

CHAPTER 1: THE PROBLEM

Many things in the world such as weather (like: rainfall rate, season time, etc.), stock exchange value, currency value, natural condition (like, disaster possibly), price of items can be predicted. Meanwhile, many methods can be used for predicting, like genetic algorithm, neural network, support vector machine, autoregressive integrated moving average (ARIMA), regression, etc. This thesis, attempts to predict the prices of staple food materials using combination of ARIMA and regression.

1.1 Rationale

Various staple food materials can be found in Asia. Staple food materials become daily needs of Indonesian people in the market. According to [1], [2] the main staple food in Indonesia is rice. However according to Indonesian Industry and Trading Agency (called Dinas Perindustrian dan Perdagangan) of West of Java, there are other popular staple food materials in market, namely red chili, curly chili, green cayenne pepper, red cayenne pepper, onion, and garlic. In many countries, especially in Indonesia, the prices of staple food materials often fluctuate as a result of various reasons such as political turmoil, seasonal events or celebration for example national holidays which raises the costs of prices because the high demands of raw materials, fuel oil, etc., - including the weather (rainfall, temperature, moisture), and rupiah's exchange rate may affect the prices of those staple food materials, demand-supply, and import. Therefore, the price fluctuation of staple food can be a trending topic and influences the prices of some other products in market. The high fluctuation of staple food price in Asia, especially in Indonesia can make everyone's anxiety and threatening government stability. Stable price of staple food brings various benefit for developing countries, for instance, this can improve the lives of poor farmers and costumers, and avoid poverty traps [3]. This make, the government able to have good plans to minimize the fluctuation of staple food price in market, and one of government programs is to make a standard price for staple food materials, called Harga Eceran Tertinggi (HET).

There are several studies concerning with the price prediction of the staple foods using several methods and types of commodities like, the study conducted by N. Hemageetha and G.M. Nasira [4]. They applied Adaptive Neuro-Fuzzy Inference System to predict the price of tomatoes. Then, M. Subhasree and Mrs. C. Arun Priya [5] applied Neural Network model based on Genetic Algorithm to predict the price of tomatoes. The study conducted by [6] applied some methods such as back-propagation neural network (BPNN), neural network based on genetic algorithm, radial basis function neural network (RBFNN) and the integrated prediction model. The study attempted

to predict the price of *lentinus endodes* (mushrooms). Those studies, only applied one variable, namely is the prices of staple food materials tomato, and *lentinus endodes* or mushroom. Therefore, they had some drawbacks mentioned in [6] study, that is develop the research in terms of improving the accuracy of predictions, and as well as adding more variables and other important factors that influencing staple food material prices.

There are some studies or researches using autoregressive integrated moving average (ARIMA) for forecasting. Research article [7] state that ARIMA model or methodology have been implemented on many research for forecasting and analysis to produce short term prediction result. That is because ARIMA models are relatively strong and useful than another models that have complex structural to short-run forecasting. ARIMA have been applied in many case of economics and finance field. That is because ARIMA have a set of activities for identifying, estimating, and diagnosing ARIMA models with time series data. It makes ARIMA is reputable methods in financial forecasting. Some studies have using ARIMA are study by [7] for stock price prediction by using stock data from New York Stock Exchange. Another research is [8] using ARIMA to forecasting oil palm price in Thailand and produce ARIMA models for forecasting farm price, whole sale price, and pure oil price. Another research [9] using ARIMA to predict next-day electricity prices in Spain and California electricity markets. Hanan and Rissanen in research [10] state that ARIMA can computed and combined with standard regression. On research [11] combined ARIMA and regression models for prediction of daily and monthly clearness index, also for diffuse and total diffuse solar radiation estimation.

Based on studies in [4], [5], and [6] the researchers only applied one variable, also [6] state some future works can be done, which are improving the accuracy of predictions, and adding more variables and other important factors that influencing staple food material prices, in this thesis will improve some works for predicting staple food materials prices. This thesis will applies combination of autoregressive integrated moving average (ARIMA) and regression models, which are multiple linear regression and Fourier regression, to predicting the prices of six staple food materials as dependent variable, and adding other three more variables, namely the factors that influence the prices of staple food materials as independent variables. Like [10] state that ARIMA can computed and combine with standard regression, also research [11] that have use the combination of ARIMA and regression to forecasting.

1.2 Theoretical Framework

Autoregressive integrated moving average (ARIMA) is also known as Box-Jenkins method which is often used to forecast time series data. [12], claims there are few problems for researchers to use autoregressive integrated moving average (ARIMA) models for forecasting, one of them is

many obstruction in ARIMA models are hard to be implemented. But recently, ARIMA is not difficult method to use anymore. It has been many research and experiments are done to make automate ARIMA modelling in the last 35 years. Also, there are some studies that have implemented ARIMA to seek stationarity as well as the application of seasonal and non-seasonal data, like Hanan and Rissanen in research [10], Gomez in research [13], and Liu in research [14].

Regression method is a statistic method to discover relationship between dependent variables and independent variables. A regression approach can be applied using three approaches, parametric, semiparametric, and nonparametric (see [15] for more detail). In this paper, the two regression type models are applied, i.e., multiple linear and Fourier regression model.

1.3 Conceptual Framework/Paradigm

There are four variables used in this thesis, i.e. the price of staple food materials as dependent variable (red chili, curly chili, red cayenne pepper, green cayenne pepper, onion, and garlic), fuel oil (premium and solar), and weather (rainfall rate) as 3 (three) independent variables. Figure 1.1 shows a diagram about the general description of the process that is conducted in this thesis.

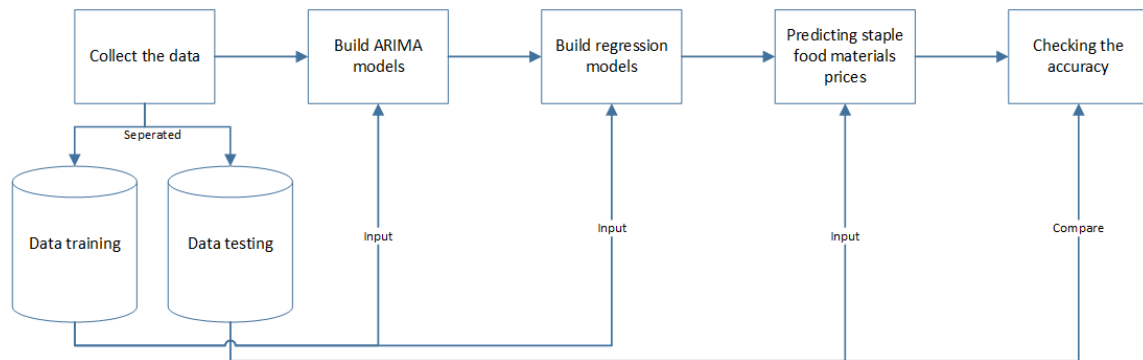


Figure 1.1 - General Description of Predicting Staple Food Material Prices

1.4 Statement of the Problem

The problem in this research that is going to be solved is how to predict the prices of staple food materials. Those prices are very volatile data with multivariate factors that influence the prices using combination autoregressive integrated moving average (ARIMA) and regression models. The other thing that must be examined is about the accuracy of prediction results, and make sure these results of price predictions of staple food materials can be used for any instance or people that needs staple food materials price predictions.

1.5 Objectives and Hypothesis

The first objective in this research is the combination of ARIMA and regression models can be used to predict the prices of staple food materials by adding other price influence factors as

independent variables. The second is the combination of ARIMA and regression models can produce a good prediction results and useful for any instance or people (farmers and consumers). Another object is to see the significance of accuracy percentage differences between ARIMA with multiple linear regression and ARIMA with Fourier regression.

The hypothesis of this research is that by adding more variables, which are staple food material prices as dependent variable, and three price fluctuation influence factors as independent variable, based on ARIMA and regression models (multiple linear regression and Fourier regression) can improve and produce higher prediction accuracy.

1.6 Assumptions

Some assumptions that must be cleared in this thesis are price influence factors to be used are the factors that the data are available, there are premium and solar price, and rainfall rate. The other thing is the selection of premium and solar price because almost all transportation vehicles that brings staple food materials using premium or solar. Also, the selection of rainfall rate because in West of Java have difference situation of rainfall than another city in Indonesia, where sometimes it rains in dry season. The last thing that must be cleared is the accuracy results percentage are obtained by comparing prediction results with the real prices which data already exists.

1.7 Scope and Delimitation

This thesis applies four variables, there are the prices of staple food materials, premium, solar, and the rainfall rate. For the data of staple food material prices are taken from Trading and Industry Agency of West Java, the prices of red chili, curly chili, red cayenne pepper, green cayenne pepper, onion, and garlic. All data are collected from 2015 until July 2019 and arranged in a week period. This thesis focuses on comparing the prediction results ARIMA with multiple linear regression results and ARIMA with Fourier regression results. The outputs in this thesis are price predictions in next 3 (three) weeks, the accuracy, and errors rate using Mean Square Error (MSE).

1.8 Importance of the Study

Based on [16], making a decision is needed, like (a) the application of policies in Indonesia and (b) formulating announcement of prices of food staples, this thesis aims for helping government to make a price standard or Harga Eceran Tertinggi (HET), and can help farmers and consumers to make decision their own business related to staple food materials prices. Another important thing to learn is how to combine ARIMA with regression models to predict staple food materials prices.