

Uso de metodologias de preenchimento de falhas para estimativas de dados de precipitação

Use of gap filling methodologies to estimate rainfall data

Uso de metodologías de llenado de fallas para estimar los datos de precipitación

Recebido: 23/04/2021 | Revisado: 01/05/2021 | Aceito: 04/05/2021 | Publicado: 17/05/2021

Allana Siqueira Dias

ORCID: <https://orcid.org/0000-0002-2851-4885>

Universidade de Pernambuco, Brasil

E-mail: allanadias123@gmail.com

Willames de Albuquerque Soares

ORCID: <https://orcid.org/0000-0003-3268-7241>

Universidade de Pernambuco, Brasil

E-mail: was@poli.br

Resumo

No presente trabalho, as metodologias para preenchimento de falhas em séries temporais de precipitação foram utilizadas, não como ferramenta auxiliar no preenchimento de lacunas provocadas por erros, mas sim como um meio de prever integralmente os dados de pluviometria para o município de Catanduva, baseando-se em dados dos municípios de São Carlos (SC), Franca (FR) e Votuporanga (VT), localizados no estado de São Paulo. Objetivando assim, analisar e avaliar a viabilidade do uso dos métodos da ponderação regional, da regressão linear simples e da regressão linear múltipla, verificando o desempenho e a correlação entre os dados estimados e os dados reais (registrados pelo posto pluviométrico de Catanduva). Foram utilizadas informações de séries históricas com 30 anos de observações mensais fornecidas pelo Instituto Nacional de Meteorologia (INMET) para a aplicação dos métodos citados. Comparando os dados estimados aos dados medidos pode-se afirmar que o método da regressão linear múltipla foi o que melhor simulou a realidade dos eventos observados e que as estimativas têm melhor desempenho para os meses menos chuvosos e quando são utilizados dados pluviométricos de regiões próximas.

Palavras-chave: Ponderação regional; Regressão linear; Série histórica.

Abstract

In the following paper, gap filling methodologies of historical series of rainfall data were applied, not as an auxiliary tool to fill gaps provoked by mistakes, but as a way to fully predict rainfall data for the city of Catanduva, based on data of the cities of São Carlos (SC), Franca (FR) e Votuporanga (VT), located in São Paulo state. Aiming to analyze and evaluate the viability of using the methods of regional weighting, simple linear regression e multiple linear regression, verifying the performance and correlation between estimated and real data (registered by Catanduva's pluviometric station). There were used historical series information with 30 years of monthly observations provided by the Instituto Nacional de Meteorologia (INMET) to apply the methods mentioned. Comparing the estimated with the measured data, it is possible to affirm that the method of multiple linear regression was the one that best simulated the reality of the observed events and the estimations have a better performance for driest months and when there are used rainfall data from nearby regions.

Keywords: Regional weighting; Linear regression; Historical series.

Resumen

En el presente trabajo, se utilizaron las metodologías para llenar vacíos en series de tiempo de precipitación, no como una herramienta auxiliar para llenar los vacíos causados por errores, sino como un medio para predecir completamente los datos de precipitación para el municipio de Catanduva, con base en sobre datos de los municipios de São Carlos (SC), Franca (FR) y Votuporanga (VT), ubicados en el estado de São Paulo. Así, con el objetivo de analizar y evaluar la viabilidad de utilizar los métodos de ponderación regional, regresión lineal simple y regresión lineal múltiple, verificando el desempeño y la correlación entre los datos estimados y los datos reales (registrados por la estación de precipitación de Catanduva). Para la aplicación de los métodos mencionados se utilizó información de series históricas con 30 años de observaciones mensuales proporcionadas por el Instituto Nacional de Meteorología (INMET). Comparando los datos estimados con los datos medidos, se puede decir que el método de regresión lineal múltiple fue el que mejor simuló la realidad de los eventos observados y que las estimaciones funcionan mejor para los meses menos lluviosos y cuando los datos de precipitación de regiones cercanas son usó.

Palabras clave: Ponderación regional; Regresión lineal; Series históricas.

1. Introdução

A obtenção e interpretação de informações pluviométricas é essencial à tomada de decisões no que diz respeito a administração dos recursos hídricos, construções e a projetos de cunho socioeconômico, como: irrigação, transporte, obras de infraestrutura, geração de energia elétrica etc. O tratamento estatístico das informações pluviométricas disponíveis (juntamente com outros dados meteorológicos como velocidade do vento, temperatura do ar, etc.) possibilita a previsão de desastres, cheias e secas, por exemplo, permitindo assim, a ação de contenção de rios ou a distribuição e racionamento da água potável quando necessário (Salgueiro e Montenegro, 2017; Sena, Lucena, Ribeiro, 2017; Silva, Silva, Silva, Almeida, Araújo, 2018).

Tratando-se do campo acadêmico, os dados de precipitação são de extrema importância para os estudos de variabilidade e tendência climática (Coan, Back, Bonetti, 2015; Mello, Kohls, Oliveira, 2015; Silva, Montenegro, Souza, 2017; Coutinho, et al., 2018), para as estimativas de vazões (Cabral, Sakuragi, Silveira, 2017), para pesquisas a respeito das atividades agrícolas regionais e locais e para o planejamento territorial (Fonseca, Souza, Silva, 2016).

É possível ainda que se realizem projetos de redirecionamento de recursos hídricos (abastecimento de água), levando em consideração a pouca homogeneidade na distribuição de chuvas em escala global ou mesmo em escala regional, como é o caso do sertão nordestino brasileiro (a transposição do rio São Francisco é um exemplo de tentativa de redirecionamento). Dessa forma, é necessário que o registro convertido em séries históricas apresente uma margem de erro aceitável e uma continuidade, originando uma base de dados consistente de modo a permitir o adequado tratamento estatístico das informações e o emprego dos resultados nos estudos de viabilidade de obras (Carvalho et al., 2017; Farias, Santos & Silva, 2018).

Os dados de séries disponíveis são obtidos a partir do monitoramento e coleta de informações fornecidas por dispositivos chamados pluviômetros e pluviógrafos. No mercado, existem aparelhos desde os mais simples (de registro de dados manual) até os mais sofisticados (equipados com sensores que captam a precipitação e suas características), porém, todos eles estão sujeitos a alguma espécie de erro que pode comprometer a consistência das informações coletadas (Villazón & Willems, 2010).

A ocorrência de falhas nos dados de precipitação disponíveis (ou ausência deles) é comum e pode ocorrer por conta dos erros provocados pela ausência do operador do aparelho pluviométrico, pelo registro incorreto, por falhas mecânicas, por perda dos registros realizados ou mesmo pela falta de verba para continuidade da operação. Para superar esses empecilhos existem as chamadas metodologias para preenchimento de falhas, que permitem prever a ocorrência de chuvas em determinada região baseando-se em dados obtidos em regiões próximas (Ventura, Santana, Martins, Figueiredo, 2016; Bier, Ferraz, 2017; Bielenki Junior et al., 2018).

No estudo realizado por Ventura et al. (2016), foi avaliada a viabilidade do uso dos métodos: média aritmética simples, média móvel, regressão linear simples e regressão linear múltipla para o preenchimento de falhas em dados de temperatura, umidade relativa do ar, ponto de orvalho, pressão atmosférica e radiação solar. Os autores concluíram que para a maioria das variáveis testadas (temperatura, umidade e ponto de orvalho), a regressão linear múltipla teve ótimo desempenho e para a pressão atmosférica, a média simples apresentou resultados precisos. Enquanto que para a radiação solar, nenhum dos métodos apresentou resultado satisfatório.

Mello et al. (2017) realizaram análises dos métodos da ponderação regional, da regressão linear simples e múltipla e da ponderação regional com base em regressões lineares para o preenchimento de falhas em dados mensais de precipitação, utilizando as séries históricas fornecidas com base em oito postos pluviométricos localizados no município de Joinville (Santa Catarina). Os postos foram divididos em dois grupos, no grupo 1 utilizaram-se dados desde junho de 1987 a junho de 1996 e no grupo 2 desde agosto de 1976 a agosto de 1984. Chegou-se a conclusão que o método que apresentou melhor resultado foi o da regressão linear múltipla, seguido pela ponderação regional e por último pela ponderação regional com base em regressões lineares.

Já Bier e Ferraz (2017) Compararam seis métodos aplicados a séries de dados mensais de precipitação e temperatura média compensada entre os anos de 2003 e 2012. Nesse caso, as estações meteorológicas escolhidas foram: Porto Alegre, Bagé, Santa Maria e Lagoa Vermelha (Designação do Instituto Nacional de Meteorologia -INMET) todas localizadas no estado do Rio Grande do Sul. Os autores mostraram que os resultados obtidos para temperatura foram melhores com a utilização dos métodos de regressão linear múltipla e ponderação regional enquanto que para precipitação não houve um método a se sobressair, onde, Ponderação Regional, Interpolação do inverso da distância, Média aritmética simples e o Método da razão normal apresentaram uma discreta vantagem em comparação ao Método tradicional do Reino Unido e a Regressão linear múltipla.

Em trabalho realizado por Diaz, Pereira e Nóbrega (2018), foram avaliados os métodos da Ponderação Regional e a TRMM (Tropical Rainfall Measuring Mission) na estimativa de dados mensais de precipitação para a bacia do Rio Pajeú, no estado de Pernambuco. Para tal, foram utilizados dados de 3 estações pluviométricas desde 1998 a 2013. No estudo concluiu-se que os dois métodos apresentaram desempenho satisfatório e bastante semelhante, ou seja, ambos são alternativas viáveis para o preenchimento de falhas na região mencionada.

Oliveira et al. (2010) realizaram uma análise de algumas dessas metodologias (ponderação regional, regressão linear, ponderação regional com base em regressões lineares e vetor regional) utilizando séries históricas de precipitações totais anuais contendo 22 anos de observações de 6 diferentes postos localizados na região central do estado de Goiás. Chegaram a conclusão que o método que apresentou melhor resultado foi o da regressão linear múltipla, seguido por: vetor regional combinado com a regressão potencial múltipla, ponderação regional, vetor regional combinado com a regressão linear múltipla, regressão potencial, ponderação regional com base em regressões lineares, vetor regional combinado com a ponderação regional e, aquele que obteve o pior resultado, o vetor regional.

Em outro estudo, realizado por Villazón e Willems (2010), foram analisados os métodos da regressão linear simples e da regressão linear múltipla aplicados a dados mensais desde 1976 até 1999 de 33 estações localizadas na região da bacia do rio Pirai (afluente do rio Amazonas), na Bolívia. Chegando a conclusão que com o uso da regressão linear múltipla obteve-se uma redução de 36% no desvio padrão em detrimento da regressão linear simples.

No presente trabalho as metodologias para preenchimento de falhas em séries temporais de precipitação foram utilizadas, não como ferramenta auxiliar no preenchimento de lacunas provocadas por erros, mas sim como um meio de prever integralmente os dados de pluviometria para o município de Catanduva, localizado no estado de São Paulo. Objetivando assim, analisar e avaliar a viabilidade do uso dos métodos da ponderação regional, da regressão linear simples e da regressão linear múltipla, verificando o desempenho e a correlação entre os dados estimados e os dados reais (registrados pelo posto pluviométrico de Catanduva). Desta forma, provendo os dados de precipitação necessários aos mais diversos estudos climáticos citados anteriormente.

2. Metodologia

Para a realização da análise das metodologias de preenchimento de falhas fez-se necessária a utilização de dados de séries históricas de pluviometria e os métodos utilizados foram o da regressão linear simples, da regressão linear múltipla e da ponderação regional.

Obtenção de dados

Os dados de séries históricas utilizados foram fornecidos pelo Banco de Dados Meteorológicos para Ensino e Pesquisa (BDMEP), pertencente ao Instituto Nacional de Meteorologia (INMET). Foram coletadas informações mensais referentes a trinta anos de observações de quatro estações pluviométricas localizadas no estado de São Paulo, e então foram estimados os dados

para a estação de Catanduva, baseando-se em outros três postos próximos a ela. Os municípios escolhidos, destacados na Figura 1, foram: São Carlos, a 149 km de distâncias; Franca, a 177 Km e Votuporanga, a 130 Km.

Figura 1 - Área de estudo.



Fonte: Autores.

Para a estimativa dos dados de precipitação da Cidade Catanduva-SP, foram utilizadas técnica de preenchimento de falhas em séries temporais: regressão linear simples, regressão linear múltipla e a ponderação regional. Tais metodologia já vem sendo utilizada no preenchimento de falhas com êxito a anos (Franco et al., 2019).

O método da regressão linear simples (Equação 1), foi executado com relação a cada um dos três municípios vizinhos (Carlos-SP, Votuporanga-SP e Franca-SP) ao que desejava-se estimar (Catanduva-SP). Ou seja, obteve-se uma equação para cada município, cujo valores precipitados foram estimados mês a mês, para cada um dos 30 anos de observação dos postos pluviométricos analisados.

$$P_x = aP_z + b \quad (1)$$

O método da regressão linear múltipla (Equação 2) foi executado mês a mês com base nos dados de precipitação dos municípios de São Carlos, Votuporanga e Franca resultando em uma única equação mensal que contempla as três variáveis.

$$P_x = aP_z + bP_y + cP_w + d \quad (2)$$

Sendo P_x a precipitação a ser estimada (em Catanduva-SP), P_z , P_y e P_w as precipitações medidas nos postos vizinhos (em Carlos-SP, Votuporanga-SP e Franca-SP), e a , b , c e d os coeficientes, que devem ser ajustados anteriormente.

No método da ponderação regional, a obtenção dos dados estimados para preenchimento das falhas das séries temporais é realizada tomando por base as medições de outros postos vizinhos a ele, além de suas médias, conforma a Equação 3:

$$P_x = \frac{P_{xm}}{N_p} \sum_{i=1}^{N_p} \left(\frac{P_i}{P_{im}} \right) \quad (3)$$

Sendo P_m o valor a ser estimado, P_{xm} a precipitação média do posto com falhas, P_{im} a precipitação média de cada posto próximo ao posto com falhas, P_i a precipitação do posto próximo no mês i , e N_p o número de postos próximos ao posto com falhas que estão sendo utilizados (Dias et al, 2018).

Para analisar os dados gerados pelos modelos foram utilizados nove parâmetros estatísticos, a saber, o coeficiente de correlação (R , equação 4), o índice de concordância (d , equação 5), o índice de confiança (c , equação 6), o erro quadrático ($RMSE$, equação 7), o erro quadrático médio (EQM , equação 8), a razão dos desvios (RD , equação 9), a eficiência da modelagem (EM , equação 10) e o coeficiente de massa residual (CMR , equação 11), apresentados a seguir:

$$R = \frac{\sum(M_i - \bar{M})(E_i - \bar{E})}{\sqrt{\sum(M_i - \bar{M})^2 \sum(E_i - \bar{E})^2}} \quad (4)$$

$$d = 1 - \frac{\sum(E_i - M_i)^2}{\sum((E_i - \bar{M}) + (M_i - \bar{M}))^2} \quad (5)$$

$$c = R \cdot d \quad (6)$$

$$RMSE = \left\{ \frac{\sum(E_i - M_i)^2}{N} \right\}^{1/2} \quad (7)$$

$$EQM = \left[\frac{\sum(E_i - M_i)^2}{N} \right]^{1/2} \cdot \frac{100}{\bar{M}} \quad (8)$$

$$RD = \frac{\sum(M_i - \bar{M})^2}{\sum(E_i - \bar{M})^2} \quad (9)$$

$$EM = \frac{\sum(M_i - \bar{M})^2 - \sum(E_i - M_i)^2}{\sum(M_i - \bar{M})^2} \quad (10)$$

$$CMR = \frac{\sum M_i - \sum E_i}{\sum M_i} \quad (11)$$

Sendo E_i os dados estimados, \bar{E} a média dos dados estimados, M_i os dados medidos, \bar{M} a média dos dados medidos e N número de observações (Antonino et al., 2004; Soares et al., 2013).

3. Resultados e Discussão

As estimativas de dados de precipitação mensais para o município de Catanduva realizadas seguindo os métodos de preenchimento de falhas, foram baseadas nos dados de precipitação de três municípios do Estado de São Paulo, relativamente próximos, a saber, Franca, São Carlos e Votuporanga.

O método da regressão linear simples foi executado com relação a cada um dos três municípios vizinhos ao que desejava-se estimar. Os ajustes foram realizados mês a mês, para cada um dos 30 anos de observação dos postos pluviométricos analisados. Foram executados os parâmetros estatísticos para avaliar a correlação entre os dados estimados e os dados observados. Como foram obtidas equações com respeito a cada município, as análises estatísticas também foram realizadas com relação a cada um deles e a totalidade dos dados calculados.

Na Tabela 1, encontram-se os valores dos parâmetros estatísticos. A regressão linear simples baseada no município de Votuporanga foi a que obteve um melhor desempenho com os melhores resultados para todos os parâmetros estatísticos, com os valores de *d*, *RD* e *EM* equivalentes a 0.93, 1.32 e 0.76, respectivamente, em comparação a 0.90, 1.52 e 0.69 obtidos pela regressão simples com base no município de São Carlos e 0.90, 1.48 e 0.69 com base em Franca, sendo esses dois últimos muitíssimo semelhantes no desempenho de seus resultados, apresentando exatamente os mesmos valores para os parâmetros *d*, *c*, *EM*, e *R*, e apresentando diferenças decimais para o *RMSE*, *EQM* e *RD*.

Tabela 1 - Parâmetros estatísticos para análise dos dados obtidos pelo método da regressão linear simples.

Município	<i>d</i>	<i>c</i>	RMSE(%)	EQM(%)	RD	EM	CMR (10 ⁻⁴)	R
São Carlos	0.90	0.75	52.53	51.24	1.52	0.69	90.62	0.83
Franca	0.90	0.75	52.98	51.46	1.48	0.69	-94.48	0.83
Votuporanga	0.93	0.81	47.01	45.66	1.32	0.76	0.14	0.87

Fonte: Autores.

Nota-se que os valores do *RMSE* e *EQM* são bastante elevados, uma vez que seu valor ótimo é 0, atribui-se a isso o fato de que está sendo tratada uma enorme quantidade de dados (30 anos de observações pluviométricas mensais, para cada um dos três municípios), o que causa uma maior dispersão no comportamento dos dados obtidos com a modelagem, não comprometendo, entretanto, a confiabilidade dos resultados.

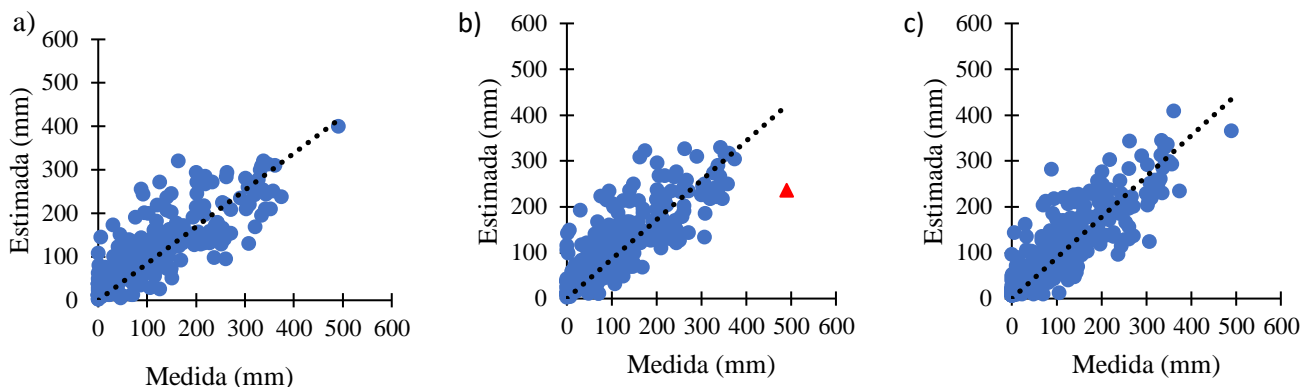
Observa-se que os valores obtidos com a aplicação dos parâmetros estatísticos indicam bons resultados e que, dentre os três municípios, aquele que proporcionou o melhor desempenho na modelagem dos dados foi o de Votuporanga que é o município mais próximo geograficamente de Catanduva (130 Km; Catanduva é o posto pluviométrico que teve seus dados estimados), logo, nesse caso, quanto menor a distância e maior a homogeneidade pluviométrica (semelhança na ocorrência dos eventos pluviométricos) entre o local do qual pretende-se estimar os dados e o local do qual os dados são extraídos para realizar tais estimativas, melhores são os resultados obtidos (mais próximos são dos resultados reais). Já os outros dois municípios (São Carlos e Franca) apresentaram um desempenho bastante semelhante com respeito aos valores obtidos com os parâmetros estatísticos, o que pode ser explicado pelo fato de que suas distâncias ao posto a ser estimado também são semelhantes (São Carlos: 149 Km; Franca: 177 Km).

Villazón e Willems (2010) ao analisarem a metodologia da regressão linear simples para estações localizadas na região da bacia do rio Pirai (afluente do rio Amazonas), na Bolívia, optaram por utilizar o método apenas para as estações mais homogêneas (em relação a pluviometria) e obtiveram, assim, coeficiente de correlação de 0.68, menor que o coeficiente apresentado entre as estações de Votuporanga e Catanduva (0.87), que foi o melhor ajuste obtido na análise aqui realizada.

Observa-se nos gráficos da Figura 2 que, no geral, os pontos apresentam pouca dispersão em torno da linha de tendência. Porém, em alguns casos, existem pontos que foram exageradamente subestimados ou superestimados. O ponto triangular vermelho destacado no gráfico (b), por exemplo, representa a ocorrência de modelagem menos eficaz para a regressão linear simples com base no município de Franca, onde o valor estimado foi de, aproximadamente, 263.26, e o real foi de 490 milímetros, correspondente ao mês de Janeiro do ano de 1994 e representando um erro relativo de 52%.

Analisando estatisticamente os dados de pluviometria obtidos a partir dos modelos, observa-se que as melhores estimativas obtidas para todos os três (regressão simples baseada em: São Carlos, Franca e Votuporanga) são relativas ao mês de junho. Quanto aos meses com pior desempenho, existe uma grande variedade, sendo Novembro para Franca, Julho para São Carlos e Outubro para Votuporanga.

Figura 2 - Precipitação medida versus estimada, pela regressão linear simples, com base nos valores medidos nos postos pluviométricos nas cidades de: São Carlos (a), Franca (b) e Votuporanga (c).



Fonte: Autores.

O mês de Junho apresenta um volume médio de chuva de 31.3 milímetros, os meses de Novembro, Julho e Outubro apresentam volumes médios de 134.3, 19.7 e 84.6 milímetros, respectivamente. Isso indica que, no geral, a modelagem apresenta melhor desempenho para os meses menos chuvosos.

O método da regressão linear múltipla foi executado mês a mês com base nos dados de precipitação dos municípios de São Carlos, Votuporanga e Franca resultando em uma única equação mensal que contempla as três variáveis.

Com a aplicação das referidas equações e posterior análise estatística com o emprego dos parâmetros já citados, foram obtidos os valores ilustrados, a seguir, na Tabela 2. Quase todos os parâmetros indicam melhores resultados que aqueles obtidos com a regressão linear simples (independente de qual município foi tomado como base para a modelagem), com exceção apenas do parâmetro CMR. Observa-se que, para os parâmetros d, c, EM, RD e R, os valores obtidos com a regressão linear múltipla foram muito próximos aos obtidos com a regressão linear simples baseada no município de Votuporanga (município que obteve melhor desempenho para a regressão simples), com diferenças de alguns décimos, já para os parâmetros RMSE e EQM houve diferença de 3.72, 3.61, respectivamente. Ressaltando que os valores do RMSE e EQM são bastante elevados devido ao fato de que está sendo tratada uma enorme quantidade de dados (30 anos de observações pluviométricas mensais), não comprometendo, entretanto, a confiabilidade dos resultados.

Tabela 2 - Parâmetros estatísticos para análise dos dados obtidos pelo método da regressão linear múltipla.

d	c	RMSE (%)	EQM (%)	RD	EM	CMR (10^{-4})	R
0.94	0.84	43.29	42.05	1.24	0.79	27.68	0.89

Fonte: Autores.

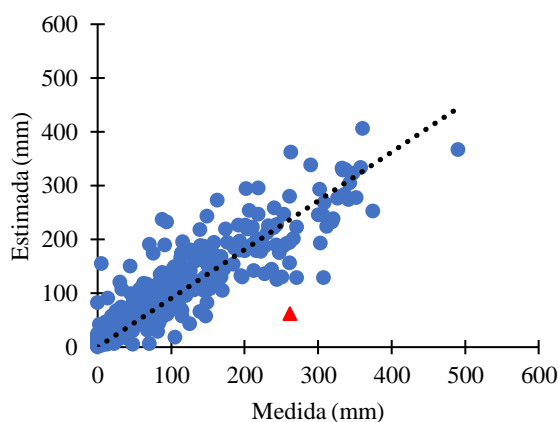
Observa-se que todos os valores obtidos com a aplicação dos parâmetros indicam um melhor desempenho do método da regressão linear múltipla com relação ao método da regressão linear simples, fato já esperado, uma vez que, a regressão múltipla combina em uma única equação as contribuições de todas as variáveis utilizadas para estimar os dados pluviométricos do posto de Catanduva. No gráfico da Figura 3, o ponto triangular vermelho destacado representa a pior estimativa obtida com o método da regressão linear múltipla, que ocorreu no mês de dezembro do ano de 1999. Nesse caso, o valor estimado foi, aproximadamente, 63.06 e o real foi 261.2, apresentando um erro relativo de 76%.

Villazón e Willems (2010) concluíram que com o uso da regressão linear múltipla obteve-se uma redução de 36% no erro quadrático (RMSE), bem como uma eficiência (EM) de 0.84 em detrimento da regressão linear simples (que apresentou eficiência de 0.59). Já na análise aqui realizada, houve uma redução de 7.91% no RMSE e a obtenção de um EM de 0.79 em comparação com um EM de 0.76 obtido na análise de regressão linear simples com base no município de Votuporanga (município que obteve melhor desempenho). Com relação aos outros dois municípios, São Carlos (SC) e Franca (FR), foi obtida uma redução no RMSE de 17.59% e 18.29%, respectivamente, e a eficiência calculada com o método da regressão linear simples também foi inferior ao da regressão linear múltipla, sendo, aproximadamente, 0.69 em ambos.

No estudo realizado por Mello, et al. (2017), a regressão linear múltipla foi o método que apresentou melhores resultados, mesmo com a utilização de diferentes parâmetros estatísticos para avaliá-los. No geral, o mês que apresentou o melhor ajuste na modelagem foi o mês de Junho e os piores foram os meses de Março e Novembro. Observa-se que o mês de junho apresenta o terceiro menor volume médio de chuva (31.3) e Março e Novembro apresentam o quarto e o quinto meses com maior volume médio, com médias de 137.3 e 134.3, respectivamente (valores muito próximos).

O método da ponderação regional foi aplicado utilizando as médias mensais dos dados de precipitação dos postos de Catanduva, Franca, São Carlos e Votuporanga para compor as equações de cada um dos 12 meses

Figura 3 - Precipitação medida versus estimada pelo método da regressão linear múltipla.



Fonte: Autores.

Com a aplicação das referidas equações e posterior análise estatística com o emprego dos parâmetros estatísticos, foram obtidos os valores ilustrados, a seguir, na Tabela 3. Observa-se que o método da regressão linear múltipla e o da ponderação

regional apresentam desempenho muito semelhante, com o mesmo valor para o parâmetro d e diferindo em alguns décimos para os parâmetros c, RD, EM e R. Para a regressão linear múltipla, foram obtidos os valores 0.84, 43.29, 42.05, 0.79 e 0.89 para os parâmetros c, RMSE, EQM, EM e R, respectivamente, em detrimento da ponderação regional com: 0.82, 45.23, 43.93, 0.77 e 0.88, o que indica um melhor desempenho da regressão múltipla. Apenas os parâmetros RD e CMR indicam desempenho superior da ponderação regional, com 1.20 e 20.88×10^{-4} , em relação a 1.24 e 27.68×10^{-4} obtidos pela regressão. Lembrando, mais uma vez, que os valores elevados obtidos com os parâmetros RMSE e EQM ocorrem devido ao grande número de dados tratados na modelagem (30 anos de observações pluviométricas mensais) o que não compromete a confiabilidade dos resultados.

Tabela 3 - Parâmetros estatísticos para análise dos dados obtidos pelo método da ponderação regional.

d	c	RMSE (%)	EQM(%)	RD	EM	CMR (10^{-4})	R
0.94	0.82	45.23	43.93	1.20	0.77	20.88	0.88

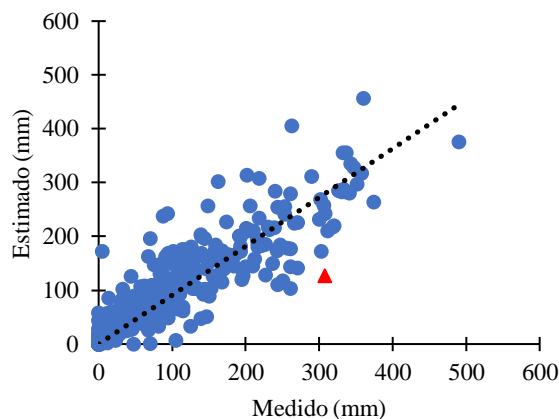
Fonte: Autores.

De acordo com a Tabela 3, o coeficiente de correlação obtido entre os dados gerados pelo método da ponderação regional e os dados reais foi de 0.88, representado assim, o segundo melhor resultado entre os métodos analisados, inferior apenas aos resultados obtidos com a regressão linear múltipla. Resultado semelhante foi obtido por Mello, et al. (2017), onde a regressão linear múltipla apresentou um desvio relativo médio de 14.4% e a ponderação regional de 20%.

No gráfico da Figura 4, o ponto triangular vermelho destacado representa o mês de novembro do ano de 1995, onde a modelagem realizada utilizando o método da ponderação regional retornou o resultando mais distante da realidade, sendo o dado estimado 127.78 e o medido 307.1, resultando num erro relativo de 58%.

Analisando os resultados obtidos, tem-se que o modelo da ponderação regional apresentou melhor desempenho no mês de Junho e pior desempenho nos meses de Março e Novembro, semelhantemente ao modelo da regressão linear múltipla. Nota-se que as estimativas apresentaram um melhor desempenho para o mês de junho, que, de acordo com os dados medidos pelo posto pluviométrico de Catanduva, é o terceiro mês com o menor volume médio de chuva, sendo assim, as metodologias apresentam melhores resultados para os meses menos chuvosos, corroborando com os resultados obtidos por Diaz, Pereira e Nóbrega (2018).

Figura 1 - Precipitação medida versus estimada pelo método da ponderação regional.



Fonte: Autores.

4. Conclusão

De acordo com os parâmetros estatísticos aplicados, todas as metodologias utilizadas para estimar os dados pluviométricos para o município de Catanduva obtiveram um desempenho satisfatório e, dentre elas, a que obteve o melhor resultado foi a da Regressão Linear Múltipla. Vale ressaltar que, embora tenha obtido o melhor resultado, a Regressão Múltipla diferiu pouco em seu desempenho com relação ao método da ponderação regional (segundo melhor método avaliado), uma vez que, dos oito parâmetros estatísticos utilizados na avaliação dos desempenhos, um deles apresentou o mesmo valor para os dois métodos (parâmetro d), enquanto outros quatro apresentaram diferenças de apenas alguns décimos (parâmetros c, RD, EM e R), sendo que os parâmetros RD e CMR foram os únicos que indicaram um melhor desempenho da Ponderação Regional.

O mês de junho (notadamente com pouco volume de chuvas comparado aos demais) foi aquele que resultou em maiores erros nas estimativas realizadas com os métodos da Regressão Linear Simples e Múltipla e Ponderação Regional, sendo assim, as metodologias apresentam melhores resultados para os meses menos chuvosos.

Quanto mais próxima está a estação a ter seus dados estimados da estação utilizada para realizar as estimativas, melhores são os resultados obtidos (no caso em que haja uma homogeneidade pluviométrica entre elas).

Sugere-se que os procedimentos metodológicos aqui utilizado seja aplicado em outras regiões, a fim de que as precipitações possam ser estimadas com o uso de equações já existentes na literatura, cuja a finalidade inicial seja apenas o preenchimento de falhas em dados meteorológicos.

Referências

- Antonino, A. C. D., Soares, W. de A., Silva, E. B. da, Lima, J. R. de S., Netto, A. M. & Lira, C. A. B. de O. (2004). Utilização do Método Inverso para a Caracterização Hidrodinâmica de um Neossolo Flúvico. *Revista Brasileira de Recursos Hídricos* 9, 81-87. 10.21168/rbrh.v9n3.p81-87
- Bielenki Junior, C., et al (2018). Alternative methodology to gap filling for generation of monthly rainfall series with GIS approach. *Revista Brasileira de Recursos Hídricos*, 23, 1-10. 10.1590/2318-0331.231820170171
- Bier, A. A., & Ferraz, S. E. T. (2017). Comparação de Metodologias de Preenchimento de Falhas em Dados Meteorológicos para Estações no Sul do Brasil. *Revista Brasileira de Meteorologia*, 32 (2), 215-226. 10.1590/0102-77863220008
- Cabral, S., Sakuragi, J., & Silveira, C. (2016). Uncertainties and errors flow estimate using hydrological modelling and precipitation by RADAR. *Ambiente e Agua*, 12 (1), 57-70. 10.4136/ambi-agua.1924
- Carvalho, J. R. P., et al (2017). Model for Multiple Imputation to Estimate Daily Rainfall Data and Filling of Faults. *Revista Brasileira de Meteorologia*, 32 (4), 575-583. 10.1590/0102-7786324006
- Coan, B. de P., Back, Á. J., & Bonetti, A. V. (2015). Precipitação mensal e anual provável no estado de Santa Catarina. *Revista Brasileira de Climatologia* 15 (10), 122-142. <http://dx.doi.org/10.5380/abclima.v15i0.38348>
- Coutinho, E. d. C., Rocha, E. J. P. da, Lima, A. M. M., Ribeiro, H. M. C., Gutierrez, L. A. C. L., Barbosa, A. J. S., Paes, G. K. A. A., Bispo, C. J. C., & Tavares, P. A. (2018). Variabilidade climática da precipitação na bacia amazônica brasileira entre 1982 e 2012. *Revista Brasileira de Climatologia* 22, 476-500. 10.5380/abclima.v22i0.46074
- Coutinho, E. R., Silva, R. M. da, Madeira, J. G. F., Coutinho, P. R. de O. dos S., Boloy, R. A. M., & Delgado, A. R. S. (2018). Application of Artificial Neural Networks (ANNs) in the Gap Filling of Meteorological Time Series. *Revista Brasileira de Meteorologia* 33, 317-328. 10.1590/0102-7786332013
- Diaz, C. C. F., Pereira, J. A. dos S., & Nóbrega, R. S. (2018). Comparação de dados estimados pelo método da Ponderação Regional (PR) e dados estimados pelo TRMM para o preenchimento de falhas de precipitação na bacia hidrográfica do Rio Pajeú. *Revista Brasileira de Climatologia* 22 (14), 324-339. 10.5380/abclima.v22i0.46911
- Fante, K. P., & Sant'anna Neto, J. L. (2016). Técnicas estatísticas para a homogeneização de dados de temperatura em séries temporais climatológicas. *Revista Brasileira de Climatologia* 18, 143-156. 10.5380/abclima.v18i0.43202
- Franco, V., Souza, E., Costa, C., Ferreira, D., Oliveira, J., Sodré, G., Kuhn, P., & Azevedo, F. (2019). Prognóstico sazonal da precipitação para o verão e outono austral da Amazônia oriental (Seasonal prognosis for the southern summer and autumn in the Eastern Amazon). *Revista Brasileira de Geografia Física*, 12 (1), 057-070. 10.26848/rbgf.v12.1.p057-070
- Fonseca, S. F. da, Souza, M. J. H. de, & Silva, A. C. (2016). Análise da precipitação pluvial do município de Pirapora-MG, entre 1961 e 2013. *Raega - O Espaço Geográfico em Análise*, 38, 35-49. 10.5380/raega.v38i0.41804
- Mello, Y. de, Kohls, W., & Oliveira, T. M. N. (2015). Análise da precipitação mensal provável para o município de Joinville (SC) e região. *Revista Brasileira de Climatologia*, 17, 246-258. 10.5380/abclima.v17i0.42975

- Mello, Y. R., Kohls, W., & Oliveira, T. M. N. (2017). Uso de diferentes métodos para o preenchimento de falhas em estações pluviométricas. *Boletim de geografia*, 35, 112-121. 10.4025/bolgeogr.v35i1.30893
- Oliveira, L. F., Fioreze, A. P., Medeiros, A. M., & Silva, M. A., (2010). Comparison of failure completion methodologies for historical series of annual rainfall. *Brazilian Journal of Agricultural and Environmental Engineering* 14, 1186-1192. 10.1590/S1415-43662010001100008
- Salgueiro, J. H. P. B., & Montenegro, S. M. G. L. (2008). Análise da distribuição espacial da precipitação na bacia do rio Pajeú em Pernambuco segundo método geostatístico. *Revista Tecnologia*, 29, 174-185.
- Sena, J. P. de O., Lucena, D. B., & Ribeiro, G. do N. (2017). Eventos extremos de precipitação no sertão paraibano: Variação espaço-temporal. *Revista Verde de Agroecologia e Desenvolvimento Sustentável*, 12, 748-755. 10.18378/rvads.v12i4.4938
- Silva, G. S., Silva, W. S., Silva, A. L., Almeida, N. V., & Araújo, L. E. (2018). Análise da precipitação da microrregião do Cariri Oriental Paraibano. *Revista de Geociências do Nordeste*, 4, 42-57.
- Silva, R. O. B. da, Montenegro, S. M. G. L., & Souza, W. M. (2017). Tendências de mudanças climáticas na precipitação pluviométrica nas bacias hidrográficas do estado de Pernambuco. *Engenharia Sanitaria e Ambiental*, 22, 579-589. 10.1590/s1413-41522017142481
- Soares, W. de A., Antonino, A. C. D., Lima, J. R. de S., Souza, E. S. de, Montenegro, S. M. G. L., & Lira, C. A. B. de O. (2013). Simulação dos Fluxos de Água e de Energia na Microbacia Hidrográfica da Represa Vaca Brava no Brejo Paraibano. *Revista Brasileira de Recursos Hídricos*, 18, 185-196. 10.21168/rbrh.v18n4.p185196
- Ventura, T. M., Santana, L. L. R., Martins, C. A., & Figueiredo, J. M. (2016). Análise da aplicabilidade de métodos estatísticos para preenchimento de falhas em dados meteorológicos. *Revista Brasileira de Climatologia*, 19, 168-177. 10.5380/abclima.v19i0.44989
- Villazón, M. F. & Willems, P. (2010). Filling gaps and Daily Disaccumulation of Precipitation Data for Rainfall-runoff model. In 4th International Scientific Conference on Water Observation and Information Systems for Decision Support. Ohrid, Republic of Macedonia. *CD-ROM* pp.1-9. Publisher: BALWOIS 2010. 25-29.