Comparative analysis of time series modeling in the prediction of mortality due to Acute Myocardial Infarction in Brazil (2010-2022)

Elisângela Ap. da Silva Lizzi[1], Vitor Hugo Amadeu Da Silva[2]

1. Departamento Acadêmico de Matemática e Programa de Pós-Graduação em Bioinformática (UTFPR-CP)
2. Departamento Acadêmico de Computação, Graduação em Engenharia da Computação (UTFPR-CP)

X-meeting 2025

INTRODUCTION

Acute Myocardial Infarction (AMI) is the leading cause of death in Brazil. It is estimated that there are 300,000 to 400,000 cases of heart attack each year, and that one death occurs in every 5 to 7 cases. AMI results from cellular necrosis in the heart muscle due to the sudden obstruction of blood flow by clots, and can affect different regions of the heart depending on the location of the occlusion.

The goal of this study is to compare the predictive performance of three time series modeling approaches: Holt-Winters exponential smoothing, SARIMA, and LSTM neural networks, in estimating mortality from AMI in Brazil, aiming to identify the most appropriate method to support public health management and health surveillance.

METHODOLOGY

The methods are based on stages by CRISP-DM(CRoss Industry Standard Process for Data Mining), in the acquisition of data through the open access repository of the Mortality Information System (SIM) of the Ministry of Health, data were collected between the years 2010 to 2022. In the pre-processing stage, structured organization was carried out, with grouping into months and calculation of standardized rates: number of deaths per 100,000 inhabitants, considering all brazilian territory. Time series models were adjusted and three approaches were considered: Holt-Winters smoothing, SARIMA models identified via autocorrelation functions (ACF/PACF) and LSTM-type neural networks with Adam optimizer, all procedures were implemented in Python program, in addition to the diagnosis of the models by adjustment metrics and white noise analysis, the out-of-sample adjustment was verified by accuracy criteria and calculation of the mean absolute percentage error (MAPE).

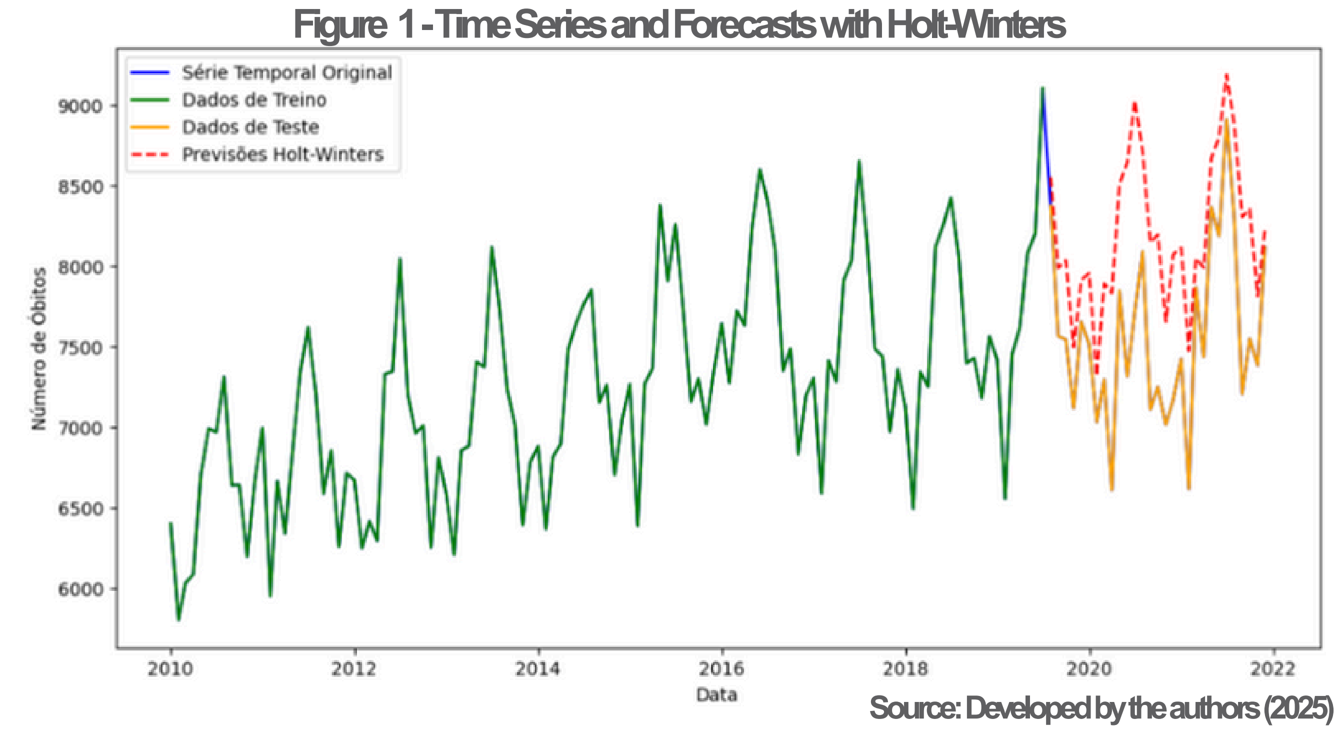
RESULTS

The results show that the smoothing based on Holt-Winters exponential (MAPE=8.5%), considering trend and seasonality components, proved to be practical and intuitive, being an approximation of temporal dynamics. The SARIMA(1,1,1)(1,1,1)12 model demonstrated a good fit to the historical series (MAPE=5.7%), being especially effective in capturing trends and seasonality, as long as the parameters are correctly identified through the autocorrelation (ACF) and partial autocorrelation (PACF) functions. It is a dynamic model that captured the oscillation in the short and medium term with good profiling to the data. The LSTM (Long Short-Term Memory) neural networks, implemented with the Adam optimizer and four optimized layers, achieved a competitive predictive performance (MAPE=6.0%), surpassing the Holt-Winters model and approaching SARIMA. This architecture has demonstrated the ability to model non-linear relationships and long-term temporal dependencies in the series, benefits absent in traditional methods. However, its 'black box' nature limits the interpretability of the results, which is a critical disadvantage in public health applications, where model transparency is essential for decision-making.

CONCLUSION

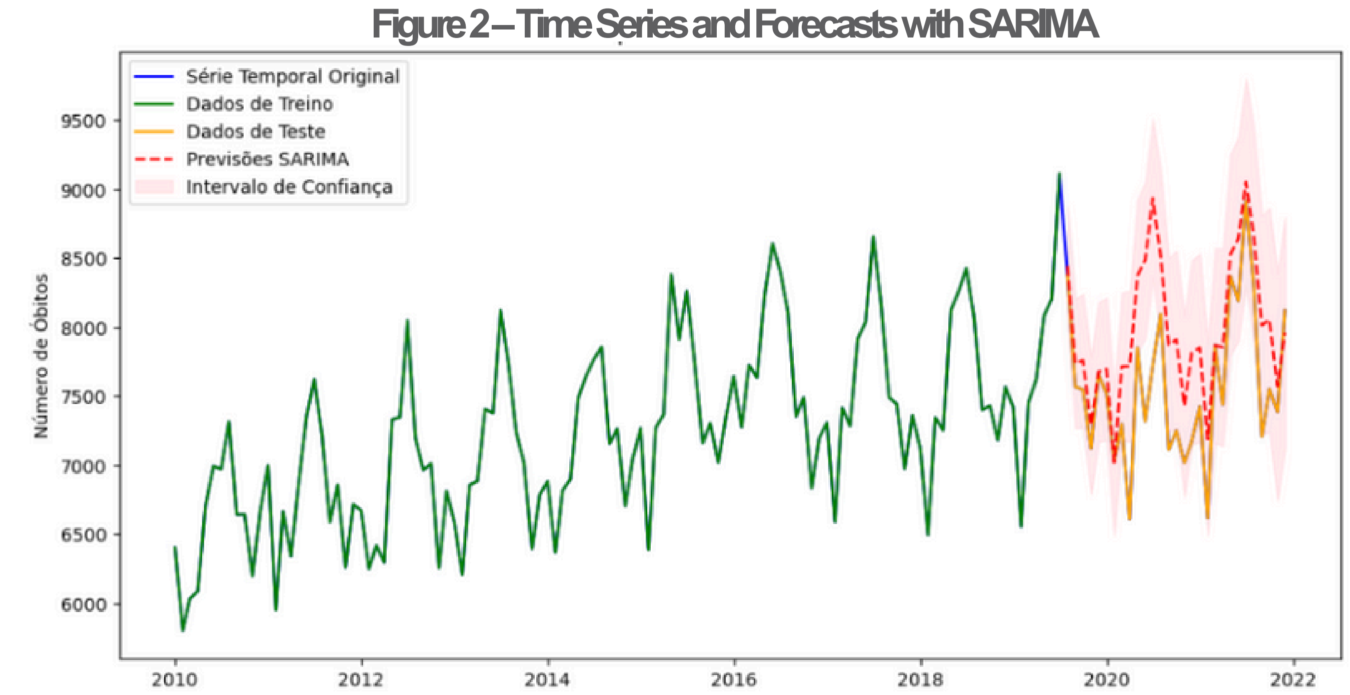
It is concluded that the SARIMA(1,1,1)(1,1,1)12 model stands out for its robustness, interpretability of parameters within the health context and better out-of-sample prediction. The Holt-Winters method is simple and fast, but it was shown to be sensitive to abrupt oscillations, while LSTM offers good prediction in series with complex patterns, but without interpretability of parameters. In terms of health surveillance, these methods support decision-making at a decentralized management and governance levels.

Figure 1 - Time Series and Forecasts with Holt-Winters



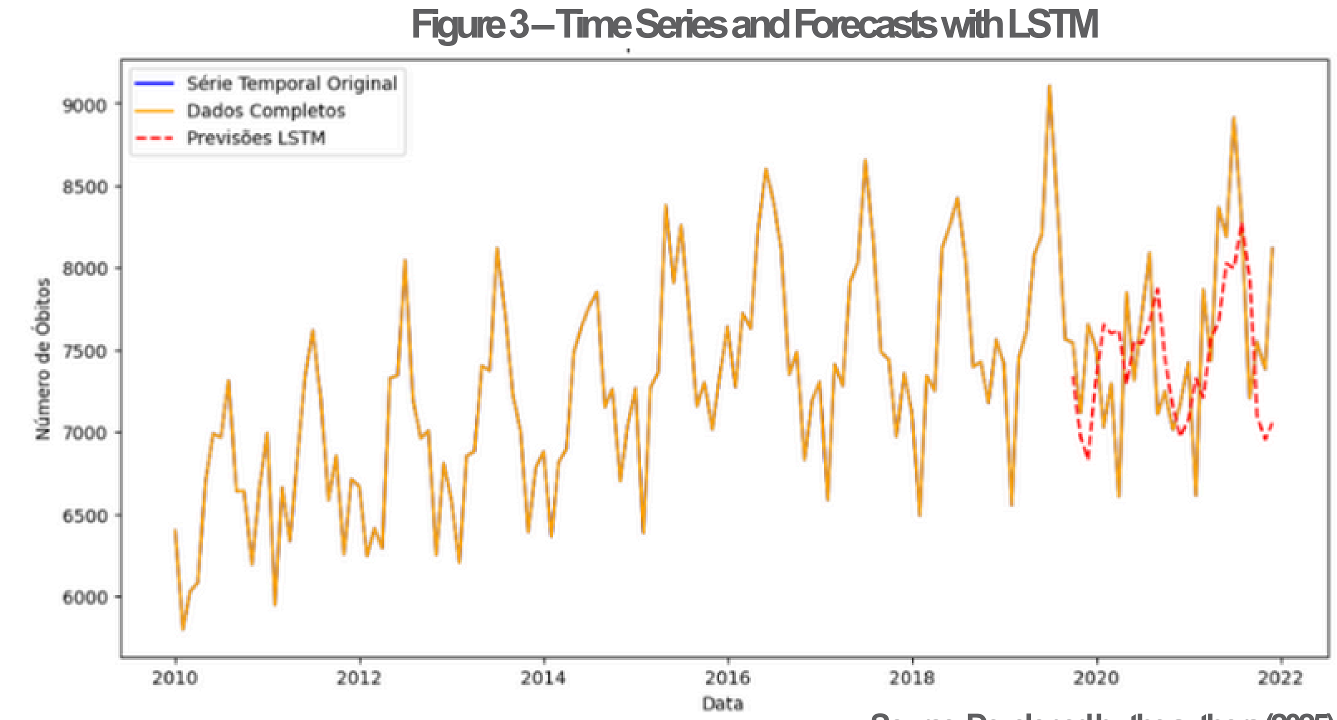
Source: Developed by the authors (2025)

Figure 2 – Time Series and Forecasts with SARIMA



Source: Developed by the authors (2025)

Figure 3 – Time Series and Forecasts with LSTM



Source: Developed by the authors (2025)

Criterion	SARIMA	Holt-Winters	LSTM (with Adam)
Model Type	Statistical modeling, seasonal autoregressive integrated moving average	Exponential smoothing with trend and seasonality	Deep Learning using recurrent neural network architecture
Modeled Components	Seasonality, moving average, and autoregressive components	Trend and seasonality	Non-linear sequential dynamics
Main Parameters	ARIMA(p, d, q); Seasonal (P, D, Q) with period *s*	Alpha (level), beta (trend), gamma (seasonality)	LSTM layers, number of units, and density
Fit to Historical Series	Good fit, dependent on correct identification via ACF/PACF	Direct and intuitive fit, easy to calibrate	Captures patterns with good performance
Short-Term Forecasting	Reliable	Good, but sensitive to sudden changes	Reliable, but require proper training
Result Interpretation	Interpretable, based on components	Intuitive, though less robust	Less interpretable, considered a "black box"
Predicted Trend Curve	Captures seasonality and cyclicality	Captures smooth patterns	Captures fluctuations and cyclicality

UTFPR UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ

bioinformática programa de pós-graduação