

Hybrid of ARIMA and Quantile Regression (ARIMA-QR) Model for Forecasting Paddy Price in Indonesia

¹Wiwik Anggraeni, ¹Faizal Mahananto, ¹Fajar Ratna Handayani, ²A. Kuntoro Boga and ³Sumaryantoe

¹Department of Information Systems, Faculty of Information and Communication Technology,
Sepuluh Nopember Institute of Technology, Surabaya, Indonesia

²Agriculture Technology Assessment Center, Riau, Indonesia

³Center of Economic Study and Agriculture Policy, Bogor, Indonesia

Abstract: The price of paddy as the main food commodity in Indonesia, from year to year continues to experience fluctuations but tends to increase over the past few years. This requires that decision makers take action to maintain price stabilization. Indonesian Bureau of Logistics (BULOG) as a decision maker needs to know the forecasting of paddy prices over the next periods in order to determine the best actions or policies. The policy can be in the form of the amount of government paddy reserves, the amount of release of stock to the market, the determination of the amount of imported paddy and the price of paddy. In this research, the price of paddy is forecasted by using the ARIMA-QR method to obtain forecasting results for the future period as well as identifying factors that influence paddy price fluctuations. In doing this forecasting, several variables are used which influence the fluctuations in paddy prices such as the price of grain basis (GKG) and world paddy prices, the amount of BULOG stock, Ied holiday and the forecasting value of paddy prices that have been done previously. The data used is monthly data starting from 2000-2015. Based on the results of the study, the price forecasting model using ARIMA and ARIMA-QR has an accuracy of 1.47% for $q = 0.25$, 1.21% for $q = 0.5$ and 1.42% at the time $q = 0.75$. This average accuracy is 0.03% lower than the ARIMA accuracy.

Key words: Forecasting, ARIMA, quantile regression, ARIMA-QR, paddy prices, fluctuations

INTRODUCTION

Paddy is a basic necessity which is one of the main food commodities for the community and the government. Therefore, it is also important for the government to make the issue of rice a priority. From several aspects of the food economy, prices are one of the important things that need attention and need special policies that regulate food prices. The increase or decrease in prices will have an impact on people's lives, both for the lives of the community and farmers as producers (Anonymous, 2018) (Wulandari, 2017; Post, 2017). The importance of policies related to the issue of food prices at the level of farmers, distributors and consumers is aimed at helping to achieve food self-sufficiency and reduce import dependence, reduce price volatility and pay attention to consumer purchasing power (Syafaat, 2017; Baderi, 2018).

Fluctuations in paddy prices on the market can be caused by an unstable market demand. For example, there is an increase in demand on national holidays such as before and during the month of Ramadan, Eid al-Fitr,

Christmas holidays, etc., (Arifin, 2013). The increasing of market demand that is not offset by increased production will cause scarcity and increase in prices for certain commodities. Things like this are not new deals anymore because people's demand for basic needs always increases every month before (Rahman, 2017; Adrian, 2013). Thus, the policy in controlling the price of food commodities should be better done by Jackson (2018) the government. One of the institutions in Indonesia that has the task of managing, controlling and maintaining the stability of paddy prices is BULOG (Indonesian Bureau of Logistics).

BULOG as a company that continues to carry out public duties from the government to carry out activities to maintain the basic purchase price for grain, stabilize prices, especially, basic prices, distribute rice to the poor (Raskin) and manage food stocks (Anonymous, 2016).

It can be seen that the price aspect and its relation to the welfare of farmers and consumers is one of the important elements that need to be analyzed in order to

formulate better policies. For this reason, there is a need for forecasting paddy prices in the future to help with the decision-making process regarding pricing policies and related policies.

Previous research uses the autoregressive integrated moving average model for seasonally adjusted headline Consumer Price Index (Jadhav *et al.*, 2017). In this study, ARIMA is suitable for seasonal data. ARIMA is also used for modeling monthly headline consumer price index and the (6,0) (0,0) model is quite adequate (Jadhav *et al.*, 2017). ARIMA is also used as one of the models in forecasting the price of crude palm oil. In agriculture price forecasting have been done by Jadhav *et al.* (2017). This research forecast the Paddy, Ragi and Maize in Karnataka State and demonstrated the of the ARIMA Model as a tool for price forecasting prices in univariate variables. common paddy forecasting in India (Darekar and Reddy, 2017a) proof that ARIMA Model is the most representative model for the price forecast of paddy in overall India. ARIMA Model is also used to forecast market price of soybean in India (Zahari *et al.*, 2018). In agriculture, price forecasting has been done and indicated that the ARIMA Model is the most adequate and efficient model (Panasa *et al.*, 2017). ARIMA is also used the Cock Model to forecast the cotton price (Darekar and Reddy, 2017b). In addition, the ARIMA Model is also suitable for forecasting the short term and the ARIMA (1,1,2) models that were found to be the best fit to forecast short term gold prices. From these studies it is shown that ARIMA is a suitable model for forecasting prices. But the drawback is that the ARIMA Model can only do forecasting with 1 variable (multivariate).

In addition, many studies related to quantile regression have been carried out. In this study QR is used more to analyze the influence of a variable on other variables. Quantile regression is used to test the effects of urban rail transit networks on non-landed private housing values in Singapore (Singh and Kumar, 2017), relationship between Islamic bond (sukuk) prices and financial (Reboredo and Naifar, 2017), to generate probabilistic load forecasts (Liu *et al.*, 2017), Quantile-based dependence relationships in EUA futures market (Tan and Wang, 2017) to examine the impact of income, economic openness and interest rates on housing prices in China (Chernozhukov *et al.*, 2017), explore the determinants of behind financial contagion between US and BRIC (Brazil, Russia, India and China) equity markets (Kocaarslan *et al.*, 2018) causality between gold commodity and gold stock prices (Reboredo and Ugolini, 2017), examines the effects of oil price shocks on stock market returns of the Gulf Cooperation Council (Nusair and Al-Khasawneh, 2018). From the research above, it can be concluded that QR is suitable to be used to analyze how the influence of a variable on other variables and can see in more detail how the influence of

independent variables in each quintile or certain forecasting points. The use of a mixed method of SARIMA and QR becomes SARIMA-QR in a study conducted by stated that the mixed method of SARIMA-QR provides better results than traditional forecasting methods which are indicated by smaller MAPE and RMSE values (Darekar and Reddy, 2017).

Based on several studies above, in this study price forecasting was carried out using the ARIMA and Quantile Regression (ARIMA-QR) method, taking into account several variables that influence paddy price fluctuations, namely the basic price of grain and world paddy prices (Anonymous, 2018; Post, 2017), the amount of BULOG stock (Wulandari, 2017; Adrian, 2017; Anonymous, 2016), national holidays (Rahman, 2017) and paddy prices in the previous period.

The main objective of this research is to find out the price of rice in the next period, so that, it can assist in making the best decisions or policies that need to be applied related to prices, so as to improve the welfare of the community.

This research has a difference with the research that has been done before. These only one method was used in previous studies. In this study, the application of hybrid models was carried out, namely the ARIMA and quantile regression models. This mixed ARIMA and QR Model is used to determine the value of paddy prices forecasting in the future while at the same time knowing the factors that influence them. It can be used as a consideration for Bulog in making policies related to paddy prices. In this study forecasting is done by involving influential variables where the variables involved are different from previous studies. In addition, paddy prices in Indonesia also have different characteristics than others. This can be seen from high production but consumption is also high.

MATERIALS AND METHODS

Data: The data used in this study are data from BULOG which includes data on paddy prices, total stock, milled dry grain prices (GKG), national holidays, world paddy prices. Data used for 15 years in a monthly period. The data is obtained in hardcopy in internal and external statistical books as well as operational books belonging to the State Logistics Agency. This next data will be divided into training data and testing data with a division of 70 and 30%, respectively. Each data plot is shown in Fig. 1.

Figure 1 shows the actual data sharing to 70% training data and 30% testing data. As much as 70% of the first actual data, namely data from January 2000

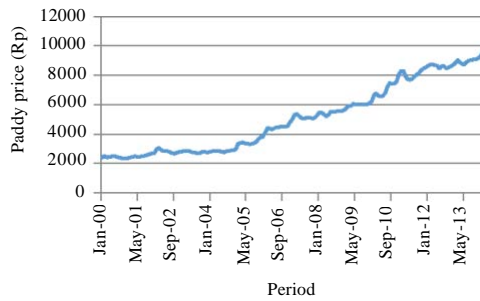


Fig. 1: Distribution of training and testing data

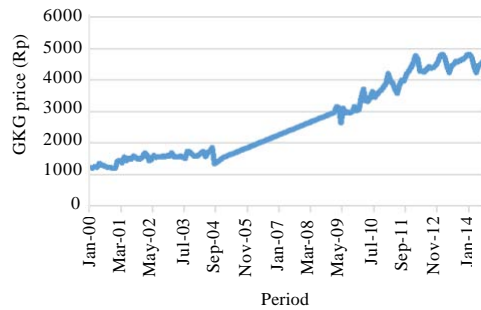


Fig. 2: Graph of milled rice grain price variable (GKG)

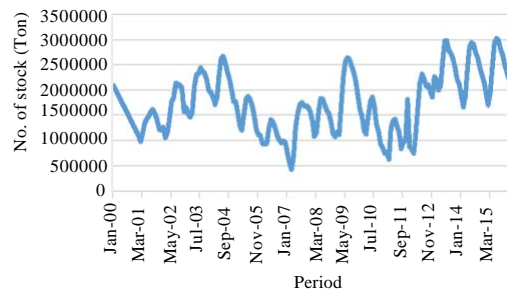


Fig. 3: Graph of variable stock quantities

December 2010 were used as training data. While the remaining 30%, namely data from January 2011 to December 2015 will be used as test data. In the graph, blue shows training data while red indicates testing data. Figure 1 shows that paddy prices tend to increase from year to year.

Besides paddy price data as the main variable in this study also involved several variables including the milled dry grain price variable (GKG) shown in Fig. 2, the number of stocks shown in Fig. 3, world paddy prices in Fig. 4 and national holidays on Fig. 5. For national holiday data only worth 0 and 1 where the value of 1 for the month in which there is a holiday that is Eid al-Fitr while the value of 0 for the ordinary month that does not have Eid al-Fitr in it.

Figure 2 shows the variable price of milled dry grain (GKG) data experiencing fluctuations but tends

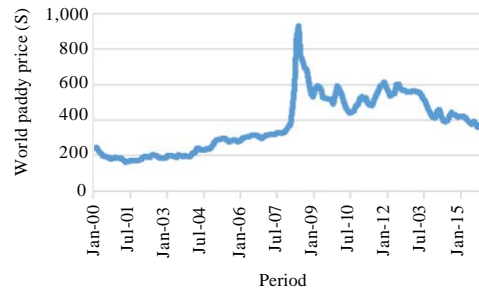


Fig. 4: Graph of world paddy price variables

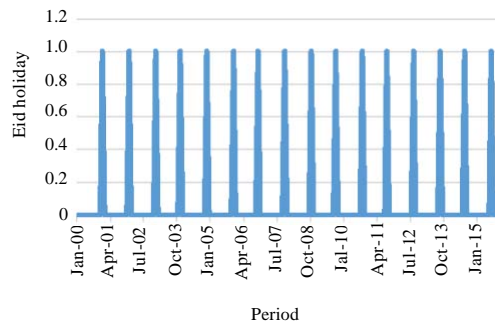


Fig. 5: Graph of national holidays

to increase from year to year. Figure 3 shows the variable data on the amount of rice stock from 2000-2015. From the graph it can be concluded that the data on the number of stocks is quite fluctuating data. Figure 4 shows the variable data on world paddy prices (prices in other countries). The price of world rice used is the price of 5% Bangkok rice. From Fig. 3 it can be seen that the world paddy price variable data during 2000-2015 experienced fluctuations with preliminary data which tended to increase later in the last few years tending to decline.

Figure 5 shows the variable data of national holidays which only have a value of 0 and 1 where the value of 1 is for the month in which there is a holiday, namely Eid al-Fitr while the value of 0 for the normal month is not Eid in it.

Stage of research: The following are the stages in the research conducted for ARIMA-QR forecasting:

- Data preparation
- Data interpolation

Data interpolation is carried out, if there are blank data found. Interpolation allows to get some data that has known the value of data in the period before and after data that is lost or empty. In this study, data interpolation

was carried out on MPD variables for certain months in 2000, 2004, 2005 and 2008. In addition, interpolation was also carried out on the variable number of stocks for June-August 2009.

Dividing the data into 70% training sets and 30% testing sets: Data that has been obtained with a monthly period for the next 15 years is divided into two parts, namely training data and test data with a ratio of 70:30. As much as 70% of the first actual data, namely data from January 2000-December 2010 were used as training data. While the remaining 30%, namely data from January 2011-December 2015 will be used as test data.

Forecasting using ARIMA: ARIMA (Autoregressive Integrated Moving Average) or Box-Jenkins Model is one of the timeseries forecasting techniques that are only based on observed variable data behavior. ARIMA has a flexible nature (following a data pattern) has a fairly high level of forecasting accuracy (Jackson *et al.*, 2018). Following the pattern of data means, if the data is not stationary, the data can be adjusted to be stationary data by doing differencing. The steps that must be taken in analyzing data with Box-Jenkins or ARIMA techniques are as follows (Nusair and Al-Khasawneh, 2018).

Test stationarity in variety and average: This stationarity test is used to determine whether or not transformation needs to be done to the existing data, so that, the data becomes stationary in the variance and determines whether or not processes need to be differenced or not in order for the data to be stationary in the mean. This is indeed in accordance with the ARIMA Model requirements.

Model identification: In the model identification phase, time series data plots are made. Then the stationary variety test and the stationary test are carried out. If the data is not stationary in a variety, it is necessary to do a transformation process. If the data is not stationary against the mean, differencing processes are carried out. The identification of the ARIMA Model can be done using the ACF and PACF functions from the stationary data to determine the initial model (determination of AR and MA orders).

Parameter estimation: Parameters are said to be significant if they have a $p < \alpha$ or $p < 0.05$. If you have found the appropriate parameters, then proceed with the diagnostic test process.

Significance test and diagnosis test: Tests of significance and diagnostic tests can be done by plotting ACF and PACF for the residuals. The model is said to pass the significance test if the probability value of all variables ≤ 0.05 and $|t\text{-statistic}| \text{ all variables} > t\text{-table}$. The t-table for testing the significance of parameters is 1.98. The diagnostic test can be seen from the p value of the q-statistic and squared residual correlogram. The value of $p > 0.05$ on the q-statistic indicates that the residuals are random or random which means the model is acceptable. The value of $p > 0.05$ in the squared residual indicates that the side is homogeneous. Models that have met the criteria can be used for forecasting.

Forecasting: After getting the best model from the results of parameter tests and diagnostic tests, the next process is to forecast using the model. Forecasting is done to estimate or predict paddy prices in the following years.

Forecasting using the Quantile Regression (QR) method: Regression is a statistical technique for determining line or curve equations by minimizing deviations between observational data and their predicted values (Panasa *et al.*, 2017). Broadly speaking, regression analysis is interpreted as an analysis of the dependence of dependent variables on independent variables. Regression analysis is interpreted as an analysis of independent variables in order to make estimates or predictions of the value of dependent variables by knowing the value of independent variables (Post, 2017).

Then, introduce quantile regression as an approach in regression analysis. The Quantile Regression (QR) Model introduced by Koenker and Bassett is to estimate the linear quantile model $Q_\tau(\theta) = x' \beta_\tau$, where x_t is a vector from regression and β is a parameter vector (Rahman, 2017).

Quantum regression can be considered as an extension of the ordinary least square model. In particular, OLS regression only estimates how predictor variables are related to the average value of response variables while quantile regression allows for a predictor model of various location/measurement of response variables (Reboredo and Naifar, 2017). Generally, steps that are done are:

Determining quantile: Quantile used in this study is 0.25, 0.50 and 0.75.

Analyze variables: Analyze variables that are related to the dependent variable and repressor (independent variables) including analyzing relationships between variables:

- Find the forecasting value
- Calculate forecasting errors
- Analysis of results and conclusion
- Analyze and compare the results of forecasting accuracy between ARIMA models and ARIMA-QR mixed models. The measure of accuracy used is (Arifin, 2013)

$$MAD = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i| = \frac{1}{n} \sum_{i=1}^n |e_i| \quad (1)$$

$$MAPE = 100 \left(\frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \right) \quad (2)$$

RESULTS AND DISCUSSION

Stationarity test

Stationary variety test: Variety stationary tests are carried out to test whether the data is stationary in variety or not. Stationary Variety Test Results are shown in Fig. 6. Figure 6 shows that the test results show that the data is not stationary in variety because of the value of rounded value 1. Therefore, transformation needs to be done. The graph after transformation is shown in Fig. 7. Figure 7 shows data on paddy prices after the transformation. The graph shows that the transformed data is stationary.

Average stationary test: The average stationary test results are shown in Table 1. Table 1 shows that the data is not stationary in the average because the probability values are ≥ 0.05 and $|t\text{-statistic}| < |t\text{-critical value}|$ then it needs to be done differencing. Results after differencing are shown in Table 2.

Table 2 shows that the data has been stationary because it meets the requirements of probability values ≤ 0.05 and $|t\text{-statistic}| > |t\text{-critical value}|$.

ARIMA modeling: ARIMA modeling is done through the stages of stationary testing, then identifying the ARIMA Model components through ACF and PACF graphs from stationary data. The component of the Auto Regression (AR) Model is obtained from the PACF graph while the Moving Average (MA) component is obtained through Table 3: Correlogram stationary ACF PACF data the ACF graph. Differencing components are obtained by calculating the number of differencing that has been done on related variables. Table 3 shows stationary ACR and PACF correlogram graphs with one-time difference. Table 3 shows the graphs of ACF and PACF from stationary data. By considering the number of graphs that

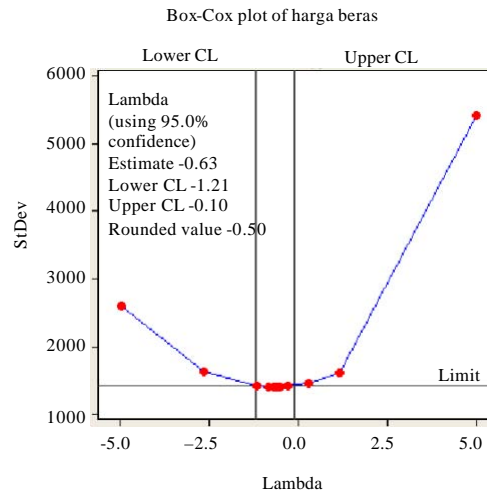


Fig. 6: Variable stationary test results

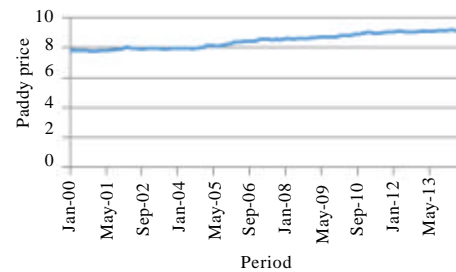


Fig. 7: Graph of transformational results

Table 1: Average stationary test results

Augmented Dickey-Fuller Unit Root Test on TRANSFORMDATA		
Null Hypothesis: TRANSFORMDATA has a unit root		
Exogenous: Constant, Linear Trend		
Lag Length: 2 (Automatic - based on SIC, maxlag=12)		
	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic:	-1.772466	0.7125
Test critical values:		
1% level	-4.030729	
5% level	-3.445030	
10% level	-3.147382	

*MacKinnon (1996) one-sided p-values.

Table 2: ADF test on level difference

Augmented Dickey-Fuller Unit Root Test on D(TRANSFORMDATA)		
Null Hypothesis: D(TRANSFORMDATA) has a unit root		
Exogenous: Constant, Linear Trend		
Lag Length: 1 (Automatic - based on SIC, maxlag=12)		
	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic:	-8.063729	0.0000
Test critical values:		
1% level	-4.030729	
5% level	-3.445030	
10% level	-3.147382	

*MacKinnon (1996) one-sided p-values.

exceed the significance line on the correlogram ACF PACF, the ARIMA Model can be identified. The results of the identification of several components of the ARIMA Model are listed in Table 4 shows the possibility of the

Table 4: Results of identification of the ARIMA Model

Variable	Model
Paddy price	ARIMA (1,1,1), ARIMA (1,1,2)
	ARIMA (1,1,3), ARIMA (2,1,1)
	ARIMA (2,1,2), ARIMA (2,1,3)

Table 5: Significance test results

Model	Prob values	t- stat	Status of significance test
ARIMA (1,1,1)			
AR(1)	0.3201	0.998	Not adequate
MA(1)	0.0169	2.42	
SIGMASQ	0.0000	13.59	
ARIMA (1,1,2)			
AR(1)	0.0000	6.41	adequate
MA(2)	0.0398	2.07	
SIGMASQ	0.0000	13.25	
ARIMA (1,1,3)			
AR(1)	0.0000	6.23	Not adequate
MA(3)	0.6062	0.51	
SIGMASQ	0.0000	13.43	
ARIMA (2,1,1)			
AR(2)	0.4302	0.79	Not adequate
MA(1)	0.0000	6.13	
SIGMASQ	0.0000	13.10	
ARIMA (2,1,2)			
AR(2)	0.0000	42.77	Adequate
MA(2)	0.0000	9.28	
SIGMASQ	0.0000	9.14	
ARIMA (2,1,3)			
AR(2)	0.4879	0.69	Not adequate
MA(3)	0.8752	0.15	
SIGMASQ	0.0000	12.65	

Table 6: Diagnostic test results

Model	The probability average of residual randomness test (Q-stat)	The test result	The probability average of residual homogeneity test (squared residual)	The test result
ARIMA (1,1,2)	0.244	Adequate	0.871	Adequate
ARIMA (2,1,2)	0.0002	Not adequate	0.821	Adequate

Table 7: ARIMA performance measurement (1,1,2)

Performance	Training (%)	Testing (%)
MAPE	1.56	0.93

ARIMA Model being formed by looking at correlograms of ACF and PACF. The next stage is the significance test and diagnostic test to get the best model. Table 5 is the result of the significance test for each model. Table 5 shows that the models that passed the significance test were only ARIMA (1,1,2) and ARIMA (2,1,2) because the probability values of all variables ≤ 0.05 and $|t\text{-statistic}| \text{ all variables} > t\text{-table}$. Models that pass the significance test can proceed to the next stage. The next stage is a diagnostic test. The model is said to pass the diagnostic test, if the probability value is ≥ 0.05 on the Q-statistic and squared residual correlogram. The diagnostic test results are found in Table 6.

Table 6 shows the results of diagnostic tests for the ARIMA (1,1,2) and ARIMA (2,1,2) models. The model

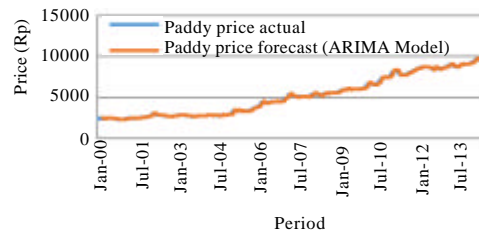


Fig. 8: Graph of actual data and forecasting

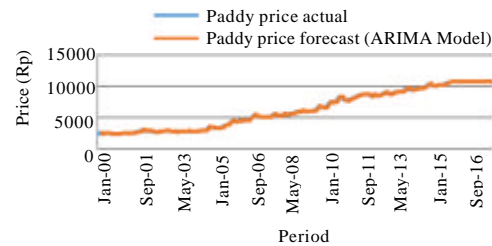


Fig. 9: Graph of forecasting future periods

that passes the test is ARIMA (1,1,2) because this model has a probability value > 0.05 on the Q-statistic and squared residual correlogram. Therefore, this model is used for forecasting. Graphs of forecasting results are shown in Fig. 8. Figure 8 is a graph of the actual data and forecasting data where the actual data graph is depicted by a blue line and forecasting data is represented by a red line. The forecasting value is very similar to the actual data value. So, this model can be said to be very good. To determine the performance of the forecasting model it is necessary to calculate the error value. The results of the calculation of the performance of the ARIMA Model (1,1,2) for training sets and testing sets are listed in Table 7. From the results of the MAPE value both in the training set and in the testing set is $< 10\%$, so, it can be said that the forecasting ability with the model is very good.

Future forecasting results use ARIMA: The best ARIMA Model chosen was used for forecasting for the next 2 years. Future forecasting results are shown in Fig. 9. Figure 9 is a graph of forecasting results for the coming period until 2017. The chart of forecasting future periods tends to resemble a straight line but in fact the forecasting value always increases even though it is not fluctuating. This might be due to the limitations of the tools used.

Modeling of quantile regression: The model chosen in the quantile regression method in this study is referring to the Trimean theory which uses quantile 0.25, 0.5 and 0.75. The relationship between variables is shown in Fig. 10 and 11.

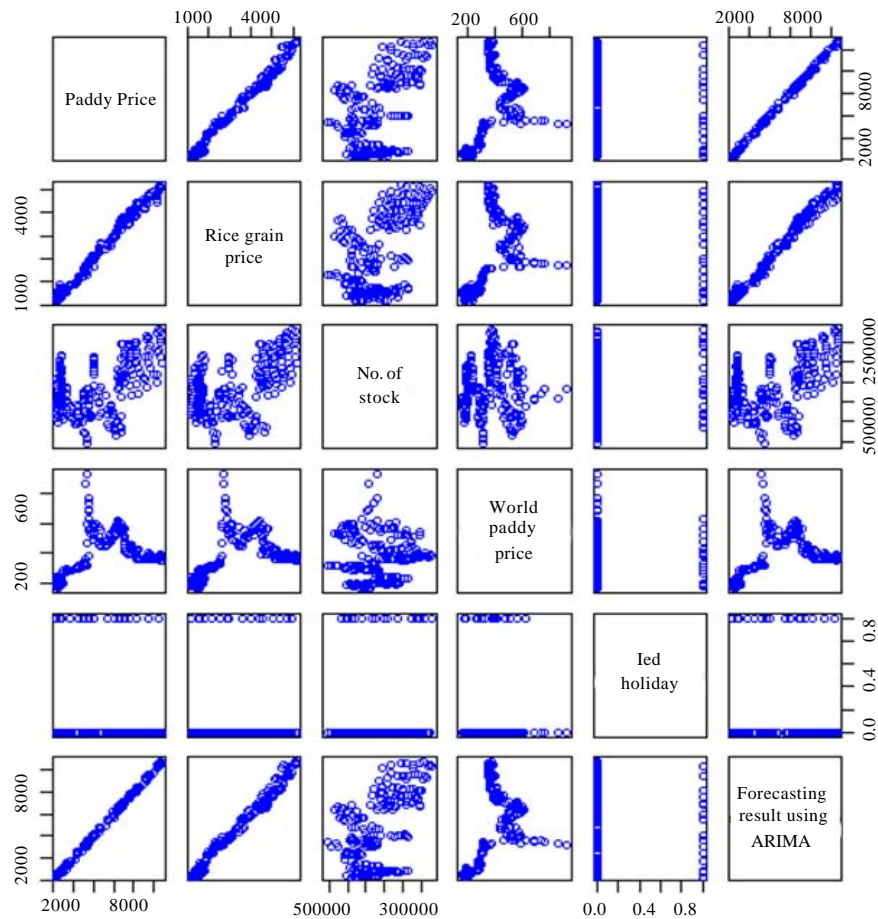


Fig. 10: Variable scatter plot

```
Call:
lm(formula = Y ~ X, data = QRData)

Residuals:
    Min       1Q   Median       3Q      Max
-358.73  -47.38   -8.16   36.71   342.05

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  3.958e+01  2.704e+01   1.464  0.14497
XGKG         2.326e-01  4.285e-02   5.429 1.78e-07 ***
Xjumlah.stock -3.759e-05  1.364e-05  -2.756  0.00645 **
XHarga.Beras.Dunia  4.898e-02  6.172e-02   0.794  0.42845
xhari.besar.nasional -2.138e+01  2.514e+01  -0.850  0.39619
Xhasil.peramalan.ARIMA  8.866e-01  2.142e-02  41.398 < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 96.21 on 184 degrees of freedom
(2 observations deleted due to missingness)
Multiple R-squared:  0.9988,    Adjusted R-squared:  0.9987
F-statistic: 2.966e+04 on 5 and 184 DF,  p-value: < 2.2e-16
```

Fig. 11: Results of OLS regression

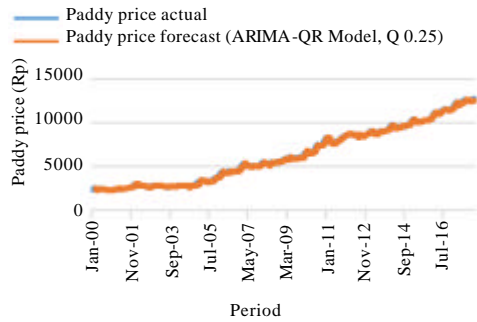


Fig. 12: Graph of forecast values with quantile 0.25

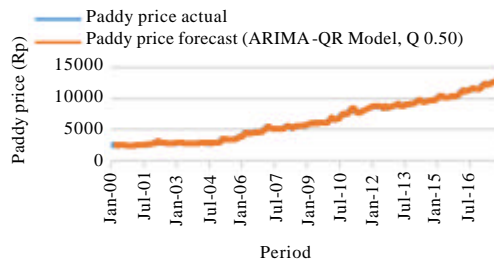


Fig. 13: Graph of forecasting values with quantile 0.50

Figure 10 is a scatter plot image that shows the relationship between variables used. The results show that the relationship between the MPD variable and the paddy price variable is a positive relationship. If the MPD price increases, it will also affect the increase in paddy prices. Likewise, on the ARIMA forecasting variable. Whereas the relationship between the national holiday variables requires further analysis because the correlation is not very visible when using scatter plots. This can happen because the data from this variable is only worth 0 and 1. The OLS regression results are shown in Fig. 11. Figure 11 shows the results of OLS regression which shows the significance of the effect of variable X (independent variable) on the Y variable (dependent variable). Variables that significantly influence paddy prices include the price of milled dry grain, the amount of rice stock and the forecasting results of the ARIMA method at different levels of significance. But among the three, the most influential is the price of milled dry grain (GKG) and ARIMA forecasting variables. MPD variable and ARIMA forecasting results influence the price of rice in the significance level of 0.001 while the variable number of stocks has a significant effect on the significance level of 0.05. Forecasting values using quantile regression with quantile 0.25, 0.5 and 0.75 are shown in Fig. 12-15.

Figure 12 is a graph of the forecast value with quantile 0.25. The graph shows that the graph of forecasting results is very similar to the actual data. Figure 13 is a graph of the forecast value with quantile

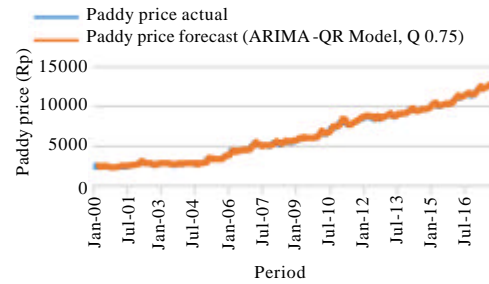


Fig. 14: Graph of forecasting values with quantile 0.75

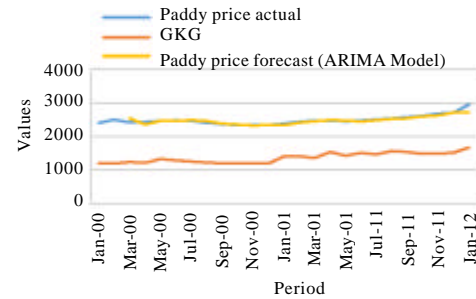


Fig. 15: Fluctuations in paddy prices and significant influence variables on 50% of the initial data

Table 8: Results of calculation of the QR Model error value

Models	MAPE
ARIMA-QR (Q = 0.25)	1.475709
ARIMA-QR (Q = 0.50)	1.212932
ARIMA-QR (Q = 0.75)	1.423903

Table 9: Comparison of method performance

Models	MAPE
ARIMA (1,1,2)	1.346806
ARIMA-QR (Q = 0.25)	1.475709
ARIMA-QR (Q = 0.50)	1.212932
ARIMA-QR (Q = 0.75)	1.423903

0.50. The graph shows that the graph of forecasting results is very similar to the actual data. Figure 12-14 are graphs of forecasting values with quantile 0.25, 0.50 and 0.75. The three graphs above show that the actual forecasting and data values are very similar. If an error value is calculated, it produces a small error value. So that, the modeling can be said to be a good model. The results of the calculation of error values for forecasting actual data are in Table 8.

To find out the forecasting performance used, it is necessary to compare the results of the calculation of the error values listed in Table 8. The error value calculation results are carried out for the entire actual data from 2000-2015 of the several quantile values used, quantile is a quantile that has the smallest MAPE value, this is because in quantile or at 50% the data under paddy prices is much influenced by significant variables. Fluctuations that occur in significant variables along with paddy price

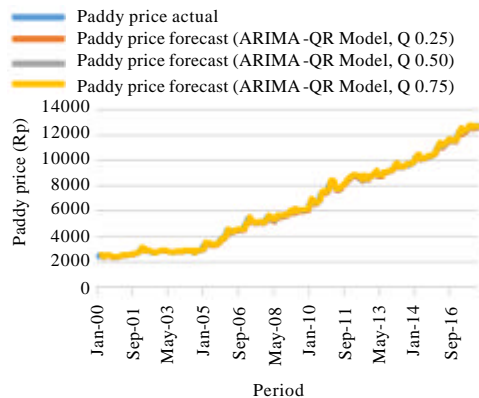


Fig. 16: Graph of comparison of actual data and forecasting values using ARIMA and ARIMA-QR

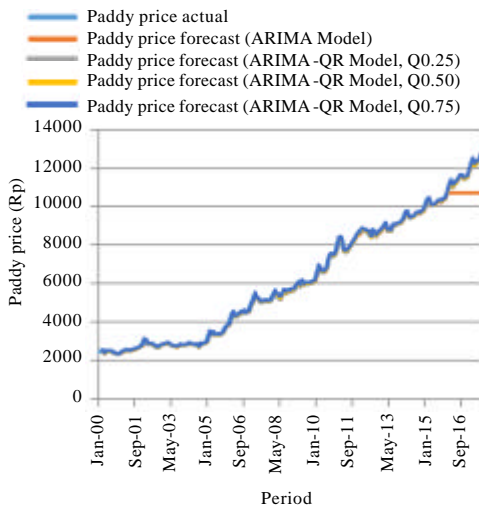


Fig. 17: Graph of actual data values and forecasting values using ARIMA and ARIMA-QR for future periods

fluctuations as shown in Fig. 15. Figure 15 shows that paddy price fluctuations occur along with fluctuations in the variables that are significant at 50% of the initial data. This shows that at 50% of the data below the price of rice is much influenced by significant variables.

Comparison of the performance of the ARIMA and ARIMA-QR methods: Comparison between the ARIMA method alone and the ARIMA-QR mixed method with a MAPE comparison for the overall actual data. Figure 16 shows a graph of the actual data, ARIMA and ARIMA-QR forecasting values for the overall actual data.

Figure 16 shows that the forecasting value uses both the ARIMA method and the ARIMA-QR mixed method,

both of which give results similar to the actual data. So, to further clarify the performance comparison, it is necessary to calculate the error value. A smaller MAPE value indicates that the method can produce a better forecasting value. Table 9 is a performance comparison table between the ARIMA and ARIMA-QR methods.

From the results of the MAPE values above, it can be concluded that for the overall actual data, the ARIMA method can be said to be very good to use because it has a MAPE value below 10%. However, when compared with the ARIMA-QR method, there is a lower MAPE value that is at quantile 0.50 or median.

To get the forecasting value for the future period, the ARIMA-QR method has limitations because basically the regression method is only to identify the influence of certain variables on paddy prices. So, to get the results of forecasting in 2017 by using ARIMA-QR, it is necessary to forecast the value of each variable first until 2017. In this study, forecasting each variable is done using the Holt Winter method, so that, the fluctuations in variable forecasting results follow actual data fluctuations. After obtaining the results of forecasting each variable, then forecasting using ARIMA-QR is done to get the value of forecasting the price of rice. In the ARIMA-QR method, variables that are significantly or not significantly known for the price of rice can be known. This cannot be obtained, if only using the ARIMA method.

Whereas the advantages of using the ARIMA method alone can directly predict paddy prices until 2017 even though the data is only up to 2015. Figure 17 is a graph of the actual data values and forecasting values using ARIMA and ARIMA-QR until 2017. Figure 17, we can see that the graph of the ARIMA-QR method can more closely follow price fluctuations. This is because forecasting is indeed done for each variable involved.

ARIMA and ARIMA-QR models for forecasting paddy prices have been obtained. ARIMA modeling results can be said to be good because it has a small MAPE value of 1.56% for training sets and 0.93% for testing sets. Forecasting using the ARIMA Model in this case study is better than the ARIMA-QR method with quantile 0.25 and 0.75. Whereas the QR ARIMA with quantile 0.50 (median) has a MAPE value that is better than the ARIMA method.

The ARIMA and ARIMA-QR models each have advantages and limitations. For the ARIMA Model, even though the error value is small but the forecasting value graph tends to resemble a straight line and the value fluctuations are not seen significantly. Whereas in the ARIMA-QR, the forecasting graph results more fluctuate following the actual data. Of course for analysis needs

and compared to the actual data, the ARIMA-QR is better. However, ARIMA_QR requires the prediction of each influential variable first.

With the quantile regression model, it can also be seen how the relationship between the MPD price variable, the number of stocks, world paddy prices, national holidays and price forecasting values use ARIMA to increase paddy prices on certain quantiles, namely quantile 0.25, 0.50 and 0.75. From this study it was found that there are several variables that influence paddy prices including the price of milled dry rice, the amount of rice stock in BULOG'S warehouse and the results of price forecasting using ARIMA. But the variable that most influences the price is the price of milled dry grain (GKG) variable and the price forecasting result using ARIMA. An increase in these variables will affect the increase in paddy prices. While the holiday variable is declared not to have a significant effect.

CONCLUSION

For this reason in the subsequent research, it is necessary to do a more in-depth analysis of why the national holiday variable is stated not to have a significant effect on prices while in reality at times these prices can rise. The national holiday variables used in this study are only limited to the Eid-ul-Fitr holidays, so, the possibility of causing not much effect on the price of rice. For further research, this variable can be developed for data on other national holidays, so that, it can be seen more how it affects prices.

ACKNOWLEDGEMENT

The research was funded by Institut Teknologi Sepuluh Nopember under the 2018 Research Based Community Service scheme.

REFERENCES

- Adrian, O., 2017. [Why is every year before Ramadan the price of goods up?]. Kompasiana.Com, Indonesia. (In Indonesian)
- Anonymous, 2016. [Overview of Bulog Public Corporation]. Indonesian Bureau of Logistics, South Jakarta, Indonesia. (In Indonesian) <http://bulog.co.id/sekilas.php>
- Anonymous, 2018. [Inflation makes the price of needs go up, apparently this is the reason]. Money Smart Indonesia, South Jakarta, Indonesia. (In Indonesian) <https://www.moneysmart.id/penyebab-inflasi-yang-bikin-harga-naik-turun/>
- Arifin, B., 2013. [Fluctuations in agricultural commodity prices]. Kompas, Kompas Gramedia. (In Indonesian) <https://regional.kompas.com/read/2013/01/07/03032863/Fluktuasi.Harga.Komoditas.Pertanian>
- Baderi, F., 2018. [The importance of maintaining the stability of rice prices in Indonesia]. Harian Ekonomi Neraca, Jakarta, Indonesia. (In Indonesian) <http://www.neraca.co.id/article/96676/pentingnya-menjaga-kestabilan-harga-beras-di-indonesia>
- Chernozhukov, V., C. Hansen, K. Wuthrich, C. Hansen and K. Wuthrich, 2017. Instrumental Variable Quantile Regression. In: Handbook of Quantile Regression, Koenker, R., V. Chernozhukov, X. He and L. Peng, (Eds.). CRC Press, Boca Raton, Florida, USA., ISBN:9781498725286, pp: 1-25.
- Darekar, A. and A.A. Reddy, 2017c. Cotton price forecasting in major producing states. Econ. Affairs, 62: 373-378.
- Darekar, A. and A.A. Reddy, 2017b. Predicting market price of soybean in major India studies through ARIMA model. J. Food Legumes, 30: 73-76.
- Darekar, A. and A.A. Reddy, 2017a. Forecasting of common paddy prices in India. J. Rice Res., 10: 71-75.
- Jackson, E.A., 2018. Comparison between static and dynamic forecast in autoregressive integrated moving average for seasonally adjusted headline consumer price index. SSRN. Electron. J., 1: 1-14.
- Jackson, E.A., A. Sillah and E. Tamuke, 2018. Modelling monthly Headline Consumer Price Index (HCPI) through seasonal Box-Jenkins methodology. Intl. J. Sci., 7: 51-56.
- Jadhav, V., B.V. Reddy and G.M. Gaddi, 2017. Application of ARIMA model for forecasting agricultural prices. J. Agric. Sci. Technol., 19: 981-982.
- Kocaarslan, B., U. Soytaş, R. Sari and E. Ugurlu, 2018. The changing role of financial stress, oil price and gold price in financial contagion among US and BRIC markets. Intl. Rev. Finance, Vol. 2018, 10.1111/irfi.12189
- Liu, B., J. Nowotarski, T. Hong and R. Weron, 2017. Probabilistic load forecasting via quantile regression averaging on sister forecasts. IEEE. Trans. Smart Grid, 8: 730-737.
- Nusair, S.A. and J.A. Al-Khasawneh, 2018. Oil price shocks and stock market returns of the GCC countries: Empirical evidence from quantile regression analysis. Econ. Change Restructuring, 51: 339-372.
- Panasa, V., R.V. Kumari, G. Ramakrishna and S. Kaviraju, 2017. Maize price forecasting using Auto Regressive Integrated Moving Average (ARIMA) model. Intl. J. Curr. Microbiol. Appl. Sci., 6: 2887-2895.

- Post, S., 2017. [Fluctuating prices of basic commodities, make traders confused]. Sriwijaya, Tangerang, Indonesia. (In Indonesian) <http://palembang.tribunnews.com/2018/03/14/naik-turunnya-harga-bahan-pokok-bikin-pedagang-bingung>
- Rahman, A., 2017. [Prices of basic foods ahead of the fasting month]. Harian Metro, Kuala Lumpur, Malaysia. (In Indonesian)
- Reboredo, J.C. and A. Ugolini, 2017. Quantile causality between gold commodity and gold stock prices. *Resour. Policy*, 53: 56-63.
- Reboredo, J.C. and N. Naifar, 2017. Do Islamic bond (Sukuk) prices reflect financial and policy uncertainty? A quantile regression approach. *Emerging Markets Finance Trade*, 53: 1535-1546.
- Singh, K. and A. Kumar, 2017. Autoregressive integrated moving average model for gold price forecasting: Evidence from the Indian market. *Indian J. Res. Capital Markets*, 4: 33-43.
- Syafaat, N., 2017. [Review of rice price policies]. Detiknews, Jakarta, Indonesia. (In Indonesian) <https://news.detik.com/read/2017/08/01/131654/3581807/103/mengkaji-ulang-kebijakan-harga-beras>.
- Tan, X.P. and X.Y. Wang, 2017. Dependence changes between the carbon price and its fundamentals: A quantile regression approach. *Appl. Energy*, 190: 306-325.
- Wulandari, D., 2017. [These 4 factors that cause the price of material price hike]. Economy of Indonesia, Jakarta, Indonesia. (In Malay) <https://ekonomi.bisnis.com/read/20170412/12/644833/ini-4-faktor-pe-nyebab-kenaikan-harga-bahan-pokok>
- Zahari, F.Z., K. Khalid, R. Roslan, S. Sufahani and M. Mohamad *et al.*, 2018. Forecasting natural rubber price in Malaysia using arima. *J. Phys. Conf. Ser.*, 995: 1-8.