

Proceeding Series of the Brazilian Society of Computational and Applied Mathematics

Estudo matemático e estatístico sobre a análise do volume das precipitações pluviométricas na cidade de São Paulo e sua previsibilidade usando o modelo ARIMA

Giselle Couto Falcão¹

Programa de Pós Graduação em Modelagem Matemática Computacional, CEFET-MG, Belo Horizonte, MG

Edgar Lacerda Aguiar²

Programa de Pós Graduação em Modelagem Matemática Computacional, CEFET-MG, Belo Horizonte, MG

Resumo. A previsibilidade das inundações tem sido um desafio recorrente nas últimas décadas em grandes cidades, principalmente em São Paulo. Nas estações de primavera e verão determinadas regiões do Brasil sofrem com constantes alagamentos, enchentes e inundações. Tais fenômenos ocorrem principalmente devido ao aumento das precipitações intensas, associada a falta de investimento dos órgãos responsáveis em prevenção. Este trabalho busca analisar as previsões do volume máximo de precipitações do bairro Jardim de São Paulo, as análises dos dados foram realizadas utilizando os métodos estatísticos não paramétricos de séries temporais não estacionárias reescaladas, para a observação de sazonalidade. Com objetivo de localizar uma possível linearidade das médias móveis auto regressivas no período sazonal. Tendo como modelo de apresentação estrutural a simulação dos dados em sistemas complexos de previsibilidade e instabilidade, o qual terá como base, dados históricos mensais dos volumes das precipitações dos meses de novembro, dezembro e janeiro coletados pela estação 83781 no Mirante de Santana. Por meio de métodos estatísticos e análises realizadas, foram observados resultados que ratificam a hipótese de crescimento significativo nas séries temporais, do volume máximo das precipitações históricas da Capital considerando o período de 58 anos, objeto desse trabalho.

Palavras-chave. Ajuste Exponencial, Análise tendência, Autocorrelação, ARIMA, Clima, Precipitações São Paulo.

1 Introdução

O volume das precipitações vem sendo amplamente discutido no meio acadêmico e científico. Embora seja um assunto climático, a análise dos dados referente as precipitações, oferecem um campo muito interessante, para o uso de diversos modelos e ferramentas matemáticas, que permitem a investigação de padrões, comportamento cíclico ou estocástico,

¹giselle_falcao@hotmail.com

²edgarlaguiar@gmail.com

entre outras informações que podem ser extraídas das séries. Com o uso de ferramentas e modelos adequados para a análise de cada série, considerando sua especificidade. Alguns estudos conseguiram identificar um crescimento linear, em séries relacionadas aos grandes volumes de precipitações [6], [1] [3], [11].

Neste sentido, as pesquisas relacionadas ao volume de precipitação, utilizam em sua maioria, testes não paramétricos, como Dickey-Fuller e Phillips-Perron, uma vez que estes testes, permitem uma análise mais rigorosa quanto ao comportamento da série, que em muitos casos apresentam um comportamento totalmente estocástico. [4], [13], [5].

Deste modo, após estudos para a identificar a melhor estratégia a ser adotada para o trabalho, foram selecionadas técnicas de dimensionamento e comportamento fractal. Estas ferramentas, são usadas para modelar diversos fenômenos naturais, tais como o volume da precipitação, temperatura, níveis do mar, ventos, entre outros. O dimensionamento permite o uso de métodos e modelos matemáticos que descrevem o modelo em ajustes de transformadas de Fourier (TF). Os ajustes possibilitam o reescalonamento da série, em busca de períodos com maior incidência de precipitações no domínio do tempo, o que permite uma verificação mais eficiente quanto a sua previsibilidade e ciclicidade [9], [12], [7].

O uso de estratégias de redimensionamento, associada ao ajuste de séries temporais com comportamento não estacionário, permite elaborar um modelo matemático capaz de ajustar de forma concisa os dados da série temporal. Deste modo o ajuste assegura a possibilidade pontual de uma análise de tendência da série, que pode ser feita por meio do modelo Holt-Winter. Este modelo representa uma ferramenta eficaz, que garante as características das séries quanto a sazonalidade, estacionariedade e principalmente a normalidade [8] [10].

Associar o ajuste e a análise de tendência ao estudo residual da série. Pode garantir a observação de resultados capazes de produzirem inferências relevantes sobre os dados, bem como subsidiar elaboração dos parâmetros necessários para desenvolvimento de outros ajustes e modelos. Os parâmetros, desenvolvidos por esta associação, auxiliam os modelos, tais com: médias móveis (SMA), autocorrelação, autocovariância, que por sua vez, são ferramentas relevantes para o desenvolvimento de outros métodos regressivos. Bem como a observação do ruído da série, que permite a investigação da adequação do modelo, ratificando sua confiabilidade. [2] [3].

Deste modo, o tratamento analítico da série temporal, após ajustes, reescalonamento e transformações adequadas, pode ser submetido a uma análise de tendência cuidadosa e seletiva, realizada com o auxílio do modelo auto-regressivo integrado de médias móveis (ARIMA). Este modelo é comumente usado na econometria, pois permite a análise de tendência em séries temporais, baseando-se nos parâmetros: nível de defasagem da autocorrelação da série, médias anteriores e posteriores e a diferença entre elas. Com estas características o modelo torna-se uma ferramenta muito interessante para observação de longas séries temporais, além de auxiliar no cálculo das previsões do modelo ajustado [14].

O modelo ARIMA, permite a interação entre outros métodos, que propiciam ajustes, resíduos e transformações. A interação entre os métodos, foi a estratégia utilizada neste estudo em busca de evidências, sobre o crescimento do volume de precipitação na cidade de São Paulo. Tendo como foco principal a média, desvio padrão e variância e suas respectivas flutuações ao longo da série histórica do volume das precipitações da cidade. Os dados

foram coletados mensalmente, dos registros históricos da precipitação do Instituto Nacional de Meteorologia ¹ (INMET), colhidos durante 58 anos na estação 83781.

2 Metodologia

A estação 83781 localizada no Mirante de Santana no Bairro Jardim, é a principal estação do estado. Ela é composta por vários sensores isolados que registram continuamente os parâmetros meteorológicos como a pressão atmosférica, temperatura, umidade relativa do ar, precipitação, radiação solar, direção e velocidade do vento, etc. Tais dados são anotados por um observador a determinado intervalo e enviado a um centro coletor por um meio de comunicação.

Os dados foram submetidos inicialmente ao teste de normalidade, estacionaridade e sazonalidade dos altos volumes de precipitações, este último resultado, foi detectado devido a autocorrelação da série. Após as execuções e análises dos testes, a série foi submetida a TF, em busca de um estudo espectral dos dados. A partir destes resultados foram extraídos os períodos de maior incidência de chuva, que correspondem aos meses de novembro, dezembro e janeiro dos últimos 58 anos.

Em sequência foi utilizado o método aditivo de Holt-Winters, que garante uma suavização exponencial da série, com uma equação de previsão e três de suavização, dentre estas tem-se: uma para o nível, uma para a tendência e uma para o componente sazonal. O parâmetro de suavização é dado pelo modelo apresentado na equação (1).

$$s_t = \gamma(1 - \alpha)(y_t - \beta_{(t-1)} - b_{(t-1)}) + [1 - \gamma(1 - \alpha)]s_{(t-m)}, \quad se \quad 0 \leq \gamma \leq 1 \quad (1)$$

Já para a utilização do ARIMA temos t , que representa um índice pertencente aos números reais, o qual o modelo pode ser descrito pela equação (2), a seguir. Para esta análise é importante salientar que o processo ARIMA, possui como parâmetros iniciais: (p, d, q) .

$$\left(1 - \sum_{i=1}^{p'} \alpha_i L^i\right) = \left(1 - \sum_{i=1}^{p'-d} \phi_i L^i\right) (1 - L)^d \cdot \left(1 - \sum_{i=1}^{p'} \alpha_i L^i\right) = \left(1 - \sum_{i=1}^{p'-d} \phi_i L^i\right) (1 - L)^d. \quad (2)$$

Foi conferido uma maior representatividade do volume de precipitações no trimestre correspondente aos meses de novembro, dezembro e janeiro. O agrupamento destes meses permitiu a identificação de um comportamento crescente nas séries que indicam o volume máximo de precipitação. Ao utilizarmos os resultados obtidos por meio do modelo ARIMA, foi possível ajustar a série com auxílio TF. Este ajuste permitiu realizar uma análise de tendência de forma associada aos resultados do modelo HoltWinters.

¹www.inmet.gov.br/

3 Resultados

A partir dos métodos realizados e gráficos desenvolvidos, pode-se afirmar a não estacionariedade da série, bem como sua não normalização. No entanto, após as transformações e ajustes foi possível realizar a decomposição da série, e observar a flutuação da SMA, que por sua vez, permitiu a indicação, crescente no volume das precipitações dos meses novembro, dezembro e janeiro. De forma conjunta, a análise autocorrelacional da série, possibilitou a observação de persistência de períodos com altos volumes das precipitações após períodos extensos de baixo volume de precipitação em São Paulo.

Deste modo, com uso de um ajuste exponencial aditivo (AEA), conforme demonstrado na equação (1), associado a um ajuste residual (AR), foi possível a identificar uma defasagem auto-correlacionada na série, de 5 anos. Esta análise indica inicialmente a ciclicidade dos altos volumes de precipitação ao longo dos anos.

Interessados em estabelecer um modelo para previsão da série objeto deste estudo, que se adéque à suas características assegurada pela associação dos ajustes AEA e AR. A série foi submetida a TF, para preservar sua estacionariedade e normalidade, informações essenciais para cálculo ARIMA, que conta como ponto de partida as informações obtidas pelo modelo Holt-Winters modificado.

De tal modo que, o modelo ARIMA utilizando os parâmetros iniciais (9,0,0) e (1,0,1), permite verificar na defasagem da série, quanto na defasagem residual, um crescimento dentro das médias esperadas para previsão, considerando a estimativa calculada pelo modelo ARIMA, em que foi detectado um coeficiente $p = 0.687564$. O coeficiente calculado, indica que a persistência dos altos volumes de precipitação, é sazonal, ou seja, obedecem uma periodicidade de aproximadamente 6-8 anos entre períodos de alto volume e baixo volume de precipitações. O aumento citado anteriormente, pode ser identificado por meio de análise de tendências dos dados históricos coletados.

Para uma verificação mais minuciosa dos componentes não observáveis da série, a mesma foi submetida a uma decomposição multiplicativa, conforme observado na figura 1, esta decomposição consiste na integração multiplicativa do efeito sazonal, tendência, ciclicidade e resíduos. O efeito multiplicativo desta comparação permite a verificação dos volumes mínimos e máximos nos meses de novembro, dezembro e janeiro. Ao analisar o gráfico pode-se observar picos nos componentes: sazonal, tendências e resíduos da série, no entanto o componente ciclicidade não demonstra grande diferença ao longo da série histórica. A decomposição é uma informação relevante para a escolha de método de análise de tendência mais adequado às características dos dados.

Após as análises dos dados, seleção dos meses de maior relevância para o estudo, transformações, tanto de linearização quanto e exponenciação, bem como os ajustes realizados na série, foi possível descrever um modelo que se aproxima muito bem da série original, é possível observar uma sobreposição em quase toda série. Como pode ser observado no gráfico da Figura 2, em que a linha preta indica a série original e a em vermelho o modelo proposto para ajuste da série.

Por meio do modelo elaborado, foi possível aplicar uma análise de tendências utilizando os dados obtidos no modelo ARIMA, juntamente com os parâmetros estabelecidos com auxílio de métodos matemáticos, tais como, Holt-Winters e suas modificações. Após

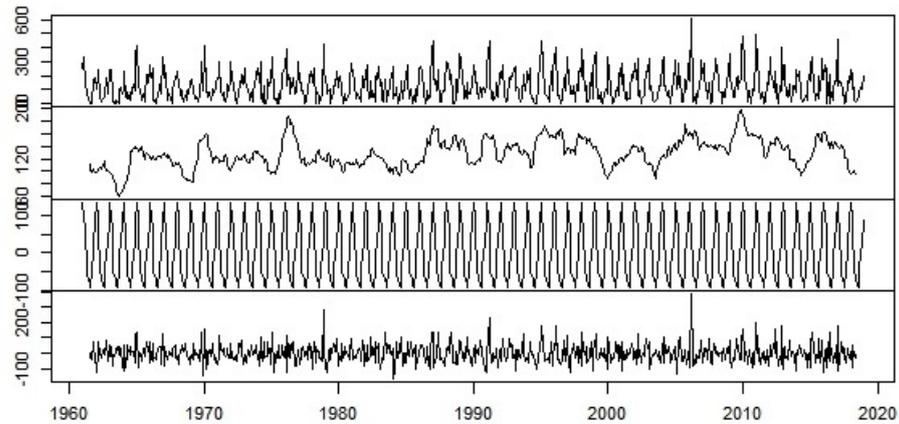


Figura 1: Decomposição da série.

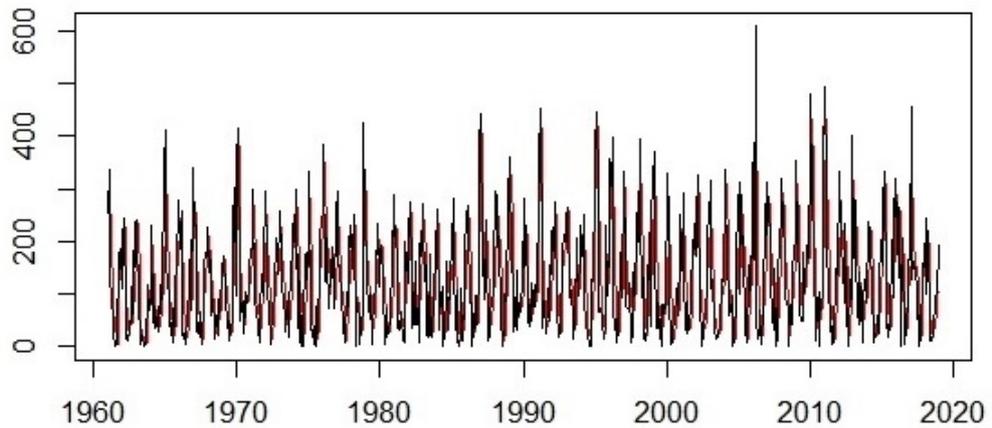


Figura 2: Ajuste da série do volume de precipitação.

o desenvolvimento da análise de tendência, a mesma foi validada, em janelas deslizantes, correspondente ao período de 5 anos da série. Deste modo foi possível estabelecer com maior precisão o período representado no gráfico exibido na figura 3, uma previsão para os volumes máximos, exposto pela linha azul, na qual consta os valores médios esperados para o volume máximo de precipitação. Utilizando o respaldo matemático e as estratégias descritas, foi possível observar uma tendência ascendente para os grandes volumes de precipitações em São Paulo.

Deste modo, pode-se inferir após a integração das análises, juntamente com o teste de resíduo, que o modelo utilizado para descrever a série, possui um ajuste com validação de 96% da série original, podendo ser utilizado para fazer previsões com considerável nível de significância. De forma semelhante, o modelo para a análise de tendência, conferiu uma alta significância para futuras previsões. O modelo ofereceu uma validação para toda a série, com média de 95% de confiança. No entanto, ao parametrizar valores máximos

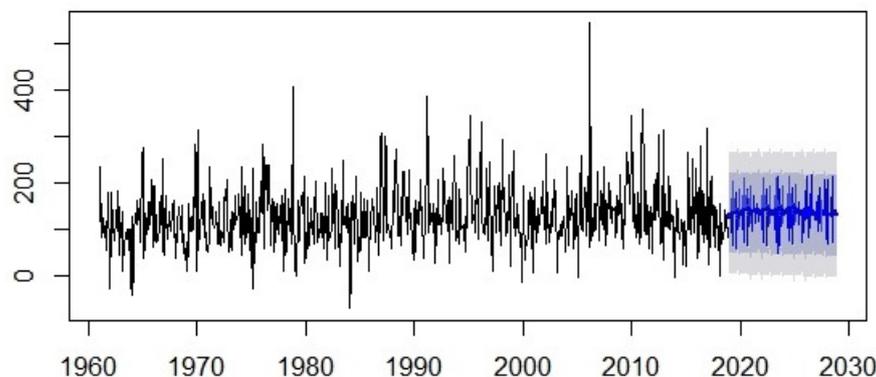


Figura 3: Análise de Tendência.

e mínimos, conforme pode ser observada no gráfico da figura 3, na região colorida com cinza claro e escuro, que abrange o período estabelecido entre 2010 a 2023, a margem de confiabilidade dos resultados foi superior a 95% de confiança.

4 Conclusão

A partir dos resultados obtidos e das discussões realizadas, foi possível identificar uma sazonalidade no volume das precipitações ocorridas nos meses de novembro, dezembro e janeiro, quando estes são agrupados e analisados separadamente em relação aos demais meses do ano. Estes resultados foram obtidos após serem observados os últimos 57 anos (1961-2018). A sazonalidade encontrada por meio dos métodos escolhidas, garante 95% de confiança quanto a ciclicidade de período com média de 7,2 anos, entre os volumes máximo das precipitações em São Paulo de acordo com a estação 83587. Além da observação de sazonalidade, foi possível verificar por meio de ajustes e análises de tendências, as previsões do volume máximo das precipitações para os próximos 20 anos.

Devido a elevada confiabilidade das análises na aplicação das tendências das futuras precipitações, o modelo juntamente com suas previsões podem vir a auxiliar em diversas áreas como a construção civil, hidráulica, entre outras. Essas métricas são muito importantes para os governantes atuais e futuros, se prepararem e investirem em infraestrutura com objetivo de minimizar os danos causados por enchentes que possam vir a ocorrer em São Paulo.

Referências

- [1] A.J.Back. *Aplicação de análise estatística para identificação de tendências climáticas*. Pesquisa agropecuária brasileira, v. 36, n. 5, 717-726, 2001.

- [2] T. Bollerslev, et al. *A conditionally heteroskedastic time series model for speculative prices and rates of return*. Review of economics and statistics, v. 69, n. 3, 542-547, 1987.
- [3] A. C. Braga, et al. *Sobre a Detecção de Autocorrelações em Séries Temporais: Uma Comparação Objetiva entre Análise de Flutuações, Transformações Wavelet e Análise Entrópica*. v. 3, n. Iii, 1-7, 2015.
- [4] W. J. Conover. *Practical Nonparametric Statistics*. New York, Wiley, 1980
- [5] D. A. Dickey, and W.A. Fuller. *Distribution of the Estimators for Autogressive Time Series with A Unit Rot*, Journal of the American Statistical Association, 74, 427-431, 1979.
- [6] A. B. Favaretto. *Estimativa do Expoente de Hurst de Série Temporais de Chuvas do Estado de São Paulo Usando as transformadas de Fourier, Wavelets e Análise R/S*. [s.l.] Universidade Estadual Paulista, 2004.
- [7] B.V. Gnedenko, and A.N. Kolmogorov. *Limit Distributions for Sums of Independent Random Variables*, Addison-Wesley: Reading, 1968.
- [8] P. S. Kalekar. *Time series forecasting using holt-winters exponential smoothing*. Kanwal Rekhi School of Information Technology, v. 4329008, n. 13, 2004.
- [9] H. M. Oliveira, *Análise de Fourier e Wavelets: sinais estacionário e não estacionários*. Editora Universitária da UFPE, 2007.
- [10] J.C. Sprott. *Chaos and Time-Series Analysis*, Oxford University Press: Oxford, 2003.
- [11] X. Tan, T.Y.Gan. *Multifractality of Canadian precipitation and streamflow*. International Journal of Climatology, v. 37, 1221-1236, 2017.
- [12] H. C. S. Thom. *Some methods of climatological analysis*. Genève : World Meteorological Organization, 1966. 54 (WMO. Technical Note, 81).
- [13] P. Yu, T.Pao-Shan, Y. Tao-Chang, and K.Chun-Chao. *Evaluating long-term trends in annual and seasonal precipitation in Taiwan*. Water Resources Management 20.6 (2006): 1007-1023.
- [14] G.P. Zhang, *Time series forecasting using a hybrid ARIMA and neural network model*. Neurocomputing, v. 50, 159-175, 2003.