Prediction of Inflation in Indonesia Using Nonparametric Regression Approach Based on Local Polynomial Estimator

V Fibriyani¹, N Chamidah²,

¹Faculty of Economics, Universitas Merdeka Pasuruan, Indonesia
²Department of Mathematics, Faculty of Science and Technology, Airlangga University, Indonesia

Corresponding author: nur-c@fst.unair.ac.id

ABSTRACT
Inflation is one of macroeconomics indicator can describe the economic development of the country. Inflation is one of the important factors, the high inflation can disturb the economy a country so it has been concerned by the government. Inflation can be caused by various factors, one of them come from the money supply. Inflation in Indonesia has a high variance so it needs a model based on high degree estimator. In this research, we predict inflation using by two approaches i.e. nonparametric (local model using local polynomial estimator) and parametric (global model). Inflation predicting can is used to prepare government policies to keep inflation at a stable. The Mean Absolute Percentage Error (MAPE) is used to know the accuracy predicted value. For nonparametric approaches, i.e. the local linear model has a MAPE as 4,933% and local quadratic model has a MAPE as 4,692%, both of them is highly accurate to predict. But, for the parametric approach (global model) has a MAPE as 29.43%. Based on MAPE, we conclude that the best model is the local quadratic model at predicting inflation.

Keywords: inflation, global model, local model, local polynomial estimator, MAPE.

1. INTRODUCTION
Inflation is a macroeconomic indicator that has an essential role in a country so that inflation caused a much influence in economic growth. The rate of inflation that continues to increase can become an obstacle in a country's economic growth. High inflation can cause a decreased prosperity level and reduced economic growth. In 1998, Indonesia experienced high inflation. One of the causes was the uncontrolled printing of money, which more than 3 currencies, so that the amount of money in circulation was huge [1]. In general, inflation is defined as the increasing price of goods and services, which is a basic need of society or a decrease in the country's currency [2]. An increase in the price of just one or two goods cannot be called inflation but a general and continuous rise in prices and growth in the price of an item to expand (causing an increase in other goods). Deflation is the opposite of inflation [3].

High inflation caused an increase in the price of imported commodities. It increased the amount of foreign debt as the impact of the rupiah exchange towards foreign currencies, especially the US dollar. However, inflation conditions that are too low also affect a country's economic downturn. Thus, it needs an effort to keep inflation in Indonesia at a low and stable level. Stable inflation is a prerequisite for sustained economic growth, thereby increasing society's welfare [4]. Inflation control is considering that a high and unstable rate of inflation will negatively impact society's social conditions. High inflation can cause a continuous decrease in people's income. Besides, uneven inflation creates uncertainty to make decisions in consumption for economic factors and investment and production, reducing economic growth [3].

As a monetary authority, Bank Indonesia regulates the amount and allocation of the money supply to achieve macroeconomic objectives. There are several causes of high inflation; one is too much money circulation [5]. Kesavarajah and Amirthalingam [6] said that the consequences of expansive monetary policy made Sri Lanka in inflation during the post-liberalization period. In order of the situation to preserve price stability, the money supply is one of the vital policy instruments. Amin [7] explained that the weak implementation of monetary policy increases by Central Bank in Bangladesh made inflation. There is a fluctuating movement at an inflation rate; it can be said that inflation has a high variance, so appropriate modeling is needed. Inflation is an essential indicator in macroeconomics, so it requires inflation prediction as a reference of government in making policies and controlling economic growth. One of the statistical techniques used to describe a functional relationship between response variables and predictor variables in regression analysis; it can also use prediction. In regression analysis, a function can be estimated through two approaches, which are parametric and nonparametric. One of the parametric approaches used to predict is the linear regression analysis, as has been done by Yolanda [8]. Some other researchers...
who use Linear Regression in predicting include Ismail et al. [9], Elsiddig [10], and Moriyama [11].

Nonparametric approaches are used if the functional relationship between the response variable and the predictor variable does not assume a specific form. Nonparametric regression approach has high flexibility because the regression function is not specified in a particular form but is assumed smooth. It can be estimated using certain smoothing methods based on data patterns, as described in [13]. Several estimators estimate nonparametric regression curves like Kernel estimator, local linear, local polynomial, spline, and Fourier series. Several research related to regression modeling using a nonparametric approach includes designing child growth charts based on local polynomial estimators by Chamidah et al. [14]. Local polynomial estimators are a popular new approach. The local polynomial estimator depends on two parameters: the order of local polynomial fit and the smoothing parameter, namely bandwidth. The effect of those parameters is increasing variance and reducing bias if a lower-order fit and smaller bandwidth. But, it's reducing variance and increasing if a lower order fit and large bandwidth. Local polynomial estimators have a special case. That is, if the order is equal to zero, then it is called the Kernel estimator, and if the order is similar to one, then it is called a local linear estimator.

Many previous studies on using various estimators in nonparametric regression approaches, such as using kernel estimators by Chamidah and Saifudin [15] and Lestari et al. [16]. The modeling used local linear estimator is explained by Chamidah and Rifada [17, 18], and Chamidah et al. [19] and using penalized spline estimators is done by Chamidah et al. [20]. Also, a non-parametric regression approach based on the local polynomial estimator is applied to the classification of cystic and tumor diseases, as explained by Chamidah et al. [21].

The nonparametric regression approach is more complicated than a parametric approach. There is no guarantee that a detailed process causes a better model. In this paper, the researcher discussed the nonparametric regression approach using the local polynomial estimator and parametric regression approach using the global model. Researchers apply the model to predict inflation in Indonesia using inflation data from September 2004 until August 2019.

2. MATERIAL AND METHODS

2.1. Research Materials

In this research, the Indonesian inflation act as a response variable and money supply as a predictor. The researcher gets the data from Bank Indonesia (https://www.bi.go.id). There are 180 observations from September 2004 until August 2019. Researchers used 80% from total observations, starting from the first until 144th observation as in sample data, to build the estimation model and used 145th observation until 180th as out sample data (20% of data) for prediction of inflation.

2.2. Regression Modeling

2.2.1. Parametric Regression

Parametric regression is a regression approach to determine the pattern of relationship between the response variable and predictor variable where the shape of the curve is known. The equation of global model using linear regression as follows [22].

\[ y_t = \beta_0 + \beta_1 x_{t1} + \varepsilon_t \quad t = 1,2,...,n \]  

(1)

In the matrix, we obtain the regression model:

\[ \mathbf{y} = \mathbf{X}\beta + \mathbf{\varepsilon} \sim N(\mathbf{0}, \sigma^2\mathbf{I}) \]

To estimate a regression coefficient \( \beta \), we use OLS (Ordinary Least Square) to minimize a sum of square error. We have estimator

\[ \hat{\beta} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y} \]  

(2)

2.2.2. Local Polynomial

Given a response nonparametric regression:

\[ y_t = m(x_t) + \varepsilon_t \quad t = 1,2,...,n \]  

(3)

Where \( m(x_t) \) is an unknown smooth function, and \( \varepsilon_t \) is a random error with assumed zero mean and \( \sigma^2 \) variance. To estimate function \( m(x_t) \) in equation (3) using local polynomial estimator, function \( m(x_t) \) can be approximated by Taylor expansion as follow by:

\[ m(x_t) \approx \sum_{k=0}^{p} \frac{m^{(k)}(x_0)}{k!} (x - x_0)^k = \sum_{k=0}^{p} \beta_k (x_0)^k (x - x_0)^k \]  

(4)

With \( m^{(k)}(x_0) \) is the \( -k \) derivative of \( m(x_0) \) against \( x_0 \), \( x \in (x_0 - h, x_0 + h) \). If equation 4 written in form matrix, we obtain

\[ m(x_t) = x_{t0}^T \hat{\beta}'(x_0) \]  

(5)
with $x_{x_0} = [1 \ (x-x_0) \ \ldots \ (x-x_0)^p]$, $x \in (x_0-h, x_0+h)$ and $\beta^*(x_0) = [\beta_0^*(x_0) \ \beta_1^*(x_0) \ \ldots \ \beta_p^*(x_0)]^T$.

Based on equation (5), we can write an equation (3) to be:

$$y = x_{x_0}^T \beta^*(x_0) + \varepsilon \tag{6}$$

To estimate $\beta^*(x_0)$ based on local polynomial estimator, from $n$ sample paired data $(x_t, y_t)_{t=1}^n$ so equation (6) is define:

$$y_1 = \beta_0^*(x_0) + \beta_1^*(x_0)(x_1-x_0) + \beta_2^*(x_0)(x_1-x_0)^2 + \ldots + \beta_p^*(x_0)(x_1-x_0)^p + \varepsilon_1 \\quad (7)$$

Expression (7) in matrix form

$$y^* = X_{x_0}^T \beta^*(x_0) + \varepsilon^* \tag{8}$$

with

$$X_{x_0} = \begin{bmatrix} 1 & (x_1-x_0) & \ldots & (x_1-x_0)^p \\ \vdots & \vdots & \ddots & \vdots \\ 1 & (x_n-x_0) & \ldots & (x_n-x_0)^p \\ \end{bmatrix}, \quad y^* = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}, \quad \varepsilon^* = \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix} \tag{9}$$

Obtaining $\beta^*(x_0)$ estimator is done by minimizing the Weighted Least Square (WLS) criteria below:

$$Q(x_0) = (y^* - X_{x_0}^T \beta^*(x_0))^T K_{h^*}(x_0) (y^* - X_{x_0}^T \beta^*(x_0)) \tag{10}$$

with

$$K_{h^*}(x_0) = \begin{bmatrix} K_{h^*}(x_1-x_0) & 0 & \ldots & 0 \\ 0 & K_{h^*}(x_2-x_0) & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & K_{h^*}(x_n-x_0) \end{bmatrix} \tag{11}$$

$K_{h^*}(x_0)$ is a matrix contain a weighted function that $K_{h^*}(\cdot)$ is Kernel function [13] with bandwidth $h^*$ is defined as follows:

$$K_{h^*}(x) = \frac{1}{h} K \left( \frac{x}{h} \right); \quad -\infty < x < \infty \text{ and } h^* > 0$$

In this study, we use types of Kernel function is Gaussian Kernel is defined as follows [23]:

$$K(x) = \frac{1}{\sqrt{2\pi}} \exp \left( -\frac{x^2}{2} \right)$$

To estimate model (1) from the sample data $(x_t, y_t)_{t=1}^n$, we use a local polynomial estimator defined as follows expression (10) with $K_{h^*}(x)$ is Gaussian Kernel. The fitted value $\hat{\beta}^*(x_0)$ is $\hat{\beta}^*(x_0)$, which is substituted in equation (10), minimizing $Q(x_0)$. Estimating $\beta^*(x_0)$ ads to determine the first derivative of equation (10) against $\beta^*(x_0)$ and then equalizing to zero, so

$$\hat{\beta}^*(x_0) = (X_{x_0}^T K_{h^*}(x_0) X_{x_0})^{-1} X_{x_0}^T K_{h^*}(x_0) y^* \tag{12}$$

Based on equations (5) and (12), the local polynomial estimator for $\hat{m}(x_t)$ is defined:

$$\hat{m}(x_t) = X_{x_0}^T (X_{x_0}^T K_{h^*}(x_0) X_{x_0})^{-1} X_{x_0}^T K_{h^*}(x_0) y^* \tag{13}$$

Bandwidth $h$ is a smoothing parameter that controls the smoothness of the curve. The optimum bandwidth selection is based on minimum Cross-Validation (CV) value as follows [24]:

$$h_{opt} = \arg\min_{h} CV(h)$$
\( h_{opt} = \arg \min_n \frac{1}{n} \sum_{i=1}^{n} \left( \hat{h}_{n,i}^{(h)} (x_i) - y_i \right)^2 \) \hspace{1cm} (14)

2.2.3. The Goodness of Fit

The goodness of fit of the model used in this research is Mean Absolute Percentage Error (MAPE). MAPE is used to measure the accuracy of model estimate values expressed in terms of the mean absolute percentage of errors. Mathematically it can be written [26]:

\[
MAPE = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{\hat{y}_i - y_i}{y_i} \right| \times 100
\] \hspace{1cm} (15)

Following this is a table of typical MAPE values and their interpretation:

| MAPE  | Interpretation     |
|-------|--------------------|
| < 10  | Highly Accurate    |
| 10 – 20 | Accurate           |
| 20 – 50 | Reasonable        |
| > 50  | Inaccurate         |

2.3. Research Methods

The stages of this research as follows:
1. Describing the characteristics of inflation and money supply in Indonesia
2. Making scatter plot about the relationship between inflation and money supply
3. Determining the optimal bandwidth each order (local linear and local quadratic) for nonparametric regression approach using local polynomial estimator by sample data with minimum CV criteria as Eq. (14).
4. Determining the accuracy of predicting is use MAPE in Eq. (15) from each nonparametric regression model
5. Estimating inflation using global model refers to Eq. (2).
6. Determining the accuracy of predicting use MAPE as Eq. (15).
7. Comparing models based on MAPE

3. RESULTS AND DISCUSSION

There are 180 observations in this study, which are inflation rate data from September 2004 to August 2019. The data was divided into 2 parts, the first 80% data is a sample to build the model and the next 20% of the data as out sample for making predictions. The inflation rate (%) in this study as a response variable, while the money supply (thousand trillion Rupiah) as a predictor variable. The following are descriptive statistics of macroeconomic indicators for the period 2004 – 2019.

| Variable                          | Statistic |
|----------------------------------|-----------|
|                                  | Mean      | Varian | Min. | Max.     |
| Inflation (%)                    | 6.302     | 11.920 | 2.410| 18.380  |
| Money Supply (Thousand Trillion Rupiah) | 3.106     | 2.409  | 0.987 | 5.941   |

Based on Table 1, it is known that the average inflation rate since the beginning of the reformation era up to date is 6.302%, with a relatively high diversity of 11.920%. Indonesia experienced the highest inflation rate of 18.380% in November 2005, where there has been an increase in oil prices worldwide and international interest rates, putting pressure on domestic monetary stability. This condition resulted in domestic fuel prices and decreasing in Rupiah. To keep Rupiah from depleting, the government imposed an economic policy by increasing the BI rate to 12.5% [25]. Indonesia was at the lowest inflation level of 2.410% in November 2009. The average amount of money supply in Indonesia since the reformation era amounted to 3,106 trillion Rupiah with variations of 2,410 trillion Rupiah. Besides occurring in 2005, a sharp increase in inflation also occurred in 2008. This is also due to the global financial crisis due to the deteriorating condition of the United States economy. The shocks that happened in the superpower had an impact on the world economy, including Indonesia. Weakened due to falling oil prices when the growth of goods exports caused the global crisis in the fourth quarter of 2008 [27].

We make scatter plot to known information about the relationship between response and predictor variable. From this plot, we determine the pattern of functional relationship between inflation and money supply. The pattern of the relationship between inflation and money supply can be seen in Figure 1.
Figure 1. Scatter Plot between Inflation vs Money Supply

Figure 1. Shows that the observations points spread and not follow a certain curve shape, so we use nonparametric regression approach. This research uses nonparametric regression approach based on local polynomial estimator because this method can be modeling a high order.

3.1. Modeling of Inflation Based on Money Supply

In this study, the researcher uses a nonparametric regression approach using local polynomial estimator. The data used for modeling are 144 observations. The optimal bandwidth for local linear model is 0.031 with CV value as 1.281, and for local quadratic model is 0.046 with CV value as 1.244. Selection of optimum bandwidth using minimum CV value, as shown in Figure 2.

Figure 2. (a) For local linear model

Figure 3. Plot Bandwidth versus CV value for Local Linear and Local Quadratic Model

Figure 3 shows the estimation results from local linear model (as green line) and local quadratic model (as red line). Visually, there appears no difference between estimation of inflation using local linear model and local quadratic model. It causes both of them to obtain the estimated value that follows the data.

Figure 4. Plot of Observations and Estimation Based on Local Linear and Local Quadratic Models
Table 2. Optimum bandwidth \( (h) \) and MAPE of the nonparametric regression approach

| Degree | \( h \) optimum | CV value | MAPE  |
|--------|-----------------|----------|-------|
| 1      | 0.031           | 1.281    | 6.856 |
| 2      | 0.046           | 1.244    | 6.901 |

Shown in Table 2, local linear model has a MAPE value as 6.856\%, and local quadratic model has a MAPE value as 6.901\%. MAPE of nonparametric regression approach is smaller than 10; it means that the model has a highly accurate prediction, as Moreno et al. [26]. Based on the MAPE value, local linear model is better than local quadratic model. Modeling inflation using nonparametric regression approach based on local polynomial estimator, as follows:

Local linear model with optimum bandwidth as \( h = 0.031 \) and \( x_0 = 5.941 \):

\[
\hat{y} = 3.398 + 2.098(x - 5.941), \quad x \in (5.910; 5.972) \tag{14}
\]

Based on expression (14), if a money supply in Indonesia as 5.9 thousand trillion Rupiah, inflation is 3.312\%.

![Figure 5. Plot prediction of inflation versus money supply](image)

Modeling inflation using parametric approach (global model) is explained in the following table:

Table 3. MAPE of parametric regression approach

| Degree | MAPE  |
|--------|-------|
| 1      | 33.587 |
| 2      | 46.776 |

Shown in Table 3, global linear model has a MAPE value as 33.587. In global quadratic model has a MAPE value as 46.776. MAPE of parametric regression approach is smaller than 50; it means that model has a reasonable prediction, as explained by Moreno et al. [25]. Based on the MAPE value, global linear model is better than global quadratic model. Modeling inflation using a parametric regression approach as follows by:

\[
\hat{y} = 10.353 - 1.308x \tag{16}
\]

Based on expression (16), if a money supply in Indonesia as 5.9 thousand trillion Rupiah, inflation is 2.636\%.

Table 4. Comparing Models Based on MAPE Value

| Approach   | Degree | Model       | MAPE  |
|------------|--------|-------------|-------|
| Nonparametric | 1      | Local Linear| 6.856 |
|            | 2      | Local Quadratic | 6.901 |
| Parametric | 1      | Global Linear | 33.587 |
|            | 2      | Global Quadratic | 46.776 |

It can be seen in Table 4 that the nonparametric regression approach, both in local linear model and local quadratic model, is a model with a high ability to predict, as explained...
in Moreno [25], that a model with MAPE value of less than 10 is categorized as highly accurate predicting. In the parametric approach, both linear and quadratic models obtain MAPE values between 20 and 50, so both models are reasonably predicting models.

4. CONCLUSION

The nonparametric approach based on local polynomial estimators can adequately accommodate the fluctuating inflation data pattern in Indonesia. This is indicated by the MAPE value obtained from the model in the highly accurate predicting category. We conclude that a nonparametric model has a highly accurate prediction.

REFERENCES

[1] Atmadja, A. S. (1999). Inflasi di Indonesia: Sumber-sumber Penyebab dan Pengendaliannya. Jurnal Akuntansi dan Keuangan. 1(1): 54-67.
[2] Badan Pusat Statistik (BPS). 2019. Laporan Perekonomian Indonesia. BPS.
[3] Bank Indonesia. 2019. Bank Indonesia dan Inflasi. Bank Indonesia
[4] Ginting, A. M. (2017). Analisis Faktor-Faktor Yang Mempengaruhi Inflasi: Studi Kasus Di Indonesia Periode Tahun 2004-2014. Kajian. 21(1): 37-58.
[5] Wulan, E. R., Nurfaiza, S. (2014). Analysis of Factors Affecting Inflation in Indonesia: an Islamic Perspective. International Journal of Nusantara Islam. 2(2): 67-80.
[6] Kesavarajah, M., Amirthalingam K. (2012). The Nexus Between Money Supply and Inflation in Sri Lanka. Jaffna University International Research Conference. 232.
[7] Amin, S. B. (2011). Quantity Theory of Money and its Applicability: The Case of Bangladesh. World Review of Business Research. 1(4): 33-43.
[8] Yolanda, Y. (2017). Analysis of Factors Affecting Inflation and its Impact on Human Development Index and Poverty in Indonesia. European Research Studies Journal. XX(4B): 38 – 56.
[9] Yuhan, R. J., Sohibien, G. P. D. (2018). Relationship Between Inflation, Exchange Rate and Money Supply in Indonesia Using Threshold Vector Autoregressive (TVAR). In AIP Conference Proceedings. 2014(1): 020128.
[10] Nyoni, T. (2018). Modeling and Forecasting Inflation in Kenya: Recent Insights from ARIMA and GARCH Analysis. Dimorian Review. 5(6): 16-40.
[11] Sukono, H. Y., Suhartono, S. B., Bon, A. T. B., Supian, S. (2016). Indonesian Financial Data Modeling and Forecasting By Using Econometrics Time Series and Neural Network. Global Journal of Pure and Applied Mathematics. 12(4): 3745-3757.
[12] Sari, N. R., Mahmudy, W. F., Wibawa A. P. (2016). Backpropagation in Neural Network Method for Inflation Rate Forecasting in Indonesia. Int. J. Advance Soft Compu. Appl. 8(3): 70 – 87.
[13] Eubank, R. (1999). Spline Smoothing and Nonparametric Regression Second Edition. Marcel Dekker. New York.
[14] Chamidah, N., Budiantara, I. N., Sunaryo, S., Zain, I. (2012). Designing of Child Growth Chart Based on Multi-response Local Polynomial Modeling. Journal of Mathematics and Statistics. 8(3): 342-347.
[15] Chamidah, N., Saifudin, T. (2013). Estimation of Children Growth Curve Based on Kernel Smoothing In Multi-Response Nonparametric Regression. Applied Mathematical Sciences. 7(37-40): 1839-1847.
[16] Lestari, B., Fatmawati, B. IN, Chamidah, N.(2018). Estimation of Regression Function in Multi-Response Nonparametric Regression Model Using Smoothing Spline and Kernel Estimators. In Journal of Physics: Conference Series. 1097(1): 012091.
[17] Chamidah, N., Rifada, M. (2016). Local Linear Estimator In Bi-Response Semiparametric Regression Model For Estimating Median Growth Charts of Children. Far East Journal of Mathematical Sciences. 99(8): 1233-1244.
[18] Chamidah, N., Rifada, M. (2016). Estimation of Median Growth Curves for Children Up Two Years Old Based on Biresponse Local Linear Estimator. In AIP Conference Proceedings. 1718(1): 110001. AIP Publishing LLC.
[19] Chamidah, N., Tjahjono, E., Fadilah, A. R., Lestari, B. (2018). Standard growth charts for weight of children in east java using local linear estimator. In Journal of Physics: Conference Series. 1097(1): 012092.
[20] Chamidah, N., Zaman, B., Muniroh, L., Lestari, B (2019). Estimation of Median Growth Charts for Height of Children in East Java Province of
Indonesia Using Penalized Spline Estimator, International Conference Proceedings GCEAS.5: 68-78.

[21] Chamidah N, Gusti K H, Tjahjono E and Lestari B (2019). Improving of classification accuracy of cyst and tumor using local polynomial estimator TELKOMNIKA. 17(3): 1492-1500.

[22] Eubank, R. L., Huang, C., Maldonado, Y. M., Wang, N., Wang, S., Buchanan, R. J. (2004). Smoothing Spline Estimation in Varying-Coefficient Models. Journal of the Royal Statistical Society: Series B (Statistical Methodology). 66(3): 653-667.

[23] Hardle, W., (1990). Applied Nonparametric Regression. New York: Cambridge University Press.

[24] Bappenas. 2006. Buku Pegangan 2006: Penyelenggaraan Pemerintahan dan Pembangunan Daerah.

[25] Moreno, J. J. M., Pol, A. P., Abad, A. S., Blasco, B. C. (2013). Using the R-MAPE Index as a Resistant Measure of Forecast Accuracy. Psicothema. 25(4): 500-506.

[26] Bappenas. 2009. Buku Pegangan 2009: Penyelenggaraan Pemerintahan dan Pembangunan Daerah.