

PEMILIHAN INPUT MODEL ANFIS UNTUK DATA RUNTUN WAKTU MENGUNAKAN METODE FORWARD SELECTION DILENGKAPI GUI MATLAB

(Studi Kasus: Jumlah Penumpang Kereta Api di Wilayah Jawa Non Jabodetabek)

Tiara Sukma Valentina¹, Tarno², Alan Prahutama³

^{1,2,3} Departemen Statistika FSM Universitas Diponegoro

Tarno.stat@gmail.com

Abstract

One of the methods that is commonly used to identify a time series model and input ANFIS (*Adaptive Neuro Fuzzy Inference System*) model is PACF plot. The PACF plot shows the correlation between current observations and previous observations visually. Formally there are several methods that are known to effectively identify ANFIS inputs, one of which is the *Forward Selection* regression method. With the same concept as PACF, the process of selecting ANFIS inputs using the *Forward Selection* method is based on the order of the correlation between the predictors of the response which is indicated by the magnitude of the correlation coefficient. This study discusses the *Forward Selection* method in simulation data that has stationary characteristics, stationary with outliers, non stationary, non stationary with outliers and implements data on the number of train passengers in the Non Jabodetabek Java region. ANFIS modeling on data of the number of train passengers in the Non Jabodetabek Java region produces AIC of 15,5617, MAPE of 8,5093% and RMSE of 571,33691. The result of this study is equipped with a GUI which is useful as a tool to facilitate users in processing data.

Keywords : PACF Plot, *Forward Selection*, ANFIS, non stasionary, outlier

1. PENDAHULUAN

Pesatnya pemanfaatan program komputasi di era global berdampak pada analisis statistika dalam menggunakan *soft computing* sebagai metode pemecahan masalah data runtun waktu. Kesulitan yang dijumpai dalam menentukan aturan yang akan dimasukkan ke dalam basis kaidah sistem *fuzzy* dan kesulitan dalam merancang seberapa banyak lapisan dalam yang akan digunakan dalam struktur *neural network*, dapat di atasi dengan menggabungkan kedua sistem menjadi sistem *neuro-fuzzy* struktur ANFIS (*Adaptive Neuro-Fuzzy Inference System*) (Fariza *et al.*, 2007). Struktur jaringan ANFIS terdiri dari lima lapisan, mulai dari menentukan tingkat dan derajat keanggotaan sampai dengan menghasilkan estimasi output dan input adalah komponen penting yang perlu dipertimbangkan sebelumnya.

Pada sebagian pemodelan, pemilihan input ANFIS dilakukan dengan beberapa cara diantaranya secara random, berdasarkan model AR (*Autoregressive*) yang secara visual dapat dilihat dari plot PACF maupun pendekatan fungsi transfer. Metode *Forward Selection* adalah metode dimana peubah bebas dimasukkan menurut urutan besar pengaruhnya terhadap model dan berhenti bila semua yang memenuhi syarat telah masuk (Sembiring, 2014). Hal ini menyiratkan persamaan konsep antara PACF (*Partial Autocorrelation Function*) dan *Forward Selection* dalam menggunakan hubungan yang erat untuk melihat variabel yang berpengaruh.

Dalam penulisan ini akan difokuskan pada penggunaan metode *Forward Selection* sebagai pemilihan input model ANFIS yang diuji cobakan pada data runtun waktu yang memiliki karakteristik stasioner, stasioner dengan *outlier*, nonstasioner dan nonstasioner dengan *outlier* serta mengimplementasikan pada data jumlah penumpang kereta api di wilayah Jawa Non Jabodetabek yang dilengkapi dengan GUI Matlab.

2. TINJAUAN PUSTAKA

2.1 Analisis Regresi

Model regresi umum yang mengandung p variabel prediktor untuk i respon dituliskan dalam persamaan sebagai berikut:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_p X_{ip} + \varepsilon_i,$$

dengan Y_i adalah variabel respon pada pengamatan ke $-i$, X_{ik} adalah variabel prediktor $-k$ pada pengamatan ke $-i$ dan ε_i merupakan error atau galat pada pengamatan ke $-i$ dengan $i = 1, 2, \dots, n$ dan $k = 1, 2, \dots, p$ dimana n adalah banyaknya data.

Sebuah variasi (penyimpangan) dituliskan dalam persamaan sebagai berikut:

$$\begin{aligned} \sum_{i=1}^n (Y_i - \bar{Y})^2 &= \sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2 + \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \\ \text{JKT} &= \text{JKR} + \text{JKS} \end{aligned}$$

2.2 Forward Selection

Samosir *et al.* (2014) menjelaskan langkah-langkah dalam metode *Forward Selection* adalah sebagai berikut:

- a. Membentuk matriks koefisien korelasi Y dengan X_k dengan rumus:

$$r = \frac{\sum (X_{ik} - \bar{X}_k)(Y_i - \bar{Y})}{\sqrt{\sum (X_{ik} - \bar{X}_k)^2 \sum (Y_i - \bar{Y})^2}}$$

- b. Membentuk regresi pertama berdasarkan nilai mutlak koefisien korelasi terbesar antara Y dan X_k . Misal terpilih X_h , dilakukan uji keberartian regresi berupa uji hipotesis:

$H_0 : b_h = 0$ (X_h tidak mempunyai pengaruh yang berarti terhadap Y)

$H_1 : b_h \neq 0$ (X_h mempunyai pengaruh yang berarti terhadap Y)

Tabel 1. Analisa Variansi untuk Uji Keberartian Regresi

Sumber	Derajat Bebas	Jumlah kuadrat	Rataan Kuadrat	F_{hitung}
Regresi	$p-1$	JKR	$RKR = JKR/(p-1)$	RKR/RKS
Sisa	$n-p$	JKS	$RKS = JKS/(n-p)$	
Total	$n-1$	JKT		

dengan $F_{tabel} = F_{(p-1; n-p; 0,05)}$. Keputusan: terima H_0 bila $F_{hitung} < F_{tabel}$. Jika H_0 ditolak yang berarti X_h berpengaruh signifikan terhadap Y , proses selanjutnya yaitu mencari variabel prediktor lain yang berpengaruh terhadap Y dapat dilanjutkan. Sedangkan jika H_0 tidak ditolak maka langkah *Forward Selection* dapat dihentikan.

- c. Melakukan seleksi variabel kedua dengan memilih korelasi parsial, dengan rumus:

$$r_{Y X_k.X_h} = \frac{r_{Y X_k} - r_{Y X_h} r_{X_h X_k}}{\sqrt{(1-r_{Y X_h}^2)(1-r_{X_h X_k}^2)}}, \text{ dimana } h \neq k$$

dengan $r_{Y X_h}$ adalah koefisien korelasi Y dengan X_h dan $r_{Y X_k.X_h}$ = Koefisien korelasi parsial Y dengan X_k , mengendalikan X_h .

- d. Membentuk regresi kedua dengan memilih parsial korelasi variabel sisa terbesar untuk masuk kedalam model regresi. Uji keberartian regresi dilakukan seperti pada langkah b.

- e. Pemilihan model terbaik dalam metode *Forward Selection* dapat dilakukan dengan mempertimbangkan variabel prediktor yang masuk kedalam model berdasarkan koefisien determinasi (R^2).

2.3 ANFIS : Adaptive Neuro Fuzzy Inference System

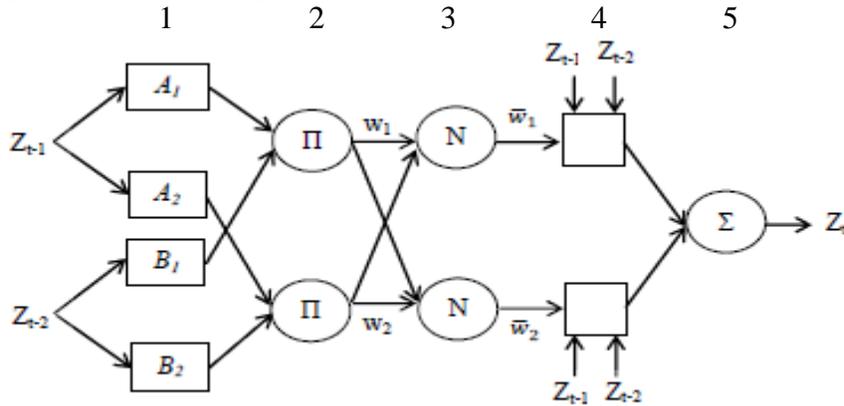
ANFIS (*Adaptive Neuro Fuzzy Inference System* atau *Adaptive Network based Fuzzy Inference System*) adalah arsitektur yang secara fungsional sama dengan *fuzzy rule base* model Sugeno. Misalkan terdapat 2 input terdiri atas Z_{t-1} dan Z_{t-2} dan sebuah output Z_t atau f dengan aturan model Sugeno orde satu dengan dua aturan *fuzzy if-then* adalah sebagai berikut:

Aturan 1 : $\underbrace{\text{If } Z_{t-1} \text{ is } A_1 \text{ and } Z_{t-2} \text{ is } B_1}_{\text{Premis}} \text{ then } \underbrace{f_1 = q_1 \cdot Z_{t-1} + r_1 \cdot Z_{t-2} + s_1}_{\text{Konsekuen}}$

Aturan 2 : $\underbrace{\text{If } Z_{t-1} \text{ is } A_2 \text{ and } Z_{t-2} \text{ is } B_2}_{\text{Premis}} \text{ then } \underbrace{f_2 = q_2 \cdot Z_{t-1} + r_2 \cdot Z_{t-2} + s_2}_{\text{Konsekuen}}$

dengan A_i dan B_i adalah nilai-nilai keanggotaan merupakan label linguistik (seperti “kecil” atau “besar”), q_i , r_i , dan s_i adalah parameter konsekuen.

Lapisan



Gambar 1. Arsitektur jaringan ANFIS untuk 2 input dengan 2 aturan

Menurut Jang *et al.* (1997), jaringan ANFIS terdiri dari lapisan-lapisan:

Lapisan 1:

Lapisan ini merupakan lapisan fuzzifikasi. Output dari tiap neuron berupa derajat keanggotaan yang diberikan oleh fungsi keanggotaan input. Misalkan fungsi keanggotaan *Gauss* diberikan sebagai:

$$\mu[Z] = e^{-\frac{1}{2}(\frac{Z-c}{\sigma})^2}$$

Dengan Z adalah input, dalam hal ini $Z = \{Z_{t-1}, Z_{t-2}\}$ dan $\{c, \sigma\}$ adalah parameter-parameter. Jika nilai parameter-parameter ini berubah, maka bentuk kurva akan ikut berubah. Parameter-parameter ini biasanya disebut dengan nama **parameter premis**.

Lapisan 2:

Lapisan ini berupa neuron tetap (diberi simbol Π) merupakan hasil kali dari semua masukan, sebagai berikut:

$$w_k = \mu_{A_k} \cdot \mu_{B_k}$$

Biasanya digunakan operator *AND*. Hasil perhitungan ini disebut *firing strength* dari sebuah aturan. Tiap neuron merepresentasikan aturan ke- k .

Lapisan 3:

Tiap neuron pada lapisan ini berupa neuron tetap (diberi simbol N) merupakan hasil perhitungan rasio dari *firing strength* ke- k (w_k) terhadap jumlah dari keseluruhan *firing strength* pada lapisan kedua, sebagai berikut:

$$\bar{w}_k = \frac{w_k}{w_1 + w_2}, \quad i = 1, 2$$

Hasil perhitungan ini disebut *normalized firing strength*.

Lapisan 4:

Lapisan ini berupa neuron yang adaptif terhadap suatu output, sebagai berikut:

$$\bar{w}_k f_k = \bar{w}_k (q_k Z_{t-1} + r_k Z_{t-2} + s_k)$$

dengan \bar{w}_k adalah *normalized firing strength* pada lapisan ketiga dan q_k , r_k , dan s_k adalah parameter-parameter pada neuron tersebut. Parameter-parameter ini biasa disebut **parameter konsekuen**.

Lapisan 5:

Lapisan ini berupa neuron tunggal (diberi simbol Σ) merupakan hasil penjumlahan seluruh output dari lapisan keempat, sebagai berikut:

$$\sum_k \bar{w}_k f_k = \frac{\sum_k w_k f_k}{\sum_k w_k} \quad (\text{Kusumadewi dan Hartati, 2006})$$

2.4 Algoritma Pembelajaran Hybrid

Pada saat parameter premis ditemukan keluaran keseluruhan akan merupakan kombinasi linier dari konsekuen parameter, yaitu:

$$\begin{aligned} f &= \frac{w_1}{w_1 + w_2} f_1 + \frac{w_2}{w_1 + w_2} f_2 \\ &= \bar{w}_1 (q_1 Z_{t-1} + r_1 Z_{t-2} + s_1) + \bar{w}_2 (q_2 Z_{t-1} + r_2 Z_{t-2} + s_2) \\ &= (\bar{w}_1 Z_{t-1}) q_1 + (\bar{w}_1 Z_{t-2}) r_1 + (\bar{w}_1) s_1 + (\bar{w}_2 Z_{t-1}) q_2 + (\bar{w}_2 Z_{t-2}) r_2 + (\bar{w}_2) s_2 \end{aligned}$$

adalah linier terhadap parameter q_1 , r_1 , s_1 , q_2 , r_2 dan s_2 .

Algoritma hibrida akan mengatur parameter-parameter konsekuen q_k , r_k dan s_k secara maju (*forward*) dan akan mengatur parameter-parameter premis a , b dan c secara mundur (*backward*).

Tabel 2. Prosedur pembelajaran *Hybrid* metode ANFIS

	Arah Maju	Arah Mundur
Parameter Premis	Tetap	<i>Gradient descent</i>
Parameter Konsekuen	<i>Least-square estimator</i>	Tetap
Sinyal	Keluaran neuron	Sinyal error

(Jang *et al.*, 1997)

2.5 Pemilihan Model Terbaik

Untuk menentukan model terbaik digunakan beberapa ukuran pengukuran antara lain:

$$AIC = n \ln \sigma_a^2 + 2M$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{\hat{z}_t - z_t}{z_t} \right| \times 100\%$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n \varepsilon_i^2}{n}} = \sqrt{\frac{\sum_{i=1}^n (z_{t(i)} - \hat{z}_{t(i)})^2}{n}}$$

3. METODOLOGI PENELITIAN

3.1 Sumber dan Variabel Penelitian

Penelitian ini menggunakan data simulasi dengan karakteristik stasioner, stasioner dengan *outlier*, nonstasioner dan nonstasioner dengan *outlier* sebanyak 150 data yang dibangkitkan menggunakan *software* Matlab R2015b untuk melihat konsistensi metode *Forward Selection* sebagai input model ANFIS. Selain itu juga digunakan data sekunder sebagai data studi kasus berupa data bulanan jumlah penumpang kereta api di wilayah Jawa Non Jabodetabek terhitung dari Januari 2006 sampai dengan Desember 2018 yang diperoleh dari website resmi BPS sebanyak 156 untuk pemodelan.

3.2 Metode Analisis

1. Pemodelan *Forward Selection*

- a. Mendefinisikan variabel respon Z_t dan variabel prediktor Z_{t-k} dengan $k=1, 2, \dots, p$.
- b. Membentuk matriks koefisien korelasi antar variabel.
- c. Menentukan koefisien korelasi terbesar dan membentuk model regresi sementara antara variabel Z_t dan Z_{t-k} .
- d. Melakukan uji F, jika variabel prediktor tidak signifikan berpengaruh maka langkah dihentikan. Apabila variabel prediktor signifikan berpengaruh proses dilanjutkan.
- e. Menghitung korelasi parsial dari variabel prediktor sisa dan ulangi langkah c.

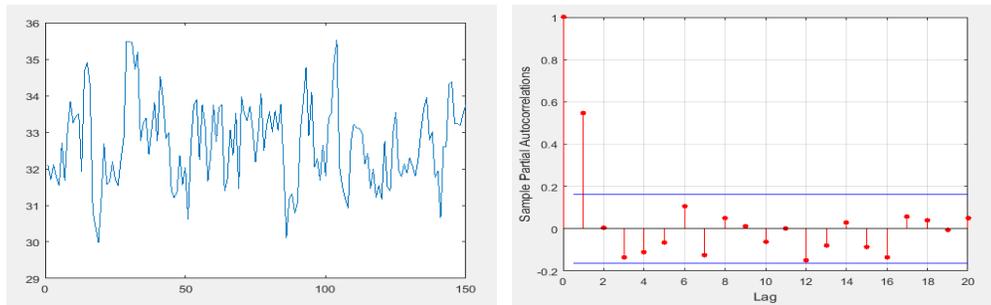
2. Pemodelan ANFIS

- a. Menentukan variabel input dari lag yang *cut off* pada plot PACF atau variabel prediktor yang signifikan berdasarkan metode *Forward Selection*.
- b. Menentukan banyaknya jumlah keanggotaan atau *cluster* yaitu 2 dan 3 serta fungsi keanggotaan yaitu *Gauss* dan *Generalized Bell (gbell)*.
- c. Melakukan pembagian input data yang bernilai *crisp* menjadi linguistik sesuai banyaknya kelompok dengan *Fuzzy C-Means* dan *Grid Partition*.
- d. Menentukan banyak iterasi (*epoch*) yang diinginkan.
- e. Menjalankan pelatihan parameter FIS (*Fuzzy Inference System*) yaitu fuzzifikasi, operasi logika fuzzy, pengaktifan derajat ternormalisasi dan defuzzifikasi.
- f. Menghitung nilai AIC, MAPE dan RMSE pada model.
- g. Melakukan pengecekan model optimal yang terbentuk berdasarkan nilai AIC, MAPE dan RMSE terkecil. Jika belum optimal maka dapat dilakukan pengulangan proses dalam menentukan jumlah dan atau fungsi keanggotaan yang berbeda.
- h. Memilih model ANFIS terbaik berdasarkan AIC, MAPE dan RMSE terkecil.

3. Hasil dan Pembahasan

Analisis pada Data Stasioner

Data yang digunakan merupakan data *random* ARIMA(1,0,0) atau AR(1) yang dibangkitkan dari program Matlab dengan $\phi_1 = 0,6$ sebanyak 150 buah. Analisis pemilihan input pada data simulasi stasioner adalah sebagai berikut:



Gambar 2. Grafik Runtun waktu data Stasioner (kiri) dan plot PACF (kanan)

Tabel 3. Hasil Seleksi Variabel Metode *Forward Selection* Data Stasioner

Langkah	Variabel Dimasukkan	p-value
1	Z_{t-1}	$4,84831 \times 10^{-12}$

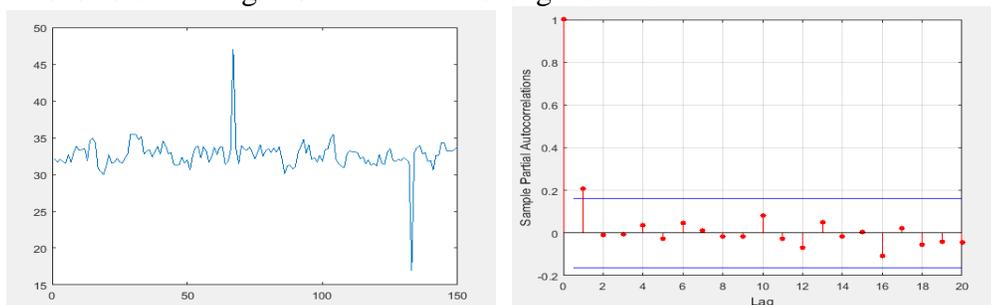
Berdasarkan plot PACF, lag yang *cut off* adalah lag ke-1. Dengan *Forward Selection* variabel prediktor pertama dan satu-satunya yang masuk dalam model adalah Z_{t-1} yang dikuatkan dengan uji F dimana $F_{hitung} = 57,4451$ lebih besar daripada $F_{tabel} = 3,9107$ yang artinya Z_{t-1} berpengaruh signifikan terhadap Z_t . Hal ini mengindikasikan performa *Forward Selection* dalam mengidentifikasi model sama dengan plot PACF yang dibuktikan dengan kesesuaiannya sebagai input ANFIS yang menghasilkan nilai AIC, MAPE dan RMSE berikut:

Tabel 4. Ringkasan Perbandingan Nilai AIC, MAPE dan RMSE Data Stasioner

Teknik	Input	Jumlah Keanggotaan 2			Jumlah Keanggotaan 3		
		AIC	MAPE	RMSE	AIC	MAPE	RMSE
FCM	Z_{t-1}	2,7761	2,3874	0,9631	2,7356	2,3070	0,9438

Analisis pada Data Stasioner dengan *Outlier*

Data yang digunakan merupakan data *random* ARIMA(1,0,0) atau AR(1) yang dibangkitkan dari program Matlab dengan $\phi_1 = 0,6$ sebanyak 150 buah dan ditentukan *outlier* dengan mengganti data ke-67 menjadi 47 dan data ke-133 menjadi 17. Analisis pemilihan input pada data simulasi stasioner dengan *outlier* adalah sebagai berikut:



Gambar 3. Grafik Runtun waktu data Stasioner dengan *outlier* (kiri) dan plot PACF (kanan)

Tabel 5. Hasil Seleksi Variabel Metode *Forward Selection* Data Stasioner dengan *Outlier*

Langkah	Variabel Dimasukkan	p-value
1	Z_{t-1}	0,01653

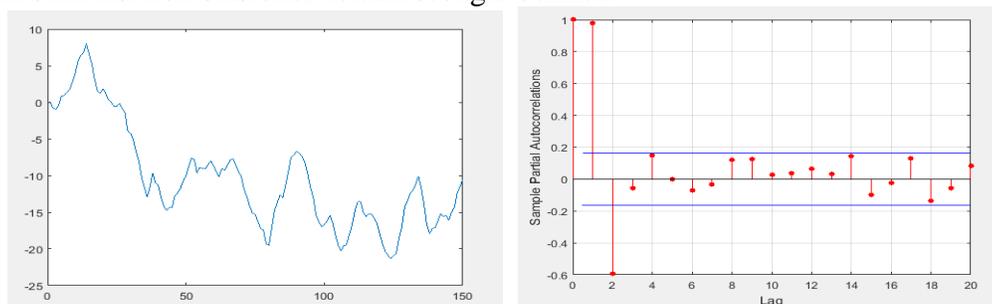
Berdasarkan plot PACF, lag yang *cut off* adalah lag ke-1. Dengan *Forward Selection* variabel prediktor pertama dan satu-satunya yang masuk dalam model adalah Z_{t-1} yang dikuatkan dengan uji F dimana $F_{hitung} = 5,8906$ lebih besar daripada $F_{tabel} = 3,9107$ yang artinya Z_{t-1} berpengaruh signifikan terhadap Z_t . Hal ini mengindikasikan performa *Forward Selection* dalam mengidentifikasi model sama dengan plot PACF yang dibuktikan dengan kesesuaiannya sebagai input ANFIS yang menghasilkan nilai AIC, MAPE dan RMSE berikut:

Tabel 6. Ringkasan Perbandingan Nilai AIC, MAPE dan RMSE Data Stasioner dengan *Outlier*

Teknik	Input	Jumlah Keanggotaan 2			Jumlah Keanggotaan 3		
		AIC	MAPE	RMSE	AIC	MAPE	RMSE
FCM	Z_{t-1}	4,1629	3,1805	1,9267	4,1249	3,0317	1,8905

Analisis pada Data Non Stasioner

Data yang digunakan merupakan data *random* ARIMA(1,1,0) atau ARI(1,1) yang dibangkitkan dari program Matlab dengan $\phi_1 = 0,6$ sebanyak 150. Analisis pemilihan input pada data simulasi non stasioner adalah sebagai berikut:



Gambar 4. Grafik Runtun waktu data Non Stasioner (kiri) dan plot PACF (kanan)

Tabel 7. Hasil Seleksi Variabel Metode *Forward Selection* Data Non Stasioner

Langkah	Variabel Dimasukkan	p-value
1	Z_{t-1}	$3,60315 \times 10^{-103}$
2	Z_{t-2}	$7,73876 \times 10^{-14}$

Berdasarkan plot PACF, lag yang *cut off* adalah lag ke-1 dan lag ke-2. Dengan *Forward Selection* variabel prediktor yang masuk dalam model adalah Z_{t-1} dan Z_{t-2} yang dikuatkan dengan uji F dimana $F_{hitung} = 5,8906$ lebih besar daripada $F_{tabel} = 3,9107$ yang artinya Z_{t-1} dan Z_{t-2} berpengaruh signifikan terhadap Z_t . Hal ini mengindikasikan performa *Forward Selection* dalam mengidentifikasi model sama dengan plot PACF yang dibuktikan dengan kesesuaiannya sebagai input ANFIS yang menghasilkan nilai AIC, MAPE dan RMSE berikut:

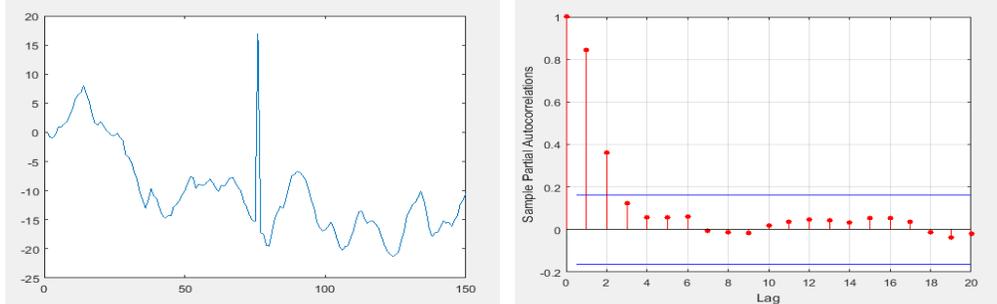
Tabel 8. Ringkasan Perbandingan Nilai AIC, MAPE dan RMSE Data Non Stasioner

Teknik	Input	FK	Jumlah Keanggotaan 2			Jumlah Keanggotaan 3		
			AIC	MAPE	RMSE	AIC	MAPE	RMSE
<i>Grid Partition</i>	Z_{t-1}, Z_{t-2}	<i>Gauss</i>	2,5721	16,456	0,8638	2,5195	18,3511	0,8414

Analisis pada Data Non Stasioner dengan *Outlier*

Data yang digunakan merupakan data *random* ARIMA(1,1,0) atau ARI(1,1) yang dibangkitkan dari program Matlab dengan $\phi_1 = 0,6$ sebanyak 150 buah dan ditentukan *outlier*

dengan mengganti data ke-67 menjadi 76 menjadi 17. Analisis pemilihan input pada data simulasi stasioner dengan *outlier* adalah sebagai berikut:



Gambar 5. Grafik Runtun waktu data Non Stasioner dengan *outlier* (kiri) dan plot PACF (kanan)

Tabel 9. Hasil Seleksi Variabel Metode *Forward Selection* Data Non Stasioner dengan *Outlier*

Langkah	Variabel Dimasukkan	p-value
1	Z_{t-1}	$1,67966 \times 10^{-33}$
2	Z_{t-2}	$1,68258 \times 10^{-5}$

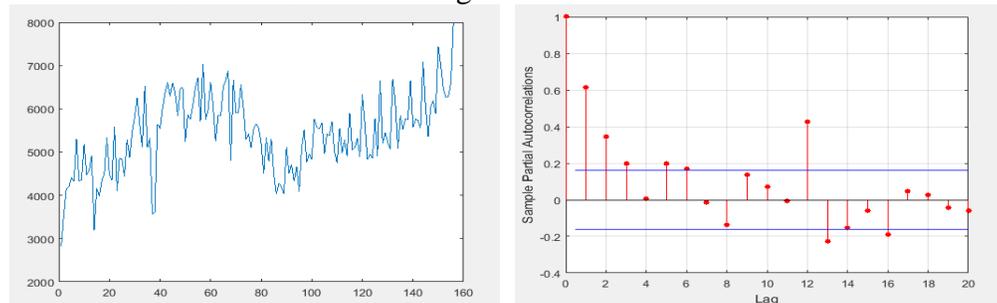
Berdasarkan plot PACF, lag yang *cut off* adalah lag ke-1 dan ke-2. Dengan *Forward Selection* variabel prediktor yang masuk dalam model adalah Z_{t-1} dan Z_{t-2} yang dikuatkan dengan uji F dimana $F_{hitung} = 5,8906$ lebih besar daripada $F_{tabel} = 3,9107$ yang artinya Z_{t-1} dan Z_{t-2} berpengaruh signifikan terhadap Z_t . Hal ini mengindikasikan performa *Forward Selection* dalam mengidentifikasi model sama dengan plot PACF yang dibuktikan dengan kesesuaiannya sebagai input ANFIS yang menghasilkan nilai AIC, MAPE dan RMSE berikut:

Tabel 10. Ringkasan Perbandingan Nilai AIC, MAPE dan RMSE Data Stasioner dengan *Outlier*

Teknik	Input	FK	Jumlah Keanggotaan 2			Jumlah Keanggotaan 3		
			AIC	MAPE	RMSE	AIC	MAPE	RMSE
<i>Grid Partition</i>	Z_{t-1}, Z_{t-2}	<i>Gauss</i>	4,8988	16,0664	2,7648	4,8730	25,5458	2,7293

Analisis pada Data Studi Kasus

Analisis pemilihan input pada data jumlah penumpang kereta api di wilayah Jawa Non Jabodetabek tahun 2006 - 2018 adalah sebagai berikut:



Gambar 6. Grafik Runtun waktu data Studi Kasus (kiri) dan plot PACF (kanan)

Tabel 11. Hasil Seleksi Variabel Metode *Forward Selection* Data Studi Kasus

Langkah	Variabel Dimasukkan	p-value
1	Z_{t-12}	$2,10532 \times 10^{-14}$
2	Z_{t-1}	$1,65899 \times 10^{-9}$
3	Z_{t-2}	0,00156448

Berdasarkan plot PACF, terdapat beberapa lag yang *cut off*. Dengan *Forward Selection* variabel prediktor yang masuk dalam model adalah Z_{t-1} , Z_{t-2} dan Z_{t-12} yang dikuatkan dengan uji F dimana $F_{hitung} = 51,4562$ lebih besar daripada $F_{tabel} = 2,6693$ yang artinya Z_{t-1} , Z_{t-2} dan Z_{t-12} berpengaruh signifikan terhadap Z_t . Terdapat beberapa perbedaan dengan konsistensi antara *Forward Selection* dan plot PACF pada variabel prediktor Z_{t-1} , Z_{t-2} dan Z_{t-12} yang digunakan sebagai input ANFIS yang menghasilkan nilai AIC, MAPE dan RMSE seperti pada Tabel 12:

Tabel 12. Ringkasan Perbandingan Nilai AIC, MAPE dan RMSE Data Studi Kasus

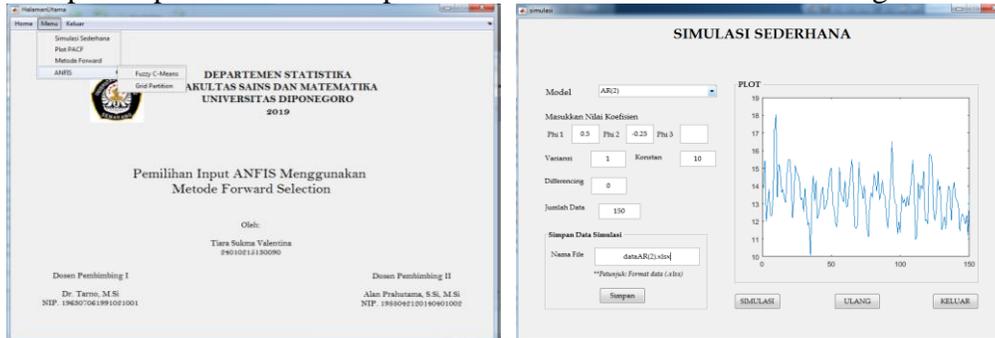
Teknik	Input	FK	Jumlah Keanggotaan 2			Jumlah Keanggotaan 3		
			AIC	MAPE	RMSE	AIC	MAPE	RMSE
<i>Grid Partition</i>	Z_{t-1}, Z_{t-12}	<i>Gauss</i>	15,5617	8,5093	571.3691	15,3807	7,7998	521,9149

Berdasarkan pemilihan model yang terbaik, diperoleh model ANFIS sebagai berikut:

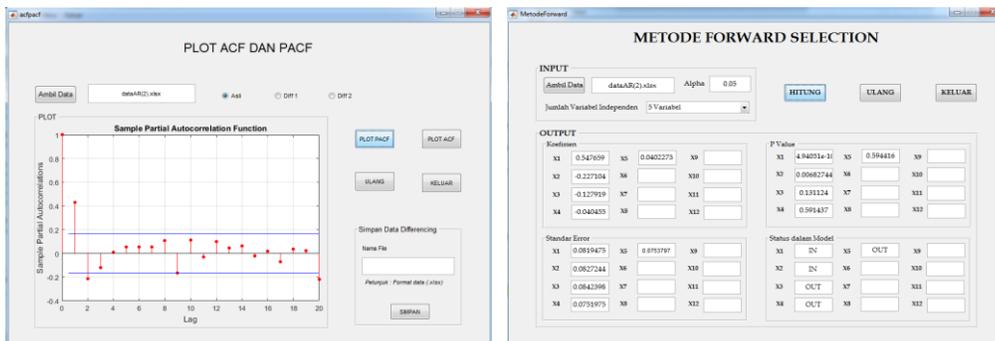
$$\hat{Z}_t = \bar{w}_1(0,8912 Z_{t-1} - 0,5185 Z_{t-12} + 2133) + \bar{w}_2(-1,085 Z_{t-1} - 0,6557 Z_{t-12} + 13150) + \bar{w}_3(0,9788 Z_{t-1} + 0,4702 Z_{t-12} - 2196) + \bar{w}_4(-1,385 Z_{t-1} + 1,389 Z_{t-12} + 7610)$$

Tampilan GUI

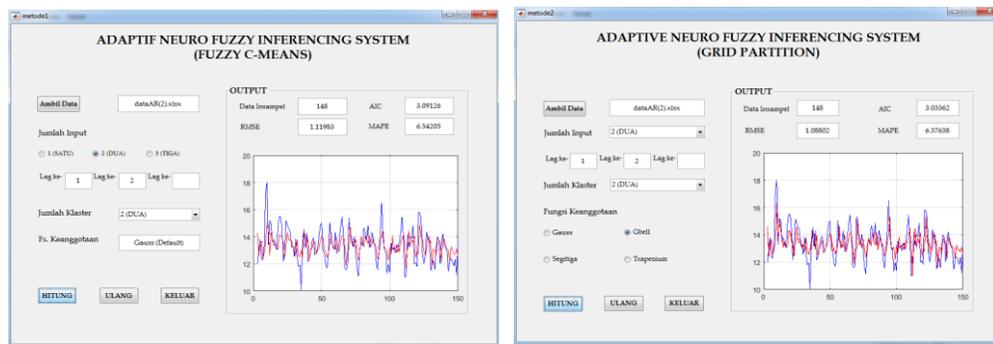
Beberapa tampilan umum hasil pembentukan GUI Matlab adalah sebagai berikut:



Gambar 7. Tampilan Halaman Utama (Kiri) dan Tampilan Hasil Simulasi Sederhana



Gambar 8. Tampilan Hasil Plot PACF (Kiri) dan Tampilan Hasil *Forward Selection* (Kanan)



Gambar 9. Tampilan Hasil ANFIS dengan FCM (Kiri) dan Tampilan Hasil ANFIS dengan Grid Partition (Kanan)

4. KESIMPULAN

- a. Hasil analisis menunjukkan bahwa metode *Forward Selection* dapat dijadikan sebagai salah satu alternatif yang efektif dalam mengidentifikasi input ANFIS.
- b. Analisis data runtun waktu menggunakan ANFIS pada data studi kasus dan 4 karakteristik data yaitu stasioner, stasioner dengan outlier, non stasioner dan non stasioner dengan outlier menunjukkan hasil yang sangat baik terlihat dari AIC, MAPE dan RMSE serta plot perbandingan data dengan output ANFIS.
- c. Analisis pada data jumlah penumpang kereta api di wilayah Jawa Non Jabodetabek menghasilkan model ANFIS terbaik menggunakan teknik *Grid Partition* dengan input Z_{t-1} , Z_{t-12} dan jumlah keanggotaan 2:

$$\hat{Z}_t = \bar{w}_1(0,8912 Z_{t-1} - 0,5185 Z_{t-12} + 2133) + \bar{w}_2(-1,085 Z_{t-1} - 0,6557 Z_{t-12} + 13150) + \bar{w}_3(0,9788 Z_{t-1} + 0,4702 Z_{t-12} - 2196) + \bar{w}_4(-1,385 Z_{t-1} + 1,389 Z_{t-12} + 7610)$$

dengan $\bar{w}_i = \frac{w_i}{\sum_{i=1}^4 w_i}$

5. Daftar Pustaka

- Algifari. 2010. *Analisis Regresi: Teori, kasus dan solusi*. Yogyakarta:BPFE.
- Chang, J.R., Wei L.Y. dan Cheng C.H. 2011. *A Hybrid ANFIS Model Based on AR and Volatility for TAIEX Forecasting*. Applied Soft Computing Vol. 11, No. 1:Hal 1388-1395.
- Jang, J.S.R., Sun, C.T. dan Mizutani, E. 1997. *Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence*. London: Prentice-Hall, Inc.
- Kusumadewi, S dan Hartati S. 2006. *Neuro Fuzzy: Integrasi Sistem Fuzzy & Jaringan Syaraf*. Yogyakarta: Graha Ilmu.
- Samosir N., Siagian, P. dan Bangun, P. 2014. *Analisa Metode Backward dan Metode Forward untuk Menentukan Persamaan Regresi Linier Berganda (Studi Kasus: Jumlah Kecelakaan Lalu Lintas di Kotamadya Medan)*. Jurnal Sainia Matematika. Vol. 2, No. 4: Hal. 345-360.
- Svalina, I., Galzina, V., Lujic, R. dan Simunovic, G. 2013. *An adaptive network-based fuzzy inference system (ANFIS) for the forecasting: The case of close price indices*. Jurnal Expert Systems with Applications Vol. 40:Hal. 6055-6063.

- Tarno *et al.* 2018. *Pemodelan Hybrid ARIMA-ANFIS untuk Data Produksi Tanaman Hortikultura di Jawa Tengah*. Media Statistika. Vol. 11, No. 1: Hal. 65-78.
- Wei, WWS. 2006. *Time Series Analysis Univariate and Multivariate Methods Second Edition*. USA: Pearson Education, Inc.