

Peramalan Curah Hujan Harian di Stasiun Ahmad Yani Kota Semarang Menggunakan *Adaptive Neuro Fuzzy Inference System* (ANFIS)

Rr. Sekar Kalaksita dan Irhamah

Jurusan Statistika, Fakultas MIPA, Institut Teknologi Sepuluh Nopember (ITS)

Jl. Arief Rahman Hakim, Surabaya 60111 Indonesia

e-mail: sekarkalaksita@gmail.com dan irhamahn@gmail.com

Abstrak—Kota Semarang sebagai ibukota Jawa Tengah merupakan salah satu kota yang berkembang di Pulau Jawa dan menjadi kota perdagangan. Dalam beberapa tahun terakhir, Kota Semarang menghadapi permasalahan yang cukup sulit, yaitu setiap tahun mengalami bencana banjir dikarenakan letaknya yang dilalui oleh jalur pantura dan merupakan kawasan rob. Perlu adanya peramalan curah hujan harian dengan akurasi yang tinggi untuk mengurangi dampak negatif seperti banjir. Metode yang digunakan untuk meramalkan jumlah curah hujan periode ke depan di Stasiun Ahmad Yani Kota Semarang adalah ARIMA dan ANFIS. Model ARIMA yang terpilih adalah ARIMA (1,1,1) dan ARIMA ([1,60],1,[1,31]) sedangkan dengan pendekatan ANFIS, *input* yang terpilih adalah jumlah curah hujan harian satu hari dan dua hari sebelumnya dengan fungsi keanggotaan *Trapezoidal* dan banyaknya MF sebesar 2. Metode ANFIS menunjukkan kinerja akurasi peramalan yang lebih baik daripada ARIMA karena menghasilkan RMSE dan MAPE yang lebih kecil.

Kata Kunci—ANFIS, ARIMA, Curah Hujan, Harian, MAPE, RMSE.

I. PENDAHULUAN

Kota Semarang sebagai ibukota Jawa Tengah merupakan kota metropolitan kelima di Indonesia setelah Jakarta, Surabaya, Bandung, dan Medan dengan 51,71 persen dari luas wilayahnya berisi perairan. Kota Semarang juga merupakan salah satu kota yang berkembang di Pulau Jawa dan menjadi kota perdagangan. Sektor transportasi juga turut men-dorong pertumbuhan dan perkembangan di Kota Semarang berjalan dengan cepat. Seiring dengan laju pembangunan Kota Semarang, pertumbuhan dan perkembangan kota telah menyebabkan perubahan tata guna lahan. Dalam beberapa tahun terakhir ini, Kota Semarang menghadapi permasalahan yang cukup sulit, yaitu setiap tahun selalu mengalami bencana banjir.

Pada tahun 2013 dan 2014, banjir di Kota Semarang telah mengacaukan seluruh transportasi lintas Jawa. Jalur Pantura telah terputus selama beberapa hari dikarenakan banjir. Selain itu, banjir juga menggenangi rel kereta api bahkan hingga masuk ke peron Stasiun Semarang dan menyebabkan arus lalu lintas kereta terhenti total. Pada awal tahun 2016, banjir setinggi 20 hingga 50 sentimeter telah merendam beberapa kawasan di Kota Semarang, seperti daerah Tawang Sari dan Kaligawe [1]. Peningkatan dan penurunan jumlah curah hujan yang tak menentu tersebut diakibatkan karena letak Kota Semarang dilalui

oleh jalur pantai utara yang sering dilanda banjir ketika musim penghujan dan merupakan kawasan rob yaitu kawasan yang terkena banjir akibat pasang naik air laut yang menggenangi daratan dan umumnya terjadi di daerah yang lebih rendah dari permukaan air laut [2]. Daerah yang terkena banjir rob adalah dataran pantai di daerah pesisir yang rendah atau daerah rawa-rawa pantai. Genangan banjir ini dapat diper-kuat dengan jumlah curah hujan yang tinggi. Maka, dapat dikatakan bahwa banjir ini dapat terjadi lebih hebat di saat musim hujan [3]. Oleh sebab itu, perlu adanya peramalan dengan tingkat akurasi tinggi agar dapat mengurangi dampak negatif yang ditimbulkan oleh intensitas curah hujan seperti longsor, hujan badai, hingga banjir.

Stasiun Ahmad Yani merupakan salah satu stasiun yang bertugas sebagai pengamat cuaca, salah satunya adalah curah hujan. Beberapa metode peramalan yang dapat digunakan pada curah hujan adalah ARIMA dan *Adaptive Neuro Fuzzy Inference System* (ANFIS). Penggabungan model ARIMA dan ANFIS ini dikarenakan dalam dunia nyata jarang ditemukan kejadian *time series* yang murni linier ataupun murni non linier. Oleh karena itu, model ARIMA digunakan untuk menyelesaikan kasus *time series* yang linier, sedangkan metode ANFIS digunakan untuk menyelesaikan kasus *time series* yang nonlinier. Beberapa penelitian mengenai ANFIS pernah dilaku-kan oleh Dewi [4] dan Faulina [5] yang menghasilkan keputus-an bahwa ANFIS memberikan hasil ramalan dengan akurasi lebih tinggi karena menunjukkan RMSE yang lebih kecil dibanding ARIMA. Oleh sebab itu, permasalahan yang dibahas dalam penelitian ini adalah meramalkan curah hujan Kota Semarang dengan stasiun pengamatan yang digunakan adalah Stasiun Ahmad Yani untuk beberapa periode ke depan menggunakan metode ARIMA dan *Adaptive Neuro Fuzzy Inference System* (ANFIS).

II. TINJAUAN PUSTAKA

A. Autoregressive Integrated Moving Average (ARIMA)

Model ARIMA merupakan model yang sangat kuat dalam peramalan jangka pendek [6]. Secara umum, model ARIMA dituliskan dengan notasi ARIMA(p,d,q) sebagai berikut.

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

Ada beberapa tahap dalam pembentukan model ARIMA. Pertama adalah mengidentifikasi model dengan cara melihat plot data. Dengan memandang suatu pengamatan Z_1, Z_2, \dots, Z_n sebagai suatu proses stokastik, maka variabel random $Z_{t_1}, Z_{t_2}, \dots, Z_{t_n}$ dikatakan stasioner apabila [7]:

$$F(Z_{t_1}, Z_{t_2}, \dots, Z_{t_n}) = F(Z_{t_1+k}, Z_{t_2+k}, \dots, Z_{t_n+k}) \quad (2)$$

Oleh sebab itu, dalam model peramalan menghendaki data yang stasioner baik dalam *mean* maupun varians. Apabila data belum stasioner dalam varians, maka perlu melakukan transformasi *Box-Cox* pada data dan apabila belum stasioner dalam *mean* maka perlu melakukan *differencing*. Setelah stasioneritas telah terpenuhi, maka dapat dilakukan penentuan model ARIMA sementara dengan melihat nilai ACF dan PACF. Fungsi Autokorelasi (ACF) merupakan suatu koefisien yang menunjukkan hubungan linier pada data *time series* antara Z_t dan Z_{t+k} dimana $k=0,1,2,\dots$ yang dirumuskan sebagai berikut [7].

$$\rho_k = \frac{cov(Z_t, Z_{t+k})}{\sqrt{var(Z_t)}\sqrt{var(Z_{t+k})}} = \frac{\gamma_k}{\gamma_0} \quad (3)$$

sedangkan Fungsi Autokorelasi Parsial (PACF) digunakan un-tuk mengukur tingkat keeratan hubungan pada data *time series* antara Z_t dan Z_{t+k} setelah pengaruh linier dari $Z_{t+1}, Z_{t+2}, \dots, Z_{t+k-1}$ telah dihilangkan yang dirumuskan sebagai berikut.

$$\phi_{kk} = \frac{Cov[(Z_t - \hat{Z}_t), (Z_{t+k} - \hat{Z}_{t+k})]}{\sqrt{Var(Z_t - \hat{Z}_t)}\sqrt{Var(Z_{t+k} - \hat{Z}_{t+k})}} \quad (4)$$

Tahapan kedua adalah melakukan estimasi parameter. Dalam penelitian ini, metode estimasi parameter yang digunakan adalah *Conditional Least Square Estimation*. Setelah dilakukan estimasi parameter, dilakukan uji signifikansi parameter dengan hipotesis sebagai berikut.

$$H_0 : \phi_p = 0 \text{ atau } \theta_q = 0$$

$$H_1 : \phi_p \neq 0 \text{ atau } \theta_q \neq 0$$

dengan statistik uji yang digunakan adalah sebagai berikut.

$$t = \frac{\hat{\phi}_p}{SE(\hat{\phi}_p)} = 0 \text{ atau } t = \frac{\hat{\theta}_q}{SE(\hat{\theta}_q)} \quad (5)$$

Apabila $|t| > t_{\alpha/2, n-p}$ atau $p\text{-value} < \alpha$ dimana n merupakan banyaknya observasi dan p merupakan banyaknya parameter yang ditaksir maka disimpulkan tolak H_0 atau dapat dikatakan bahwa parameter ϕ dan θ dalam model sudah signifikan [7].

Selanjutnya untuk mendapatkan model ARIMA terbaik, perlu memenuhi asumsi terhadap residual yaitu *white noise* dan distribusi normal. Pemenuhan asumsi *white noise* menggunakan uji *Ljung Box* dengan hipotesis sebagai berikut.

$$H_0 : \rho_1 = \rho_2 = \dots = \rho_K = 0$$

H_1 : minimal terdapat satu $\rho_k \neq 0$, dimana $k = 1, 2, \dots, K$ Statistik uji yang digunakan adalah

$$Q = n(n+2) \sum_{k=1}^K (n-k)^{-1} \hat{\rho}_k^2 \quad (6)$$

dengan n adalah banyak pengamatan dan $\hat{\rho}_k^2$ adalah residual sampel ACF pada lag ke- k . Apabila nilai $Q > X_{\alpha, K-m}^2$ atau $p\text{-value} < \alpha$ dimana m merupakan jumlah parameter orde ARMA maka didapatkan keputusan tolak H_0 atau dapat disimpulkan bahwa residual tidak memenuhi asumsi *white noise* [7].

Pemenuhan asumsi residual berdistribusi normal menggunakan uji *Kolmogorov-Smirnov* dengan hipotesis berikut.

H_0 : residual data mengikuti distribusi normal

H_1 : residual data tidak mengikuti distribusi normal

Statistik uji yang digunakan adalah

$$D = \max|D^+, D^-|, \quad (7)$$

keterangan:

$$D^+ = \max_i \left\{ \frac{i}{n} - Z_{(i)} \right\},$$

$$D^- = \max_i \left\{ Z_{(i)} - \frac{(i-1)}{n} \right\},$$

$$Z_{(i)} = F(x_i),$$

$F(x)$ adalah fungsi distribusi probabilitas dari distribusi normal,

x_i adalah sampel random dari orde ke- i , $1 \leq i \leq n$,

n adalah jumlah sampel.

Apabila nilai $|D| > D_{(1-\alpha)}$ atau $p\text{-value} < \alpha$ maka dikatakan tolak H_0 atau dapat disimpulkan bahwa residual tidak mengikuti distribusi normal [8].

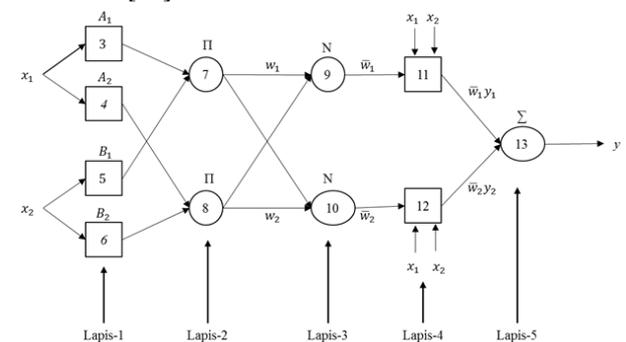
B. Adaptive Neuro Fuzzy Inference Systems (ANFIS)

Dalam pemodelan ANFIS, dikenal dengan istilah himpunan *fuzzy* dan sistem inferensi *fuzzy* [9]. Himpunan *fuzzy* merupakan himpunan dimana keanggotaan dari tiap elemennya memiliki batas yang tidak jelas. Nilai keanggotaan yang dinotasikan dengan $\mu_A(x)$ merupakan suatu nilai yang menunjukkan seberapa besar tingkat keanggotaan suatu elemen (x) dalam suatu himpunan (A). Dalam himpunan *fuzzy*, ruang *input* yang diberikan dipetakan menjadi nilai atau derajat keanggotaan (bobot) antara 0 dan 1 melalui suatu fungsi yang disebut fungsi keanggotaan [9]. Salah satu cara agar dapat digunakan untuk mendapatkan nilai keanggotaan adalah dengan melalui pendekatan fungsi seperti kurva trapesium, kurva gauss, dan kurva bentuk lonceng [9]. Sistem inferensi *fuzzy* merupakan suatu kerangka komputasi yang didasarkan oleh teori himpunan *fuzzy*, aturan *fuzzy* berbentuk IF-THEN, dan penalaran *fuzzy* [9]. Dalam penelitian ini, sistem inferensi *fuzzy* yang digunakan adalah model *fuzzy* Sugeno (TSK) orde satu karena memiliki karakteristik yaitu konsekuen tidak merupakan himpunan *fuzzy*, namun merupakan suatu persamaan linier dengan variabel sesuai dengan variabel-variabel *input*nya. ANFIS juga memungkinkan aturan-aturan untuk beradaptasi. Misalkan terdapat 2 *input* (x_1, x_2) dan satu *output* (y). Ada 2 aturan dibasis aturan model Sugeno sebagai berikut [9].

Aturan 1: *if* x_1 is A_1 and x_2 is B_1 , then $y_1 = C_{11}x_1 + C_{12}x_2 + C_{10}$

Aturan 2: *if* x_1 is A_2 and x_2 is B_2 , then $y_2 = C_{21}x_1 + C_{22}x_2 + C_{20}$

Arsitektur ANFIS terdiri atas lima lapisan dan setiap lapisan terdapat dua macam *node* yang dapat dilihat di Gambar 1 [10].



Gambar 1. Arsitektur ANFIS

Lapis ke-1 (Fuzzyfikasi): Setiap *node* i dari lapis ini adalah *node* yang bersifat adaptif, misal $x_1 = Z_{t-1}$ dan $x_2 = Z_{t-2}$, dengan fungsi sebagai berikut.

$$\begin{aligned} O_{1,i} &= \mu_{A_i}(Z_{t-1}), \\ O_{1,i} &= \mu_{B_i}(Z_{t-2}), \end{aligned} \quad (8)$$

misal menggunakan fungsi bentuk lonceng, fungsi umumnya adalah sebagai berikut.

$$\mu_A(Z_{t-1}) = \frac{1}{1 + \left| \frac{Z_{t-1} - c}{a} \right|^{2b}}, \quad (9)$$

dengan a , b , dan c merupakan parameter premis.

Lapis ke-2 (Operasi Logika Fuzzy): Setiap *node* dari lapis ini adalah *node* tetap berlabel Π dengan keluarannya adalah produk dari semua sinyal yang datang.

$$O_{2,i} = w_i = \mu_{A_i}(Z_{t-1})\mu_{B_i}(Z_{t-2}), i = 1,2. \quad (10)$$

Lapis ke-3 (Normalized Firing Strength): Setiap *node* dari lapis ini merupakan *node* tetap berlabel N . *Output* dari lapisan ini disebut kuat pengaktifan ternormalisasi.

$$O_{3,i} = \bar{w}_i = \frac{w_i}{w_1 + w_2}, i = 1,2. \quad (11)$$

Lapis ke-4 (Defuzzifikasi): Setiap *node* dari lapis ini adalah *node* adaptif dengan fungsi *node* sebagai berikut.

$$O_{4,i} = \bar{w}_i y_i = \bar{w}_i (C_{i1} Z_{t-1} + C_{i2} Z_{t-2} + C_{i0}), \quad (12)$$

dengan

\bar{w}_i : kuat pengaktifan ternormalisasi dari lapis 3,
(C_{i1} , C_{i2} , C_{i0}) : parameter konsekuen.

Lapis ke-5 (Perhitungan Output): *Node* tunggal dari lapis ini adalah *node* tetap berlabel Σ yang menghitung keluaran keseluruhan sebagai penjumlahan semua sinyal yang datang.

$$O_{5,i} = y = \sum_i \bar{w}_i y_i = \bar{w}_1 y_1 + \bar{w}_2 y_2. \quad (13)$$

ANFIS menggunakan algoritma *hybrid* yang terdiri dari 2 langkah, yaitu langkah maju dan langkah mundur [9].

1. Pada langkah maju, parameter premis yang digunakan tetap sedangkan parameter konsekuen diidentifikasi dengan menggunakan *least square estimation* (LSE).
2. Pada langkah mundur, sinyal *error* antara *output* yang diinginkan dan *output* aktual dirambatkan mundur sedangkan parameter premis diperbarui dengan metode *error backpropagation* (EBP).

C. Pemilihan Model Terbaik

Pemilihan model terbaik berdasarkan kriteria *in sample* dan *out sample*. Kriteria *in sample* menggunakan nilai AIC dan RMSE yang dirumuskan sebagai berikut [7].

$$RMSE_{in\ sample} = \sqrt{\frac{1}{T} \sum_{i=1}^T (Z_i - \hat{Z}_i)^2}, \quad (14)$$

$$AIC = T \ln(\hat{\sigma}_a^2) + 2M$$

Sedangkan kriteria *out sample* menggunakan nilai MAPE dan RMSE yang dirumuskan sebagai berikut [4].

$$RMSE_{outsample} = \sqrt{\frac{1}{T_o} \sum_{j=1}^{T_o} (Z_j - \hat{Z}_j)^2}, \quad (15)$$

$$MAPE = \frac{\sum_{j=1}^{T_o} |(Z_j - \hat{Z}_j)/Z_j|}{T_o} \times 100\%,$$

dengan

T : banyaknya data *in sample*,
 Z_i : nilai pengamatan sebenarnya pada waktu ke- i ,
 \hat{Z}_i : nilai ramalan pada waktu ke- i ,
 $\hat{\sigma}_a^2$: nilai varians untuk data *in sample*,
 M : jumlah parameter yang ditaksir.
 T_o : banyaknya data *out sample*,
 Z_j : nilai pengamatan j step ke depan,
 \hat{Z}_j : nilai ramalan j step ke depan.

D. Pengujian Nonlinieritas

Pengujian nonlinieritas menggunakan uji Terasvirta dengan hipotesis sebagai berikut [11].

H_0 : $f(x)$ adalah fungsi linier dalam x (model linier)

H_1 : $f(x)$ adalah fungsi nonlinier dalam x (model nonlinier)

$$F = \frac{(SSR_0 - SSR)/m}{SSR/(n-r-1-m)}. \quad (16)$$

dengan

SSR_0 : jumlah kuadrat *error* dari regresi $f(x)$ dengan x dimana menghasilkan residual a_t ,

SSR : jumlah kuadrat *error* dari regresi a_t dengan x dan m ,

r : jumlah variabel prediktor awal,

m : jumlah variabel prediktor kuadratik dan kubik,

n : jumlah data.

Apabila $F > F_{(n-r-1-m)}$ maka dapat dikatakan tolak H_0 .

E. Curah Hujan

Curah hujan merupakan ketinggian air hujan yang terkumpul dalam tempat yang datar, tidak menguap, tidak meresap, dan tidak mengalir. Curah hujan satu milimeter artinya adalah dalam luasan satu meter persegi tempat yang datar, tertampung air setinggi satu milimeter atau tertampung air sebanyak satu liter. Intensitas hujan merupakan banyaknya curah hujan per-satuan jangka waktu tertentu. Jadi, apabila intensitas hujan di-katakan besar, itu tandanya hujan lebat dan dapat menimbulkan banjir. Berdasarkan intensitas curah hujan dibedakan menjadi 3 yaitu hujan sedang berada diantara 20 dan 50 mm perhari, hujan lebat berada diantara 50 dan 100 mm perhari, dan hujan sangat lebat berada diatas 100 mm perhari [12].

III. METODOLOGI PENELITIAN

A. Sumber Data

Data yang digunakan merupakan data sekunder yang diperoleh dari BMKG Jakarta Pusat dimana penelitian ini menggunakan data curah hujan harian yang berasal dari Stasiun Ahmad Yani Kota Semarang pada periode 1 Januari 2013 hingga 31 Januari 2016. Data dibagi menjadi dua bagian yaitu data *in sample* pada periode 1 Januari 2013 hingga 30 September 2015 dan data *out sample* pada periode 1 Oktober 2015 hingga 31 Januari 2016.

B. Langkah Analisis Data

Langkah analisis yang dilakukan dalam penelitian ini adalah sebagai berikut.

1. Melakukan deskripsi data curah hujan harian Kota Semarang.
2. Membagi data menjadi dua bagian yaitu data *in sample* dan data *out sample*.
3. Melakukan pembentukan model peramalan dengan metode ARIMA (data *in sample*) dengan tahapan sebagai berikut.
 - a. Melakukan plot data untuk mendeteksi stasioneritas data.
 - b. Apabila data belum stasioner dalam varians perlu melakukan transformasi dan apabila belum stasioner dalam *mean* maka perlu melakukan *differencing*.
 - c. Melihat pola ACF dan PACF untuk identifikasi bentuk model ARIMA (p,d,q) sementara.
 - d. Melakukan estimasi dan pengujian signifikansi parameter.

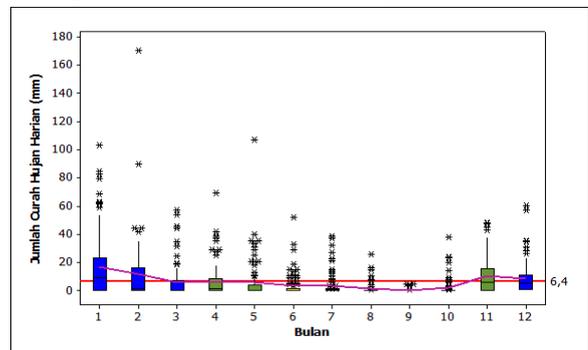
- e. Melakukan pengujian normalitas dan *white noise residual*. Apabila tidak terpenuhi maka tetap dapat dilanjutkan. Namun, semua parameter tetap harus signifikan terlebih dahulu.
 - f. Menerapkan model untuk data *in sample* dan *out sample*. Lalu dilanjutkan dengan memilih model ARIMA terbaik dengan menggunakan kriteria *in sample* dan kriteria *out sample*.
4. Mendapatkan hasil peramalan curah hujan dari model ARIMA yang telah dipilih.
 5. Melakukan peramalan untuk data curah hujan di Stasiun Ahmad Yani menggunakan metode ANFIS dengan tahapan sebagai berikut.
 - a. Membangun model peramalan pada data curah hujan (sudah dihasilkan ditahap ke-2 yaitu model ARIMA). Menentukan variabel *input* didasarkan dengan dua cara yaitu variabel yang signifikan dan variabel *Z* yang lag-lagnya signifikan di plot PACF.
 - b. Menentukan banyaknya himpunan fuzzy.
 - c. Menentukan fungsi keanggotaan yang berfungsi untuk memetakan *input* ke himpunan fuzzy.
 - d. Menentukan banyaknya aturan dan parameter yang terbentuk.
 - e. Menjalankan fungsi dari tiap lapis ANFIS (prosedur langkah maju dan langkah mundur)
 - 1) Menentukan derajat keanggotaan fuzzy. Memeta-kan *input* ke dalam himpunan fuzzy sesuai dengan fungsi keanggotaan (*membership function*).
 - 2) Menentukan aturan keanggotaan.
 - 3) Menentukan derajat pengaktifan ternormalisasi.
 - 4) Melakukan proses defuzzifikasi.
 - 5) Menghitung semua *output*.
 - 6) Melakukan prosedur langkah mundur tahapan ANFIS, yang diawali dengan menghitung *error* dari *output* langkah e(5) terhadap *output* aktual tahapan ANFIS.
 - 7) Menghitung *error* derajat pengaktifan ternormalisasi dari langkah e(3).
 - 8) Menghitung *error* derajat pengaktifan aturan fuzzy dari langkah e(2).
 - 9) Menghitung *error* derajat keanggotaan fuzzy dari langkah e(1).
 - f. Menentukan banyaknya *epoch*/iterasi untuk mendapatkan parameter-parameter ANFIS yang dapat meminimumkan *error* dengan cara mencari titik *epoch* yang konvergen.
 - g. Mendapatkan hasil ramalan ANFIS dan menghitung nilai RMSE dan MAPE hasil ramalan.
 - h. Menerapkan proses ANFIS dengan membuat kombinasi jenis *input*, jumlah dan jenis *membership function* untuk mendapatkan RMSE *outsample* yang minimum.
 6. Melakukan peramalan curah hujan dengan metode ANFIS yang terpilih.
 7. Melakukan penghitungan ukuran tingkat ketepatan prediksi MAPE dan RMSE dari hasil ramalan model ARIMA dan ANFIS di stasiun pengamatan Ahmad Yani.

8. Membandingkan ukuran akurasi ramalan model ARIMA dan ANFIS untuk data *out sample* dengan melihat nilai MAPE dan RMSE terkecil.
9. Melakukan peramalan curah hujan untuk periode ke depan dengan pemodelan yang terbaik.

IV. HASIL DAN PEMBAHASAN

A. Deskripsi Curah Hujan di Kota Semarang

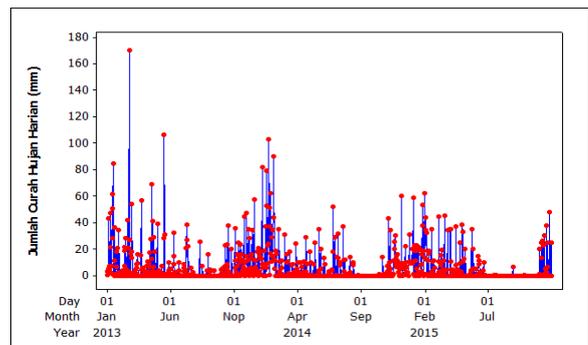
Kondisi curah hujan tiap bulannya serta pola musim yang terjadi di stasiun Ahmad Yani dari musim hujan ke musim kemarau atau sebaliknya secara lebih jelas dapat dilihat dari visualisasi *boxplot* berikut.



Gambar 2. Box Plot Data Curah Hujan Harian Ahmad Yani

Rata-rata akumulasi curah hujan di Ahmad Yani secara keseluruhan sebesar 6,413 mm. Berdasarkan Gambar 2 menunjukkan kecenderungan dari bulan Desember hingga Maret memiliki nilai rata-rata curah hujan yang lebih tinggi dibandingkan rata-rata keseluruhan sedangkan pada bulan April, Mei, dan November memiliki kecenderungan nilai rata-rata curah hujan berada di sekitaran nilai rata-rata keseluruhan dan pada bulan Juni hingga Oktober memiliki nilai rata-rata curah hujan yang cukup jauh dibanding bulan-bulan lainnya dan lebih kecil daripada rata-rata curah hujan keseluruhan. Oleh karena itu, dapat dikatakan pada bulan Desember hingga Maret menunjukkan adanya curah hujan yang tinggi atau musim hujan sedangkan bulan Juni hingga Oktober menunjukkan adanya curah hujan yang cenderung sedikit atau musim kemarau sedangkan sisanya yaitu bulan April, Mei, dan November mengindikasikan terjadinya musim pancaroba yaitu pergantian dari musim hujan ke musim kemarau atau sebaliknya.

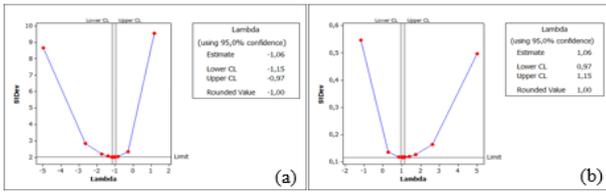
B. ARIMA



Gambar 3. Time Series Plot Data Curah Hujan Harian Ahmad

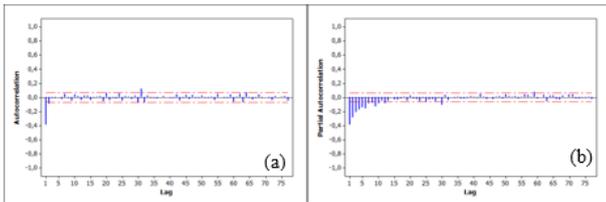
Dalam pemodelan ARIMA perlu melihat *time series plot* dari data untuk mengetahui pola dan karakteristik dari data tersebut. Hasil *time series plot* ditunjukkan di Gambar 3. Kondisi curah hujan harian di Stasiun Ahmad Yani secara visual menunjukkan pola non musiman (*non seasonal*) serta fluktuasi data yang belum stasioner baik

dalam varians maupun dalam *mean* karena antara titik satu dengan titik lainnya sangat bervariasi dan tidak berada di sekitar nilai *mean*. Namun, untuk lebih jelasnya dapat melakukan analisis dengan melihat plot Box-Cox dan plot ACF.



Gambar 4. Box-Cox Awal (a) dan Setelah Transformasi (b) Data Curah Hujan di Ahmad Yani

Dengan menggunakan transformasi $1/z_t$ terlihat di Gambar 4(b) telah stasioner dalam varians karena memiliki *rounded value* sebesar 1 sehingga dapat dilanjutkan dengan pengecekan apakah telah stasioner dalam *mean* atau belum. Terlihat pada Gambar 5(a), setelah dilakukan *differencing* pada orde ke-1, didapatkan hasil data curah hujan harian di stasiun Ahmad Yani telah stasioner sehingga dapat ditentukan model ARIMA sementara dengan melihat plot ACF dan PACF pada Gambar 5.



Gambar 5. Plot ACF (a) dan PACF (b) Setelah Transformasi dan Differencing 1 Data Curah Hujan Harian di Ahmad Yani

Selanjutnya, diperoleh 7 model ARIMA sementara yang semua parameternya signifikan dan ditunjukkan dalam Tabel 1 berikut.

TABEL 1. MODEL ARIMA SEMENTARA

Model ARIMA	Signifikansi Parameter	Asumsi White Noise	Normalitas
(0,1,2)	Signifikan	Terpenuhi	Tidak Normal
(1,1,1)	Signifikan	Terpenuhi	Tidak Normal
(1,1,[2])	Signifikan	Tidak	Tidak Normal
([1,31],1,1)	Signifikan	Terpenuhi	Tidak Normal
([1,60],1,[1,31])	Signifikan	Terpenuhi	Tidak Normal
([1,31,60],1,1)	Signifikan	Terpenuhi	Tidak Normal
(3,1,0)	Signifikan	Tidak	Tidak Normal

Semua model ARIMA sementara yang didapatkan pada Tabel 1 menghasilkan asumsi residual yang tidak berdistribusi normal. Hal ini tidak dapat dilakukan deteksi *outlier* dikarenakan histogram dari residual cenderung leptokurtik (membentuk kurva yang runcing) dan nilai dari kurtosisnya (derajat keruncingan suatu distribusi) cukup besar sehingga kemungkinan model data curah hujan di Stasiun Ahmad Yani adalah nonlinier. Namun, dalam penelitian ini tetap dilanjutkan karena parameternya telah signifikan. Untuk mengetahui model ARIMA yang terbaik, maka menggunakan kriteria *in sample* dan *out sample* yang dihasilkan dalam Tabel 2.

TABEL 2. KRITERIA IN DAN OUT SAMPLE CURAH HUJAN HARIAN STASIUN AHMAD YANI

Model ARIMA	In sample		Out sample	
	RMSE	AIC	RMSE	MAPE
(0,1,2)	14,588	-761,236	14,338	44,67%
(1,1,1)	14,598	-761,886	14,347	44,60%
(1,1,[2])	14,583	-735,654	14,336	44,64%

([1,31],1,1)	14,553	-769,288	14,218	47,64%
([1,31],1,[2])	14,550	-739,961	14,298	46,00%
([1,60],1,1)	14,600	-764,822	14,401	45,67%
([1,60],1,[1,31])	14,555	-768,828	13,999	49,18%
([1,31,60],1,1)	14,554	-772,630	14,280	46,14%
(3,1,0)	14,107	-794,873	14,143	47,98%

Tabel 2 menginformasikan bahwa model ARIMA terbaik berdasarkan kriteria *in sample* adalah ARIMA(3,1,0) sedangkan berdasarkan kriteria *out sample*, ARIMA yang terbaik adalah ARIMA ([1,60],1,[1,31]) dan ARIMA(1,1,1) karena memiliki nilai RMSE, AIC, dan MAPE terkecil. Oleh karena itu, model matematis dari ARIMA(3,1,0) adalah sebagai berikut dimana data yang digunakan merupakan data yang telah ditransformasi.

$$Z_t^* = 0,437Z_{t-1}^* + 0,9571Z_{t-2}^* + 0,6102Z_{t-3}^* + 0,2156Z_{t-4}^* + a_t.$$

Berikut model matematis dari ARIMA([1,60],1,[1,31]).

$$Z_t^* = 1,2066Z_{t-1}^* - 0,2066Z_{t-2}^* - 0,0676Z_{t-60}^* + 0,0676Z_{t-61}^* + a_t - 0,9309a_{t-1} + 0,0305a_{t-31}.$$

Berikut model matematis dari ARIMA(1,1,1).

$$Z_t^* = 1,206Z_{t-1}^* - 0,206Z_{t-2}^* + a_t - 0,9269a_{t-1}.$$

C. Pengujian Nonlinieritas

Berdasarkan perhitungan yang dilakukan, data curah hujan harian di Stasiun Ahmad Yani menghasilkan *p-value* sebesar 0,0005997. Dengan $\alpha=0,05$, didapatkan keputusan tolak H_0 karena *p-value* < α atau dapat dikatakan bahwa model jumlah curah hujan harian di Stasiun Ahmad Yani adalah nonlinier. Oleh karena itu, dapat dilanjutkan dengan menggunakan pendekatan ANFIS.

D. ANFIS

Input dalam tahap ANFIS didapatkan dari model ARIMA yang terbaik dengan data yang telah dilakukan transformasi. Pada pendekatan ini, *input* yang dipilih berdasarkan *lag-lag* yang signifikan pada plot PACF tanpa mengikutsertakan *lag-lag* pada plot ACF atau dalam model MA (*Moving Average*) sehingga *input* yang digunakan adalah sebagai berikut.

TABEL 3. INPUT ANFIS DI STASIUN AHMAD YANI

Model ARIMA	Input dalam ANFIS
(1,1,1)	Z_{t-1}^*, Z_{t-2}^*
([1,60],1,[1,31])	$Z_{t-1}^*, Z_{t-2}^*, Z_{t-60}^*, Z_{t-61}^*$
(3,0,0)	$Z_{t-1}^*, Z_{t-2}^*, Z_{t-3}^*, Z_{t-4}^*$

Setelah mendapatkan *input* yang akan digunakan, tahapan selanjutnya adalah menentukan banyaknya fungsi keanggotaan (*membership function*) dan jenis fungsi keanggotaan (*membership function*). Adapun jenis *membership function* yang digunakan adalah *Gauss*, *Trapezoidal*, dan *Generalized Bell* sedangkan banyaknya *membership function* yang digunakan sebesar 2 dan 3. Untuk meramalkan curah hujan periode selanjutnya, dipilih kombinasi antara jenis *input*, jumlah *membership function* (MF), dan jenis *membership function* yang meminimumkan RMSE dan MAPE. Berikut ini merupakan nilai RMSE dan MAPE dari masing-masing *input*, jumlah *membership function*, dan jenis *membership function* dengan pendekatan ANFIS.

TABEL 4. HASIL RMSE OUT SAMPLE MODEL ANFIS DI STASIUN AHMAD YANI

Banyak MF	Input	Jenis MF		
		G.Bell	Gauss	Trap
2	Z_{t-1}^*, Z_{t-2}^*	12,97	12,97	12,96
	$Z_{t-1}^*, Z_{t-2}^*, Z_{t-60}^*, Z_{t-61}^*$	13,75	13,74	13,75

	$Z_{t-1}^*, Z_{t-2}^*, Z_{t-3}^*, Z_{t-4}^*$	12,99	13,00	13,02
	Z_{t-1}^*, Z_{t-2}^*	12,97	12,99	12,99
3	$Z_{t-1}^* Z_{t-2}^* Z_{t-60}^*, Z_{t-61}^*$	14,09	14,07	15,06
	$Z_{t-1}^*, Z_{t-2}^*, Z_{t-3}^*, Z_{t-4}^*$	13,01	13,02	13,04

Tabel 4 dan 5 menunjukkan nilai RMSE dan MAPE terkecil berdasarkan jenis *input* dan jenis keanggotaan terdapat pada Z_{t-1}^* dan Z_{t-2}^* atau dapat dikatakan *input* yang terpilih adalah jumlah curah hujan harian satu hari dan dua hari sebelumnya sehingga dapat dilanjutkan dengan meramalkan curah hujan periode selanjutnya.

TABEL 5. HASIL MAPE *OUT SAMPLE* MODEL ANFIS DI STASIUN AHMAD YANI

Banyak MF	Input	Jenis MF		
		G.Bell	Gauss	Trap
2	Z_{t-1}^*, Z_{t-2}^*	31,76	31,43	31,12
	$Z_{t-1}^* Z_{t-2}^* Z_{t-60}^*, Z_{t-61}^*$	34,94	34,99	40,79
	$Z_{t-1}^*, Z_{t-2}^*, Z_{t-3}^*, Z_{t-4}^*$	41,74	36,60	41,59
3	Z_{t-1}^*, Z_{t-2}^*	34,15	30,29	34,44
	$Z_{t-1}^* Z_{t-2}^* Z_{t-60}^*, Z_{t-61}^*$	45,97	44,60	57,66
	$Z_{t-1}^*, Z_{t-2}^*, Z_{t-3}^*, Z_{t-4}^*$	41,70	40,24	41,88

E. Peramalan Curah Hujan Harian Periode Kedepan

Setelah melalui beberapa tahapan, Tabel 2, Tabel 4, dan Tabel 5 menunjukkan metode ANFIS memiliki nilai RMSE dan MAPE lebih kecil dibanding ARIMA sehingga metode ANFIS merupakan metode yang lebih baik dalam meramalkan curah hujan harian di stasiun pengamatan Ahmad Yani Kota Semarang. Berikut ini adalah hasil ramalan 7 periode kedepan dengan pendekatan ANFIS.

TABEL 6. HASIL RAMALAN DENGAN MODEL ANFIS DI STASIUN AHMAD YANI

Periode	Ahmad Yani
1 Februari 2016	2,09 mm
2 Februari 2016	1,75 mm
3 Februari 2016	2,00 mm
4 Februari 2016	2,15 mm
5 Februari 2016	2,07 mm
6 Februari 2016	2,01 mm
7 Februari 2016	2,04 mm

V. KESIMPULAN DAN SARAN

Kesimpulan yang didapatkan adalah karakteristik dari curah hujan di stasiun Ahmad Yani Semarang menunjukkan pada bulan Desember hingga Maret mengindikasikan adanya curah hujan yang tinggi sedangkan bulan Juni hingga Oktober menunjukkan adanya curah hujan yang cenderung sedikit dan sisanya menunjukkan musim pancaroba. Dalam metode ARIMA, dilakukan transformasi dan *differencing* orde ke-1 sehingga data curah hujan harian di Stasiun Ahmad Yani

telah stasioner. Model ARIMA yang terpilih adalah ARIMA(1,1,1) dan ARIMA([1,60],1,[1,31]) serta *input* yang terpilih adalah jumlah curah hujan harian satu hari dan dua hari sebelumnya dengan fungsi keanggotaan *Trapezoidal* dan banyaknya MF sebesar 2. Berdasarkan RMSE dan MAPE, peramalan dengan metode ANFIS menunjukkan kinerja yang lebih baik dibandingkan ARIMA karena memiliki nilai RMSE dan MAPE yang lebih kecil.

Saran dalam penelitian ini adalah untuk penelitian selanjutnya diharapkan kedepannya dapat dicoba tahapan ANFIS dengan menggunakan *input* data asli dan tanpa memenuhi asumsi stationer, serta dapat dicoba cara penentuan *input* yang lain.

DAFTAR PUSTAKA

- Admin. (2016, February 01). *Hujan Semakin Deras, Banjir Semarang kembali Mengintai*. Retrieved from ACT: <http://blog.act.id/hujan-semakin-deras-banjir-semarang-kembali-mengintai/>
- Bakti, L. M. (2010). *Kajian Sebaran Potensi ROB di Kota Semarang dan Usulan Penangannya*. Semarang: Tesis Jurusan Teknik Sipil Universitas Diponegoro.
- Nurhayati, E. P. (2012). Dampak Rob Terhadap Aktivitas Pendidikan dan Mata Pencarian di Kelurahan Bandarharjo Kecamatan Semarang Utara. *Journal of Educational Social Studies*, 66-71.
- Dewi, A. N. (2011). *Adaptive Neuro-Fuzzy Inference System untuk Peramalan Kecepatan Angin Di Bandara Juanda Surabaya*. Surabaya: Tugas Akhir Jurusan Statistika Institut Teknologi Sepuluh Nopember.
- Faulina, R. (2011). *Adaptive Neuro-Fuzzy Inference System untuk Peramalan Kecepatan Angin Rata-Rata Harian di Sumenep*. Surabaya: Tugas Akhir Jurusan Statistika Institut Teknologi Sepuluh Nopember.
- Hanke, J. E., & Winchern, D. W. (2005). *Business Forecasting 8th ed*. New Jersey: Prentice Hall.
- Wei, W. W. (2006). *Time Series Analysis Univariate and Multivariate Methods, 2nd ed*. New York: Pearson.
- Allen, A. O. (1978). *Probability, Statistics, and Queueing Theory With Computer Science Applications*. New York: Academic Press.
- Kusumadewi, S., & Hartati, S. (2006). *Neuro Fuzzy Integrasi Sistem Fuzzy dan Jaringan Syaraf*. Yogyakarta: Graha Ilmu.
- Widodo, T. S. (2005). *Sistem Neuro Fuzzy untuk Pengolahan Informasi, Pemodelan, dan Kendali*. Yogyakarta: Graha Ilmu.
- Terasvirta, T., Lin, C.-F., & Granger, C. W. (1993). Power of The Neural Network Linearity Test. *Journal of Time Series Analysis*, vol. 14, 209-220.
- Suriadikusumah, A. (2007). *Analisis Curah Hujan Perhitungan & Penggunaannya*. Bandung: Rakayasa Sains.