



UNIVERSIDADE ESTADUAL DA PARAÍBA  
CENTRO DE CIÊNCIAS E TECNOLOGIA  
DEPARTAMENTO DE ESTATÍSTICA

ANTONIO LEOPOLDO CARDOSO SABINO

**UMA APLICAÇÃO DE CADEIAS DE  
MARKOV DE ALTA ORDEM COM DADOS  
DE PRECIPITAÇÃO NO ESTADO DA  
PARAÍBA**

CAMPINA GRANDE - PB

2019

ANTONIO LEOPOLDO CARDOSO SABINO

**UMA APLICAÇÃO DE CADEIAS DE MARKOV DE  
ALTA ORDEM COM DADOS DE PRECIPITAÇÃO NO  
ESTADO DA PARAÍBA**

Trabalho de Conclusão de Curso apresentado ao curso de Bacharelado em Estatística do Departamento de Estatística do Centro de Ciências e Tecnologia da Universidade Estadual da Paraíba em cumprimento às exigências legais para obtenção do título de bacharel em Estatística.

Orientador: Prof. Dr. Sílvio Fernando Alves Xavier Júnior

CAMPINA GRANDE - PB

2019

É expressamente proibido a comercialização deste documento, tanto na forma impressa como eletrônica. Sua reprodução total ou parcial é permitida exclusivamente para fins acadêmicos e científicos, desde que na reprodução figure a identificação do autor, título, instituição e ano do trabalho.

S116a Sabino, Antonio Leopoldo Cardoso.  
Uma aplicação de cadeias de Markov de alta ordem com dados de precipitação no Estado da Paraíba [manuscrito] / Antonio Leopoldo Cardoso Sabino. - 2019.  
30 p. : il. colorido.  
Digitado.  
Trabalho de Conclusão de Curso (Graduação em Estatística) - Universidade Estadual da Paraíba, Centro de Ciências e Tecnologia , 2019.  
"Orientação : Prof. Dr. Sílvio Fernando Alves Xavier Júnior , Coordenação do Curso de Estatística - CCT."  
1. Cadeia de Markov. 2. Alta Ordem. 3. Modelos de Markov Ocultos. I. Título  
21. ed. CDD 519.5

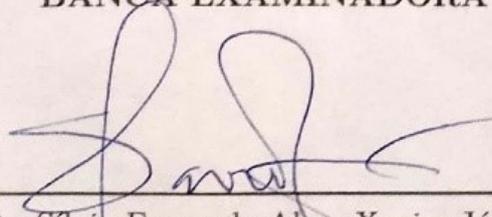
Antonio Leopoldo Cardoso Sabino

# UMA APLICAÇÃO DE CADEIAS DE MARKOV DE ALTA ORDEM COM DADOS DE PRECIPITAÇÃO NO ESTADO DA PARAÍBA

Trabalho de Conclusão de Curso apresentado ao curso de Bacharelado em Estatística do Departamento de Estatística do Centro de Ciências e Tecnologia da Universidade Estadual da Paraíba em cumprimento às exigências legais para obtenção do título de bacharel em Estatística.

Trabalho aprovado em 4 de DEZEMBRO de 2019.

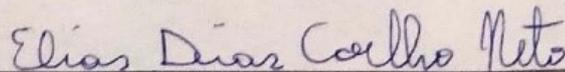
## BANCA EXAMINADORA



Dr. Sílvio Fernando Alves Xavier Júnior  
Universidade Estadual da Paraíba

*DMEsteves*

Dra. Divanilda Maia Esteves  
Universidade Estadual da Paraíba



Dr. Elias Dias Coelho Neto  
Universidade Estadual da Paraíba

# Agradecimentos

A Deus por ter me dado saúde e força para superar todos os obstáculos que surgiram tanto na vida social, quanta na acadêmica.

A Universidade Estadual da Paraíba por me proporcionar essa grandiosa oportunidade acadêmica, seu corpo docente, direção e administração que oportunizaram a janela que hoje vislumbro um horizonte superior, obtido pela confiança no mérito e ética aqui presentes.

Ao meu orientador Prof. Dr. Sílvio Fernando Alves Xavier Júnior, pelo suporte no pouco tempo que lhe coube, pelas suas correções e incentivos.

Aos meus pais Maria Edileuza Sabino Cardoso e Jose Cardoso e minha irmã Vitoria Cristina Cardoso Sabino, pelo amor, incentivo e apoio incondicional.

E a todos meus colegas e amigos que direta ou indiretamente fizeram parte da minha formação, o meu muito obrigado.

*“A estatística é a gramática da ciência.”*

*(Karl Pearson)*

*“Essencialmente, todos os modelos estão errados, mas alguns são úteis.”*

*(George Box)*

*“A verdadeira ciência ensina sobretudo a duvidar e a ser ignorante.”*

*(Miguel de Unamuno)*

# Resumo

O presente trabalho aborda alguns dos principais conceitos de Cadeia de Markov discreta em três estágios: modelos de Markov Clássico, Ocultos e Multivariados de Alta Ordem. Introduziu-se uma forma de se obter as probabilidades de transição Clássica e de Alta Ordem e a partir delas gerar as respectivas matrizes de transição com suas probabilidades. O algoritmo foi implementado no ambiente de programação computacional estatística e gráfica R e RStudio, onde foi analisado um banco de dados da precipitação de chuva diária da cidade de Campina Grande e João Pessoa no período de 1963 a 2017 da estação de Campina Grande-PB de código OMM: 82795 e João Pessoa-PB de código OMM: 81918 - INMET(Instituto Nacional de Meteorologia). O estudo teve o principal foco em avaliar estatisticamente utilizando a modelagem de cadeias de Markov no âmbito climático para responder se os  $k$  dias têm influência no clima do próximo dia. Dessa forma, a modelagem avaliou o comportamento da precipitação da cidade de Campina Grande e João Pessoa a partir da precipitação dos dias anteriores com a utilização dos três estados definidos. Resultados preliminares evidenciaram a influência do estado (seco, úmido e chuvoso) de  $k$  dias anteriores na ocorrência do estado do dia subsequente.

**Palavras-chave:** Cadeia de Markov, Alta Ordem e Modelos de Markov Ocultos.

# Abstract

This paper discusses some of the main concepts of discrete Markov Chain in three stages: Markov Classic, Hidden and High Order Multivariate models. A way to get the Classic and High transition probabilities has been introduced. Order and from them generate the respective transition matrices with their probability. The algorithm was implemented in the statistical computational programming environment R and RStudio where a database of daily rainfall was analyzed from Campina Grande and João Pessoa in the period from 1963 to 2017 of the Campina Grande-PB (OMM code: 82795) and João Pessoa-PB (OMM code: 81918) - INMET (National Institute of Meteorology). The study focused primarily on evaluating statistically using climate modeling of Markov chains to answer if  $k$  days have influence on the weather of the next day. Thus, modeling evaluated the precipitation behavior of the city of Campina Grande and João Pessoa from the precipitation of the previous days using the three defined states. Preliminary results showed the influence of the state (dry, wet and rainy) of  $k$  previous days in the occurrence of the subsequent day state.

**Key-words:** Markov Chain, High Order and Hidden Markov Models.

# Lista de ilustrações

Figura 1 – Gráfico da precipitação diária na cidade Campina Grande, PB . . . . .	19
Figura 2 – Gráfico da precipitação diária na cidade João Pessoa, PB . . . . .	20
Figura 3 – Estados da matriz de transição em três estados de Campina Grande, PB	22
Figura 4 – Estados da matriz de transição em três estados de João Pessoa, PB . .	22
Figura 5 – Gráfico de transição da Cadeia de Markov de Campina Grande, PB . .	23
Figura 6 – Gráfico de transição da Cadeia de Markov de João Pessoa, PB . . . . .	23

# Lista de tabelas

Tabela 1 – Matriz de Transição com suas Probabilidades . . . . .	11
Tabela 2 – Teste de Normalidade de Anderson-Darling . . . . .	18
Tabela 3 – Análise Descritiva da Precipitação Diária(mm) em Campina Grande, PB.	18
Tabela 4 – Análise Descritiva da Precipitação Diária(mm) em João Pessoa, PB. .	19
Tabela 5 – Matriz de transição com três estados de Campina Grande, PB . . . . .	20
Tabela 6 – Matriz de transição com três estados de João Pessoa, PB. . . . .	21
Tabela 7 – Critérios de Informação para João Pessoa . . . . .	24
Tabela 8 – Critérios de Informação para Campina Grande . . . . .	24
Tabela 9 – Matriz de transição de Alta Ordem para Campina Grande com $k = 2$ .	25
Tabela 10 – Matriz de transição de Alta Ordem para João Pessoa com $k = 3$ . . . .	26

# Sumário

<b>1</b>	<b>INTRODUÇÃO</b>	<b>10</b>
<b>2</b>	<b>METODOLOGIA</b>	<b>11</b>
<b>2.1</b>	<b>Cadeias de Markov</b>	<b>11</b>
<b>2.2</b>	<b>Modelos de Markov Ocultos</b>	<b>12</b>
<b>2.3</b>	<b>Teste de Anderson Darling</b>	<b>12</b>
<b>2.4</b>	<b>Critério de informação</b>	<b>12</b>
2.4.1	Critérios de informação de Akaike (AIC)	13
2.4.2	Critérios de informação de Akaike Corrigido (AICc)	13
2.4.3	Critério de informação de (AKICc)	13
2.4.4	Critério de informação (CAIC)	13
2.4.5	Critério de informação Bayesiano (BIC)	13
2.4.6	Critério de informação (KIC)	14
2.4.7	Critério de informação de (KICC)	14
2.4.8	Critério de informação de (HQ)	14
<b>2.5</b>	<b>Cadeias de Markov Multivariadas de Alta Ordem</b>	<b>14</b>
<b>3</b>	<b>MATERIAIS E MÉTODOS</b>	<b>16</b>
<b>4</b>	<b>RESULTADOS E DISCUSSÃO</b>	<b>18</b>
<b>4.1</b>	<b>Análise Descritiva</b>	<b>18</b>
<b>4.2</b>	<b>Cadeias de Markov</b>	<b>20</b>
<b>4.3</b>	<b>Cadeias de Markov de Alta Ordem</b>	<b>24</b>
<b>5</b>	<b>CONCLUSÃO</b>	<b>27</b>
	<b>REFERÊNCIAS</b>	<b>28</b>

# 1 Introdução

Na literatura várias pesquisas no âmbito climático foram desenvolvidas no estado da Paraíba, particularmente as regiões de Campina Grande e João Pessoa, em que foram constatadas mudanças climáticas significativas. O estado da Paraíba está incluído em uma das maiores problemáticas, que se pode observar mundialmente que é a crise hídrica associada a grande variabilidade climática (SANTOS; CUNHA; RIBEIRO-NETO, 2019).

É evidente que na maioria das situações envolvendo a variabilidade climática tem-se atribuído a culpabilidade ao ser humano. Em contrapartida, alguns pesquisadores discordam afirmando que trata-se apenas um ciclo natural do clima que ocorre devido a vários fatores, como mudança nos valores de temperatura da superfície do mar (ASSIS et al., 2018).

Segundo Rodrigues et al. (2017) esses grandes períodos de seca estão associados à irregularidade dos índices baixos de precipitação, e também altos valores de chuvas devido a uma assomada de precipitação, responsável por possíveis enchentes e desmoronamentos de casas, diminuição na produção agrícola e rios com baixos níveis de acordo com Ferreira et al. (2017).

Os autores Silva e Clarke (2004) sugerem que uma grande massa de dados no decorrer dos anos pode refletir a variabilidade do clima de uma região, objetivamente em razão das chuvas intensas. Não obstante, Grimm e Tedeschi (2004) avaliaram que os fenômenos El Niño e La Niña são co-influentes no aparecimento de grandes picos de chuva no Brasil, essencialmente na região Nordeste, em conjunto com o deslocamento da ZCIT (Zona de Convergência Intertropical).

A região Nordeste do Brasil tem clima semiárido com curta estação chuvosa, de março a abril, e grandes variações de precipitação de ano para ano. Em sua longa história, as secas impactaram severamente a população humana. A região foi destacada como uma região problemática devido a grande variação climática (HASTENRATH, 2012). A Paraíba é um dos estados brasileiros que apresenta uma escassez acentuada de água e possui clima semiárido na maioria do seu território. Em consonância com (NÓBREGA, 2014), o clima fica mais seco à medida que atinge o continente. A variação climática interna também se reflete em diferentes tipos de solo e vegetação (XAVIER et al., 2018).

Em função desse quadro climático volátil, esse trabalho tem o objetivo em estimar a variabilidade da precipitação em Campina Grande e João Pessoa, identificando os dias que foram outliers. Para a predição dos valores futuros utilizaremos modelos de cadeia de Markov que têm sido amplamente utilizado em precipitação (JALE et al., 2019; AMORIM et al., 2019), velocidade do vento (SANTOS et al., 2019).

## 2 Metodologia

### 2.1 Cadeias de Markov

Os modelos de cadeias de Markov têm sido amplamente utilizados para simular séries temporais discretas. Algumas aplicações incluem processos hidrológicos (FU; LI; HUANG, 2012; JALE et al., 2019; AMORIM et al., 2019). No presente trabalho será utilizada a cadeia de Markov de primeira a  $k$ -ésima ordem para estudar o comportamento das ocorrências de chuva nas cidades de Campina Grande e João Pessoa, Paraíba.

Tabela 1 – Matriz de Transição com suas Probabilidades

		Dia Atual (j)			Total
		Seco (s)	Úmido (u)	Chuvoso (c)	
Dia anterior (i)	Seco (s)	$P_{ss}$	$P_{su}$	$P_{sc}$	$n_s$
	Úmido (u)	$P_{us}$	$P_{uu}$	$P_{uc}$	$n_u$
	Chuvoso (c)	$P_{cs}$	$P_{cu}$	$P_{cc}$	$n_c$

As entradas  $P_{ij}$  da matriz de probabilidade de transição P na Tabela ?? são definidas por:

- $P_{ss} = P(s|s)$ : Probabilidade de uma dia seco ser precedido por um dia seco;
- $P_{su} = P(u|s)$ : Probabilidade de uma dia úmido ser precedido por um dia seco;
- $P_{sc} = P(c|s)$ : Probabilidade de uma dia chuvoso ser precedido por um dia seco;
- $P_{us} = P(s|u)$ : Probabilidade de uma dia seco ser precedido por um dia úmido;
- $P_{uu} = P(u|u)$ : Probabilidade de uma dia úmido ser precedido por um dia úmido;
- $P_{uc} = P(c|u)$ : Probabilidade de uma dia chuvoso ser precedido por um dia úmido;
- $P_{cs} = P(s|c)$ : Probabilidade de uma dia seco ser precedido por um dia chuvoso;
- $P_{cu} = P(u|c)$ : Probabilidade de uma dia úmido ser precedido por um dia chuvoso;
- $P_{cc} = P(c|c)$ : Probabilidade de uma dia chuvoso ser precedido por um dia chuvoso.

Sujeito á condição de que a soma das probabilidades de cada linha seja igual a um, isto é,  $P_{ss} + P_{su} + P_{sc} = 1$  ,  $P_{us} + P_{uu} + P_{uc} = 1$  e  $P_{cs} + P_{cu} + P_{cc} = 1$ .

As probabilidades da matriz de transição podem ser estimadas a partir das correspondentes frequências absolutas observadas  $n_{ij}$  de dias estando em um estado particular  $j$  precedido por um estado  $i$ . Os estimadores de máxima verossimilhança de  $P_{ij}$ ,  $i, j = \{s, u, c\}$ ,  $\hat{P}_{ij}$ , são dados por:

$$\hat{P}_{ij} = \frac{n_{ij}}{\sum_{j=s}^c n_{ij}} \quad (2.1)$$

## 2.2 Modelos de Markov Ocultos

Este tipo particular de modelo é formado por dois processos estocásticos. A primeira parte é uma cadeia de Markov que não é diretamente observável. A segunda parte é um modelo de mistura que é observável, com essa parte é que se forma O Modelo Markov Oculto (MMO).

Esses tipos de modelo foram se desenvolvendo exponencialmente ao longo dos anos para se adaptar a maior quantidade de aspectos teóricos e aplicações, iniciado por Bickel, Ritov et al. (1996), Santos et al. (2006)

## 2.3 Teste de Anderson Darling

O teste de Anderson-Darling utiliza um valor  $a$  que pode identificar qual a melhor distribuição que se ajusta aos dados. Os determinados valores de  $a$  são obtidos a partir da estatística  $AD$ , dado na Eq. 2.2 abaixo, que é manipulada para testar se uma amostra de dados é oriunda de uma população com uma distribuição estabelecida. Quando essa distribuição for a função densidade de probabilidade normal, as hipóteses a serem testadas são dadas na Eq. 2.3. Como exemplo, é fácil testar se um determinado banco de dados é proveniente de uma distribuição Normal para utilização de algum método estatístico que usufrui do pressuposto de normalidade (SCHOLZ; STEPHENS, 1987).

$$AD = -n - \frac{1}{n} \sum_{i=1}^n [2i - 1] [\ln(p_{(i)}) + \ln(1 - p_{(n-i+1)})], \quad (2.2)$$

$$\begin{cases} H_0 : \text{A amostra segue uma distribuição normal.} \\ H_1 : \text{A amostra não segue uma distribuição normal.} \end{cases} \quad (2.3)$$

## 2.4 Critério de informação

Se uma boa estimativa para a log verossimilhança esperada puder ser obtida através dos dados observados, esta estimativa poderá ser utilizada como um critério para comparar modelos. Assim, um modo de comparar  $n$  modelos é simplesmente comparar os pesos da função maximizada  $L(\hat{\theta}_i)$ . Entretanto, tal método não fornece uma verdadeira comparação,

haja vista que, em não conhecendo o verdadeiro modelo  $g(x)$ , primeiramente o método da máxima verossimilhança estima os parâmetros de cada modelo  $g_i(x), i = 1, 2, \dots, n$ , e posteriormente são utilizados os mesmos dados para estimar  $E_g [\log f(x|\hat{\theta})]$ , isto introduz um viés em  $L(\hat{\theta}_i)$ , sendo que, a magnitude deste viés varia de acordo com a dimensão do vetor de parâmetros.

Deste modo, os critérios de informação são construídos para avaliar e corrigir o viés( $b(g)$ ) da função suporte. Segundo Konishi e Kitagawa (2008), um critério de informação tem a forma que observa-se na Eq. 2.4:

$$CI(X_n, \hat{G}) = -2 \sum_{i=1}^n \log f(X_i | \hat{\theta}(X_n)) + 2(b(G)). \quad (2.4)$$

#### 2.4.1 Critérios de informação de Akaike (AIC)

Akaike (1974) mostrou que o viés é dado assintoticamente por  $p$ , em que  $p$  é o número de parâmetros a serem estimados no modelo, e definiu seu critério de informação como está descrito na Eq. 2.5:

$$AIC = -2 \log L(\hat{\theta}) + 2(p) \quad (2.5)$$

#### 2.4.2 Critérios de informação de Akaike Corrigido (AICc)

Bozdogan (1987) propôs a seguinte correção para o AIC na Eq. 2.6:

$$AIC_c = -2 \log L(\hat{\theta}) + 2(p) + 2 \frac{p(p+1)}{n-p-1} \quad (2.6)$$

#### 2.4.3 Critério de informação de (AKICc)

O critério de informação de AKIC corrigido foi descrito na Eq. 2.7 (SOBRAL; BARRETO, 2016):

$$AKIC_C = -2 \log f(x_n | \hat{\theta}) + \frac{(p+1)(3n-p-2)}{n-p-2} + \frac{p}{n-p} \quad (2.7)$$

#### 2.4.4 Critério de informação (CAIC)

Critério de informação AIC Consistente (CABRAL et al., 2012), que é descrito na Eq. 2.8 :

$$CAIC = -2 \log(L) + (\log(n) + 1)K \quad (2.8)$$

#### 2.4.5 Critério de informação Bayesiano (BIC)

Proposto por Schwarz et al. (1978) é dado pela Eq. 2.9:

$$BIC = -2 \log f(x_n | \theta) + p \log(n) \quad (2.9)$$

### 2.4.6 Critério de informação (KIC)

O critério de informação de Kullback foi descrito na Eq. 2.10 (YAMANOUCHI, 2017):

$$KIC = -2 \log f(x_n | \theta) + 3p \quad (2.10)$$

### 2.4.7 Critério de informação de (KICC)

O critério de informação de Kullback Corrigido é descrito na Eq. 2.11 (SARAH, 2015):

$$KICC = n(\log \left| \hat{\Sigma} \right| + q) + \frac{d(3n - p - q - 1)}{n - p - q - 1} \quad (2.11)$$

### 2.4.8 Critério de informação de (HQ)

O critério de informação de Hannan-Quinn é dado pela Eq. 2.12 (MICHE; LENDASSE, 2009).

$$HQ = N \log L(\hat{\theta}) + 2p \log(\log(N)) \quad (2.12)$$

## 2.5 Cadeias de Markov Multivariadas de Alta Ordem

O modelo de Cadeias de Markov Multivariadas de Alta Ordem (HOMMC) é utilizado para modelar o comportamento de múltiplas sequências categóricas geradas por fontes semelhantes (CHING; NG; FUNG, 2008). Suponha que haja sequências categóricas e cada uma tem estados possíveis em  $M$ . Na  $n$ -ésima ordem  $MMC$  a probabilidade do estado na  $j$ -ésima sequência no tempo  $t = r + 1$  depende da distribuição de probabilidade do estado de todas as sequências (incluindo-se às vezes)  $t = r, r - 1, \dots, r - n + 1$ .

$$x_{r+1}^{(j)} = \sum_{k=1}^s \sum_{h=1}^n \lambda_{jk}^{(h)} x_{r-h+1}^{(k)}, \quad \begin{cases} j=1,2,\dots,s \\ r=n-1,n,\dots \end{cases} \quad (2.13)$$

com distribuição inicial  $x_0^{(k)}, x_1^{(k)}, \dots, x_{n-1}^{(k)}$  ( $k = 1, 2, \dots, s$ ), onde:

- $\lambda_{jk}^{(h)} \geq 0$ , onde  $\lambda_{jk}^{(h)}$  é o vetor de pesos.
- $1 \leq j$
- $k \leq s$
- $1 \leq h \leq n$ , onde  $h$  é número de passos.
- $\sum_{k=1}^s \sum_{h=1}^n \lambda_{jk}^{(h)} = 1$
- $j = 1, 2, \dots, s$

A seguir tem-se a representação simplificada do modelo que facilita a compreensão resultado do método *fitHighOrderMultivarMC* (MAINTAINER; SPEDICATO, 2019).

$$\begin{pmatrix} X_{r+1}^{(1)} \\ X_{r+1}^{(2)} \\ \cdot \\ \cdot \\ X_{r+1}^{(s)} \end{pmatrix} = \begin{pmatrix} B^{11} & B^{12} & \cdot & \cdot & B^{1s} \\ B^{21} & B^{22} & \cdot & \cdot & B^{2s} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ B^{s1} & B^{s2} & \cdot & \cdot & B^{ss} \end{pmatrix} \begin{pmatrix} X_r^{(1)} \\ X_r^{(2)} \\ \cdot \\ \cdot \\ X_r^{(s)} \end{pmatrix}$$

$$B^{ii} = \begin{pmatrix} \lambda_{ii}^{(1)} P_1^{(ii)} & \lambda_{ii}^{(2)} P_2^{(ii)} & \cdot & \cdot & \lambda_{ii}^{(n)} P_n^{(ii)} \\ I & 0 & \cdot & \cdot & 0 \\ 0 & I & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & I & 0 \end{pmatrix}_{mn*mn}$$

$$B^{ij} = \begin{pmatrix} \lambda_{ij}^{(1)} P_1^{(ij)} & \lambda_{ij}^{(2)} P_2^{(ij)} & \cdot & \cdot & \lambda_{ij}^{(n)} P_n^{(ij)} \\ 0I & 0 & \cdot & \cdot & 0 \\ 0 & 0 & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & 0 & 0 \end{pmatrix}_{mn*mn}$$

### Representação dos Parâmetros

Por exemplo:  $P_h^{(ij)}$  é representado como  $Ph(i, j)$  e  $\lambda_{ij}^{(h)}$  como  $\text{Lambda}(i, j)$ , onde  $P_4^{(15)}$  pode ser reescrito por  $P4(1, 5)$  e  $\lambda_{34}^{(1)}$  por  $\text{Lambda}(3, 4)$ .

### 3 Materiais e Métodos

Neste trabalho foram utilizados 14913 observações de precipitação diária de uma estação meteorológica localizada na cidade de Campina Grande-PB e 18117 observações de precipitação diária localizada na cidade de João Pessoa-PB, no período compreendido entre 02/02/1963 e 13/12/2017, coletados todos os dias as 12:00 horas. A análise descritiva indica que em todo o estudo choveu um total acumulado de 33190,20mm em Campina Grande e em um total acumulado de João Pessoa 95553.50mm.

Existe uma grande variabilidade dos dados de João Pessoa em comparação com Campina Grande. Há indícios de uma grande dispersão nos dados, ou seja, temos vários dias que apresentam um alto índice de precipitação o que não é naturalmente comum acontecer no dia a dia, na região de estudo. O Banco de Dados de precipitação diária utilizado foi obtido do INMET referente aos anos de 1963 a 2017 da estação de Campina Grande-PB de código OMM: 82795 e João Pessoa-PB de código OMM: 81918 (METEOROLOGIA, 2017).

Primeiramente foi feita uma análise descritiva da precipitação das duas cidades para observar qual era o comportamento dos dados e, posteriormente, foi feita a categorização da variável precipitação de quantitativa contínua para qualitativa nominal. A precipitação diária foi categorizada de acordo com o trabalho Ratan e Venugopal (2013) uma vez que o Departamento Meteorológico da Índia afirma que a precipitação diária tem um limiar de 2.5 mm para considerar um dia como chuvoso. Por outro lado, Ali e Mubarak (2017) sugerem que uma precipitação abaixo de 5.00 mm é inapropriada, esse nível de precipitação é muito insignificante pois iria evaporar logo antes de se adentrar ao solo. Fixou-se um limiar de 2.5 mm para o ponto de corte e assim categorizar a precipitação em estados seco, úmido e chuvoso, como é descrito a seguir.

- Se o valor foi menor ou igual a 2.5 mm o dia era determinado como seco(S);
- Se o valor fosse de 2.51 mm até 5.0 mm dia era determinado como úmido(U);
- Se o valor for acima de 5.1 mm o dia era determinado chuvoso(C).

Após a categorização da variável foi possível a estimação das cadeias de Markov Clássica e de Alta Ordem com a utilização do software R (TEAM, 2018) e com o auxílio de um pacote chamado *Markov Chains* criado por **Giorgio Spedicato** (SPEDICATO et al., 2016). Entretanto, referente à citação das Cadeias de Markov de Alta Ordem era necessário que se estimasse um  $k$ , que é dessa forma, foram gerados 7 modelos, ou seja,  $k = 0, 1, 2, \dots, 6$ . O modelo identificado com os menores valores nos critérios de informação

é definido como o modelo  $k$  de Alta Ordem escolhido, indicando quantidade de dias anteriores necessário para explicar o clima do próximo dia.

## 4 Resultados e discussão

### 4.1 Análise Descritiva

Com o objetivo de observar a distribuição da precipitação utilizou-se o teste de normalidade de Anderson-Darling (AD) para as duas cidade em estudo. O resultado do teste comprovou estatisticamente pelos seus respectivos  $p$ -valores que a precipitação diária não se distribui normalmente. A hipótese nula foi rejeitada ao nível de 5% de significância, ou seja  $\alpha = 5\%$  ( $H_0$ ). A Tabela 2 apresenta o resultado do teste AD.

Tabela 2 – Teste de Normalidade de Anderson-Darling

Estatística	Campina Grande	João Pessoa
AD	3246.4	3562.7
P-valor	2.20E-16	2.20E-16

Na tabela 3 tem-se alguns números referente a precipitação diária para a cidade de Campina Grande, PB.

Tabela 3 – Análise Descritiva da Precipitação Diária(mm) em Campina Grande, PB.

Parâmetros	Precipitação
1° Quartil	0
3° Quartil	1.10
Mediana	0
Média	2.23
Variância	42.63
Somatório	33190.2
Máximo	110.1
Mínimo	0

Como anteriormente foi constatado que a variável precipitação não segue normalidade, a média não será utilizada, pois ela não representa a verdadeira média da variável em estudo. Como a maioria dos dias são secos na cidade de Campina Grande-PB a mediana foi 0 ou seja 50% dos dias da série em estudo são secos.

Tabela 4 – Análise Descritiva da Precipitação Diária(mm) em João Pessoa, PB.

Parâmetros	Precipitação
1º Quartil	0
3º Quartil	3.80
Mediana	0.10
Média	5.27
Variância	193.09
Somatório	95553.50
Máximo	194
Mínimo	0

A Tabela 4 ilustra a análise descritiva para a variável precipitação na cidade de João Pessoa. É perceptível que o cenário para a precipitação é diferente, pois 25% dos dias tem uma precipitação acima de 3.80 mm e uma precipitação total acumulada de 95553.50mm.

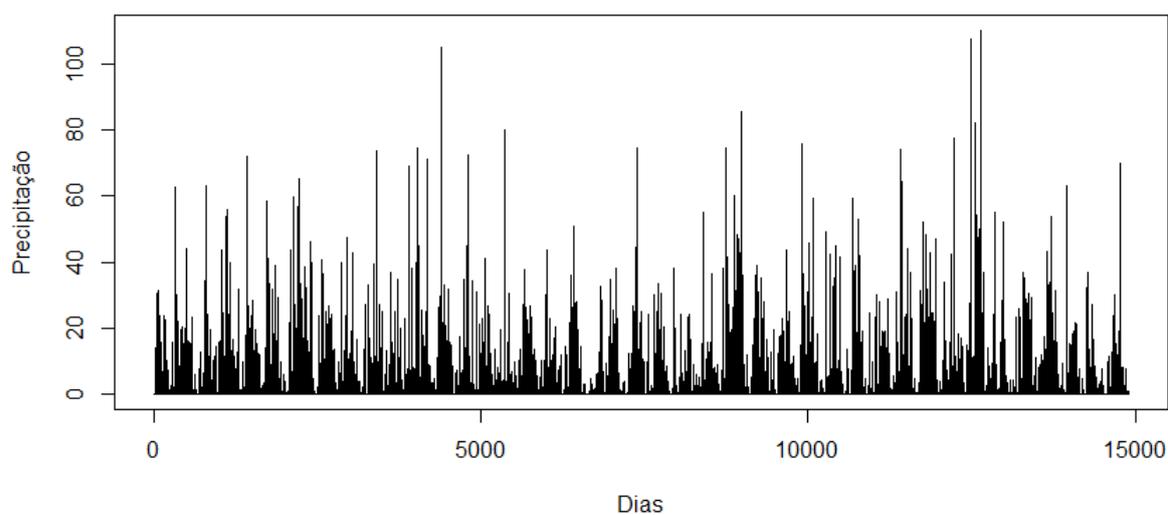


Figura 1 – Gráfico da precipitação diária na cidade Campina Grande, PB

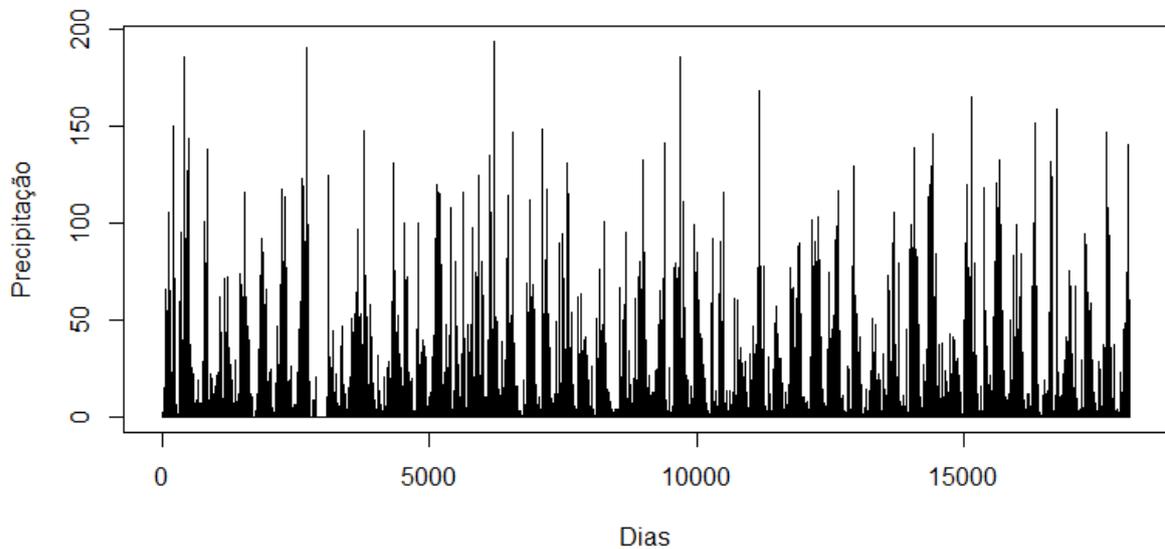


Figura 2 – Gráfico da precipitação diária na cidade João Pessoa, PB

As Figuras 1 e 2 evidenciam graficamente o comportamento da precipitação diária nas cidades de Campina Grande e João Pessoa. Pode-se observar que a chuva durante esses anos tem indícios de sazonalidades na série, e também vários dias que tiveram um alto índice de precipitação. Em Campina Grande ocorreu um dia que choveu cerca de 110 mm e em João Pessoa houve dia que ocorreu índice de 193 mm de chuva, um valor consideravelmente alto se comparado a Campina Grande.

## 4.2 Cadeias de Markov

Neste capítulo será mostrado a seguir alguns resultados referente a Cadeia de Markov.

Tabela 5 – Matriz de transição com três estados de Campina Grande, PB

		Dia Atual (j)			Total
		Seco (s)	Úmido (u)	Chuvoso (c)	
Dia anterior (i)	Seco (s)	0.87	0.05	0.09	1
	Úmido (u)	0.67	0.1	0.23	1
	Chuvoso (c)	0.57	0.13	0.31	1

Na Tabela 5 observa-se que a matriz de transição estimada para a cidade de Campina Grande apresenta maiores indícios que o dia subsequente seja seco, uma vez que

independente da ocorrência de algum dos três estados climáticos as maiores probabilidades indicam que o dia subsequente seja seco.

Tabela 6 – Matriz de transição com três estados de João Pessoa, PB.

		Dia Atual (j)			Total
		Seco (s)	Úmido (u)	Chuvoso (c)	
Dia anterior (i)	Seco (s)	0.8	0.06	0.14	1
	Úmido (u)	0.6	0.1	0.3	1
	Chuvoso (c)	0.44	0.11	0.45	1

Na Tabela 6 nota-se que a matriz de transição estimada para a cidade de João Pessoa há evidências que dado o dia tenha sido seco tem-se uma probabilidade que o dia posterior seja seco. Por outro lado, se o dia anterior foi chuvoso isso muda pois a chance que ele seja chuvoso novamente é de 45%.

Figura 3 – Estados da matriz de transição em três estados de Campina Grande, PB

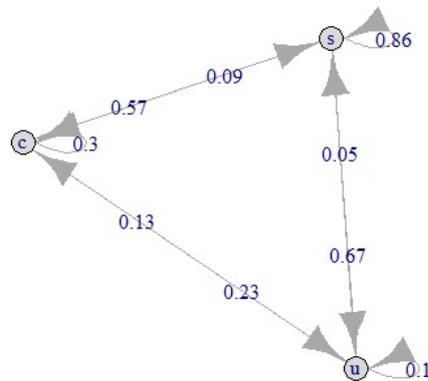
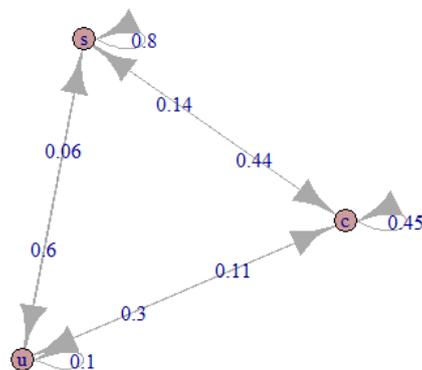


Figura 4 – Estados da matriz de transição em três estados de João Pessoa, PB



A Figura 3 descreve graficamente como é o comportamento de cada estado da matriz de transição, relatando suas respectivas probabilidades com a chance que ele vá do seu estado para o seu estado vizinho e não volte. Para um melhor entendimento, será explicado o estado seco por esse gráfico, no círculo do estado seco (s) existe uma linha que sai dele e volta para ele com um valor de 0.86, essa informação reflete que se está no estado seco a chance que ele fique no mesmo estado seco é alta em torno de 86%, já as outras duas linhas que saem do estado seco e vão para o úmido e chuvoso tem probabilidade baixa de acontecer, 0.05 que vá para o úmido e subsequentemente 0.09 para o o chuvoso.

Na Figura 4 pode-se observar a representação gráfica dos estados das matrizes de

transição de João Pessoa com suas respectivas probabilidades entre os estados. No estado chuvoso ele tem 45% de chance de não sair do próprio estado, já a chance que ele transite para outro estado é baixo se for para o estado úmido em torno de 11% e consequentemente 44% de chance que ele vá para o estado seco uma chance razoavelmente alta.

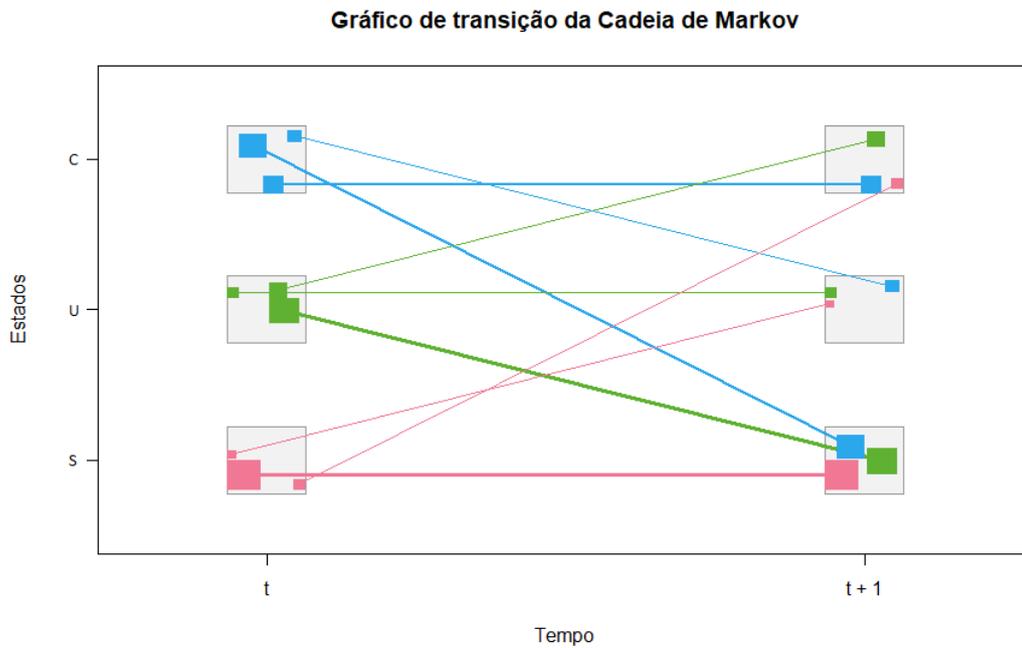


Figura 5 – Gráfico de transição da Cadeia de Markov de Campina Grande, PB

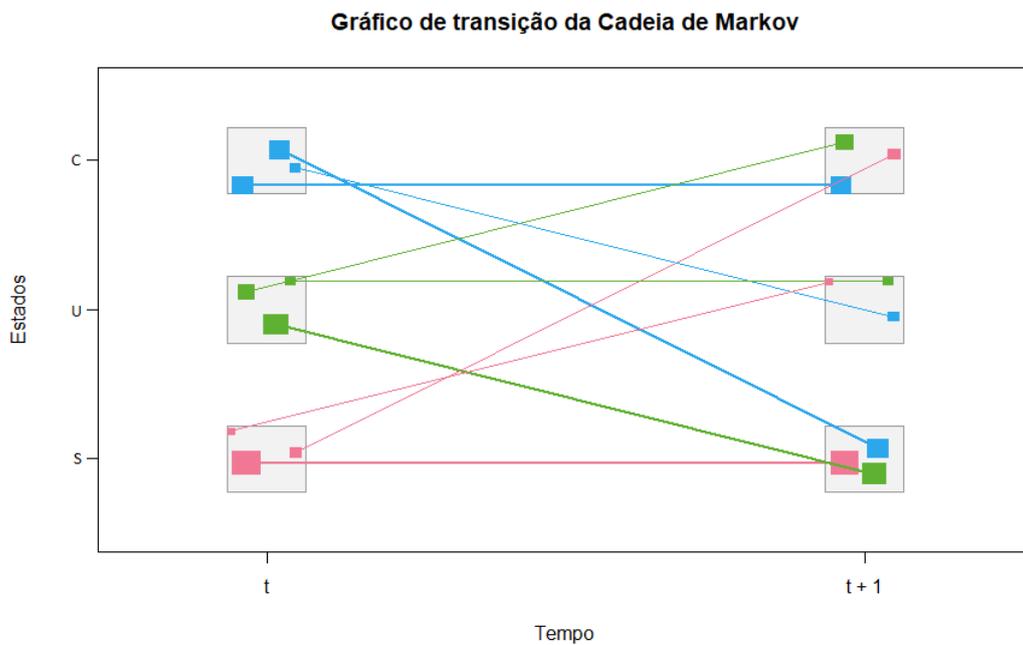


Figura 6 – Gráfico de transição da Cadeia de Markov de João Pessoa, PB

As Figuras 5 e 6 apresentam os gráficos representados em linhas coloridas as

probabilidade dado que dia anterior foi um dos três estados o próximo dia seja seco, úmido ou chuvoso. As cores das linhas foram determinadas da seguinte forma: seco recebeu a cor rosa, úmido a cor verde e chuvoso a cor azul. A partir da espessura da linha explica-se a ocorrência de um determinado evento ocorrer com uma probabilidade baixa ou alta. Uma linha muito fina reflete em uma probabilidade baixa e quanto mais espessa for a linha maior a probabilidade do evento acontecer, como pode ser observado na Figura 7 que se refere a cidade de Campina Grande dado que o dia foi seco a espessura da linha rosa é muito espessa ir para que o próximo dia seja seco, ou seja, a probabilidade desse evento é muito alta.

### 4.3 Cadeias de Markov de Alta Ordem

Tabela 7 – Critérios de Informação para João Pessoa

k	BIC	AIC	AICc	CAIC	KIC	KICc	AKICc	HQ
0	-37315.81	-3194.07	-5976.19	-32943.81	-7569.07	-10984.46	-10352.99	-14411.33
1	-39261.43	-5170.91	-7947.14	-34893.43	-9541.91	-12950.13	-12319.93	-16377.91
2	<b>-39451.98*</b>	-5455.12	-8213.71	<b>-35095.98*</b>	-9814.12	-13200.89	-12574.49	-16631.33
3	-39332.89	-5616.99	<b>-8323.09*</b>	-35012.89	<b>-9939.99*</b>	<b>-13262.92*</b>	<b>-12647.89*</b>	<b>-16700.83*</b>
4	-38519.86	<b>-5646.86*</b>	-8199.39	-34307.86	-9861.86	-12997.88	-12416.08	-16453.61
5	-35792.36	-5448.05	-7573.49	-31904.36	-9339.05	-11954.51	-11466.01	-15423.51
6	-27110.13	-4351.90	-5471.11	-24194.13	-7270.90	-8654.48	-8391.14	-11833.49

Tabela 8 – Critérios de Informação para Campina Grande

k	BIC	AIC	AICc	CAIC	KIC	KICc	AKICc	HQ
0	-38588.56	-5317.69	-8945.53	-34216.56	-9692.69	-14125.18	-13322.95	-16359.64
1	-39495.67	-6255.24	-9875.08	-35127.67	-10626.24	-15049.06	-14248.49	-17287.09
2	<b>-39513.44*</b>	-6364.33	<b>-9960.22*</b>	<b>-35157.44*</b>	<b>-10723.33*</b>	<b>-15117.23*</b>	<b>-14321.62*</b>	<b>-17365.87*</b>
3	-39267.99	-6392.84	-9917.52	-34947.99	-10715.84	-15023.74	-14242.89	-17303.46
4	-38451.97	<b>-6398.69*</b>	-9715.55	-34239.97	-10613.69	-14670.39	-13932.82	-17036.55
5	-35739.75	-6152.11	-8895.30	-31851.75	-10043.11	-13405.05	-12788.31	-15971.67
6	-27058.45	-4867.72	-6285.85	-24142.45	-7786.72	-9534.91	-9206.19	-12232.39

As Tabelas 7 e 8 indicam as escolhas para o  $k$  via Critérios de Informação para as cidades de João Pessoa e Campina Grande, respectivamente. É evidenciado que o  $k$  escolhido pelos Critérios de Informação que foram significativos estão destacados em vermelho. Na cidade Campina Grande o  $k = 2$ , em João Pessoa o  $k = 3$ , ou seja, a maioria desses oito critérios converge que o estado climático referente a precipitação do próximo dia é influenciado por dois dias anteriores, em Campina Grande e três dias anteriores em João Pessoa.

Tabela 9 – Matriz de transição de Alta Ordem para Campina Grande com  $k = 2$ 

	s	u	c
ss	0.88	0.04	0.08
su	0.68	0.1	0.22
sc	0.58	0.12	0.3
us	0.8	0.07	0.13
uu	0.65	0.09	0.26
uc	0.59	0.1	0.31
cs	0.76	0.08	0.16
cu	0.63	0.11	0.25
cc	0.53	0.14	0.33

A Tabela 9 apresenta que a aplicação do novo método Cadeia de Markov de Alta Ordem gerou essa matriz informando todos os possíveis casos sobre a premissa de dois dias anteriores, que foi o  $k$  escolhido pelos critérios de informação mostrado antes. Os valores ainda corroboram com os resultados de uma cadeia de Markov Clássica, mas indicam muito mais informação onde observa-se que dado todas as combinações tem-se uma maior chance para que o próximo dia seja seco.

Tabela 10 – Matriz de transição de Alta Ordem para João Pessoa com  $k = 3$ 

	s	u	c
sss	0.85	0.05	0.1
ssu	0.66	0.11	0.23
ssc	0.48	0.11	0.41
sus	0.75	0.08	0.18
suu	0.67	0.1	0.23
suc	0.41	0.13	0.46
scs	0.69	0.08	0.23
scu	0.55	0.13	0.32
scc	0.44	0.11	0.45
uss	0.76	0.08	0.16
usu	0.62	0.08	0.3
usc	0.43	0.09	0.48
uus	0.65	0.1	0.25
uuu	0.67	0.08	0.25
uuc	0.38	0.17	0.45
ucs	0.66	0.09	0.25
ucu	0.58	0.04	0.38
ucc	0.42	0.13	0.45
css	0.71	0.07	0.22
csu	0.49	0.1	0.42
csc	0.4	0.1	0.51
cus	0.67	0.08	0.25
cuu	0.49	0.06	0.45
cuc	0.45	0.11	0.44
ccs	0.63	0.09	0.28
ccu	0.52	0.1	0.38
ccc	0.4	0.11	0.49

A Tabela 10 apresenta a matriz de transição de Alta Ordem para a cidade de João Pessoa com  $k = 3$  é maior que a de Campina Grande com  $k = 2$ , refletindo várias possíveis combinações com três dias anteriores, ou seja, supondo que aconteceu três dias secos a chance que o dia subsequente seja seco extremamente alta em torno de 85%, já úmido é 0.5% e 1% para chuvoso. Portanto, se determinada combinação de dias anteriores muda para três dias chuvosos a chance que o próximo dia seja seco varia drasticamente passando para apenas 40% de chance que seja seco, 11% de ser úmido e 49% que seja chuvoso.

## 5 Conclusão

O presente trabalho visou averiguar se o clima de um dia é influenciado pelo clima dos dias anteriores, a partir disso verificou-se que o clima costuma mudar muito entre os três estados seco, úmido e chuvoso, e assim saber se vai ter um longo período de um ou de outro estado nas cidades de Campina Grande e João Pessoa utilizando Cadeias de Markov Clássica e de Alta Ordem. Os testes de adequação de modelos comprovaram estatisticamente que a Cadeia de Markov pode estimar a probabilidade de ocorrência de um dia de chuvoso, seco ou úmido baseando-se nos estados dos dias anteriores.

Os resultados sugerem que deve ser utilizado como melhor ordem  $k = 2$  para a cidade de Campina Grande-PB e um  $k=3$  para a cidade de João Pessoa-PB, ou seja, as duas cidades tem climas diferentes pois a cada 2 dias de determinada combinação de dias anteriores o clima muda entre os três estados possíveis, já em João Pessoa o clima muda após três dias de determinada combinação de dias anteriores.

O estado que chamou muita atenção foi a combinação da ocorrência de dias secos nas duas cidades, uma vez que se esse ciclo de dias secos acontece em João Pessoa-PB a chance que o próximo dia seja seco é alta, isto é, em torno de 85% , 5% de ser úmido, e conseqüentemente de ser chuvoso por volta de 10% apenas. Já em Campina Grande acontecendo dois dias secos consecutivos a chance que o próximo dia seja seco é extremamente alta, isto é, em torno de 88% , 4% de ser úmido, e conseqüentemente de ser chuvoso por volta de 8%, como se observar nos respectivos valores das probabilidades nas Tabelas 9 e 10. Entretanto, ainda é possível observar todas as possíveis combinações entre seco, úmido e chuvoso e partir disso avaliar a chance do próximo dia ser algum desses três climas.

Estes valores obtidos pelos dois tipos de processos estocástico, Cadeia de Markov Clássica e de Alta Ordem, corroboram com as características de um clima semiárido para a cidade de Campina Grande e de um clima tropical para a cidade de João Pessoa. O trabalho abriu um leque de investigações para mais cidades da Paraíba proporcionando uma compreensão mais abrangente da variabilidade climática via precipitação diária.

## Referências

- AKAIKE, H. A new look at the statistical model identification. *IEEE transactions on automatic control*, Ieee, v. 19, n. 6, p. 716–723, 1974. Citado na página 13.
- ALI, M.; MUBARAK, S. Effective rainfall calculation methods for field crops: An overview, analysis and new formulation. *Asian Research Journal of Agriculture*, Sciencedomain International, v. 7, n. 1, p. 1–12, jan 2017. Citado na página 16.
- AMORIM, J. F. de; SILVA, A. S. A. da; XIMENES, P. d. S. M. P.; MENEZES, R. S. C. Avaliação do inverse distance weighting para estimar a precipitação na mesorregião agreste do Estado de Pernambuco, Brasil. *Sigmae*, v. 8, n. 2, p. 340–347, 2019. Citado 2 vezes nas páginas 10 e 11.
- ASSIS, J. M. O. de; SOUZA, W. M. de; KOCH, H.; SOBRAL, M. d. C. M. Influência dos oceanos pacífico e atlântico tropicais sobre os índices climáticos da precipitação na bacia do submédio São Francisco. *Revista Brasileira de Climatologia*, v. 23, 2018. Citado na página 10.
- BICKEL, P. J.; RITOV, Y. et al. Inference in hidden markov models i: Local asymptotic normality in the stationary case. *Bernoulli*, Bernoulli Society for Mathematical Statistics and Probability, v. 2, n. 3, p. 199–228, 1996. Citado na página 12.
- BOZDOGAN, H. Icomp: A new model-selection criterion. In: *1. Conference of the international federation of classification societies*. [S.l.: s.n.], 1987. p. 599–608. Citado na página 13.
- CABRAL, W. B.; SOUZA, A.; ALEXANDRINO, E.; TORAL, F.; SANTOS, J.; CARVALHO, M. Características estruturais e agronômicas da *brachiaria brizantha cv. xaraés* submetida a doses de nitrogênio. *Revista Brasileira de Zootecnia*, v. 41, n. 4, p. 846–855, 2012. Citado na página 13.
- CHING, W.-K.; NG, M. K.; FUNG, E. S. Higher-order multivariate markov chains and their applications. *Linear Algebra and its Applications*, Elsevier, v. 428, n. 2-3, p. 492–507, 2008. Citado na página 14.
- FERREIRA, P. dos S.; GOMES, V. P.; GALVÍNCIO, J. D.; SANTOS, A. M. dos; SOUZA, W. M. de. Avaliação da tendência espaço-temporal da precipitação pluviométrica em uma região semiárida do Estado de Pernambuco. *Revista Brasileira de Climatologia*, v. 21, 2017. Citado na página 10.
- FU, D.; LI, Y.; HUANG, G. A fuzzy-markov-chain-based analysis method for reservoir operation. *Stochastic environmental research and risk assessment*, Springer, v. 26, n. 3, p. 375–391, 2012. Citado na página 11.
- GRIMM, A.; TEDESCHI, R. Influência de eventos El Niño e La Niña sobre a frequência de eventos extremos de precipitação no Brasil. In: *Congresso Brasileiro de Meteorologia*. [S.l.: s.n.], 2004. v. 13. Citado na página 10.

- HASTENRATH, S. *Climate dynamics of the tropics*. [S.l.]: Springer Science & Business Media, 2012. v. 8. Citado na página 10.
- JALE, J. da S.; JÚNIOR, S. F. A. X.; XAVIER, É. F. M.; STOŠIĆ, T.; STOŠIĆ, B.; FERREIRA, T. A. E. Application of markov chain on daily rainfall data in Paraíba-Brazil from 1995-2015. *Acta Scientiarum. Technology*, Editora da Universidade Estadual de Maringá-EDUEM, v. 41, p. e37186, 2019. Citado 2 vezes nas páginas 10 e 11.
- KONISHI, S.; KITAGAWA, G. *Information criteria and statistical modeling*. [S.l.]: Springer Science & Business Media, 2008. Citado na página 13.
- MAINTAINER, G. A. S.; SPEDICATO, G. A. Markovchain-package 3. *Package ‘markovchain’*, URL <http://github.com/spedygiorgio/markovchain/>, 2019. Citado na página 15.
- METEOROLOGIA, I.-I. N. de. Banco de dados meteorológicos para ensino e pesquisa. *BDMEP*, 2017. Citado na página 16.
- MICHE, Y.; LENDASSE, A. A faster model selection criterion for op-elm and op-knn: Hannan-quinn criterion. In: *ESANN*. [S.l.: s.n.], 2009. v. 9, p. 177–182. Citado na página 14.
- NÓBREGA, V. *Tópicos em Composição: Estrutura, Formação e Acento (Topics in Compounding: Structure, Formation and Stress)*. Tese (Doutorado) — MA. Thesis, University of São Paulo, São Paulo, 2014. Citado na página 10.
- RATAN, R.; VENUGOPAL, V. Wet and dry spell characteristics of global tropical rainfall. *Water Resources Research*, Wiley Online Library, v. 49, n. 6, p. 3830–3841, 2013. Citado na página 16.
- RODRIGUES, L. de O.; SOUZA, W. M. de; COSTA, V. S. de O.; PEREIRA, M. L. T. Influência dos eventos de El Niño e La Niña no regime de precipitação do agreste de Pernambuco (influence of El Niño and La Niña events on rainfall of agreste Pernambuco). *Revista Brasileira de Geografia Física*, v. 10, n. 6, p. 1995–2009, 2017. Citado na página 10.
- SANTOS, A. V.; DIMURO, G. P.; BARBOZA, L. V.; COSTA, A. C. d. R.; REISER, R. H. S.; CAMPOS, M. A. Probabilidades intervalares em modelos ocultos de Markov. *Trends in Applied and Computational Mathematics*, v. 7, n. 2, p. 361–370, 2006. Citado na página 12.
- SANTOS, F. S. dos; NASCIMENTO, K. K. F. do; JALE, J. da S.; FERREIRA, T. A. E. Análise de velocidade do vento em petrolina-pe usando Cadeias de Markov de ordem superior. *Sigmas*, v. 8, n. 2, p. 385–392, 2019. Citado na página 10.
- SANTOS, S. R. Q. D.; CUNHA, A. P. M. do A.; RIBEIRO-NETO, G. G. Avaliação de dados de precipitação para o monitoramento do padrão espaço-temporal da seca no Nordeste do Brasil. *Revista Brasileira de Climatologia*, v. 25, 2019. Citado na página 10.
- SARAH, I. A. *Kullback’s Information Criterion Correction (KICC) Untuk Seleksi Model Regresi Linear Multivariat*. Tese (Doutorado) — Universitas Gadjah Mada, 2015. Citado na página 14.

- SCHOLZ, F. W.; STEPHENS, M. A. K-sample Anderson–Darling tests. *Journal of the American Statistical Association*, Taylor & Francis Group, v. 82, n. 399, p. 918–924, 1987. Citado na página 12.
- SCHWARZ, G. et al. Estimating the dimension of a model. *The annals of statistics*, Institute of Mathematical Statistics, v. 6, n. 2, p. 461–464, 1978. Citado na página 13.
- SILVA, B. C. D.; CLARKE, R. T. Análise estatística de chuvas intensas na bacia do rio São Francisco. *Revista Brasileira de Meteorologia*, v. 19, n. 3, p. 265–272, 2004. Citado na página 10.
- SOBRAL, T. L.; BARRETO, G. Utilização dos critérios de informação na seleção de modelos de regressão linear. *Proceeding Series of the Brazilian Society of Computational and Applied Mathematics*, v. 4, n. 1, 2016. Citado na página 13.
- SPEDICATO, G. A.; KANG, T. S.; YALAMANCHI, S. B.; YADAV, D.; CORDÓN, I. The markovchain package: a package for easily handling discrete markov chains in r. *Accessed Dec*, 2016. Citado na página 16.
- TEAM, R. C. R: A language and environment for statistical computing. r foundation for statistical computing, Vienna, Austria. 2012. URL <http://www.R-project.org>, 2018. Citado na página 16.
- XAVIER, S. F. A.; JALE, J. da S.; STOSIC, T.; SANTOS, C. A. C. dos; SINGH, V. P. An Application of Sample Entropy to Precipitation in Paraíba State, Brazil. *Theoretical and Applied Climatology*, Springer Science and Business Media LLC, v. 136, n. 1-2, p. 429–440, may 2018. Citado na página 10.
- YAMANOUCHI, T. K. *Seleção de modelos lineares mistos utilizando critérios de informação*. Tese (Doutorado) — Universidade de São Paulo, 2017. Citado na página 14.