ANALISIS DATA INFLASI INDONESIA MENGGUNAKAN METODE FOURIER DAN WAVELET *MULTISCALE AUTOREGRESIVE*

Suparti[1], Rukun Santoso[1], Alan Prahutama[1], Hasbi Yasin[1], Alvita Rachma Devi[2]
[1]*Statistics Department, Diponegoro University, Semarang Indonesia*
[2]*Independent Researcher*

Email : suparti702@gmail.com

Abstrak. Analisis regresi merupakan metode statistika untuk mengetahui hubungan antara variabel prediktor dan variabel respon. Pendekatan regresi dapat dilakukan dengan pendekatan parametrik dan nonparametrik. Pendekatan parametrik ketat dengan asumsi dan harus dipenuhi untuk mendapatkan model yang baik. Sementara pendekatan nonparametrik tidak ketat dengan asumsi karena metode tersebut didasarkan pada pendekatan kurva yang tidak diketahui bentuknya. Pendekatan nonparametrik dapat dilakukan dengan beberapa pendekatan diantaranya metode Fourier dan Wavelet. Metode Fourier merupakan metode yang didasarkan pada deret cosinus atau sinus. Metode Fourier sangat sesuai untuk data yang mengalami pola berulang atau stasioner. Sedangkan pada pemodelan wavelet tidak hanya terbatas pada data berulang atau stasioner saja, akan tetapi juga mampu memodelkan data yang tidak stasioner. Pada penelitian ini dimodelkan nilai Inflasi di Indonesia dari Januari 2007 sampai Agustus 2017. Variabel responnya adalah nilai inflasi, sedangkan variabel prediktornya adalah waktu. Metode Fourier dengan K=100 menghasilkan MSE sebesar 0,846216 dan R^2 sebesar 80,12%. Model Wavelet menggunakan Multiscale Autoregresive dengan filter Haar, J=4 dan A_j = 2 mempunyai MSE sebesar 0,312 dengan R^2 sebesar 96,91%. Pada model Fourier dengan K=100 diperlukan parameter sebanyak 102 buah sedangkan model wavelet dengan J=4 dan A_j = 2 hanya diperlukan parameter sebanyak 10 buah. Jadi model wavelet sangat efisien dengan kinerja yang lebih bagus dibandingkan dengan model Fourier.

Kata Kunci: Inflasi, nonparametrik, Fourier, Wavelet, MSE

1. Pendahuluan

Analisis regresi merupakan metode statistika untuk memodelkan hubungan antara variabel prediktor dengan variabel respon. Pendekatan pemodelan regresi dapat dilakukan dengan tiga pendekatan antara lain : parametrik, nonparametrik dan semiparametrik. Model regersi parametrik merupakan model regresi dengan bentuk kurva diketahui. Salah satu model regresi parametrik adalah regresi linier. Dalam model regresi linier ada asumsi model yang harus dipenuhi yaitu residual harus independen, berdistribusi normal dengan mean 0 dan varian konstan. Sedangkan model regresi nonparametrik merupakan model regresi dengan bentuk kurva diketahui dan model regresi semiparametrik merupakan gabungan dari model regresi parametrik dan nonparametrik.

Beberapa metode terkait dengan model regresi nonparametrik antara lain model regresi kernel, spline, polinomial lokal, Fourier dan wavelet. Suparti (2005) telah melakukan penelaahan tentang pemodelan regresi nonparametrik menggunakan model Fourier dan wavelet berdasarkan nilai IMSE (Integrated Mean Square Error). Diperoleh hasil bahwa model regresi wavelet mempunyai konvergensi IMSE lebih cepat dari model Fourier. Telaah yang telah dilakukan oleh Suparti (2005), pemodelan wavelet yang digunakan adalah model wavelet linier dengan menggunakan analisis multiresolusi dari transformasi wavelet diskrit (Discrete Wavelet Transform/DWT). Dalam tansformasi wavelet diskrit ada keterbatasan dalam hal banyaknya data yang diolah. DWT membatasi banyaknya data yang diolah harus sebanyak 2^J dengan J interger. Untuk mengatasi keterbatasan DWT digunakan MODWT (Maximal Ovelaping Discrete Wavelet Transform) yang mampu diterapkan dalam hal banyaknya data tidak berupa 2^J.

Data inflasi merupakan data runtun waktu yang berfluktuasi naik turun tidak membentuk pola tertentu. Telah dilakukan pemodelan parametrik data inflasi Indonesia menggunakan data inflasi tahunan bulan Desember 2006 – Desember 2011, tak ada model Box Jenkins (baik AR, MA maupun ARIMA) yang sesuai karena asumsi independensi error tidak dipenuhi (Suparti, 2013). Karena data inflasi tidak membentuk pola tertentu, maka akan dilakukan pemodelan nonparametrik. Pemodelan data inflasi menggunakan model nonparametrik kernel, spline dan polinomial lokal telah dibahas dalam Suparti (2013a), Suparti (2013b) Suparti,dkk (2014)). Dalam pemodelan nonparametrik, estimasi model akan menyesuaikan sendiri pola datanya.

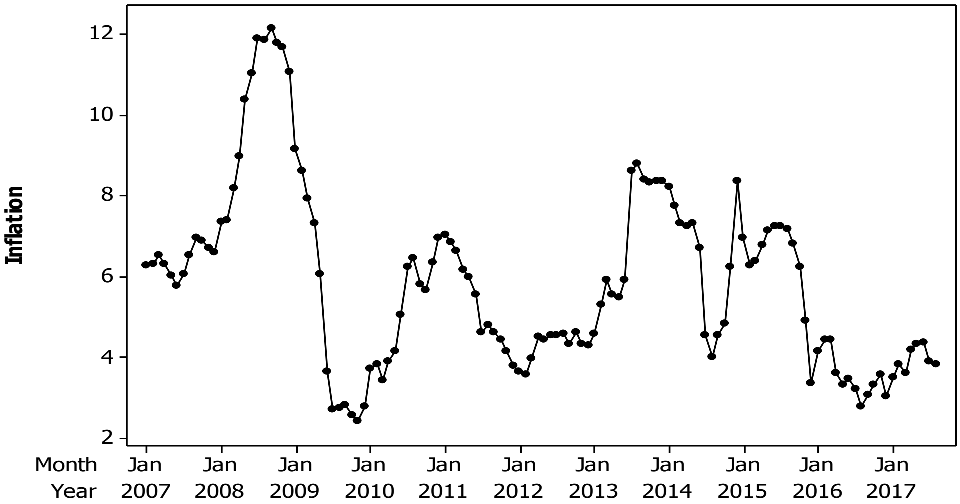
Dalam paper ini dibahas model nonparametrik menggunakan model Fourier dan model wavelet runtun waktu dengan MODWT. Model wavelet runtun waktu mengadopsi model runtun waktu Autoregresive (AR(p)) dengan variabel bebasnyaya merupakan variabel lag dari koefisien MODWT dan modelnya dikenal dengan model *Multiscale Autoregresive* (Renaud et al, 2003). Performa model akan diukur menggunakan ukuran MSE dan R^2. Model yang baik akan memberikan nilai MSE yang kecil dan R^2 yang besar.

2. Metode Penelitian

Data penelitian yang digunakan merupakan data sekunder yang diambil dari situs resmi Bank Indonesia dengan variabel penelitian adalah nilai inflasi tahunan (yoy) Indonesia pada bulan Januari 2007 – Agustus 2017. Pada pemodelan Fourier, variabel responnya adalah nilai inflasi dan variabel prediktornya adalah waktu. Sedangkan pada model wavelet MAR, variabel responnya adalah nilai inflasi dan variabel prediktornya adalah lag dari koefisien MODWT. Dalam pemodelan Fourier dicobakan beberapa nilai K mulai dari 1 sampai 100 dan dihitung nilai MSE dan R^2 untuk setiap K dengan K adalah banyaknya koefisien fungsi cosinus dalam model. Sedangkan dalam model wavelet *multiscale autoregresive* digunakan J=4 danAj=2 dengan filter wavelet Haar (D2) dan dihitung MSE dan R^2 . Kita lakukan perbandingan dengan melihat nilai MSE. Model terbaiknya adalah model dengan MSE terkecil dan R^2 terbesar.

3. Hasil dan Pembahasan

Data inflasi tahunan (yoy) Indonesia pada bulan Januari 2007 – Agustus 2017 disajikan pada Gambar 1. Pola data mengalami fluktuatif naik turun dan berulang pada suatu interval. Nilai inflasi tertinggi mencapai 12,14 terjadi pada bulan september 2008. Sedangkan nilai inflasi terendah mencapai 2,41 terjadi pada November 2009. Nilai inflasi yang naik turun sangat fluktuatif menyebabkan ketidakstabilan perekonomian. Rata-rata inflasi mencapai 5,84 dengan standar deviasi 2,23.



Gambar 1. Scatter Plot Data Inflasi Indonesia

3.1 Pemodelan Regresi Fourier

Deret Fourier merupakan deret kombinasi fungsi sinus dan cosinus (Odgen,1997). Deret Fourier sangat sesuai dengan pola data yang mengalami kejadian berulang-ulang atau periodik. Dalam regresi Fourier selain meggunakan kombinasi aditif fungsi sinus dan cosinus, dapat juga digunakan kombinasi aditif fungsi linier dan fungsi sinus atau cosinus. Bilodeau (1992) memberikan pemodelan regresi nonparametrik dengan pendekatan kombinasi fungsi linier dan fungsi cosinus. Kombinasi ini diharapkan dapat memisahkan antara tren data dan fluktuasi data. Dengan mengambil kombinasi fungsi linier dan salah satu fungsi sinus atau cosinus akan memberikan efisiensi estimasi. Bilodeau (1992) memberikan pendekatan kombinasi aditif fungsi linier dan cosinus sebagai berikut:

$$f(t) = \frac{1}{2}\alpha_0 + \gamma t + \sum_{k=1}^{K} \alpha_k \cos kt \qquad (1)$$

Model regresi nonparametrik secara umum adalah

$$y = f(t) + \varepsilon \qquad (2)$$

Jika dipunyai data sebanyak n yaitu (t_i,y_i), i=1,2,...,n dengan f(t) merupakan kurva yang tidak diketahui bentuknya dan f(t) didekati dengan menggunakan deret Fourier (1) maka

$$f(t_i) = \frac{1}{2}\alpha_0 + \gamma t_i + (\alpha_{i1} \cos t_i + \alpha_{i2} \cos 2t_i + \ldots + \alpha_{iK} \cos Kt_i) \,.$$

Dalam bentuk matriks model (2) dapat ditulis

$$\begin{aligned} \mathbf{Y} &= \mathbf{f(t)} + \boldsymbol{\varepsilon} \\ &= \mathbf{A}\boldsymbol{\theta} + \boldsymbol{\varepsilon} \qquad (3) \end{aligned}$$

Dengan

Y= $[y_1, y_2, \ldots, y_n]^T$, $\boldsymbol{\theta} = [\phi \quad \gamma \quad \alpha_{11} \quad \alpha_{12} \quad \cdots \quad \alpha_{1K}]^T$ dengan $\emptyset = \frac{n}{2}\alpha_0$, $\varepsilon = [\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n]^T$ dan

$$\mathbf{A} = \begin{bmatrix} 1 & t_1 & \cos t_1 & \cos 2t_1 & \cdots & \cos K t_1 \\ 1 & t_2 & \cos t_2 & \cos 2t_2 & \cdots & \cos K t_2 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & t_n & \cos t_n & \cos 2t_n & \cdots & \cos K t_{n1} \end{bmatrix}.$$

Estimasi model regresi nonparametrik (3) dengan menggunakan OLS (*Ordinary Least Square*) diperoleh $\hat{\boldsymbol{\theta}} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{y}$ dan $\hat{Y} = A\hat{\theta}$ (4)

Dalam penentuan model Fourier,performa model didasarkan hanya pada banyaknya jumlah K dengan K merupakan banyaknya fungsi cosinus yang terbentuk. Pemodelan data Gambar 1 menggunakan model Fourier (1) dengan beberapa nilai K disajikan pada Gambar 2. Gambar 2 dapat dsimpulkan bahwa semakin banyak nilai K, estimasi modelnya semakin mendekati pola datanya. Semakin banyak nilai K menunjukkan semakin banyak nilai parameter yang harus diestimasi sehingga modelnya semakin kompleks. Pada dasarnya, akan dicari model regresi Fourier yang parsimoni dan menghasilkan nilai akurasi yang bagus. Model parsimoni artinya model yang sederhana, tidak kompleks. Nilai akurasi yang digunakan adalah MSE dan R^2.

Berdasarkan Gambar 2, nilai estimasi yang mendekati data aktual adalah untuk K=100, dengan parameter yang dibutuhkan sebanyak 102. Sementara untuk nilai K=1 tidak menghasilkan estimasi model yang bagus. Telah dicobakan untuk nilai K berkisar 2, 4, dan 5 akan tetapi hasil estimasinya tidak bagus, seperti menggunakan nilai K=1. Kurva estimasinya disajikan dalam Gambar 2 dan nilai MSE dan R^2 disajikan dalam Tabel 1.

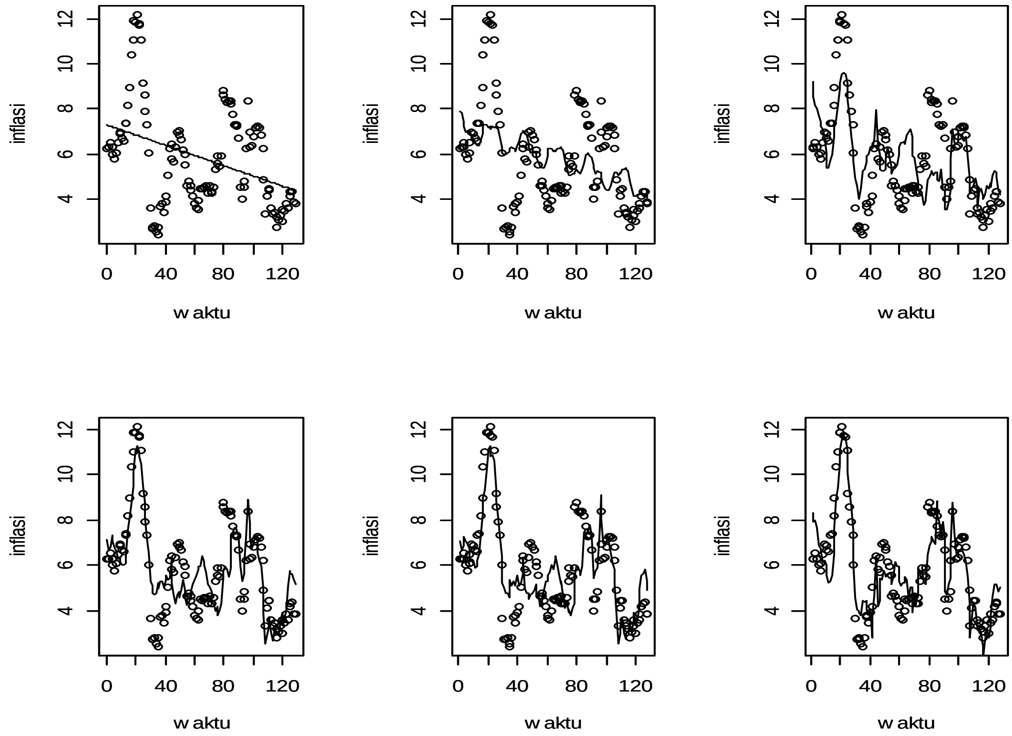
Untuk nilai K=1 sampai K= 60, R^2 yang dihasilkan tidak begitu bagus. Untuk nilai K=70, R^2 yang dihasilkan diatas 50%, akan tetapi untuk nilai K=70 sampai K=80 peningkatan nilai R^2 tidak begitu signifikan. Untuk nilai K=90 sampai K=100 terjadi peningkatan nilai K yang cukup signifikan.

Misal diambil nilai K=5 maka estimasi model regresi Fourier adalah sebagai berikut:

$$\hat{y} = \frac{7{,}3}{2} - 0{,}023t - 0{,}0117\cos t - 0{,}01\cos 2t + 0035\cos 3t - 0{,}013\cos 4t - 0{,}068\cos 5t$$

Tabel 1. Nilai MSE dan R^2 Model Regrsi Fourier

Nilai K	R^2	MSE
K=1	0,1431	4,2375
K=5	0,1437	4,2344
K=10	0,1774	4,0678
K=40	0,3252	3,3369
K=60	0,4058	2,9378
K=70	0,6601	1,6780
K=80	0,6641	1,6607
K=90	0,6985	1,4900
K=100	0,8012	0,9831



Gambar 2. Scatter plot dan estimasi deret Fourier untuk nilai K= 1,10,50,70,80,100

Keterangan : oooooo : data —— : kurva estimasi model

3.2 Pemodelan Wavelet *Multiscale Autoregresive*

Secara umum pemodelan wavelet *Multiscale Autoregressive* adalah metode pemodelan menggunakan transformasi wavelet, dalam hal ini menggunakan MODWT. Dengan adanya *multiscale decomposition* seperti wavelet, ada manfaat yang didapatkan yaitu secara otomatis memisahkan komponen-komponen data, seperti komponen trend dan komponen irregular pada data. Oleh karena itu, metode ini dapat digunakan untuk melakukan prediksi pada data stasioner maupun yang tidak stasioner (Suhartono,dkk).

Misalkan ada suatu sinyal stasioner $y = (y_1, y_2, \dots, y_t)$ dan diasumsikan akan dilakukan prediksi nilai y_{t+1}. Ide dasar yang digunakan adalah menggunakan koefisien-koefisien yang didapat dari hasil dekomposisi MODWT yaitu $w_{j,t-2^j(k-1)}$ dan $v_{J,t-2^J(k-1)}$, dengan $k = 1,2, \ldots, A_j$ dan $j = 1,2, \ldots, J$ (Renaud et al, 2003). Model prediksinya mengacu pada model *Autoregresive* (AR (p)), yaitu $\hat{y}_{t+1} = \sum_{k=1}^{p} \hat{\phi}_k \, y_{t-(k-1)}$.

Dengan menggantikan variabel independen y_t, y_{t-1},... $y_{t-(p-1)}$ dengan koefisien dari dekomposisi wavelet, Renaud *et al.* (2003) memberikan model prediksi AR menjadi model *Multiscale Autoregressive* (MAR) :

$$\hat{y}_{t+1} = \sum_{j=1}^{J} \sum_{k=1}^{A_j} \hat{a}_{j,k} \, w_{j,t-2^j(k-1)} + \sum_{k=1}^{A_j} \hat{a}_{J+1,k} \, v_{J,t-2^J(k-1)} \qquad (5)$$

dengan:

$a_{j,k}$: koefisien MAR (j=1,2,...,J dan k=1,2,...,A_j)

A_j : orde dari model MAR

$w_{j,t}$: koefisien wavelet dari data

$v_{j,t}$: koefisien skala dari data

Model MAR memiliki bentuk yang mirip dengan model regresi berganda, persamaan (5) dapat ditulis menjadi:

$$y_{t+1} = \sum_{j=1}^{J} \sum_{k=1}^{A_j} a_{j,k} \, w_{j,t-2^j(k-1)} + \sum_{k=1}^{A_j} a_{J+1k,} \, v_{J,t-2^J(k-1)} + \varepsilon_{t+1} \qquad (6)$$

Sebagai contoh, misalkan dipunyai data y=(y_1,y_2,...,y_{70}) dengan menggunakan J=4 dan A_j=2 ($k = 1,2$), rumus MAR menjadi :

$$y_{t+1} = \sum_{j=1}^{4} \sum_{k=1}^{2} a_{j,k} \, w_{j,t-2^j(k-1)} + \sum_{k=1}^{2} a_{J+1k,} \, v_{J,t-2^J(k-1)} + \varepsilon_{t+1}$$

$$y_{t+1} = a_{1,1}w_{1,t} + a_{1,2}w_{1,t-2} + a_{2,1}w_{2,t} + a_{2,2}w_{2,t-4} + a_{3,1}w_{3,t} + a_{3,1}w_{3,t-8} + a_{4,1}w_{4,t} + a_{4,1}w_{4,t-16} + a_{5,1}v_{4,t} + a_{5,2}v_{4,t-16} + \varepsilon_{t+1} \qquad (7)$$

Dalam bentuk matriks dapat ditulis

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_t \\ \vdots \\ y_{69} \\ y_{70} \end{bmatrix} = \begin{bmatrix} w_{1,0} & w_{1,-2} & w_{2,0} & w_{2,-4} & w_{3,0} & w_{3,-8} & w_{4,0} & w_{4,-16} & v_{4,0} & v_{4,-16} \\ w_{1,1} & w_{1,-1} & w_{2,1} & w_{2,-3} & w_{3,1} & w_{3,-7} & w_{4,1} & w_{4,-15} & v_{4,1} & v_{4,-15} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ w_{1,t-1} & w_{1,t-3} & w_{2,t-1} & w_{2,t-5} & w_{3,t-1} & w_{3,t-9} & w_{4,t-1} & w_{4,t-17} & v_{4,t-1} & v_{4,t-17} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ w_{1,68} & w_{1,66} & w_{2,68} & w_{2,64} & w_{3,68} & w_{3,60} & w_{4,68} & w_{4,52} & v_{4,68} & v_{4,52} \\ w_{1,69} & w_{1,67} & w_{2,69} & w_{2,65} & w_{3,69} & w_{3,61} & w_{4,69} & w_{4,53} & v_{4,69} & v_{4,53} \end{bmatrix} \begin{bmatrix} a_{1,1} \\ a_{1,2} \\ \vdots \\ a_{j,k} \\ \vdots \\ a_{5,1} \\ a_{5,2} \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_t \\ \vdots \\ \varepsilon_{69} \\ \varepsilon_{70} \end{bmatrix}$$

Matriks tersebut dapat ditulis menjadi:

$$\boldsymbol{s_1} = A_1\boldsymbol{\alpha} + \boldsymbol{\varepsilon_1} \quad (8)$$

dengan:

$\boldsymbol{s_1}$: vektor data runtun waktu berukuran 70 x 1

A_1 : matriks koefisien-koefisien wavelet berukuran 70 x 10

$\boldsymbol{\alpha}$: vektor parameter yang ditaksir berukuran 10 x 1

$\boldsymbol{\varepsilon_1}$: vektor *error* berukuran 70 x 1

Koefisien-koefisien dalam matrik A_1, ada yang berindeks negatif, nol dan positif. Koefisien dengan indeks nol dan negatif tidak terdapat pada hasil dekomposisi dengan wavelet. Pembentukan model MAR dilakukan dengan tidak menyertakan koefisien berindeks nol dan negatif, sehingga vektor $\boldsymbol{s}, \boldsymbol{\varepsilon}$ dan matriks A dimulai dari baris ke-18 dapat dimisalkan sebagai

$$\boldsymbol{s} = A\boldsymbol{\alpha} + \boldsymbol{\varepsilon} \quad (9)$$

dengan:

$\boldsymbol{s}$: vektor data runtun waktuberukuran 53 x 1

$\mathbf{A}$: matriks berisikan koefisien-koefisien wavelet berukuran 53 x 10

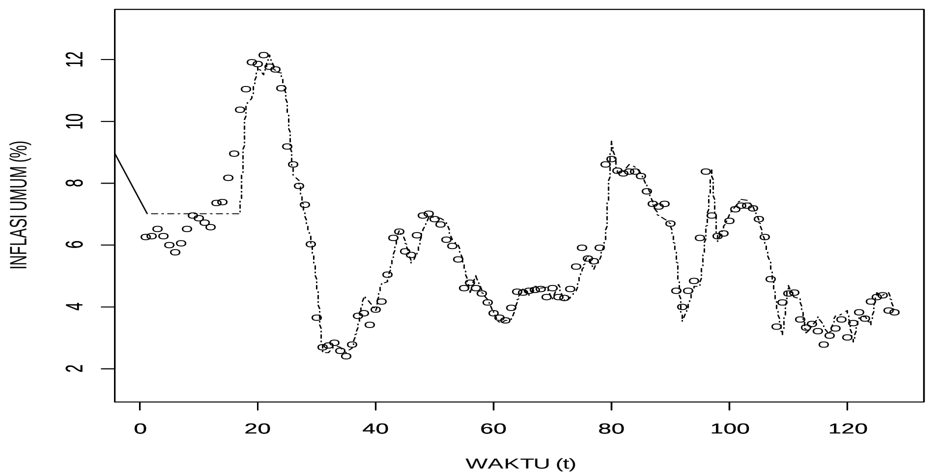
$\boldsymbol{\alpha}$: vektor dari parameter yang ditaksir berukuran 10 x 1

$\boldsymbol{\varepsilon}$ **:** vektor error berukuran 53 x 1

Untuk menaksir parameter pada model MAR dapat menggunakan metode kuadrat terkecil (*Ordinary least square*) dan dihasilkan $\hat{\boldsymbol{\alpha}} = (A^T A)^{-1} A^T \boldsymbol{s}$ dan $\hat{\boldsymbol{s}} = A\hat{\boldsymbol{\alpha}}$.

Pemodelan data Gambar 1 menggunakan Model *Multiscale Autoregressive* (MAR) dengan Filter Haar (D2) dengan J=4 dan Aj=2 dengan substitusi persamaan (7), $x_1 = w_{1,t}$, $x_2 = w_{1,t-2}$, $x_3 = w_{2,t}$, $x_4 = w_{2,t-4}$, $x_5 = w_{3,t}$, $x_6 = w_{3,t-8}$, $x_7 = w_{4,t}$, $x_8 = w_{4,t-16}$, $x_9 = v_{4,t}$ dan $x_{10} = v_{4,t-16}$
diperoleh model pada Gambar 3 dengan nilai mse = 0,312 dan R^2 = 0,9916 dengan variabel yang signifikan x_1, x_3, x_6, x_7, x_8, dan x_9.

Dengan membandingkan antara model Fourier dengan K=100 dan model wavelet MAR dengan J=4 dan j= 2 pada data inflasi diIndonesia diperoleh hasil yang terangkum dalam Tabel 2. Terlihat bahwa model wavelet multiscale autoregresive lebih efisien dari model fourier dan memberikan performa yang lebih baik karena MSE lebih kecil dan R^2 lebih besar.



Gambar 3. Model MAR dengan filter Haar (D2)

Tabel 2. Pebandingan model Fourier K model wavelet MAR

	Model Fourier denganK =100	Model wavelet MAR Dengan J=4 dan Aj = 2
Banyaknya paremeter	102	10
MSE	0,9831	0,3120
R^2	0,8012	0,9916

4. Kesimpulan

Dari analisa dan pembahasan yang telah dipaparkan pada bagian 3 memberikan kesimpulan bahwa model wavelet *multiscale autoregresive* lebih baik dari model Fourier. Model wavelet lebih efisien dari metode Fourier karena untuk membangun model dengan performa yang baik, parameter dalam model wavelet lebih sedikit dari model Fourier.

5. Ucapan Terima Kasih

Terimakasih kepada Kemenristek Dikti untuk pendanaan penelitian ini dengan skema PTUPT (Penelitian Terapan Perguruan Tinggi) untuk tahun 2018 dengan No kontrak 101-124/UN7.P4.3/PP/2018 Tanggal 5 Februari 2018 dan Departemen Statistika FSM Undip Semarang.

Daftar Pustaka

Bilodeau, M. (1992). Fourier smoother and additive models. *The Canadian Journal of Statistics, Vol. 20 (3)* , 257-269.

Odgen, R. T. (1997). *Essential Wavelets For Statistical Applications and Data Analysis.* Boston: Birkhauser.

Renaud, O., Starck, J., & Murtagh, F. (2003). Prediction based on a multiscale Decomposition. *International Journal of Wavelets, Multiresolution and Information Processing 1(2)* , 217-232.

Suhartono, Ulama, B. S.S., & Endharta, A. (2010). Seasonal Time Series Data Forecasting by Using Neural Network Multiscale Autoregresive Model. *American Journalof Applied Sciences 7(10)* , 1373-1378.

Suparti. (2013a). Analisis Data Inflasi di Indonesia Menggunakan Model Regresi Spline. *Media Statistika, Vol.6, No.1* , 1-9.

Suparti. (2013b). Analisis Data Inflasi Di Indonesia Pasca Kenaikan TDL dan BBM Tahun 2013 Menggunakan Model Regresi Kernel. *Media Statistika Vol 6 No 2* , 91-101.

Suparti. (2005). Perbandingan Estimator Regresi Nonparametrik Menggunakan Metode Fourier dan Metode Wavelet. *Jurnal Matematika, Vol.8, No.3* , 88-94.

Suparti, Warsito, B., & Mukid, M. A. (2014). Analisis Data Inflasi di Indonesia Menggunakan Model Regresi Polinomial Lokal. *IndoMS Journal on Statistics Vol 2 No 1* , 65-78.