



UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO ACADÊMICO DO AGRESTE
NÚCLEO DE TECNOLOGIA
CURSO DE GRADUAÇÃO EM ENGENHARIA CIVIL

GABRIEL MAGNO CAVALCANTE CALADO

**ANÁLISE COMPARATIVA ENTRE DISTRIBUIÇÕES DE PROBABILIDADES
PARA GERAÇÃO DE SÉRIES SINTÉTICAS DE PRECIPITAÇÕES DIÁRIAS EM
CIDADES DO AGRESTE DE PERNAMBUCO**

Caruaru

2021

GABRIEL MAGNO CAVALCANTE CALADO

**ANÁLISE COMPARATIVA ENTRE DISTRIBUIÇÕES DE PROBABILIDADES
PARA GERAÇÃO DE SÉRIES SINTÉTICAS DE PRECIPITAÇÕES DIÁRIAS EM
CIDADES DO AGRESTE DE PERNAMBUCO**

Trabalho de Conclusão de Curso apresentado à Coordenação do Curso de Engenharia Civil do Campus Agreste da Universidade Federal de Pernambuco – UFPE, na modalidade de artigo científico, como requisito parcial para obtenção do grau de Bacharel em Engenharia Civil.

Área de concentração: Recursos Hídricos

Orientadora: Prof. Dr. Artur Paiva Coutinho

Caruaru

2021

Dedico este trabalho à minha mãe, Maria de Fátima, por ensinar e compartilhar comigo o mais belo sentimento que existe, o amor.

AGRADECIMENTOS

Agradeço a Deus pelo dom da vida e por me proporcionar força, saúde, sabedoria e paciência para todos os dias buscar a realização dos meus sonhos.

A minha mãe Maria de Fátima, por toda a dedicação, carinho e incentivo que tem me dado ao longo de todos esses anos.

Ao meu pai Gerson, que de sua forma me ensinou que grandes virtudes da vida são a honestidade e o prazer no trabalho.

Aos meus irmãos Gedson e Gersyane, pelo convívio, conselhos e todo o apoio que tem demonstrado na realização dos meus anseios.

Aos amigos Ananda Barbosa, Maria Renata, Tallys Celso e Wagner Leal, em especial, pelo companheirismo, conselhos, incentivos e por compartilharem comigo os bons momentos e, também, aqueles difíceis ao longo da graduação.

Ao meu amigo Maciel Tavares e minha professora do ensino médio Lucilene Lima, pela confiança e por me apresentarem os inúmeros caminhos e possibilidades que a pesquisa científica pode proporcionar a um estudante.

Aos amigos e familiares de menor convivência, pelos bons momentos, conversas e ajuda nessa longa jornada.

Ao meu orientador Artur Coutinho, pela oportunidade de ser seu aluno de IC e TCC, pelos conselhos, ensinamentos e por me apresentar as diversas possibilidades de atuação existentes na área de recursos hídricos.

A Mayke Feitosa e Daniel Detzel, autores que tomei como base para a realização deste trabalho, por toda a atenção aos meus e-mails com dúvidas e disposição em ajudar no que foi preciso.

Aos membros da banca, pela disposição em participar deste trabalho contribuindo para sua melhoria.

A todos aqueles que de forma direta ou indiretamente contribuíram para minha formação e me ajudaram a distinguir um bom de um mal profissional.

Combati o bom combate, terminei a corrida, guardei a fé. Agora me está reservada a coroa da justiça, que o Senhor, justo Juiz, me dará naquele dia; e não somente a mim, (...)

2 Timóteo 4:7-8

Análise comparativa entre distribuições de probabilidades para geração de séries sintéticas de precipitações diárias em cidades do Agreste de Pernambuco.

Comparative analysis between probability distributions for generation of synthetic daily precipitation series in cities in the Agreste region of Pernambuco.

Gabriel Magno Cavalcante Calado¹

RESUMO

O acesso aos recursos hídricos é um fator indispensável no desenvolvimento de uma região. Assim, o conhecimento de variáveis climatológicas relacionadas às precipitações é essencial para o desenvolvimento de obras de infraestrutura hídrica para garantir a produção e o abastecimento da população durante todo o ano. As séries sintéticas figuram-se como aliadas para estudos de implantação dessas obras, quando os dados históricos de precipitações são curtos e/ou apresentam muitas falhas. Diante disso, o objetivo deste trabalho é montar procedimentos de geração das séries sintéticas de precipitações diárias para cidades do agreste de Pernambuco, avaliando qual e o que melhor se adapta. Nesse sentido, são utilizados processos estocásticos markovianos associados às distribuições de probabilidade Gama e Exponencial Mista que incorporam conceitos do método de Monte Carlo. Aplicou-se o modelo em estações pluviométricas nas cidades de Alagoinha, Canhotinho e Limoeiro. Após a análise da ordem das cadeias, pode-se generalizar a primeira para o conjunto de estudo. Pelo teste de KS a maioria das diferenças encontradas ficaram abaixo do valor crítico e através do cálculo do RMSE e MAE obteve-se bons resultados, com valores muito próximo de zero (ideal), com destaque para a distribuição exponencial, com desempenho superior na maioria dos casos analisados. Na comparação das séries geradas com os dados observados, em relação às médias, desvios e extensão dos períodos chuvosos, foram atingidos resultados próximos do esperado. O modelo teve baixo desempenho somente na geração das chuvas máximas, apresentando grandes diferenças nos meses em que ocorreram as maiores precipitações e tendências de subestimativa pela função Gama, que para sua confirmação, bem como a generalização do melhor método para a região dever-se-ia aumentar o conjunto de análise.

¹ Graduando em Engenharia Civil pela Universidade Federal de Pernambuco. E-mail: gabrielm2c@hotmail.com

Palavras-chave: cadeias de Markov; distribuição de probabilidades; Método de Monte Carlo.

ABSTRACT

Access to water resources is an indispensable factor in the development of a region. Thus, knowledge of climatological variables related to rainfall are essential for the development of water infrastructure works to ensure the production and supply of the population throughout the year. The synthetic series appear as allies for studies on the implementation of these works, when the rainfall historical series is short and/or has many flaws. Therefore, the objective of this work is to set up procedures for the generation of synthetic series of daily precipitations to cities in the agreste region of Pernambuco, evaluating which is the best fit. In this sense, Markovian stochastic processes are used associated with the Gamma and Mixed Exponential probability distributions that incorporate concepts of the Monte Carlo Method. The model was applied in rainfall stations in the cities of Alagoinha, Canhotinho and Limoeiro. After an analysis of the order of the chains, the first could be generalized to the study set. By KS test, most of the differences found were below the critical value and by calculating of the RMSE and MAE, good results were obtained, with values very close to zero (ideal), with the emphasis on the Exponential distribution, with superior performance in most cases analyzed. Comparing the generated series with the observed data, in relation to the means, deviations and extension of the rainy periods, results close to the expected were achieved. The model had poor performance only in generating maximum rainfall, showing large differences in the months when the highest rainfall occurs and underestimation trends by the Gamma function, which for its confirmation, as well as the generalization of the best method for the region, the analysis set should be increased.

Keywords: Markov chain; probability distributions; Monte Carlo Method.

DATA DE APROVAÇÃO: 31 de agosto de 2021.

1 INTRODUÇÃO

A evolução humana sempre esteve vinculada ao uso dos recursos naturais existentes disponíveis. Apesar disso, estudos climáticos indicam que as ações antrópicas têm grande responsabilidade pela intensificação do aquecimento global nos últimos tempos. De acordo com Cirilo (2008), na região semiárida, isso sinaliza para diminuição da oferta hídrica com elevação da evaporação nos corpos d'água e diminuição do volume neles escoado, redução da recarga dos aquíferos e da realimentação da vazão dos rios, minoração e concentração da precipitação em ainda menor espaço de tempo. No entanto, de acordo com Conti e Schroeder (2013), estima-se que 90% das águas precipitadas nessa região não são aproveitadas. Isso, torna o conhecimento de variáveis climatológicas essencial nessa região.

Os parâmetros das chuvas, segundo Kaiser (2006), quando utilizados nos modelos de transformação chuva-vazão permitem estimar o deflúvio em um sistema de drenagem. Eles ainda podem ser utilizados em obras de contenção, reservatórios e outros. Construções desse tipo são úteis em situações críticas, pois podem minimizar prejuízos econômicos e sociais à população, como em eventos de cheias, e, do mesmo modo, em épocas de escassez de chuva, acumular água garantindo a segurança hídrica para produção e o consumo humano.

Os estudos relativos à caracterização das precipitações são normalmente desenvolvidos a partir de séries de dados históricos. Esses por sua vez, são coletados de uma extensa rede de pluviógrafos e pluviômetros distribuídos em todo o território brasileiro. Conforme Detzel (2009), em geral, análises estatísticas alusivas às precipitações encontram limitações relacionadas ao tamanho das séries disponíveis, que em sua maioria são insuficientes. Uma alternativa para superar essas limitações é a geração sintética de precipitações, baseada em propriedades determinísticas e estocásticas, aplicadas a várias escalas de tempo.

Os modelos determinísticos preveem, na maioria dos casos, resultados únicos para um conjunto de circunstâncias de determinada problemática. Porém, principalmente nas situações em que existe a presença de incertezas, esse tipo de modelo não consegue auxiliar na busca pelos resultados requeridos devido à quantidade de variáveis e recursos que interferem na análise. Nessas condições, podem ser utilizados os modelos probabilísticos ou estocásticos, em que são considerados, além dos dados de entrada, fatores aleatórios levando a resultados estimados das características pretendidas do sistema estudado.

Os processos estocásticos desempenham um papel importante na elucidação de muitas áreas das ciências naturais e da engenharia (KARLIN & TAYLOR, 1998). Nesse contexto, se

encaixam os processos hidrológicos que são definidos, de acordo com Aparício (1989) apud Laura (1997), como aqueles gerados por fenômenos físicos de extrema complexidade que, geralmente, impossibilitam uma estimativa confiável de sua magnitude com base em leis físicas que resultam em modelos extremamente complexos e de difícil controle. Segundo WMO – World Meteorological Organization (2009), esses eventos climatológicos são caracterizados por uma grande variabilidade, aleatoriedade e incerteza quanto à sua ocorrência e dimensão, especialmente os eventos considerados extremos, como secas, enchentes e tufões. Portanto, configuram-se como claros exemplos de processos em que os modelos probabilísticos têm grande utilidade.

Os procedimentos de obtenção de chuvas sintéticas, são capazes, a partir de modelos probabilísticos, de determinar, com bastante precisão, uma série histórica de chuvas para determinadas localidades. Comumente, eles são classificados em não paramétricos, semiparamétricos e paramétricos, de acordo com as características do processo utilizado. Em tal caso, os paramétricos, segundo Detzel (2009), são os mais utilizados devido à facilidade de ajuste baseada em dois pontos, o primeiro deles é a ocorrência da precipitação, geralmente obtida por meio de processos estocásticos markovianos ou de alternância de eventos, e o segundo é a lâmina precipitada, obtida por meio de distribuições estatísticas. De acordo com Wilks (1998), uma série sintética, nesse sentido, seria o resultado da associação de séries de dados reais com números aleatórios produzidos por algoritmos computacionais, a fim de gerar sequências de variáveis aleatórias que se assemelham aos dados climáticos reais.

Os processos markovianos, portanto, auxiliam na determinação da ocorrência ou não de precipitação em um dia requerido. Essa probabilidade pode ser baseada apenas no fato de o dia atual de análise seja seco ou chuvoso, caracterizando assim uma cadeia de primeira ordem, ou ainda depender de uma observação mais profunda englobando os dias anteriores, definindo as ordens superiores. Na maioria dos estudos, a primeira ordem satisfaz bem as condições gerando bons resultados, mas em regiões específicas onde existe, por exemplo, muitos dias secos ou chuvosos consecutivos, necessita-se de uma análise mais detalhada como descrita no estudo de Azevedo e Leitão (1990), para o estado da Paraíba.

Richardson (1981), em seu trabalho, aconselha que sejam utilizadas as distribuições de probabilidade com maiores números de parâmetros, a exemplo da distribuição Gama e da Exponencial Mista. Em sua pesquisa, Progênio (2019) reúne os principais resultados de diversos autores a respeito da utilização de diferentes metodologias de cálculo, dentre elas, a Exponencial Mista destaca-se nos trabalhos de Wan et al. (2005) e no de Kits e Virgens Filho

(2015), enquanto a distribuição Gama distingue-se no de Baú et al. (2013) e Dash (2012). Segundo Wilks (2006), elas destacam-se pelo fato de alguns mecanismos físicos apresentarem, em geral, mais de um processo gerador, fazendo com que uma distribuição estatística simples seja incompleta. Logo, as distribuições mistas, por exemplo, apresentam um bom grau de flexibilidade, refletindo diretamente nos resultados a serem obtidos.

Em geral, a determinação de chuvas sintéticas permite, diante das muitas limitações das séries históricas, a obtenção de parâmetros de caracterização climática das precipitações. Usualmente, os resultados desses modelos são bastante válidos, indicando que podem ser utilizados com segurança no desenvolvimento de estudos de implantação de obras hídricas. Isso posto, é de grande interesse para a formação do engenheiro civil o conhecimento desse tipo de metodologia, pois pode ser capaz de contornar situações em sua carreira em que não há dados suficientes para embasar o desenvolvimento e execução de suas obras.

Dessa forma, o objetivo deste trabalho é estudar os modelos paramétricos para geração de séries sintéticas de precipitações diárias e replicar em áreas específicas do Agreste de Pernambuco. Busca-se ainda determinar a melhor ordem da cadeia de Markov para a localidade de estudo, além de reproduzir e automatizar o modelo proposto por meio de programação computacional, aplicar o modelo desenvolvido, reunir, avaliar, comparar e apresentar os resultados obtidos no procedimento.

2 METODOLOGIA

2.1 Caracterização da Área de Estudo

A mesorregião do Agreste de Pernambuco apresenta extensão territorial de 24.549,5 km². É composta por 71 municípios distribuídos em 3 microrregiões conhecidas como Agreste meridional, central e setentrional. Em sua maior extensão, caracteriza-se por apresentar o bioma da caatinga, clima semiárido e solos rasos. Devido a estes fatores, apresenta períodos chuvosos muito curtos e longos de estiagem, o que torna a infraestrutura de armazenamento de água de essencial importância para a população daquela região.

Para coleta dos dados foi consultado a base da APAC – Agência Pernambucana de Águas e Clima e da ANA – Agência Nacional de Águas e Saneamento Básico. Os critérios adotados para a escolha das localidades basearam-se na quantidade de anos disponíveis das séries históricas de precipitações, menor quantidade de falhas nos dados, para efeito comparativo

com as séries sintéticas geradas e posição estratégica. Diante desses fatores, os postos mais bem avaliados e consequentemente escolhidos para a análise são mostrados na Tabela 1.

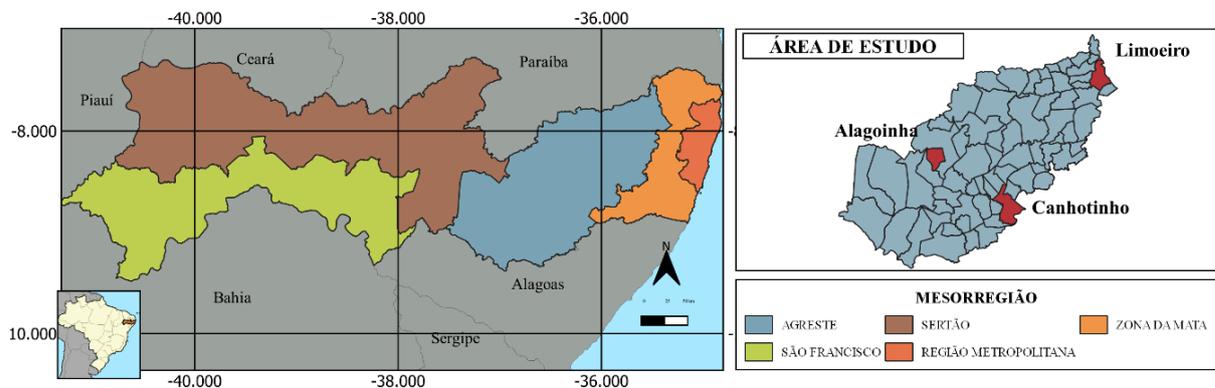
Tabela 1 – Cidades e Respectivos postos Analisados

Cidade	Posto	Latitude	Longitude	Período	Extensão	Operação
Alagoinha	140	-9,07	-36,77	1963-1996	34 anos	APAC
Canhotinho	214	-8,92	-36,12	1964-1991	28 anos	APAC
Limoeiro	735100	-7,87	-35,45	1979-2018	40 anos	ANA

Fonte: Autor (2021)

As cidades selecionadas representam cada uma das microrregiões que compõem o Agreste, sendo Alagoinha do Central, Canhotinho do Meridional e Limoeiro do Setentrional. Além de disso, elas também têm localizações interessantes com a transição do Agreste para o Sertão, influência da altitude, e a transição da Zona da Mata para o Agreste, respectivamente. Uma análise com essa conjuntura é a observação da ocorrência de variação da ordem da cadeia de Markov a ser considerada no processo à medida que a posição geográfica e clima modificam.

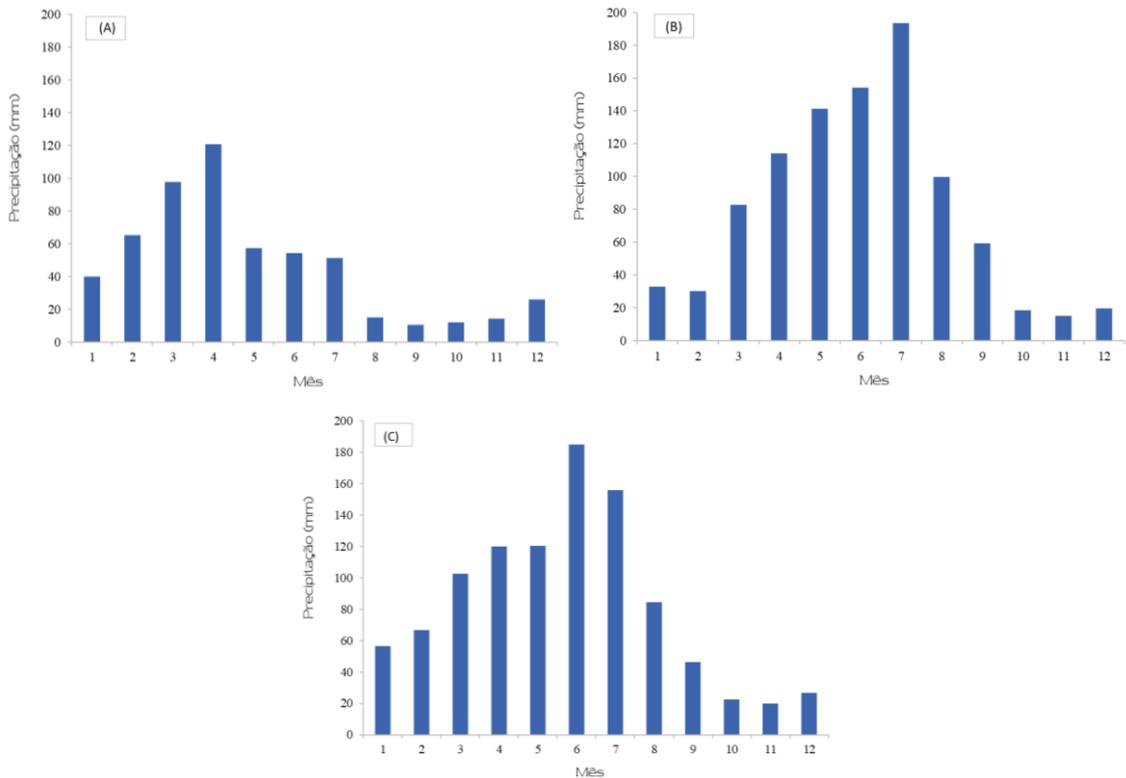
Figura 1 – Divisão regional de Pernambuco e Área de Estudo



Fonte: Autor (2021)

Na Figura 2 é mostrado o comportamento das precipitações médias mensais ao longo do ano nas localidades de estudo no período indicado na Tabela 1.

Figura 2 – Precipitações Médias Mensais (A) Alagoíinha, (B) Canhotinho e (C) Limoeiro

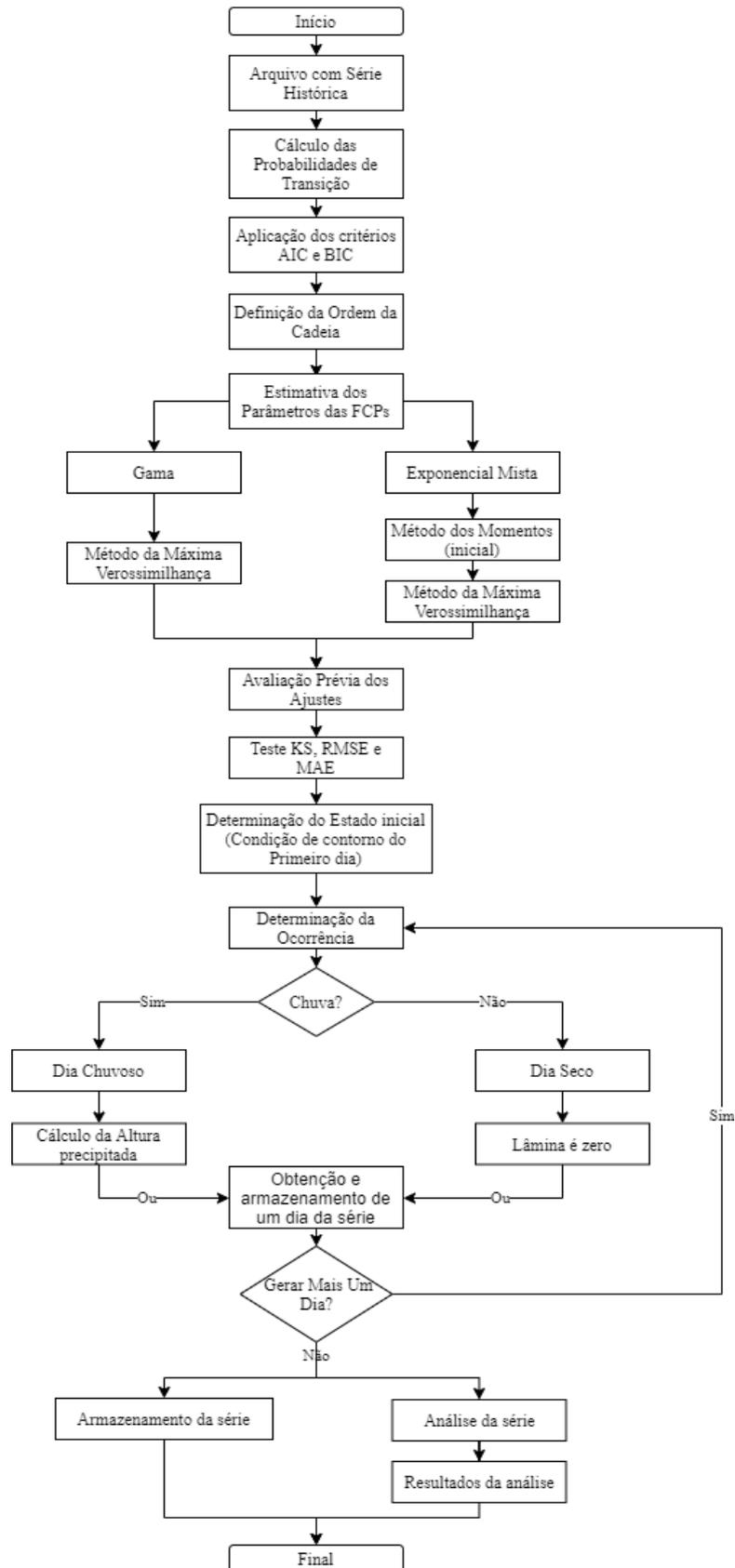


Fonte: Autor (2021)

2.2 Esquematização do Procedimento

O processo de geração das séries sintéticas apresenta um certo grau de complexidade. Como proposto pela bibliografia, ele é dividido nas etapas de definição da ocorrência e determinação das lâminas, ambas apresentam diversos passos para a progressão complementando-se no procedimento. Além disso, devem ser implementadas ferramentas para análise de consistência e qualidade dos resultados. Ainda assim, para o procedimento pode ser feita uma representação esquemática conforme mostrado na Figura 3. Ela facilita o entendimento da sistemática utilizada para geração das séries sintéticas requeridas. É importante salientar que todo o procedimento é automatizado a partir de algoritmos de programação desenvolvidos com o auxílio do software MATLAB®.

Figura 3 – Fluxograma do Procedimento



Fonte: Adaptado de Detzel (2009)

2.3 Determinação da Ordem da Cadeia de Markov

De acordo com Detzel (2009), a propriedade fundamental de um processo estocástico markoviano é que a evolução futura do sistema depende somente do estado atual e nunca dos estados passados e, dessa forma, o sistema não tem “memória” que lhe possibilite usar informações quanto ao seu comportamento atual. Porém, a ocorrência das precipitações é um fenômeno que de forma muito improvável não dependerá de um acontecimento anterior. A quantificação dessa dependência não pode ser generalizada em decorrência da constante modificação dos fenômenos atmosféricos ao longo das estações do ano, localização geográfica e ciclo hidrológico local, influenciados pelas características morfoclimáticas.

Apesar desse impasse, os processos markovianos, quando os parâmetros de estado e de tempo assumem intervalos discretos, são capazes de adquirir propriedades de cadeia. Diante desse fato, as variáveis podem ser referenciadas no tempo e, em tal caso, a ordem da cadeia passa a representar o quão influente é a ocorrência do fenômeno no tempo passado na previsão da variável requerida. Para a chuva, fenômeno de interesse deste trabalho, isso pode ser entendido como o espaço de tempo anterior e a sua caracterização, em seco ou chuvoso, influencia na determinação de ocorrência ou não de precipitação em um dia analisado.

Para a determinação da ordem ótima do modelo da cadeia de Markov, em uma localidade, pode-se utilizar o Critério de Informação Akaike (AIC) (AKAIKE, 1974) e o Critério de Informação de Bayes (BIC) (SCHWARZ, 1978). Ambos são baseados nas funções de verossimilhança para as probabilidades de transição dos modelos da cadeia ajustados. Em seu trabalho, Wilks (2006), indica essas funções com s estados e de ordem zero, um e dois para as cadeias de Markov, expandindo ainda para a terceira ordem:

$$L_0 = \sum_{j=0}^{s-1} N_j(k) \cdot \ln[p_j(k)] \quad (1)$$

$$L_1 = \sum_{i=0}^{s-1} \sum_{j=0}^{s-1} N_{ij}(k) \cdot \ln[p_{ij}(k)] \quad (2)$$

$$L_2 = \sum_{h=0}^{s-1} \sum_{i=0}^{s-1} \sum_{j=0}^{s-1} N_{hij}(k) \cdot \ln[p_{hij}(k)] \quad (3)$$

$$L_3 = \sum_{g=0}^{s-1} \sum_{h=0}^{s-1} \sum_{i=0}^{s-1} \sum_{j=0}^{s-1} N_{ghij}(k) \cdot \ln[p_{ghij}(k)] \quad (4)$$

Em que: N é o número de dias da série histórica, com as transições g, h, i e j , e P é a probabilidade de transição.

Os índices g, h, i e j , utilizados nas equações de 1 a 4, podem assumir duas configurações (estados), sendo elas “0” para dias secos e “1” para dias chuvosos. Assim, de acordo com a ordem e a posição em que aparecem indicam o estado do dia analisado e dos precedentes. Para tal, mostra-se indispensável o estabelecimento da lâmina mínima para o dia ser considerado chuvoso. Em face do exposto, é importante a ponderação entre o clima da região, bem como a menor unidade medida utilizada pelo pluviômetro (mm) e referências externas, como os trabalhos de Moreno-Pérez et al. (2014), Santos (2017) e de Zakaria e Deni (2016), diante disso indicou-se o valor de 0,1 mm como o limiar dessa caracterização. A partir dessa convenção, pode-se indicar as configurações para o número de dias, N , sejam eles para ordem zero, um, dois ou três como mostrado na Tabela 2.

Tabela 2 – Configurações dos Dias para a Cadeia de Markov

ORDEM ZERO											
		Anterior									
Atual											
				-							
S		S									
Representação		N ₀									
C		C									
Representação		N ₁									
PRIMEIRA ORDEM											
		Anterior		S	C						
Atual				S	C						
S		SS		SS	SC						
Representação		N ₀₀		N ₀₁	N ₀₁						
C		CS		CC	CC						
Representação		N ₁₀		N ₁₁	N ₁₁						
SEGUNDA ORDEM											
		Anterior		SSS	SSC	SCS	SCC				
Atual				SSS	SSC	SCS	SCC				
S		SSS		SSC	SCS	SCC	SCC				
Representação		N ₀₀₀		N ₀₀₁	N ₀₁₀	N ₀₁₁	N ₀₁₁				
C		CSS		CSC	CCS	CCC	CCC				
Representação		N ₁₀₀		N ₀₁₁	N ₁₀₁	N ₁₀₁	N ₁₁₁				
TERCEIRA ORDEM											
		Anterior		SSS	SSC	SCS	CSS	SCC	CSC	CCS	CCC
Atual				SSS	SSC	SCS	CSS	SCC	CSC	CCS	CCC
S		SSSS		SSSC	SSCS	SCSS	SSCC	SCSC	SCCS	SCCC	SCCC
Representação		N ₀₀₀₀		N ₀₀₀₁	N ₀₀₁₀	N ₀₁₀₀	N ₀₀₁₁	N ₀₁₀₁	N ₀₁₁₀	N ₀₁₁₁	N ₀₁₁₁
C		CSSS		CSSC	CSCS	CCSS	CSCC	CCSC	CCCS	CCCC	CCCC
Representação		N ₁₀₀₀		N ₁₀₀₁	N ₁₀₁₀	N ₁₁₀₀	N ₁₀₁₁	N ₁₁₀₁	N ₁₁₁₀	N ₁₁₁₁	N ₁₁₁₁

Fonte: Adaptado de Pro gênio (2019)

Assim, quaisquer configurações de dias descritas na Tabela 2 podem ser lidas como no

exemplo que se segue:

- N_{0011} é o número de dias secos antecedidos por dia seco, chuvoso e chuvoso, respectivamente.

Já as probabilidades de transição, calculadas para a função de verossimilhança, apresentam formulações que seguem, sempre, uma mesma linha de construção. Comumente, elas diferem apenas no que diz respeito aos fatores do número de dias, relativo à ordem e à sequência dos seus estados. Desse modo, o cálculo dessas probabilidades pode ser exemplificado pelas expressões 5, 6, 7 e 8, respectivamente.

$$p_0 = \frac{N_0}{N} \quad (5)$$

$$p_{00} = \frac{N_{00}}{N_{00} + N_{10}} \quad (6)$$

$$p_{100} = \frac{N_{100}}{N_{100} + N_{000}} \quad (7)$$

$$p_{1010} = \frac{N_{1010}(k)}{N_{0010} + N_{1010}} \quad (8)$$

Assim, observa-se que o padrão de construção para a obtenção dessas probabilidades é:

$$P_{(m)(n)} = \frac{N_{(m)(n)}(k)}{N_{0(n)}(k) + N_{1(n)}(k)} \quad (9)$$

Em que: p é probabilidade calculada, m é a ocorrência de chuva (“1”) ou dia seco (“0”) e o n tem dimensão de acordo com a ordem representada como uma sequência de 0’s e 1’s (secos ou chuvosos) e deve sempre apresentar a mesma sequência dos dias precedentes para todos os fatores da situação analisada.

Com as expressões mostradas, obtêm-se os parâmetros necessários para a determinação da ordem da cadeia de Markov. Tanto o modelo AIC como o modelo BIC são baseados no princípio da parcimônia, caracterizado por incorporar uma penalidade proporcional ao aumento da cadeia analisada. Assim, ambos privilegiam ordens mais simplistas devido à menor quantidade de parâmetros a serem estimados, sendo o modelo mais adequado aquele que produz menores valores de AIC e BIC. Para o critério Akaike, a expressão de cálculo é:

$$AIC(k) = -2 * L_m + 2 * s^m (s-1) \quad (10)$$

Já para o critério Bayes, a formulação é a seguinte:

$$\text{BIC}(k) = -2 * L_m + s^m (\ln n) \quad (11)$$

Em que: para ambos os métodos, m é a ordem da cadeia analisada, o valor de L correspondente é dado pelas expressões 1, 2, 3 e 4, de acordo com a ordem, s é o número de estados, que neste trabalho são dois (seco e chuvoso), e n é o tamanho da amostra.

Apesar de serem modelos bastante semelhantes, divergindo apenas na segunda parcela, correspondente à penalidade, segundo Katz (1981), o critério AIC tende a superestimar a ordem da cadeia a ser empregada. De acordo com o mesmo autor, o critério BIC consegue proporcionar resultados mais consistentes. Diante disso, para este trabalho, quando os resultados destes testes divergirem, sendo o AIC para mais, será sempre utilizado o BIC.

2.3.1 Determinação da Ocorrência da Precipitação

A ocorrência ou não da chuva baseia-se nos conceitos do método de Monte Carlo. A partir da ordem ótima da cadeia determinada são coletadas as probabilidades de ocorrência que, por sua vez, caracterizam-se como elementos básicos, indicando o acontecimento ou não da chuva. O procedimento depende, então, da quantidade de dias anteriores e de suas respectivas configurações, eles são relacionados a uma probabilidade crítica, que de acordo com a conformação, assume uma daquelas calculadas e a ocorrência é determinada por comparação direta dela com o número randômico do método utilizado.

A nível de exemplo, para a primeira ordem, tem-se que a probabilidade crítica pode assumir o valor de p_{11} ou p_{10} . Assim, é utilizado o recurso de geração de um número aleatório, u , no intervalo $[0, 1]$, que é comparado às probabilidades críticas, caso o dia anterior tenha sido seco, compara-se com p_{10} , se u for maior, logo o dia é considerado seco, caso contrário, será chuvoso. Já se o dia anterior for chuvoso, deve-se comparar com p_{11} seguindo a mesma linha de pensamento descrita. Dessa forma, se o dia for chuvoso, ele assumirá o valor dado pelas equações de distribuição de probabilidades, sendo esta a segunda parte do processo, caso não, será atribuído o valor de zero e passa-se a determinação e um novo dia da série.

2.4 Obtenção das Lâminas de Chuva

Para determinar a estrutura do gerador climático, utilizado para a obtenção das quantidades de chuva precipitadas, são empregados os modelos paramétricos. Esse tipo de metodologia apresenta boa precisão e facilidade de ajuste, embora não proporcione um bom nível de detalhamento do processo envolvido. Existem diferentes distribuições probabilísticas utilizadas para determinar a altura precipitada, dentre elas, pode-se destacar a distribuições do tipo Weibull, tipo Gama e tipo Exponencial. Para o uso neste estudo foram delimitadas a Gama à dois parâmetros e a Exponencial Mista, devido a maior quantidade de parâmetros, como indicado por Richardson (1981) e Wilks (2006).

2.4.1 Distribuição Exponencial Mista

A função distribuição de probabilidades exponencial mista é representada por:

$$f_x(x) = \frac{\alpha}{\beta_1} \exp\left(\frac{-x}{\beta_1}\right) + \frac{(1-\alpha)}{\beta_2} \exp\left(\frac{-x}{\beta_2}\right); \quad (12)$$

$x > 0; 0 \leq \alpha \leq 1; \beta_1 \geq \beta_2 > 0$

Sua função de probabilidades acumuladas é obtida a partir da integração analítica da equação 12. Ela é resolvida para x como uma função de $F(x)$, e representada pela equação 13.

$$F_x(x) = \int_0^x f_x(x) dx = 1 - \alpha \cdot \exp\left(\frac{-x}{\beta_1}\right) - (1-\alpha) \cdot \exp\left(\frac{-x}{\beta_2}\right); \quad (13)$$

Em que: x representa o valor assumido pela precipitação, α é o fator de probabilidade de proporção da mistura de exponenciais, para a localidade, que determina as frequências de como o parâmetro da distribuição com maior média (β_1) e com menor média (β_2) são combinados para geração do próximo valor no x .

A determinação dos valores de β_1 , β_2 e α é baseada no método da máxima verossimilhança. O procedimento para obtenção desses parâmetros, em maiores detalhes, pode ser visto em trabalhos mais aprofundados como o de Detzel (2009). Em síntese, consiste na maximização do logaritmo natural aplicado na verossimilhança da equação 12, representado por:

$$\ln L(x, \alpha, \beta_1, \beta_2) = \sum_{i=1}^n \ln \left[\frac{\alpha}{\beta_1} \exp\left(\frac{-x}{\beta_1}\right) + \frac{(1-\alpha)}{\beta_2} \exp\left(\frac{-x}{\beta_2}\right) \right] \quad (14)$$

Diante da não viabilidade da derivação desta equação, pode-se obter a maximização a partir do cálculo iterativo. Esse procedimento é realizado a partir da utilização do algoritmo EM (Expectância-Maximização) (DEMPSTER, LAIRD & RUBIN, 1977; WILKS, 2006), alternativo ao método de Newton-Raphson, adaptado a casos específicos como nos de probabilidades de precipitação. Para tal, é utilizada a equação 15, além da 12, 13 e 14.

$$P(f_{1,2}|x) = \frac{\alpha \cdot f_1(x, \beta_1)}{\alpha \cdot f_1(x, \beta_1) + (1-\alpha) \cdot f_2(x, \beta_2)}, \quad i=1,2,\dots,n \quad (15)$$

Com elas, podem ser determinados os novos valores de β_1 , β_2 e α a partir de:

$$\alpha_j = \frac{1}{n} \sum_{i=1}^n P(f_{1,2}|x) \quad (16)$$

$$\beta_{1,j} = \frac{1}{n \cdot \alpha_j} \sum_{i=1}^n P(f_{1,2}|x) \cdot x \quad (17)$$

$$\beta_{2,j} = \frac{1}{n \cdot (1-\alpha_j)} \sum_{i=1}^n [1 - P(f_{1,2}|x)] \cdot x \quad (18)$$

A convergência é analisada a partir da diferença dos valores do log-verossimilhança entre iterações sucessivas. Usualmente, é atribuído um valor mínimo para essa diferença como critério de parada. É necessário, ainda, para inicializar o cálculo, a atribuição de valores iniciais para β_1 , β_2 e α , que podem ser quaisquer, pois o método garante a convergência.

Apesar da garantia de convergência dada pelo método, pode-se automatizar uma atribuição dos parâmetros iniciais de forma que sejam relativamente próximos dos parâmetros obtidos ao final do procedimento. Isso é feito visando ao menor gasto computacional, proporcionando uma menor quantidade de iterações até a determinação final dos parâmetros requeridos. O mecanismo para determinar essa estimativa foi descrito por Rider (1961) e é baseado no método dos momentos, sua formulação é descrita por:

$$6(2m_1^2 - m_2) \cdot \beta_1^2 + 2 \cdot (m_3 - 3m_1m_2) \cdot \beta_1 + 3 \cdot m_2 - 2 \cdot m_1 \cdot m_3 = 0 \quad (19)$$

Em que: m_1 é a média, m_2 é a variância e m_3 é a assimetria dos dados dos dias considerados chuvosos, em suma, são os momentos amostrais.

Diante disso, β_1 e β_2 são as raízes da equação 19, baseados ainda nas condições impostas pela equação 12, e α é obtido pela expressão que se sucede:

$$\alpha = \frac{m_1 - \beta_2}{\beta_1 - \beta_2} \quad (20)$$

Para a determinação das alturas precipitadas, utiliza-se a inversão da função densidade de probabilidades acumuladas. De acordo com Wilks (1998), inicialmente deve ser comparada uma variável aleatória uniforme no intervalo $[0,1]$ com a probabilidade de mistura α . A média β_1 é escolhida para o dia t , se esta variável uniforme for menor que α , caso contrário, utiliza-se o valor de β_2 . Como a função densidade de probabilidade exponencial pode ser integrada analiticamente, a função de distribuição cumulativa resultante pode ser resolvida para x como uma função de $F(x)$, que segue uma distribuição uniforme (LINDGREN, 1976). Outra variável uniforme aleatória $[0,1]$, digamos v_t , é então produzida e usada para gerar a quantidade de precipitação sintética por inversão (BRATLEY et al., 1987), usando:

$$r(t) = r_{\min} - \beta \cdot \ln(v_t) \quad (21)$$

Em que: β é β_1 ou β_2 , de acordo com o procedimento descrito, e r_{\min} é a acumulação de precipitação diária abaixo da qual um dia é registrado como seco.

2.4.2 Distribuição Gama

Segundo Dash et. al. (2012), a função densidade de probabilidades da distribuição gama é:

$$F_x(x) = \frac{x^{\alpha-1} \cdot e^{-\frac{x}{\beta}}}{\beta^\alpha \cdot \Gamma(\alpha)} \quad (22)$$

$$x > 0; \alpha, \beta > 0$$

Diante das dificuldades de integração da equação 22, a função de distribuição acumulativa de probabilidades pode ser representada, segundo Thom (1958), pela função gama incompleta assumindo a seguinte conformação:

$$F_x(x) = \frac{1}{\beta^\alpha \cdot \Gamma(\alpha)} \int_0^x u^{\alpha-1} \cdot e^{-\frac{u}{\beta}} \quad (23)$$

Em que: α e β são os parâmetros de forma e de escala da função Gama, respectivamente, x é a quantidade de chuva, Γ é a função Gama de α e u é a variável utilizada na integração.

Os parâmetros α e β , assim como, na distribuição exponencial mista, podem ser determinados a partir do método da máxima verossimilhança com a utilização das equações 24 e 25.

$$\alpha = \frac{1 + \sqrt{1 + \frac{4 \cdot A}{3}}}{4 \cdot A} \quad (24)$$

$$\beta = \frac{\bar{X}}{\alpha} \quad (25)$$

Em que: \bar{X} é a média precipitada no período e A é o estimador da função verossimilhança para a média precipitada obtidos a partir de:

$$A = \ln(\bar{X}) - \frac{1}{n} \sum_{j=1}^n \ln(x_j) \quad (26)$$

Em que: x_j é a quantidade de chuva no mês e n é o tamanho da amostra.

De forma também similar à distribuição exponencial, o valor da lâmina precipitada é obtido a partir da aplicação da função inversa. Diante disso, além da utilização dos citados parâmetros, ainda se configura necessário um número aleatório, diga-se, c . Segundo Chen (2005), as aproximações para a densidade e inversa da função de distribuição gama não estão disponíveis na literatura. Assim, as lâminas serão obtidas a partir do fornecimento dos três dados de entrada, α , β e c ao *software*, o qual retorna à precipitação requerida.

2.5 Avaliação dos Ajustes

De acordo com Santos (2017), alguns testes estatísticos convencionais não podem ser aplicados para validar ajustes de modelos estocásticos aos valores sequenciais de chuva diária, dado que as observações não são totalmente independentes (suposição de validade da cadeia de Markov de primeira ordem). Assim, foram aplicados ajustes de qualidade relacionados às Funções Cumulativas de Probabilidades (FCPs), correlacionando-as com a distribuição empírica de Weibull. Para avaliar a capacidade de geração das séries do modelo proposto, utiliza-se de técnicas de estatística descritiva básica, com a determinação e comparação da média, desvio padrão, quantidade de dias chuvosos e dos registros de chuvas máximas.

Na verificação da qualidade do ajuste, utiliza-se o teste de Kolmogorov-Smirnov (KS), que verifica a aderência da distribuição de probabilidades. O teste consiste na análise da diferença absoluta entre os dados da precipitação pluviométrica acumulada observada e os da precipitação acumulada gerada pelo modelo. O comparativo é feito a partir da máxima diferença observada no período de análise. Apresenta então a seguinte formulação:

$$DN_{\max} = \max |F_x(i) - F_x(x)| \quad (27)$$

Em que: $F_x(i)$ é a função de probabilidade acumulada empírica e $F_x(x)$ é a função densidade de probabilidades acumuladas teóricas.

O limite da diferença, DN_{crit} , é obtido a partir do estabelecimento do nível de significância do estudo. Neste trabalho, optou-se por uma significância de 1%, devido à grande variabilidade dos dados diários. Existem tabelas de onde é possível obter esse valor crítico, porém ele pode ser calculado a partir da equação definida por Stephens (1974):

$$DN_{\text{crit}} = \frac{K_\alpha}{\sqrt{n} + 0,12 + 0,11 \cdot \sqrt{n}} \quad (28)$$

Em que: K_α assume o valor de 1,628 para o nível de confiabilidade de 99% e n é o número de elementos da amostra.

Diante disso, é possível perceber que o teste KS assume duas hipóteses:

- a) Hipótese nula (H_0): não existem diferenças entre as séries de precipitação;
- b) Hipótese alternativa (H_1): existe alguma diferença entre as séries de precipitação;

Logo, quando o valor de DN_{max} é inferior ao valor crítico DN_{crit} , aceita-se H_0 , indicando que série é confiável e, caso contrário, rejeita-se H_0 , e a série gerada não atende aos requisitos.

Além do teste KS, pode-se estender a avaliação do ajuste pelos critérios do *mean absolute error* (MAE) e *Root mean square error* (RMSE). De acordo com Chai e Draxler (2014), eles são regularmente empregados como forma de avaliar o desempenho e eficiência do modelo em simular o fenômeno estudado. Baseiam-se, então, no cálculo das distâncias entre as probabilidades esperadas e as observadas. Para tal, utilizam-se as probabilidades observadas e esperadas pelo gráfico Probabilidade-Probabilidade (*P-P plot*). Assim, a mais adequada FCP será a que obtiver menores valores nos testes com seus dados mais próximos da linha 1:1.

$$MAE = \frac{\sum_{i=1}^n |O_i - P_i|}{n} \quad (29)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (O_i - P_i)^2}{n}} \quad (30)$$

Em que: n é a quantidade de dados observados, O_i corresponde ao dado propriamente observado e P_i são os dados simulados ou preditos.

3 RESULTADOS E DISCUSSÃO

Inicialmente, foi feita a determinação das probabilidades de transição entre os dias secos e chuvosos por meio do procedimento representado pelas expressões de 5 a 9. Assim, foram obtidas as probabilidades de ordem zero até a terceira ordem, totalizando 30 valores por mês para cada uma das estações utilizadas. Em seguida, foram aplicados os testes referentes à determinação da cadeia de Markov. Tanto para o teste Akaike como para o Bayes, como mencionado, foi considerada a melhor ordem aquela que produziu os menores valores a partir das equações 10 e 11, que tem por base as expressões de 1 a 4. Na Tabela 3 são reunidos e resumidos os valores determinados das cadeias.

Tabela 3 – Ordem ótima da cadeia de Markov pelo teste AIC e BIC

Teste Mês	Posto 140 (Alagoinha-PE)		Posto 214 (Canhotinho-PE)		Posto 735100 (Limoeiro-PE)	
	AIC	BIC	AIC	BIC	AIC	BIC
Janeiro	2	1	3	1	3	1

Fevereiro	2	1	3	1	2	1
Março	2	1	2	1	3	1
Abril	3	1	3	1	2	1
Maio	2	1	3	1	3	1
Junho	2	2	1	1	3	1
Julho	2	1	2	2	2	1
Agosto	1	1	1	1	2	1
Setembro	1	1	3	1	3	1
Outubro	1	1	3	1	1	1
Novembro	2	1	3	1	2	1
Dezembro	2	1	1	1	3	1

Fonte: Autor (2021)

Com os resultados obtidos, mostrados na Tabela 3, é possível perceber que em apenas 9 casos, 25.0% das ocorrências, o valor da melhor ordem foi igual para ambos os testes. Na maioria dos casos, o teste AIC demonstrou que a ideal seria superior àquelas determinadas pelo teste BIC, reafirmando, para essa dimensão de análise, a tendência de superestimativa analisada por Katz (1981). Diante disso, foram escolhidas as ordens determinadas pelo critério Bayes e, como em 34 ocorrências, 94.44% dos casos, a ordem ótima foi a primeira, adotando-se esta então para todas as análises feitas neste trabalho. Em relação à hipótese de análise de variação da ordem da cadeia associada às localidades, comentada no item 2.1, não foi possível perceber uma grande tendência de alteração, pelo teste AIC, os padrões dos três postos são semelhantes, enquanto pelo teste BIC mais a Oeste surge ocasionalmente a segunda ordem, a exemplo do mês de junho para Alagoinha e julho para Canhotinho.

Diante da consideração da utilização da primeira ordem para todo o conjunto de análise, sintetizou-se na Tabela 4 as probabilidades de transição entre estados utilizadas na construção das séries sintéticas. As probabilidades p_{00} e p_{01} não são representadas por serem complementares às demais indicadas.

Tabela 4 – Probabilidades de transição de primeira ordem

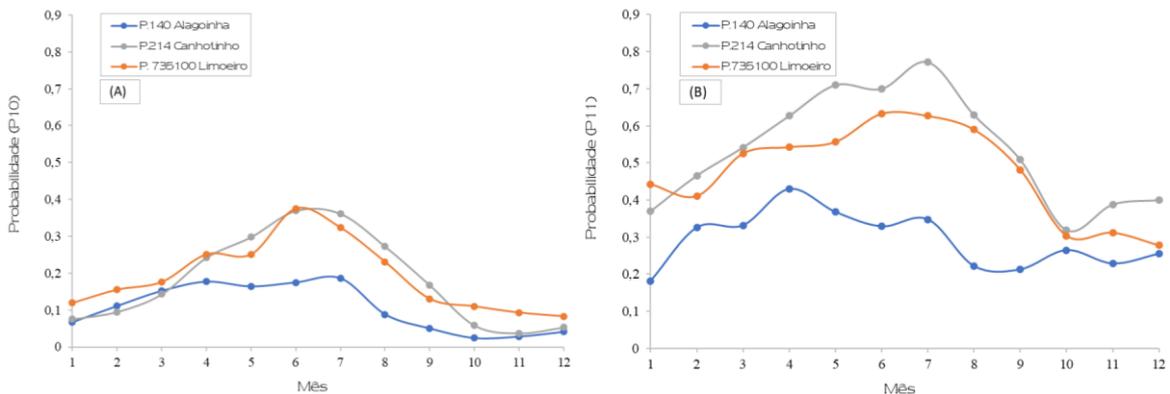
Mês \ Prob.	Posto 140 (Alagoinha-PE)		Posto 214 (Canhotinho-PE)		Posto 735100 (Limoeiro-PE)	
	P ₁₀	P ₁₁	P ₁₀	P ₁₁	P ₁₀	P ₁₁
Janeiro	0,067	0,182	0,075	0,370	0,119	0,443
Fevereiro	0,111	0,326	0,095	0,466	0,155	0,411
Março	0,152	0,332	0,143	0,542	0,176	0,526
Abril	0,177	0,430	0,242	0,627	0,251	0,543
Maio	0,164	0,368	0,298	0,710	0,251	0,557
Junho	0,175	0,329	0,369	0,700	0,375	0,633

Julho	0,186	0,348	0,361	0,772	0,323	0,627
Agosto	0,088	0,222	0,272	0,629	0,230	0,590
Setembro	0,050	0,213	0,167	0,509	0,130	0,482
Outubro	0,024	0,265	0,058	0,319	0,110	0,304
Novembro	0,028	0,229	0,037	0,388	0,093	0,312
Dezembro	0,041	0,255	0,053	0,400	0,083	0,278

Fonte: Autor (2021)

Em face do exposto na Tabela 4, é possível observar quais os períodos chuvosos e secos ao longo do ano. Isso decorre do fato de que essas características são intrínsecas às probabilidades, evidenciado no fato de que quanto maiores forem os valores mais alta será a probabilidade de transição para a ocorrência de um dia chuvoso. Nas Figuras 4A e 4B são apresentadas essas observações de forma mais compreensível.

Figura 4 – Probabilidade de ocorrência de dia chuvoso (A) Após dia seco (B) Após dia chuvoso



Fonte: Autor (2021)

Graficamente, observa-se que as regiões de pico ascendente se configuram como as épocas consideradas chuvosas na região, enquanto as descendentes são os períodos secos, que corroboram com o que foi mostrado na Figura 2. Outra importante informação obtida da figura é que as cidades de Limoeiro e de Canhotinho apresentam os períodos bem distintos por terem ambas as curvas bastante acentuadas. Em contraponto, Alagoinha apresenta uma variação bem mais suave ao longo dos meses, não demonstrando relação clara com os meses chuvosos. Essa característica pode ser entendida, portanto, pela localização da cidade no limite da transição entre a mesorregião do Agreste e do Sertão, sendo influenciada por características climáticas diferentes ao longo dos anos, com tendências de menor ocorrência de chuvas. Enquanto isso, Canhotinho e Limoeiro localizam-se em regiões de característica mais úmida, ambas com tendências à predominância de chuvas nas épocas típicas.

Posteriormente, foi feita a divisão dos dados de acordo com seus meses correspondentes.

Com essa estratégia, considera-se as características, particularidades e respeitam-se as sazonalidades intrínsecas em cada um. Portanto, foram obtidas subséries mensalmente referenciadas, com extensão de acordo com o número de anos de estudo, consideradas em decorrência disso como séries estacionárias.

Cada uma das subséries mensais foi submetida, então, ao procedimento descrito pelas equações de 12 a 20 e 24 a 26, para a determinação dos parâmetros das distribuições. Como pode ser observado por meio da quantidade de equações envolvidas em um processo e em outro, a obtenção dos índices referentes à distribuição Exponencial Mista é bem mais extensa e tem um custo computacional bem maior que o da distribuição Gama. Vale salientar que os critérios de parada utilizados na execução do algoritmo EM teve o erro menor que 10^{-4} ou um número máximo de cinco mil interações. Os resultados são apresentados nas Tabelas 5 e 6.

Tabela 5 – Parâmetros da distribuição de probabilidades da distribuição Exponencial Mista

Parâm. Mês	Posto 140 (Alagoinha-PE)			Posto 214 (Canhotinho-PE)			Posto 735100 (Limoeiro-PE)		
	α	β_1	β_2	α	β_1	β_2	α	β_1	β_2
Janeiro	0,333	36,779	7,235	0,180	10,560	9,776	0,074	42,223	7,768
Fevereiro	0,416	25,169	9,7936	0,216	7,579	6,957	0,623	14,483	5,891
Março	0,369	25,664	11,738	0,0861	21,462	10,210	0,322	20,273	8,498
Abril	0,248	32,276	12,010	0,658	9,889	9,588	0,457	14,500	8,630
Mai	0,123	19,023	7,491	0,746	9,029	8,786	0,319	13,517	9,303
Junho	0,137	20,327	6,895	0,697	9,362	9,082	0,386	18,536	8,120
Julho	0,231	12,173	6,071	0,747	10,324	10,005	0,226	19,998	8,100
Agosto	0,103	7,855	4,416	0,735	7,496	7,307	0,076	20,341	6,396
Setembro	0,211	13,634	3,630	0,562	7,796	7,556	0,082	19,127	6,468
Outubro	0,356	25,000	4,777	0,102	9,599	7,175	0,173	7,869	4,618
Novembro	0,853	15,096	1,995	0,036	55,886	6,962	0,246	12,128	3,279
Dezembro	0,386	26,676	8,844	0,213	8,071	7,6943	0,355	14,607	4,607

Fonte: Autor (2021)

Em breve análise dos parâmetros referentes à distribuição Exponencial Mista, nota-se a existência de uma grande variabilidade dos dados tanto em um mesmo posto quanto entre diferentes. Isso está de certa forma relacionado à variabilidade das condições climáticas à que cada um dos postos está sujeito, devido às diferentes localizações, à influência das estações do ano e do período chuvoso e seco, visto que cada mês é analisado de forma independente dos demais. Outra importante observação é que o parâmetro β_1 apresenta seus valores mínimos em épocas de menores probabilidades de chover e, assim como β_2 , nos meses de maior tendência de ocorrência de precipitações apresentam valores crescentes e/ou estáveis.

Tabela 6 – Parâmetros da distribuição de probabilidades da distribuição Gama

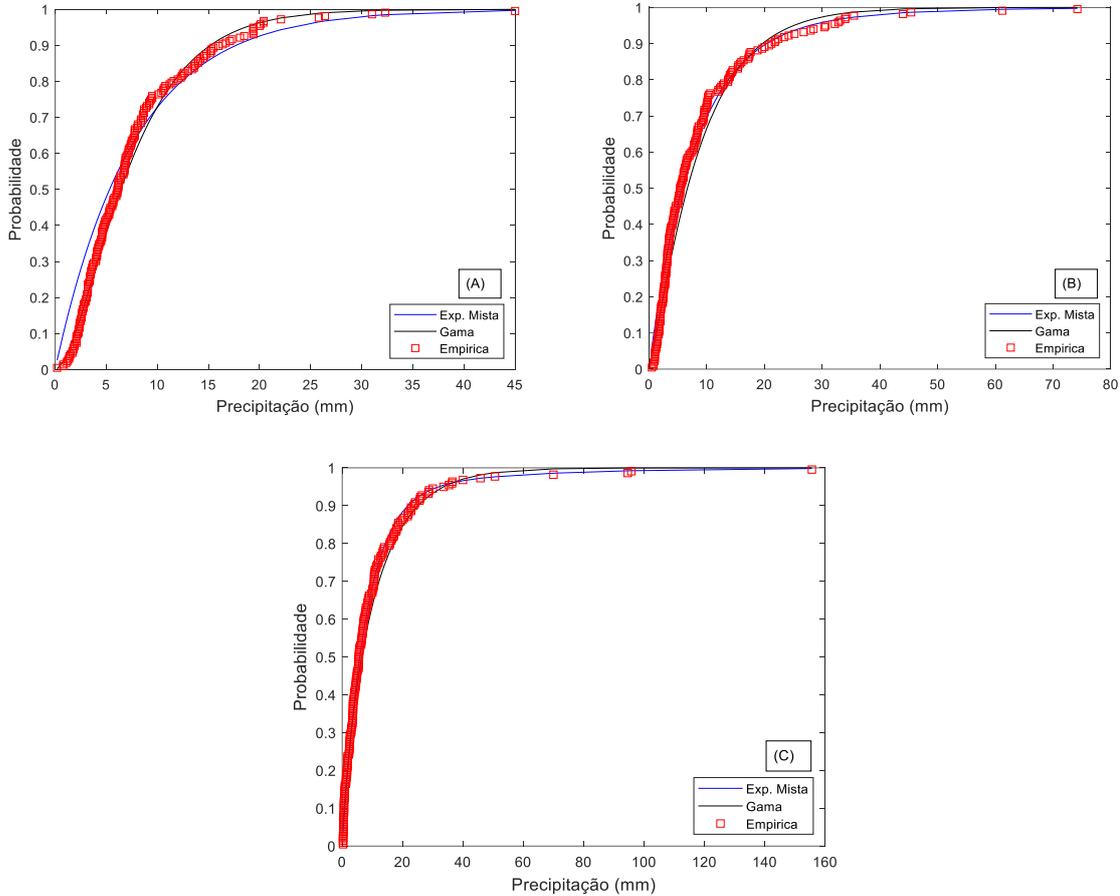
Prob. Mês	Posto 140 (Alagoinha-PE)		Posto 214 (Canhotinho-PE)		Posto 735100 (Limoeiro-PE)	
	α	β	α	β	α	β
Janeiro	0,836	20,698	1,586	6,322	0,801	12,940
Fevereiro	1,030	15,821	1,605	4,455	0,990	11,409
Março	0,932	18,200	1,449	7,775	1,048	11,764
Abril	1,024	16,694	1,694	5,793	1,116	10,165
Maio	1,215	7,369	1,688	5,322	1,165	9,155
Junho	1,131	7,765	1,742	5,337	1,062	11,453
Julho	1,205	6,234	1,637	6,269	1,051	10,291
Agosto	1,360	5,539	1,989	3,752	1,206	6,195
Setembro	1,153	5,061	1,912	4,041	1,193	6,317
Outubro	0,825	14,961	1,591	4,732	1,094	4,764
Novembro	1,088	12,439	1,388	6,400	0,935	5,869
Dezembro	1,099	14,574	1,959	4,025	0,936	8,852

Fonte: Autor (2021)

No que concerne aos parâmetros da distribuição Gama, por inspeção, é possível identificar que α apresenta seus maiores valores em épocas com menor probabilidade de chuva, enquanto β tem esse comportamento, ou de estabilidade, em épocas mais chuvosas. Isso entra em contraponto ao que alguns autores, como Moreira et al. (2010), observaram que os maiores valores de α foram encontrados em épocas de maiores quantidades de precipitações. No entanto, isso pode ser decorrente da condição climática da região, com predominância de dias secos ao longo do ano, fato este que para confirmação dever-se-ia adotar uma maior amostra para sua análise e confirmação. Como β indica o grau de dispersão entre os dados estudados, segundo Thom, (1958), os valores indicam que a distribuição gama pode ser utilizada com segurança pois são inferiores a 100.

Em decorrência da determinação dos parâmetros citados, pode-se fazer a análise de adequação das distribuições de probabilidade. A partir da metodologia de diagnóstico qualitativo adotada, foi possível determinar e representar graficamente como se comporta a frequência de não excedência das intensidades de precipitações para cada mês nas estações abordadas neste estudo. O procedimento foi realizado com base nos dados de ocorrência de chuva de cada mês, durante o período determinado para o estudo. Na Figura 5, é representado o comportamento de um destes casos estabelecidos para cada um dos postos analisados, as demais são apresentadas no Apêndice A.

Figura 5 - Frequência de não excedência (A) Setembro – Canhotinho (B) Maio – Alagoinha (C) Janeiro – Limoeiro

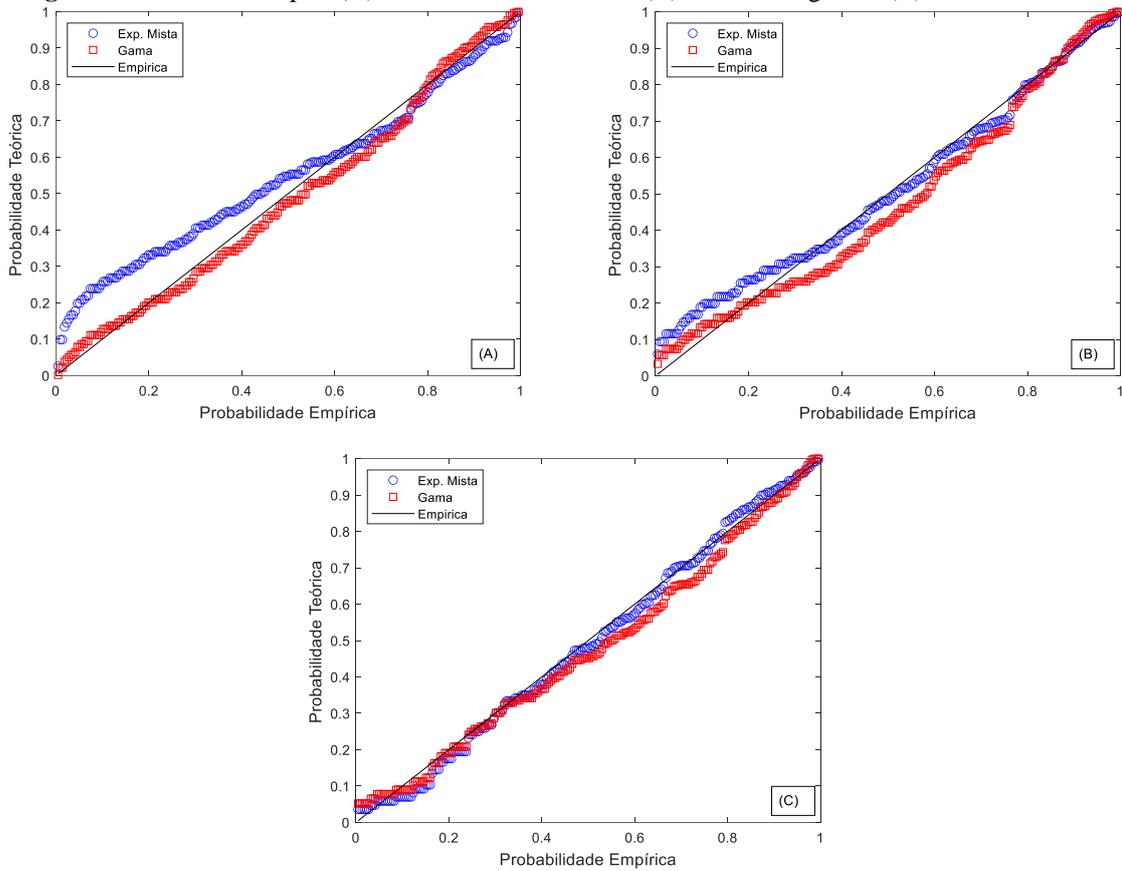


Fonte: Autor (2021)

A frequência de não excedência apresenta a probabilidade de o evento não ocorrer. Logo, a partir da interpretação de cada figura, é possível inferir para cada mês a expectativa de uma chuva de lâmina qualquer, digamos x , não ocorrer. Tomando a Figura 5B, temos que a probabilidade de uma chuva de 80 mm não ocorrer é de praticamente 100 % para Alagoinha no mês de maio, ou, então, ela possui um tempo de retorno muito maior que o observado. Outro aspecto que merece relevância são as linhas formadas pelas distribuições de probabilidades que, nesse nível de detalhe, podem, de forma visual, indicar qual aquela que representa com maior precisão os dados reais.

De forma similar, foi possível, também, a construção dos gráficos probabilidade-probabilidade. O P-P *Plot* representa uma visão mais detalhada que os já citados, em níveis comparativos, do quão próximo ao real, as distribuições são e, a partir disso, foi possível determinar qual aquela que melhor se adequa. Na Figura 6, são representados um dos gráficos PP para cada uma das cidades trabalhadas, os demais são dispostos no Apêndice B.

Figura 6 – Gráficos P-P plot (A) Setembro – Canhotinho (B) Maio – Alagoinha (C) Janeiro – Limoeiro



Fonte: Autor (2021)

Com base na Figura 6, percebe-se que ambas as distribuições apresentam uma boa aderência aos dados. Existe, no entanto, uma leve vantagem na distribuição exponencial mista para o posto de Alagoinha e Limoeiro, enquanto a Gama destaca-se no posto de Canhotinho. Outro fato importante é que à medida que a quantidade de dias chuvosos diminui, o ajuste destoa cada vez mais, fato que pode ser entendido pela menor quantidade de informações que foram consideradas para a determinação dos parâmetros, levando a menores precisões.

Avançando nas verificações, utilizou-se dos dados que foram empregados para a constituição dos gráficos apresentados para se fazer a avaliação dos ajustes. A análise teve como base as equações de 27 e 28. Na Tabela 7, são apresentados os valores calculados para o teste de Kolmogorov-Smirnov.

Tabela 7 – Teste de Kolmogorov-Smirnov

Esp. Mes	Posto 140 (Alagoinha-PE)			Posto 214 (Canhotinho-PE)			Posto 735100 (Limoeiro-PE)		
	DN _{crit}	EM	Gama	DN _{crit}	EM	Gama	DN _{crit}	EM	Gama
Janeiro	0,163	0,116	0,163	0,151	0,121	0,085	0,099	0,057	0,700
Fevereiro	0,125	0,082	0,067	0,133	0,132	0,079	0,095	0,033	0,057
Março	0,104	0,045	0,058	0,101	0,133	0,088	0,080	0,070	0,104

Abril	0,094	0,060	0,075	0,081	0,126	0,054	0,071	0,064	0,074
Mai	0,099	0,093	0,087	0,070	0,132	0,050	0,069	0,050	0,052
Junho	0,100	0,090	0,106	0,068	0,142	0,055	0,059	0,073	0,092
Julho	0,095	0,101	0,091	0,064	0,129	0,056	0,061	0,059	0,076
Agosto	0,140	0,085	0,067	0,075	0,169	0,052	0,069	0,102	0,067
Setembro	0,185	0,144	0,180	0,099	0,165	0,066	0,093	0,075	0,071
Outubro	0,251	0,101	0,131	0,174	0,138	0,094	0,111	0,052	0,077
Novembro	0,240	0,085	0,121	0,208	0,201	0,118	0,121	0,072	0,111
Dezembro	0,195	0,015	0,118	0,173	0,133	0,066	0,128	0,075	0,101

Fonte: Autor (2021)

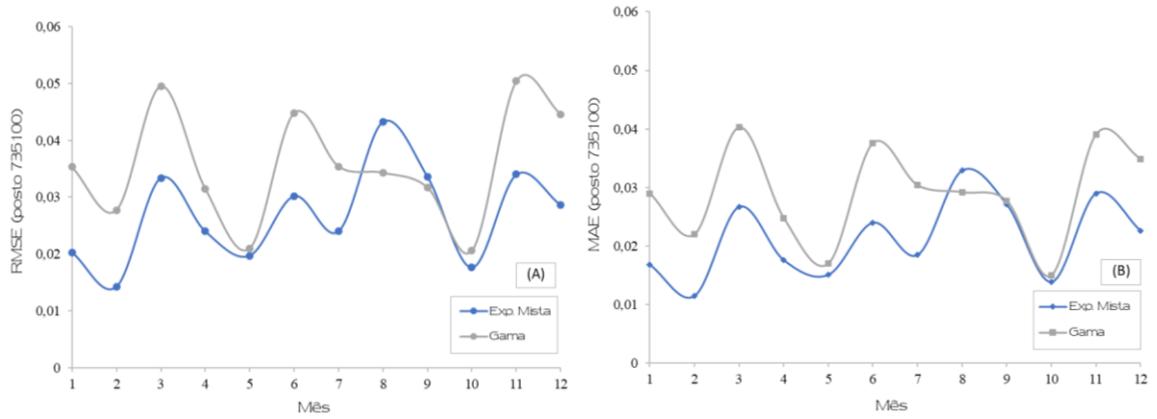
Pela Tabela 7, conclui-se que a distribuição Exponencial Mista realmente foi a mais adequada para o posto 140 e para o 735100. Para o primeiro, somente no mês de julho houve uma inconsistência dentro da significância considerada, mas que não a impossibilita seu uso visto a proximidade dos valores, a mesma ocorrência foi verificada para a distribuição Gama, no mês de junho, que também pode ser considerada adequada para esse posto. No segundo caso, referente ao posto da cidade de Limoeiro, em dois momentos, a condição para o teste não foi aceita, no caso do mês de junho, podendo ser desconsiderada. Já, no mês de agosto, apresenta uma diferença considerável. Esse comportamento, todavia, pode ser decorrente da grande variação que o mês apresenta por se configurar como transitório entre a época chuvosa e a seca. No caso da função Gama, em 4 ocasiões, o teste não foi aceito, sendo neste caso melhor a utilização da primeira distribuição.

Fato notável ocorreu no posto 214, em que a distribuição Gama se mostrou bastante adequada para todos os meses. Nessa mesma linha de análise, a Exponencial apresentou diversas inconsistências, sendo adequada somente em quatro meses durante o ano. Uma hipótese para explicar essa ocorrência pode ser o indicativo de microclima presente na cidade de Canhotinho, adquirindo características de localidades adjacentes visto que é localizada na divisa com a mesorregião Leste do estado de Alagoas que, segundo Barros (2012), apresenta predominância de morros e colinas que influenciam na umidade relativa, com média de 70%, e maior pluviosidade. Ela ainda é próxima da cidade de Garanhuns-PE que, de acordo com Barbosa et. al (2016) compreende características ora referentes ao clima tropical quente e úmido, ora ao clima tropical quente e seco, com a altitude também influenciando nas temperaturas médias que se configuram como mais amenas.

Também foi realizado o cálculo para a obtenção do MAE com a utilização da equação 29 e do RMSE, pela equação 30. Os valores obtidos, para ambas as distribuições foram reunidos e apresentados nas Figuras 7, 8 e 9, correspondentes a cada um dos postos para melhor

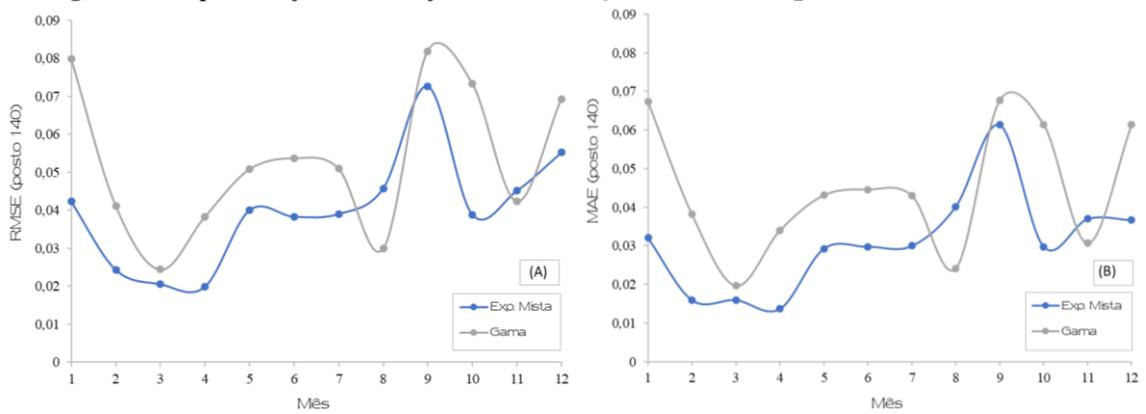
visualização do comportamento.

Figura 7 – Representação da Variação da Análise Qualitativa em Limoeiro (A) RMSE (B) MAE



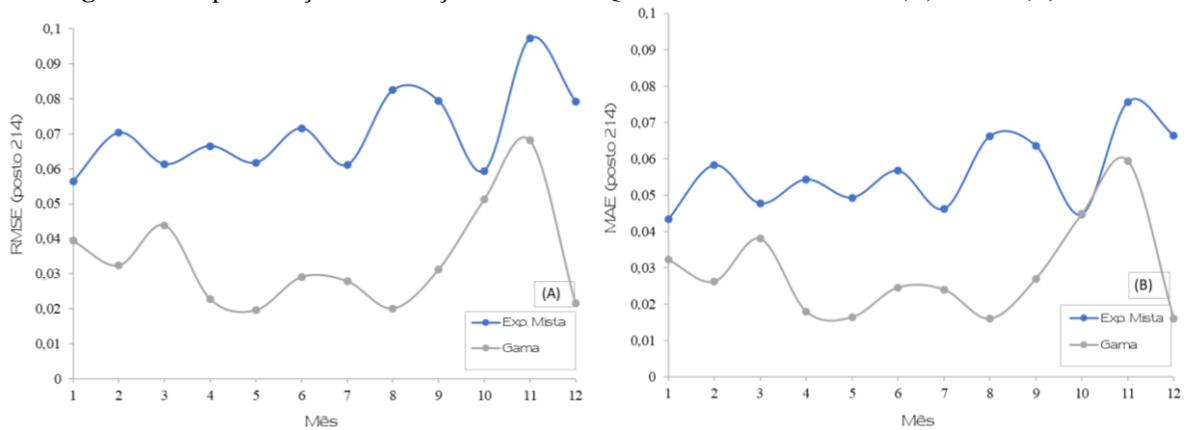
Fonte: Autor (2021)

Figura 8 – Representação da Variação da Análise Qualitativa em Alagoinha (A) RMSE (B) MAE



Fonte: Autor (2021)

Figura 9 – Representação da Variação da Análise Qualitativa em Canhotinho (A) RMSE (B) MAE



Fonte: Autor (2021)

Diante dessa nova análise, podem-se confirmar as afirmações propostas. Apesar de próximas, a Exponencial Mista mostrou-se mais adequada para o posto 140, apresentando valores de RMSE e MAE médios de 0,04 e 0,031, respectivamente, e para o posto 735100, com RMSE médio de 0,027 e MAE de 0,021. Para na estação 214, a Gama, novamente, demonstrou ser a mais apropriada com um RMSE médio de 0,034 e MAE de 0,029.

Após a análise dos ajustes foram geradas as séries sintéticas. Assim, foram produzidas, para cada um dos meses dos postos trabalhados, 1.000 séries. De acordo com Detzel (2009), isso é necessário, pois o Método de Monte Carlo requer a geração de um conjunto grande, na intenção de se obter resultados distintos e tornar possível a criação de cenários. Baseado nisso, tem-se maior visão do comportamento do modelo e se ele, de forma geral, segue as tendências dos dados originais. No confronto dos resultados foram utilizadas comparações consonantes com a estatística descritiva básica, como apresentadas na Tabela 9, para o posto de Alagoinha, os demais podem ser observados no Apêndice C.

Tabela 8 – Comparação entre Séries Sintéticas e a Série Original da Cidade de Alagoinha

Mês	Série	Média	D. Padrão (mm)	Total P. (mm)	Chuva Máx. (mm)	Dias Chuv.	Dias Secos
Janeiro	Original	17,31	24,67	1367,10	138,50	79	975
	EM	16,97	25,02	1340,68	139,42	79	975
	Gama	17,43	18,57	1366,58	94,10	78	976
Fevereiro	Original	16,30	18,41	2217,30	97,00	136	825
	EM	16,36	19,29	2241,24	117,80	137	824
	Gama	16,44	16,09	2230,29	89,22	136	825
Março	Original	16,97	18,74	3325,50	103,20	196	858
	EM	16,99	19,18	3334,58	125,91	196	858
	Gama	17,09	17,48	3313,24	103,47	194	860
Abril	Original	17,10	20,67	4104,50	142,00	240	780
	EM	17,08	20,84	4142,67	151,84	243	777
	Gama	17,19	16,90	4129,49	103,53	240	780
Maio	Original	8,96	10,15	1952,40	74,20	218	836
	EM	8,98	10,23	1958,69	73,56	218	836
	Gama	8,98	8,08	1941,98	47,84	216	838
Junho	Original	8,78	10,78	1852,60	87,00	211	809
	EM	8,86	10,81	1856,93	81,12	210	810
	Gama	8,83	8,20	1840,83	48,70	209	811
Julho	Original	7,51	8,12	1750,60	52,00	233	821
	EM	7,56	8,25	1772,20	56,80	234	820
	Gama	7,58	6,83	1756,42	40,98	232	822
Agosto	Original	4,81	4,87	519,90	28,30	108	946
	EM	4,89	4,96	521,32	27,76	107	947
	Gama	4,82	4,05	509,81	21,08	106	948

Setembro	Original	5,84	7,39	356,00	38,20	61	959
	EM	5,87	7,86	358,26	43,69	61	959
	Gama	5,89	5,34	356,45	25,50	61	959
Outubro	Original	12,40	17,09	409,10	74,00	33	1021
	EM	12,18	17,68	403,31	78,93	33	1021
	Gama	12,58	13,15	401,94	54,54	32	1022
Novembro	Original	13,54	12,92	487,30	53,20	36	984
	EM	13,27	14,27	482,93	60,44	36	984
	Gama	13,51	12,59	487,06	53,70	36	984
Dezembro	Original	16,01	18,16	880,80	81,20	55	999
	EM	15,90	19,49	880,60	97,38	55	999
	Gama	16,05	14,96	883,55	70,16	55	999

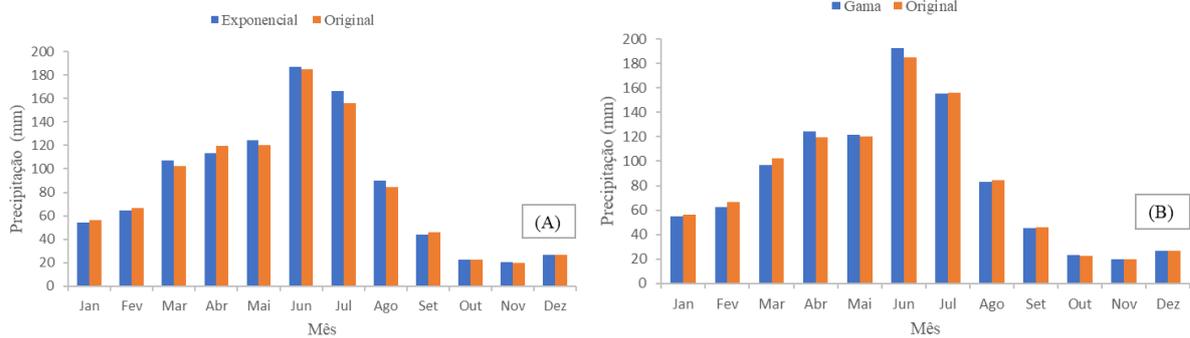
Fonte: Autor (2021)

Em relação à quantidade de dias secos e chuvosos, verificou-se que a Exponencial Mista foi capaz de simular esse aspecto da melhor forma. Repara-se que a Gama apresenta uma leve tendência de superestimação de dias secos, alegada por esta ocorrência em 22 dos 36 meses trabalhados. No caso mais extremo, foram encontrados 8 dias secos a mais na série gerada pela Gama e 7 pela Exponencial, para o mês de agosto, na estação de Canhotinho, mais uma vez isso pode ser decorrente da mudança da época chuvosa para a época seca, agravada pela alta variabilidade de ano para ano. Logo, entende-se que a Exponencial apresentou valores mais próximos aos originais, mas fato que não descaracteriza o uso da Gama.

No tocante as médias das precipitações dos dias considerados úmidos de cada mês, ambas as distribuições apresentaram desempenho satisfatórios. A maior diferença encontrada na Exponencial Mista teve valor de apenas 0,35 mm, 2.0 %, e na Gama 0,18 mm, 1.43 %. Em tese, um fato observado foi que, no conjunto de estudo, os valores médios obtidos nas séries geradas pela distribuição Gama foram melhores e mais próximos à série original, o que, também, não desestimula o uso da outra, pois a diferença obtida é em décimos de milímetros.

Como não é possível a comparação dia a dia da série real e sintética, uma das geradas foi escolhida ao acaso para realizar a verificação do comportamento com base nas precipitações médias mensais. Na Figura 10, A e B, é mostrado o comportamento das séries geradas com uso da função Exponencial e Gama em relação aos dados originais na cidade de Limoeiro, os demais estão representados no Apêndice D. Verificou-se, como esperado, baseado nos dados mostrados na Tabela 8, 9 e 10 que a configuração das chuvas é extremamente similar aos dados reais, o que mostra confiabilidade na sua utilização.

Figura 10 – Comparativo das Médias Mensais Originais em Limoeiro e Sintéticas (A) Exponencial (B) Gama



Fonte: Autor (2021)

Na questão posta em relação ao desvio padrão, a Exponencial Mista novamente mostrou-se superior. Para as simulações feitas, o desvio encontrado foi mais próximo ao dos dados reais, o que indica boa representatividade. Já a distribuição Gama em todos os casos apresentou um valor menor que aquele da base comparativa, sendo este um ponto de atenção, pois sua variabilidade não representa, em sua totalidade, a observada. No posto de Canhotinho, em que foi considerada a mais adequada, percebe-se que o desvio é relativamente mais próximo ao esperado, mas nos demais, o valor registrado destoa em maior grau.

Para o total precipitado, que corresponde à soma de toda a lâmina para o referido mês nos anos da série, o resultado é bastante semelhante aos da média. A função Gama apresentou menores variações em relação ao valor real, com dados mais consistentes, podendo ser decorrente do menor desvio padrão observado. As diferenças máximas observadas foram, para ambas as distribuições, no posto de Canhotinho, sendo 81.45 mm, 2.54%, para a Exponencial Mista e 57.37 mm para a Gama, 2.05 %, reafirmando a consistência de ambas.

Por fim, verificou-se a característica das chuvas máximas, em que foram encontradas uma maior variabilidade dos dados. Percebeu-se que a distribuição Gama apresentou uma tendência de subestimativa dessas chuvas, ocorrendo em 31 dos 36 meses analisados. Outro importante dado coletado é que as maiores diferenças foram verificadas em meses que historicamente ocorrem chuvas de grande intensidade na região, como em janeiro e dezembro. A Exponencial Mista, assim, apresentou resultados mais próximos ao esperado.

Além disso, foi analisada ainda a correlação cruzada entre as séries sintéticas e reais. Em todos os eventos analisados, os valores resultantes foram praticamente nulos, da ordem de 10^{-3} . Logo, as séries são independentes entre si, caracterizando a não necessidade da série original completa para que se possa desenvolver e/ou expandir a série sintética.

4 CONCLUSÃO

Diante do exposto, é possível afirmar que o intuito deste trabalho em estudar os modelos paramétricos para geração de séries sintéticas de precipitações diárias e sua consequente aplicação em áreas específicas do Agreste de Pernambuco foi alcançado. Os métodos mostraram-se adequados para o uso na região, diante do conjunto de dados utilizados. Isso, portanto, instrumentaliza, com essa importante ferramenta, os estudos hidrológicos que necessitam de dados históricos de precipitações em seu desenvolvimento, principalmente aqueles com tempo de retorno de maior escala.

Foi confirmada a tendência de superestimativa do critério Akaike para a ordem da cadeia de Markov. A primeira ordem, predominantemente determinada pelo critério Bayes, foi generalizada e satisfaz as condições para o uso no estudo. O procedimento de geração das séries delimitou-se para o estudo a distribuição Gama a 2 parâmetros e a Exponencial Mista. Pelo teste de Kolmogorov-Smirnov, *mean absolt error* e *Root mean square error*, constatou-se que para o posto da cidade de Alagoinha e de Limoeiro, a distribuição Exponencial melhor se aderiu aos dados, enquanto no posto de Canhotinho a Gama mostrou desempenho superior.

Com a geração das séries por meio da metodologia de Monte Carlo, observou-se que ambas as distribuições estariam adequadas. A Exponencial Mista, no entanto, é mais vantajosa no sentido de ter resultados mais fidedignos aos observados, fato confirmado pela melhor representatividade da série em termos de desvios. É importante frisar que na limitação de desempenho computacional disponível, a Gama atende bem aos requisitos com menor custo computacional e modelagem mais simples.

Para análise da tendência de subestimativa da chuva máxima pela função Gama, é interessante prolongar o conjunto de estudo. A questão das chuvas máxima é de extrema importância, pois sua insuficiência pode levar ao subdimensionamento de obras de drenagem urbana e de contenção de enchentes, por exemplo.

A observação somente desses três postos não indica a utilização de um método ou outro para o Agreste. A generalização poderá ocorrer somente com a análise dos demais postos da região, devido à possível e provável aparição de microclimas regionais que deem preferência a um método em detrimento do outro, como é apontado para a cidade de Canhotinho.

Outra sugestão para ampliar o estudo desenvolvido é a implementação de outras distribuições como Gumbel e a Generalizada de Valores Extremos (GEV), principalmente para verificação da chuva máxima. Outra proposta é utilizar a Gama a 3 parâmetros visto que

com apenas 2 ela apresentou bons resultados, muito próximo da Exponencial Mista, podendo com o aumento da quantidade de informações ser capaz de representar o fenômeno de uma melhor forma.

REFERÊNCIAS

- AKAIKE, H. A new look at the statistical model identification. **IEEE Transactions on Automatic Control**, Brookline, AC-19, 6, p. 716-723, 1974.
- AZEVEDO, P. V.; LEITÃO, M. M. V. B. R. Aplicação de processos em cadeia de Markov às precipitações diárias no estado da Paraíba. **Revista Brasileira de Meteorologia** 5 (1), p. 389-402, 1990.
- BARBOSA, V. V.; SOUZA, W. M.; GALVÍNCIO, J. D.; COSTA, V. S. O. Analysis of climate variability in the city of Garanhuns, Pernambuco - Brazil. **Revista Brasileira de Geografia Física**, v. 9, p. 353-367, 2016.
- BARROS, A. H. C.; FILHO, J. C. A.; SILVA, A. B.; SANTIAGO, G. A. C. F. **Climatologia do Estado de Alagoas**. Boletim de Pesquisa e Desenvolvimento / Embrapa Solos. Recife, 32 p., 2012.
- BAÚ, A. L.; DE AZEVEDO, C. A. V.; BRESOLIN, A. A. Modelagem da precipitação pluvial diária intra-anual da Bacia Hidrográfica Paraná III associada aos eventos ENOS. **Revista Brasileira de Engenharia Agrícola e Ambiental-Agriambi**, v. 17, n. 8, 2013.
- BRATLEY, P., FOX, B.L., SCHRAGE, L.E., 1987. **A Guide to Simulation**. Springer-Verlag, Berlin, 397 p., 1987.
- CHAI, T; DRAXLER, R. R. Root mean square error (RMSE) or mean absolute error (MAE)?—Arguments against avoiding RMSE in the literature. **Geoscientific model development**, v. 7, n. 3, p. 1247-1250, 2014.
- CHEN, Z.Y. The S-system computation of non-central gamma distribution. **Journal of Statistical Computation and Simulation**, New York, v. 75, n. 10, p. 813-829, Oct. 2005.
- CIRILO, J. A. Políticas públicas de recursos hídricos para o semiárido. **Estudos avançados**. Vol. 22, n. 63, pp. 61-82, 2008
- CONTI, I. L.; SCHROEDER, E. O. **Convivência com o semiárido brasileiro: autonomia e protagonismo social**. Editora IABS. Brasília-DF, 2013.
- DEMPSTER, A. P.; LAIRD, N. M.; RUBIN, D. B. Maximum likelihood from incomplete data via EM algorithm. **Journal of the Royal Statistical Society**, Gainesville, p. 1-38, 1977.
- DASH, P. R. A Markov Chain modelling of daily precipitation occurrences of Odisha. **International Journal of Advanced Computer and Mathematical Sciences**. v.3, p482-486, 2012.

DETZEL, D. H. M. **Geração de séries sintéticas de precipitação**. Dissertação (Mestrado em Engenharia de Recursos Hídricos) Departamento de Hidráulica e Saneamento, Setor de Tecnologia, Universidade Federal do Paraná UFPR. Curitiba, p. 183. 2009.

KAISER, Ilza Machado. **Avaliação de métodos de composição de campos de precipitação para uso em modelos hidrológicos distribuídos**. 2006. Tese (Doutorado em Hidráulica e Saneamento) - Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2006. doi:10.11606/T.18.2006.tde-17042006-222747.

KARLIN, S.; TAYLOR, H. M. **An introduction to stochastic modeling**. 3rd ed. New York: Academic Press. 1998.

KATZ, W. R. On some criteria for estimating the order of a Markov Chain. **Technometrics**, Alexandria, 23, 3, p. 243-249, 1981.

KIST, A.; VIRGENS FILHO, J. S. Análise probabilística da distribuição de dados diários de chuva no estado do Paraná. **Ambiente & Água – An Interdisciplinary Journal of Applied Science**, v. 10, n. 1, 2015.

LAURA, A. A. **Estudo do uso de séries parciais e anuais na análise de frequência de chuvas intensas aplicando modelos de distribuição de probabilidade**. Dissertação (Mestrado em Engenharia Civil) – Universidade Federal do Ceará, Fortaleza, 1997.

LINDGREN, B.W. **Statistical Theory**. Macmillan, London, 614 p., 1976.

MOREIRA, P. S. P., et al., Distribuição e probabilidade de ocorrência de chuvas no município de Nova Maringá-MT. **Revista de Ciências Agroambientais**, Alta Floresta, v.8, n.1, p.9- 20, 2010.

MORENO PÉREZ, M. F; WOOLHISER, D. A; ROLDÁN CAÑAS, J. Effects of parameter perturbation on daily precipitation models in Southern Spain using the NAO index. **International Journal of Climatology**, v. 34, n. 8, p. 2556-2572, 20 14.

PROGÊNIO, M. F. **Modelagem estocástica de função cumulativa de probabilidades de precipitação diária na região hidrográfica Tocantins-Araguaia (RHTA)**. 119 f. Dissertação (Mestrado) – Universidade Federal do Pará, Instituto de Tecnologia, Belém, 2019. Programa de Pós-Graduação em Engenharia Civil. Disponível em: <<http://repositorio.ufpa.br/jspui/handle/2011/11148>>. Acesso em: 24/03/2021.

RICHARDSON, C. W. Stochastic simulation of daily precipitation, temperature, and solar radiation. **Water Resources Research**, Washington, DC, 17, 1, p. 182-190, 1981.

RIDER, P. R. The method of moments applied to a mixture of two exponential distributions. **Aeronautical Research Laboratories**, Wright-Patterson Air Force Base, 32, 1, p. 143-147, 1961.

SANTOS, V. C. **Probabilidade de ocorrência de chuvas e sua variação espacial e temporal na Bacia Hidrográfica do Rio Tapajós**. Dissertação de Mestrado – Universidade

Federal do Pará. Instituto de Tecnologia. Programa de Pós-Graduação em Engenharia Civil, Belém, 2017a.

SCHWARZ, G. Estimating the dimension of a model. **The Annals of Mathematical Statistics**, Bethesda, 6, 2, p. 461-464, 1978.

STEPHENS, M. A. EDF statistics for goodness of fit and some comparisons. **Journal of the American statistical Association**, v. 69, n. 347, p. 730-737, 1974.

THOM, H. C. S. A note on the gamma distribution. **Monthly Weather Review**, v. 86, p. 117-22, 1958.

WAN, H; ZHANG, X; BARROW, E. M. Stochastic modelling of daily precipitation for Canada. **Atmosphere-Ocean**, v. 43, n. 1, p. 23-32, 2005.

WILKS, D S. Multisite generalization of a daily stochastic precipitation generation model. **Journal of Hydrology**, Amsterdam, 210, p. 178-191, 1998.

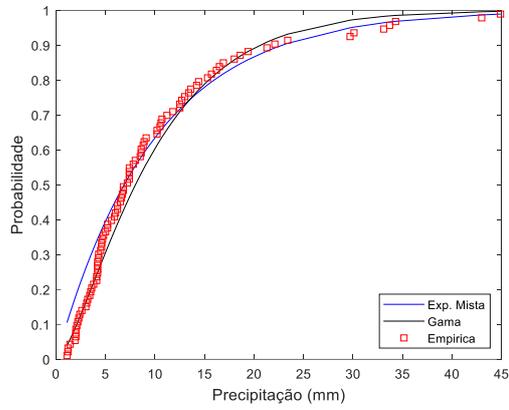
WILKS, D.S. **Statistical methods in the atmospheric sciences**. 2nd. ed. San Diego: Academic Press, 2006.

WMO – World Meteorological Organization. **Guide to hydrological practices**. N.168. Vol. 2. Management of Water Resources and Application of Hydrological Practices, 2009.

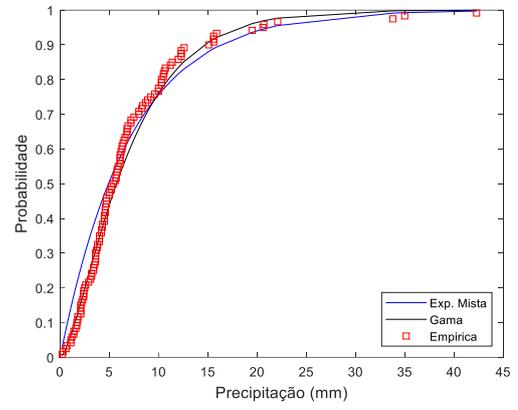
ZAKARIA, N. Z; DENI, S. M. Application of Alternative Geometric Distribution and Markov Chain Models for Fitting Sequences of Wet and Dry Days in Peninsular Malaysia. **International Journal of Engineering and Management Research (IJEMR)**, v. 6, n. 1, p. 110-119, 2016.

APÊNDICE A

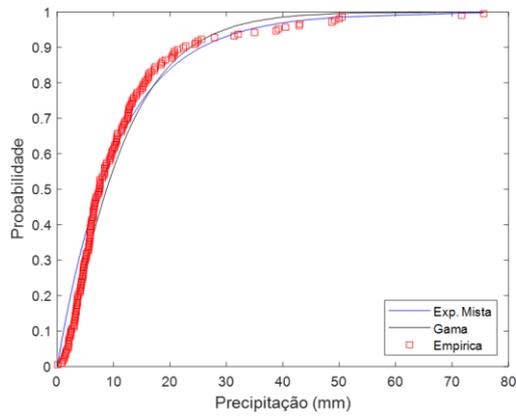
Figura 11 – Frequência de Não excedência (Canhotinho-PE)



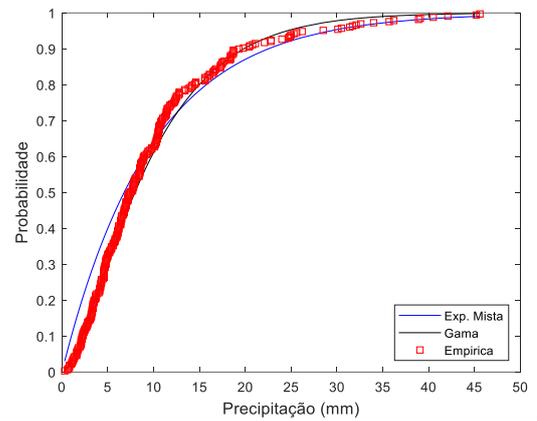
Janeiro



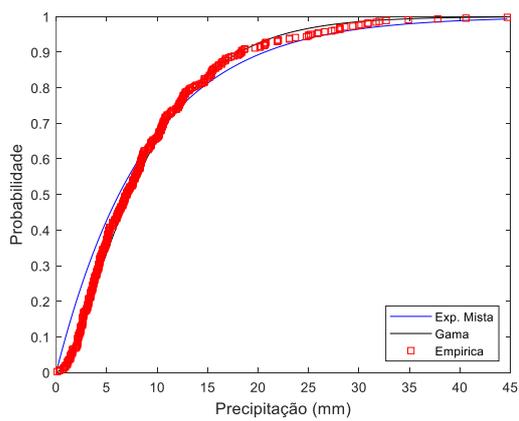
Fevereiro



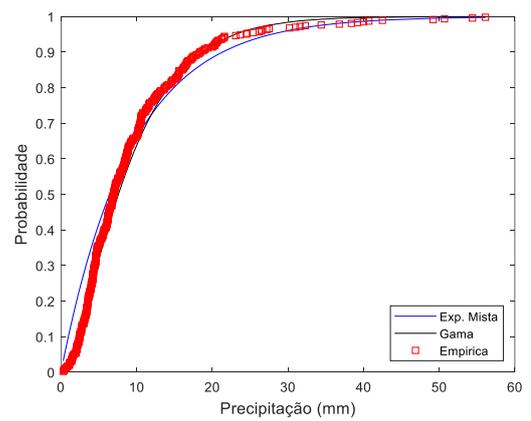
Março



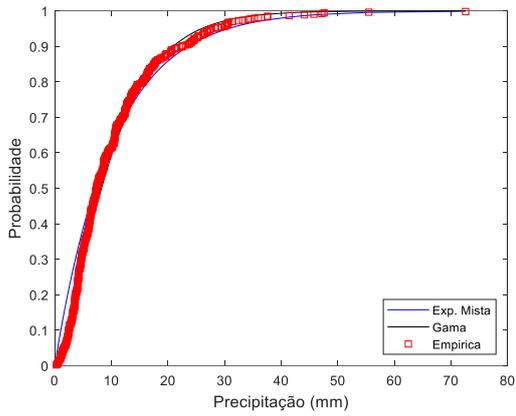
Abril



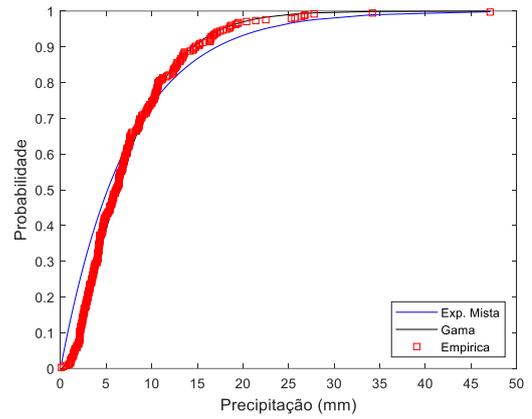
Maio



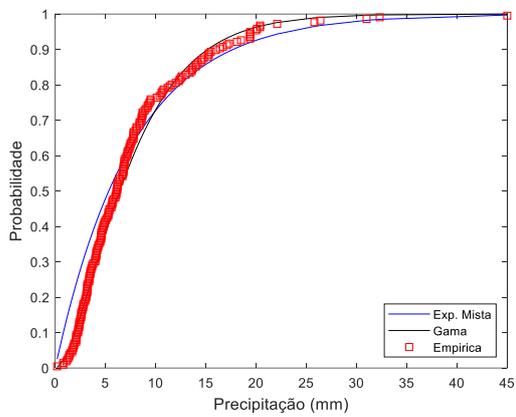
Junho



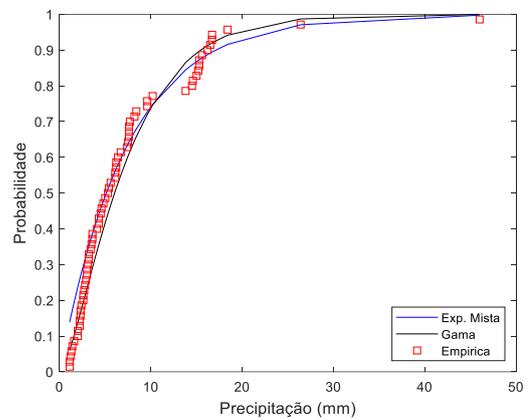
Julho



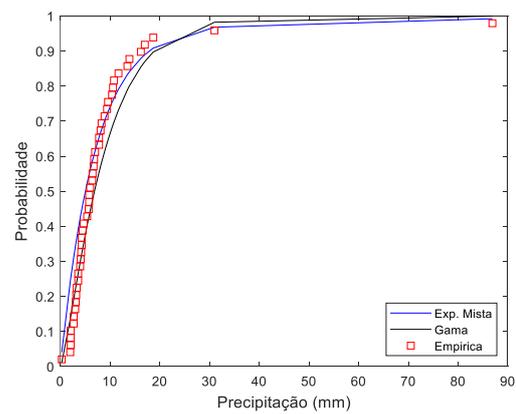
Agosto



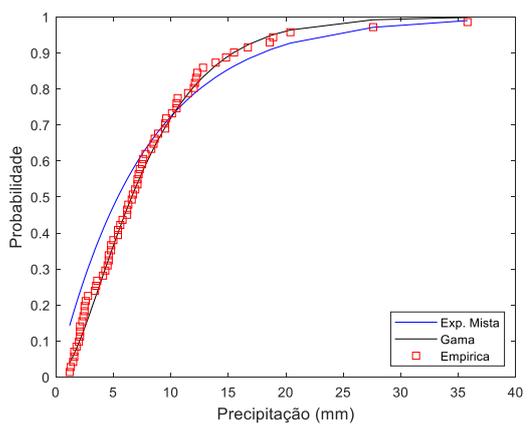
Setembro



Outubro



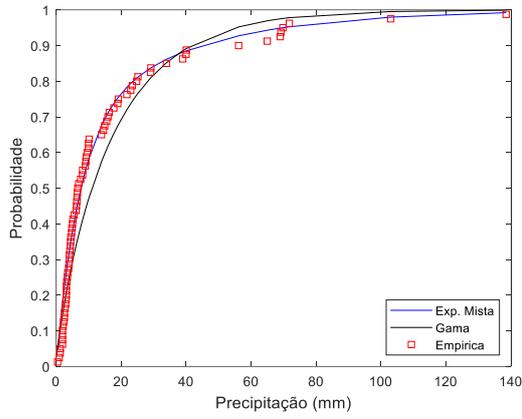
Novembro



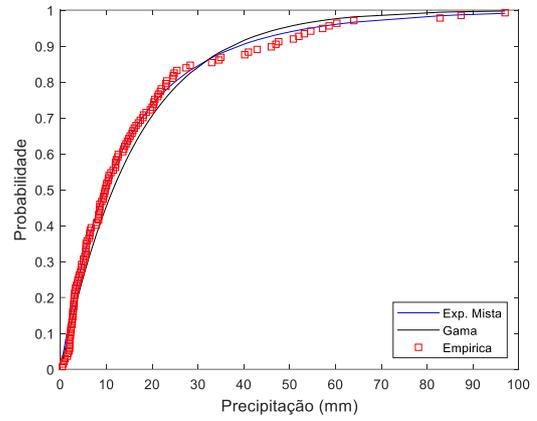
Dezembro

Fonte: Autor (2021)

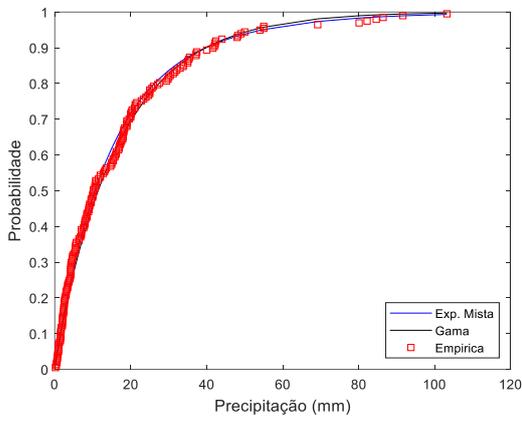
Figura 12 – Frequência de Não Excedência (Algoinha-PE)



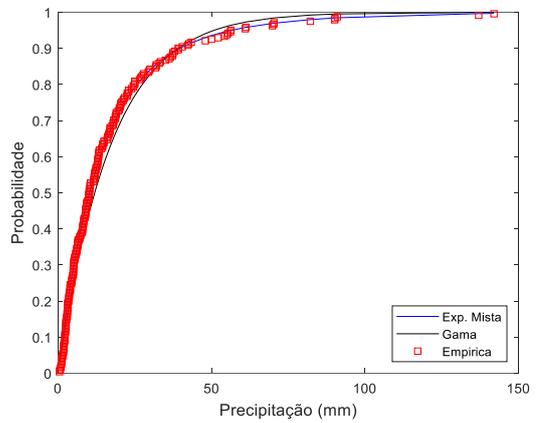
Janeiro



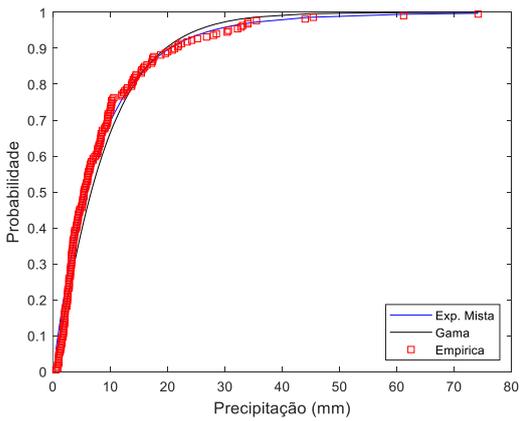
Fevereiro



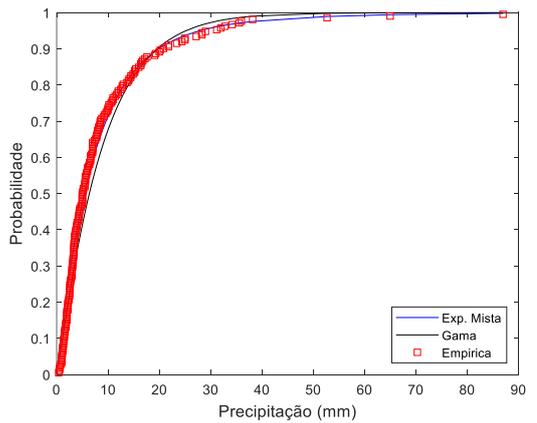
Março



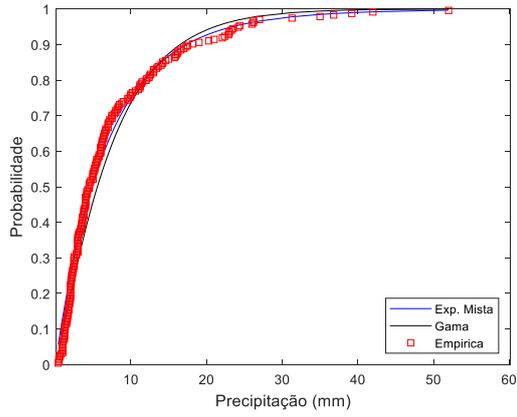
Abril



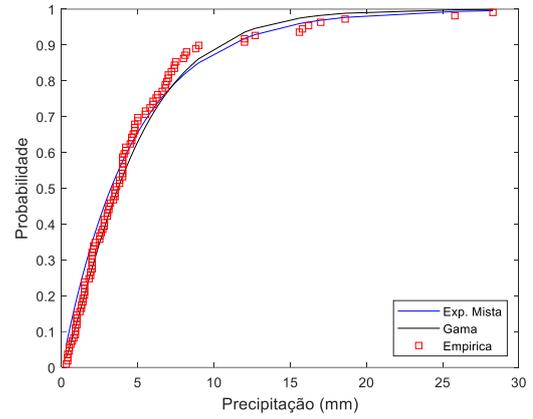
Maio



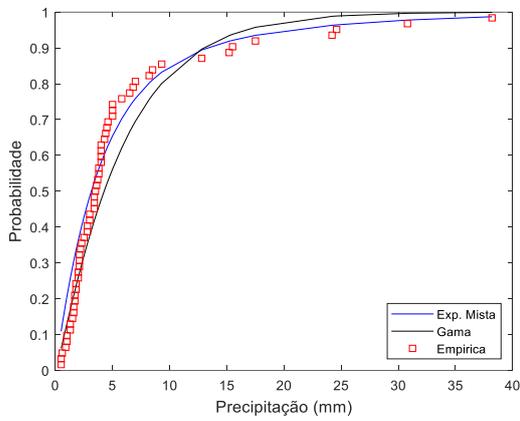
Junho



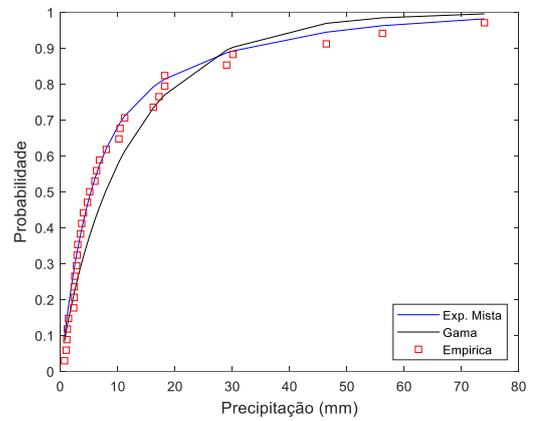
Julho



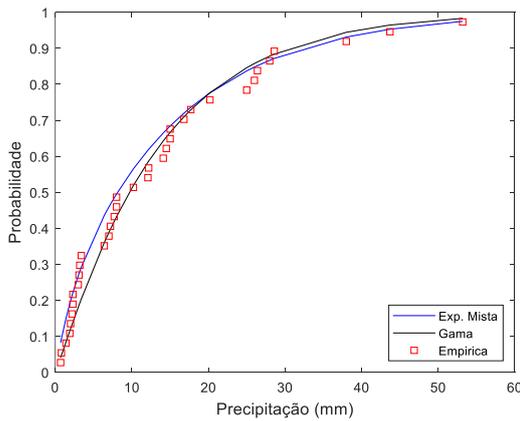
Agosto



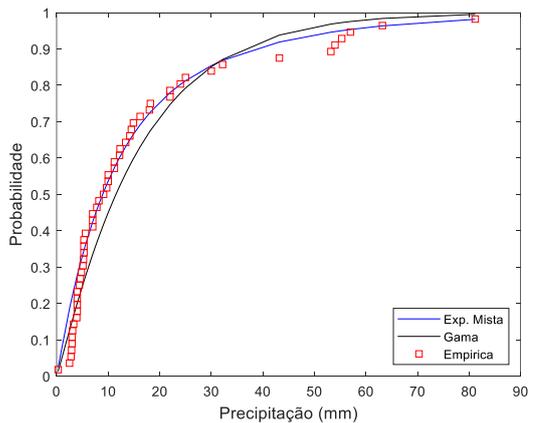
Setembro



Outubro



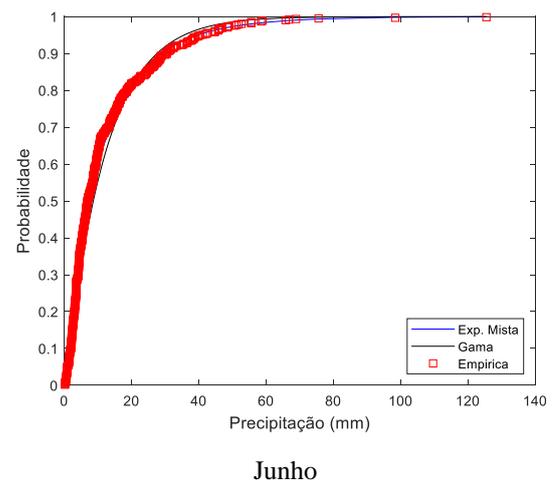
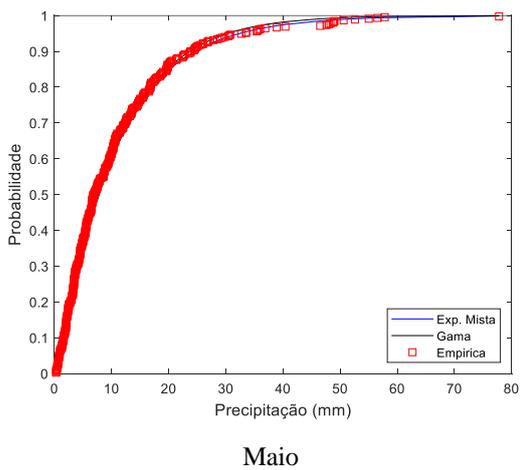
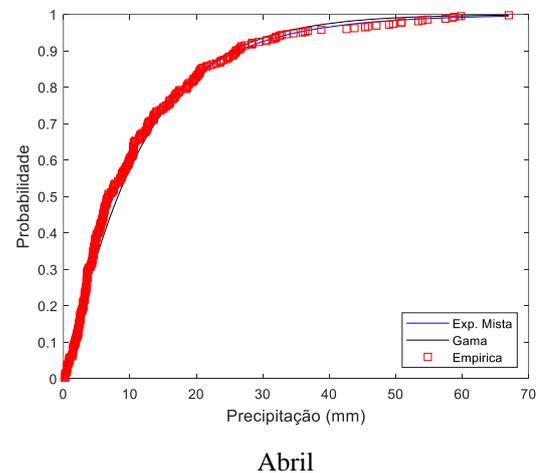
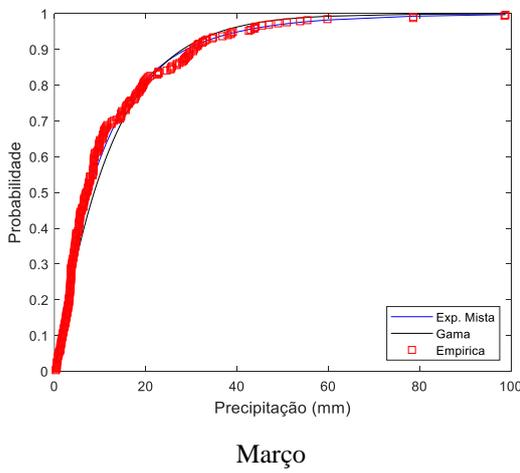
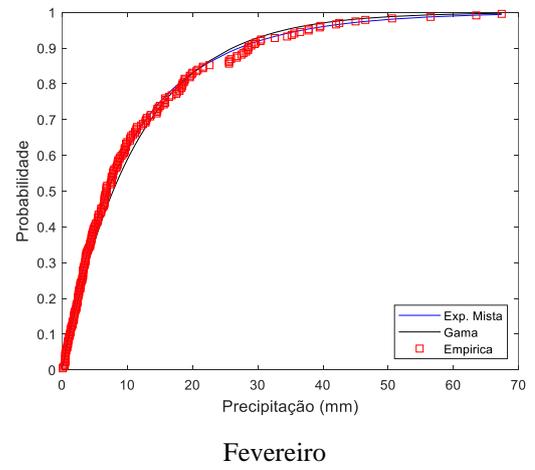
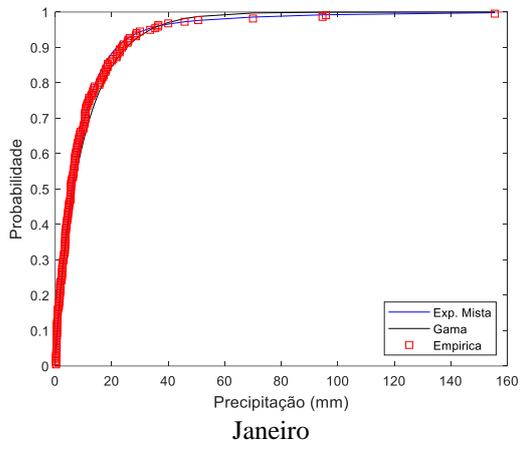
Novembro

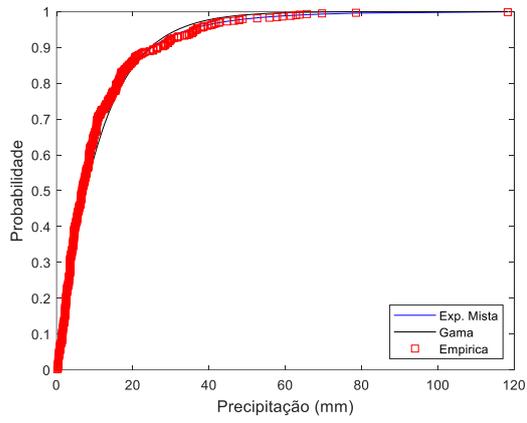


Dezembro

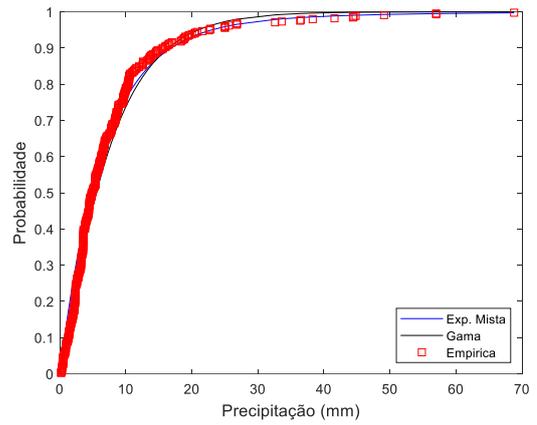
Fonte: Autor (2021)

Figura 13 – Frequência de Não Excedência (Limoeiro-PE)

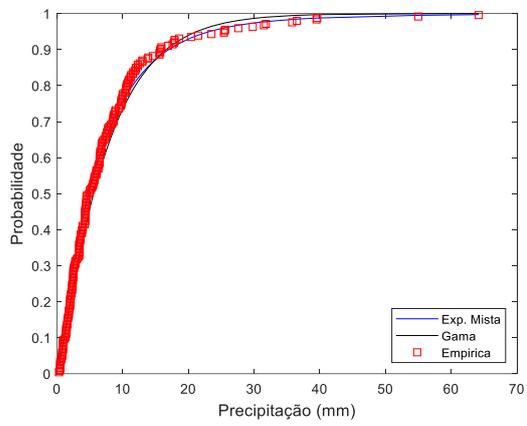




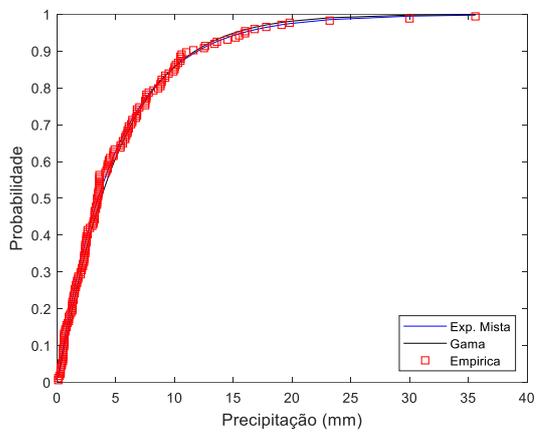
Julho



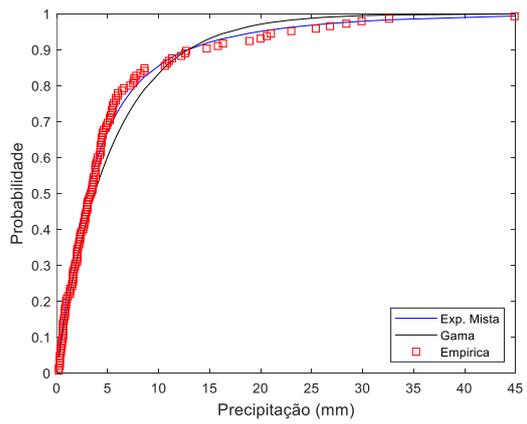
Agosto



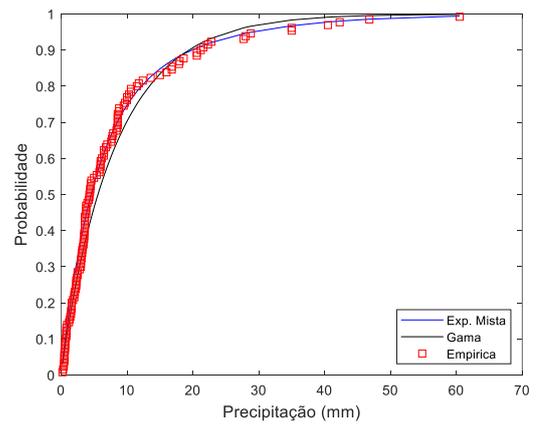
Setembro



Outubro



Novembro

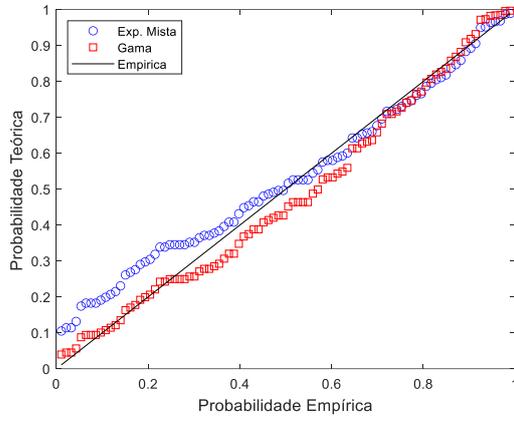


Dezembro

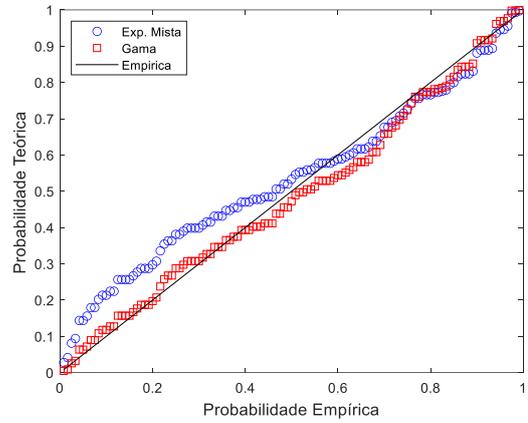
Fonte: Autor (2021)

APÊNDICE B

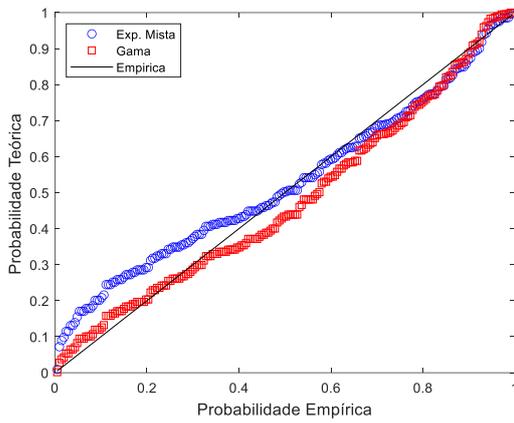
Figura 14 – Gráficos P-P plot (Canhotinho-PE)



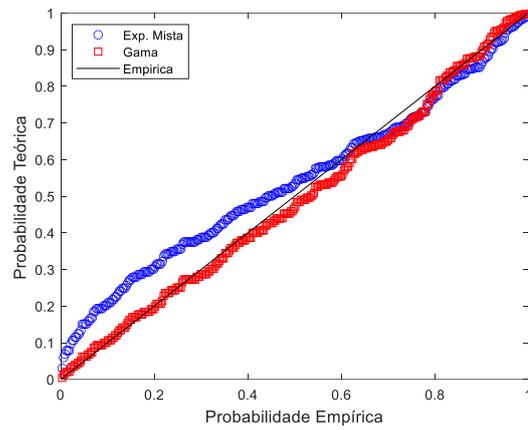
Janeiro



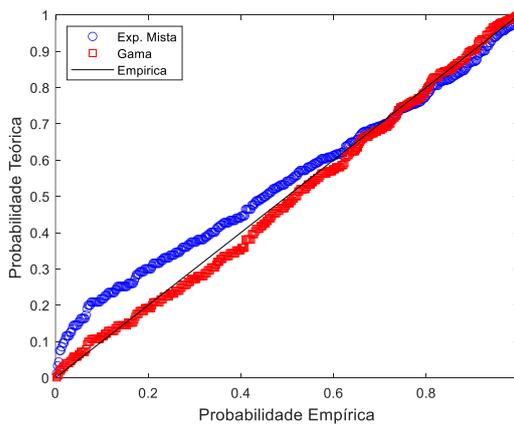
Fevereiro



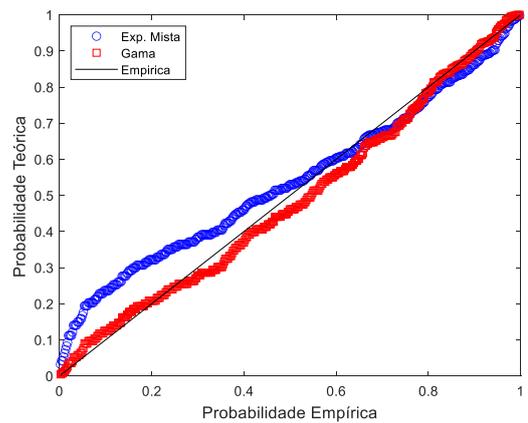
Março



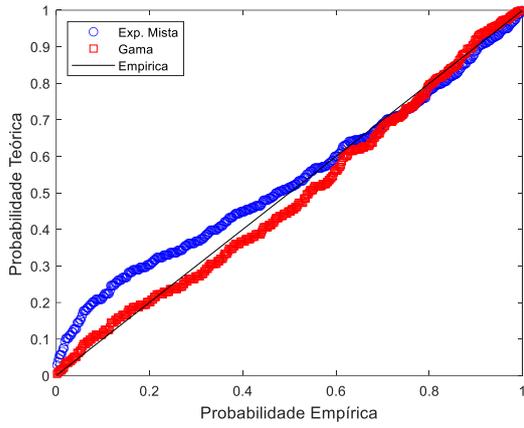
Abril



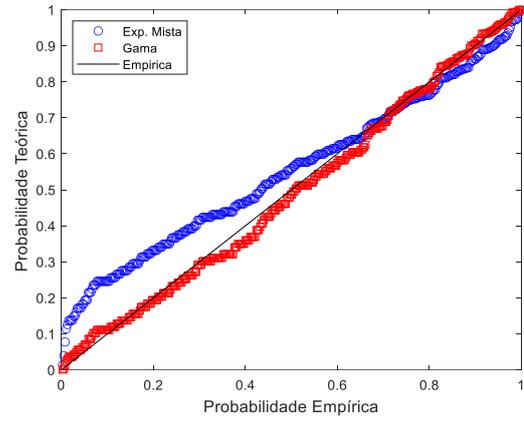
Maio



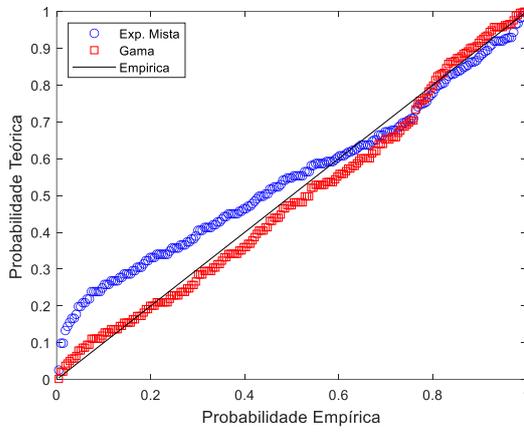
Junho



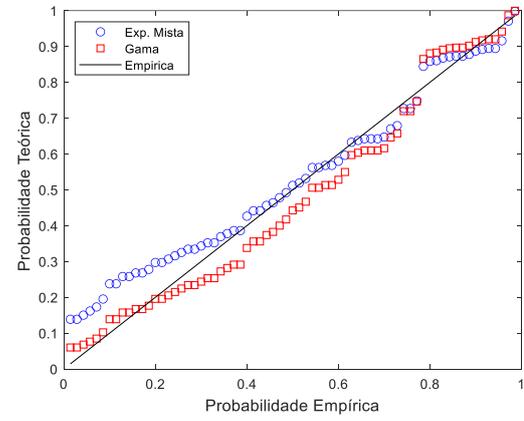
Julho



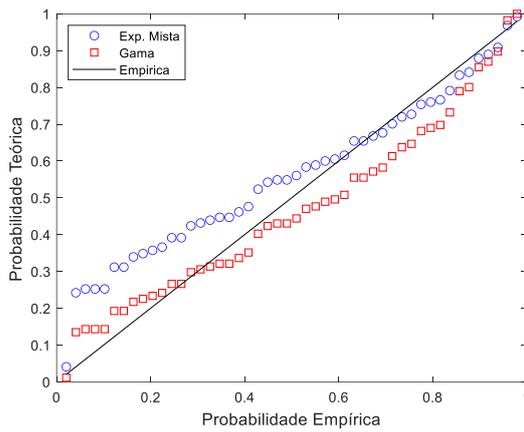
Agosto



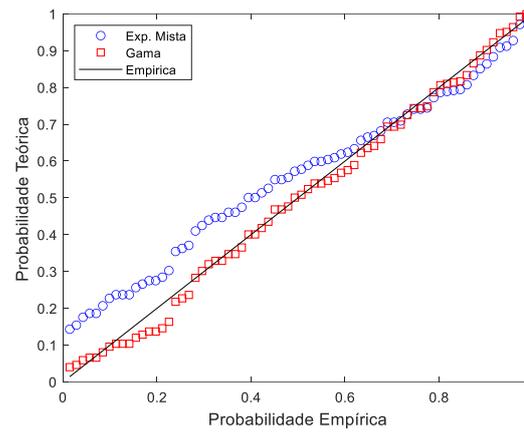
Setembro



Outubro



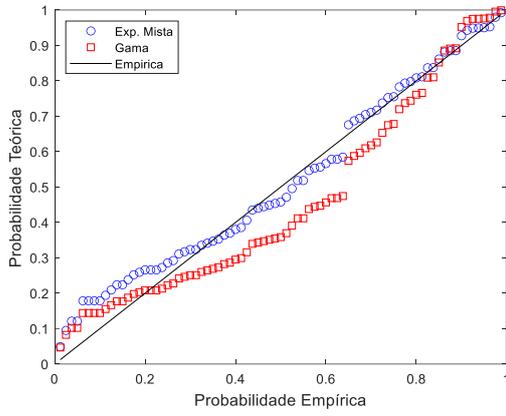
Novembro



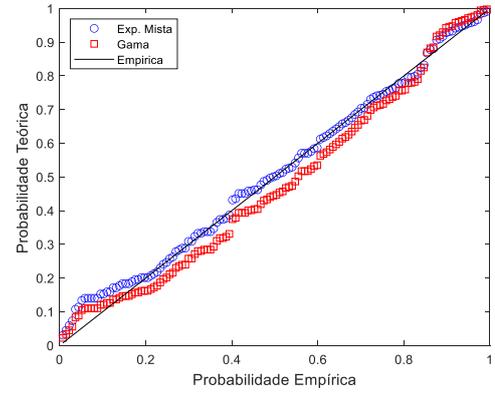
Dezembro

Fonte: Autor (2021)

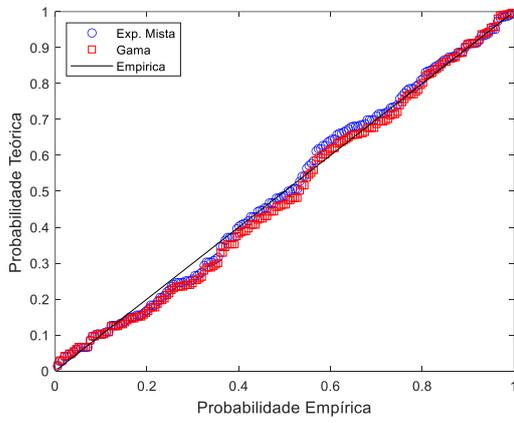
Figura 15 – Gráficos P-P *plot* (Alagoinha-PE)



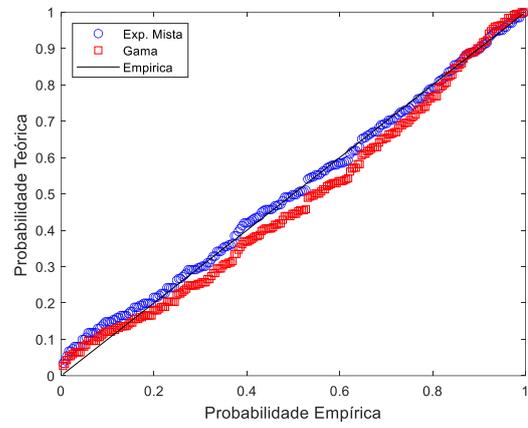
Janeiro



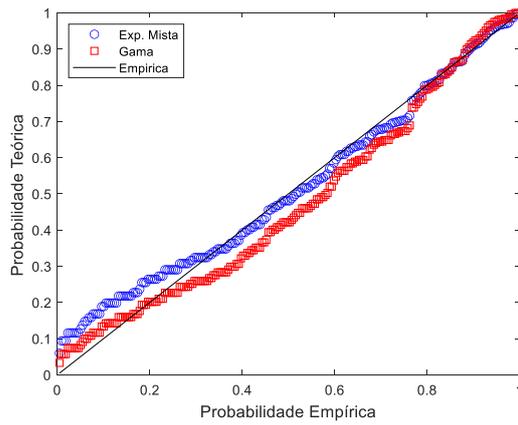
Fevereiro



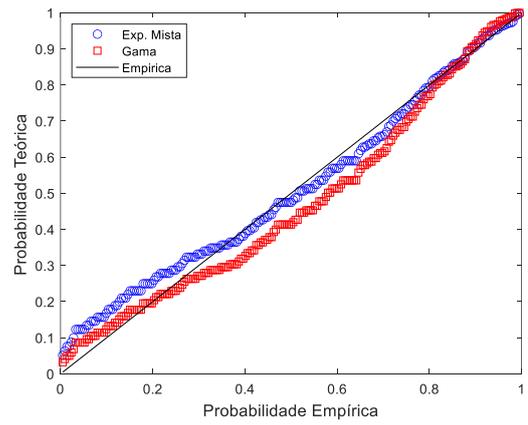
Março



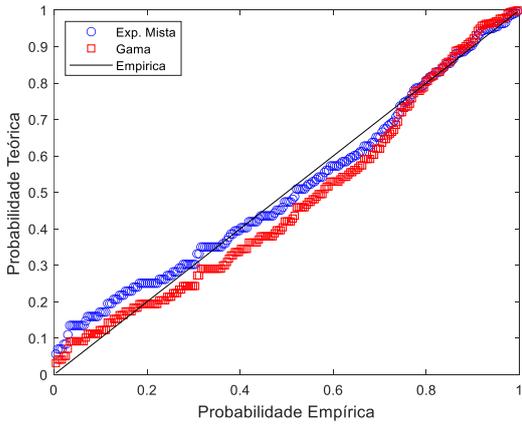
Abril



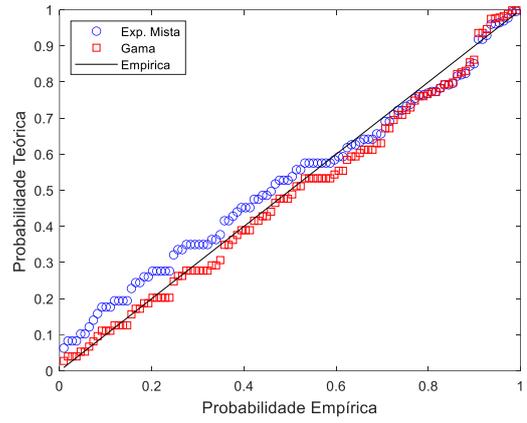
Maio



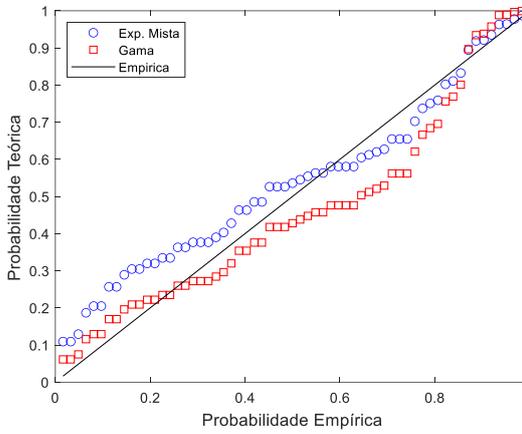
Junho



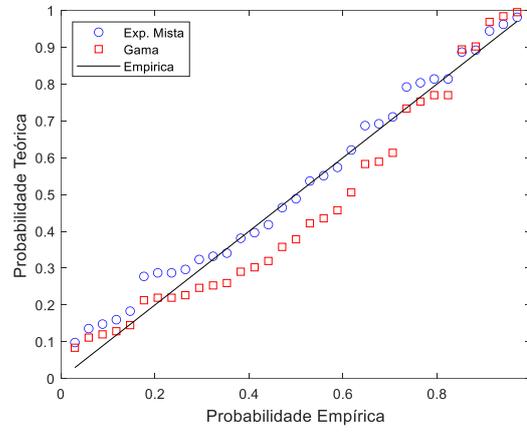
Julho



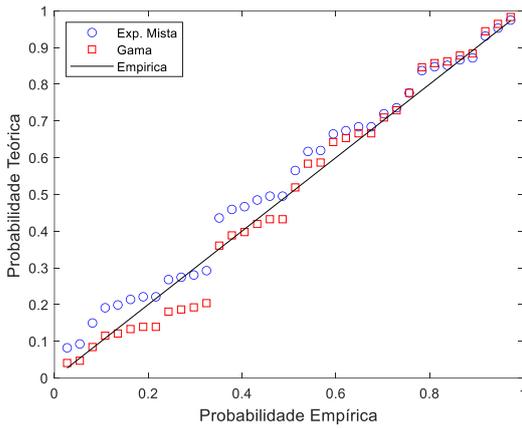
Agosto



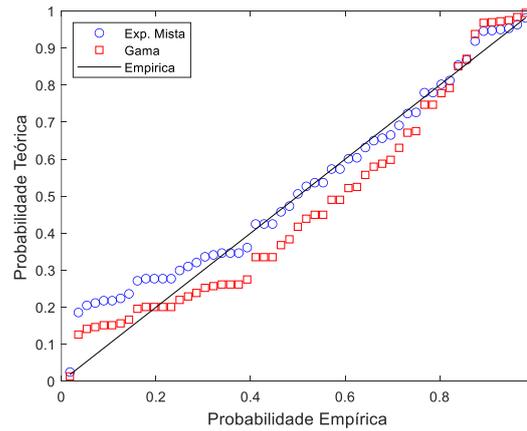
Setembro



Outubro



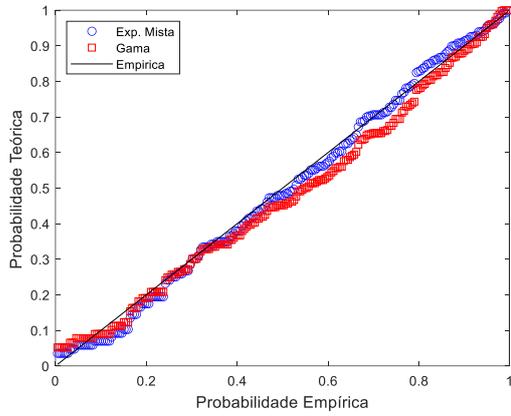
Novembro



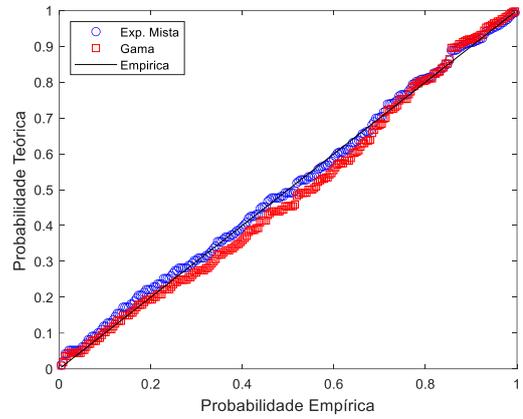
Dezembro

Fonte: Autor (2021)

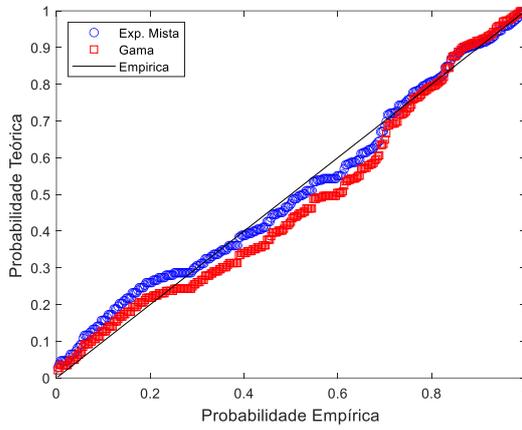
Figura 16 – Gráficos P-P plot (Limoeiro-PE)



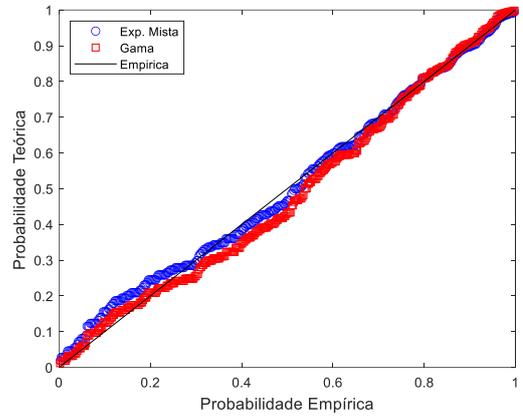
Janeiro



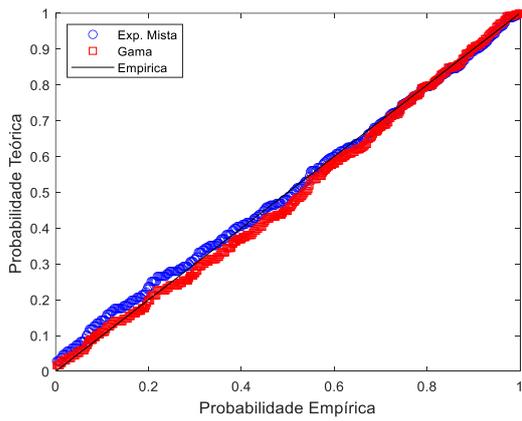
Fevereiro



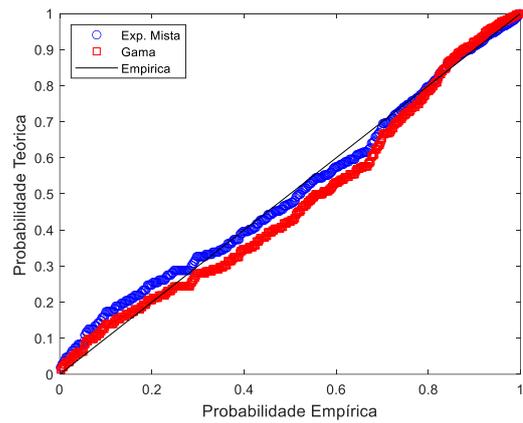
Março



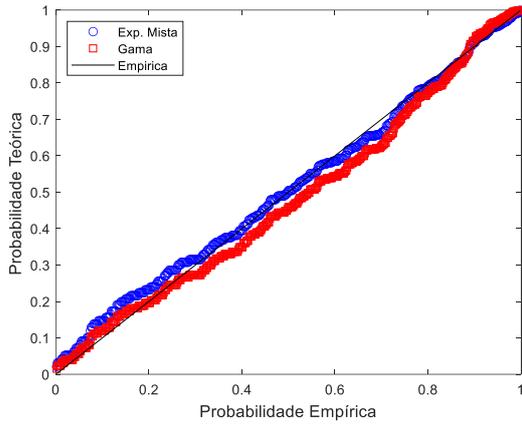
Abril



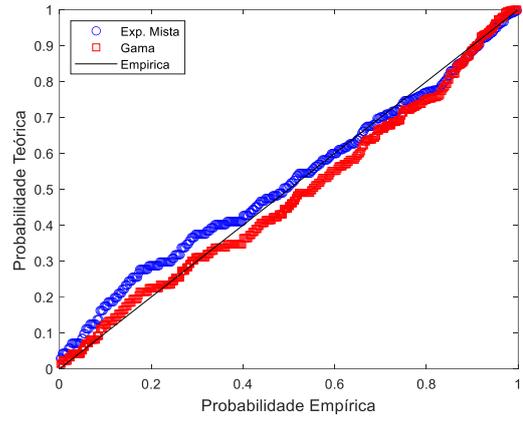
Maio



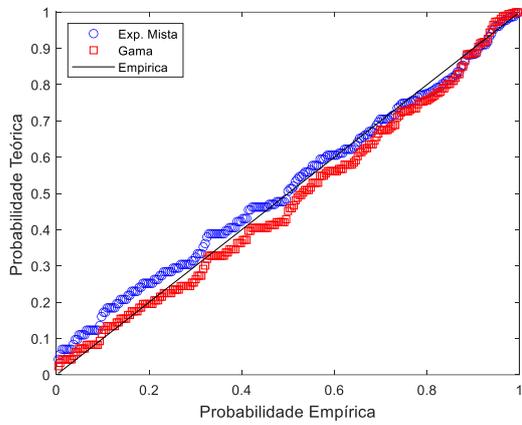
Junho



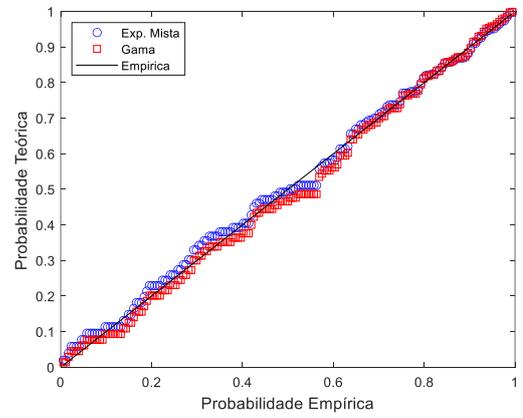
Julho



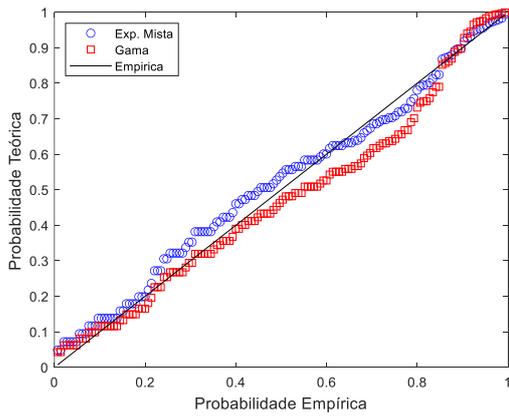
Agosto



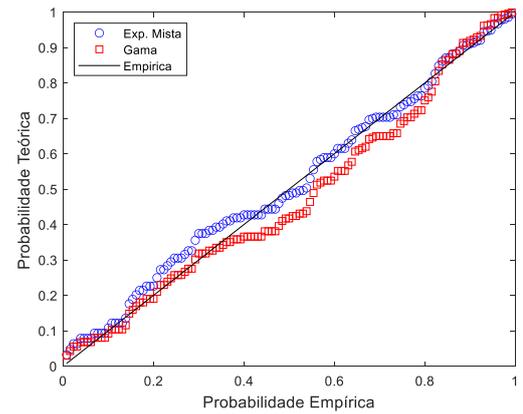
Setembro



Outubro



Novembro



Dezembro

Fonte: Autor (2021)

APÊNDICE C

Tabela 9 – Comparação entre séries sintéticas e a série original da cidade de Canhotinho

Mês	Série	Média	DP (mm)	TP (mm)	CM (mm)	DC	DS
Janeiro	Original	10,03	9,13	922,3	44,9	92	776
	EM	9,95	9,73	923,24	50,1	93	775
	Gama	9,99	7,85	921,39	40,23	92	776
Fevereiro	Original	7,15	6,65	851	42,3	119	672
	EM	7,21	7,1	862,93	38,62	120	671
	Gama	7,16	5,58	852,18	29,63	119	672
Março	Original	11,23	11,55	2314,2	75,6	206	662
	EM	11,28	11,89	2333,01	79,09	207	661
	Gama	11,25	9,32	2324,42	53,69	207	661
Abril	Original	9,82	8,24	3200,1	45,6	326	514
	EM	9,9	9,75	3281,55	62,16	332	508
	Gama	9,83	7,52	3257,18	46,31	331	509
Maio	Original	8,99	7,33	3954,5	44,7	440	428
	EM	9,06	8,92	3995,38	59,31	441	427
	Gama	8,99	6,9	3949,75	44,36	439	429
Junho	Original	9,3	8,12	4314	56,1	464	376
	EM	9,38	9,24	4341,89	62,12	463	377
	Gama	9,29	7,02	4307,76	44,97	464	376
Julho	Original	10,26	8,93	5418,8	72,6	528	340
	EM	10,33	10,22	5497,62	69,92	532	336
	Gama	10,27	8,01	5470,11	52,27	533	335
Agosto	Original	7,46	5,74	2791,3	47,1	374	494
	EM	7,55	7,4	2773,62	47,85	367	501
	Gama	7,46	5,27	2733,91	32,69	366	502
Setembro	Original	7,73	6,26	1661,2	45	215	625
	EM	7,79	7,69	1654,43	46,12	212	628
	Gama	7,71	5,56	1641,08	32,11	213	627
Outubro	Original	7,53	7,21	519,5	46	69	799
	EM	7,48	7,3	510,95	35,8	68	800
	Gama	7,53	5,91	516,8	29,11	69	799
Novembro	Original	8,89	12,72	426,6	86,9	48	792
	EM	8,73	12,7	411,18	72,39	47	793
	Gama	8,92	7,49	425,61	34,21	48	792
Dezembro	Original	7,89	6,21	552	35,8	70	798
	EM	7,9	7,68	555,25	37,66	70	798
	Gama	7,85	5,54	541,37	27,04	69	799

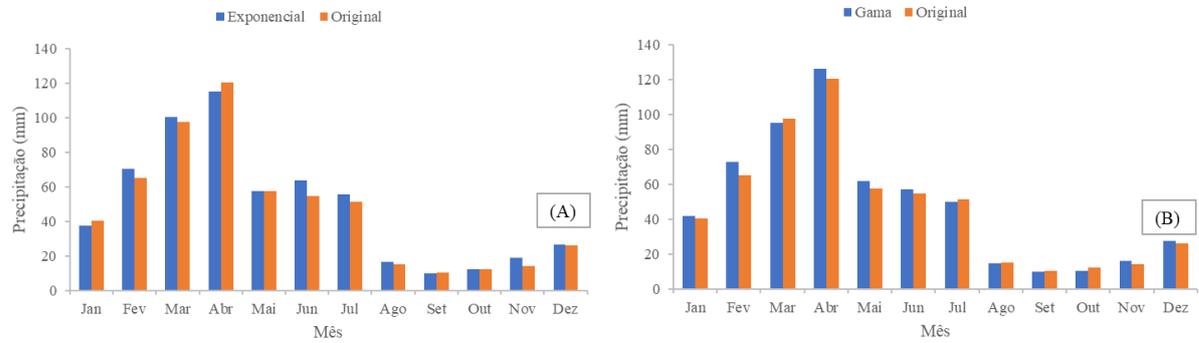
Fonte: Autor (2021)

Tabela 10 – Comparação entre séries sintéticas e a série original da cidade de Limoeiro

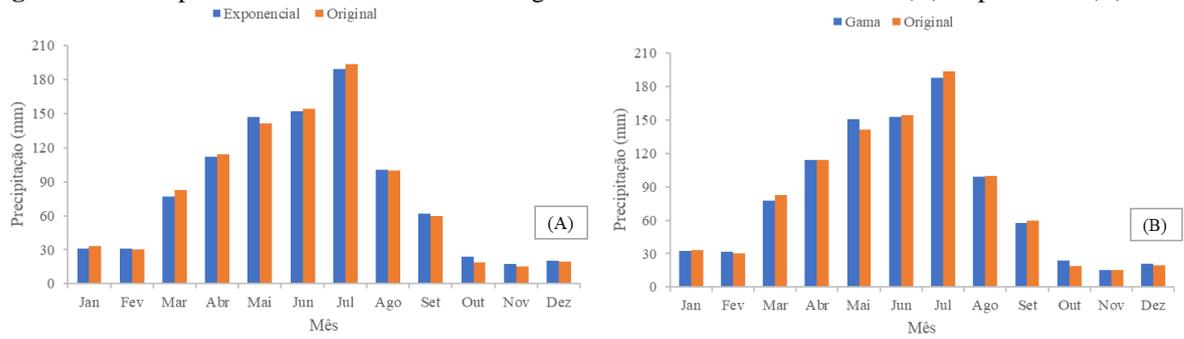
Mês	Série	Média	DP (mm)	TP (mm)	CM (mm)	DC	DS
Janeiro	Original	10,36	16,31	2258,7	155,6	218	1022
	EM	10,42	15,93	2270,81	141,26	218	1022
	Gama	10,61	11,57	2255,7	70,52	213	1027
Fevereiro	Original	11,3	12,11	2665,8	67,4	236	894
	EM	11,31	12,56	2668,76	80,28	236	894
	Gama	11,43	11,34	2654,73	69,6	232	898
Março	Original	12,32	14,36	4103,9	98,8	333	907
	EM	12,31	14,38	4153,49	106,48	337	903
	Gama	12,43	12,03	4155,58	76,66	334	906
Abril	Original	11,34	11,82	4796,9	67	423	777
	EM	11,4	11,98	4852,7	84,77	426	774
	Gama	11,42	10,73	4820,7	70,46	422	778
Maio	Original	10,67	10,92	4810,7	77,8	451	789
	EM	10,75	10,96	4806,23	78,02	447	793
	Gama	10,7	9,84	4772,26	64,97	446	794
Junho	Original	12,16	13,92	7391,8	125,5	608	592
	EM	12,21	14,01	7409,07	112,06	607	593
	Gama	12,24	11,79	7388,93	81,1	604	596
Julho	Original	10,81	12,66	6239,1	118,3	577	663
	EM	10,89	12,81	6272,59	108,62	576	664
	Gama	10,89	10,52	6219,18	73,03	571	669
Agosto	Original	7,47	8,52	3369,5	68,8	451	789
	EM	7,56	9,04	3377,27	83,28	446	794
	Gama	7,51	6,77	3337,06	44,44	444	796
Setembro	Original	7,53	8,54	1846	64,2	245	955
	EM	7,59	8,79	1841,49	68,37	242	958
	Gama	7,6	6,87	1817,29	41,43	239	961
Outubro	Original	5,21	5,42	901,1	35,6	173	1067
	EM	5,28	5,42	835,03	33,17	169	1071
	Gama	5,27	4,93	880,36	27,88	167	1073
Novembro	Original	5,49	7,15	790,5	44,9	144	1056
	EM	5,55	7,5	788	49,85	142	1058
	Gama	5,63	5,66	785,35	31,68	139	1061
Dezembro	Original	8,28	10,21	1068,6	60,5	129	1111
	EM	8,27	10,38	1065,07	64	129	1111
	Gama	8,42	8,52	1070,29	46,46	127	1113

Fonte: Autor (2021)

APÊNDICE D

Figura 17 – Comparativo das Médias Mensais Originais em Alagoinha e Sintéticas (A) Exponencial (B) Gama

Fonte: Autor (2021)

Figura 18 – Comparativo das Médias Mensais Originais em Canhotinho e Sintéticas (A) Exponencial (B) Gama

Fonte: Autor (2021)

GABRIEL MAGNO CAVALCANTE CALADO

**ANÁLISE COMPARATIVA ENTRE DISTRIBUIÇÕES DE PROBABILIDADES
PARA GERAÇÃO DE SÉRIES SINTÉTICAS DE PRECIPITAÇÕES DIÁRIAS EM
CIDADES DO AGRESTE DE PERNAMBUCO**

Trabalho de Conclusão de Curso apresentado à Coordenação do Curso de Engenharia Civil do Campus Agreste da Universidade Federal de Pernambuco – UFPE, na modalidade de artigo científico, como requisito parcial para obtenção do grau de Bacharel em Engenharia Civil. Defesa realizada por vídeoconferência.

Área de concentração: Recursos Hídricos

Aprovado em 31 de Agosto de 2021.

BANCA EXAMINADORA

Prof. Dr. Artur Paiva Coutinho (Orientador)
Universidade Federal de Pernambuco

Prof. Dr. Alessandro Romário Echevarria Antunes (Avaliador Interno)
Universidade Federal de Pernambuco

MSc. Lucas Ravellys Pyrrho de Alcântara (Avaliador Externo)
Universidade Federal de Pernambuco