ISSN: 2339-2541
JURNAL GAUSSIAN, Volume 12, Nomor 3, Tahun 2023, Halaman 352 - 361
Online di: https://ejournal3.undip.ac.id/index.php/gaussian/

PEMODELAN DATA LONGITUDINAL MENGGUNAKAN REGRESI POLINOMIAL LOKAL PADA KELOMPOK SAHAM PERUSAHAAN PENYEDIA JASA TELEKOMUNIKASI DENGAN GUI R

Gustyas Zella Noer Rachma[1*], Suparti[2], Rukun Santoso[3]
[1,2,3] Departemen Statistika, Fakultas Sains dan Matematika, Universitas Diponegoro
***e-mail : gustyas99@gmail.com**

DOI: 10.14710/j.gauss.12.3.352-361

Article Info:
Received: 2022-09-04
Accepted: 2022-12-10
Available Online: 2024-02-13

Keywords:
Stocks; Longitudinal Data; Local Polynomial; MSE; GUI

Abstract: The economic development of a country can be seen based on the capital market that was growing and developing. One of the most popular capital market instruments is stocks. Stocks based on market capitalization groups include longitudinal data. One of the statistical methods for longitudinal data modelling is nonparametric regression which has no modelling assumptions requirement. This research models monthly stock prices using a nonparametric local polynomial method with the selection of the best model which has minimum value of Mean Square Error (MSE). The data was divided into 2 parts, namely in sample data from November, 2018 to June, 2021 to form a model and out sample data from July, 2021 to February, 2022 used for evaluation of model performance by Mean Absolute Percentage Error (MAPE) values. The best model is the local polynomial model with Biweight kernel function of degree 5, local point of 4, *bandwidth* of 37, and MSE value of 0.03481085. MAPE out sample of data value is 31.13%, which indicating that the model has sufficient forecasting. In this research arrange Graphical User Interface (GUI) by using R software with shiny package is built to make display output data analyzing more easy and more interactive.

1. PENDAHULUAN

Perkembangan perekonomian dari suatu negara dapat dilihat berdasarkan pasar modal yang sedang tumbuh dan berkembang secara baik. Pasar modal menyediakan kegiatan jual-beli surat berharga serta kegiatan terkait instrument keuangan yang lainnya. Saham adalah surat berharga yang digunakan sebagai sebuah bukti penyertaan atau kepemilikan seseorang atau badan usaha pada suatu perusahaan (Purnomo *et al.*, 2013). Berdasarkan nilai kapitalisasinya, saham terbagi menjadi tiga kelompok, diantaranya kelompok saham kapitalisasi besar (*big-cap*), saham kapitalisasi sedang (*mid-cap*), dan saham kapitalisasi kecil (*small-cap*).

Metode regresi nonparametrik yaitu metode regresi yang digunakan apabila tidak diketahui bentuk kurva atau plot regresinya. Regresi nonparametrik tidak memerlukan asumsi pemodelan yang berarti regresi nonparametrik sangat objektif dan fleksibel (Budiantara, 2009). Data longitudinal adalah sebuah pengamatan sebanyak n subjek yang saling independent, dan setiap subjek dilakukan secara berulang dalam waktu yang berbeda yang mana saling dependen (Wu dan Zhang, 2006).

Pemodelan dengan data longitudinal akan lebih tepat apabila digunakan metode pendekatan regresi nonparametrik dikarenakan lebih fleksibel dibandingkan regresi parametrik (Wu dan Zhang, 2006). Ada beberapa metode regresi nonparametrik yang dapat digunakan untuk mengestimasi suatu kurva regresi, salah satunya yaitu menggunakan metode polinomial lokal. Salah satu kelebihan dari metode polinomial lokal yaitu mempunyai kemampuan untuk melakukan adaptasi dengan data dan berarti membagi data ke dalam wilayah tertentu lalu dilakukan estimasi pada wilayah yang telah ditetapkan nilainya tersebut (Fan dan Gijbels, 1997). Metode optimasi untuk mencari model polinomial lokal terbaik dapat menggunakan *Mean Square Error* (MSE).

Penelitian oleh (Khalid, 2015) yaitu melakukan pemodelan harga saham perusahaan Astra Ortoparts Tbk., Astra Graphia Tbk., dan Mahaka Media Tbk. dengan metode regresi nonparametrik data longitudinal berdasarkan estimator polinomial lokal. Penelitian yang dilakukan Khalid (2015) memodelkan data longitudinal pada regresi nonparametrik polinomial lokal dengan bantuan *software* R berbasis *Command Line Interface* (CLI). Penelitian ini bertujuan untuk mendapatkan model regresi nonparametrik polinomial lokal terbaik pada data longitudinal data harga saham bulan November 2018 sampai dengan Februari 2022 dan akan dikembangkan dengan *Graphical User Interface* (GUI) program R. Perhitungan regresi nonparametrik polinomial lokal termasuk rumit dan banyak prosesnya, maka dibutuhkan *software* untuk menyelesaikannya, salah satunya adalah *software* R. Salah satu pengembangan dalam penelitian ini adalah disusun *Graphical User Interface* (GUI) program R untuk memudahkan analisis data dan menampilkan output yang lebih menarik.

2. TINJAUAN PUSTAKA

Bursa Efek Indonesia menjelaskan bahwa saham (*stock*) adalah salah satu instrument pasar keuangan yang paling terkenal. Saham merupakan surat berharga yang bias dibeli atau dijual oleh perorangan atau Lembaga dimana pasar tempat surat tersebut diperjual-belikan (Hadi, 2013). Sehubungan dengan perkembangan pasar keuangan, saham dapat terbagi berdasarkan besarnya nilai kapitalisasi pasar. Saham berdasarkan nilai kapitalisasi pasar menurut Hadi (2013) terbagi dalam 3 kelompok yaitu: kelompok kapitalisasi tinggi (nilai kapitalisasi lebih dari 40 triliun rupiah), kelompok kapitalisasi sedang (nilai kapitalisasi antara 1 sampai 40 triliun rupiah), dan kelompok kapitalisasi rendah (nilai kapitalisasi kurang dari 1 triliun rupiah).

Data longitudinal adalah data yang didapat berdasarkan pengamatan berulang pada setiap subjek dalam waktu yang berbeda. Pada data longitudinal, pengukuran antar subjek diasumsikan independen, namun pengukuran di antara subjek yang sama adalah dependen (Wu dan Zhang, 2006). Misalkan t_{ij} merupakan waktu pengukuran subjek ke-*i* pada waktu ke-*j*, y_{ij} merupakan variabel respon yang diukur pada waktu t_{ij}, dan x_{ij} merupakan variabel prediktor yang diamati pada waktu t_{ij}, n merupakan banyaknya subjek, dan n_i merupakan banyaknya pengamatan subjek ke-*i*, maka diberikan data longitudinal sebagai berikut:

$$\left(t_{ij}, x_{ij}, y_{ij}\right), \quad i = 1,2,\dots,n \ ; \quad j = 1,2,\dots,n_i \qquad (1)$$

Model regresi nonparametrik pada data longitudinal sebagai berikut: (Wu dan Zhang, 2006)

$$y_{ij} = \eta(x_{ij}) + \varepsilon_{ij}, \quad i = 1,2,\dots,n \ ; \quad j = 1,2,\dots,n_i \qquad (2)$$

dengan $\eta(x_{ij})$ merupakan suatu fungsi yang tidak diketahui bentuk kurvanya berdasarkan data longitudinal dan ε_{ij} merupakan suatu error pengukuran dalam subjek ke-*i* pada waktu ke-*j*.

Salah satu metode untuk mengestimasi model regresi polinomial lokal yaitu menggunakan metode WLS (*Weighted Least Square*) yaitu metode estimasi parameter yang digunakan untuk menentukan nilai taksiran parameter yang meminimumkan jumlah kuadrat residual dengan dilakukan pembobotan pada suatu faktor yang tepat yang lebih dekat dengan titik lokalnya. Pembobot yang dapat digunakan untuk memperoleh estimasi polinomial lokal yaitu fungsi kernel (Eubank, 1988). Suatu fungsi kernel $K(x)$ disebut fungsi kernel jika K fungsi yang riil, kontinu, terbatas, simetris, serta memenuhi $\int_{-\infty}^{\infty} K(x)\, dx = 1$.

Menurut Ogden (1997) ada beberapa jenis fungsi kernel diantaranya:

1. Kernel Gaussian

$$K(x) = \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}}, -\infty < x < \infty$$

2. Kernel Segitiga

$$K(x) = \begin{cases} 1 - |x|, |x| \leq 1 \\ 0, \quad x \ yang \ lain \end{cases}$$

3. Kernel Epanechnikov

$$K(x) = \begin{cases} \frac{3}{4}(1 - x^2), |x| \leq 1 \\ 0, x \ yang \ lain \end{cases}$$

4. Kernel Biweight

$$K(x) = \begin{cases} \frac{15}{16}(1 - x^2)^2, |x| \leq 1 \\ 0, x \ yang \ lain \end{cases}$$

Regresi nonparametrik polinomial lokal merupakan metode regresi nonparametrik dengan suatu fungsi regresi $\eta(x)$ yang diestimasi menggunakan fungsi polinomial. Polinomial lokal mengadopsi dari perluasan deret Taylor di sekitar titik lokal x_0. Deret Taylor menyatakan bahwa setiap fungsi mulus (*smooth*) bisa didekati menggunakan polinomial dari beberapa derajat. Misal x_0 adalah titik lokal yang sudah ditentukan sehingga fungsi η akan diestimasi menggunakan estimator Kernel.

Berdasarkan deret Taylor, $\eta(x_{ij})$ pada persamaan (2) bisa didekati menggunakan polinomial berderajat p sebagai berikut (Wu dan Zhang, 2006):

$$\eta(x_{ij}) = \eta(x_0) + \eta'(x_0)\left(x_{ij} - x_0\right) + \frac{\eta''(x_{ij})}{2!}\left(x_{ij} - x_0\right)^2 + \cdots + \frac{\eta^{(p)}(x_0)}{p!}\left(x_{ij} - x_0\right)^p$$

$$i = 1,2, \dots, n \quad ; \; j = 1,2, \dots, n_i \tag{3}$$

Misalkan $\beta_r(x_0) = \frac{\eta^{(r)}(x_0)}{r!}, r = 0,1,2, \dots, p$, sehingga persamaan (3) bisa dituliskan menjadi:

$$\boldsymbol{\eta}\left(x_{ij}\right) \approx \beta_0 + \left(x_{ij} - x_0\right)\beta_1 + \cdots + \left(x_{ij} - x_0\right)^p \beta_p \tag{4}$$

Dari persamaan (2), dapat dinyatakan dalam bentuk matriks sebagai berikut :

$$\boldsymbol{Y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon} \tag{5}$$

dengan matriks basis polinomial lokal sebagai berikut :

$$\boldsymbol{X_1} = \begin{bmatrix} 1 & (x_{11} - x_0) & (x_{11} - x_0)^2 & \cdots & (x_{1n_i} - x_0)^p \\ 1 & (x_{12} - x_0) & (x_{12} - x_0)^2 & & (x_{1n_i} - x_0)^p \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & (x_{1n_i} - x_0) & (x_{1n_i} - x_0)^2 & \cdots & (x_{nn_i} - x_0)^p \end{bmatrix}$$

$$\boldsymbol{X_2} = \begin{bmatrix} 1 & (x_{21} - x_0) & (x_{21} - x_0)^2 & \cdots & (x_{2n_i} - x_0)^p \\ 1 & (x_{22} - x_0) & (x_{22} - x_0)^2 & & (x_{2n_i} - x_0)^p \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & (x_{2n_i} - x_0) & (x_{2n_i} - x_0)^2 & \cdots & (x_{nn_i} - x_0)^p \end{bmatrix}$$

$$\boldsymbol{X_n} = \begin{bmatrix} 1 & (x_{n1} - x_0) & (x_{n1} - x_0)^2 & \cdots & (x_{nn_i} - x_0)^p \\ 1 & (x_{n2} - x_0) & (x_{n2} - x_0)^2 & & (x_{nn_i} - x_0)^p \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & (x_{nn_n} - x_0) & (x_{nn_n} - x_0)^2 & \cdots & (x_{nn_n} - x_0)^p \end{bmatrix}$$

$\boldsymbol{\beta_1} = \left[\beta_{01}, \beta_{11}, \beta_{21}, \dots, \beta_{p1}\right]^T; \boldsymbol{\beta_2} = \left[\beta_{02}, \beta_{12}, \beta_{22}, \dots, \beta_{p2}\right]^T; \boldsymbol{\beta_n} = \left[\beta_{0n}, \beta_{1n}, \beta_{2n}, \dots, \beta_{pn}\right]^T$

$\boldsymbol{\varepsilon_1} = \left[\varepsilon_{11}, \varepsilon_{12}, \varepsilon_{13}, \dots, \varepsilon_{1n_1}\right]^T; \boldsymbol{\varepsilon_2} = \left[\varepsilon_{21}, \varepsilon_{22}, \varepsilon_{23}, \dots, \varepsilon_{2n_2}\right]^T; \boldsymbol{\varepsilon_n} = \left[\varepsilon_{n1}, \varepsilon_{n2}, \varepsilon_{n3}, \dots, \varepsilon_{nn_n}\right]^T$

$\boldsymbol{Y_1} = \left[Y_{11}, Y_{12}, Y_{13}, \dots, Y_{1n_1}\right]^T; \boldsymbol{Y_2} = \left[Y_{21}, Y_{22}, Y_{23}, \dots, Y_{2n_2}\right]^T; \boldsymbol{Y_n}\left[Y_{n1}, Y_{n2}, Y_{n3}, \dots, Y_{nn_n}\right]^T$

Sehingga model regresi nonparametrik polinomial lokal dapat diberikan sebagai berikut:

$$\begin{bmatrix} \boldsymbol{Y_1} \\ \boldsymbol{Y_2} \\ \vdots \\ \boldsymbol{Y_n} \end{bmatrix} = \begin{bmatrix} \boldsymbol{X_1} & 0 & \cdots & & 0 \\ 0 & \boldsymbol{X_2} & & & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & & \cdots & \boldsymbol{X_n} \end{bmatrix} \begin{bmatrix} \boldsymbol{\beta_1} \\ \boldsymbol{\beta_2} \\ \vdots \\ \boldsymbol{\beta_n} \end{bmatrix} + \begin{bmatrix} \boldsymbol{\varepsilon_1} \\ \boldsymbol{\varepsilon_2} \\ \vdots \\ \boldsymbol{\varepsilon_n} \end{bmatrix} \quad (6)$$

Parameter $\boldsymbol{\beta}$ dalam persamaan (5) merupakan parameter polinomial lokal yang dapat diestimasi dengan kriteria *Weighted Least Square* (WLS), yaitu dengan meminimumkan :

$$L = \sum_{i=1}^{n} \sum_{j=1}^{n_i} (y_{ij} - x_{ij}\beta)^2 K_h(x_{ij} - x_0) \quad (7)$$

dengan $K_h(x) = \frac{1}{h} K\left(\frac{x}{h}\right); -\infty < x < \infty \; dan \; h > 0$

$\boldsymbol{K_h}$ dituliskan dalam bentuk matriks diagonal bobot sebagai berikut :

$$\boldsymbol{K_h} = \begin{pmatrix} \boldsymbol{K_{h1}} & 0 & \dots & 0 \\ 0 & \boldsymbol{K_{h2}} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \boldsymbol{K_{hn}} \end{pmatrix}$$

dengan $\boldsymbol{K_{h1}} = \begin{pmatrix} \frac{1}{h} K\left(\frac{(x_{11} - x_0)}{h}\right) & 0 & \dots & 0 \\ 0 & \frac{1}{h} K\left(\frac{x_{12} - x_0}{h}\right) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{h} K\left(\frac{x_{1n_i} - x_0}{h}\right) \end{pmatrix}$;

$$\boldsymbol{K_{h2}} = \begin{pmatrix} \frac{1}{h} K\left(\frac{(x_{21} - x_0)}{h}\right) & 0 & \dots & 0 \\ 0 & \frac{1}{h} K\left(\frac{x_{22} - x_0}{h}\right) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{h} K\left(\frac{x_{2n_2} - x_0}{h}\right) \end{pmatrix};$$

$$\boldsymbol{K_{hn}} = \begin{pmatrix} \frac{1}{h} K\left(\frac{(x_{n1} - x_0)}{h}\right) & 0 & \dots & 0 \\ 0 & \frac{1}{h} K\left(\frac{x_{n2} - x_0}{h}\right) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{h} K\left(\frac{x_{nn_n} - x_0}{h}\right) \end{pmatrix}$$

dengan K merupakan fungsi kernel dan h merupakan *bandwidth*. *Bandwidth* h digunakan untuk menentukan efek dari titik data sesuai dengan jarak antara x_{ij} dan x_0.

Penyelesaian persamaan (7) dilakukan dengan penjabaran sebagai berikut :

$$\begin{aligned} \boldsymbol{L} &= (\boldsymbol{Y} - \boldsymbol{X\beta})^T \boldsymbol{K}_h (\boldsymbol{Y} - \boldsymbol{X\beta}) \\ &= \boldsymbol{Y^T K_h Y} - \boldsymbol{Y^T K_h X\beta} - \boldsymbol{\beta^T X^T K_h Y} + \boldsymbol{\beta^T X^T K_h X\beta} \end{aligned}$$

Fungsi $\boldsymbol{L}$ diminimumkan dengan cara diturunkan parsial terhadap matriks $\boldsymbol{\beta}$ kemudian disamadenganka 0 sehingga diperoleh estimasi untuk $\widehat{\boldsymbol{\beta}}$ sebagai berikut:

$$\widehat{\boldsymbol{\beta}} = (\boldsymbol{X^T K_h X})^{-1} \boldsymbol{X^T K_h Y}$$

Estimasi parameter $\boldsymbol{\beta}$ dapat digunakan apabila fungsi **L** telah mencapai minimum dengan menghitung turunan keduanya sebagai berikut:

$$\frac{\partial^2 \boldsymbol{L}}{\partial \widehat{\boldsymbol{\beta}}^2} = \frac{\partial}{\partial \widehat{\boldsymbol{\beta}}} \left(-2\boldsymbol{X^T K_h Y} + 2\boldsymbol{X^T K_h X \widehat{\beta}} \right)$$

$$\frac{\partial^2 \boldsymbol{L}}{\partial \widehat{\boldsymbol{\beta}}^2} = 2\boldsymbol{X^T K_h X} = \mathbf{Z}$$

dengan **Z** matriks semidefinit positif. Suatu matriks simetri **Z** berukuran n x n dengan $\boldsymbol{Z_i}$ adalah submatriks utama yang pertama dari matriks **Z** dengan $i = 1,2, \dots, n$, maka **Z** disebut matriks semidefinit positif jika dan hanya jika $det(\boldsymbol{Z_i}) \geq 0$ (Mulyadi, 2009).

Suatu hal yang perlu dilihat dalam menentukan model polinomial lokal adalah penentuan *bandwidth h*. Penentuan *bandwidth* yang sangat besar menyebabkan plot estimasi model akan terlihat menjauhi plot data asli maka menjadi sangat mulus (*oversmoothing*), bias pemodelan cenderung besar, dan keragamannya cenderung kecil. Penentuan *bandwidth* yang sangat kecil menyebabkan plot estimasi model yang tidak mulus (*undersmoothing*), bias pemodelan cenderung kecil, dan keragamannya cenderung besar. Selain *bandwidth*, penentuan derajat polinomial lokal juga perlu dilihat. Derajat polinomial yang besar maka akan mengurangi bias pemodelannya namun mengakibatkan keragaman yang cenderung besar. Pemilihan *bandwidth* dan derajat polinomial yang optimal sangat diperlukan sehingga bias dan keragaman seimbang supaya didapatkan estimasi yang cukup baik (Suparti dan Prahutama, 2016).

Model terbaik polinomial lokal merupakan model dengan *bandwidth*, titik waktu, bobot fungsi kernel, dan derajat polinomial yang optimal. Penentuan model terbaik pada penelitian ini menggunakan metode *Mean Square Error* (MSE). MSE merupakan nilai rata-rata dari kuadrat residual model yang dirumuskan sebagai berikut:

$$MSE = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n_i}(y_{ij}-\hat{y}_{ij})^2}{\sum_{i=1}^{n} n_i} \qquad (8)$$

Evaluasi ketepatan peramalan model menggunakan ukuran *Mean Absolute Percentage Error* (MAPE) dari data *out sample*. MAPE yaitu persentase rata-rata nilai absolut *error* dibagi data aktualnya yang dirumuskan sebagai berikut (Makridakis *et al*., 1995):

$$MAPE = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n_i}\left|\left(\frac{y_{ij}-\hat{y}_{ij}}{y_{ij}}\right)x100\%\right|}{\sum_{i=1}^{n} n_i} \qquad (9)$$

Nilai MAPE adalah MAPE < 10% artinya model mempunyai kinerja peramalan yang sangat baik, 10% ≤ MAPE < 20% artinya model memiliki kemampuan peramalan yang baik, 20% ≤ MAPE < 50% artinya model memiliki kemampuan peramalan cukup baik, dan MAPE > 50% artinya model memiliki kemampuan peramalan yang buruk (Chang *et al*., 2013).

Software R adalah salah satu bahasa pemrograman yang dapat digunakan untuk melakukan analisis statistik dan analisis grafik. Pengguna *software* R bukan hanya dapat melakukan komputasi statistika saja, tetapi juga dapat memvisualisasikan hasil analisisnya dalam bentuk GUI (Graphical User Interface). Pembuatan *Graphical User Interface* pada penelitian ini menggunakan *package Shiny* program R. *Shiny* terdiri dari tiga komponen yaitu *User Interface* (UI), *Server*, dan *ShinyApp*. Pada bagian *User Interface* berguna untuk panel kontrol, input data, dan penyajian output. *Server* bertugas mendefinisikan program inti dari aplikasi untuk analisis data sesuai pilihan pengguna, *shinyApp* bertugas untuk menjalankan seluruh aplikasi dari komponen *server* dan UI yang telah didefinisikan sebelumnya.

3. METODE PENELITIAN

Penelitian ini menggunakan data harga saham bulanan yang diperoleh dari website resmi *yahoo finance*. Data saham perusahaan PT Tower Bersama Infrastructure Tbk. (TBIG) yang tergolong dalam kelompok kapitalisasi besar dengan nilai kapitalisasi pasar sebesar 72,049 triliun rupiah pada periode 15 Agustus 2022, perusahaan PT XL Axiata Tbk. (EXCL) yang tergolong dalam kelompok kapitalisasi sedang dengan nilai kapitalisasi pasar sebesar 27,613 triliun rupiah pada periode 15 Agustus 2022, dan perusahaan PT Protech Mitra Perkasa Tbk. (OASA) yang tergolong dalam kelompok kapitalisasi kecil dengan nilai kapitalisasi pasar sebesar 207,988 milyar rupiah pada periode 15 Agustus 2022. Data dibagi menjadi dua yaitu data *in sample* dari bulan November 2018 sampai Juni 2021, dan data *out sample* dari bulan Juli 2021 sampai Februari 2022. Variabel yang akan digunakan dalam penelitian ini yaitu data harga saham bulanan periode November 2018 sampai dengan

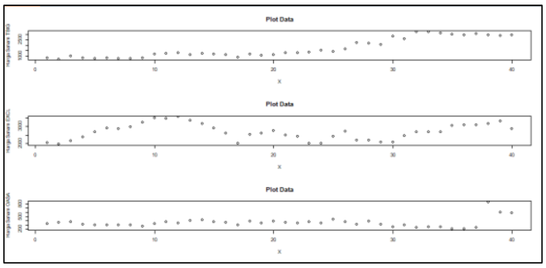
Februari 2022 pada waktu ke-*ij* atau x_{ij} sebagai variabel respon, dan data pada waktu ke-*i*(*j*-*p*) atau $x_{i(j-p)}$ sebagai variabel prediktor berdasarkan lag signifikan.

Penelitian ini menggunakan *software* Rstudio dengan bantuan GUI R dari paket *shiny*. Adapun tahapan analisis data sebagai berikut :

1. Membagi data *in sample* dan *out sample*.
2. Menghitung analisis deskriptif dan eksplorasi data dengan membuat plot data.
3. Menentukan lag maksimum signifikan (p) data *in sample* melalui plot PACF untuk menentukan variabel prediktor pada model nonparametrik polinomial lokal data longitudinal.
4. Membuat plot data $x_{i(j-p)}$ dan x_{ij} yang ditentukan berdasarkan lag signifikan pada PACF.
5. Menghitung nilai MSE pada kombinasi model derajat 1 sampai 5 dengan pembobot fungsi kernel Gaussian, Segitiga, Epanechnikov, dan Biweight, nilai *bandwidth* serta titik lokal yang telah ditentukan.
6. Memilih model polinomial lokal data longitudinal terbaik berdasarkan nilai MSE minimum.
7. Menentukan ketepatan peramalan model terbaik berdasarkan nilai MAPE *out sample*.
8. Menyiapkan aplikasi GUI R.

4. HASIL DAN PEMBAHASAN

Identifikasi awal pemodelan, data *in sample* terlebih dahulu dibuat plot untuk melihat pergerakan data. Selain itu, dilakukan juga pengecekan terhadap lag yang berpengaruh melalui plot PACF. Plot data harga penutupan saham bulanan pada perusahaan Tower Bersama Infrastructure Tbk., XL Axiata Tbk., dan Protech Mitra Perkasa Tbk. dapat dilihat pada Gambar 1.



Gambar 1 Plot Data Harga Saham Bulanan

Gambar 1 memperlihatkan data pada masing-masing subjek mengalami fluktuasi, plot yang terbentuk cenderung menyebar dan tidak membentuk suatu pola. Statistik deskriptif data harga saham bulanan dapat dilihat pada Tabel 1.

Tabel 1 Deskripsi Harga Penutupan Saham Bulanan

Perusahaan	Rata-Rata	Varian	Maksimum	Minimum
Tower Bersama Infrastructure Tbk. (TBIG)	1683,1	770595,7	3210	720
XL Axiata Tbk. (EXCL)	2684,5	210938,2	3550	1980
Protech Mitra Perkasa Tbk. (OASA)	348,9	13282,5	845	199

Plot PACF data *in sample* signifikan pada lag 1 maka pemodelan regresi nonparametrik polinomial lokal menggunakan variabel prediktor $x_{i(j-p)}$ dan respon x_{ij}. Pemilihan model terbaik dilakukan dengan percobaan dari sekumpulan titik lokal (x_0) yang berada pada interval tertentu yang akan digunakan. Diketahui data minimum 0,199 dan data maksimum 3,21 sehingga nilai titik lokal dapat dicobakan pada interval 1 sampai dengan 4 dengan selisih titik lokal 0,1, *bandwidth* dicobakan mulai dari 1 sampai 40 dengan selisih *bandwidth* 1 dan derajat polinomial dicoba derajat 1, 2, 3, 4, dan 5. Sedangkan untuk pembobot kernelnya dicobakan kernel Gaussian, kernel Segitiga, kernel Epanechnikov, dan kernel Biweight. Pemilihan model polinomial lokal terbaik menggunakan nilai *Mean Square Error* (MSE).

Model terbaik yaitu model dengan fungsi pembobot kernel Biweight, derajat polinomial 5, nilai *bandwidth* 37 dan titik lokal 3. Hasil estimasi parameter model dapat dilihat pada Tabel 2.

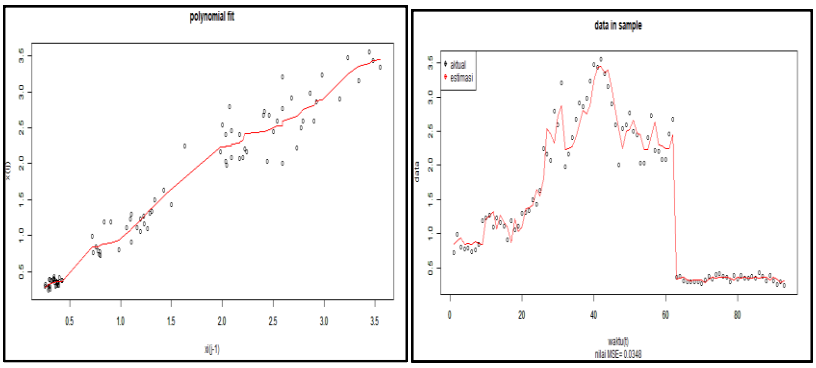
Tabel 2 Estimasi Model Polinomial Lokal Derajat 5

Subjek	Variabel	Hasil Estimasi
1	$\hat{\beta}_0$	1,8445592
	$\hat{\beta}_1$	-6,6626483
	$\hat{\beta}_2$	-14,5474091
	$\hat{\beta}_3$	-12,7471370
	$\hat{\beta}_4$	-5,3063902
	$\hat{\beta}_5$	-0,8406668
2	$\hat{\beta}_0$	2,9141151
	$\hat{\beta}_1$	1,3462149
	$\hat{\beta}_2$	1,0157608
	$\hat{\beta}_3$	-1,0989505
	$\hat{\beta}_4$	-3,5324363
	$\hat{\beta}_5$	-2,0917459
3	$\hat{\beta}_0$	8,0099425
	$\hat{\beta}_1$	-5,6429853
	$\hat{\beta}_2$	-1,6201599
	$\hat{\beta}_3$	8,2582348
	$\hat{\beta}_4$	4,9539248
	$\hat{\beta}_5$	0,7794224

Model polinomial lokal yang terbentuk pada tiap-tiap subjek adalah:

1. Subjek 1:
   $\hat{x}_{1j} = 1{,}8445592 - 6{,}6626483(x_{i(j-1)} - 3) - 14{,}5474091(x_{i(j-1)} - 3)^2 - 12{,}7471370(x_{i(j-1)} - 3)^3 - 5{,}3063902(x_{i(j-1)} - 3)^4 - 0{,}8406668(x_{i(j-1)} - 3)^5$
2. Subjek 2:
   $\hat{x}_{2j} = 2{,}9141151 + 1{,}3462149(x_{i(j-1)} - 3) + 1{,}0157608(x_{i(j-1)} - 3)^2 - 1{,}0989505(x_{i(j-1)} - 3)^3 - 3{,}5324363(x_{i(j-1)} - 3)^4 - 2{,}0917459(x_{i(j-1)} - 3)^5$
3. Subjek 3:
   $\hat{x}_{3j} = 8{,}0099425 - 5{,}6429853(x_{i(j-1)} - 3) - 1{,}6201599(x_{i(j-1)} - 3)^2 + 8{,}2582348(x_{i(j-1)} - 3)^3 + 4{,}9539248(x_{i(j-1)} - 3)^4 + 0{,}7794224(x_{i(j-1)} - 3)^5$

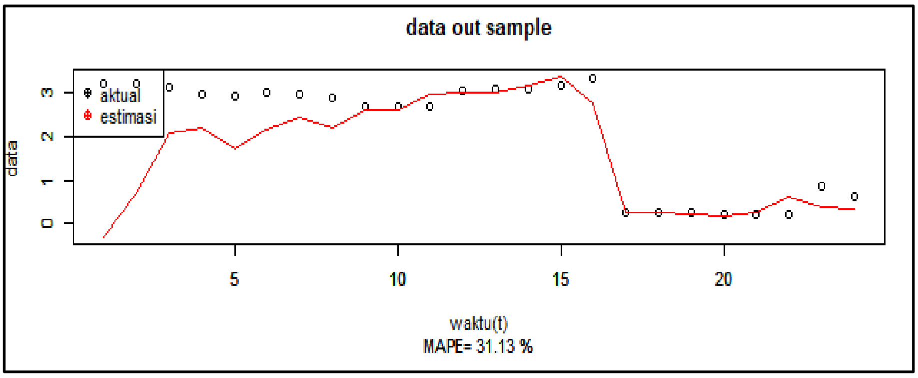
Plot data modifikasi x_{ij} dan $x_{i(j-1)}$ terhadap nilai estimasinya dan plot data asli terhadap estimasinya dapat dilihat pada Gambar 2.



(a) (b)

Gambar 2 (a) Plot Data Modifikasi dan Estimasinya
(b) Plot Data Asli dan Estimasinya

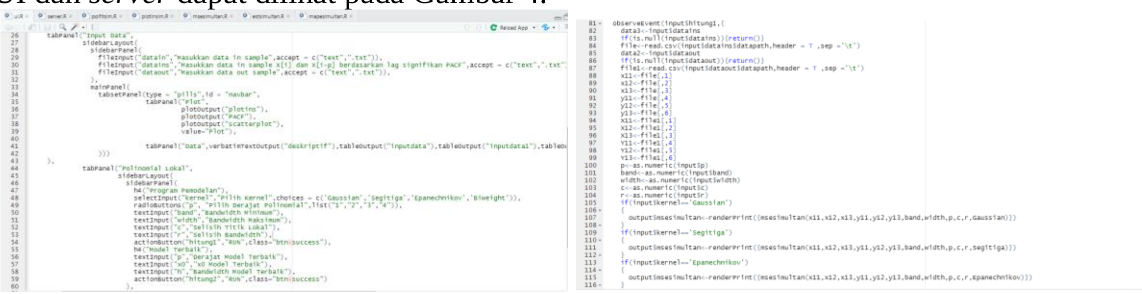
Gambar 2 terlihat bahwa plot data modifikasi (lag) dan data asli dapat didekati oleh data estimasinya dengan baik. Hal tersebut membuktikan bahwa model polinomial lokal dapat digunakan dalam pemodelan data longitudinal yaitu pada data harga saham bulanan pada kelompok saham telekomunikasi.



Gambar 3 Plot Data Aktual *Out Sample* dan Estimasinya

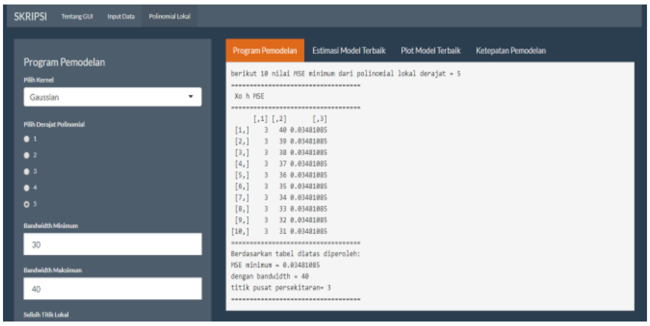
Model polinomial lokal data longitudinal menghasilkan plot estimasi yang baik untuk data *in sample*. Berdasarkan Gambar 3, secara visual plot data aktual *out sample* dapat didekati dengan baik oleh data estimasinya dan nilai MAPE *out sample* sebesar 31,13%. Nilai MAPE diantara 20% sampai 50% menunjukkan bahwa model memiliki kemampuan peramalan yang cukup baik. Hal tersebut membuktikan bahwa model polinomial lokal derajat 5 memiliki kemampuan yang cukup baik secara *in sample* dan *out sample*.

Penyusunan GUI R meliputi bagian *user interface* yang mengatur input dan output data menggunakan perintah ui<-fluidPage(). Tiap input yang dilakukan menggunakan bantuan perintah-perintah pada UI harus diberi identitas. Proses pengolahan yang dijalankan oleh GUI akan didefinisikan dalam fungsi *shinyApp* yaitu pada *server*. Sintaks tampilan program UI dan *server* dapat dilihat pada Gambar 4.



Gambar 4 Tampilan Program *UI* dan *Server*

Setelah objek UI dan *server* selesai dibentuk maka langkah selanjutnya yaitu *running* aplikasi dengan fungsi shinyApp(ui,server) dan aplikasi dapat digunakan. Tampilan GUI untuk pencarian model polinomial lokal terbaik dapat dilihat pada Gambar 5.



Gambar 5 Tampilan GUI R Polinomial Lokal

5. KESIMPULAN

Model nonparametrik polinomial lokal terbaik yaitu model derajat 5 dengan pembobot fungsi kernel Biweight, *bandwidth* 37, dan titik lokal 3 serta nilai MSE sebesar 0,03481085. Dalam menentukan model terbaik, pemilihan *bandwidth* yang tepat lebih baik dibandingkan pemilihan fungsi kernel. Pada evaluasi ketepatan pemodelan polinomial lokal terbaik menggunakan nilai MAPE *out sample* yang dihitung berdasarkan data aktual *out sample* dan data estimasinya menghasilkan nilai MAPE sebesar 31,13% yang mana termasuk dalam kategori model yang memiliki kemampuan peramalan yang cukup baik. Pembuatan GUI R pada penelitian ini merupakan pengembangan penelitian sebelumnya untuk metode nonparametrik polinomial lokal data longitudinal. Pembuatan GUI R dapat memudahkan dalam proses *input* data dan *running* program serta menghasilkan *output* yang lebih menarik daripada analisis berbasis *Command Line Interface* (CLI) dan memudahkan dalam proses analisis data.

DAFTAR PUSTAKA

Budiantara, I. N., 2009. *Spline dalam Regresi Nonparametrik dan Semiparametrik: Sebuah Pemodelan Statistika Masa Kini dan Masa Mendatang.* Surabaya: Institut Teknologi Sepuluh November (ITS).

Chang, P.-C., Wang, Y.-W., & Liu, C.-H. 2007. *The Development of a Weighted Evolving Fuzzy Neural Network for PCB Sales Forecasting*. Expert Systems with Application, *32*(88–89).

Eubank, R. L., 1999. *Nonparametric Regression and Spline Smoothing.* Texas: Departement of Statistics Southern Methodist Dallas University.

Fan, J. dan Gijbels, I., 1997. *Local Polynomial Modelling and Its Apllications.* London: Chapman and Hall.

Gujarati, D., 2003. *Basic Econometrics.* New York: Mc.Graw-Hill.

Hadi, N., 2013. *Pasar Modal: Acuan Teoritis dan Praktis Investasi di Keuangan Pasar Modal.* Yogyakarta: Graha Ilmu.

Khalid, I., 2015. PEMODELAN REGRESI NONPARAMETRIK DATA LONGITUDINAL MENGGUNAKAN POLINOMIAL LOKAL (Studi Kasus: Harga Penutupan Saham pada Kelompok Harga Saham Periode Januari 2012 – April 2015). *Gaussian*, 4(3), pp. 527-532.

Makridakis, Wheelwright & Mcgee, 1995. *Metode dan Aplikasi Peramalan.* Jakarta: Erlangga.

Mulyadi, S. (2009). *Matriks Kuasidefinit*. Institut Pertanian Bogor.
Ogden, R. T., 1997. *Essential Wavelets for Statistical Applications and Data Analysis.* Boston: Birkhauser.
Purnomo, D., S., Serfiyani, C. Y. & Haryani, I., 2013. *Buku Pintar: Pasar Uang dan Pasar Valas.* Jakarta: Kompas Gramedia.
Suparti, dan Prahutama, A. 2016. P*emodelan Regresi Nonparametrik Menggunakan Pendekatan Polinomial Lokal Pada Beban Listrik Di Kota Semarang*. Media Statistika, *9*(2), 85–93.
Widarjono, A., 2010. *Analisis Statistika Multivariate Terapan.* Yogyakarta: UPP STIM YKPN.
Wu, H. dan Zhang, J.-T., 2006. *Nonparametrik Regression Methods for Longitudinal Data Analysis.* New York: John Wiley and Sons,Inc.