UNIVERSITI TEKNOLOGI MARA

APPLICATION OF SEASONAL AUTOREGRESSIVE INTEGRATED MOVING AVERAGE (SARIMA) MODEL IN PREDICTION OF DENGUE CASES IN KUANTAN

ZULKIFLI BIN ABDUL HADI

Thesis submitted in fulfillment
of the requirements for the degree of
Bachelor of Environmental Health and Safety
(Hons.)

Faculty of Health Sciences

January 2020

ACKNOWLEDGEMENT

In the name of Allah, The Most Gracious, The Most Merciful.

Assalamualaikum and Alhamdulillah, all praise to Allah S.W.T The Supreme Lord of the Universe. Peace and blessing to Nabi Muhammad S.A.W., all prophets and their families. I praise Allah S.W.T. for the strength and His blessings in which has enabled me the opportunity to complete this thesis.

Thousands of thanks to my parents Mr. Abdul Hadi bin Shamsuddin and Mrs.
my wife Mrs. Nurul Akhillah binti Hassan Basri for their great love and support. I wish to express my sincere appreciation to my supervisor Prof. Madya Rodziah binti Ismail for being instrumental in shaping my studies and key figures whose constructive criticism and advice contributed immensely towards the completion of this study. Without her great assistance, support and patience, this thesis would not be completed on time. Not to forget, I would like to thank all the lecturers in the Department of Environmental Health and Safety, Faculty of Health Sciences who always share their thoughts, knowledge and advice throughout my study in UiTM Puncak Alam. Only God can reward all of you with goodness.

Last but not least to my course mates from HS 243, I would like to thank them for their help, support and encouragement from the start until I finished this thesis. Without the assistance and advice of those mentioned above, whether is perceive or not, I would not be able to finish this work as it should be.

TABLE OF CONTENTS

ABSTRACT

A retrospective study was carried out using epidemiological data in Kuantan, Pahang. The confirmed dengue cases from the year 2011 to 2018 was retrieved and analyzed using time series analysis. The time series model could potentially provide useful information that could be further used to facilitate the planning of public health interventions in an effort to minimize dengue outbreaks. The objective of this study was first, to assess the application of Seasonal Autoregressive Integrated Moving Average (SARIMA) model to predict dengue cases for 2019 by analyzing the trend of confirmed dengue cases (2011-2018); secondly, to compare the model's parameters accuracy to be used in the prediction of monthly dengue cases in Kuantan for the year 2019 and lastly, to perform the forecast of dengue cases using SARIMA with Box - Jenkins approach. The model was fitted with monthly confirmed dengue cases (2011-2018) and validation of the prediction was made using the dengue cases from January to December 2018. The study revealed that SARIMA $(0, 1, 0)\ (3, 0, 2)_{12}$ was the best fit model and could be used to extrapolate the cases up to twelve months in advance. The prediction of the cases in 2019 was relatively close to the actual cases within the confidence interval limit. Thus, the model derived from this study has the capability to not only forecast but also anticipate the future dengue cases. This would in turn enhance the current intervention program which is vital in minimizing the burden of the disease in Kuantan specifically.

Keywords: *Dengue Fever, Time Series Analysis, Prediction, SARIMA*

CHAPTER ONE

INTRODUCTION

1.1 Background of Study

Dengue has now become an important public health problem in Malaysia (Bujang, et al., 2017). Planning for the disease prevention and control intervention is one of an important step is reducing the risk of new infection. Planning also can guide or help to develop an adequate intervention plan which can help to reduce the dengue fever burden in the country. The planning process will be more adequate if there is supported with a statistical method which based on a strong scientific basis. Analysis of the time series is one of the techniques that can be used to interpret the dengue fever occurrence by understanding the possibilities of the disease in the future by studying its previous trend. An evaluation of the time series is a continuous sequence of numerical data points. In investing, a time series represents the movement of the specified data points, such as the price of a commodity, over a specified time period with regularly reported data points. There is no minimum or maximum amount of time to be included, requiring the data to be obtained in a way that offers the details the policymaker or analyst is searching for. Most often, it's relates to trend analysis, cyclical fluctuation analysis and issues of seasonality. For instance, in economic closing stock's price can be analyse daily, weekly and monthly basis. Delving a bit deeper, the analysis also can shows whether the stock's time series show any seasonality to determine if it goes through peak and troughs at regular time each year. The onset dengue infection in Malaysia first dated back in 1901, following the transmission from Singapore to Penang (Skae, 1980). First epidemic outbreak was recorded in 1973 which caused a total of 969 cases and 54 death (Wallace, et al., 1980). The epidemic situation continued to deteriorate, with the nation's city dwellers rapidly infested with disease. A sequence of outbreaks of 1,487 cases with 54 deaths reported in 1973, 2,200 cases with 104 deaths in 1974 and 3,006 cases reported in 1982,