SOLAR ENERGY TIME SERIES ANALYSIS VIA MARKOV CHAINS

Gabriel Kelab Sigaud, Dep. Ind. Eng., PUC-Rio, + 55 (21) 98195-0088, gabrielsigaud@gmail.com
Marianne Bechara Elabras Da Motta Veiga, Dep. Ind. Eng., PUC-Rio, +55 (21) 98026-3068, mbemveiga@gmail.com
Fernando Luiz Cyrino Oliveira, Dep. Ind. Eng., PUC-Rio, cyrino@puc-rio.br
Gustavo de Andrade Melo, Dep. Ind. Eng., PUC-Rio, gustavo.melo.rio@gmail.com

Overview

Faced with a scenario of concern about climate change, countries are carrying out the energy transition, thus ceasing to use fossil energy sources and increasing the use of renewable sources (Malar, 2022). According to the International Renewable Energy Agency (2023), the planet had an increase in renewable energy capacity, in 2022, of 13% compared to the previous year. Brazil has been following this transformation in the global energy matrix. According to the 2023 National Energy Balance, 47.4% of the Brazilian internal energy supply, in 2022, came from renewable sources.

In this context, PV solar energy is a source that deserves to be highlighted. In 2022, it was responsible, in Brazil, for 3.6% of the internal energy supply. Furthermore, between 2021 and 2022, there was an 82.4% growth in installed capacity, being the fastest growing in the country (EPE, 2023). With the increase in its use in Brazil, its characteristics, such as intermittency and random fluctuations, will increasingly affect the country's energy generation. Given this scenario, the use of time series modeling and simulation methods to study this impact is important for planning plants and the Brazilian Electric System (SEB).

Aiming to contribute to this theme, the objective of this work is to analyze the characteristics of photovoltaic energy generation in different climatic seasons (summer, autumn, winter, and spring) in two regions of Brazil with different solar incidences. For this, the time series discretization approach was used for modeling via Markov Chains, a methodology widely used in the literature for analyzing electrical energy time series. It is worth highlighting that this work presents relevant differences compared to the literature. Firstly, until the authors know it, data that has not yet been studied is used. Furthermore, these data are from two plants located in regions with considerably different characteristics and were divided by the climatic seasons of the year, which allowed for both geographic and temporal comparisons.

Methods

The methodology used was based on the work of Melo (2022) and Ma et al. (2020). Melo (2022) sought to show the complementary spatial and temporal character between variable renewable energies through modeling and joint stochastic simulation of solar and wind energy. To do this, he used two methodologies and carried out three applications, using databases from plants located in the Northeast of Brazil. Both methodologies use Markov chain modeling, Monte Carlo simulation to obtain scenarios, and the k-means technique to perform data clustering.

Ma et al. (2020) proposed a methodology for aggregating data from photovoltaic solar energy time series through clustering via k-means, Markov Chains and Monte Carlo simulation. Based on the proposed k-means-MCMC methodology, initially, the energy generation data must be grouped following the ideal number of clusters and then the transition matrix must be assembled. Finally, from this matrix, energy scenarios are generated via simulation.

Thus, the methodology applied in this work can be divided into three main stages. The first is related to data pre-processing, covering data acquisition, analysis, and processing. Time series of daily photovoltaic energy generation from the Nova Olinda (Piauí) and Guaimbê (São Paulo) complexes were used, obtained from the National Electric System Operator (ONS, 2022) for a period of four years, from 06/21/2018 to 06/20/2022. The methodology is applied separately to the Nova Olinda Complex and the Guaimbê Complex. So that the time series could be analyzed by climate season, they were subdivided into four subsets: Summer, Autumn, Winter, and Spring.

In the second stage, data processing is carried out, involving modeling via Markov Chains, and obtaining results, such as stationary distribution, recurrence time and first passage time. In order to group observations with greater similarities, the subsets of the solar energy generation time series, divided by climate season, were discretized into Markovian states independently. The clustering method used was k-means (MacQueen, 1967), in which a k number of clusters are pre-specified, and k initial centroids (mean value of the clusters) are defined. The k-means method was used in conjunction with the elbow method to find the ideal number of clusters. Next, the daily state transition matrices were created. To analyze the properties of the matrices, three measures of interest were calculated: (i) stationary distribution, which represents the distribution of states in which the chain will stabilize; (ii) recurrence time, which translates the expected number of periods for a system in state i to return to that state again; (iii) and first passage time, which is the expected number of periods for a system in state i to pass through state j for the first time (Chung, 1960).

In the last stage, post-processing the data, the results obtained were analyzed for comparison purposes both between climatic seasons and between plants. At this stage, the main purposes were: to identify the most frequent states of each season; compare the recurrence times of the most extreme energy generation states; and compare the first passage times between the states with the highest and lowest energy generation in each climate season.

Results

After obtaining the model results, it was possible to analyze and interpret the generated values and better understand the behavior of the time series of daily photovoltaic energy generation, mainly based on stationary distributions and recurrence and first passage times.

Analyzing the time series of the Nova Olinda and Guaimbê complexes, the differences in the variability of the average photovoltaic energy generation throughout the year become evident, as Nova Olinda presents seasonality with a higher average generation in winter and lower in summer, while the averages of Guaimbê are closer in all-weather seasons. In the case of Nova Olinda, the reason for subdividing the series by climatic seasons to carry out the modeling is more evident. However, despite the Guaimbê Complex presenting more homogeneous monthly averages, the results for stationary distributions and recurrence and first passage times were significantly different in each season, as previously analyzed. Therefore, the subdivision by climate season proved to be relevant for both plants.

Another interesting fact is that the climatic seasons affect each region differently as well, with similarities between different seasons in the two regions. For example, the greatest concentration of stationary probabilities in upper central states is a case present in summer and spring in the Nova Olinda Complex, but also occurs in the autumn in the Guaimbê Complex. The autumn in Nova Olinda is similar to the winter in Guaimbê, because the states with lower generations have stationary probabilities significantly lower than the others and higher probabilities in the higher states. Meanwhile, winter in Nova Olinda and spring in Guaimbê are the seasons with the most balanced stationary probabilities between the states.

Furthermore, analyzing the first passage times, other similarities were found. The cases in which the first transition time from the highest generation state to the lowest generation state was longer than the opposite was in autumn in Nova Olinda and in the summer in Guaimbê. The opposite happened in summer and spring in Nova Olinda and in autumn and winter in Guaimbê. The winter of Nova Olinda and the spring of Guaimbê had the closest first passage times when comparing the most extreme states.

Finally, the sector stakeholders can use this type of analysis to support their different levels of planning, based on the calculation of probabilities of possible scenarios of low or high photovoltaic generation by region and climate season. Furthermore, the detailed study of the characteristics of renewable sources brings greater security in supplying energy demand in the country.

Conclusions

Brazil has been going through a process of changing its energy matrix, increasing the use of renewable energy. In this context, photovoltaic solar energy has stood out due to the significant growth in its participation in the country. Therefore, intermittency characteristics and random fluctuations of this source have a greater impact on the national energy supply scenario. The results obtained showed significant differences in solar energy generation between regions and between climatic seasons, which highlighted the relevance of the comparative study carried out. By analyzing and better understanding the specificities of each location and season, plants and SEB can plan more efficiently regarding energy generation, analyzing the probabilities of the occurrence of states with different generation values. It is suggested as a contribution to future studies: apply the methodology used in photovoltaic energy plants located in other regions of Brazil and the world, and apply this methodology to other renewable energy sources, seeking to better understand their behavior.

References

CHUNG, Kai Lai. **Markov Chains with Stationary Transition Probabilities**. Berlin, Heidelberg, Springer, 1960.
EPE. **Balanço Energético Nacional 2023**. Ministério de Minas e Energia, 2023.
IEA. **Renewable Energy Market Update** - June 2023, 2023. Available in: https://www.iea.org/reports/renewable-energy-market-update-june-2023. Access: Aug. 13th, 2023.
MA, M.; YE, L.; LI, J.; LI, P.; SONG, R.; ZHUANG, H. **Photovoltaic Time Series Aggregation Method Based on K-means and MCMC Algorithm**. Asia-Pacific Power and Energy Engineering Conference, v. 2020-September, n. 9220338, 2020.
MACQUEEN, J. **Some methods for classification and analysis of multivariate observations**. In Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Oakland, CA, USA, p. 14, 1967.
MALAR, J. P. **Conheça os tipos de energia renovável e quais são usados no Brasil**. CNN Brasil, São Paulo, Janeiro de 2022. Available in: https://www.cnnbrasil.com.br/business/conheca-os-tipos-de-energia-renovavel-e-quais-sao-usados-no-brasil/. Access: Oct. 16th, 2022.
MELO, Gustavo de Andrade; CYRINO Oliveira, Fernando Luiz; LOURO, Paula Medina Maçaira. **Simulação estocástica conjunta de energias renováveis. Rio de Janeiro**, 2022. 84p. Master's thesis – Department of Industrial Engineering, Pontifícia Universidade Católica do Rio de Janeiro.
ONS, 2022. Available in: <http://www.ons.org.br>. Access: Sep. 18th, 2022.