JOURNAL OF MARITIME LOGISTICS
Universiti Malaysia Terengganu

Journal of Maritime Logistics is an international multidisciplinary journal. It collects and publishes scholarly articles on topics related to the maritime industry. It provides the results of the latest research and analysis on foreland, seaports and the hinterland. It also explores substantial topics, including logistics, policies, operations, management and navigation related to three main agenda — ocean activities, seaports and the hinterland.

eISSN: 2805-5195 Journal homepage: https://journal.umt.edu.my/index.php/jml/index

TIME – SERIES MODELLING OF FOOD SECURITY INFLATION IN MALAYSIA USING AN ARIMA MODEL AS A MACHINE LEARNING APPROACH

Qistina Mohamad Rosni*[1] and Mohamad Rosni Othman*[2, 3]

To cite this article: Qistina Mohamad Rosni and Mohamad Rosni Othman (2022): Time – series modelling of Food Security Inflation in Malaysia using an ARIMA Model as a Machine Learning Approach, Journal of Maritime Logistics.
DOI: https://doi.org/10.46754/jml.2022.12.005

Journal of Maritime Logistics
https://journal.umt.edu.my/index.php/jml

TIME – SERIES MODELLING OF FOOD SECURITY INFLATION IN MALAYSIA USING AN ARIMA MODEL AS A MACHINE LEARNING APPROACH

Qistina Mohamad Rosni*[1] and Mohamad Rosni Othman[2,3]

[1]Faculty of Computer Science & Information Technology, University of Malaya. Malaysia [2]Faculty of Maritime Studies, Universiti Malaysia Terengganu, Malaysia. [3]Malaysian Maritime Logistic and Transport Centre, Universiti Malaysia Terengganu, Malaysia.

ABSTRACT

After air and water food, is the most important thing that we as humans need to survive and getting wholesome food is becoming more and more difficult. Food security refers to the accessibility and availability of the food resources. A household is considered food-secure if there is no starvation in every family member. With the burgeoning population in the world nowadays, food security become a significant problem across the globe by every country and international organizations. The objective of this study was to forecast the prices of food by category based on the 7 years data from time-series consumer price index reported by the Department of Statistics Malaysia (DOSM) from 2014 to 2021. The study considered Autoregressive Integrated Moving Average (ARIMA) processes to forecast the future trend of the food prices. The ARIMA model for forecasting food prices showed good agreement and stationery concerning the observed data on food prices based on the Augmented Dickey Fuller (ADF). The results show the ARIMA model to be a suitable method for analyzing statistics. In data-poor food prices situations, this method can support potential evaluations of future food prices for decision making and effective management.

Keywords: ARIMA: Machine learning, time-series, food prices; food security

Introduction

Climate change is putting food production at risk. Long-term research shows that the global average temperature is increasing significantly due to the large amount of greenhouse gas emissions, and hydrological extremes become more frequent (Banerjee *et al*., 2021, Wasco & Sharma, 2017, Yun *et al*., 2021). It will seriously hinder sustainable development and the economy. The results of climate simulations by the Intergovernmental Panel on Climate Change (IPCC) show that the global average temperature is expected to increase 2°C in the next 100 years (Vicente-Serrano *et al*., 2014). The frequency and magnitude of future floods will change along with the climate (Chen & Grasby, 2014, Duan *et al*., 2017), which may lead to increased flood risk at reservoirs and putting agriculture activities in danger.

Further, the world population is growing steadily and is increasingly urbanized. Technology is evolving incessantly, and the economy is more and more globalized. At the same time, there are worrying global trends in malnutrition, including a rapid rise in overweight and obesity, even as forms of undernutrition persist. The way

*Corresponding author: 17204252@siswa.um.edu.my

food is produced, distributed and consumed worldwide has also changed dramatically. This vastly different world calls for new ways of thinking about hunger and food insecurity, which is the second of the United Nations' Sustainable Development Goal (SDG2) – a world without hunger, food insecurity and malnutrition.

After air and water, food is the most important thing humans need to survive and getting wholesome food is becoming more and more difficult. A staggering 75% of our calories come from just 12 crops and 5 animals, leading to reduced diversity in our nutrition, crops and animals that are vulnerable to climate volatility, pests and diseases, and farmers with limited choices of crops to grow. There are actually 300,000 plants potentially fit for consumption, so there is great opportunity for us to eat a greater variety of crops (FAO). A person is food secure when they have regular access to enough safe and nutritious food for normal growth and development and an active and healthy life. This is supported due to the availability of food and/or sufficient resources to obtain food. In addition, the stress of living with uncertain access to food and going on periods without food can lead to physiological changes that can contribute to overweight and obesity. Children facing hunger, food insecurity and undernutrition today may have a higher risk of overweight, obesity and chronic diseases like diabetes later in life. In many countries, undernutrition and obesity coexist and both can be consequences of food insecurity. In Malaysia, the rise in food prices accelerated April 2022's year-on-year inflation to 2.3%. This caused food ceiling prices to rise, which will lead to high levels of poverty (DOSM, 2022).

In addition, the stress of living with uncertain access to food and going on periods without food can lead to physiological changes that can contribute to overweight and obesity. Children facing hunger, food insecurity and undernutrition today may have a higher risk of getting chronic diseases like diabetes later in life. In many countries, undernutrition and obesity coexist and both can be consequences of food insecurity. Thus, the objective of this study is to forecast the Food Security Inflation in Malaysia using ARIMA Model and Machine Learning Approach to develop a clear strategy on ensuring the sustainability of food resources. Figure 1 shows that, between 702 and 828 million people in the world faced hunger in 2021.

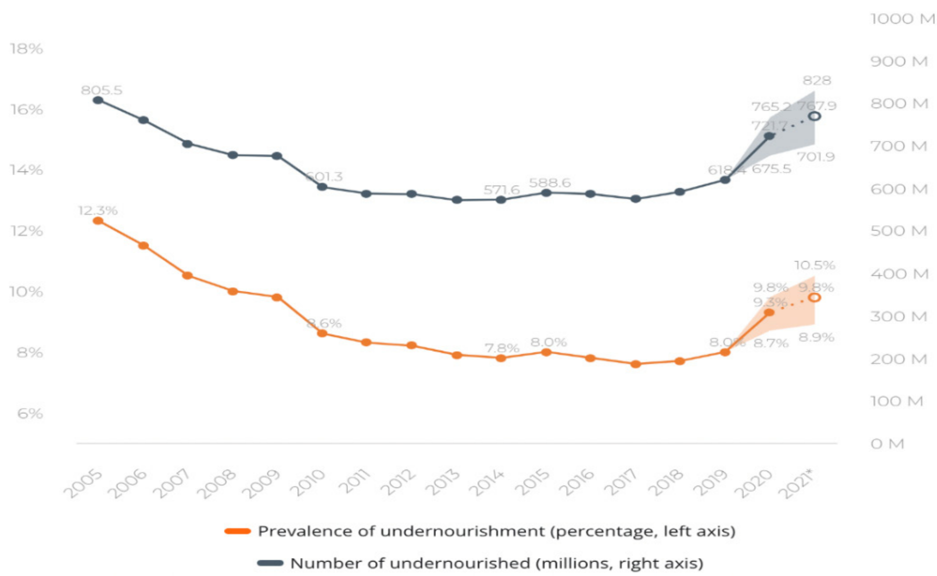


Figure 1: Prevalence of Undernourishment (PoU)

*Notes: *Projected values for 2021 in the figure are illustrated by dotted lines. Shaded areas show lower and upper bounds of the estimated range. Source: FAO, IFAD, UNICEF, WFP and WHO. 2022. The State of Food Security and Nutrition in the World 2022.*

The Prevalence of Undernourishment (PoU) is FAO's traditional indicator to monitor hunger at the global and regional level and is based on country data on food availability, food consumption and energy needs. It estimates the adequacy of a population's dietary energy intake. Historically, the number of hungry people in the world (between 702 and 828 million) has been derived using this approach. PoU estimates cannot be sufficiently disaggregated to be able to identify specific vulnerable populations within countries, which is a limitation of monitoring the ambitious goal of Zero Hunger, which part of an agenda that aims to leave no one behind and guaranteeing food security for the population.

Global, regional, and country-level estimates of the PoU and of the prevalence of moderate or severe food insecurity are disseminated annually in the State of Food Security and Nutrition in the World flagship report as well as on FAOSTAT and the UN Global SDG Indicators Data Platform. Given the above, and the data available from the DOSM for food prices, the objective of this paper is to forecast the Food Security Inflation in Malaysia using ARIMA Model and Machine Learning Approach.

Literature Review

The article, “Malaysia to set up chicken ‘stockpiles’ to overcome shortage, published inThe Straits Times, shows how food insecure Malaysia is due to the pandemic, which saw supply of chicken running low, leading to price increases. The supply of chicken, which is a major source of protein for a majority of Malaysians, is now at a critical state as many people are unable to obtain this resource. On the other hand, the wealthy are more likely to be able to obtain this resource compared with the poor, as mentioned in the article, "Food Security and Unsustainable Agriculture in Malaysia" published in The Sun Daily, showing that the Malaysian government needs to formulate a proper strategy for food distribution. The purpose of this study is to forecast food prices so that the government can prepare food resources to prevent shortages in case of unexpected events.

Methodology

In this research, we used the Machine Learning approach through the Autoregressive Integrated Moving Average (ARIMA) model to predict the prices of food by category based on past data. The methodology framework used in this paper is highlighted in Figure 1. This is to ensure that people will know the trends of food prices and will save up or be aware of incoming price increment to certain food categories. This will also allow people to stock up food resources or use them wisely to avoid the food shortage.

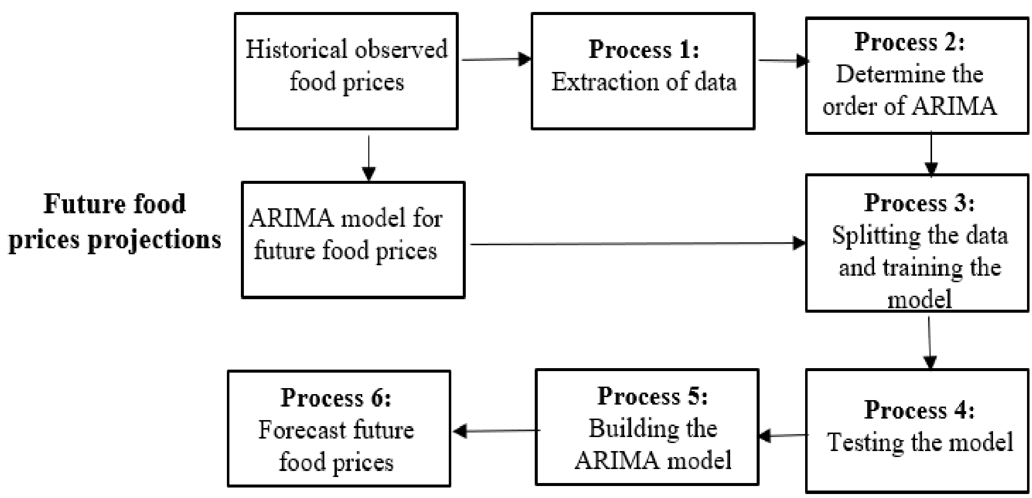
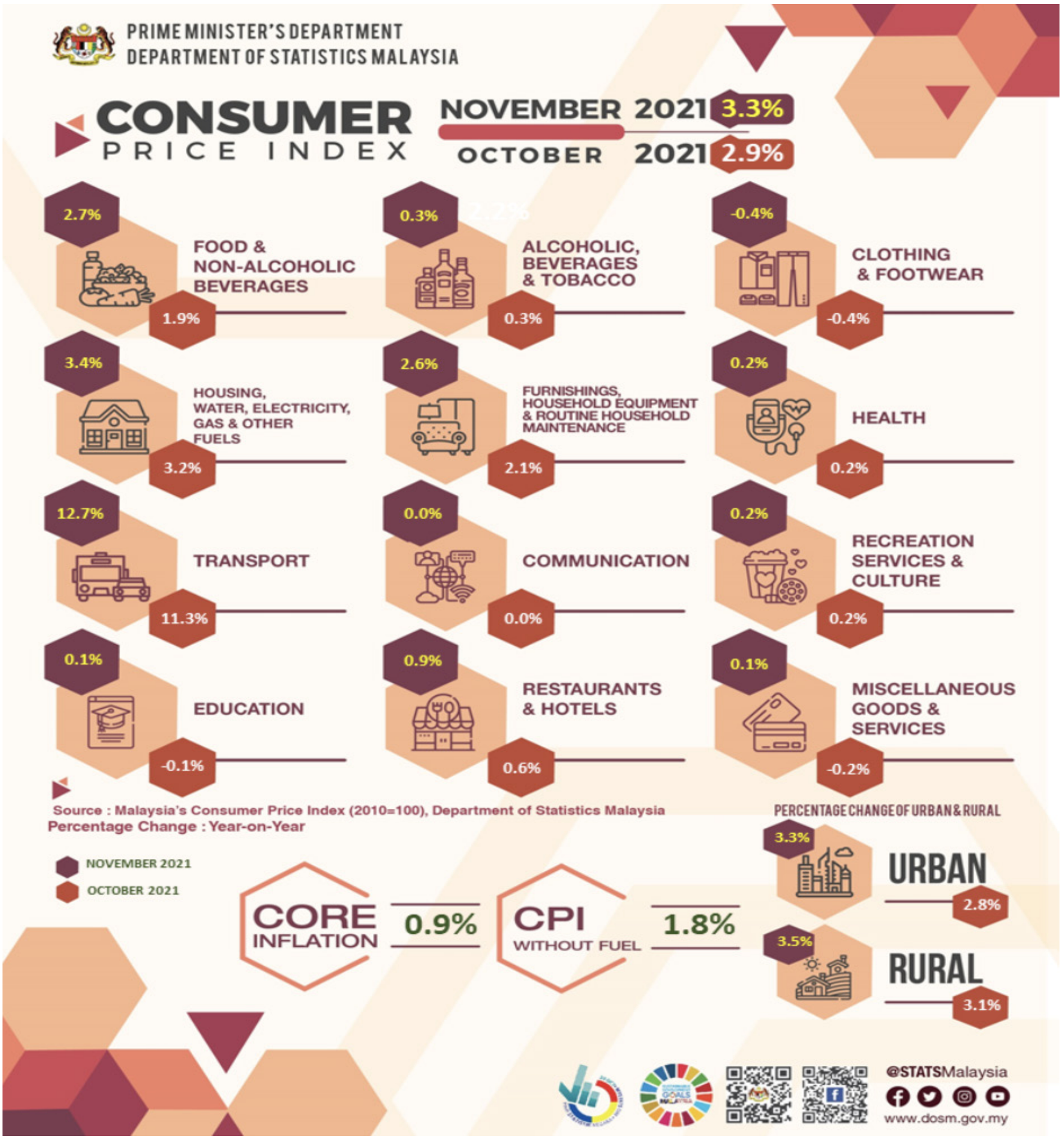


Figure 1: Steps for forecasting of food prices using machine learning

We used the ARIMA Model based on data input for Machine Learning process to forecast the future value of the food prices. In this study, the ARIMA Model was divided into six (6) processes to forecast food price trends (see Table 1a).

We used and analysed the data of the price index (see Table 1) from the Department of Statistics Malaysia (DOSM) based on the food and non-alcoholic beverages category from the year 2014-2021. Only data on food and non-alcoholic beverages was used because we wanted to analyze food prices without considering the other predictors, as the ARIMA Model supports Uni Variate Time Series Forecasting. This is due to the fact that we are modeling the pattern initially before putting other factors into the forecast. The timeframe is set between 2014 and 2021. From the data analysis, a pattern of food prices (see Figure 1) is found, and this information can be used as an input for the Machine Learning process such as ARIMA Model to forecast the future trend of food prices.



(Source: DOSM, 2021)

Data on the price index for food and non-alcoholic beverages from the year 2014 to 2021 was uploaded into Jupyter Notebook to begin our program as it will be tailored around this dataset (see Figure 2). f Jupyter Notebook is used to allow specification or a customized model of the project as it supports machine learning modules, such as the ARIMA package, which will be used to predict the food price trends in the upcoming years.

In [153]:
```
import pandas as pd
import numpy as np

df = pd.read_excel("Food & Non-Alcoholic Beverages.xlsx",
                   index_col="month", parse_dates=True)
df = df.dropna()
print("Shape of data: ", df.shape)
df.head()
```

```
Shape of data:  (94, 1)
```

Out[153]:

	per_cent_change
month	
2014-03-01	3.9
2014-04-01	3.6
2014-05-01	3.8
2014-06-01	3.7
2014-07-01	3.6

Figure 2: Uploading process of the dataset into the Jupyter Notebook platform to begin coding the program for prediction

Below (see Figure 3) is the r data plot showing the percentage of change of food prices f by utilising the Matplotlib library. By visualising our data, we can clearly see the trends of food prices over the years.

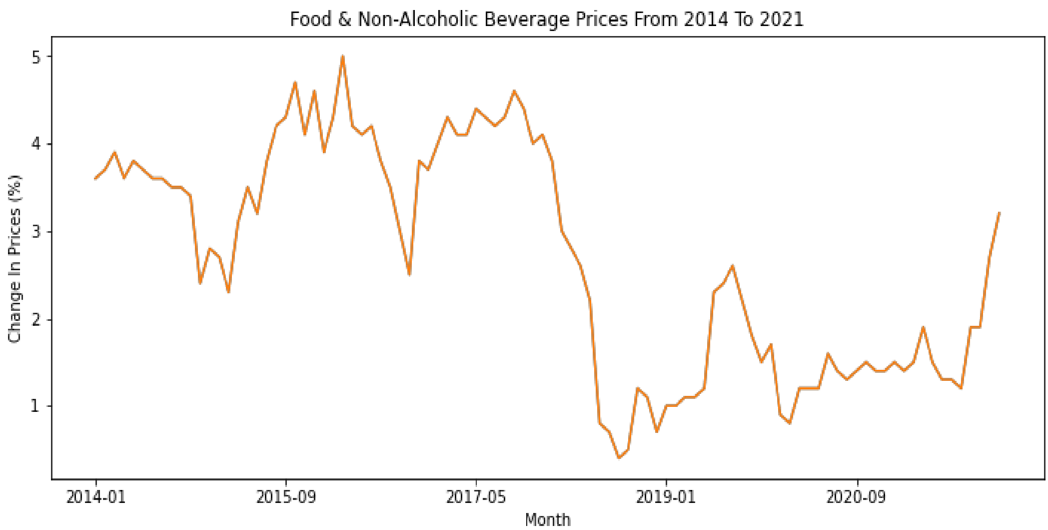


Figure 3: Output for the percentage of food and non-alcoholic beverages from the year 2014 to 2021 in graph form

Results and Discussion

Extraction of the data in process 1, shows that the null hypothesis for the ADF test is that the time series is non-stationary (see Figure 4). P-value is greater than the significance level (0.05), hence differencing is needed. Therefore, null hypothesis is accepted.

```
In [32]: ad_test(df["per_cent_change"])

         1. ADF :  -1.0312577586755933
         2. P-Value :  0.7416937522956668
         3. No. of Lags :  12
         4. No. of Observations used for ADF Regression and Critical Values Calculation :  83
         4. Critical Values :
                  1% :  -3.5117123057187376
                  5% :  -2.8970475206326833
                  10% :  -2.5857126912469153
```

Figure 4: The ADF test for the dataset

Process 2, is to determine the order of ARIMA (p,d,q) and finding the value of the d-parameter (see Figure 5). Since our p-value is larger than 0.05, differencing is needed. The first order of differencing shows much stationary data as the plot is fairly closer to the significance value. However, the 2nd order of differencing is not needed as more plots migrate to the negative value which results in over differencing of data making our d-parameter (I) 1.

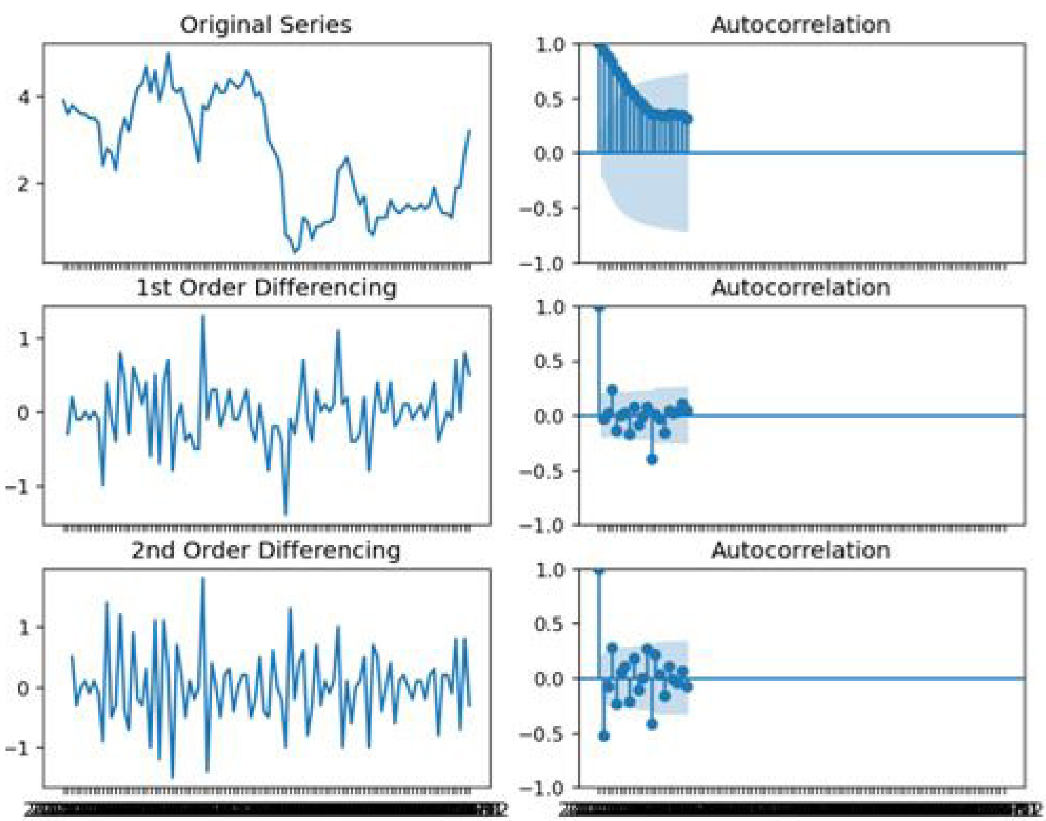


Figure 5: The order of differencing for dataset

In finding the value of the p-parameter, the value of p-parameter is obtained via the Partial Autocorrelation test (PACF). We found that two of the lags are out of the significance limit so our optimal value of our p (AR) is two (see Figure 6).

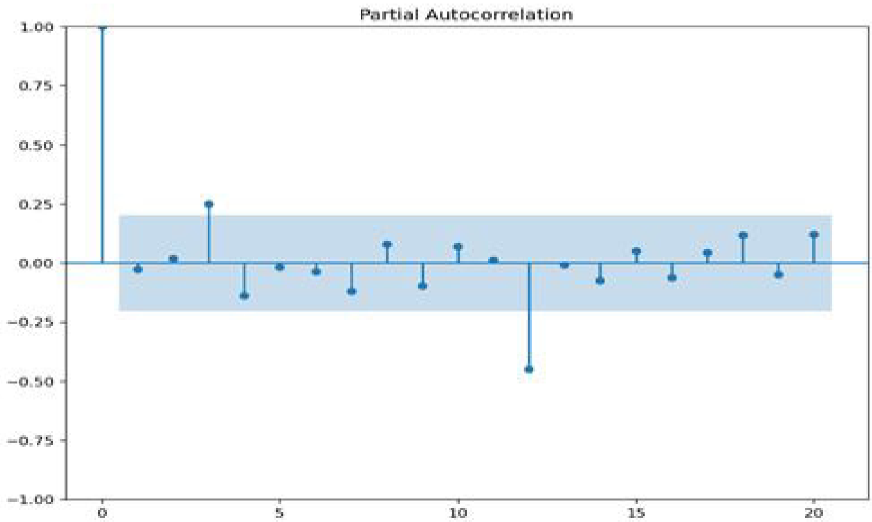


Figure 6: Partial Autocorrelation for the dataset

The value of q-parameter is obtained via the Autocorrelation test (ACF). Two of the lags are out of the significance limit so our optimal value of the q-parameter (MA) is two (see Figure 7).

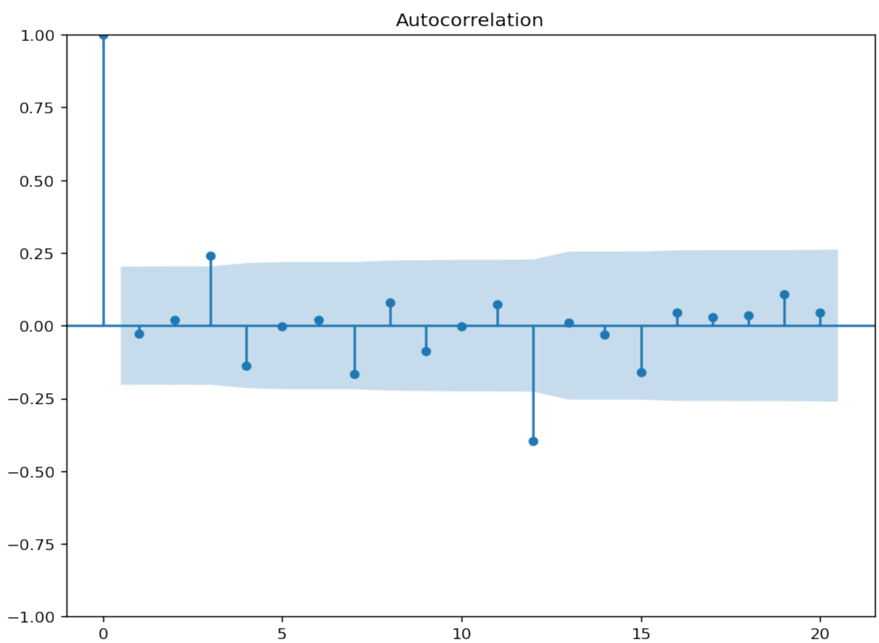


Figure 7: Autocorrelation for the dataset

Similarly, we can verify our findings with the auto ARIMA function (see Figure 8). Thus, to determine the order of ARIMA (p,d,q), we obtained value 2 for the p-parameter; value 1 for the d-parameter; and, value 2 for the q-parameter. Therefore, the order of our ARIMA is (2, 1, 2). These values are to be keyed in when building our ARIMA model.

```
Performing stepwise search to minimize aic
 ARIMA(2,1,2)(0,0,0)[0] intercept   : AIC=107.109, Time=0.12 sec
 ARIMA(0,1,0)(0,0,0)[0] intercept   : AIC=110.014, Time=0.01 sec
 ARIMA(1,1,0)(0,0,0)[0] intercept   : AIC=111.923, Time=0.01 sec
 ARIMA(0,1,1)(0,0,0)[0] intercept   : AIC=111.927, Time=0.02 sec
 ARIMA(0,1,0)(0,0,0)[0]             : AIC=108.023, Time=0.01 sec
 ARIMA(1,1,2)(0,0,0)[0] intercept   : AIC=113.677, Time=0.07 sec
 ARIMA(2,1,1)(0,0,0)[0] intercept   : AIC=115.165, Time=0.05 sec
 ARIMA(3,1,2)(0,0,0)[0] intercept   : AIC=108.870, Time=0.13 sec
 ARIMA(2,1,3)(0,0,0)[0] intercept   : AIC=113.166, Time=0.07 sec
 ARIMA(1,1,1)(0,0,0)[0] intercept   : AIC=113.921, Time=0.04 sec
 ARIMA(1,1,3)(0,0,0)[0] intercept   : AIC=111.745, Time=0.05 sec
 ARIMA(3,1,1)(0,0,0)[0] intercept   : AIC=110.806, Time=0.04 sec
 ARIMA(3,1,3)(0,0,0)[0] intercept   : AIC=110.173, Time=0.16 sec
 ARIMA(2,1,2)(0,0,0)[0]             : AIC=105.128, Time=0.07 sec
 ARIMA(1,1,2)(0,0,0)[0]             : AIC=111.693, Time=0.06 sec
 ARIMA(2,1,1)(0,0,0)[0]             : AIC=113.165, Time=0.03 sec
 ARIMA(3,1,2)(0,0,0)[0]             : AIC=106.885, Time=0.12 sec
 ARIMA(2,1,3)(0,0,0)[0]             : AIC=111.173, Time=0.05 sec
 ARIMA(1,1,1)(0,0,0)[0]             : AIC=111.932, Time=0.02 sec
 ARIMA(1,1,3)(0,0,0)[0]             : AIC=109.746, Time=0.04 sec
 ARIMA(3,1,1)(0,0,0)[0]             : AIC=108.807, Time=0.03 sec
 ARIMA(3,1,3)(0,0,0)[0]             : AIC=108.177, Time=0.13 sec

Best model:  ARIMA(2,1,2)(0,0,0)[0]
```

Figure 8: The Auto ARIMA summary for the dataset

In process 3, we split the data and training the model. In this process, we allocated 90% of our data for training and the remaining 10% for testing our model (see Figure 9).

```
from statsmodels.tsa.arima_model import ARIMA

print(df.shape)
train = df.iloc[:-10]
test = df.iloc[-10:]
print(train.shape, test.shape)

(96, 1)
(86, 1) (10, 1)
```

Figure 9: The process of splitting the data for training and testing of the model

Figure 10 shows the summarization of the model training with 90% data. We found that AIC is minimized from 107.109 to 92.848 which shows that our model and the data it trained on is balanced.

Out[303]:

SARIMAX Results

Dep. Variable:	per_cent_change	**No. Observations:**	86
Model:	ARIMA(2, 1, 2)	**Log Likelihood**	-41.424
Date:	Fri, 14 Oct 2022	**AIC**	92.848
Time:	02:33:21	**BIC**	105.061
Sample:	01-01-2014	**HQIC**	97.761
	- 02-01-2021		
Covariance Type:	opg		

	coef	std err	z	P>|z|	[0.025	0.975]
ar.L1	-1.4482	0.066	-21.778	0.000	-1.579	-1.318
ar.L2	-0.9191	0.065	-14.099	0.000	-1.047	-0.791
ma.L1	1.5849	0.176	9.030	0.000	1.241	1.929
ma.L2	0.9717	0.185	5.255	0.000	0.609	1.334
sigma2	0.1506	0.035	4.276	0.000	0.082	0.220

Ljung-Box (L1) (Q):	0.33	**Jarque-Bera (JB):**	3.69
Prob(Q):	0.57	**Prob(JB):**	0.16
Heteroskedasticity (H):	0.68	**Skew:**	-0.27
Prob(H) (two-sided):	0.32	**Kurtosis:**	3.87

Figure 10: The summarization of the trained model

In process 4, we test the model. After the training process, we now can test our model on the actual data. We can see that the ARIMA Prediction values are fairly close to the values of the actual data (see Figure 11).

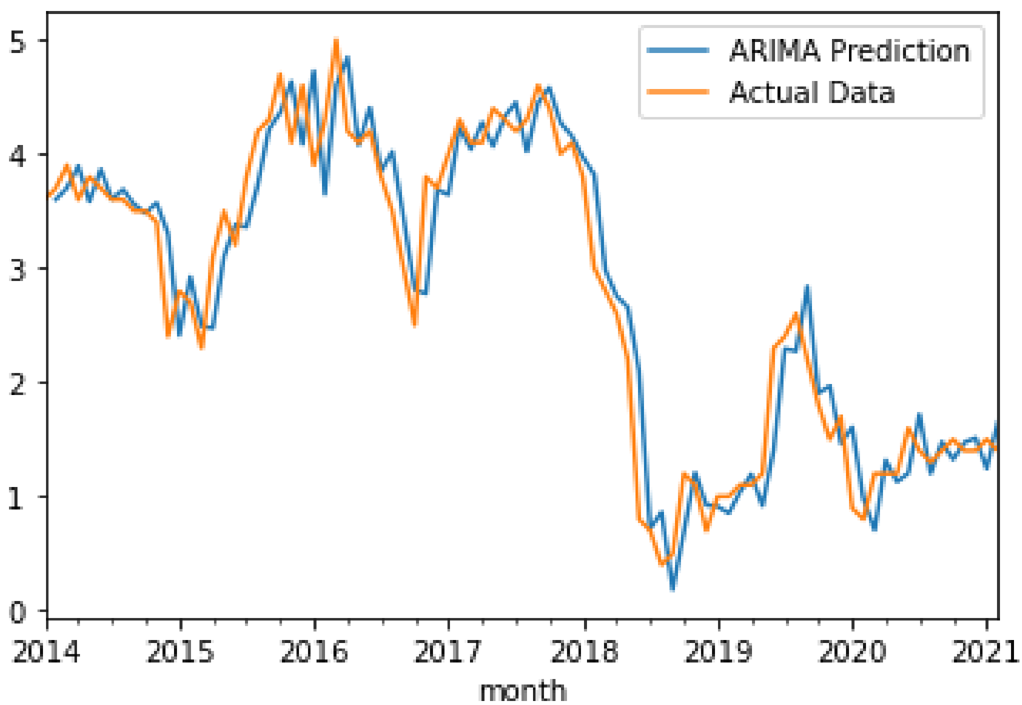


Figure 11: The Autocorrelation for the dataset

Process 5 is building the ARIMA model. Since the model is well trained from the previous values, we are able to make future predictions, that is, tested with our remaining 10% data. The result obtained is artificially good as it matches the trend of price fluctuation on the actual data (see Figure 12).

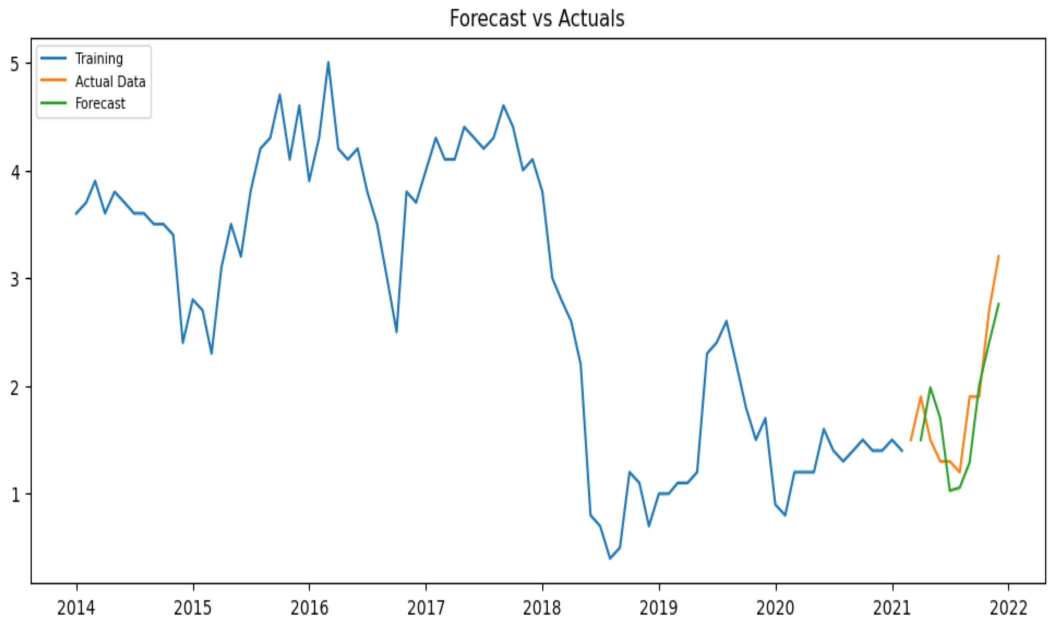


Figure 12: The summarization of the trained model

Process 6 is the last one where we forecast the food trends (see Figure 13). With proper training, we can now utilize our model to forecast trends in food prices from January 2022 to July 2024 onward.

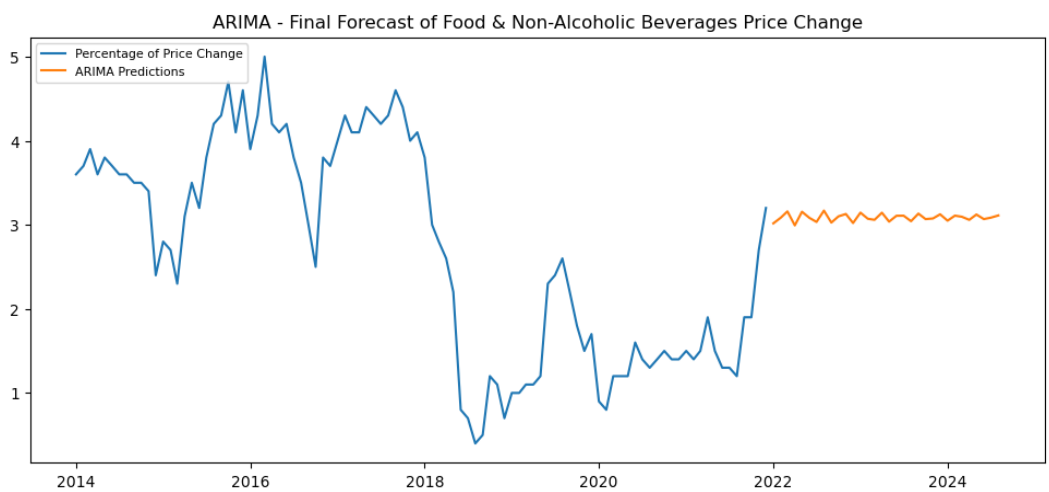


Figure 13: The forecasting of food prices

Figure 14, shows the scaled-up value of the forecasting result, and with the implementation of Machine Learning with the ARIMA Model we are able to visualize future data trained from past data. The motivation or the target of this study is to forecast the prices of food, but for the initial part of the study, we are looking for a pattern of food prices in the future comparing to the pattern of food prices in the past years, hence no specific values are needed here.

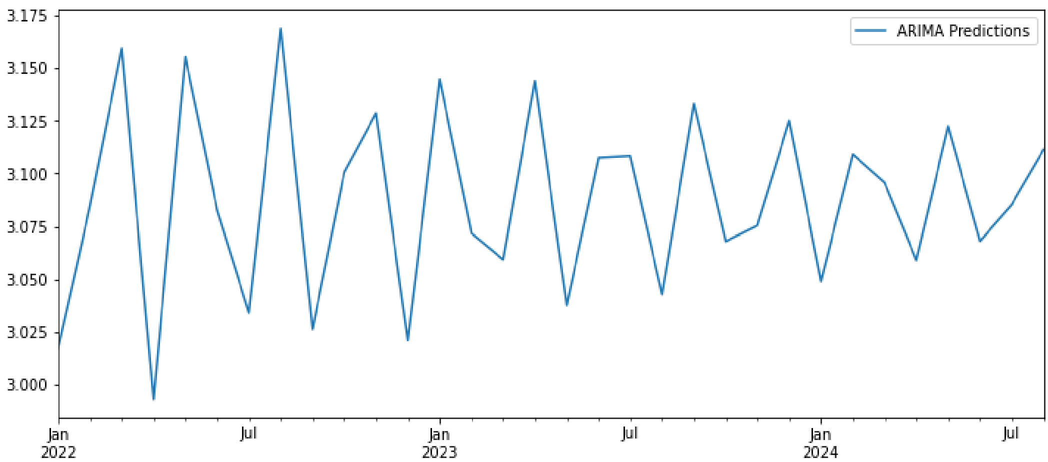


Figure 14: The scaled-up of the forecasting result

Conclusion

In this paper, we investigated the food security inflation in Malaysia through combining ARIMA model and Machine Learning approach. The future food prices series projected in this study are visualized possible future food prices scenario extracted by the ARIMA model. From the findings, in January 2023 the food price trend is shown to soar compared to in January 2024, where it is shown to drop. This is due to the external factors that contribute to the fluctuation of food price pattern. This output is based on Machine Learning predicting food price trends which can guide new strategy on overcoming food shortage and maintain food security.

An ARIMA model that can reflect the statistical characteristics of past and future food prices has been proposed, in which the initial values are assumed to obey the probability distribution derived from historical food prices. The proposed model can be applied to generate random food prices series in the future under different scenarios, which were further used to derive the food prices. We should consider and apply the strategies that have been introduced in this study to provide better understanding and solutions for the world's most pressing problems.

Disclosure Statement

No potential conflict of interest was reported by the author(s). The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Armstrong, J. (1985). *Long range forecasting: From crystal ball to computer*. John Wiley & Sons, New York.

Adler, M., & Ziglio, E. (1996). *Gazing into Oracle: Delphi method and its Application to Social Policy and Public Health*. Jessica Kingsley Publisher.

Banerjee, C., & Sharma, A., (2021). Decline in terrestrial water recharge with increasing global temperatures. *Science of the Total Environment*, *764*, 142913. https://doi.org/10.1016/j.scitotenv.2020.142913

Chen, Z., & Grasby, S.E., (2014). Reconstructing river discharge trends from climate variables and prediction of future trends. *Journal of Hydrology*, *511*, 267–278. https://doi.org/10.1016/j.jhydrol.2014.01.049

Department of Statistics Malaysia. (2021). *Analysis of Annual Consumer Price Index 2021*. https://www.dosm.gov.my/v1/index.php?r=column/cthemeByCat&cat=490&bul_id=VnZCRlBNbVVOWURoNWdObWhLYlJZdz09&menu_id=bThzTHQxN1ZqMVF6a2I4RkZoNDFkQT09

Department of Statistics Malaysia Official Portal. (2002). Sustainning Tourism. Consumer Price Index Malaysia. (2019). https://www.dosm.gov.my/v1/index.php?r=column/archiveByCat&menu_id=V1NxcDJhaEpYUUcxVG55a0h1eFZ5Zz09&arktype=1&cat_id=106&page=1

Department of Statistic, Malaysia (DOSM). (2022). *Malaysia's consumer prices rose 2.3% in April as food inflation accelerated*. shorturl.at/bhyS6

Duan, J. G., Bai, Y., Dominguez, F., Rivera, E., & Meixner, T., (2017). Framework for incorporating climate change on flood magnitude and frequency analysis in the upper Santa Cruz River. *Journal of Hydrology*, *549*, 194–207. https://doi.org/10.1016/j.jhydrol.2017.03.042

Food security and unsustainable agriculture in Malaysia. (2022, April 7). *The Sundaily* Retrieved June 10, 2022, from https://www.thesundaily.my/home/food-security-and-unsustainable-agriculture-in-malaysia-LK9048692

Malaysia to set up chicken "stockpiles" to overcome shortages. (2022, June 9). *The Straits Times*. https://www.straitstimes.com/asia/se-asia/malaysia-to-set-up-chicken-stockpiles-to-overcome-shortages

Market Watch. (May 25, 2022). *Malaysia's Consumer Prices Rose 2.3% in April as Food Inflation Accelerated.* https://www.marketwatch.com/story/malaysia-s-consumer-prices-rose-2-3-in-april-as-food-inflation-accelerated-271653451205

Qistina M. R. (2022). Time – series modelling of Food Security Inflation in Malaysia using an ARIMA Model as a Machine Learning Approach. *The International Conference of Advanced Intelligent Maritime Safety and Technology 2022 (Ai-MAST 2022)*. 13-14 October 2022, Shandong Jiaotong University, Weihai, China.

Rachel R. (February 25, 2022). Increases in food prices cause concern for the poor. *The Borgen Project*. https://borgenproject.org/increases-in-food-prices/

Vicente-Serrano, S. M., Azorin-Molina, C., Sanchez-Lorenzo, A., Revuelto, J., L ´opez- Moreno, J. I., Gonz´alez-Hidalgo, J. C., Moran-Tejeda, E., & Espejo, F., (2014). Reference evapotranspiration variability and trends in Spain, 1961–2011. *Global and Planetary Change*, *121*, 26-40

Wasko, C., & Sharma, A., (2017). Global assessment of flood and storm extremes with increased temperatures. *Scientific Reports*, *7*(1), 1–8. https://doi.org/10.1038/s41598-017-08481-1.

Yun, X., Tang, Q., Li, J., Lu, H., Zhang, L., & Chen, D., (2021). Can reservoir regulation mitigate future climate change induced hydrological extremes in the Lancang-Mekong River Basin? *Science of the Total Environment*, *785*, 147322. https://doi.org/10.1016/j.scitotenv.2021.147322