

MODEL ARFIMA UNTUK ANALISIS DATA KECEPATAN ANGIN DI BANDARA INTERNASIONAL AHMAD YANI

Mohammad Jefrie Ilham Akbar✉, Iqbal Kharisudin

Jurusan Matematika, FMIPA, Universitas Negeri Semarang, Indonesia
Gedung D7 Lt. 1, Kampus Sekaran Gunungpati, Semarang 50229

Info Artikel

Sejarah Artikel:
Diterima Juli 2019
Disetujui Juli 2019
Dipublikasikan Agustus 2019

Keywords:
Long Memory, ARFIMA, GPH, R/S

Abstrak

Penelitian ini bertujuan untuk pemodelan ARFIMA pada data kecepatan angin di Bandara Internasional Ahmad Yani dalam frekuensi harian. Pada penelitian ini dilakukan identifikasi *long memory* pada data kecepatan angin di Bandara Internasional Ahmad Yani yang memberikan hasil bahwa data diindikasikan mempunyai sifat *long memory*. Selanjutnya dilakukan pembentukan model ARFIMA(p, d, q) dengan menentukan nilai estimasi parameter d menggunakan metode *Geweke and Porter-Hudak* (GPH) dan *Rescaled Range Statistics* (R/S). Model terbaik dipilih berdasarkan nilai AIC terkecil. Model terbaik untuk ARFIMA(p, d, q) dengan $\hat{d}_{gph} = 0,346$ adalah model ARFIMA(0, $d, [2]$) dengan nilai AIC -1384,527. Model terbaik untuk ARFIMA(p, d, q) dengan $\hat{d}_{R/S} = 0,224$ adalah model ARFIMA(0, $d, 1$) dengan nilai AIC -1385,62. Tingkat akurasi peramalan didasarkan pada nilai RMSE, MAE, dan MAPE. Nilai error validasi peramalan model ARFIMA(0, $d, [2]$) dengan $\hat{d}_{gph} = 0,346$ adalah RMSE 0,74966754, MAE 0,593468, dan MAPE 10,45511. Nilai error pada validasi peramalan model ARFIMA(0, $d, 1$) dengan $\hat{d}_{R/S} = 0,224$ adalah RMSE 0,7898162, MAE 0,6295811, dan MAPE 10,75531.

Abstract

Abstract The purposes of this research were to model ARFIMA on data of wind speed at Ahmad Yani International Airport in a daily frequency. In this research, identification of long memory on data of wind speed in Ahmad Yani International Airport given the result that the data is indicated to have a long memory. Then ARFIMA(p, d, q) model is formed by determining the estimated value of the parameter d using the Geweke and Porter-Hudak (GPH) and Rescaled range Statistics (R/S) method. The best model is chosen based on the smallest AIC value. The best model for ARFIMA(p, d, q) where $\hat{d}_{gph} = 0,346$ is ARFIMA(0, $d, [2]$) with AIC -1384,527. The best model for ARFIMA(p, d, q) where $\hat{d}_{R/S} = 0,224$ is ARFIMA(0, $d, 1$) with AIC -1385,62. The forecast accuracy is based on the RMSE, MAE, and MAPE value. Error value of forecasting validation ARFIMA(p, d, q) where $\hat{d}_{gph} = 0,346$ is RMSE 0,74966754, MAE 0,593468, and MAPE 10,45511. Error value of forecasting validation ARFIMA(p, d, q) where $\hat{d}_{R/S} = 0,224$ is RMSE 0,7898162, MAE 0,6295811, and MAPE 10,75531.

How to cite:

Akbar, MJI. & Kharisudin, I. 2018. Model Arfima Untuk Analisis Data Kecepatan Angin Di Bandara Internasional Ahmad Yani. *UNNES Journal of Mathematics*. 8(2):89-101.

PENDAHULUAN

Data runtun waktu (*time series*) merupakan serangkaian observasi terhadap suatu variabel yang diambil secara beruntun berdasarkan interval waktu yang tetap (Wei, 2006). Analisis runtun waktu merupakan salah satu prosedur statistika yang diterapkan untuk memodelkan dan meramalkan struktur probabilitas keadaan yang akan datang dalam rangka pengambilan keputusan. Beberapa metode dalam statistika yang dapat digunakan untuk peramalan suatu data runtun waktu antara lain *Autoregressive* (AR), *Moving Average* (MA), ARIMA, SARIMA, *exponential smoothing* maupun Jaringan Syaraf Tiruan (JST). Metode-metode tersebut biasanya digunakan untuk menyelesaikan masalah data dengan ketergantungan atau dependensi jangka pendek (*short memory*).

Proses *long memory* telah diamati dalam banyak hal seperti dalam bidang industri, meteorologi, ekonomi dan keuangan. *Long memory* merupakan salah satu fenomena dalam runtun waktu, dimana data runtun waktu mempunyai dependensi jangka panjang yang artinya jika di antara pengamatan yang terpisah jauh, masih mempunyai korelasi yang tinggi. Sifat *long memory* pada data juga ditunjukkan oleh plot *Autocorrelation Function* (ACF) dari data tidak turun secara eksponensial melainkan turun secara lambat atau hiperbolik. Salah satu model runtun waktu yang dapat menggambarkan dependensi jangka panjang adalah model *Autoregressive Fractionally Integrated Moving Average* (ARFIMA).

Model *Autoregressive Fractionally Integrated Moving Average* (ARFIMA) pertama kali dikembangkan oleh Granger dan Joyeux (1980) yang merupakan pengembangan dari model *Autoregressive Integrated Moving Average* (ARIMA). Hosking (1981) memperkenalkan model *Autoregressive Fractionally Integrated Moving Average* (ARFIMA) untuk memodelkan data *long memory*. Model ARIMA hanya dapat menjelaskan data *short memory*, sedangkan model ARFIMA mempunyai kelebihan dapat melakukan pembedaan (*differencing*) untuk nilai yang telah ditentukan meliputi bilangan riil atau pecahan ini berbeda dengan ARIMA yang mempunyai pembeda bilangan bulat. Dengan demikian ARFIMA mampu menjelaskan data runtun waktu baik *short memory* maupun *long memory*.

Sudah banyak penelitian-penelitian terkait *long memory* dan ARFIMA. Penelitian-penelitian tersebut antara lain, Mohamed (2016) menggunakan model ARFIMA dalam meramalkan total nilai dari *traded securities* di

Republik Arab dari Mesir, yang mana dalam penelitian ini menggunakan beberapa metode estimasi parameter. Kharisudin dkk (2017) melakukan penelitian terkait model *long memory* FARIMA dan aplikasinya pada pemodelan data aset returns di Indonesia. Walid dkk (2015) melakukan identifikasi model *long memory* pada penggunaan beban listrik di Jawa-Bali. Ibrahim dkk (2011) melakukan pemodelan ARFIMA pada data harian rata-rata suhu di kota Sokoto. Paul (2017) meneliti terkait model *long memory* pada data minimum dan maksimum temperatur di India.

Cuaca merupakan aspek penting dalam aktivitas penerbangan. Cuaca mempunyai dua sisi dampak yang berbeda bagi operasional penerbangan. Pada satu sisi kondisi cuaca bisa memberi dampak ekonomis berupa efisiensi pada pesawat terbang, Pada sisi lain cuaca memberi ancaman pada keselamatan penerbangan. Adapun faktor cuaca yang mempengaruhi aktivitas penerbangan antara lain angin, tekanan udara, awan, hujan, kabut dan salju.

Informasi kecepatan angin di permukaan landasan pacu merupakan salah satu faktor penting dalam proses lepas landas (*take-off*) dan pendaratan (*landing*) pesawat. Kecepatan angin dapat mempengaruhi daya angkat pesawat serta menghindari pesawat tergelincir dan keluar dari landasan pacu.

Dalam penelitian ini penulis akan mengidentifikasi apakah data kecepatan angin di Bandara Internasional Ahmad Yani diindikasikan adanya *long memory* atau tidak. Selanjutnya penulis menggunakan model *Autoregressive Fractionally Integrated Moving Average* (ARFIMA) dalam menganalisis data kecepatan angin di Bandara Internasional Ahmad Yani.

Angin merupakan gerakan udara yang sejajar dengan permukaan bumi. Udara bergerak dari daerah bertekanan tinggi ke daerah bertekanan rendah. Gerak rata-rata dari angin dipermukaan bumi merupakan sirkulasi umum dari atmosfer. Pemahaman sirkulasi umum atmosfer dapat didekati oleh pemeriksaan pola angin utama dan pemeriksaan berbagai daerah bertekanan rendah dan tinggi. Angin terjadi disebabkan oleh adanya beda tekanan, sedangkan kecepatan angin ditunjukkan oleh kecuraman beda tekanan. Jika beda tekanan besar maka angin menjadi kencang. Sebaliknya jika beda tekanan kecil maka angin juga melemah (Tjasyono, 2004). Kecepatan angin merupakan jarak tempuh angin atau pergerakan udara per satuan waktu yang dinyatakan dalam satuan

meter per detik (m/s), kilometer per jam (km/jam), dan mil per jam (mil/jam atau knot).

Long memory memiliki arti bila fungsi autokorelasi antara Z_t dengan Z_{t+k} turun secara hiperbolik dan lag yang signifikan semakin banyak maka dapat diidentifikasi adanya dependensi jangka panjang dalam data yang merupakan ciri dari memori jangka panjang (*long memory*). Suatu runtun waktu stasioner $Z_t, t = 1, 2, \dots, N$ tergolong ke dalam *long memory* ketika mempunyai fungsi korelasi $\rho(k)$ berlaku $k \rightarrow \infty$ adalah sebagai berikut.

$$\lim_{k \rightarrow \infty} \frac{\rho(k)}{C_p(k)^{2d-1}} = 1 \quad (1)$$

Adapun C_p adalah konstanta dan $d \in (0, 0.5)$ menunjukkan parameter *long memory*. Korelasi dari suatu proses *long memory* akan cenderung turun perlahan secara hiperbolik (Beran, 1994).

Selain melihat dari plot ACF, adanya *long memory* pada data runtun waktu bisa diketahui melalui nilai *Hurst* (H) yang diperoleh dengan menggunakan statistik R/S (Hurst, 1951). Apabila nilai $H = 0.5$, maka data tersebut memiliki sifat *short memory*. Apabila berada nilai H pada interval $0 < H < 0.5$, maka data tersebut memiliki sifat *intermediate memory*. Dan apabila nilai H pada interval $0.5 < H < 1$, maka data tersebut memiliki sifat *long memory*. Identifikasi *long memory* dengan menentukan nilai *Hurst* (H) diperoleh melalui perhitungan dengan metode *Rescaled Range Statistics* (R/S) dengan langkah-langkah sebagai berikut (Safitri dkk, 2019):

1. Menghitung *mean* (\bar{Z}).

$$\bar{Z} = \frac{1}{T} \sum_{t=1}^T Z_t, \quad t = 1, 2, \dots, T \quad (2)$$

2. Menghitung *adjusted mean*.

$$Z_t^{adj} = Z_t - \bar{Z}, \quad t = 1, 2, \dots, T \quad (3)$$

3. Menghitung deviasi kumulatif.

$$Z_t^* = \sum_{t=1}^T Z_t^{adj}, \quad t = 1, 2, \dots, T \quad (4)$$

4. Menghitung rentang deviasi kumulatif.

$$R_t = \max(Z_1^*, Z_2^*, \dots, Z_t^*) - \min(Z_1^*, Z_2^*, \dots, Z_t^*), \quad t = 1, 2, \dots, T. \quad (5)$$

5. Menghitung standar deviasi.

$$s_t = \sqrt{\frac{1}{T} \sum_{t=1}^T (Z_t - \bar{Z})^2}, \quad t = 1, 2, \dots, T \quad (6)$$

6. Menghitung *Rescaled Range* (R/S).

$$(R/S)_t = \frac{R_t}{S_t}, \quad t = 1, 2, \dots, T \quad (7)$$

7. Menentukan nilai *Hurst* (H) melalui statistik R/S dari data runtun waktu.

$$(R/S)_t = c \cdot t^H, \quad t = 1, 2, \dots, T \quad (8)$$

$$\log(R/S)_t = \log c + H \log t \quad (9)$$

dengan c suatu konstanta dan H nilai

Hurst.

8. Selanjutnya nilai *Hurst* (H) dapat diperoleh menggunakan metode *Ordinary Least Square* (OLS).

$$H = \frac{\sum_{j=1}^T (X_j - \bar{X})(Y_j - \bar{Y})}{\sum_{j=1}^T (X_j - \bar{X})^2} \quad (10)$$

dengan, $X_j = \log t$ dan $Y_j = \log(R/S)_t$.

Proses ARFIMA(p, d, q), didefinisikan dengan p dan q adalah bilangan bulat non negatif dan $0 < d < 0.5$, merupakan runtun waktu stasioner dengan fungsi autokorelasi yang mana lagnya turun secara lambat (Brockwell dan Davis, 2016). Model *Autoregressive Fractionally Integrated Moving Average* (ARFIMA) dapat mengatasi kelemahan model ARIMA, dimana ARIMA hanya dapat menjelaskan data jangka pendek dengan *differencing* (d) bilangan bulat. Moulines dan Soulier (1999) mengungkapkan bahwa model ARFIMA merupakan model yang mampu menjelaskan data runtun waktu baik berupa data jangka pendek maupun jangka panjang dengan *differencing* (d) bernilai pecahan. Bentuk umum dari model ARFIMA(p, d, q) sebagai berikut (Wei, 2006):

$$\phi_p(B)(1-B)^d Z_t = \theta_q(B) a_t \quad (11)$$

dengan,

d : parameter pembeda (bilangan pecahan)

$\phi_p(B) = 1 - \phi_1 B - \dots - \phi_p B^p$ polinomial

autoregressive ke- p

$\theta_q(B) = 1 - \theta_1 B - \dots - \theta_q B^q$ polinomial

moving average ke- q

$a_t \sim \text{IIDN}(0, \sigma^2)$

$(1-B)^d = \Delta^d = \sum_{k=0}^{\infty} \binom{d}{k} (-1)^k$ operator

pembeda pecahan.

Untuk suatu nilai d bernilai pecahan, operator fraksional diferensi $(1-B)^d$ didefinisikan sebagai berikut:

$$(1-B)^d = 1 + \sum_{k=1}^{\infty} \frac{\Gamma(-d+k)}{\Gamma(-d)k!} B^k \quad (12)$$

Pada persamaan (12) untuk berbagai nilai k , ekuivalen dengan:

$$\begin{aligned} (1-B)^d &= 1 + \sum_{k=1}^{\infty} \frac{\Gamma(-d+k)}{\Gamma(-d)k!} B^k \\ (1-B)^d &= 1 + \frac{\Gamma(-d+1)}{\Gamma(-d)1!} B^1 + \frac{\Gamma(-d+2)}{\Gamma(-d)2!} B^2 \\ &+ \frac{\Gamma(-d+3)}{\Gamma(-d)3!} B^3 + \dots \end{aligned}$$

$$\begin{aligned}
(1-B)^d &= 1 + \frac{\Gamma(-d)}{\Gamma(-d-1)!} B^1 + \frac{\Gamma(-d+1)}{\Gamma(-d-2)!} B^2 \\
&+ \frac{\Gamma(-d+1)}{\Gamma(-d-3)!} B^3 + \dots \\
(1-B)^d &= 1 - dB - \frac{d(1-d)B^2}{2} - \\
&\frac{d(1-d)(2-d)B^3}{6} + \dots
\end{aligned} \quad (13)$$

Metode estimasi parameter d yang akan digunakan adalah metode *Geweke and Porter-Hudak* (GPH) dan *Rescaled Range Statistics* (R/S). Metode GPH pertama kali diusulkan oleh Geweke dan Porter-Hudak pada tahun 1983, dimana parameter *differencing* (d) dapat diestimasi secara konsisten dari regresi kuadrat terkecil yang diperoleh dari penaksiran persamaan logaritma *spectral density*. Tahapan pertama pada metode GPH adalah menentukan nilai frekuensi harmonik ω_j untuk setiap observasi.

$$\omega_j = (2\pi \cdot j/T) \quad (14)$$

dengan, $j = 1, 2, \dots, m$.

Bandwidth optimal m dibatasi sampai $m = g(T) = [T^{0.8}]$. Tahap selanjutnya adalah menentukan nilai periodogram dengan metode GPH, yang bentuk periodogramnya ditentukan melalui persamaan berikut.

$$I_Z(\omega_j) = \frac{1}{2\pi} \left\{ \gamma_0 + 2 \sum_{t=1}^{T-1} \gamma_t \cos(t \cdot \omega_j) \right\} \quad (15)$$

dengan, $\omega_j \in (-\pi, \pi)$ dan γ_t adalah nilai autokovarian dari *lag* ke- t . Kemudian nilai dari logaritma natural periodogramnya dijadikan sebagai variabel respon Y_j untuk regresi spektral.

$$Y_j = \ln(I_Z(\omega_j)). \quad (16)$$

Selanjutnya, untuk variabel prediktornya ditentukan dari persamaan berikut

$$X_j = \ln \left(\frac{1}{4 \sin^2(\omega_j/2)} \right). \quad (17)$$

Sehingga dengan persamaan regresi linear $Y_j = \beta_0 + \beta_1 X_j + a_j$, nilai estimasi parameter d ditentukan dengan metode *least square* seperti pada persamaan berikut.

$$\hat{\beta}_1 = \hat{d} = \frac{\sum_{j=1}^m (X_j - \bar{X})(Y_j - \bar{Y})}{\sum_{j=1}^m (X_j - \bar{X})^2} \quad (18)$$

dengan, $j = 1, 2, \dots, m$.

Metode estimasi *Rescaled Range Statistics* (R/S) juga sering disebut sebagai uji statistik *Hurst* (H). Penentuan nilai *Hurst* (H) dengan menggunakan estimasi R/S bisa dilihat pada persamaan (2) sampai dengan (10). Selain digunakan untuk melihat indikasi adanya dependensi jangka panjang (*long memory*) pada data runtun waktu, statistik R/S juga dapat digunakan untuk estimasi parameter d . Setelah

nilai *Hurst* (H) diperoleh maka estimasi parameter d menggunakan metode R/S diperoleh dengan persamaan (Beran, 1994).

$$\hat{d} = H - 0,5 \quad (19)$$

METODE

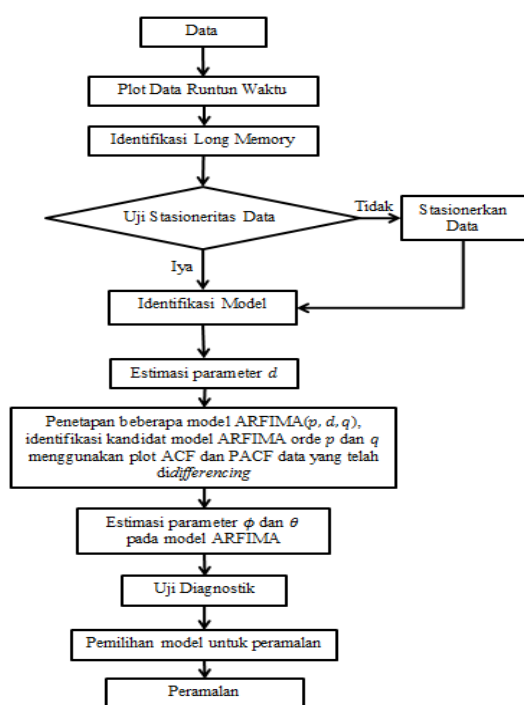
Data yang digunakan dalam penelitian kali ini merupakan data sekunder. Data yang digunakan adalah data kecepatan angin di Bandara Internasional Ahmad Yani dalam frekuensi harian yang diperoleh dari *website* NNDC *Climate Data Online*. *Climate Data Online* (CDO) merupakan data iklim dunia yang dikelola oleh *National Climatic Data Center* (NCDC) milik NOAA. Data iklim dunia yang disediakan CDO-NCDC merupakan hasil kerjasama NOAA dengan penyelenggara cuaca dan iklim di seluruh dunia. Periode data kecepatan angin yang diambil mulai awal bulan Juni 2018 hingga akhir Oktober 2019, sehingga terdapat 518 data.

Data yang sudah terkumpul dibagi menjadi dua yaitu data *in sample* untuk penentuan model dan *out sample* untuk validasi peramalan. Penelitian ini menggunakan *software* R Studio dalam pengolahan data. Tahapan analisis dalam penelitian ini adalah sebagai berikut:

1. Mempersiapkan data yang akan digunakan dalam penelitian.
2. Melakukan eksplorasi data dengan melihat plot runtun waktu data awal serta melihat statistika deskriptif dari data.
3. Memeriksa kestasioneran data dengan melihat plot data awal serta melakukan uji *Box-Cox* apabila data tidak stasioner terhadap *varians*. Selanjutnya melakukan uji *Augmented Dickey-Fuller* (ADF) untuk melihat kestasioneran dalam *mean*.
4. Identifikasi indikasi *long memory* pada data dengan melihat plot ACF. Jika pola plot ACF turun lambat atau hiperbolik, maka terdapat indikasi *long memory*. Indikasi *long memory* juga bisa diketahui dengan melihat nilai *Hurst* (H). Selain itu juga bisa dilihat dari plot periodogram dari data.
5. Pembentukan model ARFIMA dilakukan dengan metode semiparametrik. Menetapkan estimasi parameter d terlebih dahulu dengan menggunakan metode *Geweke and Porter-Hudak* (GPH) dan *Rescaled Range Statistics* (R/S). Selanjutnya penentuan beberapa model ARFIMA(p, d, q) berdasarkan plot ACF dan PACF data yang telah dilakukan *differencing* dengan nilai d yang telah diperoleh. Kemudian estimasi parameter ϕ

dan θ pada model yang sudah ditetapkan menggunakan metode *Exact Maximum Likelihood*.

6. Melakukan pemilihan model yang signifikan. Suatu parameter model dikatakan signifikan apabila nilai statistik $|z| > z$ tabel (1,96).
7. Pengujian diagnostik untuk menguji kelayakan model meliputi uji asumsi nilai residual *white noise* dan berdistribusi normal. Pengecekan *White Noise* menggunakan uji *L-Jung Box* sedangkan pengecekan distribusi normal menggunakan uji *Kolmogorov Smirnov*.
8. Pemilihan model ARFIMA yang digunakan untuk peramalan berdasarkan nilai *Akaike's Information Criterion* (AIC) terkecil.
9. Model ARFIMA yang telah ditentukan digunakan untuk evaluasi hasil peramalan. Kemudian dicari nilai erornya menggunakan *Root Mean Square Error* (RMSE), *Mean Absolute Error* (MAE), dan *Mean Absolute Percentage Error* (MAPE).

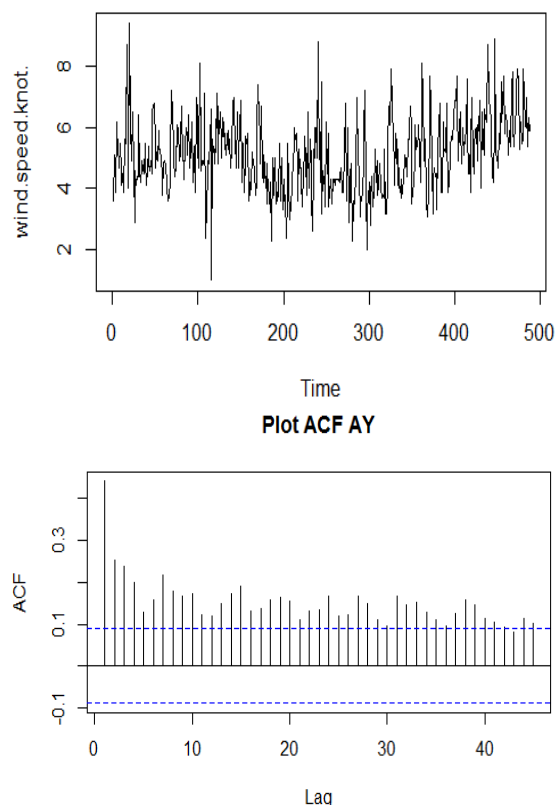


Gambar 1. Diagram Alir Penelitian
HASIL DAN PEMBAHASAN

Deskripsi Data

Data yang digunakan dalam penelitian kali ini merupakan data sekunder. Data yang digunakan adalah data kecepatan angin di Bandara Internasional Ahmad Yani dalam frekuensi harian dalam satuan knot. Periode data kecepatan angin yang diambil mulai awal

bulan Juni 2018 hingga akhir Oktober 2019, sehingga terdapat 518 data. Data dibagi menjadi dua bagian yaitu data in sample dan data out sample. Pembagian data sekitar 95% untuk data in sample dan sisanya untuk data out sample. Sebanyak 487 data in sample digunakan untuk membentuk model dan 31 data out sample untuk menguji validitas dari model yang terpilih. Berikut adalah deskripsi runtun waktu data in sample kecepatan angin di Bandara Internasional Ahmad Yani yang disajikan dalam Gambar 2. elastis pada Gambar 1.



Gambar 2. Plot Data Runtun Waktu Kecepatan Angin di Bandara Internasional Ahmad Yani

Identifikasi Long Memory

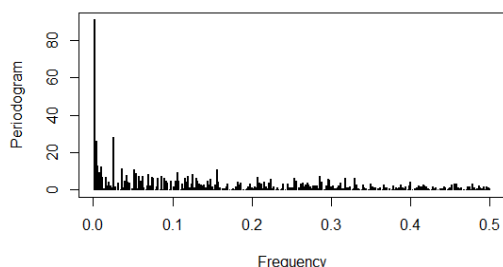
Berdasarkan Gambar 2 (b) data kecepatan angin di Bandara Internasional Ahmad Yani menunjukkan autokorelasi yang *lag*nya turun pelan menuju 0 atau tidak turun secara eksponensial, namun belum begitu terlihat turun secara hiperbolik. Sehingga *long memory* pada data masih cukup sulit untuk diidentifikasi (Rizal dkk, 2016). Selanjutnya dilakukan uji *Hurst* untuk melihat adanya indikasi *long memory*. Nilai *Hurst* diperoleh dari perhitungan menggunakan metode *Rescaled Range Statistics* (R/S). Dengan bantuan aplikasi R Studio nilai *Hurst* untuk data kecepatan angin di Bandara Internasional Ahmad Yani tersebut disajikan dalam Tabel 1.

Tabel 1. Nilai *Hurst*
Nilai

<i>Hurst</i>	0,724
--------------	-------

Berdasarkan Tabel 1 diketahui bahwa nilai *Hurst* (H) dari ketiga data berada pada selang 0,5 sampai dengan 1 ($0,5 < H < 1$). Hal ini mengindikasikan bahwa data mempunyai pola dependensi jangka panjang atau *long memory*.

Selanjutnya identifikasi pola *long memory* dengan melihat plot periodogram dari data. Plot periodogram untuk data data kecepatan angin di Bandara Internasional Ahmad Yani ditunjukkan pada Gambar 3.



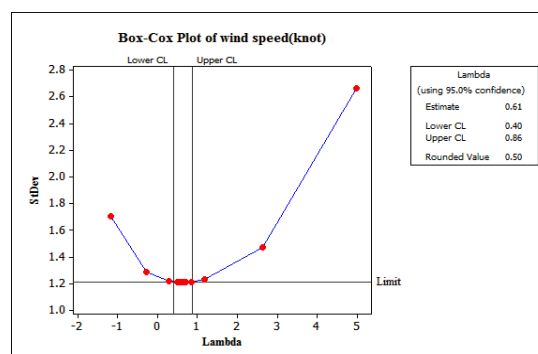
Gambar 3. Plot Periodogram

Berdasarkan Gambar 3, dapat diketahui bahwa plot periodogram untuk frekuensi yang mendekati nol meningkat menuju nilai yang sangat besar tetapi berhingga (Prafitia dkk, 2010). Oleh sebab itu, dapat diindikasikan data mempunyai pola *long memory*.

Uji Stasioneritas

Secara visual dari plot data runtun waktu kecepatan angin pada Gambar 2 (a) mengindikasikan bahwa data kecepatan angin tidak stasioner dalam varians karena fluktuasi data yang cenderung berubah dan tidak konstan. Untuk memperkuat dugaan, data kecepatan angin yang tidak stasioner terhadap varians bisa dilihat pada plot *Box-Cox* yang disajikan dalam Gambar 4.

Berdasarkan plot *Box-Cox* yang ditunjukkan pada Gambar 4, terlihat bahwa *rounded value*-nya bernilai 0,50 dengan batas bawah bernilai 0,40 dan batas atas 0,86 yang tidak melewati nilai 1 sehingga dapat dikatakan bahwa data belum stasioner terhadap varians. Untuk membuat data memenuhi asumsi stasioner terhadap varians maka dilakukan transformasi *Box-Cox* dengan menggunakan transformasi $\sqrt{Z_t}$.

Gambar 4. Plot *Box-Cox* Kecepatan Angin di Bandara Internasional Ahmad Yani

Berdasarkan plot *Box-Cox* yang ditunjukkan pada Gambar 4, terlihat bahwa *rounded value*-nya bernilai 0,50 dengan batas bawah bernilai 0,40 dan batas atas 0,86 yang tidak melewati nilai 1 sehingga dapat dikatakan bahwa data belum stasioner terhadap varians. Untuk membuat data memenuhi asumsi stasioner terhadap varians maka dilakukan transformasi *Box-Cox* dengan menggunakan transformasi $\sqrt{Z_t}$.

Analisis selanjutnya melakukan uji kestasioneran data dalam *mean*. Untuk mengetahui kestasioneran pada data dilakukan uji *Augmented Dickey-Fuller*. Uji ini didasarkan pada uji hipotesis untuk melakukan pemeriksaan terhadap kestasioneran data deret waktu. Hipotesis pengujian *Augmented Dickey-Fuller* adalah sebagai berikut :

H_0 : data *time series* tidak stasioner

H_1 : data *time series* stasioner

Hasil uji kestasioneran data transformasi kecepatan angin di Bandara Internasional Ahmad Yani dengan menggunakan *Augmented Dickey-Fuller* ditunjukkan pada Tabel 2.

Tabel 2. Hasil Uji *Augmented Dickey-Fuller*

Nilai	
Dickey-Fuller	-5,165
p-value	0,01

Berdasarkan Tabel 2, uji *Augmented Dickey-Fuller* diperoleh *p-value* 0,01 yang kurang dari taraf nyata yang digunakan sebesar 0,05 ($\alpha = 5\%$), sehingga keputusan menolak H_0 , kesimpulan menyatakan bahwa data sudah stasioner dalam *mean*.

Pembentukan Model ARFIMA

Pembentukan model ARFIMA dilakukan dengan metode semiparametrik. Menetapkan estimasi parameter d terlebih dahulu dengan menggunakan metode *Geweke and Porter-Hudak*

(GPH) dan *Rescaled Range Statistics* (R/S). Selanjutnya penentuan beberapa model ARFIMA(p, d, q) berdasarkan plot ACF dan PACF data yang telah dilakukan *differencing* dengan nilai d yang telah diperoleh. Kemudian estimasi parameter ϕ dan θ pada model yang sudah ditetapkan menggunakan metode *Exact Maximum Likelihood*.

Estimasi Parameter d

Nilai parameter *differencing* d ditentukan dengan metode *Geweke Porter-Hudak* (GPH) dan *Rescaled Range Statistics* (R/S). Estimasi parameter d dengan metode estimasi GPH dan R/S dihitung dengan bantuan aplikasi R Studio. Parameter d dari data kecepatan angin di Bandara Internasional Ahmad Yani ditunjukkan pada Tabel 3 sebagai berikut.

Tabel 3. Parameter d

	Estimasi Parameter d
<i>Geweke and Porter-Hudak</i> (GPH)	0,346
<i>Range Scaled Statistics</i> (R/S)	0,224

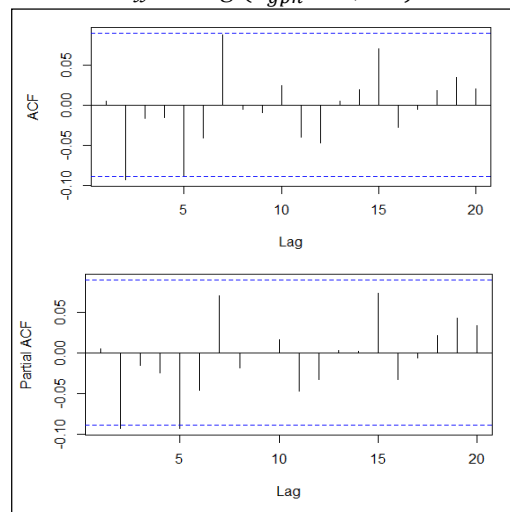
Dari hasil estimasi parameter d yang ditunjukkan pada Tabel 3, menunjukkan bahwa data mempunyai nilai d yang terletak pada selang $0 < d < 0.5$, menyatakan bahwa proses berkorelasi panjang stasioner dengan adanya ketergantungan positif antar pengamatan yang terpisah jauh yang ditunjukkan dengan autokorelasi positif dan turun lambat serta mempunyai representasi *moving average* orde tak hingga.

Identifikasi Beberapa Model ARFIMA (p, d, q) Berdasarkan Plot ACF dan Plot PACF

Identifikasi Model ARFIMA(p, d, q) untuk masing-masing data ditetapkan berdasarkan plot ACF dan PACF setelah *didifferencing* sesuai nilai d dari perhitungan menggunakan estimasi GPH dan R/S. Penentuan beberapa model ARFIMA(p, d, q) ditentukan dengan melihat *lag* yang melewati batas dari plot ACF dan PACF.

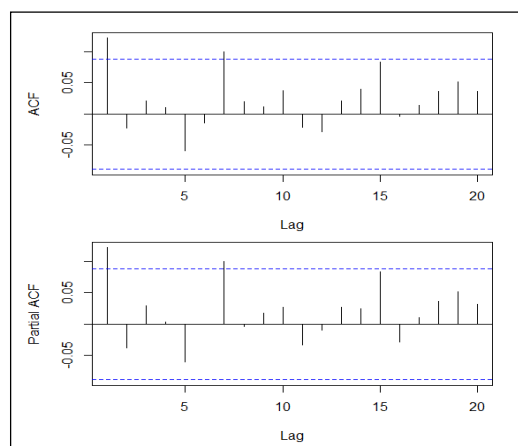
Plot ACF dan PACF setelah *didifferencing* dengan $\hat{d}_{gph} = 0,346$ dari perhitungan menggunakan estimasi GPH ditunjukkan pada Gambar 5.

Gambar 5. Plot ACF dan Plot PACF
Differencing ($\hat{d}_{gph} = 0,346$)



Berdasarkan Gambar 5 pada plot ACF *lag* yang terlihat jelas melewati batas adalah *lag* ke-2 sehingga dapat diidentifikasi model MA yang terbentuk adalah MA([2]). Kemudian berdasarkan plot PACF *lag* yang melewati batas adalah *lag* ke-2 dan *lag* ke-5, sehingga dapat diidentifikasi model AR yang terbentuk adalah AR([2]), AR([5]), dan AR([2,5]). Dengan demikian model tentatif ARFIMA dengan $\hat{d}_{gph} = 0,346$ yang dapat terbentuk adalah ARFIMA([2], $d, 0$), ARFIMA([5], $d, 0$), ARFIMA([2,5], $d, 0$), ARFIMA(0, $d, [2]$), ARFIMA([2], $d, [2]$), ARFIMA([5], $d, [2]$), dan ARFIMA([2,5], $d, [2]$).

Plot ACF dan PACF setelah *didifferencing* dengan $\hat{d}_{R/S} = 0,224$ dari perhitungan menggunakan estimasi R/S ditunjukkan pada Gambar 6.



Gambar 6. Plot ACF dan Plot PACF
Differencing ($\hat{d}_{R/S} = 0,224$)

Berdasarkan Gambar 6, model tentatif ARFIMA yang dapat terbentuk dengan $\hat{d}_{R/S} = 0,177$ adalah ARFIMA(1,d,0), ARFIMA([7],d,0), ARFIMA([1,7],d,0), ARFIMA(1,d,1), ARFIMA([7],d,1), ARFIMA([1,7],d,1), ARFIMA(1,d,[7]), ARFIMA([7],d,[7]), ARFIMA([1,7],d,[7]), ARFIMA(1,d,[1,7]), ARFIMA([7],d,[1,7]), ARFIMA([1,7],d,[1,7]), ARFIMA(0,d,1), ARFIMA(0,d,[7]), dan ARFIMA(0,d,[1,7]).

Estimasi Parameter Model Tentatif ARFIMA(p, d, q)

Dari hasil estimasi parameter, dipilih model yang signifikan yaitu model yang memiliki nilai statistik $|z| > z$ tabel (1,96). Dari estimasi model tentatif ARFIMA(p, d, q) dengan $\hat{d}_{gph} = 0,346$ yang ditentukan, model yang menghasilkan semua parameter yang signifikan adalah model ARFIMA([2],d,0), ARFIMA([5],d,0), ARFIMA([2,5],d,0), dan ARFIMA(0,d,[2]). Sedangkan dari estimasi model tentatif ARFIMA(p, d, q) dengan $\hat{d}_{R/S} = 0,224$ yang ditentukan, model yang menghasilkan semua parameter yang signifikan adalah model ARFIMA(1,d,0), ARFIMA([7],d,0), ARFIMA([1,7],d,0), ARFIMA([7],d,1), ARFIMA(1,d,[7]), ARFIMA(0,d,1), ARFIMA(0,d,[7]), dan ARFIMA(0,d,[1,7]).

Uji Diagnostik Model

Dari model signifikan yang telah diperoleh, selanjutnya dilakukan uji diagnostik. Hasil uji diagnostik pada model ARFIMA(p, d, q) dengan $\hat{d}_{gph} = 0,346$ yang signifikan, memberikan hasil bahwa residual model-model tersebut dinyatakan *white noise*. Demikian juga hasil uji diagnostik pada model ARFIMA(p, d, q) dengan $\hat{d}_{R/S} = 0,224$ yang signifikan, memberikan hasil bahwa model-model tersebut dinyatakan *white noise*.

Pemilihan Model Terbaik

Pemilihan model terbaik ARFIMA ditentukan berdasarkan kriteria nilai *Akaike's Information Criterion* (AIC) terkecil. Nilai AIC untuk model ARFIMA disajikan pada Tabel 4.

Berdasarkan Tabel 4, untuk model ARFIMA(p, d, q) dengan $\hat{d}_{gph} = 0,346$ model yang terpilih adalah diperoleh model ARFIMA(0,d,[2]) dengan nilai AIC terkecil yaitu -1384,527. Untuk model ARFIMA(p, d, q) dengan $\hat{d}_{R/S} = 0,224$ model yang terpilih adalah model ARFIMA(0,d,1) dengan nilai AIC terkecil yaitu -1385,629. Model-model terpilih tersebut selanjutnya digunakan akan digunakan untuk proses peramalan.

Tabel 4. Nilai AIC Model ARFIMA

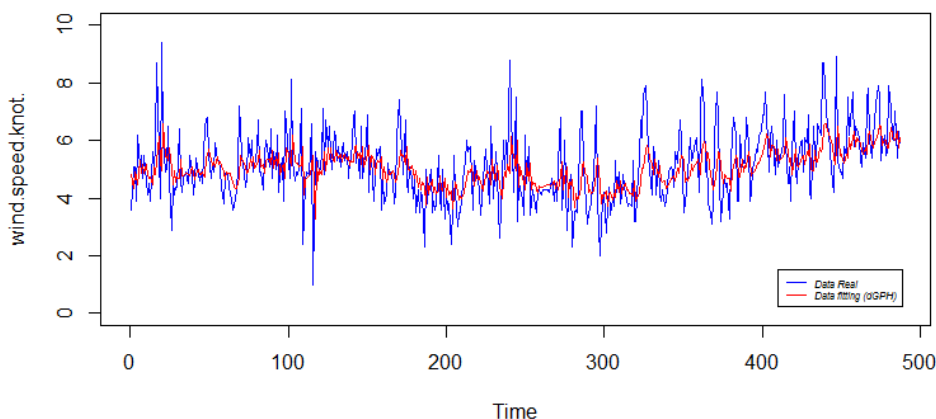
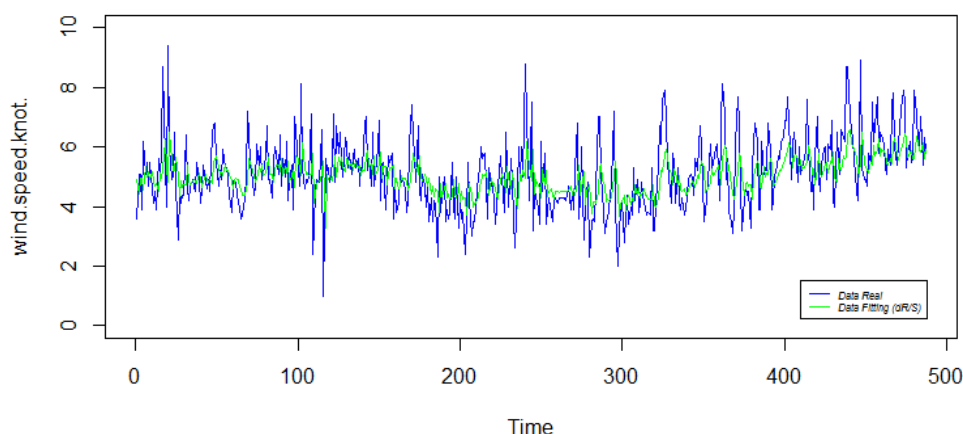
Estimasi Parameter d	Model ARFIMA	AIC
$\hat{d}_{gph} = 0,346$	ARFIMA([2], d,0)	-1384,319
	ARFIMA([5], d,0)	-1377,984
	ARFIMA([2,5], d,0)	-1382,458
	ARFIMA(0, d,[2])	-1384,527
$\hat{d}_{R/S} = 0,224$	ARFIMA(1, d,0)	-1385,107
	ARFIMA([7], d,0)	-1370,539
	ARFIMA([1,7], d,0)	-1378,198
	ARFIMA([7], d,1)	-1376,641
	ARFIMA(1, d,[7])	-1375,941
	ARFIMA(0, d,1)	-1385,629
	ARFIMA(0, d,[7])	-1370,200
	ARFIMA(0, d,[1,7])	-1378,464

Peramalan

Model ARFIMA yang terpilih selanjutnya digunakan untuk melakukan penyesuaian (*fitting*) dengan data *in sample*. Pada penelitian ini digunakan *Root Mean Square Error* (RMSE), *Mean Absolute Error* (MAE) dan *Mean Absolute Percentage Error* (MAPE) untuk menghitung akurasi model yang signifikan. Hasil perhitungan tingkat akurasi data *fitting* disajikan dalam Tabel 5. Plot data aktual dan data *fitting* menggunakan model terpilih ARFIMA(0,d,[2]) dengan $\hat{d}_{gph} = 0,346$ untuk data kecepatan angin di Bandara Internasional Ahmad Yani disajikan dalam Gambar 7. Plot data aktual dan data *fitting* menggunakan model terpilih ARFIMA(0,d,1) dengan $\hat{d}_{R/S} = 0,224$ untuk data kecepatan angin di Bandara Internasional Ahmad Yani disajikan dalam Gambar 8.

Tabel 5. Nilai Akurasi *Fitting* Model ARFIMA

Estimasi Parameter d	Model ARFIMA	RMSE	MAE	MAPE
$\hat{d}_{gph} = 0,346$	ARFIMA(0, d , [2])	1,072906	0,8101154	17,20635
$\hat{d}_{R/S} = 0,224$	ARFIMA(0, d , 1)	1,074054	0,8123931	17,30652

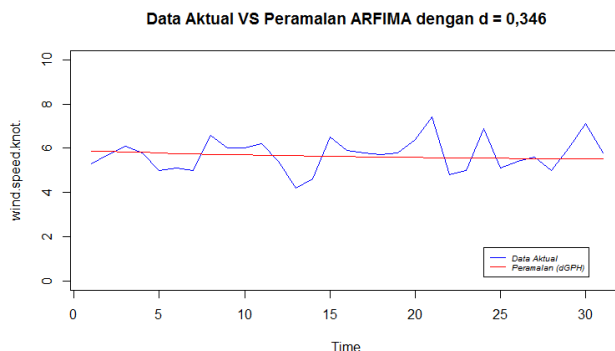
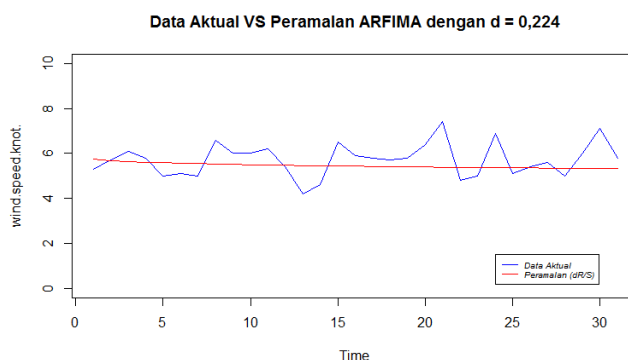
Data Real VS Data Fitting ARFIMA dengan $d = 0,346$ Gambar 7. Plot Data *Fitting* ARFIMA(0, d , [2]) dengan $\hat{d}_{gph} = 0,346$ Data Real VS Data Fitting ARFIMA dengan $d = 0,224$ Gambar 8. Plot Data *Fitting* ARFIMA(0, d , 1) dengan $\hat{d}_{R/S} = 0,224$

Berdasarkan Tabel 5, nilai error pada *fitting* model ARFIMA (0, d , [2]) dengan $\hat{d}_{gph} = 0,346$ adalah RMSE 1,072906, MAE 0,8101154, dan MAPE 17,20635. Nilai error pada *fitting* model ARFIMA(0, d , 1) dengan $\hat{d}_{R/S} = 0,224$ adalah RMSE 1,074054, MAE 0,8123931, dan MAPE 17,30652. Nilai MAPE pada kedua model menunjukkan kriteria tingkat kemampuan peramalan yang baik. Nilai error pada *fitting* model ARFIMA (0, d , [2]) dengan $\hat{d}_{gph} = 0,346$

menunjukkan error yang lebih kecil dibanding model ARFIMA(0, d , 1) dengan $\hat{d}_{R/S} = 0,224$. Selanjutnya model ARFIMA yang terpilih akan digunakan dalam peramalan dan validasi pada data *out sample*. Nilai perhitungan nilai akurasi peramalan kecepatan angin di Bandara Internasional Ahmad Yani disajikan dalam Tabel 6. Plot dari data aktual dan ramalan menggunakan model terpilih ARFIMA (0, d , [2]) dengan $\hat{d}_{gph} = 0,346$ disajikan pada

Tabel 6. Nilai Akurasi Peramalan Model ARFIMA

Estimasi Parameter d	Model ARFIMA	RMSE	MAE	MAPE
$\hat{d}_{gph} = 0,346$	ARFIMA(0, d , [2])	0,74966754	0,593468	10,45511
$\hat{d}_{R/S} = 0,224$	ARFIMA(0, d , 1)	0,7898162	0,6295811	10,75531

Gambar 9. Plot Data Aktual dan Ramalan Model ARFIMA(0, d , [2]) dengan $\hat{d}_{gph} = 0,346$ Gambar 10. Plot Data Aktual dan Ramalan Model ARFIMA(0, d , 1) dengan $\hat{d}_{R/S} = 0,224$

Berdasarkan Tabel 6, nilai eror validasi peramalan model ARFIMA (0, d , [2]) dengan $\hat{d}_{gph} = 0,346$ adalah RMSE 0,74966754, MAE 0,593468, dan MAPE 10,45511. Nilai eror pada validasi peramalan model ARFIMA(0, d , 1) dengan $\hat{d}_{R/S} = 0,224$ adalah RMSE 0,7898162, MAE 0,6295811, dan MAPE 10,75531. Nilai MAPE pada kedua model menunjukkan kriteria tingkat kemampuan peramalan yang baik. Nilai eror pada validasi peramalan model ARFIMA (0, d , [2]) dengan $\hat{d}_{gph} = 0,346$ menunjukkan eror yang lebih kecil dibanding model ARFIMA(0, d , 1) dengan $\hat{d}_{R/S} = 0,224$.

Perbandingan dengan Model Lain

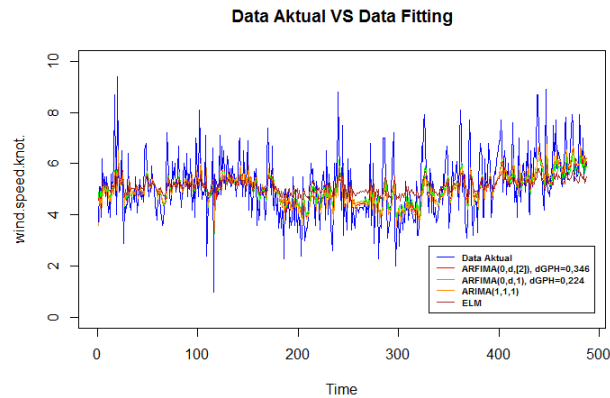
Dalam Penelitian ini penulis menambahkan model lain untuk perbandingan dengan $\hat{d}_{R/S} = 0,224$ disajikan pada Gambar 10. dan sebagai

pelengkap dalam penelitian. Penulis menggunakan model *Autoregressive Integrated Moving Average* (ARIMA) dan *Extreme Learning Machine* (ELM) sebagai komparasi atau perbandingan. Dalam perhitungan penulis menggunakan fungsi yang tersedia dalam program R Studio. Hasil perhitungan tingkat akurasi data *fitting* dengan model ARIMA dan ELM ditunjukkan pada Tabel 7. Plot dari data aktual dan data *fitting* menggunakan model menggunakan model ARFIMA(0, d , [2]) dengan $\hat{d}_{gph} = 0,346$, ARFIMA(0, d , 1) dengan $\hat{d}_{R/S} = 0,224$, ARIMA(1,1,1), dan model ELM disajikan pada Gambar 11.

Berdasarkan Tabel 7, nilai akurasi model ARIMA(1,1,1) untuk *fitting* data kecepatan angin di Bandara Internasional Ahmad Yani menunjukkan nilai eror yang paling kecil dibanding model yang lain.

Tabel 7. Nilai Akurasi *Fitting Model* ARFIMA, ARIMA, dan ELM

Model	RMSE	MAE	MAPE
ARFIMA(0, d , [2]), $\hat{d}_{gph} = 0,346$	1,072906	0,8101154	17,20635
ARFIMA(0, d , 1), $\hat{d}_{R/S} = 0,224$	1,074054	0,8123931	17,30652
ARIMA(1,1,1)	1,069348	0,802783	16,99219
ELM	1,121014	0,874623	19,0393

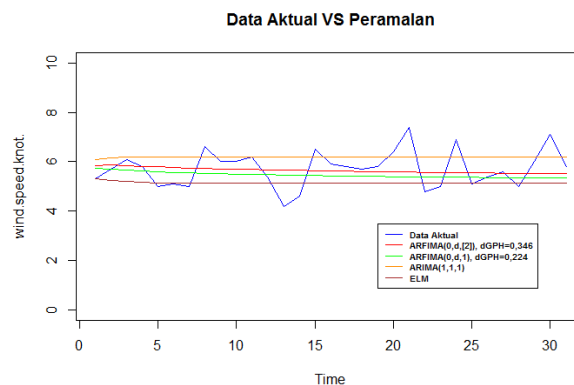
Gambar 11. Plot Data Aktual dan *Fitting Model* ARFIMA, ARIMA, dan ELM

Selanjutnya akan dilihat nilai akurasi hasil peramalan untuk masing-masing model. Nilai akurasi hasil peramalan dari masing-masing model disajikan pada Tabel 8. Plot dari data aktual dan ramalan menggunakan model

ARFIMA(0, d , [2]) dengan $\hat{d}_{gph} = 0,346$,
 ARFIMA(0, d , 1) dengan $\hat{d}_{R/S} = 0,224$,
 ARIMA(1,1,1), dan model ELM disajikan pada Gambar 12.

Tabel 8. Nilai Akurasi Peramalan Model ARFIMA, ARIMA, dan ELM

Model ARFIMA	RMSE	MAE	MAPE
ARFIMA(0, d , [2]), $\hat{d}_{gph} = 0,346$	0,74966754	0,593468	10,45511
ARFIMA(0, d , 1), $\hat{d}_{R/S} = 0,224$	0,7898162	0,6295811	10,75531
ARIMA(1,1,1)	0,8605522	0,7110627	13,41102
ELM	0,9197016	0,7235949	11,90351



Gambar 12. Plot Data Aktual dan Ramalan Model ARFIMA, ARIMA, dan ELM

Berdasarkan Tabel 8, nilai akurasi model ARFIMA(0, d , [2]) dengan $\hat{d}_{gph} = 0,346$ untuk peramalan data kecepatan angin di Bandara Internasional Ahmad Yani menunjukkan nilai

error yang paling kecil dibanding model lebih kecil dibanding model ARFIMA(0, d , 1) dengan $\hat{d}_{R/S} = 0,224$, ARIMA(1,1,1), dan model ELM.

PENUTUP

Hasil estimasi parameter d pada ARFIMA untuk data kecepatan angin di Bandara Internasional Ahmad Yani dengan metode *Geweke and Porter-Hudak* (GPH) dan *Rescaled Range Statistics* (R/S), memberikan hasil yaitu $\hat{d}_{gph} = 0,346$ dan $\hat{d}_{R/S} = 0,224$. Model terbaik untuk ARFIMA(p, d, q) dengan $\hat{d}_{gph} = 0,346$ adalah model ARFIMA(0, d , [2]) dengan nilai AIC -1384,527. Model terbaik untuk ARFIMA(p, d, q) dengan $\hat{d}_{R/S} = 0,224$ adalah model ARFIMA(0, d , 1) dengan nilai AIC -1385,629.

Hasil dari data *fitting* pada model cukup mengikuti pola pada data aktual. Nilai error pada *fitting* model ARFIMA (0, d , [2]) dengan $\hat{d}_{gph} = 0,346$ yaitu RMSE 1,072906, MAE 0,8101154, dan MAPE 17,20635 lebih kecil dibanding model ARFIMA(0, d , 1) dengan $\hat{d}_{R/S} = 0,224$. Nilai error pada validasi peramalan model ARFIMA(0, d , [2]) dengan $\hat{d}_{gph} = 0,346$ yaitu RMSE 0,74966754, MAE 0,593468, dan MAPE 10,45511 lebih kecil dibanding model ARFIMA(0, d , 1) dengan $\hat{d}_{R/S} = 0,224$.

Hasil komparasi dengan model ARIMA dan ELM memberikan hasil bahwa, nilai error dari *fitting* model ARIMA(1,1,1) yaitu RMSE 1,069348, MAE 0,802783, dan MAPE 16,99219 lebih kecil dibanding model ARFIMA(0, d , [2]) dengan $\hat{d}_{gph} = 0,346$, ARFIMA(0, d , 1) dengan $\hat{d}_{R/S} = 0,224$, dan model ELM. Sedangkan nilai error validasi peramalan model ARFIMA(0, d , [2]) dengan $\hat{d}_{gph} = 0,346$ yaitu RMSE 0,74966754, MAE 0,593468, dan MAPE 10,45511 lebih kecil dibanding model ARFIMA(0, d , 1) dengan $\hat{d}_{R/S} = 0,224$, ARIMA(1,1,1), dan model ELM.

DAFTAR PUSTAKA

- Beran, J. 1994. *Statistics for Long-Memory Processes: Monographs on statistics and applied probability*. New York: Chapman & Hall.
- Brockwell, P. J. dan Davis, R. A. 2016. *Introduction to Time Series and Forecasting (Third Edition)*. Switzerland: Springer.
- Geweke, J dan Porter Hudak, S. 1983. The Estimation and Application of Long Memory Time Series Models. *Journal of Time Series Analysis*, 4: 221 –238.
- Hurst, H.E. 1951. Long-Term Storage of Reservoirs: An Experimental Study. *Transaction on the American Society Civil Engineers*, (116): 770 – 799.
- Hosking, J. R. M. 1981. Fractional Differencing. *Biometrika*, 68(1): 165 – 176.
- Ibrahim, L. K., Asare, B. K., Yakubu, M. dan Dauda, U. 2011. Modeling of Sokoto Daily Average Temperature: A Fractional Integrated Approach. *Nigerian Journal of Basic and Applied Science* (2011), 19(1): 21 – 30.
- Kharisudin, I., Rosadi, D., Abdurakhman, dan Suhartono. 2012. Model Long-Memory FARIMA dan Aplikasinya Pada Pemodelan Data Asset Returns di Indonesia. *KNM XVI UNPAD*: 705 – 712.
- Mohamed, R. A. H. 2016. Using ARFIMA Models in Forecasting The Total Value of Traded Securities on The Arab Republic of Egypt. *IJRRAS*, 27(1): 26 – 34.
- Moulines, E. dan Soulier, P. 1999. Broadband Log-Periodogram Regression of Time Series with Longrange dependence. *Annals of Statistics*, 27(4): 1415 – 1439.
- Paul, R. K. 2017. Modelling Long Memory in Maximum and Minimum Temperature Series in India. *MAUSAM*, 68(2): 317 – 326.
- Prafitia, H. A. dkk. 2010. Long Memory pada Data Nilai Tukar Rupiah terhadap Dollar Amerika Serikat. *Jurnal*. Surabaya: ITS.
- Rizal, J., Nugroho, S., Irwanto, A. dan Debora. 2016. Analisis Kejadian Gempa Bumi Tektonik di wilayah Pulau Sumatera. *Jurnal Matematika*, 6(1): 1 – 14.
- Safitri, D., Mustafid, Ispriyanti, D., dan Sugito. 2019. Gold Price Modeling in Indonesia using ARFIMA Method, *IOP Conf. Series: Journal of Physics: Conf Series*: 1 – 11.

Tjasyono. 2004. *Klimatologi*. Bandung: ITB.

Walid, Subanar, Rosadi, D., dan Suhartono. 2015. Identification Model of Long Memory in Use Electricity Charges in Java-Bali. *International Conference on Mathematics, Science, and Education 2015 (ICMSE 2015)*:13 – 23.

Wei, W. W. S. 2006. *Time Series Analysis Univariate and Multivariate Methods Second Edition*. New York: Pearson Addison Wesley.