

METODE AUTOREGRESSIVE INTEGRATED MOVING AVERAGE (ARIMA) DAN METODE ADAPTIVE NEURO FUZZY INFERENCE SYSTEM (ANFIS) DALAM ANALISIS CURAH HUJAN

Rosita Ayu Wulandari dan Rahmat Gernowo

Departemen Fisika, Fakultas Sains dan Matematika, Universitas Diponegoro, Semarang

E-mail: rositaayu@st.fisika.undip.ac.id

ABSTRACT

Information of rainfall prediction is important for Indonesian peoples. Many statistical methods can be used in rainfall prediction, they are ARIMA (Autoregressive Integrated Moving Average) and ANFIS (Adaptive Neuro Fuzzy Inference System) methods. The purpose of this study was to compare between ANFIS method and ARIMA method to get rainfall prediction in some periods. The ARIMA method was time series data analysis often used in forecasting. While the ANFIS method was forecasting method based on rarely found time series events that are pure linear or non-linear. Based on this study, the ANFIS method has a good accuracy for time series data analysis compared with the ARIMA method. The ANFIS method has 6.9811 for the result of correlation and 87.29% for the RMSE, while result of correlation for the ARIMA method is 14.037 with 24.92% for RMSE. The ARIMA method is not good for prediction of daily data cases and non-linear data, so that the result is not actual which has a constant and flat for data prediction.

Keywords: ARIMA, ANFIS, time series data, linear, non-linear

ABSTRAK

Prediksi curah hujan merupakan informasi yang sangat penting bagi masyarakat Indonesia. Terdapat beberapa metode statistik yang dapat diaplikasikan dalam memprediksi curah hujan, yaitu metode ARIMA (*Autoregressive Integrated Moving Average*) dan metode ANFIS (*Adaptive Neuro Fuzzy Inference System*). Tujuan dari penelitian ini adalah menganalisis perbandingan metode ANFIS dengan metode ARIMA untuk memprediksi curah hujan dalam beberapa periode kedepan. Metode ARIMA merupakan metode analisis data runtun waktu yang sering digunakan untuk peramalan. Sedangkan metode ANFIS adalah metode peramalan yang didasari atas jarang ditemukan kejadian *time series* yang murni *linear* ataupun *non-linear*. Hasil penelitian ini menunjukkan bahwa metode ANFIS lebih baik untuk menganalisis data runtun waktu (*time series*) *non-linear* dibandingkan dengan metode ARIMA. Metode ANFIS memiliki keakuratan yang lebih baik yaitu dengan hasil korelasi sebesar 6,9811 dan dengan nilai RMSE 87,29%, sedangkan pada metode ARIMA dihasilkan korelasi sebesar 14,037 dan dengan nilai RMSE 24,92%. Metode ARIMA untuk kasus data harian dan data *non-linear* kurang cocok, sehingga hasil prediksi tidak mengikuti pola data aktual dan cenderung memiliki hasil yang konstan atau hasil prediksi yang cenderung *flat*.

Kata kunci: ARIMA, ANFIS, data runtun waktu, linier, non-linier

PENDAHULUAN

Indonesia merupakan negara kepulauan di Asia Tenggara yang dilintasi oleh garis khatulistiwa serta berada di antara dua benua yaitu benua Asia dan Australia, serta terletak di antara dua samudra yaitu Samudra Pasifik dan Samudra Hindia. Indonesia dikenal dengan negara yang beriklim tropis. Hal ini dapat dilihat dari letak geografis Indonesia yang berada di sekitar garis khatulistiwa. Indonesia dikelilingi hamparan laut yang sangat luas, penguapan dan pengumpulan awan-awan berlangsung sangat cepat sehingga proses hujan terjadi hampir setiap hari. Hal tersebut menjadikan Indonesia merupakan negara yang memiliki curah hujan yang tinggi.

Prediksi curah hujan merupakan informasi yang sangat penting bagi masyarakat Indonesia. Akurasi prediksi curah hujan dapat digunakan untuk menunjang berbagai kegiatan masyarakat, seperti pertanian, perkebunan dan lain sebagainya. Terdapat beberapa metode statistik yang dapat diaplikasikan untuk memprakirakan curah hujan bulanan. Salah satu metode statistik yang sudah diterapkan oleh Badan Meteorologi Klimatologi dan Geofisika dalam prakiraan curah hujan adalah metode ARIMA (*Autoregressive Integrated Moving Average*), yaitu suatu metode yang didasarkan pada nilai-nilai suatu peubah yang telah terjadi pada waktu lampau, lalu digunakan untuk menentukan pola historis data, dan kemudian digunakan untuk mengekstrapolasikan pola tersebut pada masa yang akan datang [1]. Selain metode ARIMA, terdapat metode ANFIS (*Adaptive Neuro Fuzzy Inference System*). Ide menggunakan metode ANFIS sebagai metode peramalan didasari jarang ditemukan kejadian *time series* yang murni *linear* ataupun *non-linear*.

Metode ANFIS menggunakan Jaringan Syaraf Tiruan (*neural network*) untuk mengimplementasikan sistem inferensi *fuzzy*.

Penelitian ini membandingkan metode ARIMA dan metode ANFIS dalam memprediksi curah hujan untuk beberapa periode mendatang. *Input* yang digunakan berupa data curah hujan harian. Data curah hujan merupakan data runtun waktu yang menggambarkan pola musim [2].

Menurut penelitian Tyasono (2003), ANFIS memberikan hasil akurasi yang lebih baik dibandingkan dengan ARIMA dengan nilai RMSE yang lebih kecil dibanding dengan ARIMA [3]. Dengan adanya hasil perbandingan antara kedua metode tersebut, dapat menjadi masukan kepada BMKG agar beralih menggunakan metode ANFIS untuk peramalan guna mendapatkan hasil yang lebih akurat.

DASAR TEORI

Cuaca dan Iklim

Cuaca merupakan keadaan udara pada saat tertentu pada suatu wilayah yang relatif sempit. Dalam kata lain cuaca dapat didefinisikan sebagai variasi atmosfer dalam periode pendek. Cuaca dapat berubah-ubah yang disebabkan oleh tekanan udara, suhu, angin, kelembaban udara dan juga curah hujan. Berbeda dengan iklim yang merupakan cuaca rerata suatu wilayah pada jangka panjang [3].

Hujan

Hujan dapat didefinisikan sebagai endapan yang berupa tetes cair maupun padat yang jatuh ke permukaan bumi dan berasal dari atmosfer. Sedangkan curah hujan merupakan banyaknya atau intensitas tetes cair yang jatuh ke permukaan bumi. Curah hujan sendiri merupakan tingkat hujan yang turun di suatu daerah. Banyaknya curah hujan dinyatakan dalam milimeter (mm), yang berarti hujan yang jatuh pada permukaan datar seluas satu meter persegi (m^2) setinggi

satu milimeter tidak meresap, mengalir, maupun menguap dalam selang waktu tertentu [4].

Analisis Runtun Waktu (*Time Series*) dan Peramalan

Data runtun waktu (*time series*) merupakan jenis data yang dikumpulkan sesuai dengan urutan waktu dalam rentang waktu tertentu. Adapun dasar pemikiran dalam runtun waktu yaitu pengamatan sekarang (Z_t) yang dipengaruhi oleh satu atau beberapa pengamatan sebelumnya (Z_{t-k}). Tujuan dari analisis runtun waktu yaitu untuk memahami dan menjelaskan mekanisme tertentu, meramalkan suatu nilai pada masa yang akan datang, serta dapat mengoptimalkan suatu sistem kendali [2]. Peramalan merupakan suatu kegiatan mengestimasi sesuatu yang akan terjadi pada masa berikutnya dengan rentang waktu yang relatif lama. Sedangkan ramalan merupakan suatu kondisi yang akan diprediksi terjadi pada masa mendatang. Dalam memprediksikan hal tersebut memerlukan data yang akurat pada masa lalu, agar dapat digunakan untuk melihat situasi pada masa mendatang [5].

Metode ARIMA (*Autoregressive Integrated Moving Average*)

Metode ARIMA (*Autoregressive Integrated Moving Average*) merupakan metode analisis deret berkala yang dikenal sebagai *Box-Jenkins*. Metode ini berasal dari penggabungan antara model *Autoregressive* (AR) dan *Moving Average* (MA) yang dikembangkan oleh George Box dan Gwilym Jenkins. Menurut *Box-Jenkins* metode ARIMA terdiri dari empat tahap yaitu identifikasi metode *time series*, pendugaan parameter-parameter bagi metode alternatif, pengujian metode, dan prakiraan nilai *time series* [1]. Asumsi kestasioneran merupakan asumsi yang harus dipenuhi dalam memodelkan runtun waktu. Deret nonstasioner dapat ditransformasikan ke dalam deret

stasioner dengan cara perbedaan (*differencing*) [4]. Ketidak-stasioneran dalam runtun waktu dapat meliputi *mean* yang tidak konstan, *varians* yang tidak konstan ataupun keduanya (*mean* dan *varians* tidak konstan).

Model AR (*Autoregressive*) orde p menyatakan pengamatan pada waktu ke- t yang berhubungan linier dengan pengamatan waktu sebelumnya $t-1, t-2, \dots, t-p$. Bentuk persamaan dari model AR dapat dituliskan sebagai berikut:

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \dots + \phi_p Z_{t-p} + a_t \quad (1)$$

Model MA (*Moving Average*) digunakan untuk menjelaskan suatu kejadian dimana suatu pengamatan pada waktu t dinyatakan sebagai kombinasi linier dari sejumlah *residual*. Bentuk persamaan dari model MA dapat dituliskan sebagai berikut:

$$Z_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q} \quad (2)$$

Sedangkan model ARMA merupakan gabungan dari model AR dan MA yang dapat ditulis dengan notasi ARMA (p, q). Bentuk persamaan dari model ARMA pada orde p dan q dapat dituliskan sebagai berikut:

$$Z_t = \phi_1 Z_{t-1} + \dots + \phi_p Z_{t-p} + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} \quad (3)$$

Model ARIMA sebagai fungsi dari p, d, q , dimana p sebagai orde operator dari AR, d orde *differencing*, dan q sebagai orde operator dari MA. Data *time series* yang telah stasioner setelah

mengalami *differencing* sebanyak d kali yaitu dengan menghitung selisih pengamatan dengan pengamatan sebelumnya [6].

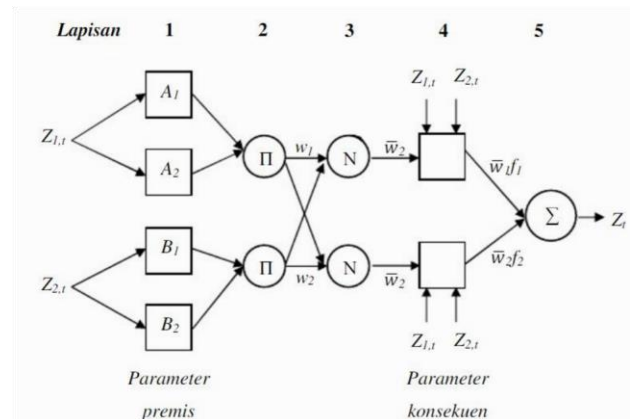
Bentuk umum persamaan model ARIMA dapat dituliskan sebagai berikut:

$$\phi_p(B)(1 - B)^d Z_t = \mu' + \theta_q(B)a_t \quad (4)$$

dengan $\phi_p(B)$ merupakan komponen AR orde p dan $\theta_q(B)$ merupakan komponen orde q .

ANFIS (Adaptive Neuro Fuzzy Inference System)

ANFIS (*Adaptive Neuro Fuzzy Inference System*) merupakan implementasi jaringan adaptif dalam *fuzzy inference system*. Sistem logika *fuzzy* bersifat adaptif, berarti bahwa sistem *fuzzy* dapat disesuaikan dengan kondisi sesuai keinginan. Arsitektur ANFIS sama dengan jaringan syaraf fungsi radial dengan sedikit batasan tertentu. Dengan kata lain, ANFIS merupakan suatu metode yang dapat melakukan penyetelan aturan yang menggunakan algoritma pembelajaran terhadap sekumpulan data. Model *fuzzy* dapat digunakan sebagai pengganti dari banyak lapisan. Dalam hal ini, sistem dibagi menjadi dua grup, Pertama adalah grup berupa jaringan syaraf dengan bobot *fuzzy* dan fungsi aktivasi *fuzzy*. Sedangkan grup yang kedua berupa jaringan syaraf dengan *input* yang di-*fuzzy*-kan pada lapisan pertama atau pada lapisan kedua. Namun, pada grup kedua, bobot-bobot pada jaringan syaraf tersebut tidak di-*fuzzy*-kan seperti pada grup yang pertama [5].



Gambar 1. Arsitektur jaringan ANFIS

Menurut Jang, dkk (1997) [7], sebagaimana pada gambar 1, jaringan ANFIS terdiri dari lapisan-lapisan sebagai berikut:

Lapisan 1: Merupakan lapisan *fuzzifikasi*, dimana pada lapisan ini setiap neuron adaptif terhadap parameter aktivasi. Output dari setiap neuron berupa derajat keanggotaan yang diberikan oleh fungsi keanggotaan input. Misal, fungsi keanggotaan *Gaussian* dapat dituliskan dalam persamaan berikut:

$$\mu(Z) = e^{-\frac{1}{2} \frac{(Z-c)^2}{\sigma^2}} \quad (5)$$

Dimana Z merupakan input, dalam hal ini $Z = \{Z_{1,t}, Z_{2,t}\}$ dan $\{\sigma, c\}$ adalah parameter-parameter. Apabila nilai parameter-parameter tersebut berubah, maka bentuk kurva akan berubah juga. Parameter ini sering disebut dengan parameter premis.

Lapisan 2: Pada lapisan ini berupa neuron tetap (Π) merupakan hasil kali dari semua masukan yang dapat dituliskan dalam persamaan berikut:

$$w_i = \mu_{A_i} \cdot \mu_{B_i} \quad (6)$$

Pada lapisan ini biasa digunakan operator *AND* dan hasil dari perhitungan ini disebut dengan *firing strength* dari sebuah

aturan. Dimana setiap neuron mempresentasikan aturan ke- i .

Lapisan 3: Pada lapisan ini berupa neuron tetap (N) dimana merupakan hasil dari perhitungan rasio dari *firing strength* ke- i (w_i) terhadap jumlah keseluruhan *firing strength* pada lapisan kedua yang dapat dituliskan dalam persamaan berikut:

$$\bar{w}_i = \frac{w_i}{\sum_{j=1}^2 w_j}, i = 1, 2 \quad (7)$$

Hasil dari perhitungan tersebut disebut dengan *normalized firing strength*.

Lapisan 4: Pada lapisan ini berupa neuron yang merupakan neuron adaptif terhadap suatu output yang dapat dituliskan pada persamaan berikut:

$$\bar{w}_i f_i = \bar{w}_i (p_i Z_{1,t} + q_i Z_{2,t} + r_i) \quad (8)$$

Dimana \bar{w}_i merupakan *normalized firing strength* pada lapisan ketiga dan; p_i , q_i , dan r_i merupakan parameter konsekuen.

Lapisan 5: Pada lapisan ini berupa neuron tunggal (Σ) dimana merupakan hasil dari penjumlahan seluruh output dari lapisan keempat yang dapat dituliskan dalam persamaan sebagai berikut:

$$\sum_i \bar{w}_i f_i = \frac{\sum_i \bar{w}_i f_i}{\sum_i \bar{w}_i f_i} \quad (9)$$

METODE PENELITIAN

Variabel yang digunakan dalam penelitian ini adalah data curah hujan harian Kota Bogor dari periode 1 September 2016 sampai 2 Desember 2017 yang terdiri dari 397 data. Data nantinya dibagi menjadi dua bagian yaitu data pelatihan (*training*) dan data pengujian (*testing*). Data pelatihan digunakan untuk pemodelan. Sedangkan data pengujian digunakan untuk pembandingan dengan hasil peramalan.

Persentase pembagian data pelatihan yaitu 70% dari jumlah data keseluruhan yaitu 278 data (1 September 2016 - 5 Agustus 2017) dan 30% untuk data pengujian sejumlah 119 data (6 Agustus 2017 - 2 Desember 2017).

Peralatan yang digunakan dalam penelitian ini yaitu berupa perangkat lunak (*software*) Matlab R2015a yang digunakan untuk membuat model prakiraan pada metode ANFIS, Minitab18 dan *Eviews* yang digunakan untuk prakiraan model dengan metode ARIMA, dan Microsoft Excel digunakan untuk menghitung korelasi dan nilai RMSE (*Root Mean Square Error*) antara hasil output dari kedua metode tersebut.

HASIL DAN PEMBAHASAN

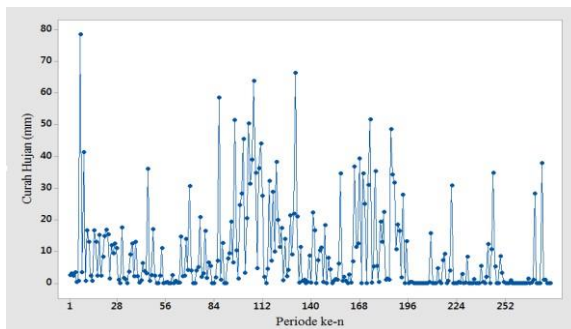
Analisis ARIMA pada Prediksi Curah Hujan

Gambar 2 menunjukkan data *real* curah hujan sebagai *input* model. Sedangkan gambar 3 merupakan uji stasioneritas dari gambar 2. Data stasioner dapat diketahui secara visual maupun formal dengan uji stasioneritas, sehingga tidak perlu dilakukan *differencing*. Uji secara visual dengan cara memperhatikan plot data runtun waktu, sedangkan untuk uji secara formal dapat dilakukan dengan uji ADF (*Augmented Dickey- Fuller*). Selanjutnya pada Gambar 4 merupakan pendugaan orde AR dan MA untuk pemodelan ARIMA dengan melihat plot ACF dan PACF.

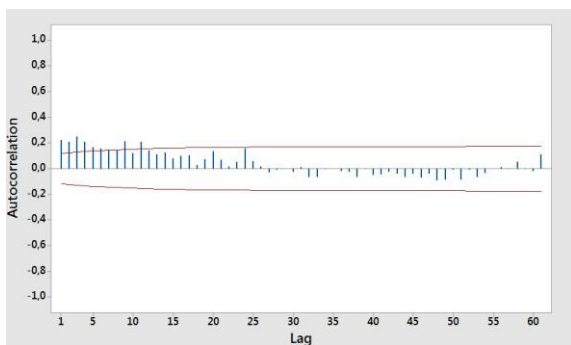
Berdasarkan plot ACF dan PACF diperoleh sepuluh model sementara dan satu model terbaik yaitu ARIMA (1,0,1) yang kemudian dilanjutkan dengan uji signifikansi, uji normalitas residual, uji independensi, uji ARCH-LM dan uji linieritas. Persamaan model terbaik ARIMA (1,0,1) dapat dituliskan sebagai berikut:

$$\dot{Z}_t = 0,471 + 0,9473 Z_{t-1} - 0,8216 a_{t-1} + a_t$$

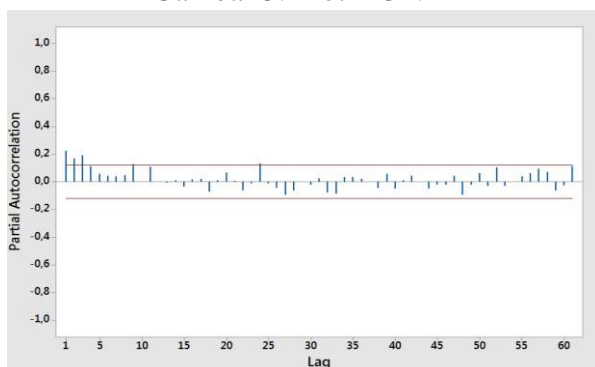
(10)



Gambar 2. Plot data curah hujan harian Kota Bogor periode 1 September 2016 - 5 Agustus 2017.

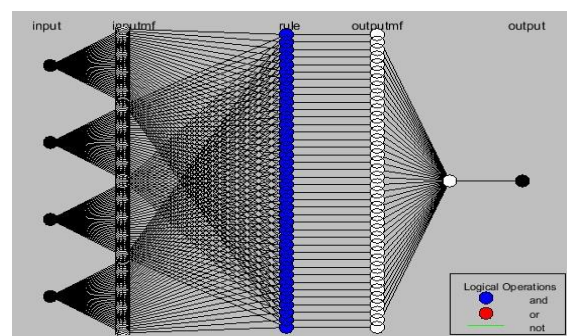


Gambar 3. Plot ACF.



Gambar 4. Plot PACF.

outputnya yaitu berupa curah hujan (mm)(Z_t). Sebuah FIS (*Fuzzy Inference System*) digunakan untuk memperoleh arsitektur ANFIS, sebelum melakukan pelatihan pada ANFIS dilakukan penentuan parameter awal terlebih dahulu yaitu menggunakan *genfis2* (*Generate Fuzzy Inference System*) dengan jenis fungsi keanggotaan (*membership function*) yang digunakan adalah *gaussmf* (*Gaussian*). Dalam *genfis2* menggunakan fungsi *subclust* untuk melakukan *clustering* data, melatih input-input data dan membangkitkan penalaran *fuzzy* dengan metode Takagi - Sugeno orde satu. Nilai radius atau sering disebut dengan *range of influence* (jari-jari) sangat berhubungan dengan penentuan input jumlah fungsi keanggotaan. Semakin kecil *range of influence* maka jumlah input pada fungsi keanggotaan akan semakin banyak yang berarti tingkat keakuratan semakin tinggi ditandai dengan *error* yang mengecil. Pada proses pelatihan menggunakan iterasi (*epoch*) sebanyak 30. Dengan menggunakan nilai *range of influence* 0,2 dihasilkan arsitektur jaringan ANFIS seperti yang terlihat pada gambar 5.



Gambar 5. Arsitektur jaringan ANFIS.

Analisis ANFIS pada Prediksi Curah Hujan

Pembentukan model ANFIS terdiri dari empat *input* dan satu *output*. *Input* dalam pembentukan model ANFIS terdiri dari suhu rata-rata ($^{\circ}\text{C}$)($Z_{1,t}$), kelembaban rata-rata (%)($Z_{2,t}$), lama penyinaran (Jam)($Z_{3,t}$) dan kecepatan angin rata-rata (*Knot*)($Z_{4,t}$). Sedangkan untuk

Terdapat rule sebanyak 40 buah dengan operator AND dan terdiri dari satu *output* dengan *output* MF yang terbentuk sebanyak 40 buah. Garis-garis yang menghubungkan antar node menunjukkan rule yang bersesuaian. Dari hasil pelatihan total

parameter yang digunakan sebanyak 520 parameter yang terdiri dari 320 parameter *non-linear* dan 200 parameter *linear*. Sehingga persamaan model terbaik dapat dituliskan sebagai berikut:

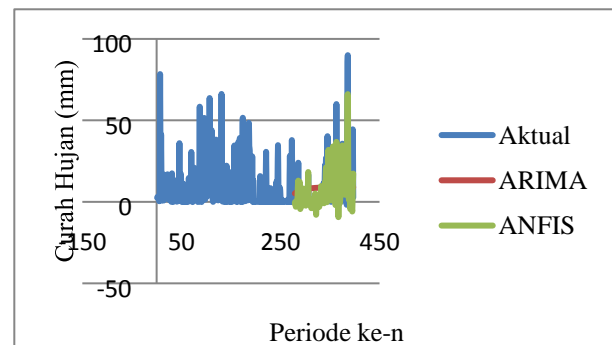
$$\begin{aligned} \dot{Z}_t = & \bar{w}_{1t}(-1,84Z_{1,t} - 0,23Z_{2,t} + 1,67Z_{3,t} \\ & + 0,57Z_{4,t} + 44,83) + \bar{w}_{2t} \\ & (69,28Z_{1,t} + 29,51Z_{2,t} + 5,81Z_{3,t} + 19,37Z_{4,t} \\ & - 4184,4) + \dots + \bar{w}_{40t} \\ & (-238,24Z_{1,t} \\ & 5,98Z_{2,t} + 26,81Z_{3,t} - 54,91Z_{4,t} + \\ & 4827,5) \end{aligned} \quad (11)$$

Analisis Hasil Metode ARIMA dan Metode ANFIS

Berdasarkan hasil yang telah diperoleh pada gambar 6, pada metode ARIMA maupun metode ANFIS dapat dilihat bahwa metode ANFIS lebih baik daripada metode ARIMA karena memiliki ketepatan yang sangat baik dalam peramalan periode pada curah hujan harian daripada metode ARIMA, dimana metode ANFIS memiliki nilai korelasi lebih besar yaitu 87,29% dengan nilai RMSE yang lebih kecil yaitu 6,8911. Sedangkan ARIMA menghasilkan korelasi sebesar 24,92% dan nilai RMSE sebesar 14,037.

Output dari metode ANFIS hampir berhimpitan dengan data aktual karena metode ANFIS mampu menganalisis data runtun waktu *non-linear* dengan mengubah parameter *nonlinear*. Dibandingkan dengan metode ARIMA yang tidak dapat menganalisis data runtun waktu *non-linear* karena kebanyakan model dari metode ARIMA adalah *linear*. Metode ARIMA untuk kasus data harian kurang cocok, sehingga hasil prediksi tidak mengikuti pola data aktual dan cenderung memiliki hasil yang konstan atau hasil prediksi yang cenderung *flat*. Dalam proses pelatihan ANFIS, data yang digunakan lebih banyak daripada data untuk pengujian karena semakin banyak data

yang digunakan untuk pelatihan maka semakin baik hasil pelatihan yang diperoleh, sehingga sistem peramalan yang diperoleh akan semakin baik pula. Sedangkan metode ARIMA yang menggunakan data terlalu banyak dapat mengakibatkan residual fungsi auto korelasinya tidak independen.



Gambar 6. Perbandingan data aktual curah hujan dengan hasil prediksi metode ARIMA dan hasil prediksi metode ANFIS

KESIMPULAN

Berdasarkan identifikasi permasalahan dan pembahasan sebelumnya, dapat disimpulkan bahwa metode ANFIS lebih baik untuk menganalisis data runtun waktu (*time series non-linear*) dibandingkan dengan metode ARIMA. Metode ANFIS memiliki keakuratan yang lebih baik untuk peramalan curah hujan yaitu dengan hasil korelasi sebesar 6,9811 dan dengan nilai RMSE 87,29%, sedangkan pada metode ARIMA dihasilkan korelasi sebesar 14,037 dan dengan nilai RMSE 24,92%. Metode ARIMA untuk kasus data harian kurang cocok, sehingga hasil prediksi tidak mengikuti pola data aktual dan cenderung memiliki hasil yang konstan atau hasil prediksi yang cenderung *flat*.

DAFTAR PUSTAKA

- [1]. Danitasari F. Perbandingan prediksi sifat hujan bulanan antara analisis komponen utama model arima dan metode probabilitas di Stasiun Meteorologi Pongtiku Tana Toraja. *Jurnal Meteorologi Klimatologi dan Geofisika*. 2015;2(2):207–215.
- [2]. Makridakis S, Wheelwright SC, McGEE VE. *Metode dan aplikasi peramalan* (diterjemahkan oleh Andriyanto US dan Basith A). Edisi kedua. Jakarta: Penerbit Erlangga; 1998.
- [3]. Tjasyono HKB, Harijono SWB. *Meteorologi indonesia volume II awan dan hujan monsun*. Jakarta: BMKG; 2012.
- [4]. Lusiani A, Habinuddin E. Pemodelan autoregressive integrated moving average (arima) curah hujan di Kota Bandung. *Jurnal Sigma-Mu*. 2011;3(2):9–25.
- [5]. Saputra AH. *Analisis data runtun waktu dengan metode adaptive neuro fuzzy inference system (ANFIS)*. Semarang: Skripsi Jurusan Statistika, Fakultas Sains dan Matematika, Universitas Diponegoro; 2012.
- [6]. Susanto Y, Ulama BSS. Pemodelan curah hujan dengan pendekatan model ARIMA, feed forward neural network dan hybrid (ARIMA-NN) di Banyuwangi. *Jurnal Sains dan Seni ITS*. 2016;5(2):145–150.
- [7]. Jang JSR, Sun CT, Mizutani E. *Neurofuzzy and soft computing: a computational approach to learning and machine intelligence*. London: PrenticeHall, Inc.; 1997.