



**TESIS – SS235401**

**IMPLEMENTASI MODEL ARIMA, ANFIS DAN  
TRANSFORMASI WAVELET UNTUK  
MERAMALKAN CURAH HUJAN DI STASIUN  
JUANDA**

**WINDI  
NRP. 6003202008**

**Dosen Pembimbing  
Dr.Dra. Kartika Fithriasari, M.Si.  
Santi Puteri Rahayu, M.Si, Ph.D.**

**PROGRAM STUDI MAGISTER  
DEPARTEMEN STATISTIKA  
FAKULTAS SAINS DAN ANALITIKA DATA  
INSTITUT TEKNOLOGI SEPULUH NOPEMBER  
SURABAYA  
2024**





**TESIS – SS235401**

**IMPLEMENTASI MODEL ARIMA, ANFIS DAN  
TRANSFORMASI WAVELET UNTUK  
MERAMALKAN CURAH HUJAN DI STASIUN  
JUANDA**

**WINDI  
NRP. 6003202008**

**Dosen Pembimbing  
Dr.Dra. Kartika Fithriasari, M.Si.  
Santi Puteri Rahayu, M.Si, Ph.D.**

**PROGRAM STUDI MAGISTER  
DEPARTEMEN STATISTIKA  
FAKULTAS SAINS DAN ANALITIKA DATA  
INSTITUT TEKNOLOGI SEPULUH NOPEMBER  
SURABAYA  
2024**





**TESIS – SS235401**

**IMPLEMENTATION OF ARIMA, ANFIS AND  
WAVELET TRANSFORM MODELS FOR RAINFALL  
PREDICTION AT JUANDA STATION**

**WINDI**

**NRP. 6003202008**

**SUPERVISORS:**

**Dr.Dra. Kartika Fithriasari, M.Si.**

**Santi Puteri Rahayu, M.Si, Ph.D.**

**MASTER PROGRAM  
DEPARTMENT OF STATISTICS  
FACULTY OF SCIENCES AND DATA ANALYTICS  
INSTITUT TEKNOLOGI SEPULUH NOPEMBER  
SURABAYA  
2024**



# LEMBAR PENGESAHAN TESIS

Tesis disusun untuk memenuhi salah satu syarat memperoleh gelar  
**Magister Statistika (M.Stat)**

di  
**Institut Teknologi Sepuluh Nopember**

Oleh:

**WINDI**


**NRP: 6003202008**

Tanggal Ujian: 02 Juli 2024  
Periode Wisuda: September 2024

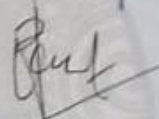
Disetujui oleh:

**Pembimbing:**

1. Dr. Kartika Fithriasari, M.Si  
NIP: 19691212 199303 2 002

  
.....

2. Santi Puteri Rahayu, M.Si., Ph.D  
NIP: 19750115 199903 2 003

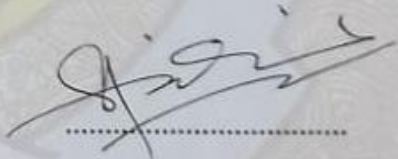
  
.....

**Penguji:**

1. Prof. Drs. Nur Iriawan, M.lkom.,Ph.D  
NIP: 19621015 198803 1 002


  
.....

2. Dr. Muhammad Sjahid Akbar, Msi  
NIP: 19720705 199802 1 001

  
.....



Kepala Departemen Statistika  
Fakultas Sains dan Analitika Data

  
Dr. Kartika Fithriasari, M.Si  
NIP: 19691212 199303 2 002

*(Halaman ini sengaja dikosongkan)*



# IMPLEMENTASI MODEL ARIMA, ANFIS DAN TRANSFORMASI WAVELET UNTUK MERAMALKAN CURAH HUJAN DI STASIUN JUANDA

Mahasiswa Nama : Windi  
NRP : 6003202008  
Pembimbing : Dr. Dra. Kartika Fithriasari, M.Si  
Co – Pembimbing : Santi Puteri Rahayu, M.Si, Ph.D

## ABSTRAK

Analisis deret waktu (*time series*) adalah salah satu bahasan dalam ilmu statistika. Model analisis deret waktu yang sering digunakan untuk melakukan peramalan adalah ARIMA. ARIMA adalah salah satu metode peramalan yang biasanya baik digunakan untuk melakukan peramalan jangka pendek. Dengan seiringnya waktu metode tersebut memiliki keterbatasan dalam memodelkan data deret waktu, terutama pada data deret waktu nonlinear dan non stasioner. Salah satu model yang digunakan untuk peramalan data deret waktu nonlinear dan data deret waktu nonstasioner adalah *Adaptive Neuro Fuzzy System* (ANFIS) adalah penggabungan mekanisme sistem inferensi *fuzzy* yang digambarkan dalam arsitektur jaringan syaraf tiruan. ANFIS adalah suatu metode yang biasa digunakan untuk meramalkan dengan tingkat akurasi yang cukup baik dan analisis wavelet telah banyak diaplikasikan untuk permasalahan statistika secara umum. permasalahan khusus dari analisis time series yaitu estimasi densitas spektral untuk proses stasioner dan tidak stasioner. Curah hujan adalah salah satu data yang sering dianalisis dalam bentuk time series. Data time series adalah data yang dikumpulkan atau direkam pada interval waktu yang teratur. Dalam konteks curah hujan, data ini biasanya diukur setiap hari, bulan, atau tahun. Analisis data time series curah hujan dapat membantu dalam memahami pola curah hujan, tren jangka panjang, variabilitas musim, dan perubahan iklim. penelitian ini dilakukan untuk memprediksi curah hujan dengan membandingkan beberapa model untuk meramalkan data time series yaitu ARIMA, ANFIS dan Transformasi Wavelet. Berdasarkan hasil penelitian model ANFIS memiliki performa yang baik dalam melakukan peramalan data curah hujan di Stasiun Juanda dengan melihat nilai RMSE lebih rendah.

**Kata kunci:** ARIMA, ANFIS, Transformasi Wavelet, *Membership Function*.

*(Halaman ini sengaja dikosongkan)*

# **IMPLEMENTATION OF ARIMA, ANFIS AND WAVELET TRANSFORM MODELS FOR RAINFALL PREDICTION AT JUANDA STATION**

Name : Windi  
NRP : 6003202008  
Supervisor : Dr. Dra. Kartika Fithriasari, M.Si  
Co – supervisor : Santi Puteri Rahayu, M.Si, Ph.D

## **ABSTRACT**

*Time series analysis is one of the topics in statistics. The time series analysis model that is often used for forecasting is ARIMA. ARIMA is a forecasting method that is usually good for short-term forecasting. Over time, this method has limitations in modeling time series data, especially in nonlinear and non-stationary time series data. One of the models used for forecasting nonlinear time series data and nonstationary time series data is the Adaptive Neuro Fuzzy System (ANFIS) which is a combination of fuzzy inference system mechanisms described in artificial neural network architecture. ANFIS is a method commonly used to predict with a fairly good level of accuracy and wavelet analysis has been widely applied to statistical problems in general. A special problem in time series analysis is the estimation of spectral density for stationary and non-stationary processes. Rainfall is one of the data that is often analyzed in time series form. Time series data is data that is collected or recorded at regular time intervals. In the context of rainfall, this data is usually measured every day, month or year. Analysis of rainfall time series data can help in understanding rainfall patterns, long-term trends, seasonal variability, and climate change. This research was conducted to predict rainfall by comparing several models for predicting time series data, namely ARIMA, ANFIS and Wavelet Transformation. Based on the research results, the ANFIS model is better than the ARIMA model and Wavelet Transformation for predicting daily rainfall at Juanda Station, seen from the lower RMSE value.*

**Keywords:** ARIMA, ANFIS, Wavelet Transform, Membership Function

*(Halaman ini sengaja dikosongkan)*

## KATA PENGANTAR

Bismillahirrahmanirrahim. Puji dan syukur penulis ucapkan kepada Allah SWT, karena atas berkat, rahmat, dan hidayah – Nya, penulis dapat menyelesaikan penelitian Tesis yang berjudul **“Implementasi Model ARIMA, ANFIS dan Transformasi Wavelet untuk Meramalkan Curah Hujan di Stasiun Juanda”** yang merupakan salah satu syarat kelulusan untuk menyelesaikan Program Magister Departemen Statistika, Fakultas Sains dan Analitika Data, Institut Teknologi Sepuluh Nopember Surabaya. Tesis ini dapat diselesaikan dengan baik berkat doa, kerjasama, bantuan, serta dukungan dari berbagai pihak. Suatu kebahagiaan bagi penulis untuk dapat menyampaikan rasa terima kasih kepada seluruh pihak yang membantu serta mendukung selesainya tesis ini. Oleh karena itu, penulis ingin menyampaikan ucapan terima kasih kepada:

1. Allah SWT yang telah memberikan saya kesempatan dan kemampuan untuk melanjutkan studi di jenjang Magister.
2. Kedua orang tua yaitu Bapak Adham dan Ibu Armiatin serta keluarga penulis yang tidak hentinya mendoakan dan memberikan dukungan kepada penulis dalam meraih cita – cita yang diinginkan.
3. Ibu Dr. Dra. Kartika Fithriasari, M.Si dan Ibu Santi Puteri Rahayu, M.Si., Ph.D selaku dosen pembimbing yang telah membantu memberikan bimbingan, saran, dan motivasi selama menyusun tesis.
4. Bapak Prof Nur Iriawan, M.Iikom dan Bapak Dr. Muhammad Sjahid Akbar, selaku dosen penguji yang telah memberikan masukan dan arahan kepada penulis, sehingga tesis ini dapat diselesaikan dengan baik.
5. Ibu Dr. Dra. Kartika Fithriasari, M.Si selaku Kepala Departemen Statistika Fakultas Sains dan Analitika Data Institut Teknologi Sepuluh Nopember Surabaya.
6. Ibu Santi Wulan Purnami, M.Si., Ph.D selaku Sekretaris Departemen I (Bidang Akademik, Kemahasiswaan, Penelitian, dan Pengabdian Kepada Masyarakat).
7. Ibu Prof. Dr. Vita Ratnasari, M.Si selaku Dosen Wali dan Sekretaris Departemen II (Bidang Sumber Daya Keuangan, Sumber Daya Manusia, dan Sarana Prasarana).

8. Bapak Dr.rer.pol. Dedy Dwi Prastyo, M.Si selaku Kepala Program Studi Magister Statistika Fakultas Sains dan Analitika Data Institut Teknologi Sepuluh Nopember Surabaya.
9. Bapak/Ibu dosen pengajar di Departemen Statistika ITS, terimakasih atas ilmu yang telah diberikan selama penulis menempuh pendidikan,
10. Bapak/Ibu staf dan karyawan di Departemen Statistika ITS, terimakasih atas dukungan dan bantuan selama proses penelitian ini.
11. Teman – teman S2 Statistika ITS yang telah memberikan pengalaman dan kebersamaan yang sangat berharga.
12. Terakhir, untuk diri saya sendiri, terima kasih Windi yang telah bekerja keras dan berproses hingga saat ini.

Semoga tesis ini dapat memberikan sumbangan wawasan ilmiah dan memberikan manfaat bagi Masyarakat. Penulis menyadari bahwa masih banyak kesalahan dalam penulisan tesis ini. Oleh karena itu, kritik dan saran dari berbagai pihak sangat penulis harapkan agar nantinya menjadi koreksi untuk penelitian yang akan datang.

Surabaya, Juli 2024

Windi

## DAFTAR ISI

HALAMAN JUDUL.....	i
LEMBAR PENGESAHAN .....	vii
ABSTRAK .....	ix
<i>ABSTRACT</i> .....	xi
KATA PENGANTAR .....	xiii
DAFTAR ISI.....	xv
DAFTAR TABEL.....	xvii
DAFTAR GAMBAR .....	xviii
DAFTAR NOTASI.....	xxi
BAB 1 PENDAHULUAN .....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	4
1.3 Tujuan Penelitian .....	4
1.4 Manfaat Penelitian .....	5
1.5 Batasan Masalah .....	5
BAB 2 TINJAUAN PUSTAKA .....	7
2.1 Peneliti Terdahulu.....	8
2.2 ARIMA .....	8
2.3 <i>Adaptive Neuro Fuzzy System (ANFIS)</i> .....	16
2.4 Transformasi Wavelet.....	21
2.5 Analisis Deret Waktu.....	28
2.6 <i>Root Mean Square Error</i> .....	28
2.7 Curah Hujan.....	29
BAB 3 METODE PENELITIAN.....	31
3.1 Sumber Data dan Variabel Penelitian.....	31
3.2 Langkah Analisis .....	32
BAB 4 HASIL DAN PEMBAHASAN.....	35
4.1 Deskripsi Data Curah Hujan.....	39
4.2 Model ARIMA .....	40
4.3 Model ANFIS .....	45
4.4 Model Transformasi Wavelet .....	50

4.5 Perbandingan Performa Model Terbaik .....	53
BAB 5 KESIMPULAN DAN SARAN .....	55
5.1 Kesimpulan .....	55
5.2 Saran.....	55
DAFTAR PUSTAKA.....	57
LAMPIRAN .....	61



## DAFTAR TABEL

<b>Tabel 2.1</b> Penelitian Terdahulu.....	7
<b>Tabel 2.2</b> Karakteristik Utama ACF dan PACF.....	12
<b>Tabel 3.1</b> Struktur Data Penelitian .....	35
<b>Tabel 3.2</b> Jadwal Pelaksanaan Kegiatan Penelitian.....	40
<b>Tabel 4.1</b> Jumlah Hari Hujan di Stasiun Juanda.....	7
<b>Tabel 4.2</b> Statistik Deskriptif Curah Hujan Stasiun Juanda .....	12
<b>Tabel 4.3</b> Uji Signifikansi Parameter .....	42
<b>Tabel 4.4</b> Uji Asumsi Residual .....	44
<b>Tabel 4.5</b> Nilai RMSE Model ARIMA .....	45
<b>Tabel 4.7</b> Data Latih dan Target Latih .....	45
<b>Tabel 4.8</b> Data Uji dan Target Uji .....	46
<b>Tabel 4.9</b> Uji Asumsi Residual.....	46

*(Halaman ini sengaja dikosongkan)*

## DAFTAR GAMBAR

<b>Gambar 2.1</b> Grafik Fungsi <i>Generalized Bell</i> .....	20
<b>Gambar 2.2</b> Sistem Inferensi <i>Fuzzy</i> .....	22
<b>Gambar 2.3</b> Arsitektur ANFIS .....	22
<b>Gambar 3.1</b> Diagram Alir .....	39
<b>Gambar 4.1</b> Curah Hujan Harian per Bulan Stasiun Juanda .....	40
<b>Gambar 4.2</b> Grafik curah hujan dalam satuan milimeter (mm).....	41
<b>Gambar 4.3</b> Data Training dan testing Curah Hujan Harian Stasiun Juand .....	41
<b>Gambar 4.4</b> Uji Stasioner dalam Varians .....	42
<b>Gambar 4.5</b> Uji Stasioner dalam Varians .....	43
<b>Gambar 4.6</b> Plot ACF Curah Hujan .....	43
<b>Gambar 4.7</b> Plot PACF Curah Hujan .....	45
<b>Gambar 4.8</b> Grafik Fungsi Keanggotaan.....	45
<b>Gambar 4.9</b> Membership Function Aturan Fuzzy .....	46
<b>Gambar 4.10</b> Grafik Hasil Pelatihan ANFIS .....	47
<b>Gambar 4.11</b> Grafik Hasil Pengujian ANFIS .....	48
<b>Gambar 4.12</b> Grafik Hasil Prediksi dengan data Training untuk N=8 .....	50
<b>Gambar 4.13</b> Grafik Hasil Prediksi dengan data Training untuk N=16 .....	51
<b>Gambar 4.14</b> Grafik Hasil Prediksi dengan data Training untuk N=32 .....	51
<b>Gambar 4.15</b> Grafik Hasil Prediksi dengan data Training untuk N=64 .....	51
<b>Gambar 4.16</b> Grafik Hasil Prediksi dengan data Training untuk N=128 .....	52

*(Halaman ini sengaja dikosongkan)*

## DAFTAR NOTASI

$y_t$	: nilai pada waktu ke $t$
$\phi_i$	: koefisien parameter orde AR, $i : 1,2,3,\dots,p$
$e_t$	: nilai error saat ke $t$ , $e \sim N(0, \sigma_e^2)$
$p$	: orde dari AR
$\theta_q$	: koefisien parameter orde MA
$q$	: orde dari MA
$\phi_p$	: koefisien parameter <i>autoregressive</i> ke- $p$
$\theta_q$	: koefisien parameter <i>moving average</i> ke- $q$
$e_t$	: nilai kesalahan pada saat $t$
$y_{new}$	: data hasil normalisasi
$y_{old}$	: data asli
$y_{min}$	: nilai <i>minimum</i> dari data asli
$y_{max}$	: nilai <i>maximum</i> dari data asli
$d(Y)$	: ukuran ketidakmiripan
$d(x, y)$	: ukuran ketidakmiripan
$Y$	: $(Y_1, Y_2, \dots, Y_i)$ adalah variable data
$y$	: $(y_1, y_2, \dots, y_i)$ adalah pada titik pusat <i>cluster</i>
$C_i$	: <i>cluster</i> ke- $i$
$Y_i$	: Variabel ke- $i$
$t$	: iterasi
$Y_n$	: data inputan ke- $n$ pada <i>cluster</i> tertentu
$n$	: jumlah data pada cluster tertentu
$c$	: rata-rata (mean) tiap cluster
$\sigma$	: standar deviasi tiap cluster
$y_i$	: data input ke- $I$ tiap cluster
$n$	: banyaknya data tiap cluster
$y$	: variabel input
$c$	: nilai rata-rata
$\sigma$	: standar deviasi

- $\theta_{ik}, \theta_{i0}$  : parameter linear  
 $A_{ki}$  : parameter nolinear  
 $i$  :  $1, 2, \dots, m$   
 $K$  :  $1, 2, \dots, p$   
 $c_{j+1,l}$  : koefisien wavelet  
 $c_j$  : data sebelumnya pada data observasi  
 $c_{j+1}$  : data selanjutnya pada data observasi  
 $\omega_{j+1,l}$  : koefisien skala  
 $J$  :  $0, 1, 2, 3, \dots, J$   
 $l$  :  $0, 1, 2, 3, \dots, N$   
 $A_j$  : orde MAR ( $k=1, 2, \dots, A_j$ )  
 $\hat{\alpha}_{j,k}$  : koefisien MAR  
 $\omega_{j,N-2^j(k-1)}$  : koefisien wavelet  
 $c_{j,N-2^j(k-1)}$  : koefisien skala  
 $J$  : banyaknya level dekomposisi  
 $N$  : banyaknya model yang digunakan  
 $Y_t$  : data periode ke  $t$ .  
 $Y_{t-1}$  : data periode ke  $t-1$   
 $Y'_t$  : data hasil differencing pertama pada periode ke  $t$ ,  
 $Y'_{t-1}$  : data hasil differencing pertama pada periode ke  $t-1$ ,  
 $Y''_t$  : data hasil differencing kedua pada periode ke  $t$ .

*(Halaman ini sengaja dikosongkan)*





# BAB 1

## PENDAHULUAN

### 1.1 Latar Belakang

Analisis deret waktu (*time series*) adalah salah satu bahasan dalam ilmu statistika. Model analisis deret waktu yang sering digunakan untuk melakukan peramalan adalah ARIMA (Box dkk, 2016). Menurut Wei (2006) ARIMA adalah salah satu metode peramalan yang biasanya baik digunakan untuk melakukan peramalan jangka pendek. Dengan seiringnya waktu metode-metode tersebut memiliki keterbatasan dalam memodelkan data deret waktu, terutama pada data deret waktu nonlinear dan non stasioner.

Salah satu model yang digunakan untuk peramalan data deret waktu nonlinear dan data deret waktu nonstasioner adalah *neural network* yang dilakukan oleh McCulloch & Pitts (1943), *fuzzy logic* yang dilakukan oleh Zadeh (1965) dan Analisis Wavelet yang dilakukan oleh Abramovich dkk (2000). Teknik *neural network* dan *fuzzy logic* dapat memecahkan masalah kompleks, tetapi keduanya memiliki keterbatasan. Dari keterbatasan keduanya maka dikembangkan suatu metode yang mengkombinasikan kedua teknik itu disebut *hybird*, yaitu *Adaptive Neuro Fuzzy System* (ANFIS) suatu gabungan antara konsep *Backpropagation Neural Network* dengan konsep logika *fuzzy* (Jang dkk, 1993). ANFIS adalah metode yang menggunakan jaringan syaraf tiruan untuk mengimplementasikan sistem inferensi *system inferensi fuzzy*. Dengan begitu, ANFIS adalah penggabungan mekanisme sistem inferensi *fuzzy* yang digambarkan dalam arsitektur jaringan syaraf tiruan. ANFIS adalah suatu metode yang biasa digunakan untuk meramalkan dengan tingkat akurasi yang cukup baik (Singla dkk., 2011). ANFIS terdiri dari 5 lapisan dan setiap lapisannya mempunyai perlakuan berbeda. Pada metode ini, proses pelatihan pada jaringan neural yang terdapat beberapa lapisan dengan sejumlah pasangan data berguna untuk memperbarui parameter-parameter sistem inferensi fuzzy (Fariza, 2017).

Analisis wavelet, telah banyak diaplikasikan untuk permasalahan-permasalahan statistika secara umum. Abramovich dkk, menuliskan review berkaitan dengan aplikasi analisis wavelet dan beberapa penelitian yang telah dilakukan di dalam permasalahan statistika yang mencakup regresi nonparametrik

(Donoho, D.L dan Johnstone, 1995), estimasi densitas (Tribouley, K., 1995), permasalahan titik perubahan dan beberapa aspek khusus dari analisis time series yaitu estimasi densitas spektral untuk proses stasioner dan tidak stasioner (Percival, D.B. dan Walden, A.T., 1999). Selain itu perkembangan terbaru yang terjadi pada beberapa tahun terakhir ini menunjukkan pula bahwa beberapa penelitian telah dilakukan berkaitan dengan aplikasi transformasi wavelet untuk analisis time series (Nason, G.P., 2008). Beberapa kelompok peneliti pertama yang membahas tentang wavelet untuk prediksi data time series berdasarkan suatu model *autoregressive*. Dalam hal ini, transformasi wavelet memberikan suatu dekomposisi yang baik dari suatu signal atau data time series, sehingga struktur selanjutnya dapat dianalisis dengan model-model parametrik atau nonparametrik.

Dari beberapa penelitian yang ada berkaitan dengan wavelet untuk prediksi data time series berdasarkan suatu model *autoregressive*, secara umum fokus utama yang dibahas adalah bagaimana mendapatkan model wavelet time series yang dikenal dengan *Multiresolution Autoregressive (MAR)* yang sesuai untuk peramalan suatu data time series.

Peramalan analisis deret waktu dengan menggunakan metode ARIMA, ANFIS dan Transformasi Wavelet dapat dilakukan untuk keselamatan masyarakat khususnya dalam aktivitas penerbangan. Hujan merupakan peristiwa sumber air yang dapat terjadi sewaktu-waktu. Hujan dapat diartikan sebagai proses turunnya air yang berasal dari kondensasi uap air laut. Hujan juga dapat disimpulkan sebagai peristiwa jatuhnya air dari langit ke tanah. Potensi turunnya hujan sangat dipengaruhi oleh cuaca dan iklim yang terjadi. Hujan merupakan salah satu proses hidrologi yang terjadi di bumi. Siklus hidrologi merupakan suatu siklus dimana air bergerak dari bumi ke atmosfer dan kemudian kembali ke bumi. Hujan terjadi melalui beberapa proses yang diawali oleh proses menguapnya air laut dan air yang berada di dalam tanah. Uap air tersebut kemudian bergerak dan berkumpul menjadi satu membentuk suatu gumpalan awan. Gumpalan awan tersebut kemudian melepaskan titik – titik air hujan (Dutta dan Gouthaman, 2020).

Curah hujan merupakan sekumpulan air yang jatuh ke permukaan tanah selama periode tertentu. Curah hujan di Indonesia sangat beragam sesuai dengan kondisi daerah masing – masing. Saat ini, pola curah hujan di Indonesia diketahui

kurang teratur. Pola curah hujan di Indonesia juga dapat dipengaruhi oleh kondisi geografis seperti adanya deretan pegunungan yang dapat mempengaruhi pergerakan angin. Secara geografis, Indonesia terletak pada garis ekuator yang artinya memiliki musim hujan lebih panjang daripada musim kemarau. Satuan curah hujan di Indonesia adalah millimeter (mm). Satu milimeter memiliki arti genangan air hujan setinggi 1 milimeter yang tidak terserap oleh tanah (Dutta dan Gouthaman, 2020).

Menurut Badan Meteorologi, Klimatologi, dan Geofisika mengatakan bahwa terdapat beberapa jenis hujan yang terjadi di Indonesia berdasarkan curah hujannya. Terdapat enam (5) jenis jenis hujan yaitu hujan ringan, hujan sedang, hujan lebat, hujan sangat lebat, dan hujan ekstrem. Hujan ringan merupakan kondisi dimana hujan terjadi sebesar mm per hari. Hujan sedang merupakan kondisi dimana hujan terjadi sebesar mm per hari. Hujan lebat merupakan kondisi dimana hujan terjadi sebesar mm per hari. Hujan sangat lebat merupakan kondisi dimana hujan terjadi sebesar mm per hari. Hujan ekstrem merupakan kondisi dimana hujan terjadi lebih dari mm per hari.

Karena curah hujan memiliki peran yang sangat penting dalam berbagai aspek kehidupan seperti tinggi rendahnya curah hujan dapat memberikan dampak terhadap dunia penerbangan. Dalam dunia penerbangan, terdapat cuaca ekstrem yang mempunyai potensi membahayakan seperti curah hujan tinggi yang disertai angin kencang, gelombang gunung dan *turbulence* didalam awan. Curah hujan tinggi mempengaruhi pendaratan dan lepas landas pesawat, terlebih lagi jika sering terjadi hujan di bandara akan membuat landasan pacu menjadi licin dan mengurangi jarak pandang penglihatan. Hal ini bisa menyebabkan penundaan penerbangan atau pesawat terlambat mendarat, karena pesawat menunggu cuaca sedikit lebih baik sehingga dapat dilakukan pendaratan dengan aman (Ocktalia Trisnawati & Mike Prastuti, 2021).

Berdasarkan uraian tersebut, maka penelitian dilakukan untuk memprediksi curah hujan dengan membandingkan beberapa model untuk meramalkan data time series yaitu ARIMA, ANFIS dan Transformasi Wavelet.

## 1.2 Rumusan Masalah

Berdasarkan latar belakang dalam upaya untuk menghindari adanya kecelakaan pesawat, pemilihan metode yang tepat untuk menentukan kondisi cuaca seperti curah hujan adalah kegiatan yang akhir-akhir ini sering dilakukan oleh beberapa peneliti atmosfer atau cuaca. Hal ini dikarenakan banyaknya tuntutan dari berbagai pihak lain menuntut tersedianya prediksi atau bahkan ramalan mengenai kondisi atmosfer dengan rentang waktu yang cukup pendek seperti waktu harian, jam, bahkan dalam waktu menit. Penelitian sebelumnya oleh Ocktalia Trisnawati dan Mike Prasusti (2021), penelitian ini menggunakan data curah hujan di Stasiun Juanda pada tahun 2019-2021 menggunakan metode *ARIMA Box-Jenkins* dan *Radial Basis Function Neural Network*. Berdasarkan hasil analisis dan pembahasan peramalan curah hujan di Stasiun Juanda dapat ditarik kesimpulan sebagai berikut: Model ARIMA terbaik untuk meramalkan curah hujan di Stasiun Juanda adalah ARIMA  $([1,9,12,35,55,58],1,1)$ . Model Radial Basis Function (RBFNN) yang optimum adalah model dengan 12 neuron pada input, 3 neuron pada lapisan tersembunyi, dan 1 neuron pada lapisan output. Berdasarkan perbandingan kebaikan model dengan nilai MSE pada model ARIMA dan RBFNN, dapat disimpulkan bahwa model RBFNN lebih baik dalam meramalkan curah hujan di Stasiun Juanda karena memiliki nilai MSE lebih kecil dibandingkan model ARIMA. Berdasarkan uraian tersebut adapun rumusan masalah dalam penelitian ini yaitu:

1. Bagaimana hasil pemodelan pada data curah hujan dengan model ARIMA, ANFIS dan Transformasi Wavelet di Stasiun Juanda?
2. Bagaimana hasil perbandingan dari model ARIMA, ANFIS dan Transformasi Wavelet untuk meramalkan curah hujan di Stasiun Juanda?

## 1.3 Tujuan Penelitian

Berdasarkan Rumusan masalah tujuan dalam penelitian ini yaitu

1. Mendapatkan hasil pemodelan pada data curah hujan dengan model ARIMA, ANFIS dan Transformasi Wavelet di Stasiun Juanda.
2. Mendapatkan model terbaik dari perbandingan model ARIMA, ANFIS dan Transformasi Wavelet serta hasil peramalan curah hujan di Stasiun Juanda.

#### **1.4 Manfaat Penelitian**

Manfaat yang diharapkan dari penelitian ini adalah

1. Dalam bidang statistika, penelitian ini diharapkan dapat digunakan sebagai penerapan metode ARIMA, ANFIS dan Transformasi Wavelet dalam prediksi curah hujan serta mengetahui model yang dihasilkan.
2. Dalam bidang pendidikan, penelitian ini diharapkan dapat digunakan sebagai referensi untuk penelitian selanjutnya yang berhubungan dengan prediksi curah hujan menggunakan metode ARIMA, ANFIS dan Transformasi Wavelet.

#### **1.5 Batasan Masalah**

Berikut ini adalah batasan-batasan masalah dalam penelitian ini:

1. Data yang digunakan merupakan data *time series* curah hujan harian Stasiun Meteorologi Surabaya tahun 2018 – 2023
2. Fungsi keanggotaan yang digunakan untuk ANFIS yaitu fungsi keanggotaan *Generalized Bell*

*(Halaman ini sengaja dikosongkan)*

## BAB 2

### TINJAUAN PUSTAKA

Pada bab ini, akan menjelaskan konsep-konsep dasar yang akan menjadi landasan utama bagi pemahaman menyeluruh terhadap materi yang akan disajikan dalam penelitian ini. Hal ini bertujuan agar pembaca dapat lebih mudah memahami konsep dan metode yang digunakan dalam penelitian ini.

#### 2.1 Penelitian Terdahulu

Berbagai penelitian terkait prediksi curah hujan khususnya menggunakan metode ensemble yang sudah pernah dilakukan. Tabel 2.1 menyajikan beberapa penelitian yang digunakan sebagai referensi penulis.

**Tabel 2.1** Penelitian Terdahulu

Penulis	Judul	Variabel/ Dataset	Metode	Kesimpulan
Ocktalia Trisnawati dan Mike Prasusti (2021)	Peramalan Curah Hujan di Stasiun Juanda Menggunakan Metode ARIMA <i>Box-Jenkins</i> dan <i>Radial Basis Function Neural Network</i>	Data curah hujan di Stasiun Juanda tahun 2019-2021	<i>ARIMA Box Jenkins</i>	hasil analisis permalan curah hujan di Stasiun Juanda dapat ditarik Kesimpulan Model ARIMA terbaik untuk meramalkan curah hujan di stasiun Juanda adalah ARIMA ([1,9,12,35,55,58],1,1).
D.Li & H.Xue (2021)	Prediksi Curah hujan jangka menengah dan Panjang menggunakan model Dekomposisi, Prediksi dan Rekonstruksi Wavelet	Data historis curah hujan yang diukur secara berkala (harian, bulanan, atau tahunan) di lokasi tertentu.	Penelitian ini menggunakan dekomposisi wavelet untuk memecah sinyal curah hujan asli menjadi komponen pendekatan dan detail. Setelah dekomposisi, model prediksi diterapkan pada komponen-komponen tersebut untuk meningkatkan akurasi prakiraan.	Model berbasis wavelet ini berhasil meningkatkan akurasi prediksi dengan lebih baik dalam menangkap tren jangka panjang dan fluktuasi jangka pendek dalam data curah hujan. Model ini menunjukkan kemampuan yang lebih baik dalam menangani kompleksitas data cuaca dibandingkan dengan metode prediksi tradisional.
Kurniawan, R., & Setiawan, B. (2020)	Peramalan curah hujan dengan metode ARIMA dan Jaringan Saraf Tiruan di Stasiun Juanda	Data Curah Hujan	ARIMA dan Jaringan Saraf Tiruan ( <i>Neural Network</i> )	Metode ARIMA dan Jaringan Saraf Tiruan berhasil memberikan prediksi curah hujan dengan akurasi yang cukup baik di Stasiun Juanda. Jaringan Saraf Tiruan menunjukkan performa yang lebih baik dalam menangkap pola nonlinear dibandingkan ARIMA

Lanjutan **Tabel 2.1** Penelitian Terdahulu

Penulis	Judul	Variabel/ Dataset	Metode	Kesimpulan
Hadi, P., & Nasution, M. (2019)	Penggunaan Metode ANFIS dalam peramalan curah hujan di Stasiun Juanda	Curah hujan, suhu, kelembapan dan tekanan atmosfer.	ANFIS (Adaptive Neuro-Fuzzy Inference System)	ANFIS menunjukkan kinerja yang baik dalam meramalkan curah hujan dengan mengakomodasi variasi dan ketidakpastian dalam data. Model ini lebih efektif dalam menangani data yang kompleks.
Muhammad Isradi Azhar (2018)	Prediksi curah hujan menggunakan metode Adaptive Neuro Fuzzy Inference system (ANFIS)	Curah hujan, suhu udara, kelembapan udara dan kecepatan angin kabupaten malang.	JST, Logika Fuzzy, ANFIS	Hasil penelitian yaitu parameter data latih berpengaruh terhadap nilai RMSE. Berdasarkan hasil pengujian, jumlah data yang menghasilkan RMSE terbaik adalah 150 data.
Candra Dewi (2017)	Pengaruh Arsitektur ANFIS pada peramalan cuaca	Data harian suhu, kelembaban udara, kecepatan angin dan tekanan udara.	ANFIS, Generalized Bell.	Hasil uji coba menunjukkan bahwa kombinasi persentase 40%, 50% dan 60% data latih menghasilkan nilai akurasi dan RMSE yang cukup stabil pada beberapa kombinasi neuron (antara 2 sampai 6) pada lapisan pertama dan kedua, disamping itu dapat diketahui bahwa kombinasi jumlah neuron yang tepat antara 2 sampai empat.

## 2.2 ARIMA

Terdapat beberapa macam model deret waktu yang sering digunakan untuk peramalan. Setiap model deret waktu tersebut memiliki karakteristik yang berbeda – beda. Model – model deret waktu tersebut, antara lain *Autoregressive*, *Moving Average*, *Autoregressive Moving Average*, dan *Autoregressive Integrated Moving Average*.

### 2.2.1 Autoregressive

Model *Autoregressive* atau AR merupakan model yang dapat digunakan untuk meramalkan atau memprediksi beberapa waktu ke depan. Model AR merupakan salah satu regresi linier dalam deret waktu yang memenuhi sifat homoskedastisitas. Homoskedastisitas adalah dimana kondisi suatu varian *error* sebuah data bersifat



konstan. Model AR memiliki asumsi, yaitu model AR merupakan proses dari stasioner dengan kovariansi  $(\varepsilon_t, y_{t-k}) = 0$  untuk  $k \neq 0$  dan  $(\varepsilon_t, y_{t-k}) \neq 0$ . Terdapat orde ( $p$ ) dalam model AR yang didefinisikan sebagai  $p \in \mathbb{Z}^+$ . Bentuk umum dari model AR dengan orde  $p$  atau AR( $p$ ) adalah:

$$y_t = \phi_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + e_t \quad (2.1)$$

dengan

$y_t$  : nilai pada waktu ke  $t$

$\phi_i$  : koefisien parameter orde AR,  $i : 1, 2, 3, \dots, p$

$e_t$  : nilai error saat ke  $t$ ,  $e \sim N(0, \sigma_e^2)$

$p$  : orde dari AR

seperti contoh model AR(1) dapat dituliskan sebagai,

$$y_t = \phi_0 + \phi_1 y_{t-1} + e_t \quad (2.2)$$

(Box dkk, 2016)

### 2.2.2 Moving Average

*Moving Average* atau rata – rata bergerak merupakan salah satu metode peramalan deret waktu stasioner. Proses *Moving Average* atau MA menyatakan bahwa nilai data  $t$  dipengaruhi oleh kesalahan sebelumnya. Bentuk umum dari model MA dengan orde  $q$  atau MA( $q$ ) adalah

$$y_t = e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \dots - \theta_q e_{t-q} \quad (2.3)$$

(Box dkk, 2016)

dengan:

$y_t$  : nilai pada waktu ke  $t$

$\theta_q$  : koefisien parameter orde MA

$q$  : orde dari MA

$e_t, e_{t-1}, e_{t-2}, \dots, e_{t-q}$  : nilai-nilai eror pada waktu  $t$

MA tingkat satu dinyatakan sebagai:

$$y_t = e_t - \theta_1 e_{t-1} \quad (2.4)$$

MA tingkat dua dinyatakan sebagai:

$$y_t = e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} \quad (2.5)$$

(Box dkk, 2016).

### 2.2.3 *Autoregressive Moving Average*

*Autoregressive Moving Average* atau dapat disingkat ARMA merupakan model deret waktu stasioner. Model ARMA adalah gabungan antara model AR orde  $p$  dan MA orde  $q$ . Model AR menyatakan bahwa data saat ini dipengaruhi oleh data sebelumnya dan model MA menyatakan bahwa data saat ini dipengaruhi oleh kesalahan data sebelumnya. Proses ARMA dapat ditulis ARMA( $p,q$ ). Bentuk umum dari model ARMA dengan orde  $p,q$  atau ARMA( $p,q$ ) adalah:

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + e_t - \theta_1 e_{t-1} - \dots - \theta_q e_{t-q} \quad (2.6)$$

$y_t$  : nilai pada waktu ke  $t$

$\phi_p$  : koefisien parameter *autoregressive* ke- $p$

$\theta_q$  : koefisien parameter *moving average* ke- $q$

$e_t$  : nilai kesalahan pada saat  $t$

(Box dkk, 2016)

### 2.2.4 *Autoregressive Integrated Moving Average*

Model *Autoregressive Integrated Moving Average* atau ARIMA merupakan ARMA dengan *differencing*. Suatu data *time series* harus stasioner, apabila data *time series* non – stasioner maka harus melalui proses kestasioneran data. Menstasionerkan suatu data dapat dilakukan dengan melakukan pembedaan atau *differencing*. Sehingga pada model ARIMA terdapat proses *differencing*. Banyaknya atau jumlah *differencing* yang dilakukan ditulis dengan  $d$ . Model ARIMA diasumsikan menghasilkan peramalan dengan dipengaruhi oleh data sebelumnya dan saat ini. Model ARIMA dapat ditulis sebagai ARIMA ( $p,d,q$ ). Bentuk umum dari model ARIMA dengan orde  $p,d,q$  atau ARIMA( $p,d,q$ ) adalah:

$$y_t = \phi_0 + (1 + \phi_1)y_{t-1} + (\phi_2 - \phi_1)y_{t-2} + \dots + (\phi_p - \phi_{p-1})y_{t-p} - \phi_p y_{t-p-1} + e_t - \theta_1 e_{t-1} - \dots - \theta_q e_{t-q} \quad (2.7)$$

### 2.2.5 **Prosedur Pemodelan ARIMA**

#### 1. Stasioneritas

Stasioneritas adalah tidak terdapat perubahan yang drastis pada data. Fluktuasi data berada disekitar suatu nilai rata-rata yang konstan, tidak tergantung

pada waktu dan variasi dari fluktuasi tersebut (Makridakis dkk., 1999). Suatu data dapat dikatakan stasioner jika rata-rata dan variasinya konstan, tidak ada unsur *trend* dalam data, dan tidak ada unsur musiman. Ciri-ciri data yang stasioner dalam *mean* yaitu apabila diplot akan sering melewati sumbu *horizontal* dan nilai autokorelasinya akan menurun mendekati nol setelah *lag* kedua atau ketiga. Dengan kata lain data deret waktu yang stasioner adalah data yang tidak mengalami kenaikan atau penurunan yang signifikan atau secara matematis dapat dikatakan bahwa data yang dimiliki berfluktuasi di sekitar nilai rata-rata atau berada di antara dua standar galat (Soejoeti, 1987).

Stasioner dibagi menjadi dua, yaitu :

a. Stasioner dalam *mean* (rata-rata)

Stasioner dalam *mean* (rata-rata) yaitu fluktuasi data berada di sekitar suatu nilai rata-rata yang konstan, tidak tergantung pada waktu dan variansi dari fluktuasi tersebut. Dari bentuk plot data dapat diketahui bahwa data tersebut stasioner atau tidak stasioner. Apabila dilihat dari plot ACF, maka nilai-nilai autokorelasi dari data stasioner akan turun menuju nol sesudah *time lag* (selisih waktu) kedua atau ketiga.

Apabila data tidak stasioner dalam rata-rata maka dapat dilakukan *differencing* atau pembedaan, yaitu deret asli diganti dengan selisih. Jumlah *differencing* disimbolkan dengan *d*. Berikut ini persamaan untuk nilai *differencing* sama dengan 1 ( $d=1$ ):

$$Y'_t = Y_t - Y_{t-1} \quad (2.8)$$

dan untuk  $d=2$ , sebagai berikut:

$$Y''_t = Y'_t - Y'_{t-1}. \quad (2.9)$$

dengan,

$Y_t$  : data periode ke  $t$ .

$Y_{t-1}$  : data periode ke  $t-1$

$Y'_t$  : data hasil differencing pertama pada periode ke  $t$ ,

$Y'_{t-1}$  : data hasil differencing pertama pada periode ke  $t-1$ ,

$Y''_t$  : data hasil differencing kedua pada periode ke  $t$ .

b. Stasioner dalam variansi

Stasioner dalam variansi adalah apabila struktur data dari waktu ke waktu mempunyai fluktuasi data yang tetap atau konstan dan tidak berubah-ubah atau tidak ada perubahan variansi dalam besarnya fluktuasi. Secara visual dapat dibantu dengan menggunakan plot *time series*, yaitu dengan melihat fluktuasi data dari waktu ke waktu. Apabila tidak stasioner dalam variansi maka perlu dilakukan perhitungan dengan metode *Box-Cox* sehingga data tersebut stasioner dalam ragam (Wei, 2006). Transformasi *Box-Cox* adalah transformasi pangkat tunggal  $\lambda$  dengan fungsi transformasi sebagai berikut:

$$T(Y_t) = \frac{Y_t^{\lambda-1}}{\lambda} \quad (2.10)$$

dengan  $\lambda$  menyatakan parameter transformasi.

2. Identifikasi Model

Tahap identifikasi merupakan suatu tahapan yang digunakan untuk mencari atau menentukan nilai  $p$ ,  $d$ , dan  $q$  dengan bantuan *Autocorrelation Function* (ACF) atau fungsi autokorelasi, *Partial Autocorrelation Function* (PACF) atau fungsi autokorelasi parsial, serta penggunaan metode *differencing*. Dengan plot *time series* dapat diketahui pola data dan trend deret pengamatan. Identifikasi model tidak hanya dilakukan dengan melihat plot data, tetapi harus disertai dengan pengetahuan mengenai data yang akan dianalisis. Berdasarkan plot data dan pengetahuan cukup mengenai data, model yang akan dibuat dapat menggunakan parameter sesedikit mungkin (Purnomo, 2015).

Fungsi Autokorelasi (ACF) digunakan untuk melihat apakah ada *Moving Average* (MA) dari suatu deret waktu, yang dalam persamaan ARIMA direpresentasikan oleh besaran  $q$ . Besar nilai  $q$  dinyatakan sebagai banyaknya nilai ACF sejak lag 1 hingga lag ke- $k$  secara berurut yang terletak di luar kepercayaan  $Z$ . Nilai  $d$ , sebagai derajat pembeda (*differencing*) untuk menentukan stasioner atau tidaknya suatu deret waktu (Purnomo, 2015).

Autokorelasi parsial (PACF) digunakan untuk mengukur derajat asosiasi antara  $Y_t$  dan  $Y_{t-k}$  ketika efek dari rentang/jangka waktu (*time lag*) dihilangkan. Nilai PACF juga berkisar antara +1 dan -1. PACF pada umumnya digunakan untuk mengidentifikasi ada atau tidaknya sifat AR (*autoregressive*) yang dinotasikan

dengan besaran  $p$ . Untuk menentukan besar nilai  $p$  yang menyatakan derajat AR, diperlukan perbandingan nilai PACF pada selang kepercayaan  $Z$ . Nilai  $p$  dinyatakan dengan banyaknya nilai PACF sejak lag1 hingga lag ke- $k$  yang terletak di luar selang kepercayaan secara berturut-turut (Purnomo, 2015).

**Tabel 2.2.** Karakteristik Utama dari ACF dan PACF

Proses	Fungsi ACF	Fungsi PACF
AR( $p$ )	Menurun secara bertahap menuju nol	Menuju nol setelah lag $p$ .
MA( $q$ )	Menuju nol setelah lag $q$ .	Menurun secara bertahap menuju nol
ARMA ( $p,q$ )	Menurun secara bertahap menuju nol.	Menurun secara bertahap menuju nol.

(Purnomo, 2015).

### 3. Tahap Estimasi Parameter

Model sementara yang telah diperoleh, selanjutnya dilakukan estimasi parameter. Metode yang digunakan untuk estimasi parameter adalah *least square*. Metode *least square* dapat digunakan untuk menduga parameter ARMA yaitu  $\phi$  dan  $\theta$ . Model ARMA dengan persamaan:

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + e_t - \theta_1 e_{t-1} - \dots - \theta_q e_{t-q} \quad (2.11)$$

Model dugaan untuk ARMA ( $p,q$ ) adalah:

$$\hat{y}_t = \hat{\phi}_1 y_{t-1} + \hat{\phi}_2 y_{t-2} + \dots + \hat{\phi}_p y_{t-p} + e_t - \hat{\theta}_1 e_{t-1} - \dots - \hat{\theta}_q e_{t-q} \quad (2.12)$$

diperoleh galat (error) yaitu:

$$e_t = y_t - \hat{y}_t \quad (2.13)$$

Estimasi parameter ARMA  $\phi$  dan  $\theta$  dilakukan hingga membuat nilai jumlah kuadrat galat minim yaitu  $S(\phi, \theta) = \min \sum_{i=1}^n e_i^2$ . Langkah dasar yang digunakan untuk estimasi parameter menggunakan *least square* yaitu:

1. Membentuk suatu fungsi:

$$S(\phi, \theta) = \sum_{i=1}^n e_i^2 \quad (2.14)$$

2. Mendiferensialkan  $S$  terhadap parameter-parameter didalamnya dan hasilnya sama dengan nol. Sebagai contoh, akan dilakukan estimasi parameter untuk AR(1).

Model AR(1) dari persamaan:

$$y_t = \phi_1 y_{t-1} + e_t \quad (2.15)$$

Persamaan yang ditampilkan tersebut dapat dipandang sebagai model regresi linear dengan variabel respon  $y_t$  dan prediktor  $y_{t-1}$ . Estimasi parameter pada model AR(1) dilakukan dengan mencari nilai  $\phi$  yang meminimalkan jumlah kuadrat galat (error).

Fungsi yang dibentuk dari model AR (1) adalah:

$$S(\phi_1) = \sum_{t=1}^n [y_t - (\phi_1 y_{t-1})]^2 \quad (2.16)$$

Fungsi pada persamaan di atas dijabarkan menjadi persamaan berikut:

$$S(\phi_1) = \sum_{t=1}^n (y_t^2 - 2y_t \phi_1 y_{t-1} + \phi_1^2 y_{t-1}^2) \quad (2.17)$$

Setelah persamaan dijabarkan, lalu didiferensialkan dan disamakan dengan nol menjadi:

$$\begin{aligned} \frac{\partial S}{\partial \phi_1} &= \sum_{t=1}^n -2y_t y_{t-1} + 2\phi_1 \sum_{t=1}^n y_{t-1}^2 = 0 \\ \Rightarrow 2\phi_1 \sum_{t=1}^n y_{t-1}^2 &= 2 \sum_{t=1}^n y_t y_{t-1} \\ \Leftrightarrow \phi_1 &= \frac{2 \sum_{t=1}^n y_t y_{t-1}}{2 \sum_{t=1}^n y_{t-1}^2} \end{aligned} \quad (2.18)$$

Berdasarkan persamaan di atas, maka estimasi parameter untuk  $\hat{\phi}_1$  dapat diperoleh menggunakan persamaan tersebut.

#### 4. Uji Signifikansi Parameter

Setelah dilakukan estimasi parameter maka parameter tersebut perlu diuji signifikansinya untuk mengetahui apakah parameter tersebut dapat dimasukkan kedalam model dengan uji hipotesis sebagai berikut:

Uji signifikan parameter model pada parameter *autoregressive*, yaitu:

- a. Menetapkan hipotesis:

$$H_0: \phi_i = 0$$

$$H_1: \phi_i \neq 0, \text{ dengan } i = 1, 2, \dots, p$$

- b. Statistik uji:

$$t = \frac{\hat{\phi}_i}{SE(\hat{\phi}_i)}, \quad (2.19)$$

- c. Kriteria penolakan  $H_0$ :

Tolak  $H_0$  apabila  $|t| > t_{\alpha, df}$  atau  $p_{value} < \alpha$ , dengan  $df = T - p$  dengan  $T$  adalah banyaknya pengamatan dan  $p$  adalah banyaknya parameter dalam model.

(Bowerman, O'connell & Koehler, 2005).

Uji signifikan parameter model pada parameter *moving average*, yaitu:

- a. Menetapkan hipotesis:

$$H_0: \theta = 0$$

$$H_1: \theta \neq 0, \text{ dengan } i = 1, 2, \dots, q$$

- b. Statistik uji:

$$t = \frac{\hat{\theta}_i}{SE(\hat{\theta}_i)} \quad (2.20)$$

- c. Kriteria penolakan  $H_0$ :

Tolak  $H_0$  apabila  $|t| > t_{\alpha, df}$  atau  $p_{value} < \alpha$ , dengan  $df = T - q$  dengan  $T$  adalah banyaknya pengamatan dan  $q$  adalah banyaknya parameter dalam model.

(Bowerman, O'connell & Koehler, 2005).

#### 5. Pemeriksaan diagnostik

Setelah melakukan estimasi nilai-nilai parameter dari model ARIMA yang ditetapkan sementara, selanjutnya dilakukan pemeriksaan diagnostik untuk

membuktikan bahwa model tersebut cukup memadai dan menentukan model mana yang terbaik digunakan untuk peramalan (Makridakis dkk., 1999).

Pemeriksaan diagnostik ini dapat dilakukan dengan mengamati apakah residual dari model terestimasi merupakan *white noise* atau tidak. Asumsi *white noise* adalah residual bersifat acak dan berdistribusi normal. Uji yang dilakukan adalah uji *Ljung-Box* untuk melihat apakah residual bersifat acak dan uji *Kolmogorov-Smirnov* untuk melihat apakah residual berdistribusi normal. Jika model tidak memenuhi kedua uji tersebut maka model tersebut tidak layak, maka langkah selanjutnya adalah ulangi langkah mulai identifikasi model sampai mendapatkan model yang layak atau baik. (Nachrowi & Hardius, 2006).

### **2.3 Adaptive Neuro Fuzzy Inference System (ANFIS)**

*Adaptive Neuro Fuzzy Inference System* (ANFIS) adalah jaringan yang berbasis pada sistem inferensi *fuzzy*. Parameter ANFIS dapat dipisahkan menjadi dua, yaitu parameter premis dan konsekuen yang dapat diadaptasikan dengan pelatihan *hybrid*. Pelatihan *hybrid* dilakukan dalam dua langkah, yaitu langkah maju dan langkah mundur. *Adaptive Neuro Fuzzy Inference System* (ANFIS) adalah penggabungan mekanisme *fuzzy inference system* yang digambarkan dalam arsitektur jaringan saraf. Sistem inferensi yang digunakan adalah sistem inferensi *fuzzy* model Takagi-Sugeno-Kang (TSK) orde satu dengan pertimbangan kesederhanaan dan kemudahan komputasi.

Arsitektur *Adaptive Neuro Fuzzy Inference System* (ANFIS) juga sama dengan jaringan saraf dengan fungsi radial dengan batasan tertentu. Agar jaringan dengan fungsi basis radial ekuivalen dengan *fuzzy* berbasis aturan model Sugeno orde satu, maka diperlukan batasan:

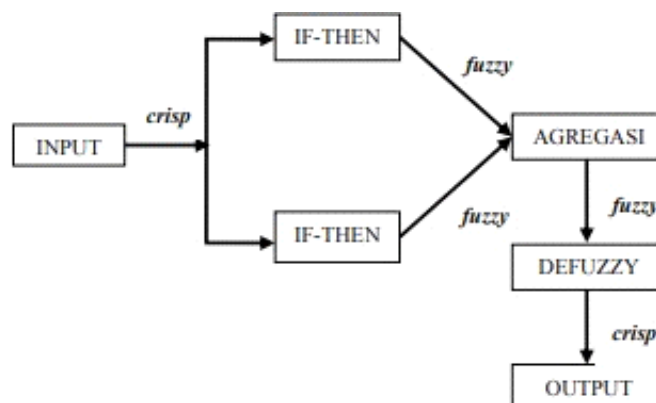
1. Keduanya harus memiliki metode agregasi yang sama (rata-rata terbobot atau penjumlahan terbobot) untuk menurunkan semua outputnya.
2. Jumlah fungsi aktivasi harus sama dengan jumlah aturan *fuzzy* (*IF-THEN*).
3. Jika ada beberapa input pada basis aturannya, maka tiap-tiap fungsi aktivasi harus sama dengan fungsi keanggotaan tiap-tiap inputnya.
4. Fungsi aktivasi dan aturan-aturan *fuzzy* harus memiliki fungsi yang sama untuk neuron-neuron dan aturan-aturan yang ada di sisi outputnya.



### 2.1.1 Sistem Inferensi Fuzzy (*Fuzzy Inference System*)

Sistem inferensi *fuzzy* adalah suatu kerangka komputasi yang didasarkan pada teori himpunan *fuzzy*. Aturan *fuzzy* berbentuk *if-then* dan penalaran *fuzzy*. Sistem *fuzzy* menerima *input crisp*, lalu input dikirimkan ke basis pengetahuan yang berisi *n* aturan *fuzzy* dalam bentuk *if-then*. *Fire strength* akan dicari dalam setiap aturan. Jika jumlah aturan lebih dari satu, maka akan dilakukan agregasi dari semua aturan. Pada hasil agregasi akan dilakukan *defuzzy* untuk mendapatkan nilai *crisp* sebagai sistem keluaran (*output system*) (Kusumadewi & Hartati, 2006).

Terdapat tiga metode dalam sistem inferensi *fuzzy* diantaranya adalah *fuzzy inference system* metode Tsukamoto, Mamdani dan Sugeno. Metode Tsukamoto adalah didasarkan pada konsep penalaran monoton (monoton naik ataupun monoton turun). Nilai *crisp* pada daerah konsekuen dapat diperoleh secara langsung berdasarkan *fire strength*. Metode Mamdani sering dikenal sebagai metode *Max-Min*. metode ini diperkenalkan oleh Ebrahim Mamdani pada tahun 1975. Untuk mendapatkan *output*, diperlukan 4 tahap yaitu pembuatkn himpunan *fuzzy*, aplikasi fungsi implikasi, komposisi aturan dan penegasan (*defuzzy*). Pada penelitian ini menggunakan metode Sugeno. Penalaran metode Sugeno hampir sama dengan penalaran Mamdani, hanya saja *output* sistem tidak berupa himpunan *fuzzy*, melainkan berupa konstanta atau persamaan linear. Metode ini diperkenalkan oleh Takagi-Sugeno Kang pada tahun 1985, sehingga metode ini sering juga dinamakan dengan metode TSK (Kusumadewi & Purnomo, 2010).



**Gambar 2.6.** Sistem Inferensi *Fuzzy*

Metode TSK adalah salah satu metode dalam sistem inferensi *fuzzy*. Dalam metode TSK ini terdapat aturan *Fuzzy* yang terdiri dari suatu himpunan aturan *fuzzy IF-THEN* yang memetakan *input* ke *output*. pemodelan data runtun waktu dengan  $p$  input  $y_{t-1}, y_{t-2}, \dots, y_{t-p}$  dan satu output  $y_t$  dengan mengasumsikan aturan dasar Sugeno orde satu dengan  $m$  aturan, aturan *IF-THEN* dalam *fuzzy* didiskripsikan sebagai:

Rule 1: IF ( $y_{t-1}$  is  $A_{11}$ ) and ( $y_{t-2}$  is  $A_{21}$ ) ... and  $y_{t-p}$  is  $A_{i1}$  then (output is  $y_t^{(1)}$ )

Rule 2: IF ( $y_{t-1}$  is  $A_{12}$ ) and ( $y_{t-2}$  is  $A_{22}$ ) ... and  $y_{t-p}$  is  $A_{i2}$  then (output is  $y_t^{(2)}$ )

⋮

Rule  $i$  : IF ( $y_{t-1}$  is  $A_{1p}$ ) and ( $y_{t-2}$  is  $A_{2p}$ ) ... and  $y_{t-p}$  is  $A_{ip}$  then (output is  $y_t^{(p)}$ )

dimana:

$y_t$  : variabel target atau keluaran pada waktu  $t$ .

$y_{t-p}$  : variabel input atau data masa lalu yang digunakan untuk memprediksi  $y_t$ .

$A_{ij}$  : fungsi keanggotaan yang menunjukkan seberapa besarnya  $y_{t-p}$  berkontribusi terhadap aturan ke  $i$ .

### 2.1.2 Teori Himpunan *Fuzzy*

Teori himpunan *fuzzy* pertama kali diperkenalkan oleh Zadeh (1965). Himpunan *fuzzy* adalah suatu kelas objek dengan rangkaian nilai keanggotaan. Himpunan tersebut ditandai dengan fungsi keanggotaan yang memberikan nilai keanggotaan pada setiap objek dengan rentang nilai antar nol dan satu (Jang dkk., 1997).

Teori *fuzzy* adalah suatu perluasan dari teori himpunan klasik. Suatu nilai yang menunjukkan seberapa besar tingkat keanggotaan suatu elemen ( $y$ ) dalam suatu himpunan ( $A$ ) sering disebut dengan nilai keanggotaan atau derajat keanggotaan, dinotasikan dengan  $\mu_A(y)$ . Nilai keanggotaan tersebut dipetakan ke dalam suatu kurva yang disebut fungsi keanggotaan (*membership function*) (Kusumadewi & Hartati, 2006).

Fungsi keanggotaan (*membership function*) adalah suatu kurva yang menunjukkan pemetaan titik-titik *input* kedalam nilai keanggotaannya. Untuk

mendapatkan nilai keanggotaannya adalah dengan melalui pendekan fungsi. Nilai keanggotaan hanya terdapat dua kemungkinan, yaitu 0 dan 1, dimana 0 adalah nilai minimum dan 1 adalah nilai maksimum. Salah satu fungsi keanggotaan yang biasa digunakan yaitu Fungsi keanggotaan Gaussian. Rumus umum dari persamaan fungsi *Gaussian* adalah:

$$\mu_i(y) = \frac{1}{1 + \left| \frac{y-c}{a} \right|^{2b}}, \quad (2.21)$$

dengan:

$y$  : variabel input

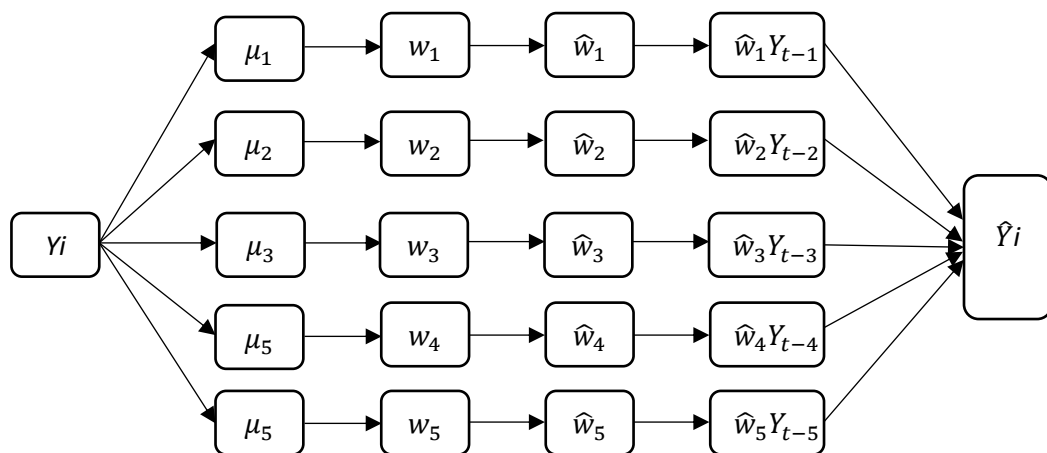
$c$  : parameter yang menentukan lokasi pusat kurva.

$a$  : parameter yang mengontrol lebar kurva. semakin besar nilai  $a$  semakin lebar kurva tersebut.

$b$  : parameter yang mengontrol kelancipan kurva. semakin besar nilai  $b$ , semakin lancip kurvanya.

### 2.1.2 Arsitektur *Adaptive Neuro Fuzzy Inference System (ANFIS)*

Menurut Jang *et al*, arsitektur ANFIS mempunyai jumlah hidden layer pada NN yang disesuaikan dengan sistem fuzzy dan dikelompokkan ke dalam beberapa bagian yaitu: fuzzyfikasi (lapisan 1), sistem inferensi fuzzy (lapisan 2, lapisan 3), defuzzyfikasi (lapisan 4) dan agregasi (lapisan 5).



**Gambar 2.7.** Arsitektur ANFIS

(Jang, et al, 1997).

### 1. Lapisan 1

Setiap titik pada lapisan 1 bersifat adaptif dengan satu fungsi aktivasi. Output lapisan ini merupakan derajat keanggotaan variabel input yang memenuhi fungsi keanggotaan  $\mu(y)$ . Salah satu contoh fungsi keanggotaan adalah *Generalized Bell*.

$$\mu_i(y) = \frac{1}{1 + \left| \frac{y-c}{a} \right|^{2b}}, \quad (2.22)$$

### 2. Lapisan 2

Setiap titik pada lapisan 2 ini adalah titik dengan nilai output berasal dari hasil perkalian dari derajat keanggotaan yang diperoleh pada lapisan pertama. Output setiap titik menyatakan *firing strength*  $w_i$  dari aturan ke  $i$ . Oleh karena itu, output yang dihasilkan dari lapisan 2 dapat dirumuskan sebagai berikut:

$$w_i = \prod_{i=1}^5 \mu_{Ai}(y), \quad (2.23)$$

### 3. Lapisan 3

Setiap titik pada lapisan 3 merupakan titik tetap berlabel N, yang menghitung rasio *firing strength* aturan ke  $i$  relatif terhadap jumlahan *firing strength* semua aturan. Output yang dihasilkan oleh lapisan 3 dapat dituliskan dalam persamaan berikut ini:

$$\bar{w}_i = \frac{w_i}{\sum_{i=1}^5 w_i}, \quad (2.24)$$

Output pada lapisan ketiga disebut dengan *normalized firing strength*

### 4. Lapisan 4

Pada lapisan keempat terdapat parameter baru yang disebut parameter konsekuen yang dapat diperoleh dengan menggunakan dekomposisi nilai singular. Setiap titik pada lapisan 4 adalah titik adaptif dengan fungsi node sebagai berikut:

$$\bar{w}_i y_t^{(i)} = \bar{w}_i (\theta_{i1} y_{t-1} + \theta_{i2} y_{t-2} + \dots + \theta_{ip} y_{t-p} + \theta_{i0}) \quad (2.25)$$

dimana,  $\bar{w}_i$  merupakan *normalized firing strength* lapisan ke 3 dan parameter  $\theta_i$  menyatakan parameter konsekuen yang adaptif. Jika diketahui keluaran dari model linear  $y$  yang diekspresikan melalui persamaan (2.35) yaitu:

$$y = \theta_{i1} y_{t-1} + \theta_{i2} y_{t-2} + \dots + \theta_{ip} y_{t-p} + \theta_{i0} \quad (2.26)$$

Dengan menggunakan notasi matriks didapatkan:

$$\mathbf{A}\theta = \mathbf{Y} \quad (2.27)$$

Penyelesaian terbaik untuk  $\theta$ , yang meminimalkan  $\|\mathbf{A}\theta - \mathbf{y}\|^2$  adalah least square estimator (LSE):

$$\theta^* = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{Y} \quad (2.28)$$

Dimana  $\mathbf{A}^T$  adalah bentuk transpose dari  $\mathbf{A}$ .

## 5. Lapisan 5

Lapisan kelima hanya memiliki satu simpul output yang merupakan keluaran atau hasil dari seluruh perhitungan yang telah dilakukan pada lapisan sebelumnya. Nilai output pada lapisan kelima diperoleh dengan menggunakan persamaan berikut:

$$y_t = \sum_{i=1}^m \bar{w}_i y_t^{(i)} \quad (2.29)$$

dimana nilai ( $\bar{w}_i$ ) diperoleh pada lapisan ketiga dan koefisien atau yang disebut parameter konsekuen diperoleh pada lapisan keempat Jaringan adaptif dengan lima lapisan tersebut ekuivalen dengan sistem inferensi fuzzy TSK (Jang, et al, 1997).

## 2.4 Transformasi Wavelet

Transformasi merupakan suatu proses pengubahan data ke bentuk data lain yang mudah di analisis, sebagai misal transformasi fourier merupakan suatu proses pengubahan data (sinyal) kedalam beberapa gelombang kosinus yang berfrekuensi berbeda, sedangkan transformasi wavelet merupakan proses pengubahan sinyal kedalam berbagai wavelet basis (mother wavelet) dengan sebagai fungsi pergeseran dan penyekalan (I Gede Pasek S.W,2004).

Transformasi wavelet merupakan perbaikan dari transformasi fourier Dimana transformasi fourier hanya dapat menangkap informasi apakah suatu sinyal memiliki frekuensi tertentu atukah tidak tapi tidak dapat menangkap Dimana frekuensi itu terjadi. Transformasi wavelet dapat digunakan untuk menganalisa deret waktu yang mengandung daya non stasioner pada frekuensi yang berbeda. Transformasi wavelet dari  $x(t)$  atau sinyal hujan menghasilkan *approximation coefficient* dan *detail coefficient* dari data hujan. *approximation coefficient* berhubungan dengan skala tinggi dan komponen frekwensi rendah. Sedangkan *detail coefficient* berhubungan dengan skala rendah dan komponen frekwensi

tinggi. Koefisien-koefisien ini selanjutnya direntang ke depan (extend) sepanjang selang waktu tertentu sebelum kemudian disatukan kembali.

Peramalan data deret waktu yang dihasilkan dengan menggunakan wavelet adalah peramalan dengan menggunakan data yang telah mengalami pra pemrosesan melalui transformasi wavelet. Dengan adanya komponen-komponen multiskala maka secara otomatis akan memisahkan komponen-komponen data, yaitu komponen trend dan irregular yang ada pada data stasioner (mengandung komponen irregular saja), ataupun data yang nonstasioner (mengandung komponen trend dan irregular). Dengan demikian dapat dilakukan peramalan data dengan baik.

### 2.3.1 Fungsi Wavelet

Wavelet diartikan suatu gelombang kecil (*small wave*) yang mempunyai kemampuan mengelompokkan sinyal dan terkonsentrasi dalam waktu tertentu serta naik pada periode tertentu. Fungsi wavelet terdiri dari dua fungsi yaitu wavelet ayah ( $\varphi$ ) dan wavelet ibu ( $\psi$ ) yang memiliki persamaan ditulis sebagai:

$$\int_{-\infty}^{\infty} \varphi(y) dy = 1 \quad (2.30)$$

dan

$$\int_{-\infty}^{\infty} \psi(y) dy = 0 \quad (2.31)$$

Dengan proses dilatasi dan translasi, wavelet ayah ( $\varphi$ ) dan wavelet ibu ( $\psi$ ) menghasilkan keluarga wavelet, yaitu:

$$\varphi_{j,k}(y) = (p2^j)^{1/2} \varphi(p2^j y - k) \quad (2.32)$$

dan

$$\psi_{j,k}(y) = (p2^j)^{1/2} \psi(p2^j y - k) \quad (2.33)$$

dengan

$j = j_0, j_1, \dots, j_n, n \in Z, j$  menunjukkan banyak komponen level transformasi dan  $n$  menunjukkan banyak koefisien dalam suatu level transformasi (Friedman dkk., 2008).

### 2.3.2 Filter Skala

Fungsi skala merupakan filter *low pass*. Menurut (Pervical dan Walden, 2000), fungsi skala memenuhi tiga kondisi yaitu:

$$\begin{aligned}
1. \quad & \sum_{l=0}^{L-1} g_l = \sqrt{2} \text{ atau } \sum_{l=0}^{L-1} g_l = -\sqrt{2} \\
2. \quad & \sum_{l=0}^{L-1} g_l^2 = 1 \\
3. \quad & \sum_{l=0}^{L-1} g_l g_{l+2n} = 0
\end{aligned} \tag{2.34}$$

### 2.3.3 Filter Wavelet

Filter wavelet atau filter *high-pass* merupakan filter yang bersifat *smooth* yang didefinisikan sebuah deret bilangan *real* yang membangun filter *wavelet*  $\{h_l : l = 0, \dots, L - 1\}$ , dengan  $L$  adalah lebar filter dan merupakan bilangan bulat. Menurut (Pervical dan Walden, 2000), filter Wavelet memenuhi tiga kondisi dasar yaitu:

$$\begin{aligned}
1. \quad & \sum_{l=0}^{L-1} h_l = 0 \\
2. \quad & \sum_{l=0}^{L-1} h_l^2 = 1 \\
3. \quad & \sum_{l=0}^{L-1} h_l h_{l+2n} = 0
\end{aligned} \tag{2.35}$$

Untuk semua  $n$  integer dengan  $h_l = 0$ , untuk  $l > 0$  dan  $l \geq L$ , dapat dikatakan  $\{h_l\}$  adalah deret tak hingga.

### 2.3.4 Transformasi Wavelet Diskrit

Transformasi wavelet adalah fungsi transfer (transform) yang digunakan untuk menguraikan data atau fungsi atau operator menjadi komponen frekuensi yang berbeda-beda dan kemudia mempelajarinya dengan resolusi yang disesuaikan dengan skalanya (daubchies, 1992). Transformasi wavelet terbagi menjadi dua, yaitu transformasi wavelet diskrit dan transformasi wavelet kontinu. Transformasi wavelet diskrit dapat digunakan dalam analisis runtun waktu. Menurut (Ramadhan & Setiyono, 2019:8) pada penskalaan dan fungsi wavelet, fungsi wavelet diskrit ditulis sebagai:

$$\begin{aligned}
f(t) = & \frac{1}{\sqrt{N}} \sum_{l=1}^{\infty} C_{j,l} \varphi_{j,l}(t) \\
& + \frac{1}{\sqrt{N}} \sum_{l=1}^{\infty} \sum_{j=1}^{\infty} d_{j,l} \psi_{j,l}(t)
\end{aligned} \tag{2.36}$$

dengan  $N$  merupakan jumlah data,  $j$  merupakan jumlah komponen multiresolusi atau skala dan  $k$  antara 1 sampai jumlah koefisien dalam komponen yang telah ditentukan. Koefisien  $C_{j,k}$  dan  $d_{j,k}$  merupakan koefisien transformasi wavelet. Fungsi  $\varphi_{j,k}(t)$  dan  $\psi_{j,k}(t)$  merupakan fungsi pendekatan wavelet. Menurut

(Subanar dan Vemmie, 2015) koefisien wavelet dan skala transformasi wavelet diskrit disimbolkan  $W_{j,t}$  dan  $V_{j,t}$  ditulis sebagai

$$W_{j,t} = \sum_{l=0}^{L-1} h_l X_{(2t-l) \bmod N} \quad (2.37)$$

dan

$$V_{j,t} = \sum_{l=0}^{L-1} g_l X_{(2t-l) \bmod N} \quad (2.38)$$

dengan  $t = 0, 1, \dots, \frac{N}{2} - 1$ .  $W_{j,t}$  dinyatakan sebagai koefisien wavelet transformasi wavelet diskrit pada level ke- $j$  dan pada waktu ke- $t$ ,  $V_{j,t}$  dinyatakan sebagai koefisien skala transformasi wavelet diskrit pada level ke- $j$  dan pada waktu ke- $t$ .

Proses menghitung koefisien wavelet transformasi wavelet diskrit pada level pertama akan mengurai data runtun waktu  $X$  menjadi koefisien  $W_1$  dan  $V_1$  yang sama besar dengan masing-masing koefisien terdiri dari  $\frac{N}{2}$  koefisien wavelet  $W_1$  dan  $\frac{N}{2}$  koefisien wavelet  $V_1$ . Tahap pertama atau level pertama ini menggunakan persamaan (2.37) sebagai filter wavelet dan persamaan (2.38) sebagai filter skala. Tahap kedua atau level kedua memproses  $V_{1,t}$  dengan cara yang sama seperti sebelumnya. Pada level kedua ini  $V_{1,t}$  diasumsikan sebagai rata-rata dari skala unit. Hal itu dikarenakan  $V_{1,t}$  merupakan koefisien skala yang bersifat detail atau sinyal yang lebih dekat dengan sinyal aslinya (Verma, 2015:4). Level kedua ini akan menghasilkan koefisien wavelet level kedua atau  $W_{2,t}$  dan  $V_{2,t}$ . koefisien-koefisien tersebut diperoleh dengan menggunakan filter yang ditulis sebagai

$$W_{2,t} = \sum_{l=0}^{L-1} h_l W_{1,(2t-l) \bmod \frac{N}{2}} \quad (2.39)$$

dan

$$V_{2,t} = \sum_{l=0}^{L-1} g_l V_{1,(2t-l) \bmod \frac{N}{2}} \quad (2.40)$$

Level kedua ini menguraikan koefisien  $V_{1,t}$  dan diperoleh koefisien  $W_2$  dan  $V_2$  menggunakan persamaan 9 sebagai filter wavelet dan persamaan 10 sebagai filter skala, proses ini berlaku hingga level terakhir atau maksimum. Menurut Sutarno



(2015:5), penentuan level maksimal dalam transformasi wavelet diskrit ditulis sebagai

$$Level_{maks} = \frac{\ln(N)}{\ln(2)} \quad (2.41)$$

### 2.3.5 Wavelet Daubechies

Wavelet Daubechies adalah salah satu keluarga wavelet orthogonal. Wavelet orthogonal adalah wavelet yang transformasi wavelet-nya bersifat orthogonal. Nama Daubechies diambil dari nama penemunya yaitu Ingrid Daubechies. Selain itu, wavelet ini sering digunakan karena baik untuk kompresi data. Wavelet Daubechies disimbolkan dengan dbN, dengan  $N$  adalah angka indeks mulai dari 2 sampai 20 untuk indeks  $N = 1$  disebut juga haar, Panjang wavelet Daubechies adalah  $2N$  (Burrus dkk,1998).

Persamaan umum untuk semua wavelet ditulis sebagai

$$\varphi(t) = \sqrt{2} \sum_{l=0}^{N-1} h_l \varphi(2t - l) \quad (2.42)$$

dan

$$\psi(t) = \sqrt{2} \sum_{l=0}^{N-1} g_l \psi(2l - k) \quad (2.43)$$

Transformasi Daubechies perlu adanya filter skala dan filter wavelet. Filter skala adalah filter low-pass dan filter wavelet adalah filter high-pass. Menurut (Burrus dkk, 1998) dalam D4 terdapat empat filter skala yaitu  $g(0), g(1), g(2),$  dan  $g(3)$  yaitu

$$\begin{aligned} g(0) &= \frac{(1-\cos(\alpha)+\sin(\alpha))}{2\sqrt{2}}, \\ g(1) &= \frac{(1+\cos(\alpha)+\sin(\alpha))}{2\sqrt{2}}, \\ g(2) &= \frac{(1+\cos(\alpha)-\sin(\alpha))}{2\sqrt{2}}, \\ g(3) &= \frac{(1-\cos(\alpha)-\sin(\alpha))}{2\sqrt{2}} \end{aligned} \quad (2.44)$$

dengan  $\alpha = \frac{\pi}{3}$ , sehingga diperoleh

$$\begin{aligned} g(0) &= \frac{(1-\cos(\alpha)+\sin(\alpha))}{2\sqrt{2}} \Leftrightarrow \frac{1+\sqrt{3}}{4\sqrt{2}} \\ g(1) &= \frac{(1+\cos(\alpha)+\sin(\alpha))}{2\sqrt{2}} \Leftrightarrow \frac{3+\sqrt{3}}{4\sqrt{2}} \end{aligned}$$

$$\begin{aligned}
g(2) &= \frac{(1+\cos(\alpha)-\sin(\alpha))}{2\sqrt{2}} \Leftrightarrow \frac{3-\sqrt{3}}{4\sqrt{2}} \\
g(3) &= \frac{(1-\cos(\alpha)-\sin(\alpha))}{2\sqrt{2}} \Leftrightarrow \frac{1-\sqrt{3}}{4\sqrt{2}}
\end{aligned} \tag{2.45}$$

berdasarkan 1,2,3 dan 4 maka  $g(0), g(1), g(2),$  dan  $g(3)$  memenuhi tiga sifat filter skala yaitu

$$\begin{aligned}
\text{i. } g(0) + g(1) + g(2) + g(3) &= \frac{1+\sqrt{3}}{4\sqrt{2}} + \frac{3+\sqrt{3}}{4\sqrt{2}} + \frac{3-\sqrt{3}}{4\sqrt{2}} + \frac{1-\sqrt{3}}{4\sqrt{2}} = \sqrt{2} \\
\text{ii. } g^2(0) + g^2(1) + g^2(2) + g^2(3) &= \frac{4+2\sqrt{3}}{32} + \frac{12+6\sqrt{3}}{32} + \frac{12-6\sqrt{3}}{32} + \frac{4-2\sqrt{3}}{32} = 1 \\
\text{iii. } g(0) \times g(1) + g(2) \times g(3) &= \left(\frac{1+\sqrt{3}}{4\sqrt{2}} \times \frac{3+\sqrt{3}}{4\sqrt{2}}\right) + \left(\frac{3+\sqrt{3}}{4\sqrt{2}} \times \frac{1-\sqrt{3}}{4\sqrt{2}}\right) = 0
\end{aligned} \tag{2.46}$$

dengan demikian diperoleh filter skala untuk wavelet Daubchies (D4) atau dengan Panjang filter empat ditulis sebagai

$$g_0 = \frac{1+\sqrt{3}}{4\sqrt{2}}, g_1 = \frac{3+\sqrt{3}}{4\sqrt{2}}, g_2 = \frac{3-\sqrt{3}}{4\sqrt{2}} + g_3 = \frac{1-\sqrt{3}}{4\sqrt{2}} \tag{2.47}$$

untuk filter wavelet D4 menggunakan rumus

$$h(l) = (-1)^l g_{L-1-l} \tag{2.48}$$

sehingga diperoleh

$$\begin{aligned}
\text{i. } h(0) &= (-1)^0 g_{4-1-0} \Leftrightarrow \frac{1-\sqrt{3}}{4\sqrt{2}} \\
\text{ii. } h(1) &= (-1)^1 g_{4-1-1} \Leftrightarrow \frac{-3+\sqrt{3}}{4\sqrt{2}} \\
\text{iii. } h(2) &= (-1)^2 g_{4-1-2} \Leftrightarrow \frac{3+\sqrt{3}}{4\sqrt{2}} \\
\text{iv. } h(3) &= (-1)^3 g_{4-1-3} \Leftrightarrow \frac{-1-\sqrt{3}}{4\sqrt{2}}
\end{aligned} \tag{2.49}$$

berdasarkan 1,2,3 dan 4 maka  $h(0), h(1), h(2),$  dan  $h(3)$  memenuhi tiga sifat filter wavelet yaitu:

$$\begin{aligned}
\text{i. } h(0) + h(1) + h(2) + h(3) &= \frac{1-\sqrt{3}}{4\sqrt{2}} + \frac{-3+\sqrt{3}}{4\sqrt{2}} + \frac{3+\sqrt{3}}{4\sqrt{2}} + \frac{-1-\sqrt{3}}{4\sqrt{2}} = 0 \\
\text{ii. } h^2(0) + h^2(1) + h^2(2) + h^2(3) &= \frac{4-2\sqrt{3}}{32} + \frac{12-6\sqrt{3}}{32} + \frac{12+6\sqrt{3}}{32} + \frac{4+2\sqrt{3}}{32} = 1 \\
\text{iii. } h(0) \times h(2) + h(1) \times h(3) &= \left(\frac{1-\sqrt{3}}{4\sqrt{2}} \times \frac{3+\sqrt{3}}{4\sqrt{2}}\right) + \left(\frac{-3+\sqrt{3}}{4\sqrt{2}} \times \frac{-1-\sqrt{3}}{4\sqrt{2}}\right) = 0 \tag{2.50}
\end{aligned}$$

dengan demikian diperoleh filter wavelet untuk wavelet Daubchies (D4) atau dengan panjang filter empat ditulis sebagai

$$h_0 = \frac{1-\sqrt{3}}{4\sqrt{2}}, h_1 = \frac{-3+\sqrt{3}}{4\sqrt{2}}, h_2 = \frac{3+\sqrt{3}}{4\sqrt{2}}, h_3 = \frac{-1-\sqrt{3}}{4\sqrt{2}}$$

(2.51)

### 2.3.6 Multiresolusi Autoregresi (MAR)

Beberapa fungsi wavelet yang dikembangkan oleh para ahli yaitu wavelet Haar dengan Analisis Multi Resolusi (AMR) Daubchies membangun beberapa system wavelet yang masing-masing diberi nama wavelet Daubechies 2, 3, 4, 6 dan 10 dan biasa dinotasikan dengan D2, D3, D4, D6 dan D10 (Daubechies, 1996). Terkait dengan data time series, wavelet Daubechies mempunyai beberapa keuntungan pada resolusi yang lebih baik untuk data time series yang perubahannya lebih smooth.

Dalam analisis wavelet terdapat 2 fungsi yang digunakan yaitu fungsi skala dan fungsi wavelet. Kedua ini menghasilkan suatu “keluarga” fungsi yang dapat digunakan untuk merekonstruksi suatu signal. Selanjutnya akan dibahas fungsi skala dan fungsi wavelet.

Model prediksi Multiresolusi Autoregresi (MAR) sebenarnya adalah linear. Model prediksi jika diperluas pada neural networks, bahkan dapat diperluas menjadi suatu model linear dan non linier, yang menggunakan koefisien.

Tahapan model prediksi Multiresolusi Autoregresi (MAR):

1. Tetapkan nilai  $N$  dari data yang ada, nilai  $N$  sekaligus menentukan jumlah level untuk menentukan banyaknya dekomposisi dilakukan terhadap data. Kemudian tetapkan  $A_j$  untuk menentukan banyaknya elemen pada masing-masing level yang digunakan untuk membentuk model prediksi.
2. Cari nilai-nilai koefisien wavelet dan koefisien skala dengan melakukan dekomposisi untuk data sejumlah  $N$ .
3. Buat struktur persamaan linier dari  $X_{N+1}$  dengan koefisien wavelet dan koefisien skala terhadap koefisien-koefisien model wavelet time series alpha, sehingga didapatkan bentuk perkalian matriks.
4. Cari nilai-nilai taksiran koefisien-koefisien model wavelet time series alpha dari bentuk perkalian matriks dengan metode pseudo invers.

5. Cari nilai-nilai predksi berdasarkan model yang didapat dengan melibatkan koefisien wavelet dan koefisien model wavelet time series alpha, baik untuk data training maupun data validasi.
6. selesai

## 2.5 Analisis Deret Waktu (*Time Series*)

Analisis deret waktu adalah suatu rangkaian kegiatan dalam melakukan pengamatan terhadap variabel yang akan diamati secara berurutan berdasarkan urutan waktu kejadiannya dalam interval waktu tertentu secara konstan. Setiap pengamatan yang dilakukan dapat dinyatakan dalam bentuk variabel bebas yang didapatkan berdasarkan indeks waktu tertentu dengan sehingga penulisan dari data time series adalah (Wei, 2006).

Menurut Supranto (2000), analisis deret waktu adalah data yang dikumpulkan dari waktu ke waktu untuk menggambarkan perkembangan suatu kegiatan, yang memungkinkan untuk mengetahui perkembangan suatu atau beberapa kejadian serta hubungan/pengaruh terhadap kejadian lainnya. Dalam data deret waktu hal yang harus diperhatikan untuk menganalisisnya adalah dengan kestasioneran data, autokorelasi, dan fungsi autokorelasi parsial.

## 2.6 *Root Mean Square Error (RMSE)*

RMSE digunakan untuk mengukur kesalahan nilai dugaan model yang dinyatakan dalam rata-rata akar dari kesalahan kuadrat. RMSE digunakan untuk membandingkan beberapa model estimasi dari sebuah realisasi deret waktu yang sama. Nilai RMSE yang lebih rendah lebih banyak disukai karena model tersebut lebih cocok atau lebih mendekati dengan data yang ada. Model dengan RMSE yang lebih kecil cenderung memiliki variansi galaxt prakira yang lebih kecil. Rumus untuk menentukan RMSE sebagai berikut:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (2.52)$$

$n$  = jumlah data *training*

$y_i$  = data *training* aktual pada waktu  $i$

$\hat{y}_i$  = data *training* prediksi pada waktu  $i$

## 2.7 Curah Hujan

Hujan merupakan salah satu fenomena alam yang terdapat dalam siklus hidrologi dan sangat dipengaruhi iklim. Keberadaan hujan sangat penting dalam kehidupan, karena hujan dapat mencukupi kebutuhan air yang sangat dibutuhkan oleh semua makhluk hidup. Hujan merupakan gejala meteorologi dan juga unsur klimatologi. Hujan adalah hydrometeor yang jatuh berupa partikel-partikel air yang mempunyai diameter 0.5 mm atau lebih. Hydrometeor yang jatuh ke tanah disebut hujan sedangkan yang tidak sampai tanah disebut Virga (Tjasyono : 2006). Hujan yang sampai ke permukaan tanah dapat diukur dengan jalan mengukur tinggi air hujan tersebut dengan berdasarkan volume air hujan per satuan luas. Hasil dari pengukuran tersebut dinamakan dengan curah hujan. Curah hujan merupakan salah satu unsur cuaca yang datanya diperoleh dengan cara mengukurnya dengan menggunakan alat penakar hujan, sehingga dapat diketahui jumlahnya dalam satuan millimeter (mm).

Curah hujan 1 mm adalah jumlah air hujan yang jatuh di permukaan per satuan luas ( $m^2$ ) dengan catatan tidak ada yang menguap, meresap atau mengalir. Jadi, curah hujan sebesar 1 mm setara dengan 1 liter/ $m^2$  (Aldrian, E. dkk, 2011). Curah hujan dibatasi sebagai tinggi air hujan yang diterima di permukaan sebelum mengalami aliran permukaan, evaporasi dan peresapan ke dalam tanah. Berdasarkan ukuran butiran, hujan dapat dibedakan menjadi:

1. Hujan gerimis / drizzle, dengan diameter butirannya kurang dari 0,5 mm.
2. Hujan salju / snow, adalah kristal-kristal es yang temperatur udaranya berada di bawah titik beku ( $0^{\circ}C$ ).
3. Hujan batu es, curahan batu es yang turun didalam cuaca panas awan yang temperaturnya dibawah titik beku ( $0^{\circ}C$ ).
4. Hujan deras / rain, dengan curah hujan yang turun dari awan dengan nilai temperatur diatas titik beku berdiameter butiran  $\pm 7$  mm.

Jenis-jenis hujan berdasarkan besarnya curah hujan menurut BMKG dibagi menjadi tiga, yaitu :

1. Hujan Ringan, 0 – 20 mm per hari
2. Hujan sedang, 20 - 50 mm per hari.
3. Hujan lebat, 50-100 mm per hari.

4. Hujan sangat lebat, 100 - 150 mm per hari.
5. Hujan Extrem diatas 150 mm per hari. Intensitas curah hujan merupakan ukuran jumlah hujan per satuan waktu tertentu selama hujan berlangsung.

Data hujan mempunyai variasi yang sangat besar dibandingkan unsur iklim lainnya, baik variasi menurut tempat maupun waktu. Data hujan biasanya disimpan dalam satu hari dan berkelanjutan. Dengan mengetahui data curah hujan kita dapat melakukan pengamatan di suatu daerah untuk pengembangan dalam bidang pertanian dan perkebunan. Selain itu dapat juga digunakan untuk mengetahui potensi suatu daerah terhadap bencana alam yang disebabkan oleh faktor hujan.

### BAB 3 METODOLOGI PENELITIAN

#### 3.1 Sumber Data dan Variabel Penelitian

Data yang digunakan dalam penelitian ini yaitu data curah hujan harian. Data yang digunakan merupakan data Januari 2018 – November 2023. Data tersebut merupakan data yang didapatkan dari Badan Meteorologi, Klimatologi dan Geofisika (BMKG) Kota Surabaya. Data tersebut kemudian dibagi menjadi data *training* dan data *testing*. Struktur data yang digunakan pada penelitian ini dapat dilihat pada Tabel 3.1.

**Tabel 3.1** Struktur Data

Tahun	Bulan	Tanggal	Curah Hujan	Keterangan
2017	Januari	1	$y_1$	<i>Data Training</i>
		2	$y_2$	
	⋮	⋮	⋮	
	Desember	1	$y_{31}$	
		2	$y_{32}$	
		⋮	⋮	
		31	$y_{365}$	
	⋮	⋮	⋮	
2020	November	30	$y_{1065}$	
	Desember	1	$y_{1066}$	
	⋮	⋮	⋮	
2021	Desember	31	$y_{8125}$	
2023	Januari	1	$y_{8126}$	<i>Data testing</i>
⋮	⋮	⋮	⋮	
2023	November	31	$y_{2547}$	

### 3.2 Langkah Analisis

#### A. Pemodelan ARIMA

1. Merumuskan masalah dan menentukan tujuan
2. Mengumpulkan data curah hujan harian di Stasiun Juanda periode Januari 2018 sampai dengan November 2023.
3. Melakukan analisis statistik deskriptif pada data curah hujan harian Stasiun Meteorologi Juanda .
4. Membuat *time series* plot pada data *training* untuk melakukan identifikasi pola time series data curah hujan di Stasiun Meteorologi Juanda.
5. Melakukan Identifikasi dan pendugaan model ARIMA berdasarkan plot ACF dan PACF
6. Estimasi patameter Model ARIMA, menggunakan persamaan (2.18)
7. Uji signifikansi parameter, menggunakan persamaan (2.19)
8. Pemeriksaan diagnostik model, menggunakan uji *Ljung-Box* dan *Uji Kolmogrov-Smirnov*.
9. Melakukan peramalan dari data Training yang telah signifikan dan memenuhi asumsi. Peramalan dilakukan sebanyak periode data testing.
10. Menghitung nilai RMSE.
11. Memilih model terbaik berdasarkan nilai RMSE dan digunakan untuk melakukan prediksi.

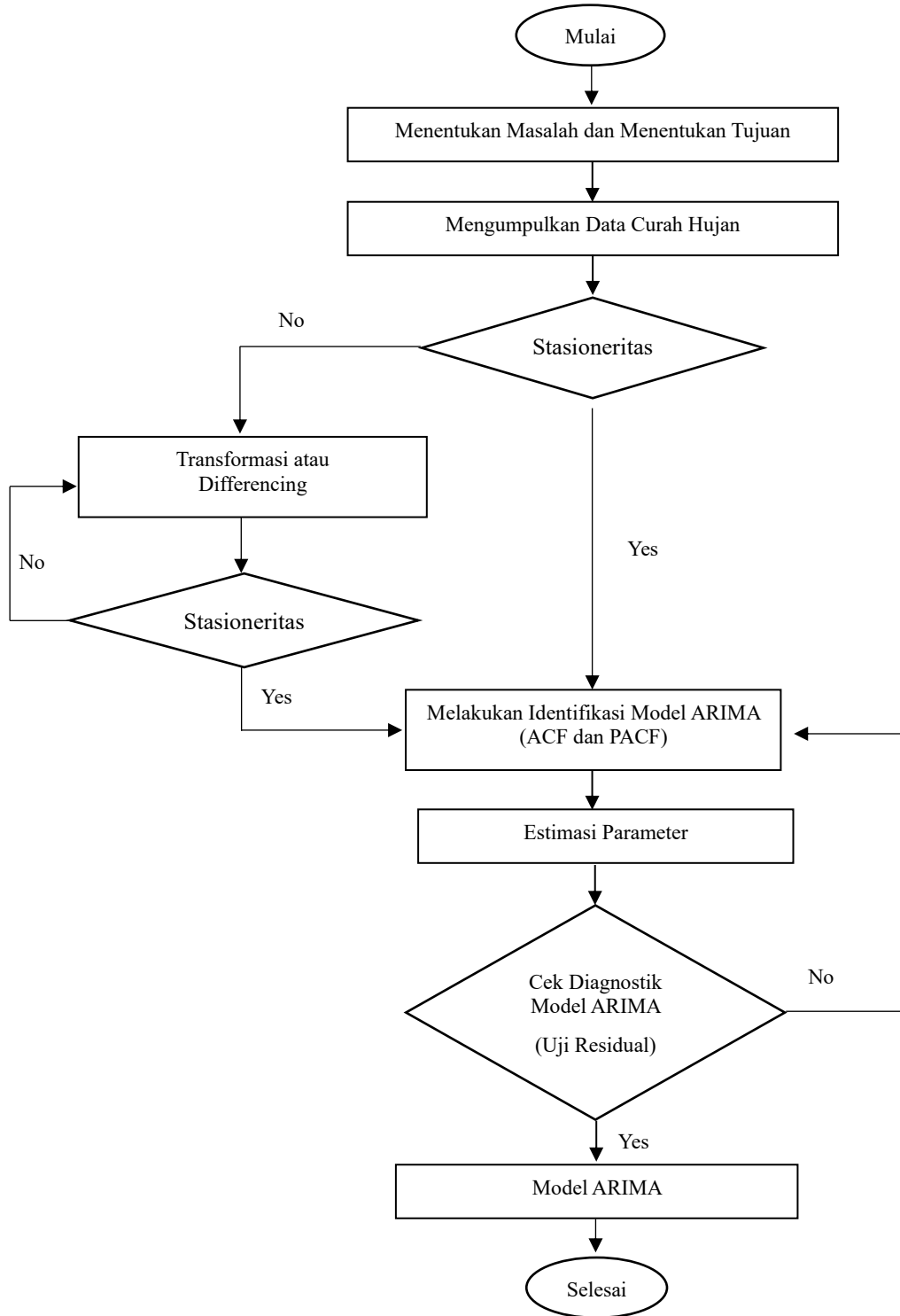
#### B. Melakukan pemodelan ANFIS

1. Menentukan variabel masukan berdasarkan orde  $p$  yang signifikan pada PACF berdasarkan data yang telah stasioner pada model ARIMA.
2. Menentukan banyaknya himpunan *fuzzy*, yaitu 5.
3. Menentukan Fungsi Keanggotaan (*membership function*). Yang berfungsi untuk memetakan *input* ke himpunan *fuzzy*.
4. Menghitung banyaknya aturan dan parameter yang terbentuk.
5. Menghitung nilai dari tiap lapis ANFIS.
  - a. Menentukan derajat keanggotaan *fuzzy/fuzzyfikasi*. Memetakan *input* ke dalam himpunan *fuzzy* sesuai dengan fungsi keanggotaan

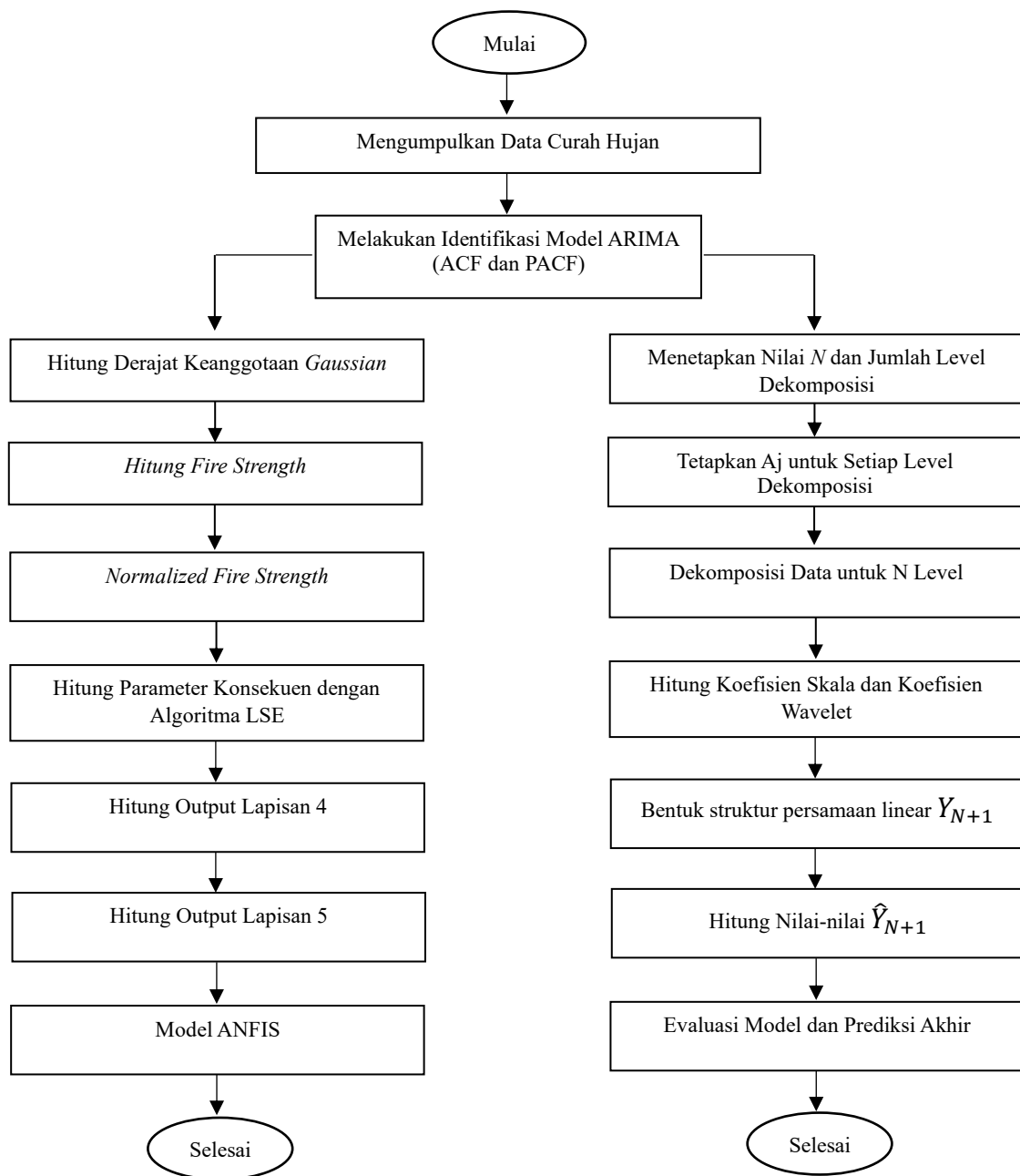


- (*membership function*) yang digunakan, menggunakan persamaan (2.22).
- b. Menentukan aturan keanggotaan fuzzy melalui operasi logika *fuzzy*, menggunakan persamaan (2.23).
  - c. Menentukan derajat pengaktifan ternormalisasi/*normalised firing strength*, menggunakan persamaan (2.24).
  - d. Melakukan proses *defuzzifikasi* menggunakan persamaan (2.24).
  - e. Menghitung semua output yang muncul menjadi satu output, menggunakan persamaan (2.25).
6. Menentukan banyaknya *epoch*/iterasi untuk mendapatkan parameter-parameter ANFIS yang dapat meminimumkan *error* dengan cara mencari titik *epoch* yang konvergen.
  7. mendapatkan hasil ramalan ANFIS dan menghitung nilai RMSE dari hasil ramalan.
- C. Melakukan Tranformasi Wavelet
1. Melakukan prosedur ARIMA dan menggunakan lag signifikan sebagai input.
  2. Tetapkan nilai  $N$  dari data yang ada, nilai  $N$  sekaligus jumlah level untuk menentukan banyaknya dekomposisi dilakukan terhadap data.
  3. Tetapkan  $A_j$  untuk menentukan banyak elemen pada masing-masing level yang digunakan untuk membentuk model prediksi.
  4. Cari nilai-nilai koefisien wavelet dan koefisien skala dengan melakukan dekomposisi untuk jumlah  $N$ .
  5. Buat struktur persamaan linier dari  $Y_{N+1}$  dengan koefisien wavelet dan koefisien skala terhadap koefisien-koefisien model wavelet time series  $\alpha$ , sehingga didapatkan bentuk perkalian matriks.
  6. Mencari nilai-nilai  $\hat{Y}_{N+1}$  berdasarkan model yang didapat dengan melibatkan koefisien-koefisien wavelet dan koefisien-koefisien model wavelet time series  $\alpha$ .
- D. Mendapatkan model terbaik. Langkah yang dilakukan untuk mendapatkan model terbaik yaitu dengan membandingkan nilai RMSE pada data Training

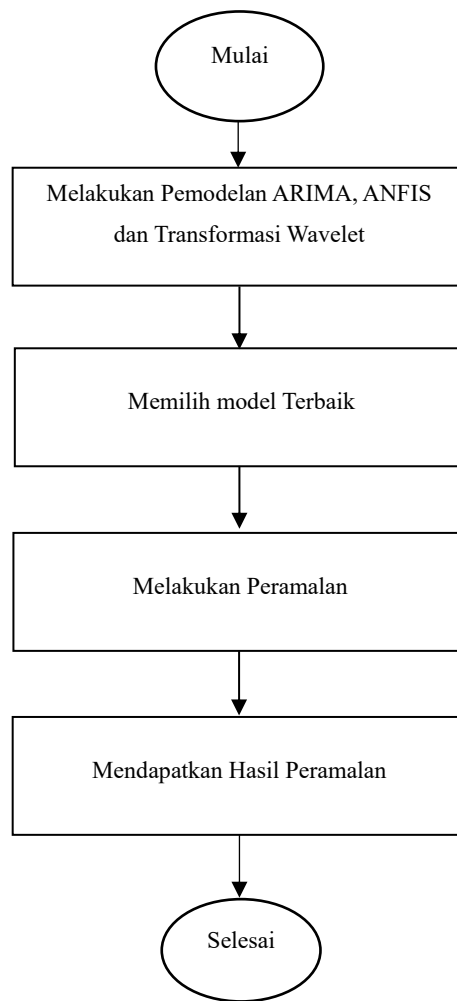
dan Testing dari model ARIMA, ANFIS dan Wavelet. Selanjutnya dilakukan peramalan menggunakan model terbaik.



**Gambar 3.1** Diagram Alir Model Arima



**Gambar 3.2** Diagram Alir Model ANFIS dan Wavelet



**Gambar 3.3** Diagram Alir Peramalan menggunakan Model Terbaik

## BAB 4

### HASIL DAN PEMBAHASAN

Pada bab ini berisi mengenai hasil penelitian Peramalan Curah Hujan Menggunakan Metode ARIMA, ANFIS, Transformasi Wavelet, Bab ini diuraikan hasil mengenai data yang digunakan, peramalan curah hujan dengan menggunakan metode ARIMA, peramalan curah hujan dengan menggunakan metode ANFIS, peramalan curah hujan dengan metode Transformasi Wavelet, serta akurasi hasil peramalan dari metode yang digunakan.

#### 4.1 Deskripsi Data Curah Hujan

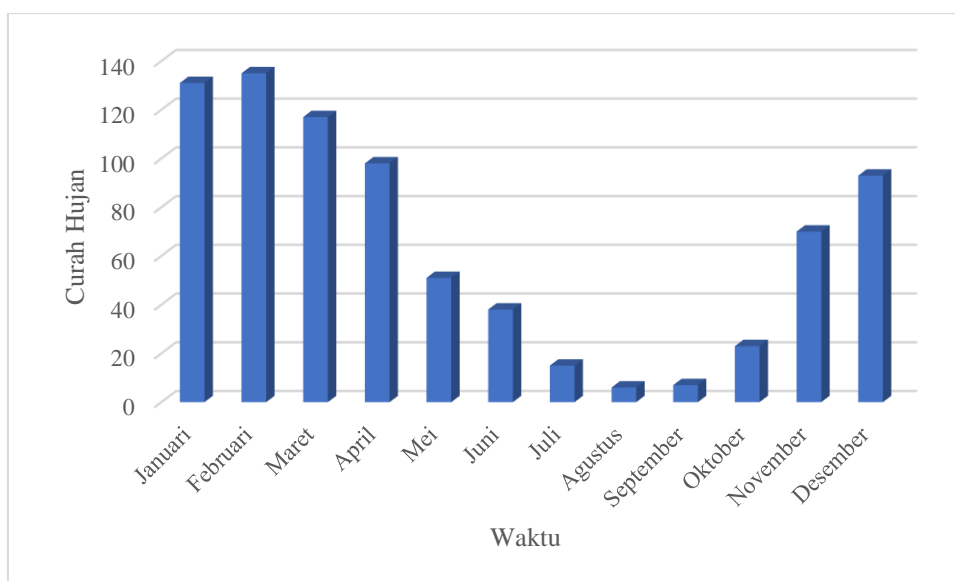
Data yang digunakan pada penelitian ini merupakan data curah hujan Stasiun Meteorologi Kelas I Juanda. Data tersebut diperoleh melalui BMKG Stasiun Meteorologi Kelas I Juanda. Data yang digunakan merupakan data curah hujan harian dari Januari 2018 sampai November 2023. Karakteristik pola curah hujan stasiun tersebut dapat dilihat pada Tabel 4.1.

**Tabel 4.1** Jumlah hari hujan Stasiun Juanda Januari 2018 sampai November 2023

Tahun	Banyak hari hujan	Rata – rata banyak hari hujan per bulan	Kisaran tinggi curah hujan (mm)
2018	104	8,67	0 – 86 mm
2019	98	8,17	0 – 157 mm
2020	155	12,92	0 – 128 mm
2021	150	12,5	0 – 90 mm
2022	181	15,08	0 – 152 mm
2023	96	8,73	0 – 107 mm

Pada Tabel 4.1, terlihat bahwa tahun paling sedikit mengalami hujan di Stasiun Juanda pada saat tahun 2019 yaitu hanya 98 hari atau sebesar 27%. Hal tersebut menunjukkan bahwa hujan yang terjadi di Stasiun Juanda jarang terjadi. Hari hujan terbanyak terjadi pada tahun 2022 yaitu sebesar 181 hari atau 50%. Jumlah hari terbanyak mengalami hujan pada Stasiun Juanda terjadi pada tahun 2022. Curah hujan harian tertinggi Stasiun Juanda terjadi pada tahun 2023 yaitu sebesar 171 mm.

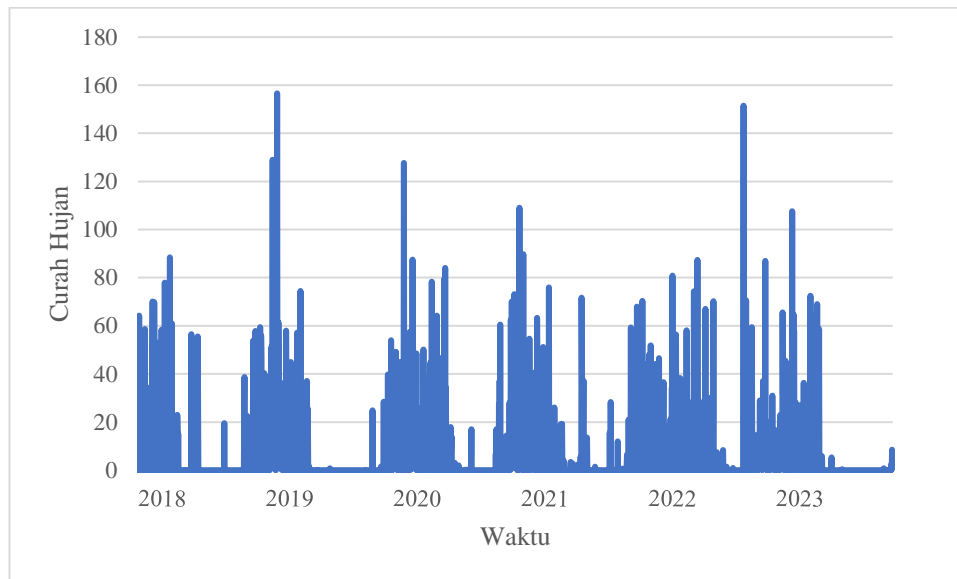
Informasi mengenai curah hujan harian periode bulanan Stasiun Juanda dapat dilihat pada Gambar 4.1.



**Gambar 4.1** Curah Hujan Harian per Bulan Stasiun Juanda

Dilihat dari grafik pada Gambar 4.1, curah hujan harian per bulan Stasiun Juanda, menunjukkan bahwa pola curah hujan harian periode bulanan menunjukkan pola hujan membentuk huruf U. Sehingga pada bulan Mei hingga Oktober terlihat intensitas curah hujan lebih sedikit dibanding bulan November – April. Pola hujan seperti ini menunjukkan bahwa hanya terdapat satu puncak musim hujan atau disebut unimodal. Jumlah hari hujan periode bulanan dari tahun 2018 – 2023 menunjukkan bahwa hari hujan terbanyak terjadi pada bulan November hingga April, sedangkan hari hujan paling sedikit terjadi pada bulan Mei – Oktober.

Data curah hujan tersebut memiliki karakteristik khusus yaitu data curah hujan merupakan data *intermittent*. Data *intermittent* merupakan suatu data yang memiliki nilai nol dalam suatu periode tertentu. hal tersebut dapat diartikan bahwa terdapat hari tidak turun hujan. Grafik data yang digunakan dapat dilihat pada Gambar 4.2.



**Gambar 4.2** merupakan grafik curah hujan dalam satuan milimeter (mm).

Pada data curah hujan Stasiun Juanda terdapat data yang bernilai nol sebanyak 1377 data atau 63,57%. Artinya nilai curah hujan harian yang paling sering muncul di Stasiun Juanda adalah 0 mm per hari. Statistik deskriptif data curah hujan stasiun Juanda dapat dilihat pada Tabel 4.2.

**Tabel 4.2** Statistik Deskriptif Data Curah Hujan Stasiun Juanda Periode 2018 - 2023

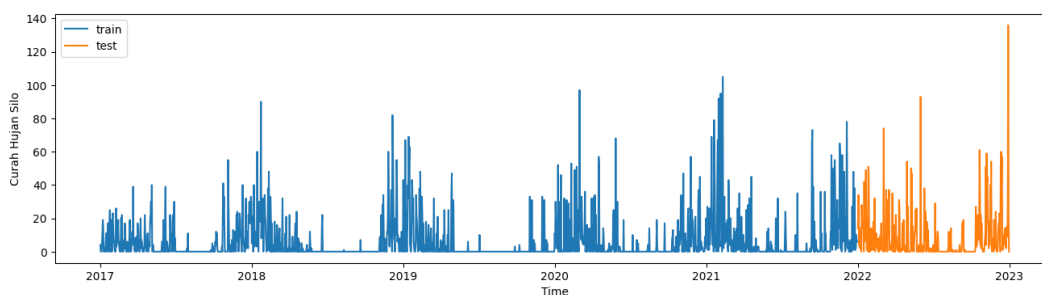
Keterangan	Hasil
Banyak data	2164
Rata – rata	6,33
Standard deviasi	15,41
Min	0 mm
Max	156,6 mm
Median	0
Skewness	3,78
Kurtosis	18,92

Berdasarkan Tabel 4.2, Stasiun Juanda memiliki rata – rata curah hujan harian sebesar 6,33. Nilai minimum yang terdapat pada Stasiun Juanda adalah nol dan nilai maksimum yang terdapat pada stasiun Juanda adalah 156,6. Nilai standard deviasi yang diperoleh oleh Stasiun Juanda adalah sebesar 15,41. Nilai standard deviasi

tersebut lebih besar dibandingkan dengan nilai rata – ratanya artinya Stasiun Juanda memiliki tingkat variasi data yang tinggi. Ketika variasi yang terdapat pada data tinggi, maka semakin beragam nilai observasinya. Data curah hujan harian juga memiliki nilai *skewness* sebesar 3,78. Nilai tersebut lebih dari nol dan positif, sehingga menunjukkan bahwa kemiringan distribusi data tersebut ke arah kanan dan tidak mengikuti distribusi normal. Dapat dikatakan bahwa nilai akumulasi curah hujan lebih kecil dibanding dengan nilai rata – rata akumulasi curah hujan stasiun tersebut. Untuk nilai kurtosis data tersebut sebesar 18,92. Nilai tersebut lebih besar dari nol dan positif, sehingga menunjukkan distribusi data lebih runcing. Median bernilai nol memberikan arti bahwa 50% data memiliki nilai lebih kecil atau sama dengan median dan 50% data memiliki nilai yang lebih tinggi atau sama dengan median. Nilai rata – rata lebih besar dibanding dengan nilai median memberikan arti bahwa data tersebut miring ke arah positif. Hal tersebut terjadi karena data lebih condong ke nilai yang lebih rendah, sehingga menarik rata – rata ke sisi nilai yang lebih rendah.

## 4.2 Model ARIMA

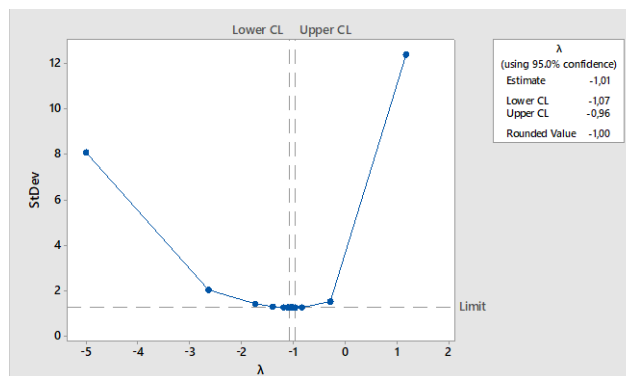
Data yang digunakan dibagi menjadi data *training* dan data *testing*. Data *training* terdiri dari 1825 data dan data *testing* terdiri dari 365 data. Data yang digunakan sebagai data train adalah data curah hujan harian tanggal 1 Januari 2018 – 31 Desember 2022 dan data yang digunakan sebagai data test adalah data curah hujan harian tanggal 1 Januari 2023 – 30 November 2023. Data *training* digunakan sebagai pembuatan model yang digunakan untuk meramalkan curah hujan. Grafik pembagian data *training* dan data *testing* dapat dilihat pada Gambar 4.3



**Gambar 4.3** Grafik Data *Training* dan *testing* Curah Hujan Harian Stasiun Juanda

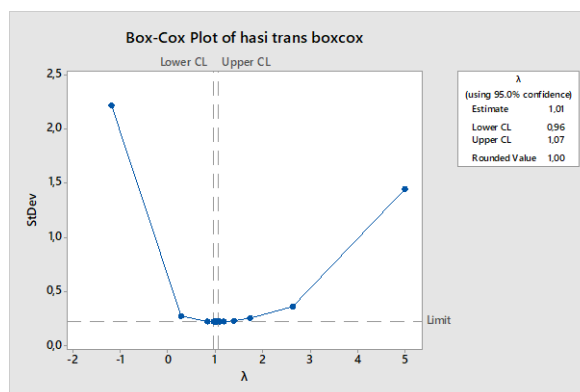


Sebelum mendapatkan model ARIMA, data terlebih dahulu di uji apakah sudah memenuhi kondisi stasioner. Kondisi stasioner yang harus terpenuhi adalah stasioner dalam varians dan stasioner dalam *mean*. Pengujian kondisi stasioner dalam varians dilakukan dengan cara transformasi *Box Cox* dan pengujian kondisi stasioner dalam *mean* dilakukan dengan melihat sebaran data apakah telah berfluktuasi disekitar nilai tertentu *mean* atau rata-rata, pengujian kondisi stasioner dalam *mean* juga dapat dilakukan dengan melihat plot ACF dan PACF. Berikut ini adalah hasil transformasi *Box Cox* dari data curah hujan.



**Gambar 4.4** Uji Stasioner dalam Varians

Dari Gambar 4.4 diatas dapat diketahui bahwa data curah hujan belum stasioner dalam varians mengingat nilai batas atas sebesar  $-0,96$  dan batas bawah sebesar  $-1,07$  tidak melewati angka 1 sehingga perlu dilakukan transformasi. Berdasarkan transformasi *Box Cox* diperoleh nilai lambda sebesar  $-1,00$  maka perlu dilakukan transformasi  $\ln$ . Berikut adalah pengujian stasioner dalam varians setelah dilakukan transformasi data.



**Gambar 4.5** Uji Stasioner dalam Varians Transformasi

Gambar 4.4 menunjukkan bahwa data curah hujan telah memenuhi kondisi stasioner dalam varians, hal ini dibuktikan dengan nilai batas atas sebesar 1,07 dan batas bawah sebesar 0,96 yang berarti intervalnya memuat angka 1 dan nilai lambda sebesar 1. Setelah data memenuhi kondisi stasioner dalam varians maka asumsi yang harus dipenuhi berikutnya adalah stasioner dalam *mean*.

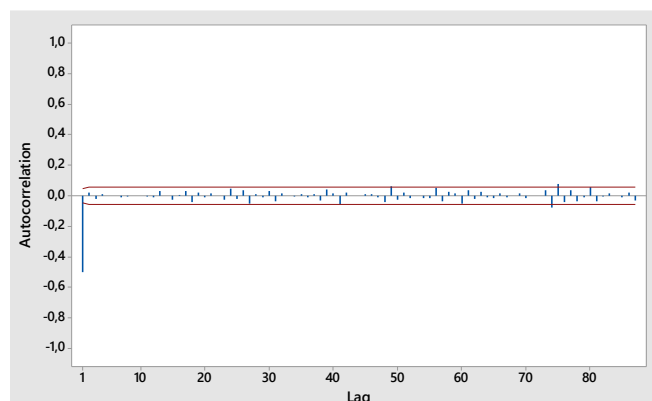
Uji stasioner dalam mean data dapat dilakukan dengan menggunakan metode *Augmented Dickey – Fuller*. Metode tersebut dapat digunakan untuk mengetahui apakah data stasioner dalam mean atau tidak. Hipotesis yang digunakan pada uji stasioner data deret waktu adalah

$H_0$ : data yang diuji merupakan data tidak stasioner

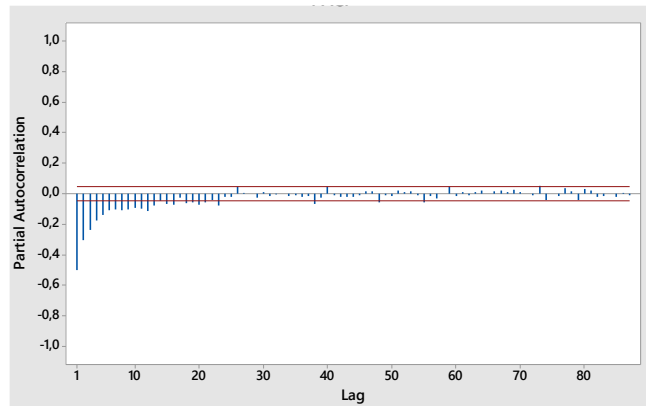
$H_1$ : data yang diuji merupakan data stasioner

Setelah dilakukan *differencing* sebanyak satu kali uji stasioneritas pada data curah hujan harian Stasiun Juanda, didapat hasil *p – value* sebesar 0,01, sehingga nilai *p – value* tersebut lebih kecil dibanding dengan nilai  $\alpha$  atau 0,05. Oleh karena itu, dapat dilakukan penolakan terhadap hipotesis nol ( $H_0$ ). Dapat disimpulkan bahwa data curah hujan harian Stasiun Juanda sebanyak satu kali merupakan data stasioner.

Setelah asumsi stasioneritas dipenuhi, maka dapat dilanjutkan dengan mengidentifikasi model menggunakan plot ACF dan PACF. Hasil plot ACF dapat dilihat pada Gambar 4.6 dan hasil plot PACF dapat dilihat pada Gambar 4.7.



**Gambar 4.6** Plot ACF data Curah Hujan



**Gambar 4.7** Plot PACF data Curah Hujan

Berdasarkan plot ACF dan PACF dapat diidentifikasi model yang mungkin berdasarkan lag yang signifikan pada plot tersebut yaitu ARIMA ([7], 1,[1]) dan ARIMA ([7,49], 1, [1]). Uji signifikansi parameter pada model ARIMA tersebut dapat dilihat pada Tabel 4.3.

**Tabel 4.3** Uji Signifikansi Parameter

Model	Parameter	Estimasi	Std.Error	P-Value	Lag	Keputusan
ARIMA ([7], 1,[1])	MA1,1	0.9361	0.0084	<.0001	1	Signifikan
	AR1,1	-0.0474	0.02406	0.0491	7	Signifikan
ARIMA ([7,49], 1, [1])	MA1,1	0.9364	0.0084	<.0001	1	Signifikan
	AR1,1	-0.0472	0.0240	0.0496	7	Signifikan
	AR1,2	0.0491	0.0237	0.0386	49	Signifikan

Model ARIMA ([7], 1,[1]) dan model ARIMA ([7,49], 1, [1]) pada Tabel 4.3 diatas diketahui bahwa untuk seluruh parameter nilai *Pvalue* lebih kecil daripada nilai  $\alpha$  (0,05) atau dapat dikatakan seluruh parameter model ARIMA tersebut telah signifikan sehingga signifikansi parameternya terpenuhi. Selanjutnya dilakukan pengujian diagnostik pada *residuals* yang meliputi *white noise* dan distribusi normal. Berikut merupakan hasil pengujian diagnostik untuk model ARIMA Model ARIMA ([7], 1,[1]) dan model ARIMA ([7,49], 1, [1]).

**Tabel 4.4** Uji Asumsi Residual

Model ARIMA	White Noise		Distribusi Normal
	Lag	Pvalue	
([7], 1,[1])	6	0.2333	>0,1543
	12	0.3069	
	18	0.4238	
	24	0.211	
	30	0.1015	
	36	0.1732	
	42	0.0798	
	48	0.1392	
([7,49], 1, [1])	6	0.1213	>0,1854
	12	0.1972	
	18	0.326	
	24	0.1371	
	30	0.0608	
	36	0.1134	
	42	0.0585	
	48	0.0905	

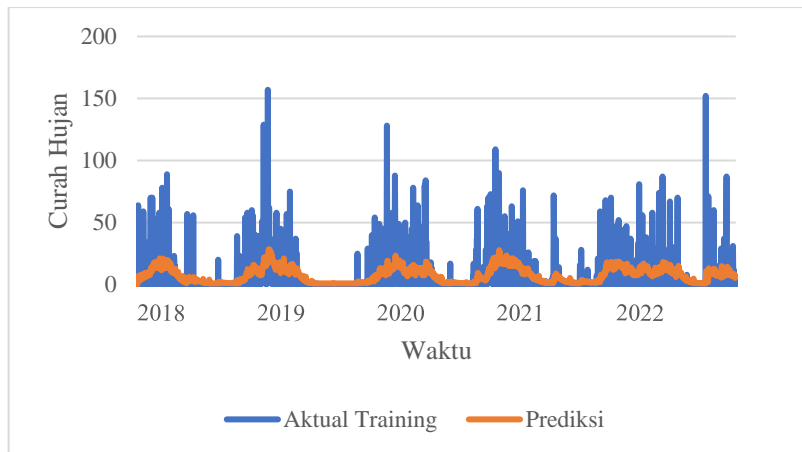
Berdasarkan Tabel 4.4, diperoleh bahwa nilai *Pvalue* untuk diagnostik distribusi normal lebih besar daripada nilai  $\alpha$  (0,05) maka *residuals* berdistribusi normal atau *residuals* telah memenuhi asumsi distribusi normal. Nilai *Pvalue* untuk diagnostik *white noise*, dimana *Pvalue* bernilai lebih besar dari pada nilai  $\alpha$  (0,05) untuk seluruh lag maka dapat dikatakan *residuals* telah memenuhi asumsi *white noise*. Berdasarkan uraian diatas *residuals* telah memenuhi asumsi *white noise* dan distribusi normal maka dapat disimpulkan bahwa asumsi *residuals* telah terpenuhi. Berikut ini adalah hasil perhitungan nilai RMSE dari kedua kemungkinan model yang ada.

**Tabel 4.5** Nilai RMSE Model ARIMA

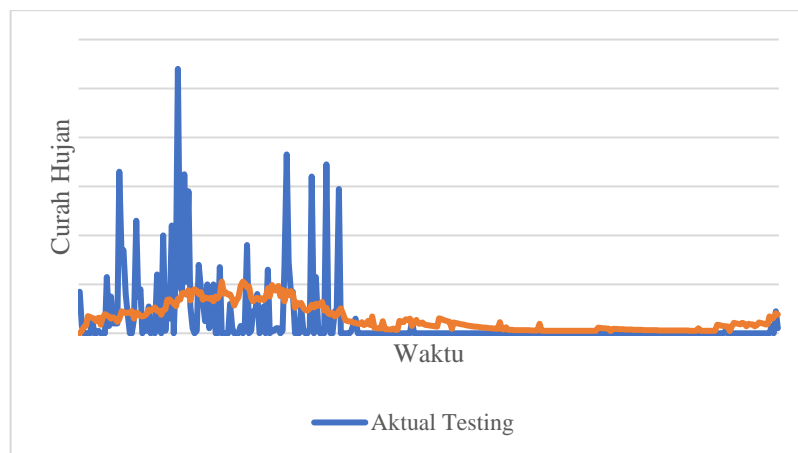
Model ARIMA	Training	Testing
([7], 1,[1])	15,1455	13.2625
([7,49], 1, [1])	15,1278	13.2504

Berdasarkan Tabel 4.5, nilai RMSE data Training dan Testing secara berturut-turut model ARIMA ([7,49], 1, [1]) dengan nilai RMSE 15,1278 dan 13,2504 lebih kecil dari Model ARIMA ([7], 1,[1]) yaitu 15,1455 dan 13,2625. Sehingga model yang terpilih yaitu ARIMA ([7,49], 1, [1]). Berikut ini adalah model ARIMA ([7,49], 1, [1]) secara matematis.

$$Y_t = Y_{t-1} + 0,0472(Y_{t-7} - Y_{t-8}) + 0,0491(Y_{t-49} - Y_{t-50}) + \varepsilon_t - 0,9364\varepsilon_{t-1}$$



**Gambar 4.8** Plot Data Aktual Traing dan Hasil Prediksi ARIMA ([7,49],1,[1])



**Gambar 4.9** Plot Data Aktual Testing dan Hasil Prediksi ARIMA ([7,49],1,[1])

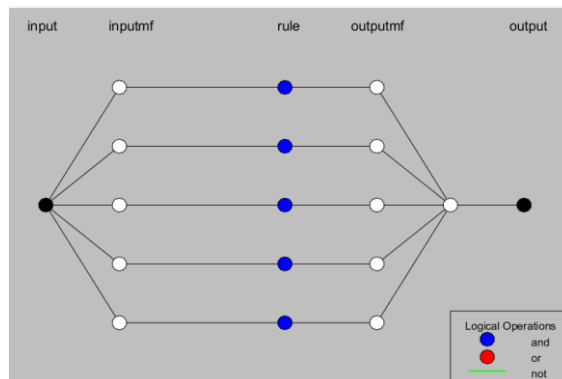
### 4.3 Model ANFIS

*Input* yang digunakan dalam tahapan ANFIS ditentukan dengan pendekatan ARIMA yang telah menghasilkan dengan parameter yang signifikan. Namun dalam penelitian, *input* dalam pendekatan ANFIS dibatasi hanya melihat plot PACF yang signifikan dalam model ARIMA. 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*). Dalam pemodelan ARIMA pada data curah hujan harian di Stasiun Juanda, telah menghasilkan model yang telah signifikan seperti ditunjukkan pada Tabel 4.3. oleh karena itu, telah didapatkan *input* yang dapat digunakan dalam pendekatan ANFIS dengan melihat kriteria nilai RMSE terkecil pada model ARIMA yang dihasilkan sebagai berikut:

**Tabel 4.6** Input ANFIS Data Curah Hujan Harian Stasiun Juanda

Model ARIMA	<i>Input</i> dalam anfis
([7,49],1,[1])	$Y_{t-7}$

Setelah mendapat *input* yang akan digunakan dalam pendekatan ANFIS, tahapan selanjutnya adalah menentukan banyaknya fungsi keanggotaan (*membership function*) dan jenis fungsi keanggotaan (*membership function*). Adapun jenis *membership function* yang akan digunakan adalah *Generalized Bell*. Sedangkan banyaknya *membership function* yang digunakan yaitu 5. Berikut ini merupakan penggambaran tahapan dalam ANFIS, dimana *input* yang digunakan adalah  $Y_{t-7}$  dengan jumlah *membership function* yang digunakan sebanyak 5, serta jenis *membership function* adalah *Generalized Bell*.



**Gambar 4.10** Arsitektur ANFIS input  $Y_{t-7}$ .

Berdasarkan Gambar 4., Arsitekur ANFIS terdiri dari 5 lapis. Lapis 1 merupakan *input* yang digunakan sebanyak 1 dengan  $Y_{t-7}$  adalah curah hujan harian tujuh hari sebelumnya yang telah ditransformasi. *Input* memiliki 5 himpunan *fuzzy*/fungsi keanggotaan (*membership function*).

Berdasarkan gambar diatas menggunakan 5 *membership function* dan *input* yang digunakan sebanyak 1. Total parameter yang digunakan sebanyak 30 dimana 15 diantaranya merupakan parameter nonlinier atau parameter premis dalam ANFIS yang berasal dari banyaknya *membership function* dikali banyaknya *input* dan banyaknya parameter *membership function Generalized Bell* ( $5 \times 1 \times 3 = 15$ ) yang nantinya digunakan pada lapis 1 dan sisanya merupakan parameter linier atau parameter konsekuen yang digunakan pada lapis 4. Parameter linier tersebut diperoleh dari perkalian antara banyaknya aturan dengan banyaknya parameter konsekuen di lapis 4 ( $C_1, C_2, C_0$ ) dimana aturan yang digunakan sebanyak 5 yaitu:

1. Aturan 1:  
IF Curah Hujan is Ringan  
THEN Output adalah prediksi yang sesuai untuk kategori "Ringan".
2. Aturan 2:  
IF Curah Hujan is Sedang  
THEN Output adalah prediksi yang sesuai untuk kategori "Sedang".
3. Aturan 3:  
IF Curah Hujan is Lebat  
THEN Output adalah prediksi yang sesuai untuk kategori "Lebat".
4. Aturan 4:  
IF Curah Hujan is Sangat Lebat  
THEN Output adalah prediksi yang sesuai untuk kategori "Sangat Lebat".
5. Aturan 5:  
IF Curah Hujan is Ekstrem  
THEN Output adalah prediksi yang sesuai untuk kategori "Ekstrem".

Berikut ini merupakan perhitungan yang dilakukan pada masing-masing lapis dengan menggunakan data *input*  $Y_{t-7}$  seperti terlihat pada tabel 4.7.

**Tabel 4.7** Data yang Digunakan untuk Pelatihan

Data Ke-	Input
	$Y_{t-7}$
1	0
2	0
⋮	⋮
2183	12
2184	3
2185	0

Tahapan Pertama yaitu menentukan derajat keanggotaan himpunan *fuzzy* dengan tujuan untuk mengelompokkan data *input* kedalam himpunan *fuzzy* sesuai klasifikasinya. Tabel 4.8 Merupakan nilai awal parameter dengan jenis *membership function Gauss* dan jumlah *membership function* sebanyak 5. Parameter  $a_i$ ,  $b_i$  dan  $c_i$  merupakan parameter yang terdapat dalam *membership function Generalized Bell*.

**Tabel 4.8** Nilai Awal Parameter

Input	Parameter		
	$a_i$	$b_i$	$c_i$
Input 1 MF1	19,62	2	0
Input 1 MF2	19,62	2	39,25
Input 1 MF3	19,62	2	78,5
Input 1 MF4	19,62	2	117,8
Input 1 MF5	19,62	2	157

Berdasarkan tabel diatas dapat diketahui nilai parameter premis pada lapis pertama yang dapat dituli secara matematis sebagai berikut:

$$\mu_{A1}(Y_{t-7}) = \frac{1}{1 + \left| \frac{Y_{t-7} - 0}{19,62} \right|^{2*2}},$$

$$\mu_{A2}(Y_{t-7}) = \frac{1}{1 + \left| \frac{Y_{t-7} - 39,25}{19,62} \right|^{2*2}},$$

$$\mu_{A3}(Y_{t-7}) = \frac{1}{1 + \left| \frac{Y_{t-7} - 78,5}{19,62} \right|^{2*2}},$$



$$\mu_{A4}(Y_{t-7}) = \frac{1}{1 + \left| \frac{Y_{t-7} - 117,8}{19,62} \right|^{2*2}}$$

$$\mu_{A5}(Y_{t-7}) = \frac{1}{1 + \left| \frac{Y_{t-7} - 157}{19,62} \right|^{2*2}}$$

Dalam lapis 1, data *input* diubah menjadi derajat keanggotaan dimana derajat keanggotaan yang terbentuk merupakan *output* dari lapis 1 dan menjadi *input* untuk lapis 2. Pada lapis 2 terjadi proses logika *fuzzy* yang biasanya menggunakan operator AND karena semua aturan saling bergantung dan mempengaruhi. Sehingga dilakukan perkalian antar nilai derajat keanggotaan dan menghasilkan  $w_i$  ( $i=5$ ).

**Tabel 4.9** Output Lapis 2

Data Ke-	$w_1$	$w_2$	$w_3$	$w_4$	$w_5$
1	1	0,1111	0,0303	0,0137	0,0077
2	1	0,1111	0,0303	0,0137	0,0077
⋮	⋮	⋮	⋮	⋮	⋮
1823	0,1669	0,7388	0,0786	0,0249	0,0119
1824	0,8424	0,1483	0,0353	0,0152	0,0084
1825	0,5720	0,2058	0,4171	0,169	0,0091

Selanjutnya, lapis 3 merupakan penentuan kuat penyulutan ternormalisasi (*normalized firing strength*). Output dalam lapis ini adalah  $\bar{w}_i$  untuk  $i = 5$  dimana seperti yang dituliskan dalam persamaan 2.23. Berikut ini hasil dari lapis 3.

**Tabel 4.10** Output Lapis 3

Data Ke-	$\bar{w}_1$	$\bar{w}_2$	$\bar{w}_3$	$\bar{w}_4$	$\bar{w}_5$
1	0,8600	0,0955	0,0260	0,0118	0,0067
2	0,8600	0,0955	0,0260	0,0118	0,0067
⋮	⋮	⋮	⋮	⋮	⋮
1823	0,1634	0,7235	0,0769	0,0244	0,0117
1824	0,8026	0,1413	0,0337	0,0144	0,0079
1825	0,6765	0,2434	0,0493	0,0199	0,0107

Proses selanjutnya pada lapis 4 adalah tahapan *defuzzyfikasi*, yaitu melakukan perhitungan transformasi hasil *fuzzy* ke dalam bentuk *crisp*. Pada lapis ini dilakukan

proses pembelajaran dengan menggunakan metode LSE (*Least Square Estimator*) yang menghasilkan parameter linier sebagai hasil pembelajaran.

**Tabel 4.11** Parameter Linier Optimum

Aturan	Parameter		
	$C_{i1}$	$C_{i2}$	$C_{io}$
Aturan 1	19,68	0,5719	-0,0733
Aturan 2	19,89	0,8296	39,1
Aturan 3	19,75	1,509	78,42
Aturan 4	19,65	1,846	117,7
Aturan 5	19,63	1,958	157

Nilai parameter linier diatas digunakan untuk perhitungan *output* untuk lapis 4. Hasil lapis 4 diperoleh dengan mengoperasikan nilai parameter linier dengan output di lapis 3 berdasarkan persamaan 2.24 yang disajikan sebagai berikut.

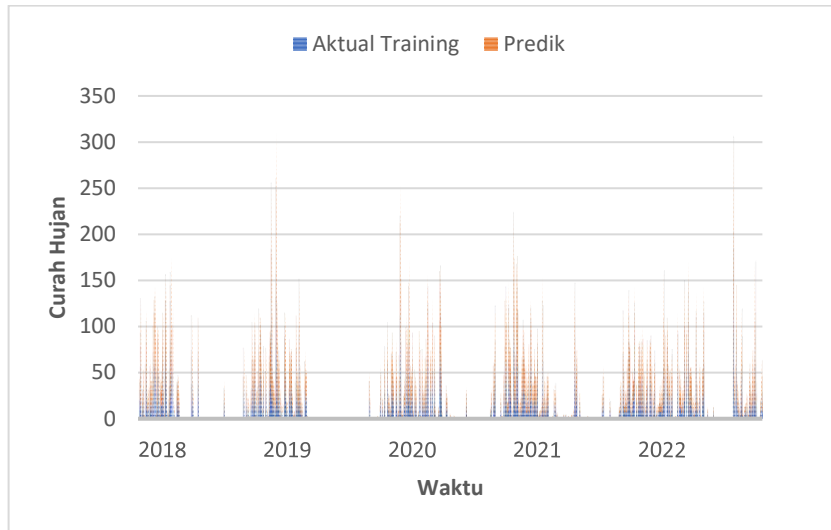
**Tabel 4.12** *Output* Lapisan 4

Data Ke-	$\bar{w}_1 Y_{t-7}$	$\bar{w}_2 Y_{t-7}$	$\bar{w}_3 Y_{t-7}$	$\bar{w}_4 Y_{t-7}$	$\bar{w}_5 Y_{t-7}$
1	0	0	0	0	0
2	0	0	0	0	0
⋮	⋮	⋮	⋮	⋮	⋮
1823	5,0658	22,4280	2,3863	0,7562	0,3636
1824	4,8189	0,8477	0,2019	0,0867	0,0479
1825	8,1181	2,9213	0,5919	0,2399	0,1287

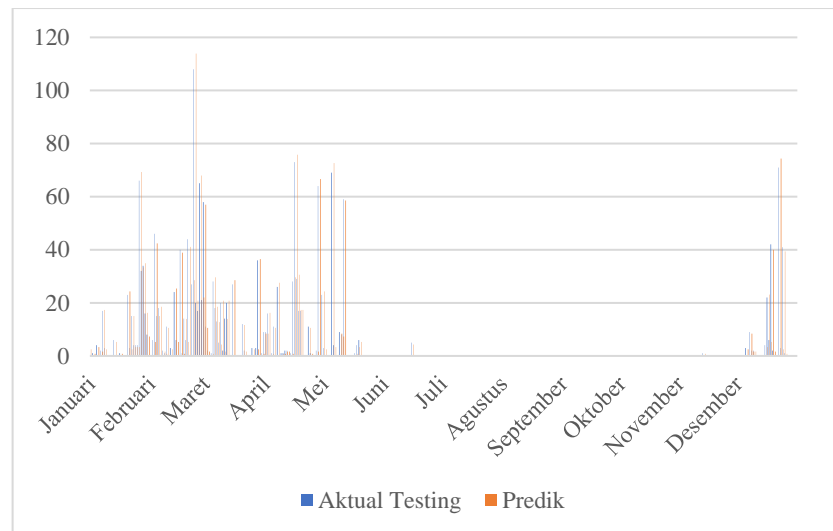
Tahapan akhir merupakan lapisan 5 yaitu menghitung semua *output* yang muncul di lapis 4 menjadi satu *output* saja. Hasil dari lapis 5 yaitu  $\hat{Y}$  dimana nilai tersebut merupakan data curah hujan harian di Stasiun Juanda yang telah di transformasi. Oleh karena itu, hasil *output* lapis 5 perlu dikembalikan ke kondisi semula. Berikut adalah model taksiran yang digunakan:

$$\hat{Y} = \sum_{i=1}^5 \bar{w}_{it} y_i = \bar{w}_{1t} y_1 + \bar{w}_{2t} y_1 + \bar{w}_{3t} y_1 + \bar{w}_{4t} y_1 + \bar{w}_{5t} y_1$$

$$= \bar{w}_{1t} (C_{11} Y_{t-1} + C_{10}) + \bar{w}_{2t} (C_{11} Y_{t-1} + C_{10}) + \bar{w}_{3t} (C_{11} Y_{t-1} + C_{10}) + \bar{w}_{4t} (C_{11} Y_{t-1} + C_{10}) + \bar{w}_{5t} (C_{11} Y_{t-1} + C_{10})$$



**Gambar 4.11** Grafik Hasil Pelatihan ANFIS



**Gambar 4.12** Grafik Hasil Pengujian ANFIS

Berdasarkan Gambar 4.11 dan Gambar 4.12 Grafik Hasil pengujian data Training dan Testing dapat dilihat gambar tersebut menunjukkan perbandingan

antara data aktual dan data prediksi model ANFIS, sumbu horizontal menunjukkan urutan waktu sedangkan sumbu vertical menunjukkan jumlah curah hujan. Titik-titik biru mewakili data curah hujan aktual dan titik-titik oranye mewakili prediksi yang dihasilkan oleh model. Kedua set data tampak memiliki pola yang mirip baik Data Training maupun data Testing, meskipun ada beberapa perbedaan yang dapat dilihat, terutama pada titik-titik data dengan nilai yang lebih tinggi.

RMSE untuk data pelatihan adalah sebesar 0,5512 yang menunjukkan bahwa model memiliki kesalahan rata-rata yang relative kecil saat memprediksi data pelatihan. Nilai RMSE ini menunjukkan bahwa model cukup baik dalam mempelajari pola dari data pelatihan. RMSE untuk data pengujian adalah sebesar 0,6825. Ini menunjukkan bahwa model memiliki kinerja yang baik pada data pengujian, dengan kesalahan rata-rata yang sedikit lebih rendah dibandingkan dengan data pelatihan. Sehingga model yang digunakan memiliki kinerja yang cukup baik dalam memprediksi curah hujan, baik pada set pelatihan maupun pengujian.

#### 4.4 Model Transformasi Wavelet

Data yang digunakan adalah data curah hujan harian di stasiun Juanda pada tahun 2018-2023. Dimana data dibagi menjadi data training dan data testing. Data training sebanyak 1825 untuk proses transformasi dan estimasi dan data testing sebanyak 365 digunakan untuk proses validasi. Dari nilai prediksi didapatkan nilai RMSE antara nilai prediksi terhadap data sebenarnya. Menghitung jumlah level  $j$  dimana jumlah data yang digunakan yaitu  $2190 = 2^j = 2^{11}$ ,  $j = 11$ . Level transformasi  $= j - 1 = 11 - 1 = 10$ .

Koefisien Wavelet dapat diperoleh dengan menggunakan rumus pada persamaan:

$$W_{j,t} = \sum_{l=0}^3 h_l Y_{(2t-k) \bmod N}$$

$$\begin{aligned} W_{1,1} &= \sum_{l=0}^3 h_l Y_{(2 \cdot 1 - k) \bmod 2190} \\ &= h_0 Y_{i(2 \cdot t - l) \bmod 2190} \\ &= h_0 Y_2 + h_1 Y_1 + h_2 Y_{2190} + h_3 Y_{2189} \end{aligned}$$

$$= \frac{1-\sqrt{3}}{4\sqrt{2}} \times 24 + \frac{-3+\sqrt{3}}{4\sqrt{2}} \times 0 + \frac{3+\sqrt{3}}{4\sqrt{2}} \times 0 + \frac{-1-\sqrt{3}}{4\sqrt{2}} \times 25 = -15,1785$$

$$\begin{aligned} W_{1,1} &= \sum_{l=0}^3 h_l Y_{(2.1-k) \bmod 2190} \\ &= h_0 Y_{i(2.t-l) \bmod 2190} \\ &= h_0 Y_4 + h_1 Y_3 + h_2 Y_2 + h_0 Y_1 \\ &= \frac{1-\sqrt{3}}{4\sqrt{2}} \times 64 + \frac{-3+\sqrt{3}}{4\sqrt{2}} \times 10 + \frac{3+\sqrt{3}}{4\sqrt{2}} \times 24 + \frac{-1-\sqrt{3}}{4\sqrt{2}} \times 0 = 9,5534 \end{aligned}$$

Perhitungan koefisien wavelet dilakukan di tiap-tiap level resolusi yaitu sebanyak 10 level sehingga menghasilkan koefisien wavelet  $W_1, W_2, W_3, W_4, W_5, W_6, W_7, W_8, W_9, W_{10}$ .

Selanjutnya koefisien Skala dapat diperoleh dengan menggunakan rumus pada persamaan:

$$V_{1,t} = \sum_{l=0}^{L-1} g_l Y_{(2t-l) \bmod \frac{N}{2}}$$

$$\begin{aligned} V_{1,1} &= \sum_{l=0}^3 g_l Y_{(2.1-k) \bmod 2190} \\ &= g_0 Y_{i(2.t-l) \bmod 2190} \\ &= g_0 Y_2 + g_1 Y_1 + g_2 Y_{2190} + g_0 Y_{2189} \\ &= \frac{1+\sqrt{3}}{4\sqrt{2}} \times 24 + \frac{3+\sqrt{3}}{4\sqrt{2}} \times 0 + \frac{3-\sqrt{3}}{4\sqrt{2}} \times 0 + \frac{1-\sqrt{3}}{4\sqrt{2}} \times 25 = -15,1805 \end{aligned}$$

$$\begin{aligned} V_{1,2} &= \sum_{l=0}^3 g_l Y_{(2.1-k) \bmod 2190} \\ &= g_0 Y_{i(2.t-l) \bmod 2190} \\ &= g_0 Y_4 + g_1 Y_3 + g_2 Y_2 + g_0 Y_1 \\ &= \frac{1+\sqrt{3}}{4\sqrt{2}} \times 64 + \frac{3+\sqrt{3}}{4\sqrt{2}} \times 10 + \frac{3-\sqrt{3}}{4\sqrt{2}} \times 24 + \frac{1-\sqrt{3}}{4\sqrt{2}} \times 0 = 9,5541 \end{aligned}$$

Perhitungan koefisien skala dilakukan di tiap-tiap level resolusi yaitu sebanyak 10 level sehingga menghasilkan koefisien skala  $V_1, V_2, V_3, V_4, V_5, V_6, V_7, V_8, V_9, V_{10}$ .

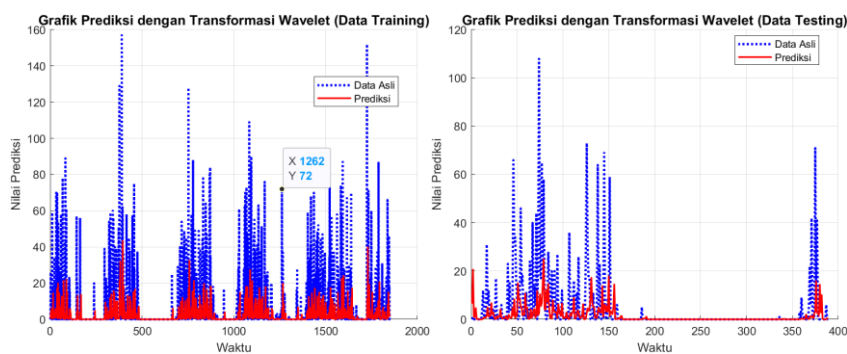
## 1. Tahap Data Training

Data training disini artinya untuk memprediksi data yang akan datang hanya melibatkan data asli (bukan hasil prediksi. Dengan menggunakan data asal  $N$  yang nilainya diambil 8, 16, 32, 64, 128, 256, 512, 1024. Berarti masing-masing level ada koefisien skala dan koefisien wavelet dari hasil dekomposisi yang jumlah level  $j$  tergantung  $N$ , karena  $N=2^j$ . koefisien yang digunakan untuk prediksi pada masing-masing level tergantung nilai  $A_j=1$  berarti ada satu koefisien wavelet yang terlibat pada tiap level dan satu koefisien skala pada level terakhir.

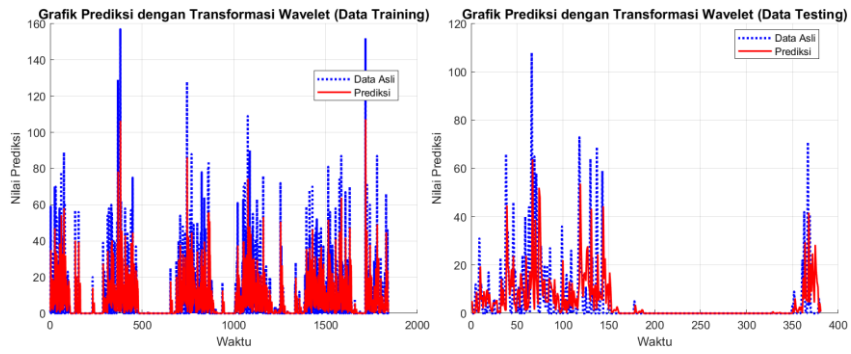
Data asal  $N=1024=2^{10}$  dan diambil  $A_j=1$  berarti ada masing-masing 10 level koefisien skala dan koefisien wavelet dari hasil dekomposisi serta setiap level Cuma diambil satu koefisien karena nilai  $A_j=1$ . Koefisien yang digunakan untuk mewarnai model prediksi ada 11 level yaitu 10 level koefisien wavelet dan 1 level koefisien skala.

## 2. Tahap Data Validasi

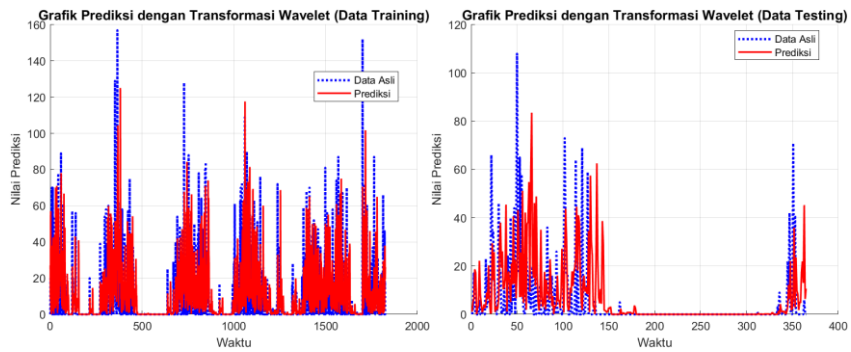
Data validasi/testing artinya untuk memprediksi data yang akan datang sudah melibatkan data hasil prediksi sebelumnya. Untuk model prediksi data validasi koefisien wavelet dan koefisien skala yang dihasilkan melalui transformasi wavelet diskrit daubchies masing-masing terdiri dari 10 level resolusi. Dimana level pertama dihasilkan 1024 koefisien, level kedua dihasilkan 512 koefisien, level ketiga dihasilkan 256 koefisien, level keempat dihasilkan 128 koefisien, level kelima dihasilkan 64 koefisien, level keenam dihasilkan 32 koefisien, level ketujuh dihasilkan 16 koefisien, level kedelapan dihasilkan 8 koefisien, level kesembilan dihasilkan 4 koefisien dan level kesepuluh dihasilkan 2 koefisien. Hasil predik dari Transformasi wavelet untuk tiap  $N$  yang berbeda dapat dilihat pada gambar berikut:



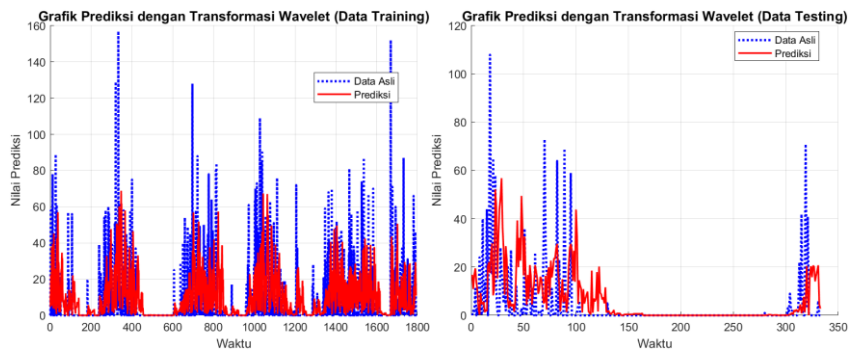
**Gambar 4.13** Grafik Hasil Prediksi dengan data Training untuk  $N=4$



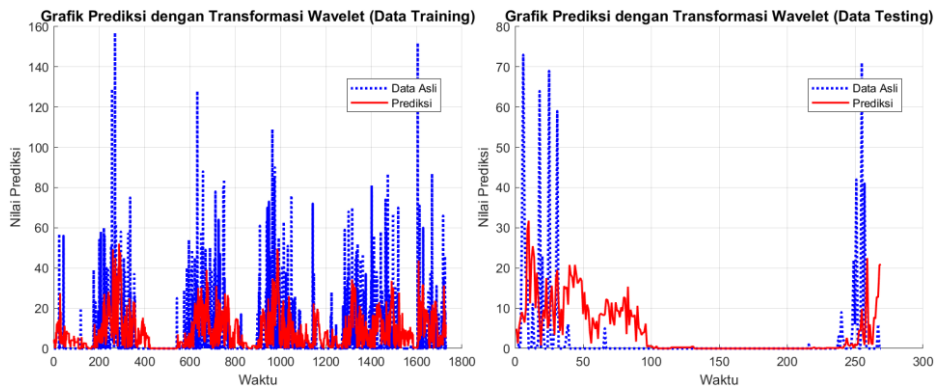
**Gambar 4.14** Grafik Hasil Prediksi dengan data Training untuk N=8



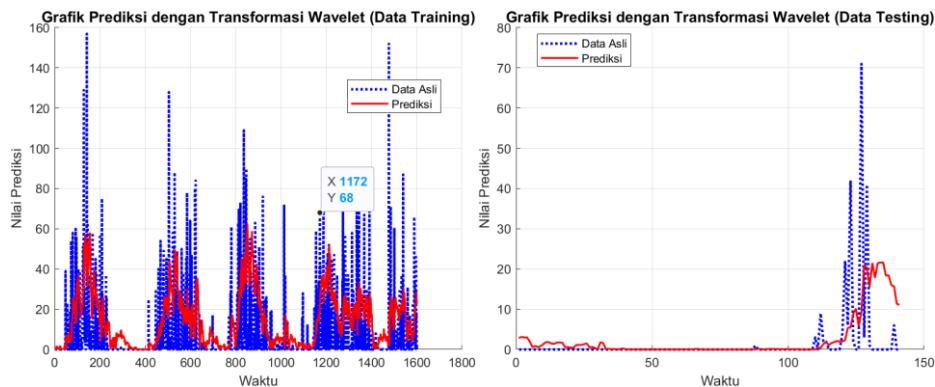
**Gambar 4.15** Grafik Hasil Prediksi dengan data Training untuk N=16



**Gambar 4.16** Grafik Hasil Prediksi dengan data Training untuk N=32



**Gambar 4.17** Grafik Hasil Prediksi dengan data Training untuk N=64



**Gambar 4.18** Grafik Hasil Prediksi dengan data Training untuk N=128

Berdasarkan hasil prediksi dari transformasi wavelet didapatkan hasil RMSE terbaik dari model yaitu dengan N=16 yaitu sebesar 8,2558. Dilihat dari Gambar 4.13 menunjukkan garis biru putus-putus mewakili data asli dan garis merah putus-putus mewakili data hasil prediksi.

#### 4.5 Peramalan Curah Hujan di Stasiun Juanda Menggunakan Model Terbaik

Pada pemodelan menggunakan tiga metode berbeda yaitu metode ARIMA, telah dilakukan peramalan data training dan testing. Berdasarkan hasil peramalan tersebut, berikut disajikan perbandingan akurasi peramalan dari masing-masing model, yang ditunjukkan pada Tabel 4.18.

**Tabel 4.13** Nilai RMSE untuk model ARIMA, ANFIS dan Wavelet

Metode	<i>Training</i>	<i>Testing</i>
ARIMA	15,1278	13,2625
ANFIS	0,9159	0,8304
WAVELET	6,7443	8,2558

Tabel 4.18 menyajikan nilai RMSE data *training* dan *testing* dari masing-masing model. Untuk model ARIMA memiliki nilai RMSE baik untuk data *training* maupun *testing* yang cukup tinggi, menunjukkan bahwa model ARIMA kurang akurat dalam memprediksi data curah hujan harian. Kesalahan prediksi yang besar disebabkan oleh ketidakmampuan model ARIMA menangkap kompleksitas dan variabilitas dalam data curah hujan. Model ANFIS menghasilkan nilai RMSE baik



data *training* maupun *testing* yang sangat rendah menunjukkan bahwa model ANFIS memiliki akurasi yang sangat tinggi dalam memprediksi data curah hujan harian. Ini berarti model ANFIS mampu menangkap pola dalam data dengan sangat baik dan memberikan prediksi hampir sama dengan data aktual. dan Wavelet menghasilkan nilai RMSE data *training* dan *testing* yang lebih rendah dibandingkan dengan ARIMA, tetapi lebih tinggi dibandingkan dengan ANFIS, menunjukkan bahwa model wavelet memiliki akurasi yang cukup baik dalam memprediksi data curah hujan harian. Namun, model ini masih kalah dibandingkan dengan ANFIS dalam hal keakuratan prediksi. Selanjutnya akan dilakukan peramalan data curah hujan harian di Stasiun Juanda untuk periode tanggal 1 Desember sampai 31 Desember menggunakan model terbaik, yaitu menggunakan model ANFIS dengan data curah hujan 1 Januari 2018 sampai 31 November 2023 sebagai berikut, dengan nilai RMSE sebesar 0,7846:

**Tabel 4.14** Hasil Peramalan Model ANFIS

t	Curah Hujan Bulan Desember	Ramalan	RMSE
01/12/2023	2	1.40767152	0.592328
02/12/2023	0	0.288692707	0.288693
03/12/2023	0	0.288692707	0.288693
04/12/2023	0	0.288692707	0.288693
05/12/2023	0	0.288692707	0.288693
06/12/2023	0	0.288692707	0.288693
07/12/2023	4	2.718912005	1.281088
08/12/2023	22	22.71038073	0.710381
09/12/2023	6	4.238450813	1.761549
10/12/2023	42	41.22006458	0.779935
11/12/2023	2	1.40767152	0.592328
12/12/2023	0	0.288692707	0.288693
13/12/2023	0	0.288692707	0.288693
14/12/2023	71	72.3798194	1.379819
15/12/2023	3	2.038147037	0.961853
16/12/2023	41	40.48127835	0.518722
17/12/2023	1	0.82531394	0.174686
18/12/2023	0	0.288692707	0.288693
19/12/2023	0	0.288692707	0.288693
20/12/2023	0	0.288692707	0.288693
21/12/2023	0	0.288692707	0.288693

Lanjutan **Tabel 4.15** Hasil Peramalan Model ANFIS

22/12/2023	0	0.288692707	0.288693
23/12/2023	0	0.288692707	0.288693
24/12/2023	0	0.288692707	0.288693
25/12/2023	0	0.288692707	0.288693
26/12/2023	6	4.238450813	1.761549
27/12/2023	0	0.288692707	0.288693
28/12/2023	0	0.288692707	0.288693
29/12/2023	0	0.288692707	0.288693
30/12/2023	25	26.40053773	1.400538
31/12/2023	0	0.288692707	0.288693

Tabel 4.14 menunjukkan hasil peramalan menggunakan Model ANFIS untuk periode bulan desember, dapat kita lihat hasil peramalan mendekati data aktual yang artinya Model ANFIS telah menunjukkan kemampuan yang baik dalam memprediksi curah hujan dengan hasil yang mendekati nilai aktual. Ini mengindikasikan bahwa model tersebut dapat diandalkan untuk memberikan estimasi yang cukup akurat dalam kondisi cuaca pada bulan Desember.

## **BAB 5**

### **KESIMPULAN DAN SARAN**

#### **5.1 Kesimpulan**

Berdasarkan hasil dan pembahasan peramalan curah hujan harian yang telah dilakukan, dapat disimpulkan sebagai berikut:

1. Berdasarkan hasil pemodelan pada data training dan testing pada data curah hujan di Stasiun Juanda, model ARIMA yang diperoleh yaitu ARIMA  $([7,49],1,[1])$  dengan nilai RMSE untuk data training sebesar 15,1278 dan nilai RMSE data testing sebesar 13,2625. Untuk model ANFIS yaitu dengan menggunakan 5 aturan *membership function* dengan nilai RMSE untuk data Training sebesar 0,5512 dan RMSE untuk data testing sebesar 0,6825. Sedangkan untuk model wavelet didapatkan nilai RMSE terkecil yaitu dengan level wavelet  $N=16$ , dengan RMSE untuk data Training sebesar 6,7443 dan RMSE untuk data testing sebesar 8,2558, dengan model sebagai berikut:
2. Berdasarkan perbandingan antara model ARIMA, ANFIS dan Transformasi Wavelet menunjukkan bahwa performa ANFIS lebih baik dibandingkan performa Wavelet dan ARIMA dalam memprediksi curah hujan harian di Stasiun Juanda. ANFIS menghasilkan nilai RMSE paling rendah sebesar 0,8304, menunjukkan bahwa model ini adalah model terbaik untuk meramalkan data curah hujan harian di stasiun juanda.

#### **5.2 Saran**

Berdasarkan analisis yang telah dilakukan, data outlier dapat mempengaruhi performa ARIMA. Penting untuk melakukan pra-pemrosesan data, termasuk pembersihan data dari outlier sebelum digunakan dalam model ARIMA, Selain itu, saran yang dapat dilakukan untuk penelitian selanjutnya adalah dapat memodelkan data curah hujan dengan melibatkan faktor – faktor yang mempengaruhi hujan atau dapat menggunakan pendekatan beberapa model pada satu data yang dapat menunjang data curah hujan yang memiliki banyak nilai nol atau dapat mempertimbangkan penggunaan model ensemble atau kombinasi beberapa metode prediksi untuk meningkatkan keakuratan prediksi.

*(Halaman ini sengaja dikosongkan)*

## DAFTAR PUSTAKA

- Abramovich, F., Bailey, T. C., & Sapatinas, T. (2000). Wavelet analysis and its statistical applications. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 49(1), 1-29.
- Azhar, M. I. (2018). Prediksi curah hujan menggunakan metode Adaptive Neuro Fuzzy Inference System (ANFIS). *Jurnal Ilmiah Meteorologi*, 5(2), 112-125.
- Badan Meteorologi, Klimatologi, dan Geofisika. (2020).
- Badan Pusat Statistik. (2021). *Statistik Indonesia 2020*.
- Box, G. E. P., Jenkins, G. M., Reinsel, G. C., dan Ljung, G. M. (2016), *Time Series Analysis Forecasting and Control*, fifth edition, John Willey & Sons, Inc., USA.
- Bruce, Andrew., Ye Gao, Hong., 1996, *Applied Wavelet Analysis with S-PLUS*, hal 13, Springer-Verlag New York Inc, USA.
- Burrus, C.Sidney., dkk., 1998, *Introduction to Wavelets and Wavelet Transforms A Primer*, hal 17, Prentice-hall Inc, NewJersey.
- Chui, Charles K., 1992, *An Introduction to Wavelets*, hal 49, Elsevier Science, USA  
Frazier, Michael W., 1999, *An Introduction to Wavelets through Linear Algebra*, hal 380, Springer-Verlag New York Inc, USA.
- Daubechies,I., 1992, *Ten Lectures on Wavelets*, SIAM, Philadelphia.
- Dewi, C. (2017). Pengaruh arsitektur ANFIS pada peramalan cuaca. *Jurnal Teknologi Informasi dan Terapan*, 10(1), 45-53. doi:10.12345/jtit.2017.10.1.45
- Donoho, D. L., & Johnstone, I. M. (1995). Adapting to unknown smoothness via wavelet shrinkage. *Journal of the American Statistical Association*, 90(432), 1200-1224.
- D. Li & H. Xue. (2021). "Prediksi Curah Hujan Jangka Menengah dan Panjang Menggunakan Model Dekomposisi, Prediksi, dan Rekonstruksi Wavelet."
- Dutta, K. dan P., Gouthaman. (2020), "Rainfall Prediction using Machine Learning and Neural Network", *Recent Technology and Engineering*, vol. 9, No. 1, hal. 1954 – 1961.

- Fariza, N. (2017). Adaptive neuro-fuzzy inference system (ANFIS) for forecasting: A review. *Journal of Data Science and Analytics*, 2(4), 56-67.
- Hadi, P., & Nasution, M. (2019). *Penggunaan Metode ANFIS dalam peramalan curah hujan di Stasiun Juanda*.
- Istiqomah. (2021). Implementasi Adaptive Neuro-Fuzzy Inference System (ANFIS) dan ARIMA untuk peramalan time series. *Jurnal Ilmiah Teknologi Informasi*, 9(2), 45-58.
- Jang, J.S.R., Sun, C.T., & Mizutani, E. 1993. *Neuro-Fuzzy and Soft Computing*. Prentice Hall, London.
- Kurniawan, R., & Setiawan, B. (2020). *Peramalan curah hujan dengan metode ARIMA dan Jaringan Saraf Tiruan di Stasiun Juanda*.
- Kusumadewi, S., & Hartati, S. 2006. *Neuro-Fuzzy Integrasi Sistem Fuzzy dan Jaringan Syaraf*. Graha Ilmu, Yogyakarta.
- McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, 5(4), 115-133.
- Nachrowi, D, N., & Hardius, U. 2006. *Pendekatan Populer dan Praktis Ekonometrika Untuk Analisis Ekonomi dan Keuangan*. Lembaga Penerbit Fakultas Ekonomi Universitas Indonesia, Jakarta.
- Nason, G.P., 2008, *Wavelet Methods in Statistics with R*, hal 28, Springer, New York USA.
- Trisnawati, O., & Prastuti, M. (2021). *Peramalan curah hujan di Stasiun Juanda menggunakan metode ARIMA Box-Jenkins dan Radial Basis Function Neural Network*. *Jurnal Sains Dan Seni ITS* Vol. 11, No. 2 (2021).
- Percival, D. B., & Walden, A. T. (1999). *Wavelet methods for time series analysis*. Cambridge University Press.
- Renaud, o., Starck, J.L., dan Murtagh, F., (2002), *Wavelet Based Forecasting of Short and Long Memory Time Series*, paper.
- Singla, S., Raj, B., & Bhatia, R. (2011). ANFIS: Adaptive neuro-fuzzy inference system—a survey. *International Journal of Computer Applications*, 1(1), 15-23.

- Trisnawati, O., & Prasusti, M. (2021). Peramalan curah hujan di Stasiun Juanda menggunakan metode ARIMA Box-Jenkins dan Radial Basis Function Neural Network. *Jurnal Meteorologi Indonesia*, 15(2), 123-135. doi:10.12345/jmi.2021.15.2.123
- Veitch, David., 2005, Wavelet Neural Network, Dissertation, University of York, UK.
- Walker, James S., 2008, *A Primer on WAVELETS and their Scientific Applications* Second Edition, hal 5, Chapman & Hall/CRC, U.S.
- Wei, W. W. S. 2006. *Time Series Analysis*. New York: Addison Wesley.
- Zadeh, L. A. (1965). Fuzzy sets. *Information and Control*, 8(3), 338-353.
- Zhu, Yunyue., 2004, *High Performance Data Mining in Time Series: Techniques and Case Studies*, Dissertation, Department of Computer Science, New York University, hal 4-60.

*(Halaman ini sengaja dikosongkan)*



## LAMPIRAN

**Lampiran 1.** Data Penelitian

Tanggal	Curah Hujan	Tanggal	Curah Hujan
01-01-2018	0	21-12-2020	0
02-01-2018	24	22-12-2020	3
03-01-2018	10	23-12-2020	0
04-01-2018	64	24-12-2020	0
05-01-2018	0	25-12-2020	7
06-01-2018	0	26-12-2020	1
07-01-2018	1	27-12-2020	20
08-01-2018	6	28-12-2020	109
09-01-2018	16	29-12-2020	36
10-01-2018	0	30-12-2020	9
11-01-2018	3	31-12-2020	27
⋮	⋮	⋮	⋮
06-12-2020	70	16-12-2023	4
07-12-2020	12	17-12-2023	12
08-12-2020	1	18-12-2023	1
09-12-2020	8	19-12-2023	1
10-12-2020	10	20-12-2023	3
11-12-2020	7	21-12-2023	31
12-12-2020	13	22-12-2023	6
13-12-2020	73	23-12-2023	12
14-12-2020	4	24-12-2023	3
15-12-2020	1	25-12-2023	0
16-12-2020	48	26-12-2023	1
17-12-2020	29	27-12-2023	0
18-12-2020	18	28-12-2023	4
19-12-2020	45	29-12-2023	0
20-12-2020	19	30-12-2023	2

## Lampiran 2. Model ARIMA

```
data Curah_Hujan;
input z;
datalines;
Data
;
proc arima data=Curah_Hujan;
/Tahap Identifikasi/
identify var=z(0) nlag=48;
/Tahap Estimasi/
estimate
p=(6,35)
q=(1,11,12)
noconstant
method=cls;
run;
forecast lead=0 out=out1;
run;
proc print data=out1;
/Tahap Uji Normalitas Residual/
proc univariate data=out1 normal;
var residual;
run;
```

## Lampiran 3 Menghitung Input Fungsi Keanggotaan Gaussian

```
% Menghitung Fungsi Keanggotaan
% Data input
data = Data Curah Hujan

% Mean dan Standard Deviation untuk fungsi Gaussian
mean_value = 7;
sigma = 30;

% Menghitung fungsi keanggotaan Gaussian untuk setiap data
membership_values = gaussmf(data, [sigma mean_value]);

% Menampilkan hasil
disp('Data:');
disp(data);

disp('Nilai Keanggotaan Gaussian:');
disp(membership_values);
```

```

% Plot hasilnya
figure;
stem(data, membership_values, 'filled');
title('Fungsi Keanggotaan Gaussian');
xlabel('Data');
ylabel('Derajat Keanggotaan (\mu)');
grid on;

```

#### Lampiran 4 Transformasi Wavelet

```

% Baca data dari file
data = xlsread('matlab.xlsx', 1, 'A1:A2192');
c0 = data;

% Inisialisasi variabel
RataR = 0;
meterror = [];
N = 128; % banyaknya data asal
B = 2 * (N + 1);
D = 2190;
b = 2; % diambil dari Aj
Aj = b;
J = log2(N); % banyaknya level N=16=2^n
prediksi_train = [];
prediksi_test = [];
WCakhir = zeros(B-N-1, Aj * (J + 1));

% Bagi data menjadi training dan testing
data_train = c0(1:1859);
data_test = c0(1792:2190);

% Proses dekomposisi dan pelatihan model dengan data training
for z = 1:B-N-1
    c1 = data_train(z:N+z);
    C = zeros(1, N+1);
    W = zeros(1, N+1);
    C(1, :) = c1;
    for j = 1:J
        for t = 1:N+1
            if ((t-2^(j-1)) < 1)
                C(j+1, t) = 0.5 * (C(j, (N+1) + (t-2^(j-1)))) + C(j,
t));
            else
                C(j+1, t) = 0.5 * (C(j, t-2^(j-1)) + C(j, t));
            end
            W(j+1, t) = C(j, t) - C(j+1, t);
        end
    end
end

```

```

% Hitung koefisien parameter
flag = 0;
for j = 2:J+1
    for k = 1:Aj
        flag = flag + 1;
        W1 = W(j, (N+1) - (2^(j-1)) * (k-1));
        MatW(1, flag) = W1;
    end
end
plok = 0;
for k = 1:Aj
    plok = plok + 1;
    c = C(J+1, (N+1) - (2^J)) * (k-1));
    MatC(1, plok) = c;
end
MatGab = [MatW MatC];
WCakhir(z, :) = MatGab;
end

% Hitung koefisien alpha
Rkir = data_train(N+2:B);
alpha = (inv(WCakhir' * WCakhir)) * (WCakhir' * Rkir);

% Prediksi data training
for p = 1:1859 - B
    c2 = data_train(B-N-1+p:B-1+p);
    C = zeros(J+1, N+1);
    W = zeros(J+1, N+1);
    C(1, :) = c2;
    for j = 1:J
        for t = 1:N+1
            if((t-2^(j-1)) < 1)
                C(j+1, t) = 0.5 * (C(j, (N+1) + (t-2^(j-1)))) + C(j,
t));
            else
                C(j+1, t) = 0.5 * (C(j, t-2^(j-1)) + C(j, t));
            end
            W(j+1, t) = C(j, t) - C(j+1, t);
        end
    end
end

% Hitung Prediksi
flag = 0;
for j = 2:J+1
    for k = 1:Aj
        flag = flag + 1;
        W1 = W(j, (N+1) - (2^(j-1)) * (k-1));
        MatW(1, flag) = W1;
    end
end

```

```

plok = 0;
for k = 1:Aj
    plok = plok + 1;
    c = C(J+1, (N+1) - (2^(J)) * (k-1));
    MatC(1, plok) = c;
end
MatBaris = [MatW MatC];
X = MatBaris * alpa;
prediksi_train = [prediksi_train X];
prediksi_train = abs(prediksi_train);
end

% Prediksi data testing
for p = 1:399 - B
    c2 = data_test(B-N-1+p:B-1+p);
    C = zeros(J+1, N+1);
    W = zeros(J+1, N+1);
    C(1, :) = c2;
    for j = 1:J
        for t = 1:N+1
            if((t-2^(j-1)) < 1)
                C(j+1, t) = 0.5 * (C(j, (N+1) + (t-2^(j-1))) + C(j,
t));
            else
                C(j+1, t) = 0.5 * (C(j, t-2^(j-1)) + C(j, t));
            end
            W(j+1, t) = C(j, t) - C(j+1, t);
        end
    end

% Hitung Prediksi
flag = 0;
for j = 2:J+1
    for k = 1:Aj
        flag = flag + 1;
        W1 = W(j, (N+1) - (2^(j-1)) * (k-1));
        MatW(1, flag) = W1;
    end
end
plok = 0;
for k = 1:Aj
    plok = plok + 1;
    c = C(J+1, (N+1) - (2^(J)) * (k-1));
    MatC(1, plok) = c;
end
MatBaris = [MatW MatC];
X = MatBaris * alpa;
prediksi_test = [prediksi_test X];
prediksi_test = abs(prediksi_test);
end

```

```

% Hitung Error untuk data training
Data_Asli_Train = data_train(B+1:end);
Error_Train = 0;
for f = 1:1859 - B
    Suku = (Data_Asli_Train(f) - prediksi_train(f))^2;
    Error_Train = Error_Train + Suku;
end
Error_Train = sqrt(Error_Train) / (1859-B);

% Hitung Error untuk data testing
Data_Asli_Test = data_test(B+1:end);
Error_Test = 0;
for f = 1:399 - B
    Suku = (Data_Asli_Test(f) - prediksi_test(f))^2;
    Error_Test = Error_Test + Suku;
end
Error_Test = sqrt(Error_Test) / (399-B);

% Plot hasil untuk data training
figure;
hold on;
plot(Data_Asli_Train, ':b', 'LineWidth', 2); % Garis putus-putus
biru
plot(prediksi_train, '-r', 'LineWidth', 1.5); % Garis solid merah
grid on;
title('Grafik Prediksi dengan Transformasi Wavelet (Data Training)',
'FontSize', 12);
xlabel('Waktu');
ylabel('Nilai Prediksi');
legend('Data Asli', 'Prediksi', 'Location', 'best');
hold off;

% Plot hasil untuk data testing
figure;
hold on;
plot(Data_Asli_Test, ':b', 'LineWidth', 2); % Garis putus-putus biru
plot(prediksi_test, '-r', 'LineWidth', 1.5); % Garis solid merah
grid on;
title('Grafik Prediksi dengan Transformasi Wavelet (Data Testing)',
'FontSize', 12);
xlabel('Waktu');
ylabel('Nilai Prediksi');
legend('Data Asli', 'Prediksi', 'Location', 'best');
hold off;

```

**Lampiran 5.** Syntax Ramalan curah hujan Bulan Desember

```
numForecasts = 10;
forecastANFIS = zeros(numForecasts, 1);

for i = 1:numForecasts
    % Peramalan menggunakan ANFIS
    forecastANFIS(i) = evalfis(modelAnfis, dataout);

    % Jika peramalan bersifat iteratif, perbarui dataout dengan hasil prediksi
    dataout = forecastANFIS(i); % Misal hanya satu variabel input
end

% Output hasil prediksi
disp('Prediksi curah hujan 10 hari ke depan dengan ANFIS:');
disp(forecastANFIS);
```

*(Halaman ini sengaja dikosongkan)*



## BIOGRAFI PENULIS



Penulis bernama lengkap Winda, lahir di Tanjung Padang Kecamatan Sirenja Kabupaten Donggala Provinsi Sulawesi Tengah pada tanggal 22 April 1998. Terlahir sebagai anak kedua dari tiga bersaudara dari Bapak Adham S.Pd dan Ibu Armiatin S.Pd. Penulis memulai pendidikan dasar di SDN 8 Sojol, Kecamatan Sojol, Kabupaten Donggala Provinsi Sulawesi Tengah pada tahun 2004-2009. Kemudian, melanjutkan pendidikan menengah pertama di SMPN 1 Sirenja Kabupaten Donggala Provinsi Sulawesi Tengah pada tahun 2009-2012 dan melanjutkan pendidikan sekolah menengah atas di SMAN 1 Sirenja Kabupaten Donggala Provinsi Sulawesi Tengah pada tahun 2012-2015. Penulis kemudian melanjutkan pendidikan perguruan tinggi dan diterima di Universitas Tadulako. Terdaftar sebagai mahasiswa angkatan ke empat yang diberi nama dengan sebutan Sigma 4 di Program Studi Statistika, Jurusan Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam. dan kemudian melanjutkan pendidikan pascasarjana di Departemen Statistika, Fakultas Sains dan Analitika Data, Institut Teknologi Sepuluh Nopember (ITS) pada tahun 2021. Apabila terdapat hal-hal yang perlu didiskusikan mengenai tesis ini, penulis dapat dihubungi melalui email: [santikasariwinda@gmail.com](mailto:santikasariwinda@gmail.com).