



**UNIVERSIDADE ESTADUAL DE FEIRA DE SANTANA**

Autorizada pelo Decreto Federal nº 77.496 de 27/04/76  
Recredenciamento pelo Decreto nº 17.228 de 25/11/2016



**PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO**  
**COORDENAÇÃO DE INICIAÇÃO CIENTÍFICA**

## **XXIV SEMINÁRIO DE INICIAÇÃO CIENTÍFICA DA UEFS** **SEMANA NACIONAL DE CIÊNCIA E TECNOLOGIA - 2020**

### **ANÁLISE DE ELEMENTOS CLIMÁTICOS DE FEIRA DE SANTANA – BA NO PERÍODO DE 1999 A 2019 COM MODELOS DE SÉRIES TEMPORAIS**

**Pedro Schmidt de Brito<sup>1</sup>; Maurício Santana Lordêlo<sup>2</sup>**

1. Estagiário de Iniciação Científica, Graduando em Agronomia, Universidade Estadual de Feira de Santana, e-mail:

[brito.ps@outlook.com](mailto:brito.ps@outlook.com)

2. Orientador, Departamento de Ciências Exatas, Universidade Estadual de Feira de Santana, e-mail:

[msslordelo@uefs.br](mailto:msslordelo@uefs.br)

**PALAVRAS-CHAVE:** Séries Temporais; Climatologia; Feira de Santana.

#### **INTRODUÇÃO**

O clima é um aspecto natural que interfere de maneira significativa na agricultura. Conhecer os elementos climáticos de uma determinada região como precipitação, temperatura, umidade relativa do ar, radiação solar e dinâmica dos ventos é essencial para os profissionais da área realizarem algumas de suas atividades (zoneamento agroclimático, calendário agrícola, etc.). Para analisar a aplicabilidade dos elementos climáticos nas atividades agrícolas que dependem de uma previsão temporal, é preciso analisar a qualidade dos dados preditivos sendo necessário aplicar modelos de séries temporais aos dados meteorológicos e validá-los por meio de testes estatísticos capazes de indicar o grau de confiabilidade das previsões (FERRAZ, 1999).

Os objetivos da análise e aplicação de séries temporais no presente trabalho se resumem em descrever o comportamento da série, buscar periodicidades relevantes e fazer o que normalmente se almeja ao analisar dados climáticos: realizar projeções futuras, convencionalmente denominadas de “*forecasting*” que numa tradução direta do inglês significa “previsão” (MORETTIN & TOLOI, 2006). Para efetuar tais previsões, é necessário utilizar um modelo que é uma descrição probabilística da série temporal. Os modelos Sazonais Autoregressivos Integrados de Médias Móveis (SARIMA) foram utilizados, pois foram os que melhor se ajustaram às séries temporais com elementos climáticos, segundo critérios de acurácia que apuraram o grau de confiabilidade dos modelos e conseqüentemente das previsões feitas a partir deles.

#### **MATERIAL E MÉTODOS OU METODOLOGIA (ou equivalente)**

No presente trabalho, o software “R” (R CORE TEAM, 2020) foi utilizado na análise dos elementos climáticos do município de Feira de Santana, que tem ponto de referência nas coordenadas geográficas latitude 12°16’S, longitude de 38°58’W e altitude 234 m (BRASIL, 1992). Esta ferramenta computacional possui pacotes, funções e outros recursos poderosos para análise estatística e gráfica dos dados climáticos secundários obtidos junto ao banco de dados históricos do Instituto Nacional de Meteorologia (INMET, 2019).

Os 756 dados quantitativos foram tabulados em uma matriz unidimensional de tempo e multivariada, cujas variáveis são três elementos climáticos: precipitação total mensal (mm); temperatura média compensada mensal (°C) e umidade relativa do ar média mensal (%). Havendo disponibilidade de dados apenas a partir do ano de 1999, a dimensão temporal teve 252 intervalos de tempo já que foram coletados dados mensais até o ano de 2019, totalizando um período de 21 anos. Após a tabulação, verificou-se a ausência de alguns dados e para solucionar a ocorrência de valores faltantes, utilizou-se a função de imputação com ajuste sazonal “na.seadec” disponível no pacote “imputeTS” (MORITZ & BARTZ-BEIELSTEIN, 2017).

Na posse dos dados tabulados e completos, construiu-se um objeto do tipo série temporal “TS” (do inglês, “Time-Series”) para cada uma das variáveis climáticas estudadas, para proceder com a análise estatística e gráfica de cada elemento do clima, individualmente. Para cada objeto TS, foi construído um gráfico básico por meio da função “plot” do pacote “stats” incluído nos pacotes básicos do próprio software. Com isso já foi possível fazer uma prévia avaliação das componentes da série, verificando se há alguma das quatro componentes das séries temporais: tendência, variações sazonais (ou sazonalidade), variações cíclicas (ou ciclos) e variações irregulares (ou ruídos).

A modelagem dos dados se deu inicialmente pela divisão de cada série temporal em duas: uma série treino com valores referentes às 240 unidades de tempo iniciais e uma série de validação com as 12 unidades de tempo. Por meio da função “auto.arima” disponível no pacote “forecast” (HYNDMAN R.J. et. al., 2020) diferentes modelos foram testados e o melhor para a série de cada variável climática foi escolhido, com base no menor valor do Critério de Informação de Akaike (AIC, do inglês *Akaike Information Criterion*). Os parâmetros SARIMA definidos são usados na função “Arima” do pacote “forecast” que é aplicada na série treino para fazer o devido ajuste.

Cada série ajustada é colocada em um novo objeto do tipo TS que é avaliada por meio da função “summary” do pacote “base” do próprio *software* que determina o valor do Erro Percentual Absoluto da Média (MAPE, do inglês *Mean Absolute Percentage Error*) sendo que apenas valores abaixo de 10% foram considerados apropriados para as previsões e a partir dele foi construído um gráfico sobrepondo curva da série treino sobre a curva da série ajustada para visualizar a semelhança entre as curvas, considerando que quanto maior a similaridade entre as curvas, melhor é o modelo e o ajuste, conseqüentemente. Após validar a modelagem, foi possível fazer a predição com a série completa utilizando a função “forecast” para obter os valores correspondentes a 12 intervalos de tempo à frente da última observação, isto é, os 12 meses de 2020.

## **RESULTADOS E DISCUSSÃO**

Seguindo a metodologia descrita, o banco de dados com os elementos climáticos tabulados foram divididos em três objetos de séries temporais distintos, cada um como 252 valores mensais sucessivos, abrangendo o período de janeiro de 1999 a dezembro de 2019, num total de 21 anos. A Figura 1 demonstra o comportamento de cada uma das séries no período em questão.

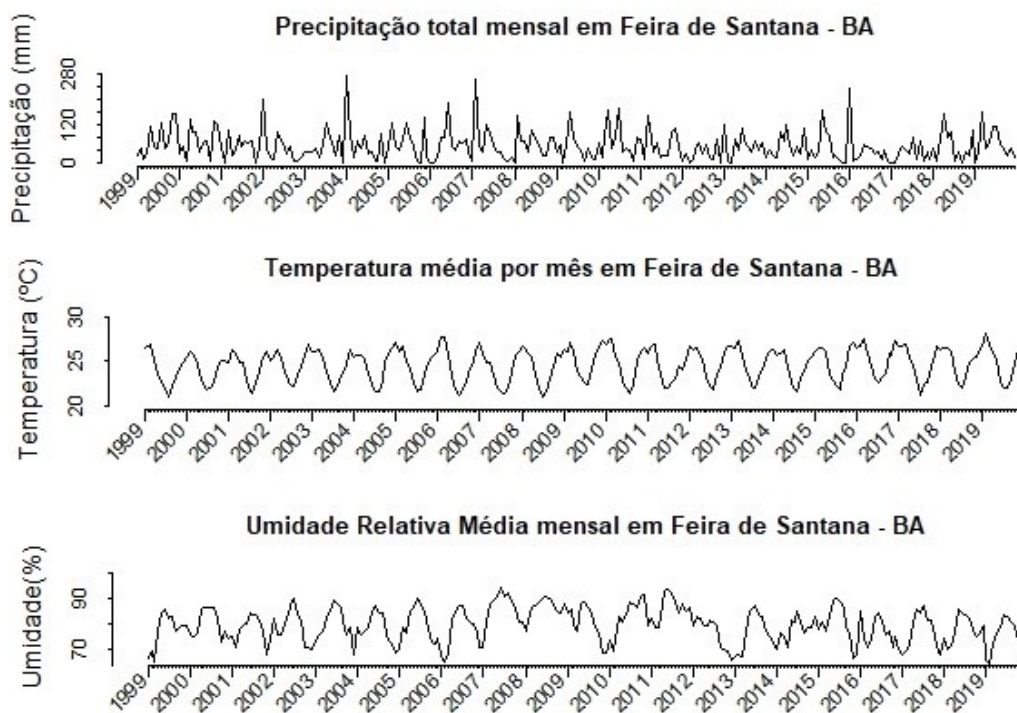


Figura 1 - Representação gráfica dos elementos climáticos no período de 1999 a 2019.

Analisando as séries pelos gráficos básicos, elas demonstram estacionariedade, não apresentam tendência e apresentam variações sazonais. Todavia, apenas a série de precipitação demonstra uma irregularidade considerável, provocada pela componente das variações aleatórias que já são esperadas no regime pluviométrico, por se tratar do município de Feira de Santana na mesorregião Agreste (zona de transição entre a Zona da Mata e o Sertão) onde ocorre má distribuição das chuvas (OLIVEIRA, 2016).

Após passar cada uma das séries na função “auto.arima” selecionou-se os modelos SARIMA (2, 0, 2) (1, 1, 0) para a série de temperatura, que registrou 464,2 no valor do AIC e SARIMA (1, 0, 2) (2, 1, 0) para a série de umidade que registrou 1391,6 para o valor do AIC. No caso da série de precipitação, o melhor modelo encontrado na função foi o SARIMA (0, 0, 0) (1, 0, 2) com AIC de 2667,4, sugerindo o ruído branco que indica aleatoriedade e falta de interdependência entre as observações.

Os parâmetros desses modelos colocados na função “Arima” são aplicados nas séries treino de cada variável climática e a série ajustada resultante tem sua acurácia medida pelo erro percentual absoluto da média, o MAPE, que na série de precipitação foi classificada como “Inf” que na linguagem estatística do R, é o valor infinito, portanto o valor é muito elevado e descarta a possibilidade de fazer previsões a partir do modelo em questão. No entanto, as séries ajustadas de temperatura média e umidade relativa do ar (URA) tiveram o MAPE de 1,8% e 3,9%, respectivamente, indicando uma acurácia excelente para previsão.

A fim de analisar a qualidade do modelo para previsão, foi feita a comparação da curva do modelo preditivo com a curva da série de validação. Os valores MAPE para definir a acurácia desse teste preditivo na comparação das curvas apontaram o valor de 2,3% e 7,1% para as séries de temperatura média e umidade relativa do ar (URA), respectivamente, o que comprova uma excelente acurácia para previsão. A Tabela 1 mostra os valores das previsões para o ano de 2020.

Tabela 1 – Sumário de previsões para as variáveis climáticas.

MÊS DE 2020	TEMPERATURA MÉDIA (°C)	URA (%)
JAN	26,4	67,8
FEV	27,3	66,5
MAR	26,6	70,0
ABR	26,1	76,2
MAI	24,5	82,8
JUN	22,5	82,8
JUL	21,9	83,7
AGO	22,5	80,7
SET	23,9	78,7
OUT	24,9	73,8
NOV	25,8	71,1
DEZ	26,4	72,2

## CONSIDERAÇÕES FINAIS

A precipitação total mensal não permite ajuste com nenhum modelo SARIMA já que os dados não apresentam dependência temporal, o que é comprovado pelo valor do erro percentual absoluto da média (MAPE) que é classificado pelo software R como “Inf” que é valor infinito, portanto muito elevado.

O modelo SARIMA (2,0,2)(1,1,0) para série da temperatura média compensada mensal e o modelo SARIMA (1,0,2)(2,1,0) para série da umidade relativa do ar média mensal foram os que melhor ajustaram as séries temporais que tiveram um MAPE de 2,3% e 7,1%, respectivamente, que são valores de acurácia excelentes para efetuar as previsões demonstradas nos resultados.

## REFERÊNCIAS

- BRASIL, Ministério da Agricultura. **Normas climatológicas**. Rio de Janeiro: MA, 1961-1990. Brasília: 1992. 84p.
- FERRAZ, Marcelo Inácio Ferreira. **Uso de modelos de séries temporais na previsão da série de precipitação pluviiais mensais no município de Lavras- MG**. Lavras: UFLA, 1999. 97p. (Dissertação - Mestrado em Estatística e Experimentação Agropecuária).
- HYNDMAN R.J. et. al. **forecast: Forecasting functions for time series and linear models**. R package version 8.12, 2020.
- INMET, Instituto Nacional de Meteorologia. **BDMEP - Banco de Dados Meteorológicos para Ensino e Pesquisa**. Disponível em: <<https://bdmep.inmet.gov.br>>. Acesso em: 7 dez. 2019.
- MORETTIN, P.A. & TOLOI, C.M. **Análise de Séries Temporais**. 2. ed. São Paulo: Edgard Blücher, 2006.
- MORITZ, S. & BARTZ-BEIELSTEIN, T. **imputeTS: Time Series Missing Value Imputation in R**. The R Journal, v. 9, n. 1, p. 207-218, 2017.
- OLIVEIRA, L. B. **Uso e manejo da água na região semiárida do nordeste do Brasil**. Anais da Academia Pernambucana de Ciência Agronômica, v. 11, p. 50–64, 2016.
- R CORE TEAM. **R: A language and environment for statistical computing**. R Foundation for Statistical Computing, Vienna, Austria, 2020.