Stochastic Dual Dynamic Programming with Markov Chain and Covariate Information: An Application to Hydrothermal Economic Dispatch

André Ramos
Davi Michel Valladão
LAMPS PUC-Rio
Departamento de Engenharia Industrial PUC-Rio

RESUMO

A técnica de programação dinâmica estocástica dual (SDDP) é uma metodologia amplamente adotada para lidar com problemas de decisão em várias etapas sob incerteza. Essa técnica utiliza hipóteses de parâmetros incertos do problema para criar políticas de decisão considerando a realização de diferentes cenários desses parâmetros estocásticos. Portanto, uma representação realista dessas incertezas é fundamental para que os modelos SDDP tenham um bom desempenho. Embora abordagens tradicionais como cadeias de Markov sejam poderosas na representação de processos estocásticos, elas podem não capturar adequadamente as incertezas que dependem fortemente de variáveis exógenas, como o tempo. Neste estudo, investigamos o impacto da incorporação da dependência da variável exógena temporal na formulação da cadeia de Markov. Exploramos como essa representação de parâmetros incertos em um modelo SDDP influencia as políticas de decisão e analisamos os resultados usando análise fora da amostra.

PALAVRAS CHAVE. Programação dinâmica dual estocástica, SDDP, Cadeias de Markov, séries temporais e problema hidrotérmico.

PM – Programação Matemática, BDA – Big Data e Analytics, EN&PG - PO na Área de Energia, Petróleo e Gás

ABSTRACT

Stochastic dual dynamic programming (SDDP) technique is a widely adopted methodology to handle multi-stage decision problems under uncertainty. This technique utilizes hypothesis of uncertain parameters of the problem to create decision policies under the realization of different scenarios of these stochastic parameters. Therefore, a realistic representation of these uncertainties is fundamental to SDDP models perform well. While traditional approaches like Markov chains are powerful in representing stochastic processes, they may not adequately capture uncertainties that depend heavily on exogenous variables, such as time. In this study, we investigate the impact of incorporating the time exogenous variable dependency into the Markov chain formulation. We explore how this representation of uncertain parameters in an SDDP model influences the decision policies and analyze the results using out-of-sample analysis.

KEYWORDS. Stochastic dual dynamic programming, SDDP, Markov chains, time series, hydrothermal problem.

PM – Programação Matemática, BDA – Big Data e Analytics, EN&PG - PO na Área de Energia, Petróleo e Gás

DOI: 10.59254/sbpo-2023-174839
https://proceedings.science/p/174839?lang=pt-br

1. Introduction

Stochastic Dual Dynamic Programming (SDDP) techniques have emerged as a powerful approach for solving real-world problems involving medium to long-term decision-making. These problems require the allocation of limited resources across multiple time frames to minimize costs, while being influenced by uncertain factors. The quality of the decision policies obtained through SDDP models is directly linked to the accuracy and reliability of simulations conducted for the uncertain variables.

It is noteworthy that within the context of a multistage problem, the number of variables pertinent to the deterministic counterpart expands exponentially in direct correlation with the escalation of stages, leading to intractable problems. Consequently, the Stochastic Dual Dynamic Programming (SDDP) framework emerges as a judicious avenue for solving this kind of problem.

This paper focuses on the hydrothermal power system model, with a specific emphasis on the Brazilian context. The model involves an independent system operator (ISO) responsible for monthly planning the economic dispatch for power plants and water stored in reservoirs. The primary objective is to minimize expected costs while meeting energy demand. Additionally, an implicit goal is to mitigate energy price spikes that may arise in scenarios with low water storage in reservoirs. However, it is crucial to note that these medium- to long-term planning decisions must take into account short-term dispatch constraints and modeling considerations. This entails considering physical limitations in transmission lines, ensuring security criteria are met, addressing uncertainties related to inflow and demand, and accounting for other relevant factors, as highlighted in prior works Warland and Mo [2016], Martins et al. [2014] and Moreira et al. [2015].

Within the framework of Stochastic Dual Dynamic Programming (SDDP) methodology, a crucial aspect involves inferring a set of states that encapsulate information pertaining to the potential outcomes of uncertain variables in the problem. These states, as described by Dowson [2020], are defined as functions of historical data that encompass all the necessary information to model a system from a specific point in time onward. To address this, Markov Chains (MC) are employed Norris [1997], wherein the states are estimated based on historical data (for example utilizing clustering techniques), and state simulations are conducted using estimated transition matrices. The application of MC facilitates the modeling and analysis of complex systems characterized by several states and transitions.

In the study by Thuener et al. Silva et al. [2021], the authors tackle a multivariate decision problem by employing a distributionally robust optimization approach. To achieve this, they utilize a machine learning technique known as Hidden Markov Model (HMM). By leveraging HMM, they estimate a Markov state (MS) that captures the temporal dynamics of the uncertain parameters (y) in the problem. This estimated MS is then used to construct an ambiguity set for the uncertain variables (Y). The overall approach is illustrated in Figure 2, which provides a simplified representation of their methodology.

Note that each decision and uncertainty depends on the current Markov state. Therefore, the first stage decisions must be prescribed simulating the actual state in each time frame, and as it is a dynamic process, the multistage decisions can be done after estimating the current Markov state (but not knowing the current values of y). Note, however, that in their approach, there is not covariate data incorporated into the model.

To provide a more comprehensive perspective of this limitation, let us consider the example in the context of seasonal variations in uncertain inflow scenarios. In such cases, we may have distinct states representing different seasons, such as a dry season and a wet season. If the transition matrix remains constant for every time instant, it assumes that the probability of transitioning

DOI: 10.59254/sbpo-2023-174839
https://proceedings.science/p/174839?lang=pt-br

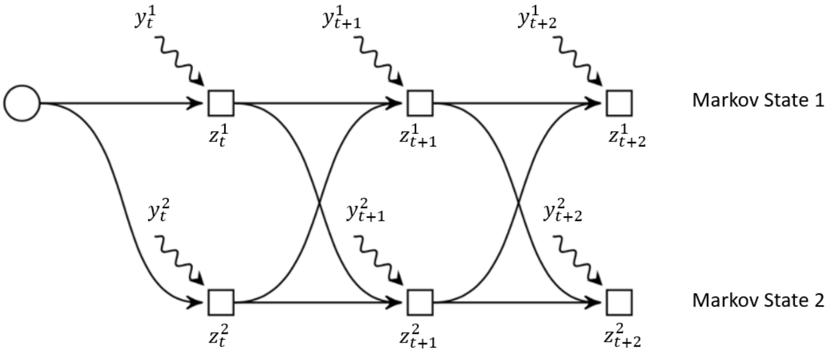


Figura 1: Adaptation of a Markovian policy graph of Dowson [2020]

from the dry season to the wet season at the beginning is the same as the probability at the end. However, this assumption fails to capture the temporal dynamics inherent in the transitions between seasons. In reality, the probabilities of transitioning from the dry season to the wet season may vary throughout the year, with different patterns observed at different stages of the seasons. Disregarding these temporal variations in transition probabilities can lead to a misrepresentation of the underlying system dynamics and undermine the accuracy of the decision-making process. Therefore, it is crucial to account for time-dependent changes in transition probabilities to ensure a more accurate assessment of the impact of uncertainties in decision models, particularly when dealing with seasonal variations.

We propose a new approach where the Markov Chain is estimated considering both the uncertain variables Y and the contextual variables X. We then introduce a Markov chained Distributionally Robust Optimization (DRO) combined with Stochastic Dual Dynamic Programming (SDDP) framework to obtain a contextual policy, where the decision z_t depends on the observed contextual variable x_t. The schematic representation of our proposed approach is shown below:

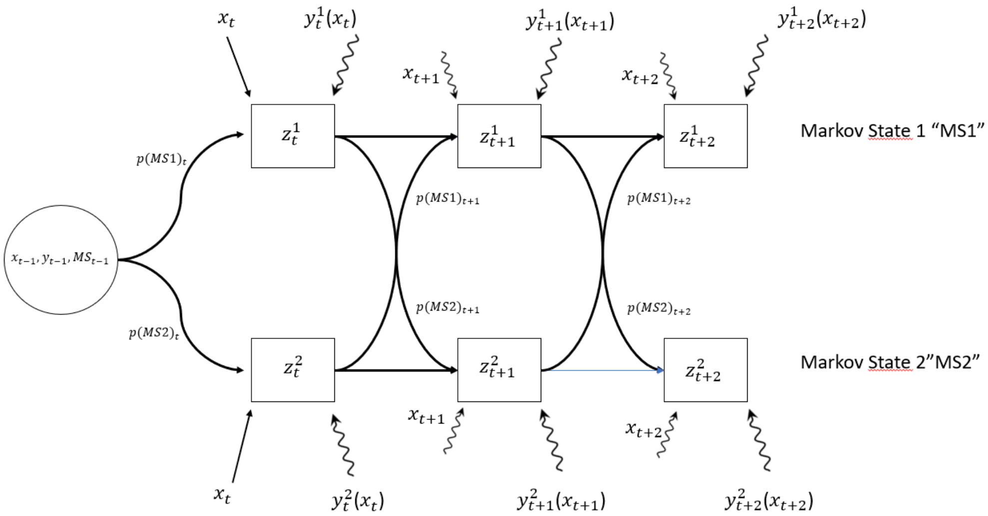


Figura 2: Adaptation of a Markovian policy graph of Dowson [2020], utilizing covariate data

DOI: 10.59254/sbpo-2023-174839
https://proceedings.science/p/174839?lang=pt-br

By incorporating this modification in the MC estimation, we can enhance the realism of the simulations for the inflow variables, as the state now captures historical dependencies. Hence, this paper makes the following key contributions:

- Illustrating the limitations of regular MC models in generating realistic simulations and their impact on decision-making in hydrothermal problems.
- Introducing historical dependency into state simulations for the uncertain inflow variable.
- Evaluating the effectiveness of our methodology using realistic data from the Brazilian power system.

2. Proposed methodology

We propose contextual stochastic dynamic programming model with Markov states

$$Q_t^j(z_{t-1}, y_t, x_t) := \min_{z_t \in Z_t(z_{t-1}, y_{t-1})} f_t(z_t, y_t) + \sum_{k=1}^{\kappa} E[Q_{t+1}^k(z_t, y_{t+1})] \cdot p_{jk}(x_t), \tag{1}$$

where z_t is the set of decision variables of the problem and Z_t represents the linear constraints on z_t at time t, y is the set of uncertain parameters at time t, x_t is the set of explanatory variables at time t, f_t is the cost function at time t. Let K represent the set of states in the Markov chain, while $\hat{P}$ symbolizes the estimated transition matrix associated with K, having dimensions of $|K| \times |K|$. Here, $\hat{p}_{jk}(x_t)$ signifies the probability of transitioning from state j to state k given the realized explanatory variable at time t. Finally, the second term of the objective function represents the future cost function given the decision at time $t-1$.

It is worth noting that SDDP frameworks can handle explanatory variant (in this case, time-varying) transition probabilities. However, traditional implementations of the problem do not incorporate this dynamic aspect in the noise term (x), potentially leading to suboptimal solutions.

To address this limitation and capture the explanatory variant state transition, we employ a time variant Markov Chain, where the state estimation is calculated as a function of exogenous features (time in this example). This approach allows us to account for the time-varying nature of the state transition and improve the overall modeling accuracy, leading to more optimal solutions. Within the framework of the proposed methodology, the exogenous variable x exclusively influences the transition probability.

We employ clustering techniques to accomplish this task. Within the historical dataset, each observation is assigned to a specific state, determined by both the realized value and the corresponding x occurrence. This classification process enables us to calculate the transition matrix through counting methods, leveraging the final categorizations.

2.1. Case Study

To test the proposed methodology, it was chosen the hydrothermal operation planning (without contingency). Therefore, we present the equivalent formulation of Equation 1 for the study case:

$$Q_t^j(v_{t-1}, a_t, x_t) := \min_{z_t, g_t, def_t, F_t} c' g_t + c'_{def} def_t + \sum_{k=1}^{\kappa} E[Q_{t+1}^k(v_t, a_{t+1})] \cdot p_{j,k}(x_t) \tag{2}$$

$$\text{s.t} \quad Ag_t + Bz_t + CF_t + def_t = d_t : \pi_t \tag{3}$$

$$v_t + Hz_t = v_{t-1} + a_t \tag{4}$$

$$(v_t, z_t, g_t^T, def_t, F_t) \in X_t, \tag{5}$$

DOI: 10.59254/sbpo-2023-174839
https://proceedings.science/p/174839?lang=pt-br

where v_t is the reservoir storage vector, a_t is the reservoir inflow at time t, z_t is a decision vector the contains both water discharged and spilled at time t for each reservoir and x_t is the exogenous variable at time t (in this case it represents the month of the year). It is important to observe that the transition probability $p_{j,k}$ is currently influenced by not only the state at time (t-1) denoted as "j"and the state at time (t) denoted as "k", but also the specific time frame of observation denoted as x_t (the month of the year). Within the context of this study, we estimate separate transition matrices for each month of the year, thereby revealing distinct state transition patterns inherent to each individual month, as anticipated in the Introduction.

Moreover, g^T represents thermal generation, def_t is a decision vector that represents the deficit at each location, c and c_{def} are a vector containing the costs of each thermal generator and the deficit cost respectively, F_t contains all other physical variables, such as energy passing through each line and nodal-phase angles and π_t is the energy spot price at time t (represented by the dual of Equation 3 constraint). A is a locational matrix that assign each thermal generation variable to its respective location. Analogously, B assigns each z variable to its locations converting the water discharge and spilled decision to energy units. C is an incidence matrix, i.e., it assigns the sense of from-to energy is transported. Therefore, Equation 3 represents the demand constraint (d_t is the demand in each location at time t). As H is also a location matrix that assigns the water discharge and spilled decisions to each location, Equation 4 represents the water balance equation. Finally, Equation 5 constraint contains other physical constraints such as bounds, ramp up and down or Kirchhoff's Voltage Law.

In this study, we focus on modeling the real aggregated configuration of the Brazilian interconnected power system. The system comprises four reservoirs representing different regions of the country, namely Southeast (SE), South (S), Northeast (NE), and North (N). Additionally, there are 10 transmission lines connecting these aggregated systems, along with a total of 95 thermal generators distributed across the subsystems. Detailed information, including generation costs, can be found in the attached data.

For this analysis, the deficit cost was fixed at R$2300/Mh, and the ramp-up and ramp-down limits were set to 20% of the maximum capacity of each thermal unit. The methodology employed in this study builds upon the work presented in Brigatto et al. [2017].

Furthermore, we consider a decision horizon of 84 months (equivalent to 7 years) for the study case, encompassing a substantial time frame to capture the medium- to long-term dynamics and associated decision-making processes.

2.2. Comparison methodology

To properly assess the dynamic representation of the monthly uncertain parameter, specifically the water inflow, we utilized the SDDP.jl framework Dowson and Kapelevich [2021]. Subsequently, we estimated several adapted Markov Chains (MC) with varying numbers of states, ranging from 1 to 3.

The training process for the adapted MC models focused on modeling a multivariate time-variant MC. In this approach, the classification of all regions is jointly estimated, and the inflow for each region and month is simulated based on random values derived from past state realizations.

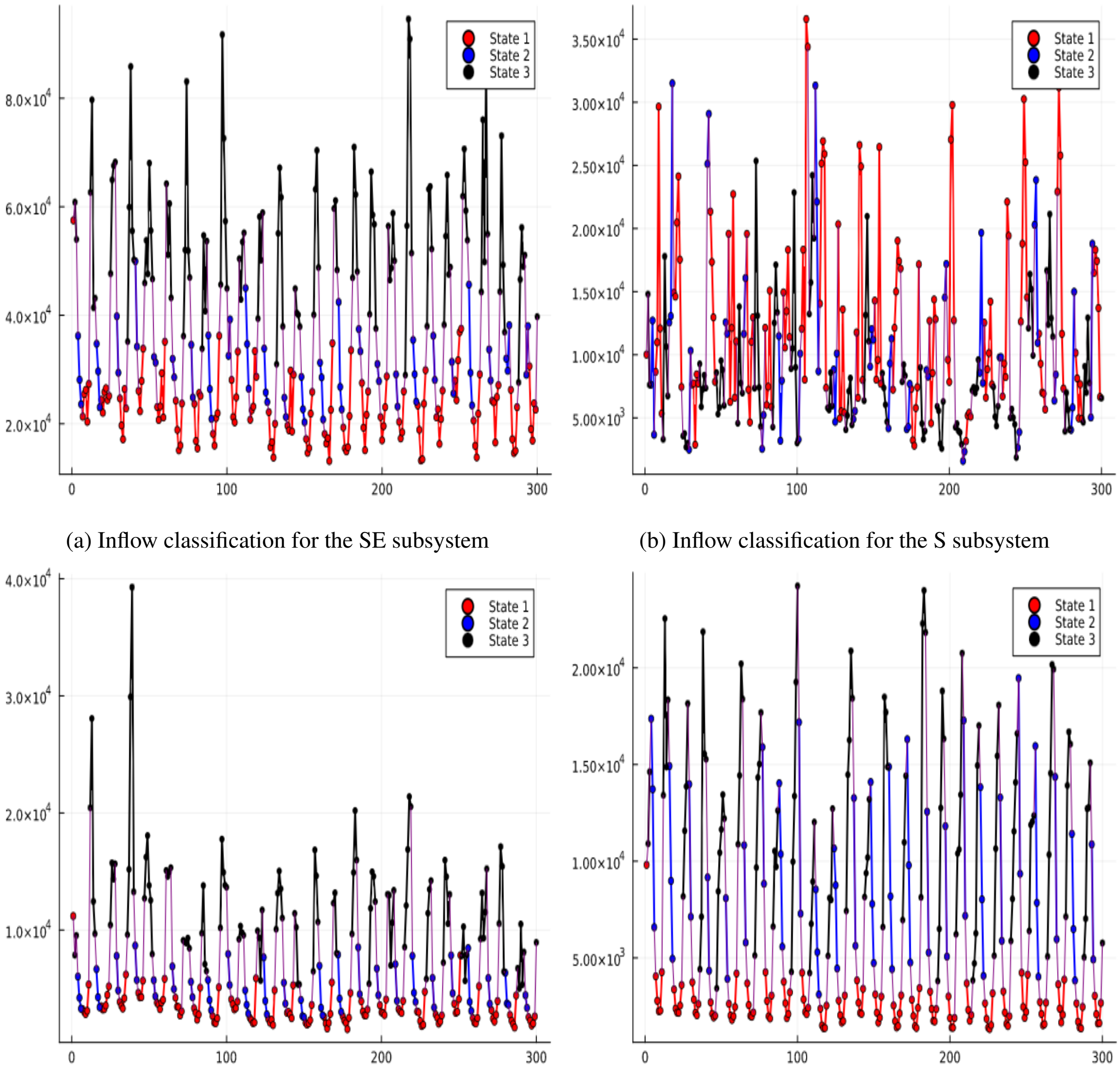
To compare the different models, we generated a set of scenarios following the time dynamics of an adapted MC with 3 states, as estimated in our proposed methodology. It is important to note that this experiment intentionally introduces a bias. The primary aim is to demonstrate the impact of not considering the correct dynamics in the estimation of the SDDP model, particularly regarding the uncertain parameter.

DOI: 10.59254/sbpo-2023-174839
https://proceedings.science/p/174839?lang=pt-br

By conducting this comparative analysis, we aim to shed light on the significance of accurately capturing the dynamic nature of the uncertain parameter during the estimation process of the SDDP model.

3. Results

In this section we present the results of the case study presented in the previous section. The first step of the methodology is to estimate the multivariate Markov states for the historical data of inflow of the four regions of the Brazilian case study. As an example we present the result of this classification for an adapted MC with 3 states:



(a) Inflow classification for the SE subsystem

(b) Inflow classification for the S subsystem

(c) Inflow classification for the NE subsystem

(d) Inflow classification for the N subsystem

Figura 3: Adapted MC classification results

It is noteworthy that the majority of regions display discernible patterns of wet and dry periods, except for the South region. However, due to the adoption of a multivariate estimation approach, the classification for the South region was influenced by the state transitions observed

DOI: 10.59254/sbpo-2023-174839
https://proceedings.science/p/174839?lang=pt-br

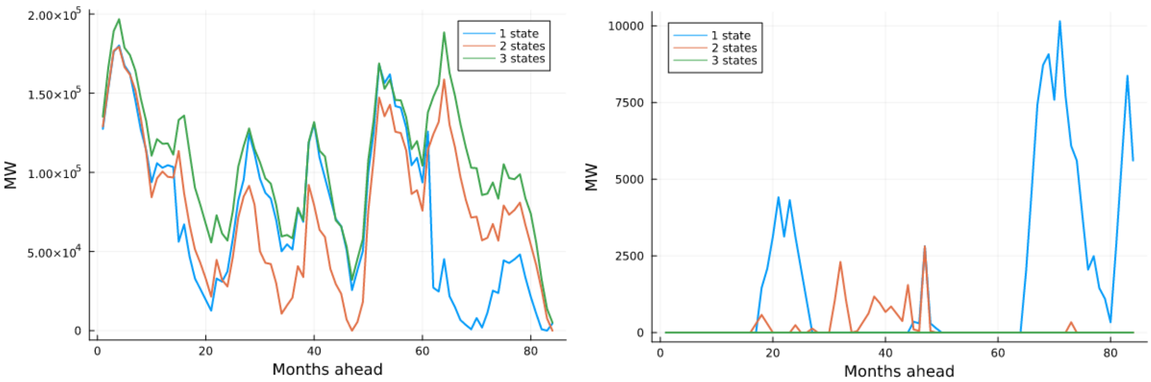
in the remaining regions. This means that the inflow classifications were not solely determined by its own specific characteristics but rather influenced by the dynamics of the other regions. This highlights the interconnectedness and interdependence of the different regions in terms of their hydrological patterns and reinforces the importance of considering the collective behavior of the system when analyzing and modeling hydrological processes.

The adapted Markov Chain (MC) implementation resulted in a simulation of a time-variant transition matrix. This enables the visualization of the classification of each historical observation into specific states, thereby capturing the temporal dynamics of the system.

3.1. Decision Policies

Finally, a comparison can be made among the different models using out-of-sample scenarios of inflow generated by employing an adapted MC with three states. It is expected that this model will yield the most favorable results, as it incorporates the true dynamics of the stochastic parameter.

To observe the impact of using the correct stochastic inflow parameter, we present in Figure 5 the variation of the reservoir volume (given in MW) for the south east region (the most relevant in the problem) for a given scenario:



(a) Variation of reservoir volume following the decision policies of different SDDP models

(b) Usage (MW) of thermal generators in the south east region

The models that consider only one or two states fail to accurately capture the dynamic nature of the uncertain parameter, resulting in inadequate water usage policies and leading to shortages.

Another policy to analyze is the utilization of thermal generators in the southeast region. Figure 4b illustrates that an incorrect prior assumption regarding the dynamics of the stochastic parameter leads to excessive and unnecessary reliance on thermal generators. This, in turn, has a significant impact on the cost function.

These two graphs highlight the issues that can arise in the hydrothermal problem. However, they also directly influence another crucial aspect of the problem, namely the energy spot price represented by the dual variable of Equation 3. Market agents aim to minimize spikes in energy prices over time as these spikes have a negative impact on the market. Unfortunately, as depicted in the figure below, when the stochastic process is inadequately represented in the SDDP formulation, these price spikes occur more frequently and with greater intensity:

3.2. Cost results

Finally, there were generated 100 scenarios of the inflow variable (following the 3 states dynamic) and all models were evaluated in relation of the objective function in this out-of-sample.

DOI: 10.59254/sbpo-2023-174839
https://proceedings.science/p/174839?lang=pt-br

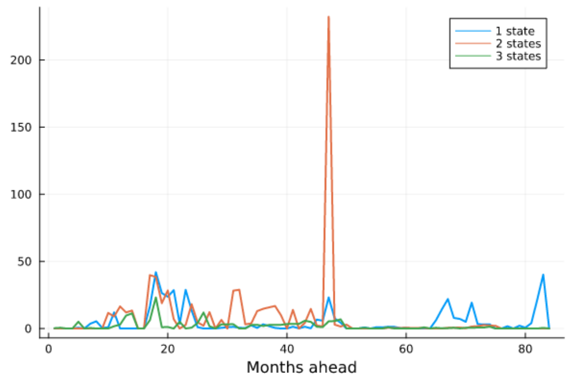


Figura 5: Energy spot price for different SDDP models in the south east region

The results are presented in the box plot graph at Figure 6:

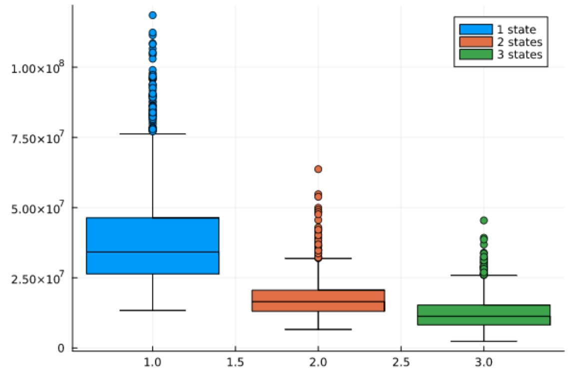


Figura 6: Model cost in out-of-sample scenarios evaluation

The boxplots clearly demonstrate that as the SDDP model incorporates a representation of the uncertain parameter that closely aligns with its real dynamic (as observed in the controlled study with 3 states), the out-of-sample results improve significantly. Conversely, when the stochastic terms are poorly represented, the worst-case scenarios become highly unfavorable. In the context of the hydrothermal example, these scenarios may result in energy shortages, leading to adverse economic conditions and social issues.

4. Conclusion

Stochastic dual dynamic programming (SDDP) problems that incorporate Markov chains, as implemented in SDDP.jl, allow for the consideration of state transition dynamics with the inclusion of explanatory variables such as time. However, traditional Markov Chain models (MC) are built under the assumption that all properties of the stochastic process depend solely on the current state. This limitation restricts their applicability in time series contexts, where the consideration of time in state transition probabilities is crucial.

In this paper, we presented a study focusing on a controlled case of the hydrothermal problem in Brazil. We constructed a time variant MC model to represent the out-of-sample scenarios and evaluated the behavior of the SDDP model under different assumptions about the stochastic process. Specifically, we tested models with 1, 2, and 3 Markov states against out-of-sample scenarios generated using a time variant MC with 3 states. As expected, the models that did not consider

DOI: 10.59254/sbpo-2023-174839
https://proceedings.science/p/174839?lang=pt-br

the correct stochastic dynamic exhibited poorer performance in terms of the objective function (overall cost). These models also led to inconsistent utilization of reservoir water, resulting in potential energy shortages and spot price spikes in certain scenarios.

As future work, it would be valuable to further test the proposed MC framework in SDDP models using backtests with real pseudo out-of-sample scenarios, comparing them against benchmark models. Additionally, developing a methodology to determine the optimal number of Markov states for stochastic processes would be important. Furthermore, extending the study to other cases where additional significant exogenous variables need to be considered in the MC estimation would be beneficial.

DOI: 10.59254/sbpo-2023-174839
https://proceedings.science/p/174839?lang=pt-br

Referências

Arthur Brigatto, Alexandre Street, and Davi M. Valladão. Assessing the cost of time-inconsistent operation policies in hydrothermal power systems. *IEEE Transactions on Power Systems*, 32(6): 4541–4550, 2017. doi: 10.1109/TPWRS.2017.2672204.

O. Dowson and L. Kapelevich. SDDP.jl: a Julia package for stochastic dual dynamic programming. *INFORMS Journal on Computing*, 33:27–33, 2021. doi: https://doi.org/10.1287/ijoc.2020.0987.

Oscar Dowson. The policy graph decomposition of multistage stochastic programming problems. *Networks*, 76, 02 2020. doi: 10.1002/net.21932.

Leonardo Martins, Aníbal Azevedo, and Secundino Soares. Nonlinear medium-term hydro-thermal scheduling with transmission constraints. *Power Systems, IEEE Transactions on*, 29:1623–1633, 07 2014. doi: 10.1109/TPWRS.2013.2296439.

Alexandre Moreira, Alexandre Street, and Jose Arroyo. An adjustable robust optimization approach for contingency-constrained transmission expansion planning. *Power Systems, IEEE Transactions on*, 30:2013–2022, 07 2015. doi: 10.1109/TPWRS.2014.2349031.

J. R. Norris. *Markov Chains*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 1997. doi: 10.1017/CBO9780511810633.

T Silva, D Valladão, and T Homem-de Melo. A data-driven approach for a class of stochastic dynamic optimization problems. 2021.

Geir Warland and Birger Mo. Stochastic optimization model for detailed long-term hydro thermal scheduling using scenario-tree simulation. *Energy Procedia*, 87:165–172, 01 2016. doi: 10.1016/j.egypro.2015.12.347.

DOI: 10.59254/sbpo-2023-174839
https://proceedings.science/p/174839?lang=pt-br