UNIVERSITI TEKNOLOGI MARA

A NOVEL HYBRID HOLT INTEGRATED MOVING AVERAGE (HIMA) MODEL FOR ENHANCED FORECAST ACCURACY IN TREND DATA SERIES

NURIN QISTINA BINTI MOHAMAD FOZI

Thesis submitted in fulfilment
of the requirements for the degree of
Master of Science
(Statistics)

Faculty of Computer and Mathematical Sciences

October 2025

ABSTRACT

Holt's method is one of the most popular forecasting techniques for time series, particularly with trend variations. Unfortunately, due to limitations of Holt's method, such as sensitive parameter selection, the linearity assumption requirement for such a model can lead to overestimation or underestimation, especially for different trend variations. This study aims to introduce a hybrid Holt's method by integrating the traditional Holt's method and the Moving Average (MA) in Box-Jenkins methodology called Holt Integrated Moving Average (HIMA) to improve forecast accuracy for different trend variations. Eighteen simulated datasets, with six different sample sizes, such as n=50, 100, 150, 500, 1000, 2500 and three different trend variations: linear, cubic, and quadratic were used to evaluate the model performance. Besides that, two real datasets, which Consumer Price Index (CPI) and PETRONAS share price, were used to validate the model performance. The model performance of the proposed model was compared with several traditional models, such as Holt's method, Damped Trend Method (DTM), and Autoregressive Integrated Moving Average (ARIMA), for both simulation and real datasets. The forecasting accuracy was evaluated in terms of Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE) and percentage of forecast accuracy. The results show that the proposed HIMA model performs best across all three trend variations. The results also consistent founded in real datasets used in this study with forecast accuracy of 99.006%, for CPI and 98.9570% for PETRONAS datasets . Hence, the proposed HIMA model results in better accurate forecast accuracy compared to traditional models.

ACKNOWLEDGEMENT

I am deeply grateful to Almighty God, who has been my strength and guide throughout this journey, granting me wisdom, perseverance, and resilience to complete this study. I would like to take a moment to acknowledge myself for the hard work, dedication, and countless hours I put into this research. It has been a challenging yet rewarding experience, and I am proud of the commitment and growth I have demonstrated. My heartfelt thanks go to my beloved parents, whose unwavering support, love, and encouragement have been my source of motivation. Their prayers and belief in my dreams have been instrumental in this journey. I am sincerely thankful to my main supervisor, Dr Nurhasniza Idham Abu Hasan with her guide to write the thesis perfectly, and co-supervisors, Mr Azlan Abdul Aziz, lead and provide the idea for my analysis and writing, and Dr Siti Meriam Zahari whose guidance and insightful feedback have been invaluable in shaping this work. Their expertise and patience have pushed me to reach new heights, and I am grateful for their mentorship. A special thank you goes to my friends and colleagues, who have supported me in various ways, whether through collaboration, constructive criticism, or simply their companionship during late-night study sessions. I would also like to express my gratitude to my lecturers and staff at Universiti Teknologi MARA (UiTM) for providing the resources and facilities necessary for this research. Their dedication and assistance have been pivotal in bringing this work to fruition. Finally, I extend my appreciation to everyone who contributed to this success in one way or another. Thank you for your kindness, encouragement, and support.

TABLE OF CONTENTS

CHAPTER 1
INTRODUCTION

1.1 Background of Research

Forecasting plays a central role in statistical and analytical processes, focusing on predicting future events by analyzing historical data, identifying patterns, and considering factors that may affect forecast accuracy (Hyndman and Athanasopoulos, 2021). Over time, various definitions of time series forecasting have emerged. Makridakis et al. (1998) define time series forecasting as using statistical models to predict future values based on past observations, emphasizing the recognition of patterns and trends in historical data. Another definition described time series forecasting involves analyzing values occur in a specific time order, highlighting the importance of capturing underlying patterns. This aspect is crucial for accurate forecasting, especially in fields like epidemiology, where time series analysis is used to predict disease incidence (Hamilton, 1994; Pan et al., 2020). In addition, time series forecasting characterized as a system behaviour that forecasts the future based on current and historical information (De Gooijer and Hyndman, 1982; Zhang et al., 2024). Conversely, Lazim (2017) defines forecasting as a statistical approach that involves analyzing data to underlying trends, and aiding decision-making for accurate future predictions. Based on the various definitions discussed, forecasting is a specialized branch of statistics that uses past or historical data to predict future events with the highest degree of accuracy, by employing the most appropriate models or methods.

Forecast accuracy is a key challenge in statistical modeling, particularly in time series analysis, where the goal is to predict values that closely match actual outcomes. According to Chambers et al. (1971), although the concept of forecasting is straightforward, accurately identifying the components of time series data is a more complex task. The choice of forecasting method depends on factors like data characteristics, the presence of time series components, available historical data, and the forecasting horizon (Sharif and Hasan, 2019; Twumasi et al., 2021). Studies have shown that selecting the right model can greatly influence forecast accuracy. For example, Shaharudin et al. (2021) used the Repetitive-Singular Spectrum Analysis (RF-SSA) model to predict daily COVID-19 cases in Malaysia, but the model's forecast was