A Comparison of Classical Time Series Models and Machine Learning LSTM Model to Forecast Paddy Production in Sri Lanka

IBI Sandaruwani[1#] and RAB Abeygunawardana[1]

[1] Department of Statistics, Faculty of Science, University of Colombo, Sri Lanka

[#] imesha4sandaruwani@gmail.com

The Autoregressive Integrated Moving Average (ARIMA) family models and the Exponential Smoothing family models are the most widely used and successful conventional techniques for univariate time series forecasting. As a result of the recent development in more sophisticated machine learning methodologies, such as the Long-Short-Term Memory modelling approach, new algorithms are being developed to evaluate and forecast time series data. The objective of this study was to identify the best time series forecasting model among classical time series models and machine learning LSTM model to forecast the annual paddy production of Sri Lanka. The results showed that the estimated error of ARIMA & Double Exponential Smoothing (DES) models is much higher than the estimated error of the preferred LSTM model based on the RMSE, MAE, and MAPE values. Hence LSTM outperforms traditional-based algorithms like ARIMA and smoothing models for forecasting the paddy production in Sri Lanka. The forecasts for paddy production from 2022 to 2024 were 4.92, 4.89, and 5.34 million Mt respectively. Researchers can use this model to forecast the paddy output of Sri Lanka, and it should be continuously improved by including new data.

Keywords: *time series, forecasting, paddy production, Sri Lanka, ARIMA, double exponential smoothing, LSTM*