Solar energy time series analysis via markov chains

Marianne Bechara Elabras da Motta Veiga[1], Gabriel Kelab Sigaud [2], Fernando Luiz Cyrino Oliveira [3], Gustavo de Andrade Melo [4]
Recibido: 18/11/2024 y Aceptado: 04/2/2025

1.- mbemveiga@gmail.com
2.- gabrielsigaud@gmail.com
3.- Industrial Engineering Department Pontifical Catholic University of Rio de Janeiro
cyrino@puc-rio.br ORCID: 0000-0003-1870-9440
4.- Industrial Engineering Department Pontifical Catholic University of Rio de Janeiro
gustavo.melo.rio@gmail.com

Resumen

Brasil, ante un escenario global de preocupación por el cambio climático, viene incrementando el uso de energías renovables, especialmente la energía solar en los últimos años. Con el crecimiento de su participación, las características de la energía solar, como la intermitencia y las fluctuaciones aleatorias, vienen afectando la planificación de la operación del Sistema Eléctrico Brasileño (SBE). Tales factores pueden ser estudiados con modelos de series de tiempo, auxiliando la planificación de plantas generadoras y SBE. Con el fin de contribuir al análisis factorial, el objetivo de esta investigación es analizar las características de la generación de energía fotovoltaica en las estaciones meteorológicas del año en dos regiones de Brasil con diferentes incidencias solares. Para ello, se aplica una metodología basada en conceptos de Cadenas de Markov para dos series de tiempo estacionarias. El trabajo se destaca por la subdivisión de las series de tiempo entre las estaciones climáticas, por el uso de datos aún no estudiados y por la presentación de la metodología y resultados en detalle. El objetivo de la investigación fue alcanzado con éxito, evidenciando las diferencias entre los modelos de generación de energía solar entre las estaciones meteorológicas y las dos regiones estudiadas.

PALABRAS CLAVE: Fuentes de Energía Renovable, Fuentes de Energía Variables, Energía Solar, Estaciones Climáticas, Cadenas de Markov, K-means

Abstract

Brazil, given a global scenario of concern with climate change, has been increasing the use of renewable energy, especially solar energy in the last years. With the growth in its participation, the characteristics of solar energy, such as intermittence and random fluctuations, have been affecting the operation planning of the Brazilian Electricity System (BES). Such factors can be studied with time series modeling, helping the planning of power plants and BES. In order to contribute to the factor analysis, the objective of this research is to analyze the characteristics of photovoltaic energy generation in the meteorological seasons of the year in two regions of Brazil with different solar incidences. For this, a methodology based on Markov Chain concepts is applied for two stationary time series. The work stands out for the subdivision of the time series between the climatic seasons, for the use of data not yet studied and for the presentation of the methodology and results in detail. The objective of the research was successfully achieved, making evident the differences between the solar energy generation models between the meteorological seasons and the two regions studied.

KEYWORDS: Renewable Energy Sources, Variable Energy Sources, Solar Energy, Climatic Seasons, Markov Chains, K-means

1. INTRODUCTION

Faced with a scenario of concern about climate change, countries are carrying out the energy transition, thus moving away from using 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. Renewable energies are considered inexhaustible, as they can always be renewed by nature, and generate considerably lower environmental impacts than non-renewable energies (EPE, 2022).

Brazil has been following this transformation in the world's energy matrix. According to the 2023 National Energy Balance, 47.4% of Brazil's domestic energy supply in 2022 came from renewable sources. In 2013, this percentage was 40.6%, that is, in 9 years, there was an increase of approximately 17% (EPE, 2023).

In this context, solar energy is a source that deserves to be highlighted. In 2022, it accounted for 3.6% of the domestic energy supply in Brazil. In addition, between 2021 and 2022, it had 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 affect even more the country's energy generation. Solar energy is generated from solar radiation, captured by photovoltaic panels. In addition to being renewable, it has the advantages of being silent, requiring little maintenance and being able to be installed in a short time (Imhoff, 2007). With the increase in its use in Brazil, its characteristics, such as intermittency and random fluctuations, will increasingly affect the country's energy generation. Considering this scenario, the use of time series modeling and simulation methods to study this impact is important for the planning of the plants and the BES.

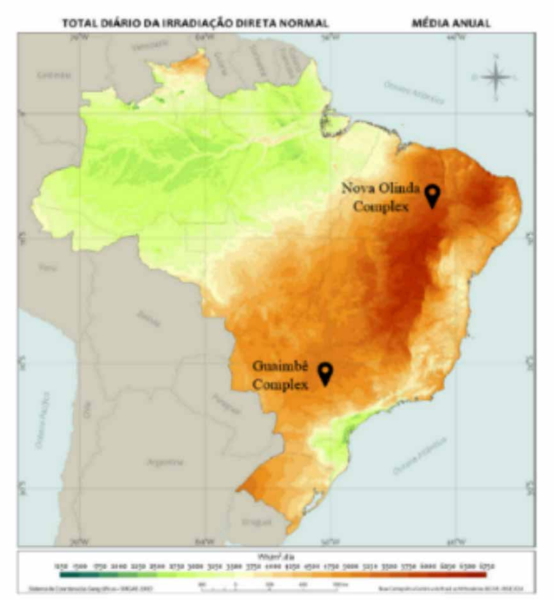
In order to contribute to this theme, the objective of this work is to analyze the characteristics of photovoltaic energy generation in different climatic seasons (summer, autumn, winter and spring) in two regions of Brazil with different solar incidences. For this, the time series discretization approach was used for Markov Chain modeling, a methodology already widely used in the literature for the analysis of electric energy time series. Furthermore, the subdivision by climatic season differs from other studies because it is based on a natural phenomenon, as opposed to monthly subdivisions, which are more frequently used, for example.

It is worth noting that this study presents relevant differentials in the literature. In the first place, to the authors' knowledge, data that have not yet been studied are used. Also, 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 both geographical and temporal comparisons.

The analysis presented in the study was carried out through two daily photovoltaic energy generation databases from ONS (National Electric System Operator): Nova Olinda Complex, located in Piauí (PI) and founded in 2017 (G1, 2017); and Guaimbê Complex, located in the state of São Paulo (SP) and inaugurated in 2019 (G1, 2019). According to Gadelha de Lima (2020), the state of Piauí has different meteorological characteristics depending on the quarter of the year, which could justify a division into four seasons.

Figure 1 shows the location of the two plants on the brazilian solarimetric map. This map is an adaptation of the one presented in the Brazilian Atlas of Solar Energy (Pereira et al., 2017) and shows the annual average of the total daily normal direct irradiation over Brazil. It is possible to perceive the difference in the averages of direct irradiation between the two locations of the plants, which is greater in the Nova Olinda Complex (Ribeira do Piauí – PI) in relation to the Guaimbê Complex (Guaimbê – SP).

Figure 1 - Brazilian Solarimetric Map - Average annual normal direct irradiation.



Source: Adapted from Pereira et al. (2017).

The applied methodology is exploratory and can be divided into three main phases. The first relates to data pre-processing, including data collection, analysis, and treatment. In the second phase, data processing is performed, involving modeling via Markov Chains and obtaining results such as stationary distribution, recurrence time, and first passage time. In the last phase, data post-processing, the results obtained were analyzed for comparison between the climatic seasons and between the plants.

2. THEORETICAL FRAMEWORK

In the literature, there are several renewable energy modeling studies that apply the concept of Markov Chains in their methodologies. Sigauke and Chikobvu (2017) performed an analysis of daily peaks of electricity demand through Markov Chains, seeking to find the stationary distribution (distribution of states in which the chain will stabilize). To do this, the authors used demand data from South Africa from 2000 to 2011. Models with two states were considered, being the positive or negative variations between the days, and with three states, where the difference between small and large positive variations was considered.

Maçaira et al. (2019), faced with a scenario of increased wind energy use in Brazil, showed that the dispatch model used in the period of their research did not consider the stochastic behavior of this energy source. The model, which sought to optimize long-term energy planning, only evaluated the future aspects of water and thermal sources. In view of this, the work proposed the wind-hydrothermal dispatch model, which incorporated wind power generation using the MCMC (Markov-Chain Monte Carlo) method to simulate energy scenarios.

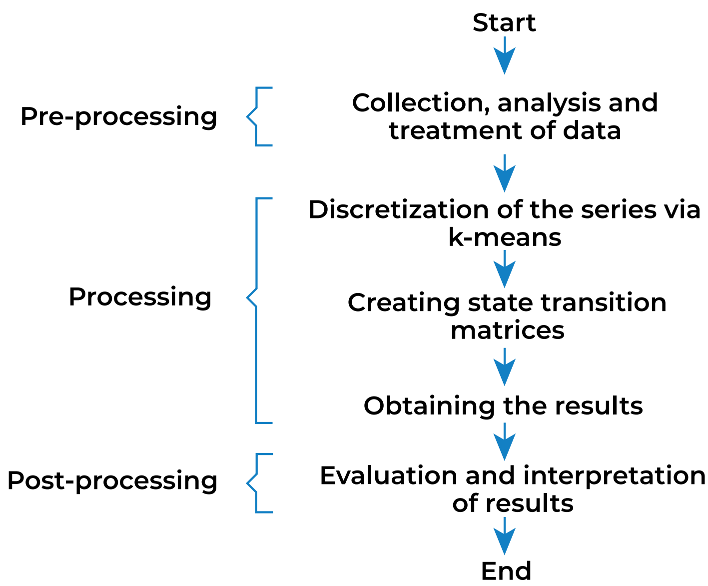
Ma et al. (2020) proposed a methodology for aggregating solar photovoltaic time series data through clustering via k-means, Markov Chains, and Monte Carlo simulation. For the authors, Markovian processes efficiently represent the transitions of photovoltaic power generation time series. Based on the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

Melo (2022) sought to show the spatial and temporal complementarity between variable renewable energies through the joint stochastic modeling and simulation of solar and wind energy. To this end, it used two methodologies and performs three applications, through databases of mills 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.

3. METHODOLOGY

A methodology based on Markov Chains was applied to modeling the time series of photovoltaic solar power generation. Figure 2 shows the flowchart with the main stages of the methodology, divided into the data's pre-processing, processing, and post-processing phases.

Figure 2 - Main steps of the methodology.



3.1. Pre-processing

The pre-processing phase consists of obtaining, analyzing, and treating data. The data of the time series of daily photovoltaic energy generation of the Nova Olinda (Piauí) and Guaimbê (São Paulo) complexes were obtained from the National Electric System Operator (ONS, 2022) for a period of four years, from 06/21/2018 to 06/20/2022, with a total of 1,461 observations for each complex. The only two variables used were date and energy generation. According to Ma et al. (2020), due to the characteristics of photovoltaic power generation data, the optimal time scale to fragment scenarios would be daily. The methodology is applied first to the Nova Olinda Complex and then to the Guaimbê Complex, so the two series are worked separately in the modeling.

A preliminary analysis of the data obtained from energy generation during the period was performed. First, to test the stationarity of the time series over the four years, Augmented Dickey-Fuller (ADF) unit root tests were carried out. The null hypothesis of the ADF test is that there are unit roots in the time series and, therefore, it would not be stationary (Dickey, D.; Fuller, 1979). The stationarity test is essential for the application of the Markov Chain concepts, because a non-stationary series depends on time, and in Markovian processes, the probabilities of transition to the next state depend only on the current state (Norris, 1998). Furthermore, non-stationarity would mean a change in the installed capacity of the plants.

To complete the pre-processing phase, a treatment of the databases is carried out so that the time series can be modeled as Markov Chains. First, the null or missing values were replaced by the averages of the month in the corresponding year, as it is an adequate estimate for the value of generation in the period, given seasonality. Then, so that the time series could be analyzed by climatic season, they were subdivided into four subsets: Summer, Autumn, Winter, and Spring.

3.2. Processing

3.2.1 Series discretization via k-means

In order to group the observations with greater similarities, the subsets of the solar energy generation time series, divided by climatic season, were discretized into markovian states independently. The clustering method used was k-means (MacQueen, 1967), as it is easily programmable and computationally economical. In the k-means method, a number k of clusters is pre-specified, and initial k centroids (average value of clusters) are defined based on a random variable. Then, the following steps are performed: Observations are assigned to the nearest centroid cluster by calculating the distance from each observation to each centroid; New k centroids are calculated from the average of intra-cluster observations; Iterations of steps 1 and 2 are performed until the centroid values do not change further The method can be summarized by the objective function (1).

$$J = \sum_{j=1}^{k} \sum_{i=1}^{n} \| x_i^{(j)} - c_j \|^2 \qquad (1)$$

However, to apply the k-means method, it is necessary to pre-define the number k of clusters. According to Fritz et al. (2020), choosing the wrong values for k can lead to poor results, and to choose the ideal number of clusters, it is common to use the elbow method, first discussed by Thorndike (1953). As the number of clusters increases, the sum of the squared error of the distance between the observations and the centroids tends to decrease (Thorndike, 1953). Hence, the elbow method helps to limit the choice of very high values for k, in which there are no relevant benefits with the addition of a new cluster. The elbow method can be used in conjunction with the k-means method to find the optimal number of clusters (Fritz et al., 2020).

To apply the elbow method using k-means, it is first necessary to perform the k-means steps for each k-value up to a chosen maximum number. Then, the sum of the intra-cluster squared error, or Within-Cluster-Sum of Squared Errors (WSS), is calculated for each clustering obtained by the k-means result. The WSS consists of the sum of the square of the euclidean distances from each observation to the centroid of the cluster to which it belongs.

Consequently, a graph can be created that presents the WSS for each value of k. So it is possible to observe the point k at which the curve presents a “fold”, like an elbow, and it can be inferred that the difference between the WSS of k and k+1 would not provide substantial gains to clustering.

3.2.2 Creating State Transition Matrices

The next step is to create the daily transition matrices of states, P. Transition matrices are composed of the transition probabilities pi,j between a state i and a state j between a period n and n+1 (Chung, 1960).

The transition probabilities and transition matrices are represented by (2) and (3), respectively.

$$P\{x_n(\omega) = j \mid x_{n-1}(\omega) = i\} = p_{i,\,j} \quad (2)$$

$$P = \left(p_{1,1}\, p_{1,2}\ \cdots\ p_{1,k-1}\, p_{1,k}\, p_{2,1}\, p_{2,2}\ \cdots\ p_{2,k-1}\, p_{2,k}\ \vdots\ \ \vdots\ \ddots\ \vdots\ \ \vdots\ p_{k-1,1}\, p_{k-1,2}\ \cdots\ p_{k-1,k-1}\, p_{k-1,k}\, p_{k,1}\, p_{k,2}\ \cdots\ p_{k,k-1}\, p_{k,k}\right) \quad (3)$$

In this step, based on Melo (2022) and Ma et al. (2020), the transition probabilities are calculated by the ratio between the number of occurrences of transitions from state i to state j and the total occurrences of transitions from state I, as represented by (4).

$$p_{i,j} = \frac{n_{i,j}}{\sum_{a=1}^{k} n_{i,a}} \qquad (4)$$

3.2.3 Obtaining the results

To analyze the properties of the transition matrices, three measures of interest were calculated: Stationary distribution (π) - represents the distribution of states in which the chain will stabilize, satisfying the equations (5) and (6); Recurrence time (mii) - the expected number of periods for a system in state i to return to that state again, as in the equation (7); First passage time (mij) - The number of periods expected for a system in state i to first passage through state j, as in the equation (8) (Chung, 1960).

$$\pi_j = \sum_{i=0}^{M} \pi_i p_{ij} \ \forall\, j = 0, 1, \ldots, M \tag{5}$$

$$\sum_{i=0}^{M} \pi_j = 1 \tag{6}$$

$$m_{ii} = \frac{1}{\pi_i} \ \forall\, i = 0, 1, \ldots, M \tag{7}$$

$$m_{ij} = 1 + \sum_{k \neq j} p_{ik} m_{kj} \tag{8}$$

Interpreting the above concepts, the measures presented are important to assist in analyzing the behavior of the Markov Chains model when the process stabilizes. With a stationary distribution, it is possible to identify the most frequent states of the system, where the process is most likely to be in the future. The recurrence time allows us to understand, for example, the average time to return to a state of maximum or minimum energy generation, while the first passage time would indicate the average transition time between these two states.

3.3. Post-processing

Finally, in the post-processing phase, the analysis and evaluation of the results obtained in the previous phase were carried out, with the objective of analyzing the characteristics of the generation of the two plants in the four climatic seasons and in regions of Brazil with different solar incidences. In this phase, the main purposes were: to identify the most frequent states of each season; to compare the recurrence times of the most extreme power generation states; and to compare the first passage times between the states of highest and lowest power generation of each climatic season.

4. DISCUSSION AND PRESENTATION OF RESULTS

In this chapter, the results of the methodology's application are presented for the two plants individually, starting with the Nova Olinda Complex (PI) and, later, addressing the Guaimbê Complex (SP). Finally, the results of the two plants are compared. All the computational steps in this chapter were performed in the R® programming language (R Development Core Team, 2009).

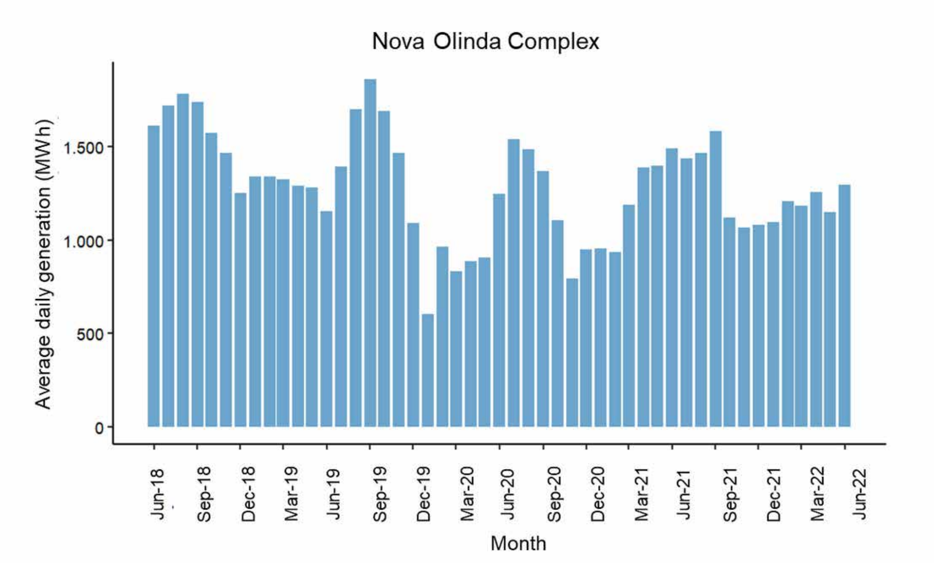
4.1. Nova Olinda Complex (Piauí)

4.1.1 Pre-processing

4.1.1.1 Collection, analysis and treatment of data

When testing the stationarity of the time series of the Nova Olinda Complex in the analyzed period, the result obtained was a p-value lower than 0.01, i.e., the null hypothesis that the time series would not be stationary is rejected. Thus, it is concluded that the time series is stationary and, therefore, the installed capacity is constant, which is fundamental for the Markov Chain modeling performed in this work. The stationarity of the time series in the period can be seen in Figure 3, which represents the average daily generation per month. In addition, the series presents considerable volatility and annual seasonality, with higher energy generation in the months of July, August, and September and lower generation in the months of December, January, February, and March, while the other months assume intermediate energy generation values. It is possible to notice greater similarities in the data in the months of the same climatic season. Due to this observation, an opportunity is identified to model the time series by subdividing it into four subsets, one for each climatic season, for a better representation of the data in each period.

Figure 3 - Average daily generation - Nova Olinda Complex.



Source: Based on data from ONS (2022).

Table 1 shows that winter has the highest daily average of energy generation in the Nova Olinda Complex in Piauí, with 1,572.87 MWh/day. Meanwhile, the summer has a daily average of 32% lower than that of winter, with 1,066.39 MWh/day, probably due to a higher number of cloudy days in this period of the year, which reduces the average daily solar radiation in the region of the plant. Furthermore, it is also possible to note that winter has the lowest standard deviation, while spring, the second season with the highest average energy generation, has the highest standard deviation, therefore, a greater dispersion of data.

Table 1: Measures of daily energy generation - Nova Olinda Complex.

	General	Summer	Autumn	Winter	Spring
Average (MWh)	1,282.78	1,066.39	1,208.70	1,572.87	1,271.72
Median (MWh)	1,323.02	1,067.53	1,246.82	1,605.86	1,285.52
Standard deviation (MWh)	394.55	357.41	325.43	269.34	427.05

4.1.2 Processing

4.1.2.1 Discretization of the series via k-means

The discretization of the photovoltaic time series was performed individually for each climatic season, so that the number of clusters and the values for the centroids were better suited specifically to each of the subsets.

The first step in the execution was to create a function that would calculate the k-means for values of k from 1 to 20. The maximum number of 20 clusters was chosen because it was verified that this is a sufficient amount to represent the data. The second step was to create a function that returned WSS for each of the 20 clusters. The third step was to apply the previously created functions to each of the subsets created. The fourth step was the application of the elbow method. With the results of the WSS calculation, a list was created that contained the ratio between the WSS of a number k and k+1 of clusters for k=1 to k=19. Then, for each of the subsets, the k-value of clusters in which the calculated ratio was greater than 0.90 was identified, i.e., the number of clusters necessary for the reduction of the sum of the intra-cluster squared error, when including a new cluster, to be less than 10%, which would not justify the addition.

Thus, the k-means result for each of the subsets found the ideal number of clusters (Table 2) and centroid values (Table 3).

Table 2: Ideal number of clusters - Nova Olinda Complex.

	Summer	Autumn	Winter	Spring
Number of clusters	11	12	8	8

Table 3: Centroids of the states - Nova Olinda Complex.

Centroids (MWh)		States											
		1	2	3	4	5	6	7	8	9	10	11	12
Seasons	**Summer**	289	545	740	874	1,018	1,137	1,239	1,350	1,453	1,557	1,819	-
	Autumn	307	624	799	916	990	1,075	1,182	1,279	1,372	1,469	1,582	1,709
	Winter	875	1,209	1,407	1,508	1,604	1,720	1,825	1,945	-	-	-	-
	Spring	309	741	980	1,187	1,407	1,634	1,814	1,969	-	-	-	-

4.1.2.2 Creating State Transition Matrices

In this step, the transition matrices of the Nova Olinda Complex (Figure 4) were constructed from the transition frequencies between the states for each subset.

Figure 4 - Transition matrices - Nova Olinda Complex.

Summer

	1	2	3	4	5	6	7	8	9	10	11
1	0,29	0,43	0,07	0,14	0,07	0,00	0,00	0,00	0,00	0,00	0,00
2	0,13	0,17	0,17	0,17	0,13	0,10	0,10	0,00	0,03	0,00	0,00
3	0,10	0,12	0,24	0,15	0,15	0,10	0,05	0,07	0,02	0,00	0,00
4	0,02	0,13	0,27	0,13	0,16	0,13	0,00	0,11	0,04	0,00	0,00
5	0,00	0,06	0,14	0,22	0,16	0,12	0,10	0,06	0,06	0,04	0,02
6	0,00	0,03	0,05	0,16	0,26	0,18	0,16	0,03	0,08	0,05	0,00
7	0,00	0,06	0,00	0,15	0,09	0,21	0,06	0,18	0,12	0,09	0,03
8	0,03	0,00	0,08	0,06	0,19	0,08	0,17	0,06	0,14	0,14	0,06
9	0,00	0,03	0,07	0,00	0,07	0,03	0,20	0,30	0,13	0,13	0,03
10	0,00	0,04	0,00	0,00	0,00	0,04	0,08	0,29	0,21	0,29	0,04
11	0,00	0,00	0,00	0,08	0,08	0,00	0,00	0,08	0,08	0,17	0,50

Autumn

	1	2	3	4	5	6	7	8	9	10	11	12
1	0,00	0,00	0,30	0,10	0,20	0,00	0,10	0,00	0,30	0,00	0,00	0,00
2	0,06	0,18	0,24	0,18	0,00	0,06	0,06	0,06	0,00	0,06	0,12	0,00
3	0,04	0,15	0,19	0,15	0,11	0,04	0,11	0,07	0,11	0,04	0,00	0,00
4	0,09	0,04	0,09	0,13	0,13	0,22	0,04	0,13	0,09	0,00	0,00	0,04
5	0,04	0,04	0,12	0,08	0,23	0,12	0,12	0,12	0,04	0,08	0,04	0,00
6	0,00	0,10	0,03	0,08	0,10	0,21	0,21	0,08	0,00	0,05	0,10	0,05
7	0,03	0,03	0,06	0,06	0,08	0,17	0,11	0,14	0,11	0,06	0,17	0,00
8	0,02	0,00	0,00	0,02	0,02	0,15	0,20	0,22	0,17	0,10	0,05	0,05
9	0,04	0,02	0,06	0,02	0,04	0,04	0,06	0,21	0,23	0,06	0,19	0,00
10	0,00	0,03	0,06	0,00	0,03	0,06	0,11	0,06	0,19	0,31	0,14	0,03
11	0,00	0,03	0,03	0,05	0,00	0,13	0,03	0,08	0,13	0,18	0,15	0,21
12	0,04	0,00	0,00	0,04	0,00	0,00	0,00	0,00	0,12	0,08	0,27	0,46

Winter

	1	2	3	4	5	6	7	8
1	0,26	0,37	0,00	0,00	0,21	0,16	0,00	0,00
2	0,09	0,25	0,14	0,25	0,14	0,07	0,05	0,02
3	0,03	0,26	0,21	0,18	0,16	0,11	0,05	0,00
4	0,06	0,10	0,25	0,17	0,19	0,15	0,04	0,04
5	0,04	0,04	0,07	0,18	0,38	0,19	0,07	0,01
6	0,03	0,07	0,07	0,08	0,22	0,32	0,17	0,03
7	0,02	0,03	0,05	0,03	0,05	0,16	0,56	0,10
8	0,00	0,00	0,00	0,06	0,03	0,06	0,22	0,63

Spring

	1	2	3	4	5	6	7	8
1	0,21	0,32	0,16	0,21	0,11	0,00	0,00	0,00
2	0,17	0,41	0,22	0,15	0,05	0,00	0,00	0,00
3	0,07	0,14	0,48	0,24	0,02	0,03	0,02	0,00
4	0,00	0,13	0,24	0,34	0,23	0,06	0,00	0,00
5	0,05	0,00	0,06	0,15	0,45	0,20	0,08	0,02
6	0,02	0,02	0,02	0,09	0,28	0,28	0,22	0,07
7	0,00	0,03	0,00	0,05	0,05	0,43	0,30	0,14
8	0,00	0,00	0,00	0,00	0,00	0,15	0,35	0,50

4.1.2.3 Obtaining the results

All the transition matrices created were classified as irreducible and ergodic, important properties for the Markov Chain to have a stationary distribution. Then, stationary distributions (Table 4), recurrence times (Table 5), and first passage times (Figure 5) were calculated.

Table 4: Stationary distribution - Nova Olinda Complex.

a) Summer	States										
	1	2	3	4	5	6	7	8	9	10	11
Centroid (MWh)	289	545	740	874	1,018	1,137	1,239	1,350	1,453	1,557	1,819
Stationary distribution	0.0401	0.0851	0.1191	0.1246	0.1390	0.1073	0.0910	0.1049	0.0827	0.0719	0.0343

b) Autumn	States											
	1	2	3	4	5	6	7	8	9	10	11	12
Centroid (MWh)	307	624	799	916	990	1,075	1,182	1,279	1,372	1,469	1,582	1,709
Stationary distribution	0.0269	0.0456	0.0693	0.0622	0.0665	0.1068	0.0997	0.1110	0.1253	0.0962	0.1162	0.0745

c) Winter	States							
	1	2	3	4	5	6	7	8
Centroid (MWh)	875	1,209	1,407	1,508	1,604	1,720	1,825	1,945
Stationary distribution	0.0503	0.1108	0.1007	0.1246	0.1829	0.1668	0.1757	0.0882

(d) Spring	States							
	1	2	3	4	5	6	7	8
Centroid (MWh)	309	741	980	1,187	1,407	1,634	1,814	1,969
Stationary distribution	0.0553	0.1199	0.1777	0.1784	0.1798	0.1419	0.0949	0.0522

Table 5: Recurrence time - Nova Olinda Complex.

a) Summer	States										
	1	2	3	4	5	6	7	8	9	10	11
Centroid (MWh)	289	545	740	874	1,018	1,137	1,239	1,350	1,453	1,557	1,819
Recurrence time (days)	25	12	8	8	7	9	11	10	12	14	29

b) Autumn	States											
	1	2	3	4	5	6	7	8	9	10	11	12
Centroid (MWh)	307	624	799	916	990	1,075	1,182	1,279	1,372	1,469	1,582	1,709
Recurrence time (days)	37	22	14	16	15	9	10	9	8	10	9	13

c) Winter	States							
	1	2	3	4	5	6	7	8
Centroid (MWh)	875	1,209	1,407	1,508	1,604	1,720	1,825	1,945
Recurrence time (days)	20	9	10	8	5	6	6	11

(d) Spring	States							
	1	2	3	4	5	6	7	8
Centroid (MWh)	309	741	980	1,187	1,407	1,634	1,814	1,969
Recurrence time (days)	18	8	6	6	6	7	11	19

Figure 5 - First passage time - Nova Olinda Complex.

Summer

	1	2	3	4	5	6	7	8	9	10	11
1	0	7	9	7	8	12	13	13	17	24	62
2	32	0	9	8	7	10	12	12	15	23	61
3	33	13	0	8	7	10	12	11	15	23	60
4	36	13	8	0	7	10	12	11	15	22	60
5	38	15	10	8	0	10	11	10	14	21	58
6	39	15	11	8	6	0	10	10	14	21	59
7	39	15	12	9	8	9	0	9	13	19	57
8	38	16	11	9	7	11	10	0	12	18	55
9	39	16	12	10	8	11	9	7	0	17	55
10	40	17	13	11	9	12	10	6	11	0	54
11	40	18	13	10	9	13	13	9	12	15	0

Autumn

	1	2	3	4	5	6	7	8	9	10	11	12
1	0	26	13	16	15	12	10	11	8	15	11	27
2	35	0	14	15	19	11	11	11	11	15	11	27
3	36	23	0	15	17	11	10	11	10	15	12	27
4	34	26	17	0	17	9	11	10	10	15	11	26
5	36	26	16	17	0	10	10	10	10	14	11	27
6	38	25	18	17	18	0	9	11	11	14	10	25
7	37	27	18	18	18	10	0	10	9	14	9	26
8	37	28	19	19	20	10	9	0	9	13	10	25
9	37	27	18	19	19	11	11	9	0	14	9	26
10	39	27	18	19	20	11	10	11	8	0	9	25
11	38	27	19	19	21	11	12	11	9	12	0	20
12	37	29	20	19	21	13	13	12	9	13	7	0

Winter

	1	2	3	4	5	6	7	8
1	0	7	13	9	6	7	14	32
2	25	0	11	7	7	8	13	31
3	26	9	0	8	7	8	13	31
4	26	11	10	0	7	7	13	30
5	26	12	12	8	0	7	12	31
6	27	12	13	9	7	0	11	30
7	29	14	14	11	9	7	0	26
8	30	15	15	11	10	9	8	0

Spring

	1	2	3	4	5	6	7	8
1	0	10	8	6	10	15	24	51
2	18	0	7	6	11	16	25	52
3	22	12	0	6	11	15	24	51
4	25	14	9	0	8	13	22	49
5	26	18	13	8	0	9	18	45
6	27	19	14	9	7	0	14	40
7	29	20	15	10	9	5	0	35
8	30	22	17	12	10	5	6	0

4.1.3 Post-processing

4.1.3.1 Evaluation and interpretation of results

After obtaining the model's results, it becomes possible to analyze and interpret the generated values and better understand the behavior of the daily photovoltaic power generation time series, mainly from the stationary distributions, recurrence times and first passage times.

From the stationary distribution, in Table 4, it is observed that the system presents higher probabilities for the states of intermediate generation values in the summer and spring seasons. Also, the probabilities decay little by little and in a similar way for the lower and higher extreme states. Another way to analyze it is by the time of recurrence of the states in Table 5. In both seasons, the recurrence times of the extreme states are significantly higher compared to the central states and very close to each other. Analyzing the extreme states, in the summer, states 1 (289 MWh) and 11 (1,819 MWh) have a recurrence of 25 and 29 days, respectively, while states 1 (309 MWh) and 8 (1,969 MWh) in spring have a recurrence of 18 and 19 days, respectively. Consequently, the tendency is for the system to remain in medium-generation states and the extremes to be rarer, with lower expectations of low or high generation in a day, especially in the summer, whose recurrence times of extreme states are even longer.

Autumn, on the other hand, has a higher probability of being in the central and upper states, with lower probabilities in states of lower energy generation, comparatively. At this season, the two states with the highest power generation, states 11 (1,582 MWh) and 12 (1,709 MWh), have recurrence times of 9 and 13 days, respectively. Meanwhile, the recurrence times of the two lowest-generation states, states 1 (307 MWh) and 2 (624 MWh), are 37 and 22 days, respectively. In addition, the recurrence time of state 1 of autumn is the longest among all states of all seasons, i.e., autumn presents the longest average period for the system to return to low levels of power generation. Hence, it appears that the system has a tendency towards higher states with a lower risk of low generation. Furthermore, by investigating the first passage times of summer and spring, it is possible to analyze that the time to leave the state of lowest energy generation and reach the state of highest generation for the first time is longer than the reverse. For example, the first passage time from state 1 (309 MWh) to state 8 (1,969 MWh) in the spring is 51 days, while the time from state 8 to state 1 is 30 days, approximately 40% shorter. Therefore, although the probabilities of the system being at each extreme are close, once the system is in a low-generation state, it will take longer to reach the higher-generation states in both seasons.

Meanwhile, when looking at winter, the first passage times between the two most extreme states, lower and upper, are close — 32 days from state 1 (875 MWh) to state 8 (1,945 MWh) and 30 days from state 8 to state 1— although their recurrence times are quite different (20 days for state 1 and 11 days for state 8). It is interesting to note that the first passage times between states with more distant generation levels may be shorter than among others with closer generations. For example, the first passage time from state 2 (1,209 MWh) to state 3 (1,407 MWh) is 11 days, while the time from state 2 to state 6 (1,720 MWh) is 8 days.

4.2. Guaimbê Complex (São Paulo)

4.2.1 Pre-processing

4.2.1.1 Collection, analysis and treatment of data

By testing the stationarity of the time series of the Guaimbê Complex in the analyzed period, it was concluded that the time series is stationary and, therefore, the installed capacity is constant. The stationarity of the time series in the period can be seen in Figure 6, which represents the average daily generation per month. In addition, the series has considerably lower volatility than that of the Nova Olinda Complex, with low variations in energy generation over the months.

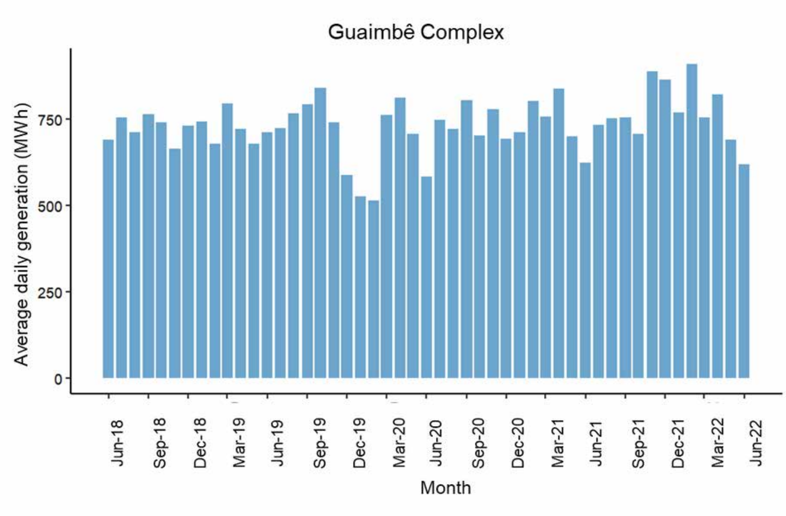


Table 6 shows that spring has the highest daily average of energy generation in the Guaimbê Complex, with 752.92 MWh/day. Meanwhile, the summer has a daily average of 5% lower than that of spring, with 717.33 MWh/day, being the lowest average for the plant. Consequently, the low variability of energy generation between the seasons of the year is evident, with all values being considerably close to the general average. Also, the standard deviation of the seasons also assumes close values.

Table 6: Measurements of daily energy generation - Guaimbê Complex.

	General	Summer	Autumn	Winter	Spring
Average (MWh)	733.22	717.33	727.17	735.38	752.92
Median (MWh)	774.37	733.33	774.09	777.22	796.82
Standard deviation (MWh)	196.43	201.25	188.33	181.12	213.40

4.2.2 Processing

4.2.2.1 Discretization of the series via k-means

The discretization of the photovoltaic energy time series for the Guaimbê Complex was performed with the same method as the Nova Olinda Complex, but in a totally independent way, using the k-means technique and the elbow method. The ideal number of clusters and centroid values are shown in Tables 7 and 8, respectively.

Table 7: Ideal number of clusters - Guaimbê Complex.

	Summer	Autumn	Winter	Spring
Number of clusters	8	9	11	8

Table 8: Centroids of the states - Guaimbê Complex.

Centroids (MWh)		States										
		1	2	3	4	5	6	7	8	9	10	11
Se as o ns	Summer	234	386	499	610	715	810	909	1,018	-	-	-
	Autumn	244	459	619	699	760	810	867	924	992	-	-
	Winter	182	324	448	545	615	679	741	791	840	900	983
	Spring	269	457	593	688	780	863	949	1,033	-	-	-

4.2.2.2 Creating State Transition Matrices

Thus, the transition matrices of states of the Guaimbê Complex were created, represented in Figure 7.

Figure 7 - Transition matrices - Guaimbê Complex.

Summer

	1	2	3	4	5	6	7	8
1	0,00	0,38	0,13	0,25	0,13	0,00	0,13	0,00
2	0,00	0,39	0,43	0,11	0,07	0,00	0,00	0,00
3	0,12	0,09	0,12	0,30	0,14	0,12	0,07	0,05
4	0,02	0,09	0,25	0,26	0,18	0,09	0,11	0,02
5	0,02	0,02	0,07	0,18	0,29	0,24	0,15	0,04
6	0,02	0,02	0,05	0,14	0,24	0,26	0,22	0,05
7	0,00	0,03	0,06	0,06	0,06	0,24	0,40	0,16
8	0,00	0,00	0,03	0,03	0,06	0,11	0,31	0,46

Autumn

	1	2	3	4	5	6	7	8	9
1	0,25	0,13	0,29	0,08	0,08	0,13	0,04	0,00	0,00
2	0,08	0,14	0,22	0,14	0,06	0,22	0,11	0,03	0,00
3	0,19	0,17	0,14	0,14	0,14	0,11	0,06	0,00	0,06
4	0,05	0,16	0,13	0,16	0,21	0,11	0,11	0,08	0,00
5	0,06	0,06	0,08	0,19	0,31	0,15	0,05	0,06	0,03
6	0,02	0,06	0,06	0,05	0,29	0,35	0,08	0,06	0,03
7	0,02	0,13	0,06	0,06	0,06	0,17	0,37	0,13	0,00
8	0,00	0,00	0,00	0,05	0,11	0,03	0,34	0,37	0,11
9	0,00	0,00	0,00	0,00	0,13	0,13	0,06	0,31	0,38

Winter

	1	2	3	4	5	6	7	8	9	10	11
1	0,21	0,00	0,14	0,14	0,21	0,14	0,00	0,00	0,07	0,07	0,00
2	0,20	0,00	0,30	0,20	0,00	0,00	0,10	0,10	0,10	0,00	0,00
3	0,07	0,07	0,07	0,07	0,07	0,14	0,14	0,07	0,00	0,14	0,14
4	0,06	0,11	0,17	0,00	0,11	0,11	0,06	0,17	0,17	0,06	0,00
5	0,09	0,04	0,04	0,13	0,17	0,09	0,22	0,17	0,04	0,00	0,00
6	0,08	0,03	0,03	0,14	0,05	0,27	0,11	0,19	0,08	0,00	0,03
7	0,00	0,05	0,05	0,04	0,04	0,16	0,39	0,16	0,07	0,04	0,02
8	0,03	0,00	0,00	0,03	0,06	0,12	0,22	0,34	0,12	0,06	0,02
9	0,00	0,02	0,00	0,02	0,07	0,02	0,07	0,24	0,40	0,16	0,02
10	0,00	0,02	0,02	0,00	0,02	0,02	0,04	0,06	0,27	0,47	0,08
11	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,39	0,61

Spring

	1	2	3	4	5	6	7	8
1	0,05	0,21	0,16	0,21	0,16	0,05	0,11	0,05
2	0,15	0,12	0,37	0,17	0,05	0,07	0,05	0,02
3	0,02	0,34	0,11	0,11	0,17	0,15	0,09	0,02
4	0,03	0,06	0,17	0,17	0,06	0,29	0,09	0,14
5	0,10	0,15	0,10	0,10	0,10	0,21	0,17	0,06
6	0,03	0,05	0,08	0,09	0,27	0,39	0,05	0,03
7	0,07	0,09	0,11	0,00	0,13	0,20	0,20	0,20
8	0,00	0,02	0,02	0,02	0,04	0,11	0,30	0,49

4.2.2.3 Obtaining the results

Then, stationary distributions (Table 9), recurrence times (Table 10), and first passage times (Figure 8) were calculated.

Table 9: Stationary distribution - Guaimbê Complex.

a) Summer	States							
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Centroid (MWh)	234	386	499	610	715	810	909	1,018
Stationary distribution	0.0227	0.0747	0.1224	0.1582	0.1560	0.1663	0.1986	0.1011

b) Autumn	States								
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Centroid (MWh)	244	459	619	699	760	810	867	924	992
Stationary distribution	0.0663	0.0956	0.1010	0.1045	0.1745	0.1698	0.1412	0.1034	0.0438

c) Winter	States										
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
Centroid (MWh)	182	324	448	545	615	679	741	791	840	900	983
Stationary distribution	0.0372	0.0266	0.0398	0.0475	0.0613	0.0969	0.1433	0.1693	0.1583	0,1489	0,0711

(d) Spring	States							
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Centroid (MWh)	269	457	593	688	780	863	949	1,033
Stationary distribution	0.0545	0.1218	0.1319	0.0997	0.1351	0.2079	0.1270	0.1221

Table 10: Recurrence time - Guaimbê Complex.

a) Summer	States							
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Centroid (MWh)	234	386	499	610	715	810	909	1,018
Recurrence time (days)	44	13	8	6	6	6	5	10

b) Autumn	States								
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Centroid (MWh)	244	459	619	699	760	810	867	924	992
Recurrence time (days)	15	10	10	10	6	6	7	10	23

c) Winter	States										
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
Centroid (MWh)	182	324	448	545	615	679	741	791	840	900	983
Recurrence time (days)	27	38	25	21	16	10	7	6	6	7	14

(d) Spring	States							
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Centroid (MWh)	269	457	593	688	780	863	949	1.033
Recurrence time (days)	18	8	8	10	7	5	8	8

Figure 8 - First passage time - Guaimbê Complex.

Summer

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
1	0	14	7	6	8	9	9	20
2	42	0	4	7	8	9	10	21
3	39	20	0	6	8	8	8	19
4	43	21	7	0	7	8	8	19
5	45	23	10	7	0	6	7	18
6	45	23	10	8	7	0	7	18
7	46	24	11	9	9	6	0	15
8	47	25	11	10	9	7	5	0

Autumn

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
1	0	10	8	10	7	8	11	17	38
2	19	0	9	10	7	7	10	16	38
3	17	10	0	10	7	8	11	17	36
4	20	10	11	0	7	8	10	16	38
5	20	12	11	9	0	8	11	16	37
6	21	12	12	11	6	0	10	15	37
7	21	11	12	11	8	7	0	14	38
8	23	13	14	12	8	9	6	0	33
9	23	14	14	13	7	8	9	9	0

Winter

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
1	0	7	9	7	8	12	13	13	17	24	62
2	32	0	9	8	7	10	12	12	15	23	61
3	33	13	0	8	7	10	12	11	15	23	60
4	36	13	8	0	7	10	12	11	15	22	60
5	38	15	10	8	0	10	11	10	14	21	58
6	39	15	11	8	6	0	10	10	14	21	59
7	39	15	12	9	8	9	0	9	13	19	57
8	38	16	11	9	7	11	10	0	12	18	55
9	39	16	12	10	8	11	9	7	0	17	55
10	40	17	13	11	9	12	10	6	11	0	54
11	40	18	13	10	9	13	13	9	12	15	0

Spring

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
1	0	8	7	9	7	7	9	15
2	17	0	6	9	8	7	10	16
3	19	7	0	10	7	7	10	16
4	19	10	8	0	8	6	9	14
5	18	9	8	10	0	6	9	15
6	19	10	9	11	6	0	10	16
7	18	9	8	12	7	6	0	13
8	20	11	10	13	8	7	6	0

4.2.3 Post-processing

4.2.3.1 Evaluation and interpretation of results

With the measurements of interest obtained for the Guaimbê Complex, the next step is to analyze the model's characteristics for each of the seasons.

In summer, the stationary probability of the system being in the state of lower power generation is the lowest (2.27%), resulting in a recurrence time of 44 days, that is, the occurrence of a state of low generation is extremely rare, and, once in this state, many transitions are expected for the return. Furthermore, state 2 (386 MWh) has the second-lowest stationary probability at 7.47%, followed by state 8 (1,018 MWh) at 10.11%. The states with the highest probabilities are the higher power plants, and the state with the highest stationary probability is state 7 (909 MWh), with 19.86%.

Looking at autumn, the system has a higher probability of being stationary in the upper central states of power generation values in the seasons, with the probabilities gradually decreasing to the lower and upper extreme states. The two states with the longest recurrence times are the extremes, states 1 (244 MWh) and 9 (992 MWh), with 15 and 23 days, respectively. Winter, on the other hand, in the Guaimbê Complex, has higher stationary probabilities for the upper states and very low probabilities for the four states with lower energy generation. An interesting case is that the recurrence time of state 2 (324 MWh) is 38 days, which is approximately 40% longer than the time of state 1 (182 MWh) 27 days. In this way, the risk of the system being in low-generation states is lower, and there is an expectation of higher energy generations, comparatively.

Spring has more balanced stationary probabilities among its eight states, with the exception of state 1 (269 MWh), which has lower power generation and a probability of only 5%.

Analyzing the times of the first passage, it can be seen that, in the summer of the Guaimbê Complex, the time to leave the state of the highest generation to the state of the lowest generation is more than double the reverse. The first passage time from state 8 (1,018 MWh) to state 1 (234 MWh) is 47 days, while from state 1 to state 8 is 20 days. Hence, this characteristic is favorable to generation because the average time to have a low generation from a high generation is high. However, autumn and winter have first passage times with the reverse logic, it takes longer to move from a state of lower generation to one of greater generation. This analysis is important because, in low-generation situations, the expected time to return to high-generation is longer. In the autumn, the first passage time from state 1 (244 MWh) to state 9 (992 MWh) is 38 days, and the reverse is 23 days. Meanwhile, in winter, the time from state 1 (182 MWh) to state 11 (983 MWh) is 62 days, and the reverse is 30 days.

4.3 Comparison of results

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 are evident since Nova Olinda presents seasonality with higher average generation in winter and lower in summer, which is the wet period, while the averages of Guaimbê are closer in all climatic seasons. In the case of Nova Olinda, the reason for subdividing the series by the climatic seasons to perform the modeling is more evident, however, although the Guaimbê Complex presents more homogeneous monthly averages, the results for the stationary distributions and recurrence and first passage times were significantly different in each season, as previously analyzed. Thus, the subdivision by climatic 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 highest concentration of stationary probabilities in upper central states is a case present in summer and spring in the Nova Olinda Complex, but it also happens in the autumn in the Guaimbê Complex. On the other hand, the autumn of Nova Olinda is similar to the winter of Guaimbê because the states of lower generations have significantly lower stationary probabilities than the others and higher probabilities in the higher states. Meanwhile, Nova Olinda's winter and Guaimbê's spring are the seasons with the most balanced stationary probabilities between the states.

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

Finally, BES can use this analysis to assist in the country's energy planning by calculating the probability of possible scenarios of low or high photovoltaic generation by region and climatic season. The detailed study of the characteristics of renewable sources brings greater security to the supply of energy demand in the country.

5. CONCLUSION

Brazil has been going through a process of changing its energy matrix and increasing the use of renewable energies non-dispatchable. In this context, photovoltaic solar energy has stood out due to the significant growth of its share in the country. Hence, its characteristics of intermittency and random fluctuations have a greater impact on the national energy supply scenario. Therefore, the study of photovoltaic generation through modeling methods is relevant, and an opportunity to contribute to the literature was found through the present work.

This work studies the generation characteristics of two photovoltaic solar power plants located in regions with solar incidences of different magnitudes and seasonalities. The methodology used was based on Markov Chains. The time series were subdivided among the climatic seasons of the year. Then, the state transition matrices were created, and the results of the measures of interest, such as stationary distribution, recurrence time, and first passage time, were investigated. Consequently, it was possible to analyze the differences between the photovoltaic energy generation in the different seasons and regions. In this way, the objective of the work was achieved in a pertinent way.

Confirming the initial hypothesis, the results showed significant differences in solar energy generation between the regions and between the climatic seasons, which evidenced the relevance of the comparative study carried out. By analyzing and better understanding the specificities of each location and season, power plants and the Brazilian Electric System can plan more efficiently about energy generation, analyzing the probabilities of the occurrence of states of different generation values.

6. REFERENCES

Chung, K. L. (1960). Markov Chains with Stationary Transition Probabilities. Springer-Verlag.

Dickey, D. & Fuller, W. (1979). Distribution of the Estimators for Autoregressive Time Series With a Unit Root. Journal of the American Statistical Association, v. 74, p. 427-431.

EPE (2022). Fontes de Energia. https://www.epe.gov.br/pt/abcdenergia/fontes-de-energia

EPE (2023). Balanço Energético Nacional 2023. Ministério de Minas e Energia.

Fritz, M., Behringer, M. & Schwarz, H. (2020). LOG-Means: efficiently estimating the number of clusters in large datasets. Proceedings of the VLDB Endowment, v. 13, n. 12.

G1 (2019). Complexos de energia solar são inaugurados em duas cidades do interior de SP. https://g1.globo.com/sp/bauru-marilia/noticia/2019/08/15/complexos-de-energia-solar-serao-inaugurados-em-duas-cidades-do-interior-de-sp.ghtml

Gadelha, M. L. (2020). Climas do Piauí: interações com o ambiente. Edufpi.

IEA (2023). Renewable Energy Market Update - June 2023. https://www.iea.org/reports/renewable-energy-market-update-june-2023

Imhoff, J. (2007). Desenvolvimento de Conversores Estáticos para Sistemas Fotovoltaicos Autônomos. Master's thesis presented to the School of Electrical Engineering of Universidade Federal de Santa Maria.

Ma, M., Ye, L., Li, J., Li, P., Song, R. & Zhuang, H. (2020). Photovoltaic Time Series Aggregation Method Based on K-means and MCMC Algorithm. Asia-Pacific Power and Energy Engineering Conference, v. 2020-September, n. 9220338.

Maçaira, P., Cyrillo, Y. M., Cyrino, F. & Souza, R. C. (2019). Including wind power generation in Brazil's long-term optimization model for energy planning. Energies, v. 12, n. 826.

Macqueen, J. (1967). 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.

Malar, J. P. (2022). Conheça os tipos de energia renovável e quais são usados no Brasil. CNN Brasil. https://www.cnnbrasil.com.br/business/conheca-os-tipos-de-energia-renovavel-e-quais-sao-usados-no-brasil

Melo, G., Cyrino, F. & Maçaira, P. (2022). Simulação estocástica conjunta de energias renováveis. Master's thesis – Departamento de Engenharia Industrial, Pontifícia Universidade Católica do Rio de Janeiro.

Nascimento, A. & Araújo, T. (2017). Maior parque solar da América Latina é inaugurado no Piauí. https://g1.globo.com/pi/piaui/noticia/maior-parque-solar-da-america-latina-e-inaugurado-no-piaui.ghtml

Norris, J. R. (1998). Markov chains. Cambridge university press.

ONS (2022). http://www.ons.org.br

Pereira, E. B., Martins, F. R., Gonçalves, A. R., Costa, R. S., Lima, F. L., Rüther, R., Abreu, S. L., Tiepolo, G. M., Pereira, S. V. & Souza, J. G. Atlas brasileiro de energia solar. 2.ed. São José dos Campos: INPE.

R Development Core Team (2009). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. http://www.R-project.org

Sigauke, C. & Chikobvu, D. (2017). Estimation of extreme inter-day changes to peak electricity demand using Markov chain analysis: A comparative analysis with extreme value theory. Journal of Energy in Southern Africa, v. 28, n. 4.

Thorndike, R. L. (1953). Psychometrika, v. 18, p. 266-267.