

## **Aplicação da classe de modelos SARIMA para previsão da precipitação acumulada mensal na cidade de Aracaju/SE**

**Otávio Enrique José de Oliveira Ramos<sup>1</sup>, Elias Silva de Medeiros<sup>2</sup>, Carolina Cristina Bicalho<sup>3</sup>**

### **1. Introdução**

A perspectiva de antever o futuro sempre encantou a humanidade. Saber o que vai acontecer antes mesmo de os primeiros sinais se manifestarem pode propiciar melhor aproveitamento dos efeitos benéficos de eventos futuros ou uma preparação antecipada de eventuais efeitos adversos. Talvez até mais importante que antecipar os resultados seja reconhecer o que pode interferir, favorável ou desfavoravelmente, nos processos em curso, para permitir o planejamento (ANTUNES; CARDOSO, 2015).

As diferentes regiões do Nordeste brasileiro (NEB) apresentam alta variabilidade temporal da precipitações. Especificamente, o NEB tem sofrido extremos do clima, com secas e as cheias dentro de um mesmo ano. A previsão climática constitui-se em uma ferramenta em desenvolvimento para atenuar os impactos negativos causados por esses eventos extremos, para isso utilizam-se tais previsões para antecipar os fenômenos e organizar ações mitigadoras (GODDARD et al., 2001).

O objetivo deste trabalho é estudar os padrões (tendência e sazonalidade), bem como ajustar modelos de séries temporais para os dados de precipitação da cidade de Aracaju, capital do estado Sergipe, localizado na NEB.

### **2. Material**

O conjunto de dados utilizado neste trabalho é referente a precipitação pluviométrica da cidade de Aracaju, capital do estado de Sergipe. As observações coletadas são com relação ao período de 33 anos, compreendidos entre os meses de janeiro de 1986 a dezembro de 2018. A variável analisada foi a precipitação, medida em milímetros (mm), acumulada mensal. Assim, a série histórica apresentava 396 registros. Com o intuito de mensurar a acurácia da previsão do modelo ajustado, foi deixado de fora da modelagem os doze meses referente ao ano de 2018.

Os dados de precipitação foram coletados na estação convencional de Aracaju, a qual fica localizada a -10,95° de longitude, -37,05° de latitude e 4,72 metros de altitude. Este conjunto de dados está disponível na Base de Dados Meteorológicos para Ensino e Pesquisa (BDMEP), cujo órgão responsável é o Instituto Nacional de Meteorologia (INMET), cujo endereço eletrônico encontra-se em: <http://www.inmet.gov.br/portal/>.

---

<sup>1</sup> UFGD/FACET. email: [otaviojateioliveira@hotmail.com](mailto:otaviojateioliveira@hotmail.com).

<sup>2</sup> UFGD/FACET. email: [eliasmedeiros@ufgd.edu.br](mailto:eliasmedeiros@ufgd.edu.br).

<sup>3</sup> UEMS. email: [carolinabicalho@gmail.com](mailto:carolinabicalho@gmail.com).

### 3. Métodos

Neste trabalho optou-se por utilizar o modelo autorregressivo integrado de médias móveis sazonal (SARIMA) que contém duas partes: uma não sazonal ( $p, d, q$ ) e uma outra parte sazonal ( $P, D, Q$ ). O modelo geral, com uma sazonalidade  $S$ , pode ser escrito da forma:

$$\phi(B)\Phi(B^S)(1-B^S)^D(1-B)^dZ_t = \theta(B)\Theta(B^S)a_t,$$

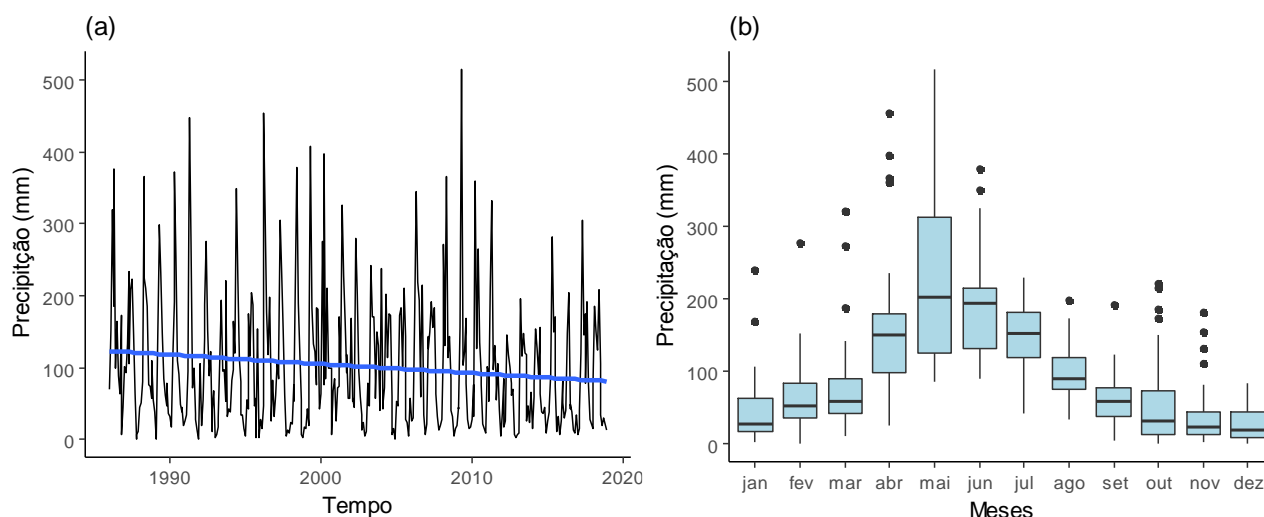
em que  $\phi(B) = 1 - (\sum_{i=1}^p \phi_i B^i)$ ,  $\Phi(B^S) = 1 - (\sum_{j=1}^P \Phi_j B^{j \cdot S})$ ,  $\theta(B) = 1 - (\sum_{k=1}^q \theta_k B^k)$  e  $\Theta(B^S) = 1 - (\sum_{w=1}^Q \Theta_w B^{w \cdot S})$ . A classe de modelos SARIMA são extremamente úteis, uma vez que são capazes de modelar padrões como tendência e sazonalidade. Consultar o trabalho de Espinosa, Prado e Ghellere (2010) para maiores detalhes da estrutura algébrica desta classe de modelos.

Para analisar se a série histórica apresentava tendência e sazonalidade foram aplicados, respectivamente, os testes de Cox Stuart (MATEUS; CAEIRO, 2014) e de Kruskal-Wallis (MYLES; DOUGLAS, 1973).

Os ajustes dos diferentes modelos da classe SARIMA deu-se por meio da metodologia Box-Jenkins, a qual é composta por um ciclo de três etapas iterativas: identificação, estimação e diagnóstico do modelo (MORETTIN; TOLOI, 2006). O teste de Ljung-Box foi aplicado nos resíduos de cada modelo ajustado com o intuito de verificar o pressuposto de que estes resíduos não são autocorrelacionados. Os critérios numéricos utilizados para seleção do modelo, mais adequado ao ajuste da série histórica, foram o AIC (Critério de Informação de Akaike) e a RMSE (Raiz Quadrada do Erro Quadrático Médio). Todas as análises estatísticas foram realizadas no software R (R DEVELOPMENT CORE TEAM, 2018), com auxílio das bibliotecas forecast (HYNDMAN et al., 2018).

### 4. Resultados

Fonte: Próprios autores.



**Figura 1**—Série temporal com linha de tendência na cor azul (a) e o gráfico de box-blot (b) para a precipitação acumulada mensal da cidade de Aracaju, no período de 1986 a 2018.

Na Figura 1(a) tem-se uma linha de tendência ajustada por regressão linear, indicando uma tendência decrescente da precipitação ao longo do tempo, sugerindo que no modelo de série temporal deve ser considerado no ajuste uma diferença não sazonal. Adicionalmente, com esta tendência decrescente, o comportamento da precipitação acumulada mensal na cidade de Aracaju, tende a ser cada vez menor nos próximos anos. Ainda nesta figura percebe-se um comportamento periódico de 12 meses, indicando que no ajuste dos modelos SARIMA deve ser inserido uma diferença sazonal na série. Entretanto, estas análises são influenciadas pelo curto período de coleta de dados, que foram de 33 anos. Na Figura 1(b) tem-se que os maiores índices de precipitação se concentraram entre os meses de abril a julho, sendo o mês de maio compreendendo os maiores registros.

O teste de Cox Stuart para verificar a existência de tendência na série histórica resultou no valor p de  $2,146 \times 10^{-4}$ , indicando que há fortes evidências estatísticas de que a série mensal da precipitação apresentou uma tendência não constante. Para esta mesma série histórica, o teste de Kruskal-Wallis forneceu um valor p de  $2,2 \times 10^{-16}$ , sugerindo a existência de sazonalidade. Ambas as indicações de tendência e sazonalidade foram consideradas ao nível de 5% de significância. Por meio destes resultados foram sugeridos diferentes modelos da classe SARIMA (Tabela 1).

**Tabela 1** – Critérios numéricos AIC e RMSE para seleção de modelo, e os respectivos valores p do teste de Ljung-Box (L-B)

SARIMA	AIC	RMSE	L-B
(1,1,0)(1,1,0) <sub>12</sub>	4461,405	96,088	0,001
(0,1,1)(0,1,1) <sub>12</sub>	4209,935	64,555	0,793
(0,1,1)(1,1,0) <sub>12</sub>	4321,825	78,893	0,624
(0,1,1)(1,1,1) <sub>12</sub>	4210,507	64,401	0,769

Fonte: Próprios autores

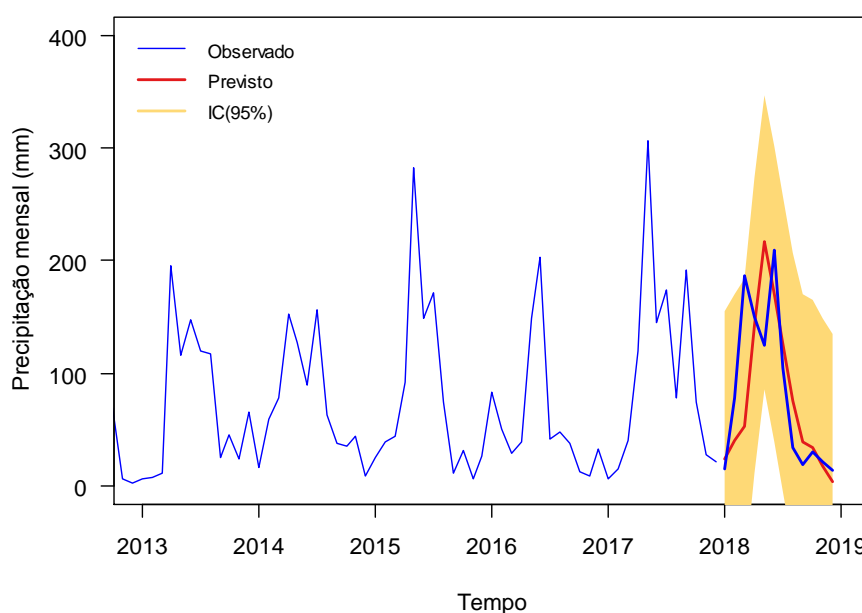
Na Tabela 1 nota-se que apenas o modelo SARIMA (1,1,0)(1,1,0)<sub>12</sub> apresentou problemas em relação aos resíduos deste modelos serem estatisticamente, ao nível de 5% de significância, autocorrelacionados. Por meio da estatística AIC o modelo mais adequado é o SARIMA (0,1,1)(0,1,1)<sub>12</sub>. Entretanto a estatística RMSE indicou que o SARIMA (0,1,1)(1,1,1)<sub>12</sub> é o que oferece, dentre os modelos sugeridos, os melhores resultados em relação a predição, embora a diferença entre estas estatísticas neste dois modelos é de apenas 0,238%. Sendo assim, levando o princípio da parcimônia, o modelo SARIMA (0,1,1)(0,1,1)<sub>12</sub> é o mais indicado para obter previsões da precipitação acumulada mensal, uma vez que apresenta apenas dois de parâmetros. Este modelo pode ser escrito da forma:

$$(1 - B^{12})^1(1 - B)^1 Z_t = [1 + 0,979B][1 + 0,995B^{12}]a_t,$$

em que, todas as estimativas desse modelo foram estatisticamente diferente de zero ao nível de 5% de significância.

Na Figura (2) é apresentada parte da série histórica da precipitação mensal acumulada na cidade de Aracaju, seguido dos valores observados e previstos pelo modelo SARIMA (0,1,1)(0,1,1)<sub>12</sub> para o ano de 2018. Nesta ano, a comparação entre os valores observados e previstos resultou em uma estatística RMSE de 51,88 mm. A maior e menor diferença entre estes valores deu-se nos meses de março e outubro, respectivamente. Análises futuras devem ser realizadas para modelar a alta variabilidade existente entre os meses, e também dentro de cada mês quando comparado os diferentes anos.

Fonte: Próprios autores



**Figura 2** - Série histórica da precipitação mensal acumulada (mm) da cidade de Aracaju e os respectivos valores previsto pelo modelo SARIMA  $(0,1,1)(0,1,1)_{12}$ , seguido do região de 95% de confiança.

## 5. Considerações finais

Os dados de precipitação acumulada mensal na cidade de Aracaju, no período de 1986 a 2018, apresentou uma tendência decrescente e uma sazonalidade de 12 meses. O ajuste da classe de modelos SARIMA a esta série histórica mostrou-se uma alternativa adequada para previsão da precipitação mensal nesta cidade.

## Referências Bibliográficas

- ANTUNES, J. L. F.; CARDOSO, M. R. A. Using time series analysis in epidemiological studies. **Epidemiol. Serv. Saúde**. 2015, vol.24, n.3.
- ESPINOSA, M. M.; PRADO, S. M.; GHELLERE, M. Uso do modelo SARIMA na previsão do número de focos de calor para os meses de junho a outubro no Estado de Mato Grosso. **Ciência e Natura**, v. 32, n. 2, p. 07-21, 2010.
- GODDARD, L.; MASON, J. S.; ZEBIAK, S. E.; ROPELEWSKI, C. R.; BASCHER, R.; CANE, M. A. Current approaches to seasonal-to-interannual climate predictions. **International Journal of Climatology**, Londres, v. 21, p. 1111-1152, 2001.

HYNDMAN, R.; BERGMEIR, C.; CACERES, G.; CHHAY, L.; O'HARA-WILD, M.; PETROPOULOS, F.; RAZBASH, S.; WANG E.; YASMEEN, F. (2018). *\_forecast: Forecasting functions for time series and linear models\_*. R package version 8.3, <URL: <http://pkg.robjhyndman.com/forecast>>.

MATEUS, A.; CAEIRO, F. An R implementation of several Randomness Tests. In: **AIP Conference Proceedings**. AIP, 2014. p. 531-534.

MORETTIN, P. A.; TOLOI, C. M. C. **Análise de séries temporais**. 2. ed. São Paulo: Edgard Blücher, 2006. 538p.

MYLES, H.; DOUGLAS, A. W. *Nonparametric Statistical Methods*. New York: John **Wiley & Sons**. 1973, p.115–120.

R CORE TEAM (2018). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.