



**UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE CIÊNCIAS EXATAS E DA NATUREZA**

MAYKE LUCAS DA SILVA BARROS

TRABALHO DE CONCLUSÃO DE CURSO

**ANÁLISE ESPACIAL APLICADA AO ESTUDO
DA PRECIPITAÇÃO PLUVIOMÉTRICA EM PERNAMBUCO**

RECIFE

2022

**UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE CIÊNCIAS EXATAS E DA NATUREZA
BACHARELADO EM ESTATÍSTICA**

MAYKE LUCAS DA SILVA BARROS

**ANÁLISE ESPACIAL APLICADA AO ESTUDO
DA PRECIPITAÇÃO PLUVIOMÉTRICA EM PERNAMBUCO**

TCC apresentado ao Curso de Estatística da Universidade Federal de Pernambuco, Centro de Ciências Exatas e da Natureza, como requisito para a obtenção do título de Bacharel em Estatística.

Orientador(a): Prof. Dr. Getúlio José Amorim do Amaral

**RECIFE
2022**

Ficha de identificação da obra elaborada pelo autor,
através do programa de geração automática do SIB/UFPE

Barros, Mayke Lucas da Silva.

Análise espacial aplicada ao estudo da precipitação pluviométrica em
Pernambuco / Mayke Lucas da Silva Barros. - Recife, 2022.
26 p. : il., tab.

Orientador(a): Getúlio José Amorim do Amaral

Trabalho de Conclusão de Curso (Graduação) - Universidade Federal de
Pernambuco, Centro de Ciências Exatas e da Natureza, Estatística -
Bacharelado, 2022.

1. Interpolação. 2. Precipitação. 3. Semivariância. 4. Validação cruzada. I.
Amaral, Getúlio José Amorim do. (Orientação). II. Título.

310 CDD (22.ed.)

MAYKE LUCAS DA SILVA BARROS

**ANÁLISE ESPACIAL APLICADA AO ESTUDO
DA PRECIPITAÇÃO PLUVIOMÉTRICA EM PERNAMBUCO**

TCC apresentado ao Curso de Estatística da
Universidade Federal de Pernambuco, Centro
de Ciências Exatas e da Natureza, como
requisito para a obtenção do título de Bacharel
em Estatística.

Aprovado em: 25 / 10 / 2022 .

BANCA EXAMINADORA



Documento assinado digitalmente

GETULIO JOSE AMORIM DO AMARAL

Data: 08/11/2022 17:38:22-0300

Verifique em <https://verificador.iti.br>

Profº. Dr. Getúlio José Amorim do Amaral (Orientador)

Universidade Federal de Pernambuco



Documento assinado digitalmente

CALITEIA SANTANA DE SOUSA

Data: 08/11/2022 16:49:56-0300

Verifique em <https://verificador.iti.br>

Profª. Dr. Caliteia Santana de Sousa (Examinadora)

Universidade Federal de Pernambuco



Documento assinado digitalmente

FERNANDA DE BASTIANI

Data: 08/11/2022 20:25:44-0300

Verifique em <https://verificador.iti.br>

Profª. Dr. Fernanda De Bastiani (Examinadora)

Universidade Federal de Pernambuco

AGRADECIMENTOS

Agradecer é reconhecer. E antes de tudo agradeço a Deus, por sempre estar comigo nessa jornada de vida. Agradeço pelas bênçãos que tem me dado e pelas oportunidades concedidas a cada dia. Agradeço pela sua infinita bondade e misericórdia e por ter me abençoado, me dando força, saúde e sempre estando comigo em todos os dias.

Agradeço a minha mãe Ildete pela educação que me deu, e que proporcionou viver esse momento. Agradeço a minha noiva e futura esposa Paula que por muitas vezes teve que entender a minha ausência para que as atividades e trabalhos pudessem ser concluídos. Pelo apoio que sempre me deu quando eu falava em desistir, estando ao meu lado me ajudando naquilo que podia. E quero dizer as duas “muito obrigado” por tudo, amo vocês!

A meu orientador, Professor Dr. Getúlio José Amorim do Amaral, por ter me ajudado na elaboração desse trabalho. Sou grato pela paciência, apoio e tranquilidade que teve durante todo o processo de construção do trabalho, mesmo diante dos muitos desafios que apareceram. Estendendo o agradecimento a todos os professores que tive o privilégio de aprender durante a graduação, e à Universidade Federal de Pernambuco por ter me proporcionado cursar a graduação.

Agradeço a Agência Pernambucana de Águas e Clima (APAC), por ter disponibilizado os dados da pesquisa, em especial a Aparecida Fernandes, carinhosamente Cida, no qual através dela tive a possibilidade de adquirir experiência durante o tempo que puder fazer parte da equipe APAC. Pela paciência que a mesma demonstrou e pela ajuda em cada momento.

Aos meus amigos, por sempre me apoiarem e estarem presente nessa caminhada. Que compartilharam comigo momentos bons e ruins durante esses “quatro” (um pouco mais) anos de convivência para integralização da graduação. Aos amigos que seguiram até o fim, e também aqueles seguiram outros caminhos. Grato a todos que de alguma forma estiveram envolvidos e que contribuíram para a realização desse objetivo.

“Posso todas as coisas naquele que me fortalece”

(Filipenses 4.13)

RESUMO

O objetivo do trabalho é estudar a variabilidade espacial da precipitação em Pernambuco por meio da estatística descritiva e de técnicas geoestatísticas. Os dados analisados foram de pluviosidade total anual, o volume acumulado da chuva que acontece anualmente no estado de Pernambuco. Para realizar a análise espacial foram utilizados alguns modelos para a função de semivariância: Exponencial, Esférico, Gaussiano e *Matérn*, $k=1$. Como forma de determinação do método mais adequado optou-se pela técnica de validação cruzada calculando as medidas *R* Quadrado (R^2), a Raiz Quadrática Média dos Erros (*RMSE*), e o Erro Absoluto Médio (*MAE*). Dentre os modelos considerados, o que melhor se ajustou aos dados foi o *Matérn* $k=1$, com o estimador de *Cressie e Hawkins*. A interpolação foi realizada pelo método da krigagem ordinária. As técnicas utilizadas se mostraram úteis para o mapeamento da precipitação pluviométrica e descrição da variável de interesse. Os resultados evidenciam que os municípios mais próximos do litoral são os que possuem uma média de precipitação maior se comparados com os municípios do interior do estado.

Palavras-chave: Interpolação, precipitação, semivariância, validação cruzada.

ABSTRACT

The objective of this work is to study the spatial variability of precipitation in Pernambuco through descriptive statistics and geostatistical techniques. The data analyzed were total annual rainfall, the accumulated volume of rain that occurs annually in the state of Pernambuco. To perform the special analysis, some models for the semivariance function were used: Exponential, Spherical, Gaussian and Matérn, $k=1$. As a way of determining the most appropriate method, we opted for the cross-validation technique, calculating the measures R-Squared (R^2), the Root Mean Square of Errors ($RMSE$) and the Mean Absolute Error (MAE). Among the models considered, the one that best fitted the data was the Matérn $k=1$, with the Cressie and Hawkins estimator. The interpolation was performed by the ordinary kriging method. The techniques used proved to be useful for the mapping of rainfall and description of the variable of interest. The results show that the municipalities closest to the coast are those that have a higher average rainfall compared to the municipalities in the interior of the state.

Keywords: Interpolation, precipitation, semivariance, cross-validation.

LISTA DE FIGURAS

Figura 1: Exemplo: Pluviômetro ordinário (esquerda) e totalizador (direita)	11
Figura 2: Exemplo de um semivariograma	13
Figura 3: Localização espacial das 162 estações meteorológicas.....	16
Figura 4: Histograma dos dados (esquerda) e box plot dos dados (direita).	18
Figura 5: Gráfico de dispersão em função da latitude(esquerda) e longitude(direita) ..	19
Figura 6: Post plot dos dados pluviosidade total anual (mm)	19
Figura 7: Modelos estimados com os estimadores de Matheron e Cressie e Hawkins .	17
Figura 8: Interpolação por krigagem da precipitação em Pernambuco	22
Figura 9: Variabilidade da interpolação por krigagem da precipitação em Pernambuco	22

LISTA DE TABELAS

Tabela 1: Informações sobre as 162 estações meteorológicas da base de dados...	15
Tabela 2: Modelos teóricos da função de semivariância	17
Tabela 3: Medidas para avaliação dos modelos ajustados.....	17
Tabela 4: Medidas descritivas da pluviosidade total anual (<i>mm</i>).....	18
Tabela 5: Estimativas dos parâmetros e medidas para seleção do modelo	21

SUMÁRIO

1 INTRODUÇÃO	10
2 REVISÃO DE LITERATURA	11
2.1 Climatologia e Precipitação pluviométrica	11
2.2 Pluviômetros.....	11
2.3 Semivariograma	12
3 MATERIAIS E MÉTODOS	14
4 RESULTADOS.....	17
5 CONCLUSÃO	23
REFERÊNCIAS.....	24

1. INTRODUÇÃO

Uma das preocupações da sociedade diz respeito às alterações no clima e suas consequências. Para ajudar o planejamento estratégico em determinada região é de fundamental importância compreender como funciona a dinâmica da variabilidade da precipitação pluviométrica (AYOADE, 2010). O monitoramento de precipitação é feito através de postos pluviométricos, os quais fornecem os dados de precipitação. Em lugares não registrados é necessário realizar a interpolação (MAZZINI; SCHETTINI, 2009; NOGUEIRA et al., 2020).

Para extrair uma estrutura de correlação espacial dos eventos chuvosos, podemos utilizar a abordagem estatística. No estudo da distribuição da precipitação pluviométrica, as técnicas da geoestatística têm se mostrado eficientes podendo descrever o comportamento da variável de interesse (MEDEIROS et al., 2018; MEDEIROS et al., 2019). Podemos utilizar os métodos de análise espacial para entender a dinâmica da precipitação pluviométrica, podendo criar estratégias para minimizar os prejuízos que podem acontecer devido aos eventos extremos (ARNBJERG-NIELSEN et al., 2013).

O objetivo do presente trabalho é analisar espacialmente a pluviosidade total anual no estado de Pernambuco, no período de 1991 a 2020. São mostradas algumas das técnicas da geoestatística mais comumente usadas na literatura. As técnicas incluem análises descritivas usuais e espaciais, ajuste e seleção de modelos para função de semivariância, interpolação pelo método de krigagem com o modelo selecionado para estimar a função de semivariância e apresentação do resultado da interpolação.

Conhecer como se comporta a precipitação pluviométrica é essencial para que medidas possam ser tomadas da melhor forma possível. Podemos utilizar o conhecimento desse comportamento para melhoria da drenagem da água em períodos chuvosos, da infraestrutura dos municípios com potencial de cheias, abastecimento de reservatórios para utilizar a água em períodos de seca, tanto para consumo quanto para irrigação de plantações, utilizando a água da melhor forma. Desta forma vemos a importância da análise espacial da precipitação pluviométrica.

2. REVISÃO DE LITERATURA

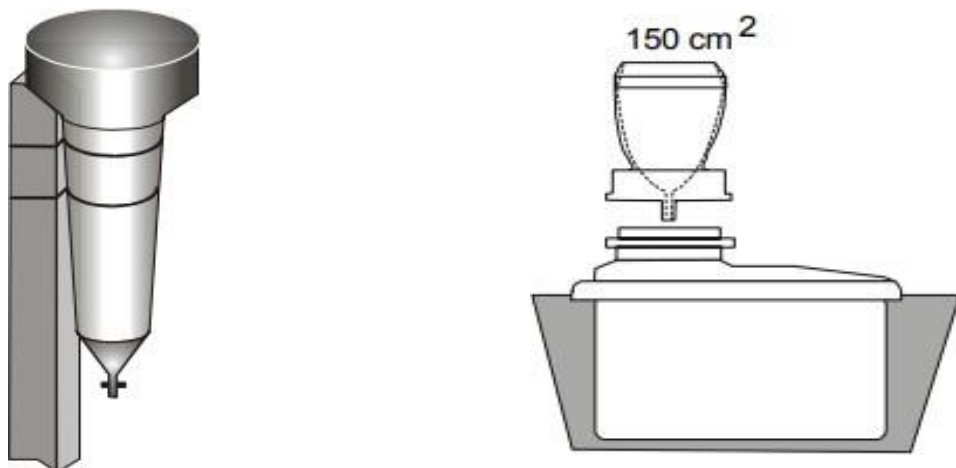
2.1. Climatologia e Precipitação pluviométrica

O estudo científico do clima e de suas mudanças é chamado de climatologia. Ele trata dos extremos e também da variabilidade dos eventos meteorológicos que acontecem durante um determinado período de tempo. Para expressar significativamente o evento meteorológico é necessário no mínimo três décadas consecutivas para um determinado local. A variabilidade da precipitação pluviométrica é um dos parâmetros estudados na climatologia, sendo esta variabilidade bastante elevada, tanto espacial quanto temporalmente. (CASTRO, 1994; FRANCISCO, 1991; SANT'ANNA NETO, 1995).

2.2. Pluviômetros

Possuindo a maior variabilidade espacial dentro do ciclo hidrológico, a precipitação depende, na maioria dos casos, de pluviômetros para sua medição (COLLISCHONN, 2006). O pluviômetro é um aparelho coletor de chuva com um reservatório interno para acumulação de água (CHEVALLIER, 2001; BERTONI & TUCCI, 2001).

Figura 1: Exemplo: Pluviômetro Ordinário (esquerda) e totalizador (direita)



Fonte: ADELMO VAREJÃO-SILVA (2005)

Na Figura 1, temos um exemplo de pluviômetro ordinário (esquerda) que normalmente possuem reservatórios capazes de acumular a precipitação pluviométrica ocorrida em 24 horas e um exemplo de pluviômetro totalizador (direita) que possuem um reservatório mais amplo, podendo acumular até vários meses (ADELMO VAREJÃO-SILVA, 2005).

2.3. Semivariograma

O semivariograma é o gráfico da distância versus a função de semivariância. A correlação espacial entre duas observações de uma variável $z(s)$ nas localizações s_1 e s_2 pode ser estimada a partir de dados observacionais. Antes, precisamos fazer suposições de estacionariedade. Uma forma comumente usada de estacionariedade é a *intrínseca*, que assume que o processo que gerou as amostras é uma função aleatória $Z(s)$ composta por uma média (constante) e um resíduo

$$Z(s) = \mu + \varepsilon(s) \quad (1)$$

E uma função de semivariância é definida como

$$\gamma(h) = \frac{1}{2}E(Z(s) - Z(s+h))^2. \quad (2)$$

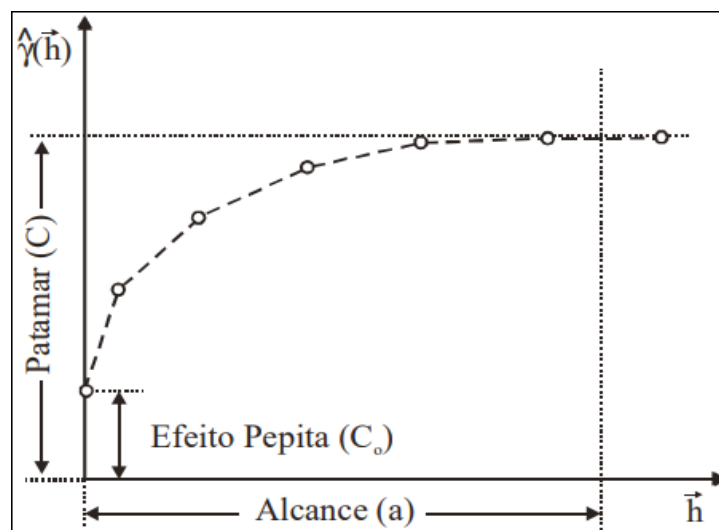
Sob essa suposição, basicamente afirmamos que a variância de Z é constante e que a correlação espacial de Z não depende da localização s , mas apenas da distância h . Então, podemos formar múltiplos pares $z(s_i)$, $z(s_j)$ que possuem vetores $\mathbf{h} = s_i - s_j$ e estimar a correlação a partir deles. Assumindo a *isotropia*, que é a independência da direção da semivariância, podemos substituir o vetor $\mathbf{h}$ pelo seu comprimento, $||h||$. O semivariograma pode ser estimado a partir de N_h pares de dados de amostra $z(s_i)$, $z(s_i + h)$ para um número de distâncias h por

$$\hat{\gamma}(h) = \frac{1}{2N_h} \sum_{i=1}^{N_h} (Z(s_i) - Z(s_i + h))^2 \quad (3)$$

Na equação 3, temos o estimador clássico de *Matheron* da função de semivariância, onde $\hat{\gamma}(h)$ é o valor do semivariograma estimado para a distância h ; $Z(s_i)$ é o valor da variável para localização s_i ; $Z(s_i + h)$ é o valor da mesma variável para alguma distância h em qualquer direção; e $N(h)$ representa o número de pares separados pelo vetor distância h . O estimador de *Matheron* não é viciado, entretanto é influenciado pela presença de valores muito diferente dos demais, valores discrepante, que são conhecidos como *outliers*. O objetivo da análise através do semivariograma é extrair uma noção de variabilidade e uma medida de correlação entre os valores em dois pontos do espaço (BIVAND, 2013). A diferença entre o Patamar (C) e o Efeito Pepita (C_0) é conhecida como Contribuição (C_1). Na Figura 02, temos um exemplo de semivariograma. Os parâmetros são:

- Alcance (a): Distância em que as amostras apresentam correlação espacial.
- Patamar (C): É o valor do semivariograma correspondente ao seu alcance, a partir do qual o semivariograma se estabiliza. A partir deste ponto, é considerado que a dependência espacial entre as amostras não existe mais.
- Efeito Pepita (C_0): Erro de medição ou de variabilidade em pequena escala. (ISAAKS E SRIVASTAVA, 1989).

Figura 2: Exemplo de um semivariograma



Fonte: CELSO; CAMARGO; CÂMARA, 2002

3. MATERIAIS E MÉTODOS

Utilizamos o software R para auxiliar na análise espacial da precipitação pluviométrica por meio de métodos geoestatísticos. Os dados foram obtidos através da APAC - Agência Pernambucana de Águas e Clima, e são referentes a precipitação medida por estações pluviométricas localizadas em alguns municípios de Pernambuco referentes ao período de 1991 a 2020. Para facilidade de visualização desconsideramos o município de Fernando de Noronha.

Nos nossos estudos utilizamos pluviosidade total anual de cada estação meteorológica nos trinta anos para realizar a análise da variabilidade da precipitação em Pernambuco. A OMM (Organização Meteorológica Mundial) recomenda que se tenha um período mínimo de trinta anos pra descrever o clima de uma região. O período a ser usado inicia-se com anos terminado em um, por exemplo: 1981, 1991, etc. (WORLD METEOROLOGICAL ORGANIZATION, 2015).

As pesquisas climáticas geram uma dificuldade em relação a obtenção de dados sem falhas ou que contenham poucas falhas. Os dados meteorológicos estão suscetíveis a problemas de dados faltantes devido à falta do observador, uma medição incorreta ou mesmo o não funcionamento do instrumento de medição. Para variáveis de valores acumulados a OMM recomenda que seja utilizado apenas os meses com informações completas. Seja X_{ij} a soma de todos os valores disponíveis para o mês i , do ano j , temos que X_{ij} é dado por:

$$X_{ij} = \sum_1^k X_{ijk} \quad (4)$$

onde k é o número de dias referente ao mês i , do ano j . Neste caso, k pode assumir 28, 29, 30 ou 31 dias (INMET, 2022). Inicialmente calculamos a acumulada para cada mês e ano. Seguimos calculando a média das acumuladas para cada um dos doze meses do ano. Por fim, calculamos a acumulada dessas médias para cada estação meteorológica formando assim a base de dados que utilizamos neste trabalho. A variável de interesse representa pluviosidade total anual. Na Tabela 1, temos as informações da longitude, latitude e nome de cada estação meteorológica.

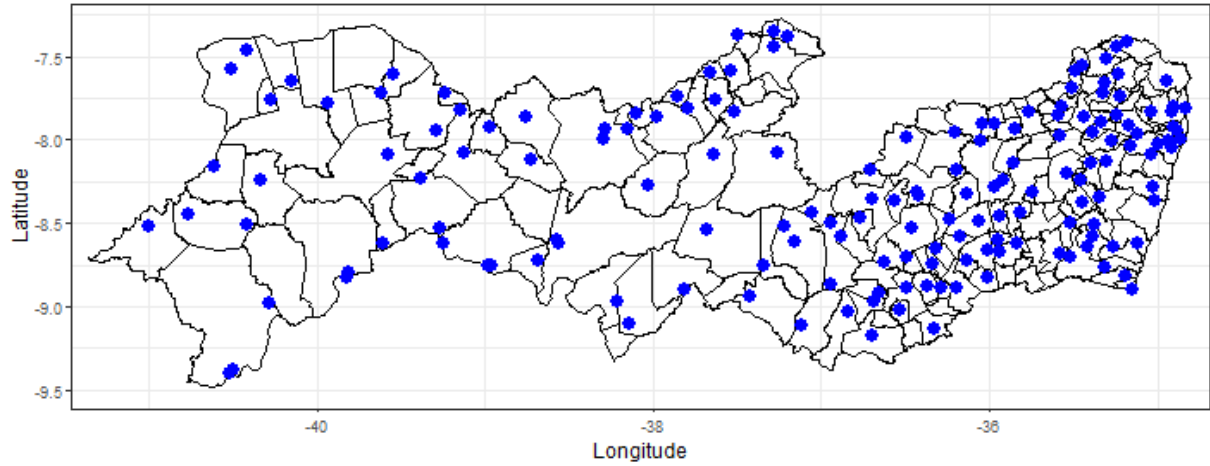
Tabela 1: Informações sobre as 162 estações meteorológicas da base de dados

Longitude	Latitude	Posto	Longitude	Latitude	Posto	Longitude	Latitude	Posto
-41,0050	-8,5167	Afrânio	-34,9000	-7,9281	Abreu e Lima	-37,6331	-7,7569	Afogados da Ingazeira
-35,9419	-8,4583	Agrestina	-35,5219	-8,7025	Água Preta	-37,0556	-8,4336	Arcoverde (INMET)
-36,7739	-8,4661	Alagoinha	-37,1197	-9,1119	Águas Belas	-38,9667	-8,7500	Belém de São Francisco (CHESF)
-35,2306	-7,6042	Aliança	-40,5058	-7,5767	Araripina	-39,2500	-8,6167	Belém de São Francisco (Ibó - CHESF)
-36,0597	-8,4906	Altinho	-40,4172	-7,4583	Araripina - PCD	-38,9822	-8,7583	Belém de São Francisco (IPA)
-35,4472	-8,3778	Amaraji	-35,1903	-8,8125	Barreiros	-36,4253	-8,3367	Belo Jardim
-36,2861	-8,8883	Angelim	-35,8406	-8,6222	Belém de Maria	-36,4267	-8,3147	Belo Jardim (Açude Bituri)
-38,0331	-8,2708	Betânia	-35,5678	-7,8017	Bom Jardim	-36,6928	-9,1725	Bom Conselho (IPA)
-39,9375	-7,7806	Bodocó	-36,5356	-9,0161	Brejão - PCD	-36,1942	-8,1819	Brejo da Madre de Deus (Fazenda Nova)
-37,2847	-7,3478	Brejinho	-35,3272	-7,7228	Buenos Aires	-35,0289	-8,2858	Cabo
-37,1547	-8,6147	Buíque	-36,2372	-8,4806	Cachoeirinha	-35,0150	-8,3678	Cabo (Barragem de Suape)
-39,2681	-8,5278	Cabrobó	-34,9922	-8,0217	Camaragibe	-35,7417	-8,3175	Camocim de São Felix
-36,3344	-8,7400	Calçados	-36,1936	-8,8825	Canhotinho	-35,2408	-7,8511	Carpina (Est. Exp. de Cana-de-Açúcar)
-38,1539	-7,9394	Calumbi	-36,6267	-8,7342	Capoeiras	-40,7672	-8,4494	Dormentes
-37,7981	-7,8078	Carnaíba	-35,9678	-8,2794	Caruaru	-40,4147	-8,5106	Dormentes (Lagoas)
-39,2361	-7,7211	Cedro	-35,9158	-8,2383	Caruaru (IPA)	-35,8469	-7,9347	Frei Miguelinho (IPA)
-35,5144	-8,5028	Cortês	-35,4592	-8,2422	Chã Grande	-34,9489	-7,6442	Goiana (Itapirema - IPA)
-35,9511	-8,6039	Cupira	-36,3283	-9,1286	Correntes	-34,9150	-7,8281	Igarassu
-37,6450	-8,0864	Custódia	-35,3886	-7,9556	Feira Nova	-35,0336	-7,8364	Igarassu (Bar.Catucá)
-37,9739	-7,8669	Flores	-35,2386	-7,4461	Ferreiros	-35,0386	-8,0922	Jaboatão dos Guararapes (Bar.Duas Unas)
-39,6158	-7,7172	Granito	-38,5667	-8,6167	Floresta (CHESF)	-35,3333	-7,8967	Lagoa de Itaenga (Barragem de Carpina)
-35,5431	-8,2006	Gravatá	-38,5756	-8,6014	Floresta (IPA)	-40,2833	-8,9833	Lagoa Grande (IPA)
-36,8439	-9,0350	Iati	-35,3878	-8,5817	Gameleira	-35,1669	-7,9178	Paudalho
-37,5128	-7,8325	Iguaraci	-36,4878	-8,8833	Garanhuns	-35,1158	-7,9686	Paudalho (Barragem de Goitá)
-40,1469	-7,6522	Ipupi	-35,2728	-8,0111	Glória do Goitá	-36,9450	-8,5014	Pedra
-37,4236	-8,9347	Itaíba	-37,6786	-8,5378	Ibimirim (IPA)	-36,9428	-8,8625	Pedra (São Pedro do Cordeiro)
-37,2006	-7,3767	Itapetim	-36,1778	-8,5828	Ibirajuba	-40,5175	-9,3992	Petrolina
-36,4953	-7,9819	Jataúba	-37,8167	-8,9000	Inajá (CHESF)	-40,5000	-9,3833	Petrolina (INMET)
-36,4928	-8,7067	Jucati	-38,6864	-8,7242	Itacuruba	-34,9353	-8,0014	Recife (Alto da Brasileira)
-36,1369	-8,7181	Jurema	-34,8378	-7,8086	Itamaracá	-34,9167	-8,0500	Recife (Várzea)
-36,3178	-8,6556	Lajedo	-35,1786	-7,4075	Itambé (IPA)	-35,8592	-8,1381	Riacho das Almas
-35,4347	-7,8658	Limoeiro	-34,9058	-7,8003	Itapissuma	-35,3769	-8,5100	Ribeirão (Fazenda Capri)
-35,5131	-7,6875	Machados	-35,5886	-7,8539	João Alfredo	-35,2683	-8,6397	Rio Formoso (Usina Cucaú)
-34,8589	-7,9989	Olinda	-35,4167	-8,6417	Joaquim Nabuco	-40,3369	-8,2406	Santa Cruz da Venerada
-39,6031	-8,6183	Orocó	-35,4497	-7,5522	Macaparana	-36,2036	-7,9533	Santa Cruz do Capibaribe
-35,5797	-8,6783	Palmares	-38,7272	-8,1158	Mirandiba	-39,8228	-8,8219	Santa Maria da Boa Vista
-36,0078	-8,6633	Panelas	-39,5500	-7,6075	Moreilândia	-39,8167	-8,8000	Santa Maria da Boa Vista (CHESF)
-35,5836	-7,9761	Passira	-35,2228	-7,7408	Nazaré da Mata	-35,9389	-8,6689	São Benedito do Sul
-34,8844	-7,9425	Paulista	-36,6583	-8,9192	Paranatama	-36,4594	-8,5272	São Bento do Una (IPA)
-36,7053	-8,1836	Poção	-39,5775	-8,0889	Parnamirim	-35,8133	-8,4328	São Joaquim do Monte
-35,3961	-8,1386	Pombos	-36,6972	-8,3531	Pesqueira	-35,1497	-8,8958	São José da Coroa Grande
-36,0119	-8,8269	Quipapá	-38,2164	-8,9742	Petrolândia	-38,7619	-7,8639	São José do Belmonte
-37,8586	-7,7406	Quixaba	-35,3475	-8,3483	Primavera	-37,2806	-7,4440	São José do Egito (Faz. Muquén)
-36,6917	-8,9694	Saloá	-39,1297	-8,0733	Salgueiro	-35,1653	-8,0403	São Lourenço da Mata (Tapacurá)
-36,5664	-8,3639	Sanharó	-40,6133	-8,1594	Santa Filomena	-35,4889	-7,5875	São Vicente Férrer
-36,3667	-8,8769	São João	-37,4997	-7,3747	Santa Terezinha	-38,2936	-7,9931	Serra Talhada
-37,6581	-7,6003	Solidão	-36,1375	-8,3283	São Caetano	-38,2889	-7,9306	Serra Talhada (IPA)
-35,7628	-7,8369	Surubim	-35,1194	-8,6194	Sirinhaém	-39,2936	-7,9492	Serrita
-37,5369	-7,5883	Tabira	-35,3142	-8,7656	Tamandaré	-39,1486	-7,8197	Serrita (Santa Rosa)
-35,3058	-7,5144	Timbaúba	-39,3772	-8,2272	Terra Nova	-37,2639	-8,0750	Sertânia
-36,0564	-8,0106	Toritama	-37,3422	-8,7547	Tupanatinga	-37,2244	-8,5144	Sertânia (Cedoca)
-40,2761	-7,7597	Trindade	-36,8806	-8,5747	Venturosa	-38,1461	-9,1058	Tacaratu (Sítio Gameleira)
-38,1061	-7,8372	Triunfo	-38,9694	-7,9286	Verdejante	-36,0383	-7,9047	Taquaritinga do Norte
-35,3183	-7,6536	Vicência	-35,9739	-7,9047	Vertentes - PCD	-35,3028	-8,1283	Vitória de Santo Antão (IPA)

Fonte: Elaboração do autor

Na Figura 3, podemos visualizar o contorno do estado de Pernambuco e seus municípios e também a distribuição espacial das 162 estações meteorológicas contidas na nossa base de dados. Podemos observar que a maioria das estações meteorológicas estão no litoral do estado. Para a análise descritivas, usamos algumas estatísticas básicas, tais como: média, mediana, máximo, mínimo e desvio padrão.

Figura 3: Localização espacial das 162 estações meteorológicas



Fonte: Elaboração do autor

Utilizamos o histograma, boxplot, postplot, para visualizar os dados e realizar uma análise descritiva dos dados. Com a utilização de um gráfico de dispersão dos dados em relação a longitude e latitude, podemos identificar possíveis tendências nos dados. Como alternativa ao estimador clássico de *Matheron*, podemos utilizar o estimador de *Cressie* e *Hawkins*, dado por:

$$\hat{\gamma}(h) = \left[\frac{1}{N_h} \sum_{i=1}^{N_h} |Z(s_i) - Z(s_i + h)|^2 \right]^{\frac{1}{2}} / 0,914 + \frac{0,988}{N_h}. \quad (5)$$

O estimador de *Cressie* e *Hawkins* é mais robusto à presença de *outliers*. (CRESSIE, 1993). Os parâmetros são os mesmos do estimador clássico de *Matheron*. Os modelos são ajustados por meio da escolha adequada dos valores dos parâmetros: Alcance (α), Efeito Pepita (C_0) e Contribuição (C_1). Utilizou-se o método de mínimos quadrados ponderados para o ajuste dos modelos, com ponderação de “*cressie*” utilizada no argumento *weights*, da função *geoR::variofit()* no software *R*. Os modelos teóricos ajustados foram: Exponencial, Esférico, Gaussiano e *Matérn*, $k = 1$. Esses modelos são apresentados na Tabela 2.

Tabela 2: Modelos teóricos da função de semivariância

Exponencial	$C_0 + C_1 \left[1 - e^{-\left(\frac{h}{a}\right)} \right] , \quad h > 0;$ $0 , \quad h = 0$
Esférico	$C_0 + C_1 \left[\frac{3}{2} \left(\frac{h}{a} \right) - \frac{1}{2} \left(\frac{h}{a} \right)^3 \right] , \quad 0 < h < a;$ $0 , \quad h = 0$ $C_0 + C_1 , \quad h > a$
Gaussiano	$C_0 + C_1 \left[1 - e^{-\left(\frac{h}{a}\right)^2} \right] , \quad h > 0;$ $0 , \quad h = 0$
Matérn	$C_0 + C_1 \left[1 - (2^{k-1} \Gamma(k))^{-1} \left(\frac{h}{a} \right)^k K_k \left(\frac{h}{a} \right) \right] , \quad h > 0;$ $0 , \quad h = 0$

Fonte: Elaboração do autor

As medidas utilizadas com o resultado da validação cruzada para seleção do modelo teórico de semivariograma foram: O R Quadrado (R^2), a Raiz Quadrática Média dos Erros ($RMSE$), e o Erro Absoluto Médio (MAE) como critério para seleção do modelo. Essas medidas são apresentadas na Tabela 3. Após a escolha do modelo, realizamos a interpolação em áreas sem informação através do método da krigagem ordinária.

Tabela 3: Medidas para avaliação dos modelos ajustados

R^2	$1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y}_i)^2}$
$RSME$	$\sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$
MAE	$\frac{1}{n} \sum_{i=1}^n y_i - \hat{y}_i $

Fonte: Elaboração do autor

4. RESULTADOS

A visualização de dados por meio de gráficos como histograma, post-plot, boxplot, é certamente um componente importante e necessário da análise exploratória de dados espaciais. Na Tabela 4, vemos algumas medidas descritivas dos dados. Podemos observar que a quantidade de pluviosidade total anual varia de 247,4 *mm* até 2155,6 *mm*. Também é possível observar que 75% dos municípios avaliados possuem pluviosidade total anual acima de 1022,8 *mm*.

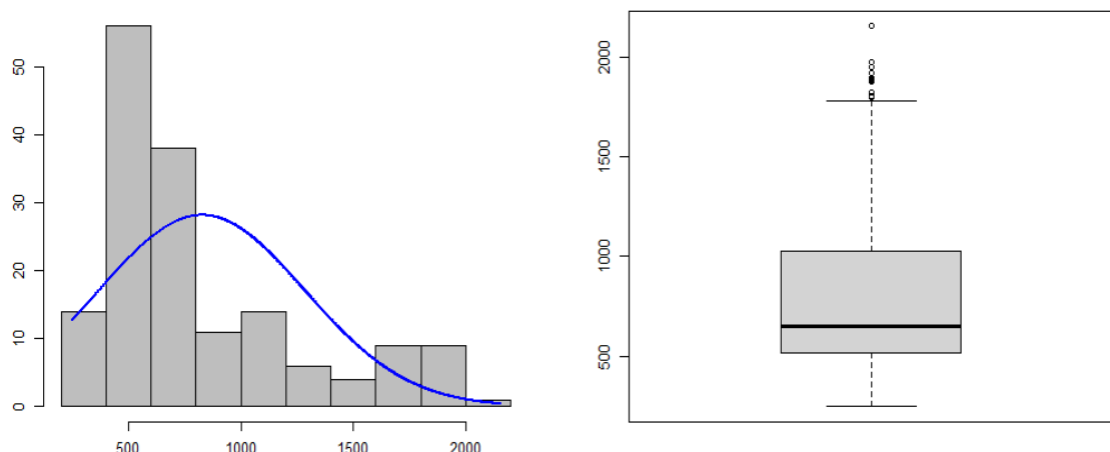
Tabela 4: Medidas descritivas da pluviosidade total anual (*mm*)

Mínimo	1º Quartil	Mediana	Média	3º Quartil	Máximo	Desvio Padrão
247,4	519,3	648,7	826,3	1022,8	2155,6	458,96

Fonte: Elaboração do autor

Na Figura 4, podemos ver o histograma dos dados (esquerda) e o box plot (direita). Podemos observar no histograma uma concentração de dados entre 400 *mm* e 800 *mm*. Visualmente os dados não aparentam seguir uma distribuição normal. Realizamos testes para verificar a normalidade dos dados. Os testes realizados através do *software R* foram: *Shapiro-Wilk*, *Kolmogorov-Smirnov*, *Jarque Bera* e *Lilliefors*. Todos os valores *p* dos testes foram abaixo de 0,05 indicando que nossos dados não seguem uma distribuição normal. No box plot observamos a presença de *outliers*. A estação meteorológica que apresentou a maior pluviosidade total anual foi Recife (Várzea) com 2155,6 *mm*.

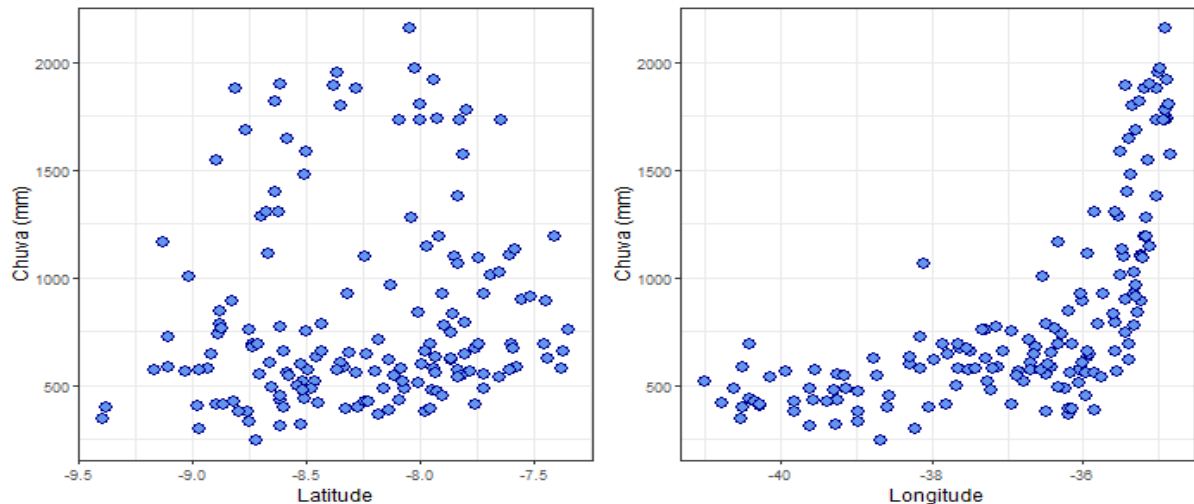
Figura 4: Histograma dos dados (esquerda) e box plot dos dados (direita).



Fonte: Elaboração do autor

Na Figura 5, observamos que os dados em relação à latitude (esquerda) aparentemente não possuem uma tendência. Já em relação à longitude (direita) apresentam uma tendência, indicando que quanto “maior” a longitude, a quantidade de precipitação pluviométrica também aumenta. No nosso caso, o “aumento” da longitude indica que estamos nos aproximando do litoral de Pernambuco.

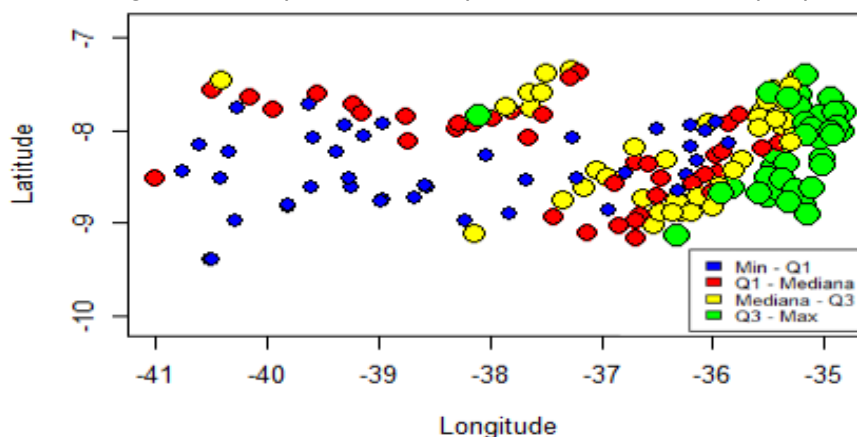
Figura 5: Gráfico de dispersão em função da latitude(esquerda) e longitude(direita)



Fonte: Elaboração do autor

Na Figura 6, temos o gráfico de post plot onde vemos uma concentração maior no litoral de valores acima do terceiro quartil. A observação acima do terceiro quartil mais distante das demais representa o município de Triunfo, onde possui pluviosidade total anual de 1067,4 mm. No interior do estado podemos observar uma maior concentração de baixa precipitação pluviométrica, sendo inferior a 520 mm.

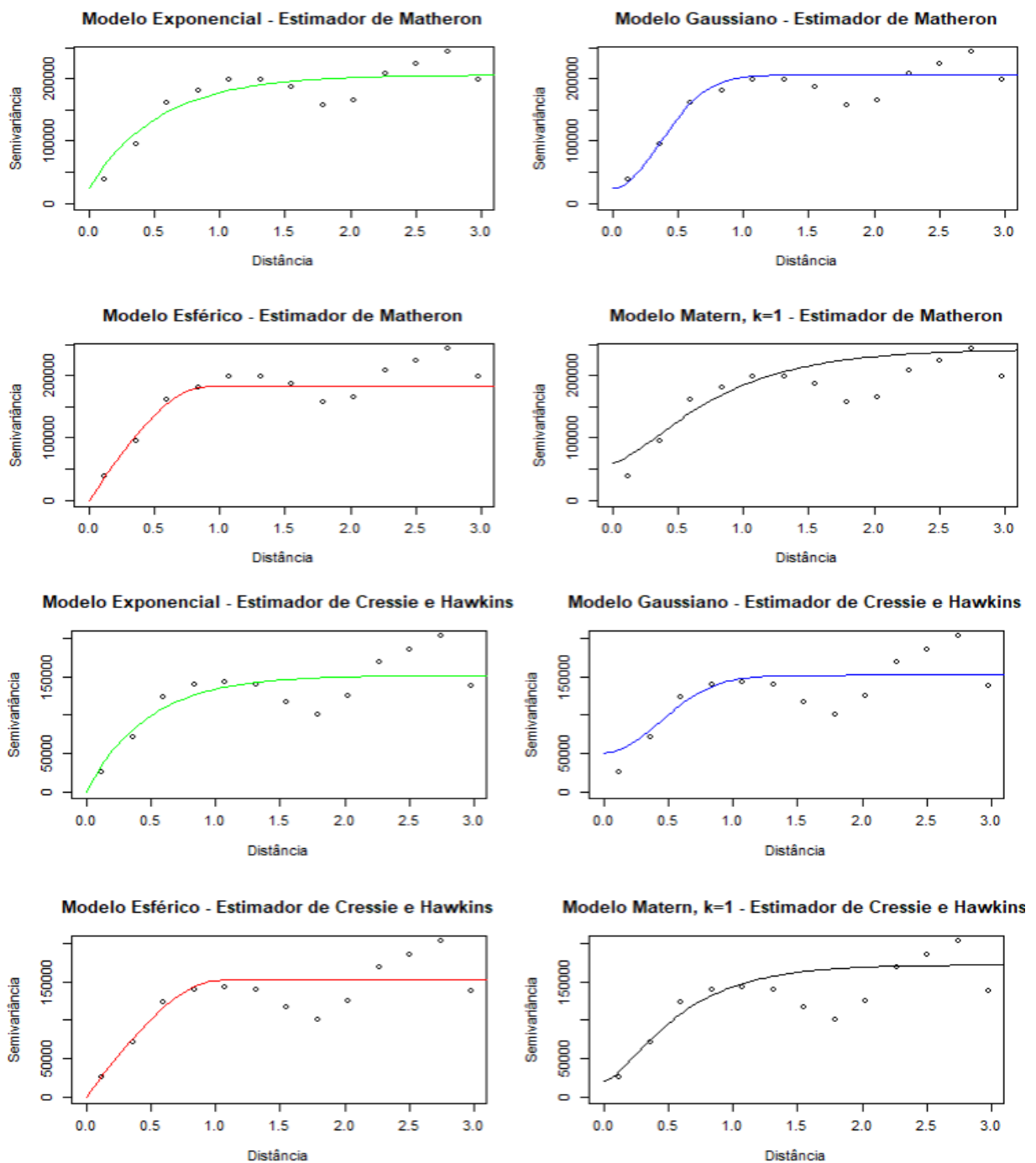
Figura 6: Post plot dos dados pluviosidade total anual (mm)



Fonte: Elaboração do autor

Na Figura 7, temos o semivariograma dos dados e o ajuste de cada modelo proposto, utilizamos o método de mínimos quadrados ponderados para o ajuste dos modelos, com ponderação de “*cressie*” no argumento *weights*, da função *variofit()* do pacote *geoR* no *software R*. O estimador utilizado foi o de *Cressie* e *Hawkins*, com um número mínimo de 30 pares para obter cada estimativa e um ponto de corte de 50% da distância máxima entre os pontos.

Figura 7: Modelos estimados com os estimadores de Matheron e Cressie e Hawkins



Fonte: Elaboração do autor

Podemos observar na Tabela 5 as estimativas dos parâmetros dos modelos propostos. O Alcance (a) no modelo *Matérn*, $k=1$ foi o maior dentre os modelos, independente do estimador utilizado. No caso da Contribuição (C_1), com o estimador de *Matheron* todos os modelos apresentaram o mesmo resultado. Os Efeitos Pepita (C_0) mostram que existe descontinuidade para distâncias menores do que a mínima distância entre os dados.

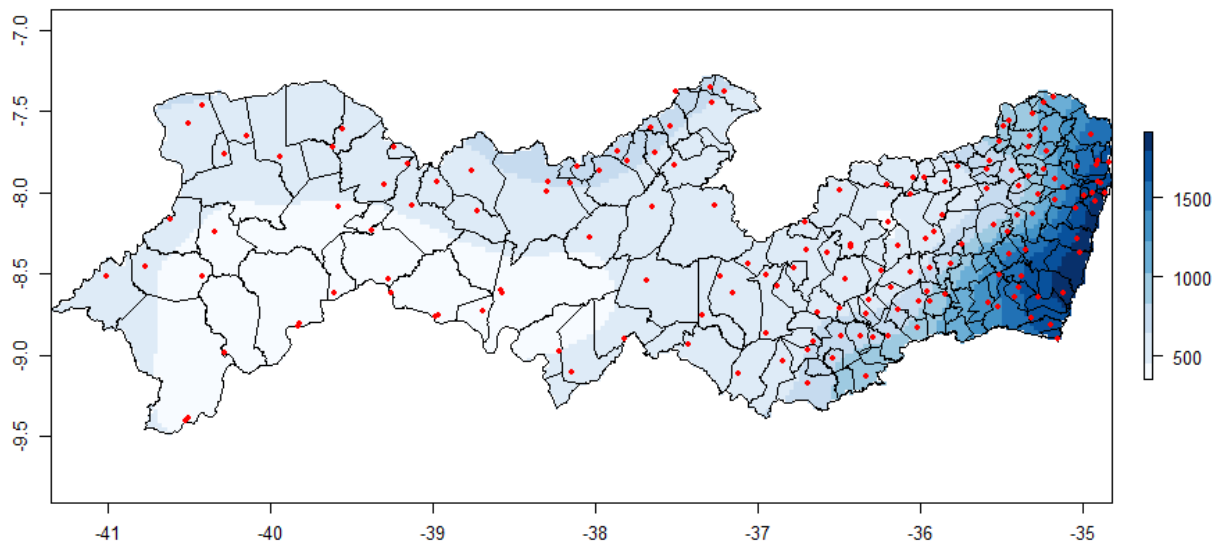
Tabela 5: Estimativas dos parâmetros e resultado das medidas para seleção do modelo

Modelo	Estimador	$\hat{C}_0$	$\hat{C}_1$	$\hat{a}$	R^2	$RSME$	MAE
Esférico	<i>Matheron</i>	0.0006	181927.1	0.898	0.895	148.4	111.1
Exponencial	<i>Matheron</i>	24256.9	181927.1	1.611	0.899	145.7	106.4
Gaussiano	<i>Matheron</i>	24256.9	181927.1	0.878	0.887	153.7	112.8
Matérn $k = 1,0$	<i>Matheron</i>	60642.4	181927.1	2.174	0.892	150.3	107.1
Esférico	<i>Cressie e Hawkins</i>	0	151978.2	1.0389	0.895	147.9	112.3
Exponencial	<i>Cressie e Hawkins</i>	2.5865	151978.2	1.4142	0.897	146.8	109.7
Gaussiano	<i>Cressie e Hawkins</i>	50659.4	101318.8	1.0475	0.881	158.1	112.6
Matérn $k = 1,0$	<i>Cressie e Hawkins</i>	20263.8	151978.2	1.6074	0.900	144.5	105.9

Fonte: Elaboração do autor

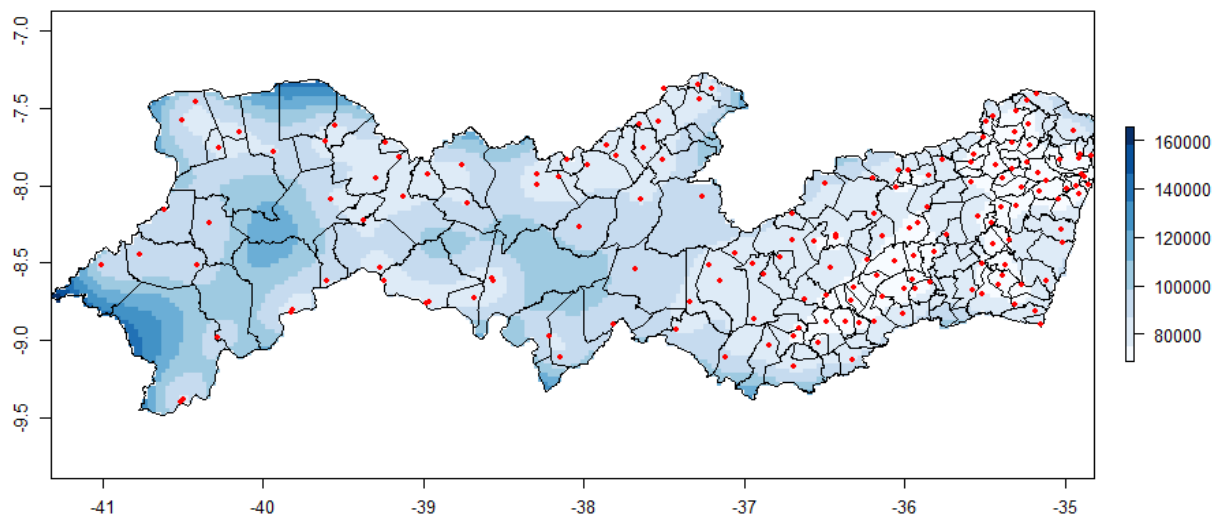
Ao compararmos as medidas usadas para seleção de modelos podemos perceber que quando foi utilizado o estimador de *Matheron*, o modelo exponencial foi o que apresentou o menor $RSME$ e MAE , e o maior R^2 . Utilizando o estimador de *Cressie e Hawkins*, o modelo que melhor se ajustou ao semivariograma foi o *Matérn* $k = 1,0$. Desta forma, para escolha do modelo a ser usado para interpolação por krigagem ordinária, optamos pelo modelo *Matérn* $k = 1,0$ com o estimador de *Cressie e Hawkins*, por ser mais robusto diante da presença de *outliers*.

Na Figura 8, temos o resultado da interpolação por krigagem ordinária. A cores mais escuras no mapa representam os lugares onde a pluviosidade total anual é maior. Os pontos vermelhos são as localizações das estações pluviométricas da nossa base de dados. Conforme as estimativas, nota-se que a partir do litoral, os níveis da precipitação pluviométrica vão decrescendo em direção ao interior do estado. Podemos observar que na mesorregião de São Francisco temos uma extensa área com baixa pluviosidade total anual.

Figura 8: Interpolação por krigagem da precipitação em Pernambuco

Fonte: Elaboração do autor

Na Figura 9, temos a variância da interpolação por krigagem ordinária, os lugares com mais “certeza” são os mais claros. Os pontos em vermelho mostram a localização das estações meteorológicas. Quanto mais distante dos nossos pontos de amostragem, maior fica a incerteza da interpolação. Podemos observar que na região onde temos uma maior concentração de estações meteorológicas a variância é pequena. Podemos observar que o modelo *Matérn* $k=1$, com o estimador de *Cressie e Hawkins*, se ajustou bem aos dados, tornando boa a interpolação realizada. Com as informações obtidas, podemos ver que a precipitação pluviométrica em Pernambuco tem sua maior ocorrência no litoral do estado em comparação ao interior.

Figura 9: Variabilidade da interpolação por krigagem da precipitação em Pernambuco

Fonte: Elaboração do autor

5. CONCLUSÃO

A utilização de técnicas de geoestatística apresentou resultados satisfatórios no mapeamento da pluviosidade total anual de 1991 a 2020 no estado de Pernambuco, que apresentaram variações significativas espacialmente. O modelo *Matérn*, $k=1$, com o estimador de *Cressie e Hawkins* foi o que melhor se ajustou aos dados pluviométricos, com base no critério de escolha utilizado. A interpolação pelo método de krigagem conseguiu representar bem a variabilidade da precipitação pluviométrica.

Vale ressaltar que em Pernambuco temos uma maior quantidade de estações disponíveis da APAC (Agência Pernambucana de Águas e Clima) do que as usadas neste trabalho. Todavia algumas apresentaram dados faltantes. A utilização de estações pluviométricas com muitos dados faltantes na análise estatística espacial pode gerar algumas consequências, como uma representação inadequada, ou imprecisão nos mapas de interpolação, maiores erros padrões das estimativas e problemas nos ajustes dos modelos.

Para trabalhos futuros, uma solução para utilizar uma quantidade maior de estações pluviométricas seria o emprego de técnicas estatísticas para imputação para dados faltantes nas estações e posteriormente a utilização das mesmas para realização da análise espacial. Outra alternativa é visualizar onde as estações meteorológicas estão localizadas, e verificar a possibilidade de instalação de uma nova estação meteorológica nos locais que não tem estação ou realizar a manutenção regular de uma estação meteorológica próxima, caso já exista.

Na análise espacial podem ser utilizados outros métodos de escolha do modelo, utilização de outros modelos não abordados neste trabalho. Pode verificar mais detalhadamente como os dados se comportam, acrescentar a parte temporal ou verificar como a precipitação se comporta ao longo do ano, analisando os dados mês a mês. Para a interpolação pode ser utilizados outros métodos para comparação dos resultados ou até mesmo variações do próprio método de interpolação de krigagem. Visando sempre na melhoria dos resultados.

REFERÊNCIAS

- ADELMO VAREJÃO-SILVA, M. **METEOROLOGIA E CLIMATOLOGIA** Recife, Pernambuco. 2005.
- ARNBJERG-NIELSEN, K., WILLEMS, P. OLSSON, J. BEECHAM, S., PATHIRANA, A., BULOW GREGERSEN, I., MADSEN, H., NGUYEN, V. **Impacts of climate change on rainfall extremes and urban drainage systems: a review**. Water Science & Technology. v.68 , p.16-28, 2013.
- AYOADE, J. O., 2010. **Introdução à climatologia para os trópicos**, 14 ed. Bertrand Brasil, Rio de Janeiro.
- BERTONI, C.; TUCCI, E.M. (2001). **“Precipitação”**, in **Hidrologia, Ciência e Aplicação**, ABRH, ed. UFRGS, Porto Alegre – RS. 2ª ed
- BIVAND, R. S.; PEBESMA, E.; GÓMEZ-RUBIO, V. **Applied Spatial Data Analysis with R: Second Edition**. [s.l: s.n.].
- CELSO, E.; CAMARGO, G.; CÂMARA, G. **A análise espacial de superfícies**. 2002.
- CRESSIE, N.A.C (1993) **Statistics for Spatial Data**. New York: Wiley.
- CASTRO, R. **Distribuição probabilística da frequência de precipitação na região de Botucatu, SP**. Botucatu. 101p. Dissertação (Mestre em Agronomia) – Faculdade de Ciências Agronômicas, Universidade Estadual Paulista, 1994.
- CHEVALLIER, P. (2001). **“Aquisição e Processamento de Dados”**, in **Hidrologia, Ciência e Aplicação**, ABRH, ed. UFRGS, Porto Alegre – RS. 2ª ed.
- COLLISCHONN, B., 2006. **Uso de precipitação estimada pelo satélite TRMM em modelo hidrológico distribuído**. Dissertação de Mestrado. Porto Alegre, Universidade Federal do Rio Grande do Sul.
- FRANCISCO, J. D. **Parâmetros pluviométricos auxiliares no planejamento de empreendimentos na região de Botucatu, SP**. Botucatu. 120p. Dissertação (Mestre em Agronomia) – Faculdade de Ciências Agronômicas, Universidade Estadual Paulista. 1991.
- INMET. Normais Climatológicas 1991-2020. **Instituto Nacional de Meteorologia - INMET**, p. 27, 2022.

ISAAKS, E. H.; SRIVASTAVA, R. M. **An introduction to applied geostatistics**. New York: Oxford University Press, 1989.

MAZZINI, P. L. F.; SCHETTINI, C. A. F. **Avaliação de metodologias de interpolação espacial aplicadas a dados hidrográficos costeiros quase-sinóticos**. Brazilian Journal of Aquatic Science and Technology, Itajaí, SC, v. 1, n. 13, p. 53-64, 2009.

MEDEIROS, E. S.; LIMA, R. R.; OLINDA, R. A.; SANTOS, C. A. C.. **Modeling Spatiotemporal Rainfall Variability in Paraíba, Brazil**. Water, v.11, n.9, p.1843, 2019. DOI: <https://doi.org/10.3390/w11091843>

MEDEIROS, E. S.; OLINDA, R. A.. **Probability maps for the rainy quarter in the State of Paraíba**. Revista Univap, v.25, n. 46, p.19-29, 2018. DOI: <http://dx.doi.org/10.18066/revistaunivap.v24i46.1902>

NOGUEIRA, DANILO & SILVA, ALEXSANDRO & SILVA, ANA. (2020). **Comparação entre métodos de interpolação espacial para a estimativa da distribuição de precipitação no ceará-brasil**. IRRIGA. 25. 131-142. 10.15809/irriga.2020v25n1p131-142.

SANT'ANNA NETO, J. H. **As chuvas no Estado de São Paulo: contribuição ao estudo da variabilidade e tendência da pluviosidade na perspectiva da análise geográfica**. 1995. Tese (Doutorado em Geografia) - Universidade de São Paulo.

WORLD METEOROLOGICAL ORGANIZATION. **New Two-Tier approach on “climate normals”**. Disponível em: <https://public.wmo.int/en/media/news/new-two-tier-approach-%E2%80%9Cclimate-normals%E2%80%9D#:~:text=Climate%20normals%20are%20presently%20updated,climate%20much%20faster%20than%20before>. Acesso em: 9 de out. 2022