branching size N_t, the corresponding policy were estimated using the SDDP algorithm. The performance of each one of these policies was assessed considering a *fixed* set of 5000 out-of-sample energy inflow scenarios generated by the same PMAR model used in the policy calculation. Preliminary analysis was based on a set of box plot graphs with 20 values for the objective function lower bound, average and conditional tail expectation (CTE) operation cost.

Figure 7 shows the box plot for the 20 values corresponding to the lower bound of the objective function for both Monte Carlo sampling and K-means reduction. One can see that, as the branching increases from $N_t = 25$ up to $N_t = 225$, the values of the lower bound costs for the Monte Carlo sampling show no trend and the associated variability decreases, as predicted by the theory. On the other hand, although the K-means reduction shows small variability for all branchings, the average values are underestimated with respect to the Monte Carlo sampling results and increase with branching. Similar behavior can be observed in the total average operation, see Figure 8a.

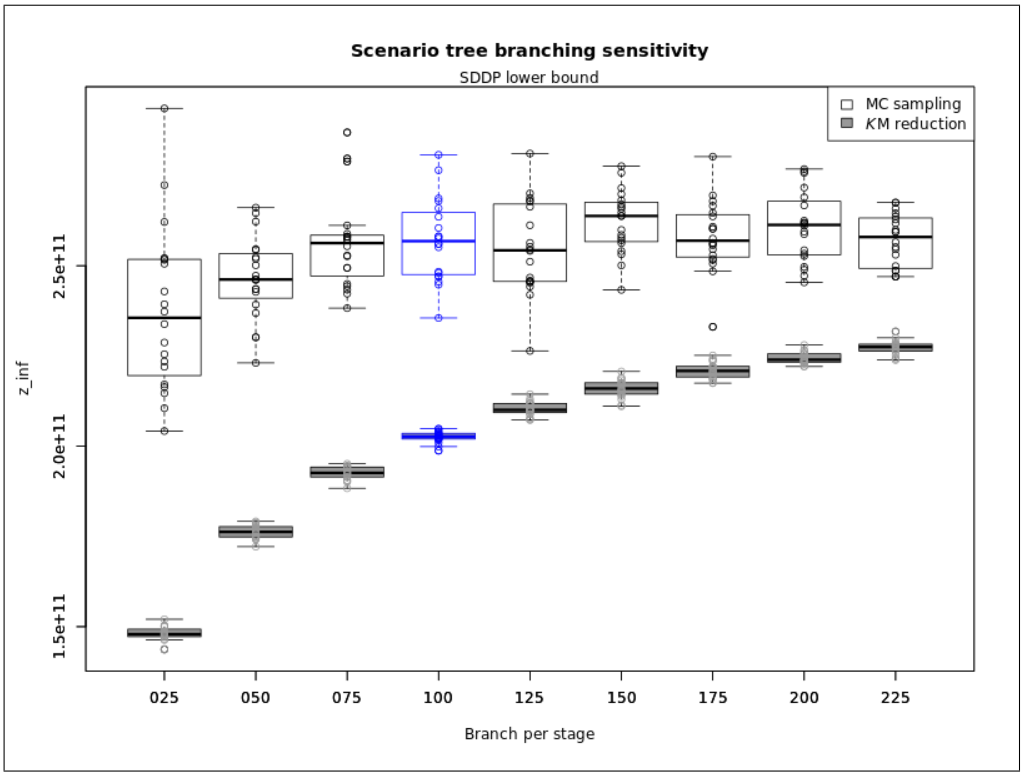


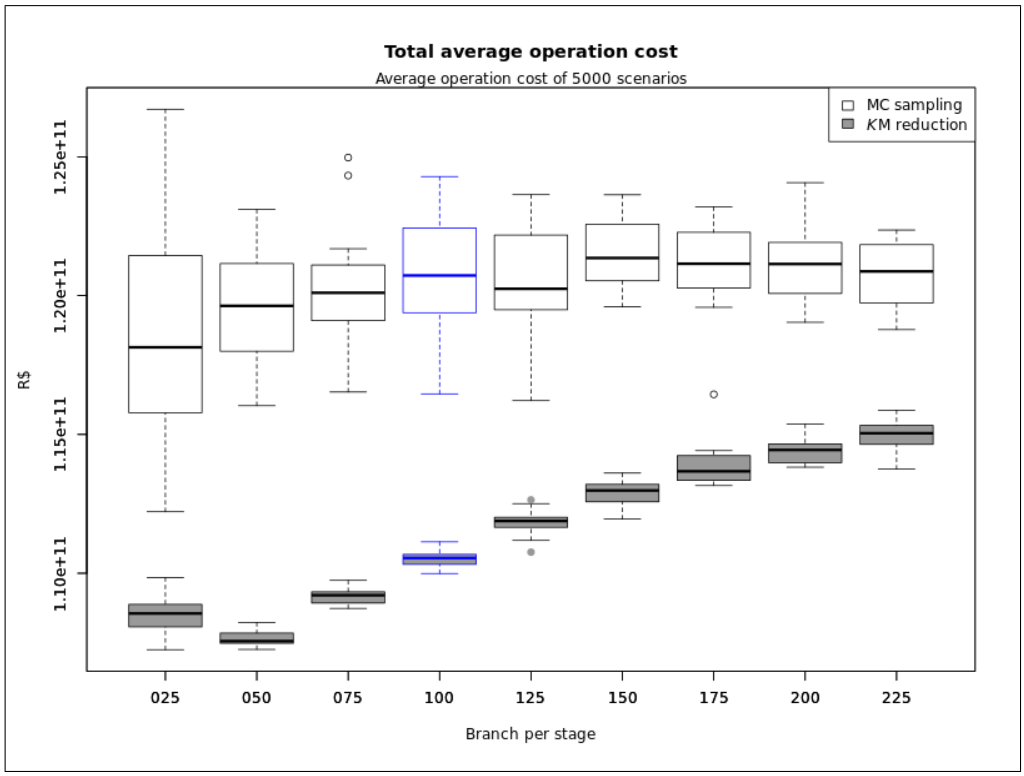
Figure 7: mean-AV@R approach objective function lower bound ($)

Considering branching $N_t = 25$, the 20 values for the lower bound of the objective function for MC sampling range from $200. \times 10^9$ to $300. \times 10^9$. In contrast, for K-means reduction the lower bound values range is about $140. \times 10^9$ to $160. \times 10^9$, see Figure 7. Table 3 shows the average and standard deviations results obtained considering Monte Carlo sampling and K-means reduction policies for the lower bound of the objective function, the operation costs and the conditional tail expectation for the 5% highest costs values, $\text{CTE}_{0.05}$.

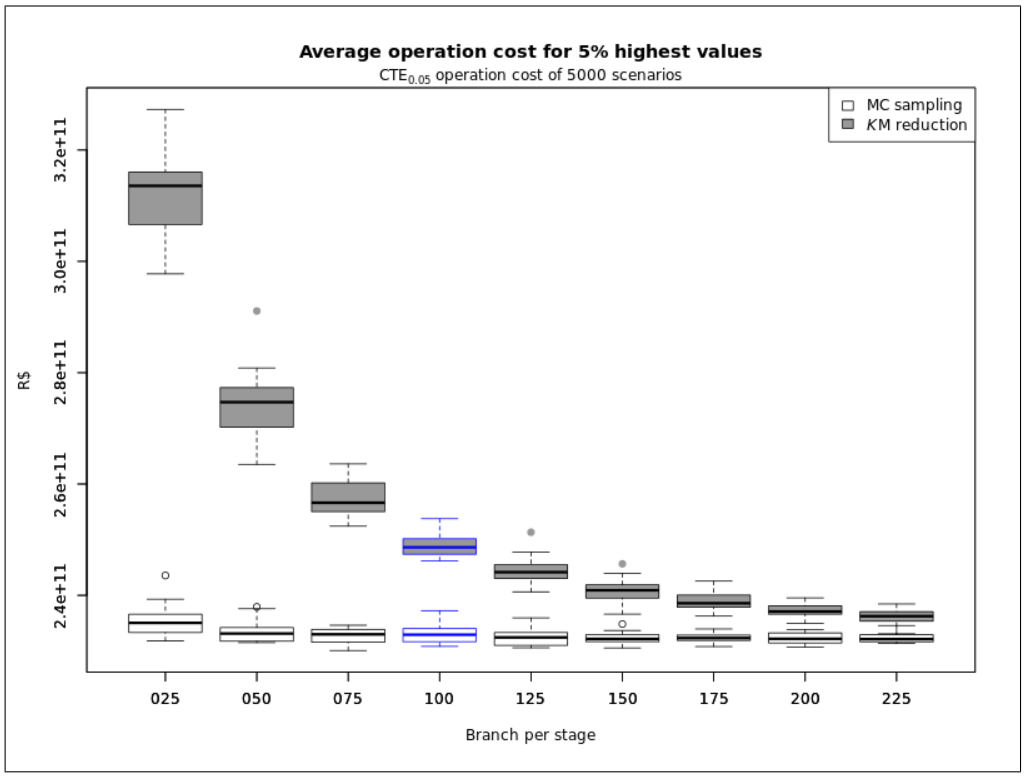
Scenario Tree		Lower Bound	Operation cost Average	$CTE_{0.05}$
Monte Carlo	average	237,512	118,646	235,382
	std dev	22,569	3,674	2,756
K-means	average	147,990	108,449	311,800
	std dev	2,039	660	7,026

Table 3: Risk averse policy evaluation for branching 25, $\lambda = .15$ and $\alpha = 0.05$ (10^6\$)

It is interesting to note in Figure 8b that $CTE_{0.05}$ for the MC sampling is smaller and relatively constant along the branching values when compared to the corresponding K-means reduction results. This can serve as a warning regarding the risk averse protection one gets with the use of the K-means reduction. For instance, considering again branching 25, one can be misled to think that using the K-means reduction the protection provided by risk averse policy results also in cheaper decisions for critical scenarios. However, while the policy derived by the use of K-means reduction incurs in relatively small increase for *average* operation costs when compared to the MC sampling (about 35% smaller), it turns out that for the more expensive scenarios the $CTE_{0.05}$ is about 30% greater than the ones incurred with the MC sampling, see Table 3. In other words, for the same branching value, the K-means reduction results in a policy that is more expensive for critical scenarios than the MC ones.



(a) Average



(b) $\text{CTE}_{0.05}$

Figure 8: mean-AV@R approach operation cost ($)

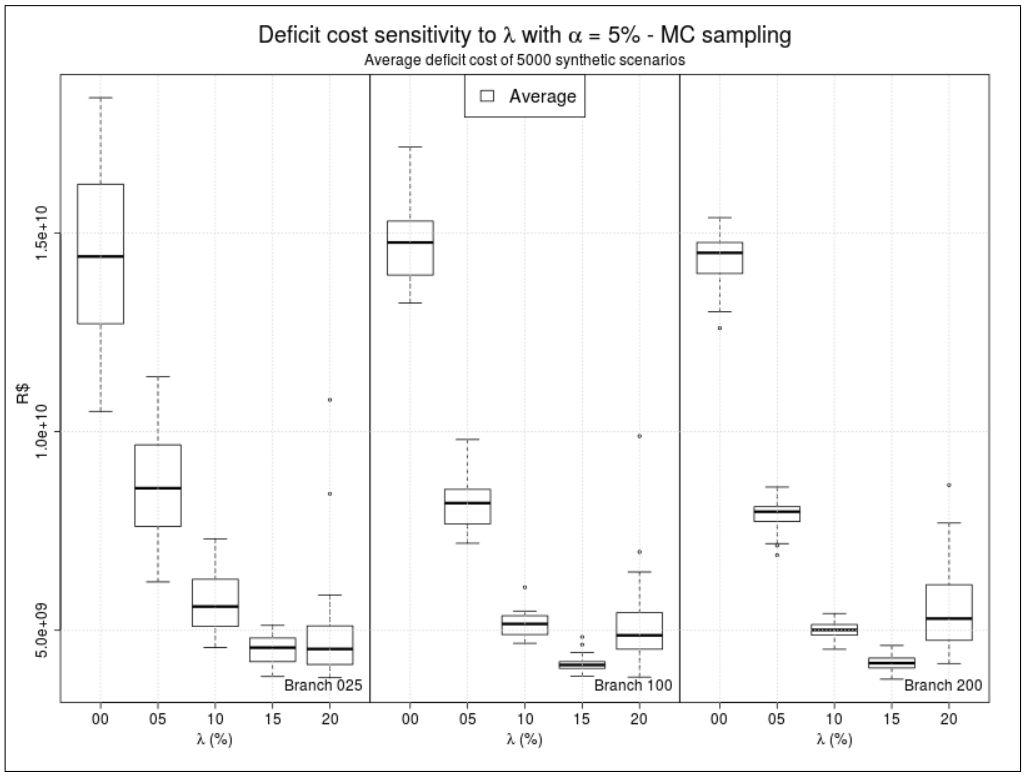
A.2 Setting the branching value

In this section, we will use both Monte Carlo sampling and K-means reduction to define a reasonable branching value that enables the discrimination among different λ's. Such branching should ideally be one value for the all λ values considered in the search procedure. That is, an unique scenario tree size should fit all the numerical experiments envisaged to define the mean-AV@R parameter choice. It is also important to point that the branching size has a great impact on policy computing time. Therefore, the chosen value for branching should represent a feasible compromise between precision and CPU effort.

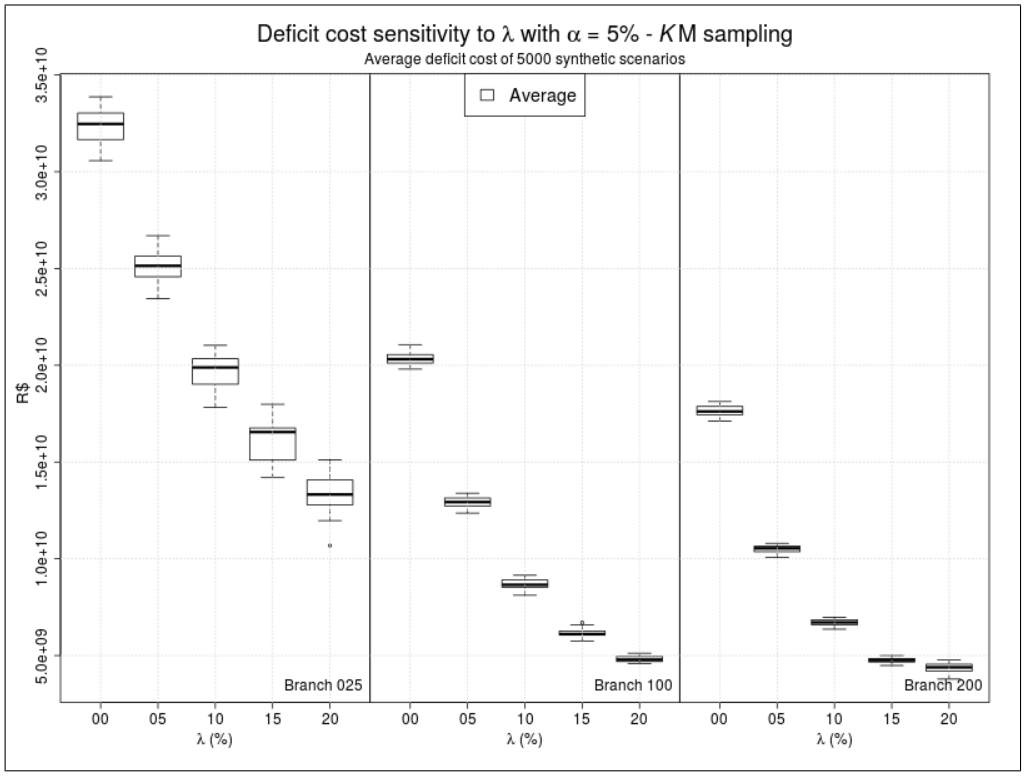
The numerical experiments consider $\lambda = 0.00$ (risk neutral), 0.10, 0.15 and 0.20 and branching values of 25, 100 and 200. Figures 9 and 10 show the average and the conditional tail expectation (5% highest values) of the total deficit costs. One can see that, for each given λ, the variability of average and $\text{CTE}_{0.05}$ costs decreases as branching increases. Also, the policies obtained with the K-means reduction result in more deficit costs than the ones associated with the Monte Carlo sampling.

Figures 11, 12 and 13 reinforces that the policies provided by the K-means reduction underestimate the necessary effort to meet the demand when compared to the ones obtained with Monte Carlo sampling.

As observed with results of the previous section, branching 25 shows large variability which reduces from branching 100 onwards. In this report, we will use branching 200 to perform the numerical experiments regarding λ choice.

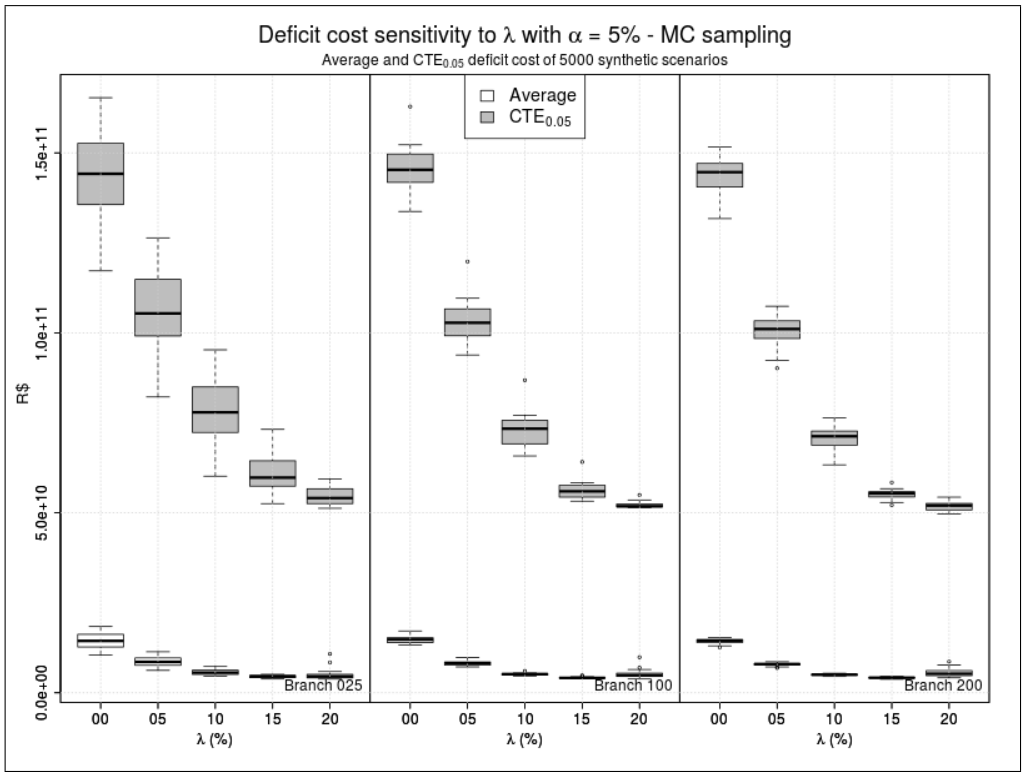


(a) Monte Carlo sampling

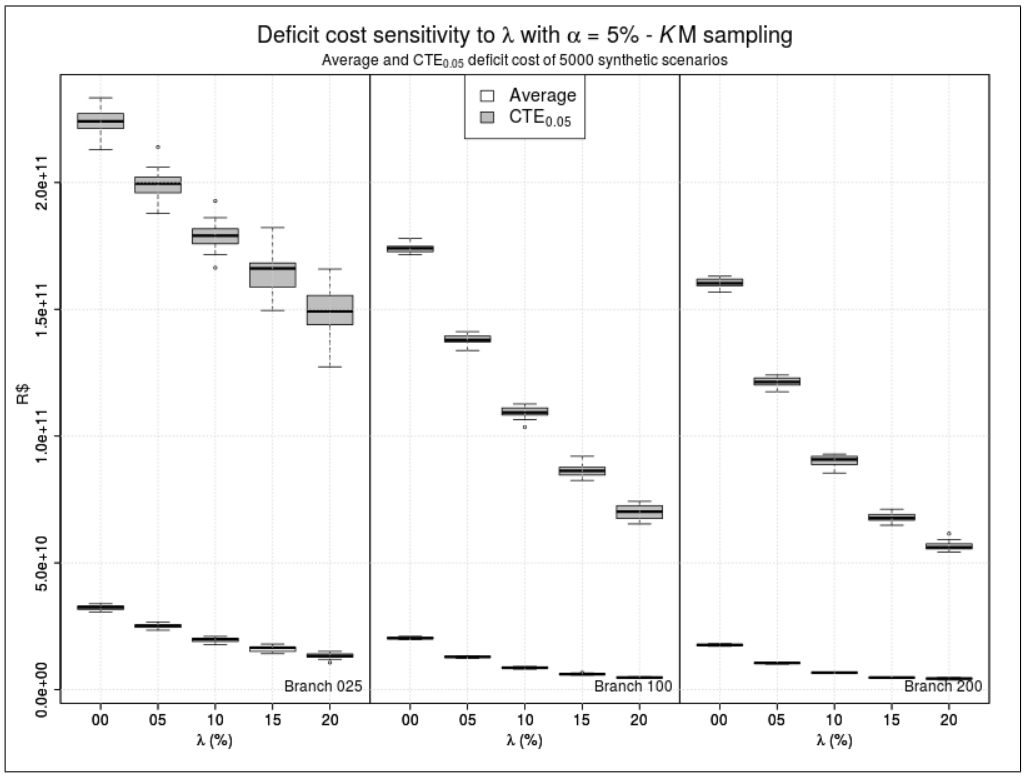


(b) K-means reduction

Figure 9: Total deficit cost, average ($)

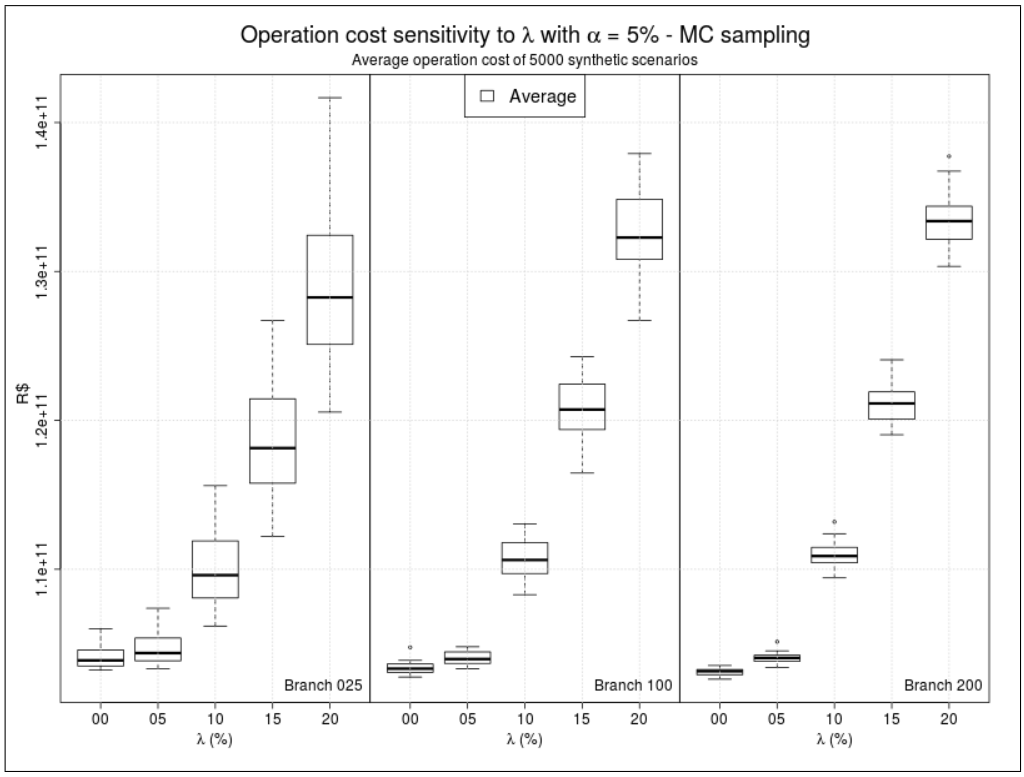


(a) Monte Carlo sampling

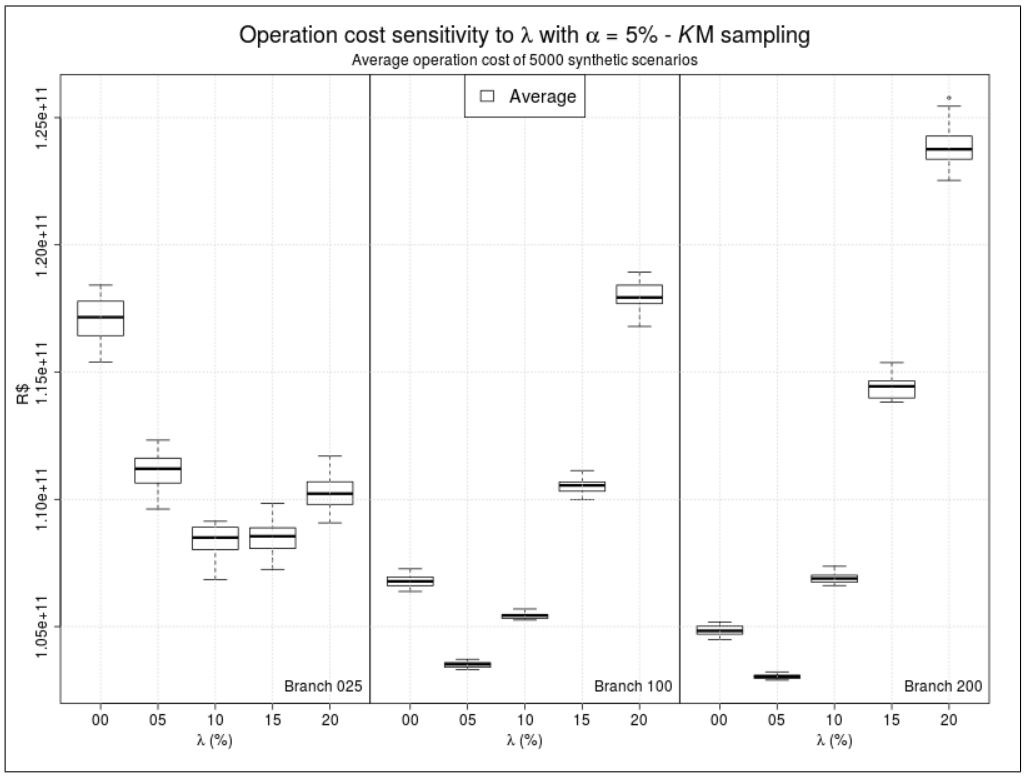


(b) K-means reduction

Figure 10: Total deficit cost, average and $\text{CTE}_{0.05}$ (\$)

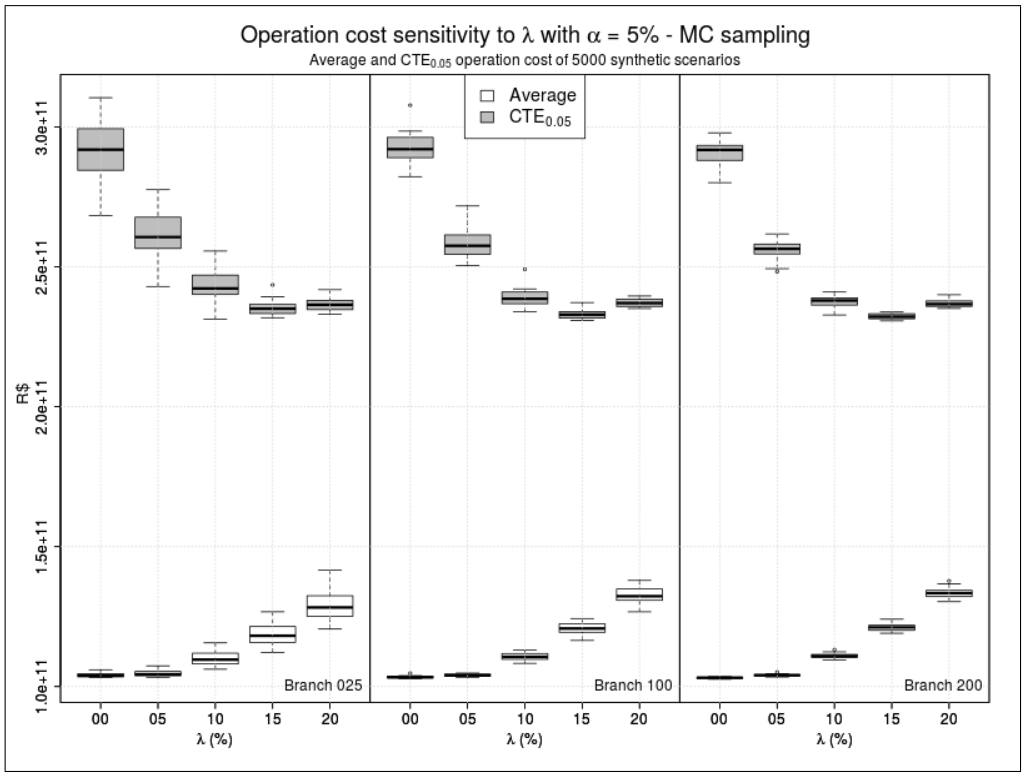


(a) Monte Carlo sampling

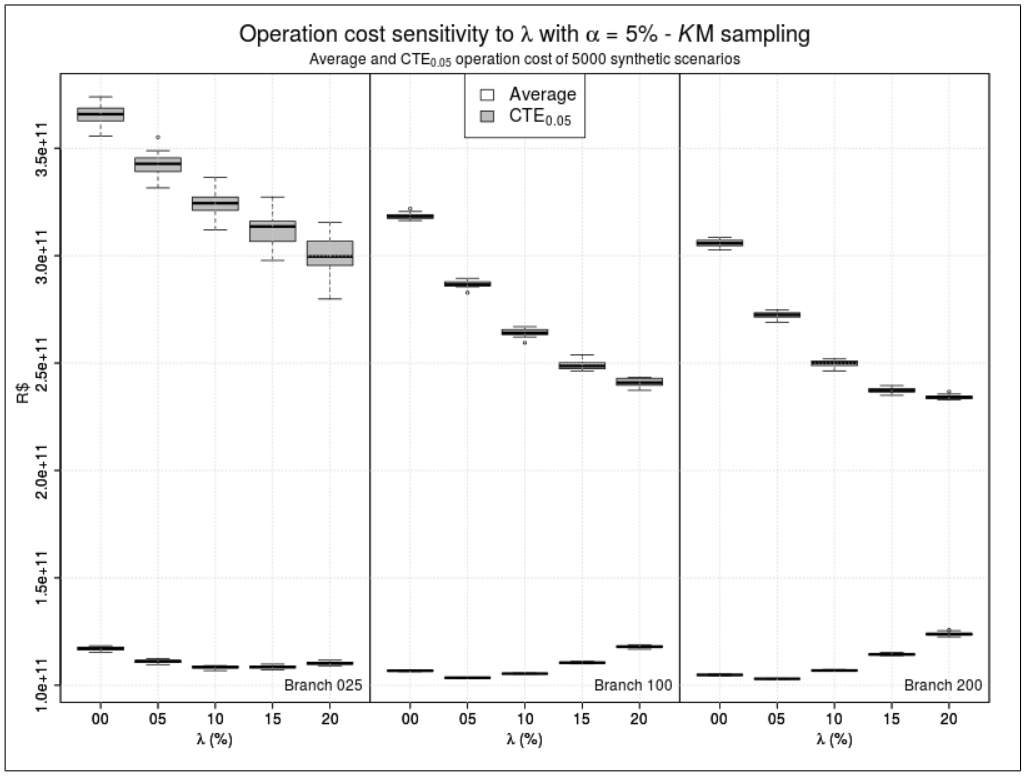


(b) K-means reduction

Figure 11: Total operation cost, average ($)

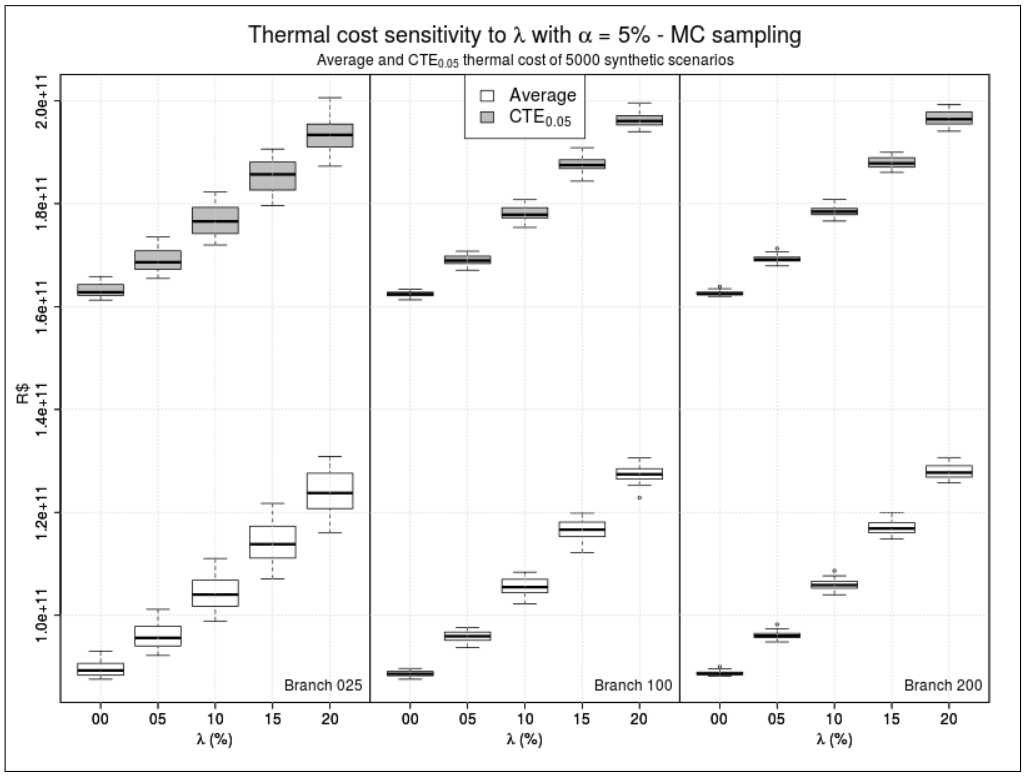


(a) Monte Carlo sampling

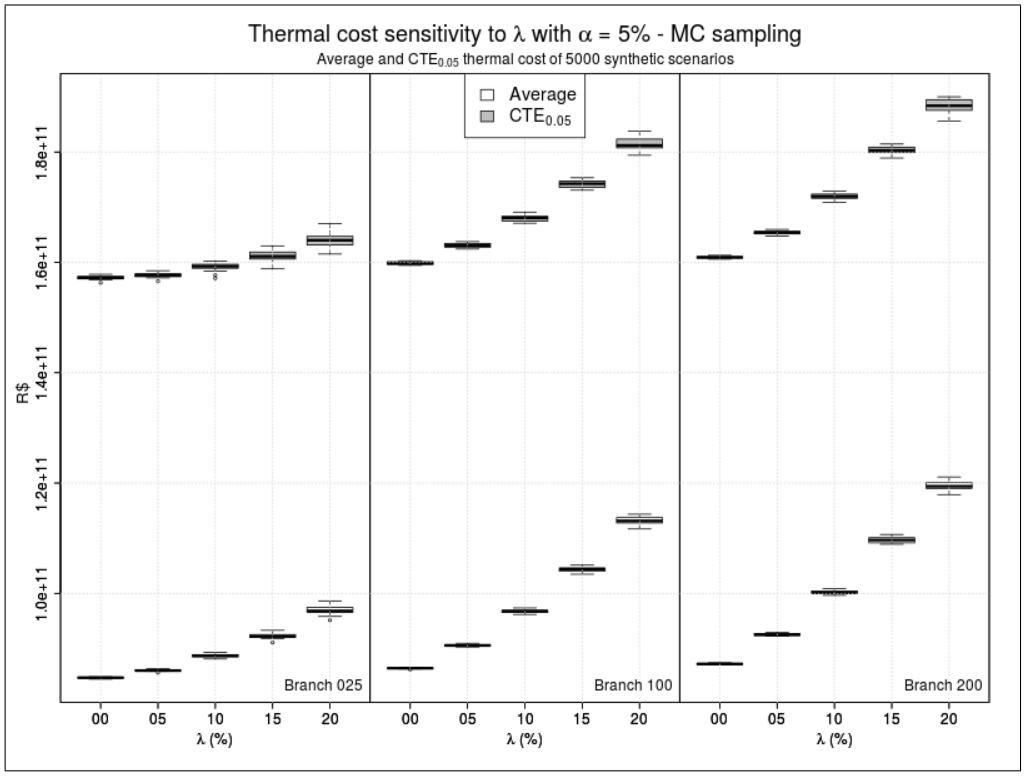


(b) K-means reduction

Figure 12: Total operation cost, average and $\text{CTE}_{0.05}$ (\$)



(a) Monte Carlo sampling



(b) K-means reduction

Figure 13: Total thermal operation cost, average and $\text{CTE}_{0.05}$ (\$)

A.3 Conclusion

In order to discriminate indexes among different (λ, α)'s performance for the policies defined with SDDP algorithm it is necessary to have sufficient precision in the numerical experiments. In this way, the scenario tree size was reassessed considering two generation methods: the Monte Carlo sampling and the K-means reduction approach.

The reported experiments show that:

- K-means objective function and total average operation costs show little variation for all λ considered. However, numerical results suggest that these values are underestimated when compared to the corresponding Monte Carlo sampling;
- MC sampling policies are more stable than the K-means reduction ones and result in better smoothing of high operation costs;
- branching equal to 25 produces large variation for policy indexes and therefore is insufficient for both operation planning and (λ, α) choice;
- although branching around 100 provides sufficient precision, in this report a value of branching equal to 200 was used as it represents a good compromise between the necessary precision and associated computational effort.

In light of this reasoning, we suggest that the Monte Carlo approach is better suited for use in the SDDP algorithm.

References

[1] G.Ch. Pflug and A. Pichler, *Multistage Stochastic Optimization*, Springer, 2014.

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

[3] Alexander Shapiro, Wajdi Tekaya, Joari Paulo da Costa and Murilo Pereira Soares, *Risk neutral and risk averse Stochastic Dual Dynamic Programming method*, Eur. J. Oper. Res., 224:375-391, 2013.