



TESIS - SS142501

***Peramalan Banyak Kasus Demam Berdarah
(DB) di Kota Surabaya Menggunakan Hybrid
Integer-valued Autoregressive Integrated
Moving Average (INARIMA) dan Radial Basis
Function Neural Network (RBFNN)***

Fajarani Juliaristi

1314201203

DOSEN PEMBIMBING

Dr. Brodjol Sutijo S.U., M.Si

Dr. Kartika Fithriasari, M.Si

PROGRAM STUDI MAGISTER

JURUSAN STATISTIKA

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM

INSTITUT TEKNOLOGI SEPULUH NOPEMBER

SURABAYA

2016



THESES - SS142501

**Forecasting The Number Of Dengue Fever
Using Hybrid Integer-valued Autoregressive
Integrated Moving Average (INARIMA) and
Radial Basis Function Neural Network (RBFNN)**

Fajarani Juliaristi

1314201203

SUPERVISOR

Dr. Brodjol Sutijo S.U., M.Si

Dr. Kartika Fithriasari, M.Si

MAGISTER PROGRAM

DEPARTMENT OF STATISTIKA

FACULTY OF MATHEMATICS AND NATURAL SCIENCES

SEPULUH NOPEMBER INSTITUTE OF TECHNOLOGY

SURABAYA

2016

**PERAMALAN BANYAK KASUS DEMAM BERDARAH (DB) DI KOTA
SURABAYA MENGGUNAKAN *HYBRID INTEGER-VALUED
AUTOREGRESSIVE INTEGRATED MOVING AVERAGE (INARIMA) DAN
RADIAL BASIS FUNCTION NEURAL NETWORK (RBFNN)***

**Tesis ini disusun untuk memenuhi salah satu syarat memperoleh gelar
Magister Sains (M.Si)**

**di
Institut Teknologi Sepuluh Nopember**

Oleh:

**FAJARANI JULIARISTI
NRP. 1314201203**

Tanggal Ujian : 21 Juli 2016

Periode Wisuda : September 2016

Disetujui oleh :


1. **Dr. Brodjot Sutijo S.U., M.Si.**
NIP. 196601251990021001

(Pembimbing I)


2. **Dr. Kartika Fithriasari, M.Si.**
NIP. 196912121993032002

(Pembimbing II)


3. **Dr. Bambang Widjanarko O., M.Si.**
NIP. 196811241994121001

(Penguji)


4. **Dr. rer pol. Heri Kuswanto, M.Si.**
NIP. 198203262003121004

(Penguji)



Direktur Program Pascasarjana,


Prof. Ir. Djauhar Manfaat, M.Sc., Ph.D.
NIP. 19601202 198701 1 001

**Peramalan Banyak Kasus Demam Berdarah (DB) di Kota
Surabaya Menggunakan *Hybrid Integer-valued Autoregressive
Integrated Moving Average (INARIMA)* dan *Radial Basis Function
Neural Network (RBFNN)***

Nama : Fajarani Juliaristi
NRP : 1314201203
Pembimbing : Dr. Brodjol Sutidjo S.U., M.Si
Dr. Kartika Fithriasari, M.Si

ABSTRAK

Metode *Hybrid* merupakan kombinasi dua atau lebih sistem dalam satu fungsi. Pada penelitian ini digunakan kombinasi *Integer-valued Autoregressive Integrated Moving Average (INARIMA)* dengan *Radial Basis Function Neural Network (RBFNN)*. *INARIMA* digunakan untuk memodelkan komponen linier dan menghasilkan nilai *residual*. Nilai *residual* yang tersebut kemudian akan di analisis menggunakan *RBFNN*. Pada model *INARIMA*, *input* yang digunakan adalah variabel acak bilangan bulat positif yang memenuhi *identically and independent distributed (iid)*. Pada model *INARIMA* juga digunakan *Binomial Thinning Operator* yang merupakan jumlahan dari variabel acak *Bernoulli*. Parameter yang belum diketahui pada model *INARIMA* akan diestimasi menggunakan Metode *Conditional Least Square (CLS)*. Kombinasi dari model *INARIMA* yang telah dijelaskan sebelumnya adalah *RBFNN*. *RBFNN* adalah salah satu model *Neural Network (NN)* yang digunakan untuk menyelesaikan permasalahan nonlinier. Pada proses pembelajaran *RBFNN* terjadi proses nonlinier dari lapisan *input* ke lapisan tersembunyi dengan menggunakan fungsi aktivasi *Gaussian*, sedangkan pada lapisan tersembunyi ke lapisan *output* terjadi proses linier. Pada proses menentukan model terbaik, digunakan Metode *Self Organizing Map (SOM)* untuk mengelompokkan data *input* menjadi beberapa kelompok sehingga nilai pusat dan varians setiap kelompok dapat dihitung. Aplikasi Metode *Hybrid INARIMA* dan *RBFNN* diterapkan untuk meramalkan banyak kasus demam berdarah di Kota Surabaya. Data yang digunakan adalah data banyak kasus demam berdarah di Kota Surabaya dari bulan Januari 2006 sampai Desember 2014. Penelitian ini menghasilkan model terbaik *Hybrid INAR(1)* dengan variabel prediktor x_t dan *RBFNN* dengan 2 *neuron input*, 2 *neuron* pada lapisan tersembunyi, dan 1 *neuron output*. Model ini menghasilkan nilai *Mean Square Error (MSE)* sebesar 371,494 dan *Root Mean Square Error (RMSE)* sebesar 19,274.

Kata Kunci: DB, *Hybrid*, *INARIMA*, *CLS*, *RBFNN*, *SOM*.

Halaman ini sengaja dikosongkan.

Forecasting The Number Of Dengue Fever Using Hybrid Integer-valued Autoregressive Integrated Moving Average (INARIMA) and Radial Basis Function Neural Network (RBFNN)

Name : Fajarani Juliaristi
NRP : 1314201203
Supervisor : Dr. Brodjol Sutidjo S.U., M.Si
Dr. Kartika Fithriasari, M.S

ABSTRACT

Hybrid method is a combination of two or more systems in one function. In this study, the combination is Integer-valued Autoregressive Integrated Moving Average (INARIMA) with Radial Basis Function Neural Network (RBFNN). INARIMA modelling the linear component of data and produce the residual. Residual value will analysed using RBFNN. In INARIMA process, the input is nonnegative integer random variable and it is identically and independent distributed (iid). In INARIMA model, there is Binomial Thinning Operator is sum of Bernoulli random variable. The unknown parameter in INARIMA model estimated using Conditional Least Square (CLS). The combination of INARIMA is RBFNN. RBFNN is the one of Neural Network (NN) model which it use to solve nonlinear problem. In RBFNN learning, there is nonlinear process from input layer to hidden layer using Gaussian activation function and linear process from hidden layer to output. To determining the optimum model, Self Organizing Map (SOM) classify the input data to several groups then the centroid value and variance each groups can be calculate. The application of Hybrid INARIMA and RBFNN method is forecast the number of dengue fever case in Surabaya from January 2006 until December 2014. This study obtain the best model is Hybrid INAR(1) with explanatory variable x_1 and RBFNN with 2 neurons in input layer, 2 neurons in hidden layer, and 1 neuron in output layer. This model produce Mean Square Error (MSE) 371.494 and Root Mean Square (RMSE) 19.274.

Keyword: Dengue Fever, Hybrid, INARIMA, RBFNN, CLS, SOM.

Halaman ini sengaja dikosongkan.

DAFTAR ISI

HALAMAN JUDUL	
LEMBAR PENGESAHAN	i
ABSTRAK	iii
ABSTRACT	v
KATA PENGANTAR	vii
DAFTAR ISI	ix
DAFTAR GAMBAR	xi
DAFTAR TABEL	xiii
DAFTAR LAMPIRAN	xv
BAB 1 PENDAHULUAN	1
A. Latar Belakang	1
B. Rumusan Masalah	5
C. Tujuan Penelitian	5
D. Manfaat Penelitian	5
E. Batasan Penelitian	6
BAB 2 KAJIAN TEORI	7
A. Distribusi <i>Poisson</i>	7
B. Prediksi (<i>Forecast</i>)	8
C. Analisis Deret Waktu (<i>Time Series Analysis</i>)	8
D. Stasioneritas	9
E. Kriteria Keباikan Model	11
F. Autokorelasi	12
G. Model <i>Hybrid</i>	16
H. Model <i>Autoregressive Integrated Moving Average (ARIMA)</i>	17
I. Model <i>Integer-valued Autoregressive Integrated Moving Average (INARIMA)</i>	18
J. <i>Conditional Least Square (CLS)</i>	20
K. <i>Neural Network (NN)</i>	21
L. <i>Self Organizing Map (SOM)</i>	24
M. <i>Radial Basis Function Neural Network (RBFNN)</i>	26
N. Demam Berdarah (DB)	30

BAB 3 METODE PENELITIAN	35
A. Analisis Variabel Penelitian	35
B. Metode Analisis	36
BAB 4 HASIL DAN PEMBAHASAN	39
A. Deskripsi Data Banyak Kasus Demam Berdarah	39
B. Deskripsi Data Iklim.....	40
C. Identifikasi Model <i>INARIMA</i>	41
D. Estimasi Parameter	44
E. Ramalan Model <i>INAR</i>	46
1. Pemodelan ARIMA Kelembaban Udara	47
2. Pemodelan ARIMA Temperatur	50
F. Ramalan Hybrid untuk Model <i>INAR(1)</i> dengan x_1	54
1. Identifikasi Model <i>RBFFNN</i>	55
2. Menentukan Nilai Pusat dan Varians	57
3. Prediksi <i>RBFFNN</i>	57
4. Prediksi <i>Hybrid</i>	60
G. Ramalan Hybrid untuk Model <i>INAR(2)</i>	62
1. Identifikasi Model <i>RBFFNN</i>	64
2. Menentukan Nilai Pusat dan Varians	65
3. Prediksi <i>RBFFNN</i>	65
4. Prediksi <i>Hybrid</i>	68
BAB 5 KESIMPULAN DAN SARAN	71
A. Kesimpulan.....	73
B. Saran	74
DAFTAR PUSTAKA	75
LAMPIRAN	
BIOGRAFI	

DAFTAR GAMBAR

Gambar 2.1.	Pola Data Stasioner.....	10
Gambar 2.2.	<i>Correlogram</i> untuk ACF	14
Gambar 2.3.	Jaringan Banyak Lapisan	22
Gambar 2.4.	Jaringan Lapisan Kompetitif	23
Gambar 2.5.	Fungsi Aktivasi Linier	23
Gambar 2.6.	Arsitektur SOM	25
Gambar 2.7.	Arsitektur Radial Basis Function Neural Network (RBFNN).....	27
Gambar 3.1.	Diagram Alur	38
Gambar 4.1.	<i>Plot</i> Banyak Kasus Demam Berdarah di Kota Surabaya	39
Gambar 4.2.	<i>Plot ACF</i> data banyak kasus DB di Kota Surabaya	42
Gambar 4.3.	<i>Plot PACF</i> data banyak kasus DB di Kota Surabaya	42
Gambar 4.4.	<i>Time Series Plot</i> Data Kelembaban Udara	47
Gambar 4.5.	<i>Box-Cox Plot</i> Kelembaban Udara Sebelum Transformasi	48
Gambar 4.6.	<i>Box-Cox Plot</i> Data Kelembaban Udara Setelah Transformasi	48
Gambar 4.7.	<i>Plot ACF</i> Data Kelembaban Udara	49
Gambar 4.8.	<i>Plot PACF</i> Data Kelembaban Udara	49
Gambar 4.9.	<i>Time Series Plot</i> Data Temperatur	51
Gambar 4.10.	<i>Plot ACF</i> Data Temperatur	52
Gambar 4.11.	<i>Plot ACF</i> Data Temperatur	52
Gambar 4.12.	<i>Plot</i> Data Aktual dengan Hasil Ramalan Model $INAR(1)$, x_t	55
Gambar 4.13.	<i>Plot ACF</i> a_t	56
Gambar 4.14.	<i>Plot PACF</i> a_t	56
Gambar 4.15.	Arsitektur Model terbaik <i>RBFNN</i>	58

Gambar 4.16. <i>Plot</i> Data Aktual dengan Data Hasil Ramalan <i>Hybrid</i>	62
Gambar 4.17. <i>Plot</i> Data Aktual dengan Hasil Ramalan Model <i>INAR(2)</i>	63
Gambar 4.18. <i>Plot ACF</i> a_t	64
Gambar 4.19. <i>Plot PACF</i> a_t	64
Gambar 4.20. Arsitektur Model terbaik <i>RBNN</i>	66
Gambar 4.21. <i>Plot</i> Data Aktual dengan Data Hasil Ramalan <i>Hybrid</i>	69

DAFTAR TABEL

Tabel 3.1. Variabel-variabel.....	35
Tabel 4.1. Hasil Estimasi <i>INAR(2)</i> dengan metode <i>CLS</i>	44
Tabel 4.2. Hasil Estimasi <i>INAR(2)</i> dengan x_1 dan x_3	45
Tabel 4.3. Hasil Estimasi Parameter untuk Kelembaban Udara	50
Tabel 4.4. Hasil Estimasi Parameter untuk Temperatur.....	53
Tabel 4.5. Nilai <i>MSE</i> dan <i>RMSE</i> untuk Model dugaan <i>INAR</i>	53
Tabel 4.6. Hasil Ramalan <i>INAR(1)</i> dengan x_1	54
Tabel 4.7. Nilai <i>MSE</i> dan <i>RMSE</i> untuk 4 sampai 7 Kelompok	57
Tabel 4.8. Bobot Pembelajaran <i>RBFNN</i>	58
Tabel 4.9. Hasil Ramalan <i>RBFNN</i>	60
Tabel 4.10. Hasil Ramalan <i>Hybrid INAR(1)</i> dengan x_1 dan <i>RBFNN</i>	61
Tabel 4.11. Hasil Ramalan <i>INAR(2)</i>	63
Tabel 4.12. Nilai <i>MSE</i> dan <i>RMSE</i> untuk 2 sampai 5 Kelompok	65
Tabel 4.13. Bobot Pembelajaran <i>RBFNN</i>	66
Tabel 4.14. Hasil Ramalan <i>RBFNN</i>	68
Tabel 4.15. Hasil Ramalan <i>Hybrid INAR(2)</i> dengan <i>RBFNN</i>	69
Tabel 4.16. Nilai <i>MSE</i> dan <i>RMSE</i> untuk 2 sampai 12 kelompok	70

Halaman ini sengaja dikosongkan.

DAFTAR LAMPIRAN

Lampiran 1. Data Banyak Kasus Demam Berdarah di Kota Surabaya	79
Lampiran 2. Data Kelembaban, Tekana Udara, Temperatur, Penyinaran Matahari, Curah Hujan, dan Kecepatan Angin	80
Lampiran 3. Sintaks Estimasi Parameter untuk model <i>INAR</i>	81
Lampiran 4. Sintaks <i>Self Organizing Maps (SOM)</i>	82
Lampiran 5. Sintaks <i>Radial Basis Function Neural Network (RBFNN)</i>	83
Lampiran 6. Nilai Pusat dan Varians model <i>RBFNN</i>	84
Lampiran 7. Uji <i>Dickey Fuller</i> untuk Stasioneritas Data Banyak Kasus DB di Surabaya	85
Lampiran 8. Hasil Ramalan Variabel Prediktor untuk Tahun 2015	86
Lampiran 9. Penjabaran Hitungan <i>Binomial Thinning Operation</i>	87

Halaman ini sengaja dikosongkan.

BAB 1

PENDAHULUAN

A. Latar Belakang

Salah satu masalah kesehatan masyarakat yang utama di Indonesia adalah penyakit Demam Berdarah (DB). DB adalah penyakit yang disebabkan oleh virus *Dengue* yang tergolong *Arthropod-Borne Virus* yang masuk ke peredaran darah manusia. DB ditularkan melalui gigitan nyamuk dari genus *Aedes* terutama *Aedes aegypti* atau *Aedes albopictus* yang membawa virus *Dengue*. DB banyak ditemukan di daerah tropis dan subtropis Asia Tenggara dan Wilayah Pasifik Barat.

Di Indonesia, DB sering menjadi penyebab terjadinya Kejadian Luar Biasa (KLB) maupun wabah. KLB merupakan kondisi timbulnya atau meningkatnya kejadian kesakitan atau kematian secara epidemiologis pada suatu daerah dalam kurun waktu tertentu. Pada tahun 2013, penderita DB yang tersebar di seluruh Indonesia sebanyak 112.511 kasus, sedangkan pada tahun 2014, kasus DB mengalami penurunan yaitu sebanyak 100.347 kasus dengan kematian sebanyak 907 orang (Kemenkes RI, 2015: 153). Demam berdarah pertama kali ditemukan di Kota Surabaya pada tahun 1968, dimana sebanyak 58 orang terinfeksi dan 24 orang diantaranya meninggal dunia (Kemenkes RI, 2010: 1).

Pada tahun 2014, Provinsi Jawa Timur tercatat sebagai provinsi tertinggi ketiga untuk banyak kasus kematian yang disebabkan DB, yaitu sebanyak 107 kasus. Berdasarkan data yang diperoleh dari tahun 2006 sampai 2014, kasus DB di Kota Surabaya selalu menempati angka tertinggi baik untuk banyak penderita dan banyak kasus kematian dibanding dengan kabupaten/kota lain di Provinsi Jawa Timur. Pada tahun 2013, di Kota Surabaya terdapat 2.196 penderita DB dengan kasus kematian sebanyak 15 orang. Selanjutnya pada tahun 2014, terjadi penurunan yaitu terdapat 816 penderita dan 17 kasus kematian (Dinkes Jatim, 2015).

Penyebaran penyakit DB dapat disebabkan oleh mobilitas penduduk yang tinggi, perkembangan wilayah perkotaan, perubahan iklim, urbanisasi, dan sebagainya. Perubahan iklim menyebabkan perubahan curah hujan, suhu, kelembaban udara dan arah angin sehingga berpengaruh terhadap perkembangbiakan vektor penyakit seperti nyamuk *Aedes*, *Malaria*, dan lain sebagainya (Kemenkes RI, 2010: 3). Curah hujan tinggi akan menyebabkan suhu lingkungan menjadi rendah sehingga kelembaban udara menjadi tinggi. Pada kondisi tersebut, perkembangbiakan nyamuk *Aedes* akan meningkat sehingga perlu diwaspadai adanya kenaikan banyak kasus DB.

Salah satu upaya yang dapat dilakukan untuk mengendalikan penyebaran kasus DB adalah dengan meramal banyaknya kasus DB pada waktu yang akan datang. Angka ramalan pada waktu yang akan datang ini dapat dijadikan acuan upaya pemerintah untuk meminimalkan penyebaran kasus DB khususnya di Kota Surabaya. Peramalan dapat dilakukan dengan beberapa metode, antara lain Metode Pemulusan (*Exponential*), Analisis Regresi, *Autoregressive Integrated Moving Average (ARIMA)*, Jaringan Saraf Tiruan (JST) dan lain sebagainya. Penelitian untuk memprediksi banyak kasus demam berdarah di Kota Surabaya pernah dilakukan oleh Mafrida (2012) dengan menggunakan Metode *Integer-valued Autoregressive INAR(1)*. Hasil penelitian tersebut menyatakan bahwa pemodelan *INAR(1)* dengan memasukkan semua variabel prediktor memberikan hasil yang lebih baik dibanding dengan hanya menggunakan variabel prediktor yang signifikan. Berdasarkan hal tersebut, peneliti ingin mengetahui apakah variabel prediktor memberikan pengaruh yang berbeda terhadap pemodelan *INAR* dengan order yang lain.

Data banyak kasus DB merupakan data deret waktu (*time series*) dimana pada suatu data deret waktu dapat mengandung struktur linier maupun nonlinier. Hal ini dapat mempersulit penentuan metode untuk menganalisis data tersebut. Berdasarkan hal tersebut, pada penelitian ini digunakan Metode *Hybrid* yaitu kombinasi dua atau lebih sistem dalam satu fungsi. Zhang (2003) mengembangkan Metode *Hybrid* yang merupakan kombinasi dari model linier dengan model nonlinier. Pada penelitian ini

digunakan kombinasi *Integer-valued Autoregressive Integrated Moving Average (INARIMA)* sebagai komponen linier dengan *Radial Basis Function Neural Network (RBFNN)* sebagai komponen nonlinier. Metode *Hybrid* sudah banyak digunakan dalam beberapa penelitian, seperti pada Tahun 2015, Khandelwal, Adhikari dan Verma menggunakan Metode *Hybrid ARIMA* dan *Artificial Neural Network (ANN)* untuk melakukan prediksi yang dipresisi dengan Metode *Discrete Wavelet Transform (DWT)*. Wiyanti dan Pulungan (2012) melakukan penelitian menggunakan Metode *Hybrid Radial Basis Function (RBF)* dan *ARIMA* untuk memprediksi Indeks Harga Perdagangan Besar (IHPB) dan inflasi komoditi Indonesia. Pada tahun 2015, Shukur dan Lee memprediksi kecepatan angin harian dengan *Hybrid Kalman Filter (KF) ANN* dan *ARIMA*.

Data banyak kasus DB merupakan data *count* yang disajikan dalam data bulanan. Data *count* adalah data yang dihitung dari banyaknya kejadian dalam interval waktu atau interval ruang. Misalkan data banyaknya kecelakaan lalu lintas di Kota Surabaya setiap bulan, banyaknya kasus kematian akibat penyakit Malaria setiap bulan. Data *count* merupakan data deret waktu diskrit, dimana data diskrit adalah data kuantitatif yang nilainya diperoleh dari hasil perhitungan dan bernilai bulat positif atau secara matematis termasuk dalam himpunan bilangan cacah. Salah satu model yang dapat digunakan untuk data *count* adalah *Integer-valued Autoregressive Integrated Moving Average (INARIMA)*. Nilai *input* yang digunakan merupakan variabel acak hasil penjumlahan atau data *count* yang dependen dengan observasi pada waktu sebelumnya. Pada model *INARIMA* juga digunakan *Binomial Thinning Operator* yang merupakan jumlahan dari variabel acak *Bernoulli*.

Penelitian mengenai model *INARIMA* telah banyak dilakukan antara lain pada Tahun 2009, Silva *et al.* melakukan prediksi pada Model *Integer-valued Autoregressive INAR(1)* dan membandingkan hasil prediksi antara Metode Klasik dan Metode *Bayes*. Pedeli dan Karlis (2013) melakukan pengembangan model *INAR(1)* menjadi *Bivariate INAR* atau *BINAR(1)*. Pada tahun 2015, Weib menguraikan cara kerja *Poisson INAR(1)* untuk data ekonomi.

Pada penelitian ini, kombinasi dari model *INARIMA* yang telah dijelaskan sebelumnya adalah *RBFNN*. *RBFNN* adalah salah satu metode pengembangan dari *Neural Network (NN)*. Model *NN* telah banyak digunakan sebagai metode penelitian seperti Fithriasari dkk. (2013) yang memodelkan *Multivariate Time Series* menggunakan *Time Delay Neural Network*. Pada Tahun 2016, Riyadi dkk. menggunakan *Cluster Time Series* sebagai *Preprocessing* untuk meramalkan konsumsi listrik. Zahrati dan Fithriasari (2016) melakukan peramalan temperatur di Semarang menggunakan *Multi-output Neural Network*. Ciri khas *RBFNN* adalah hanya memiliki satu lapisan tersembunyi, sedangkan algoritma *NN* yang lain dapat memiliki satu atau lebih lapisan tersembunyi. Menurut Park dan Sandberg (1991), *RBFNN* memiliki kelebihan yaitu memiliki struktur jaringan yang sederhana sehingga proses pembelajarannya cepat dan juga memiliki kemampuan prediksi yang lebih baik jika dibandingkan dengan *NN* yang lain. Selain dapat digunakan sebagai metode peramalan, *RBFNN* juga dapat diaplikasikan pada beberapa permasalahan lain seperti regresi dan klasifikasi (Orr, 1996). Pada proses pembelajaran *RBFNN* terjadi proses nonlinier dari lapisan *input* ke lapisan tersembunyi, sedangkan pada lapisan tersembunyi ke lapisan *output* terjadi proses linier. Pada proses *RBFNN*, data *input* dikelompokkan menjadi beberapa kelompok, dimana banyak kelompok merupakan banyaknya *neuron* pada lapisan tersembunyi. Model *RBFNN* juga telah digunakan pada banyak penelitian, seperti Zhu, Cao dan Zhu (2014) memprediksi volume lalu lintas menggunakan *RBFNN* dengan memperhatikan arus lalu lintas di persimpangan. Selanjutnya pada Tahun 2015, Sugiartawan dan Mauko melakukan optimasi *RBFNN* dengan *Growing Hierarchial Self Organizing Map* untuk mereduksi dimensi *input*. Tatar, et al. (2015) memprediksi sifat air garam di waduk menggunakan *RBFNN*.

B. Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan di atas, diperoleh rumusan masalah pada penelitian ini adalah:

1. Bagaimana hasil peramalan banyak kasus demam berdarah di Kota Surabaya menggunakan Model *INARIMA*?
2. Bagaimana hasil peramalan banyak kasus demam berdarah di Kota Surabaya menggunakan *RBFNN*?
3. Bagaimana hasil peramalan banyak kasus demam berdarah di Kota Surabaya menggunakan Model *Hybrid INARIMA* dan *RBFNN*?

C. Tujuan Penelitian

Berdasarkan rumusan masalah, tujuan penelitian ini adalah:

1. Meramalkan banyak kasus demam berdarah di Kota Surabaya menggunakan Model *INARIMA*.
2. Meramalkan banyak kasus demam berdarah di Kota Surabaya menggunakan *RBFNN*.
3. Meramalkan banyak kasus demam berdarah di Kota Surabaya menggunakan Model *Hybrid INARIMA* dan *RBFNN*.

D. Manfaat Penelitian

Manfaat yang diperoleh dari penelitian ini adalah:

1. Menambah pengetahuan mengenai model *INARIMA*, *RBFNN*, dan Model *Hybrid INARIMA-RBFNN*.
2. Memberikan suatu metode alternatif untuk melakukan peramalan pada data deret waktu.

E. Batasan Penelitian

Sebagai batasan penelitian, pada penelitian ini estimasi parameter pada Model *INARIMA* dilakukan dengan menggunakan Metode *Conditional Least Square (CLS)*. Kemudian untuk proses pengelompokkan *input* pada pembelajaran *RBFNN* digunakan Metode *Self Organizing Map (SOM)* dan fungsi aktivasi *Gaussian* pada lapisan tersembunyi. Pada proses linier *RBFNN*, estimasi parameter menggunakan Metode *Least Square* yang dibantu dengan Metode *Global Ridge Regression* untuk menentukan parameter regulasi.

BAB 2

KAJIAN TEORI

Pada bab ini diberikan beberapa teori yang digunakan dan mendukung penelitian ini. Teori-teori yang akan dibahas meliputi distribusi *Poisson*, prediksi (*forecast*), analisis deret waktu (*time series analysis*), stasioneritas, kriteria kebaikan model, autokorelasi, Model *Hybrid*, Model *Autoregressive Integrated Moving Average (ARIMA)*, Model *Integer-valued Autoregressive Integrated Moving Average (INARIMA)*, Metode *Conditional Least Square (CLS)*, *Self Organizing Map (SOM)*, *Neural Network (NN)*, *Radial Basis Function Neural Network (RBFNN)*, dan Demam Berdarah (DB).

A. Distribusi *Poisson*

Distribusi *Poisson* adalah distribusi peluang diskrit yang menyatakan peluang banyaknya peristiwa yang terjadi pada waktu tertentu jika rata-rata kejadian tersebut diketahui. Variabel acak pada distribusi *Poisson* berupa bilangan bulat positif sehingga menjadi pilihan yang baik untuk pemodelan data cacah atau data *count*.

Suatu variabel acak Y dikatakan berdistribusi *Poisson* dengan parameter $\lambda > 0$ yang dinotasikan dengan $Y \sim P(\lambda)$ jika nilai y adalah bilangan bulat positif $0, 1, 2, \dots$ dan fungsi densitas peluangnya adalah (Bartosznski dan Bugaj, 2007: 141):

$$f(y) = \frac{\lambda^y}{y!} e^{-\lambda} \quad , y = 0, 1, 2, \dots \quad (2.1)$$

dengan $E(Y) = \text{var}(Y) = \lambda$.

Model regresi *Poisson* adalah model regresi untuk data *count*. Misal diketahui variabel dependen $Y_t \sim P(\lambda_t)$ dan vektor variabel independen \mathbf{x}_t dengan $\mathbf{x}_t^T = [x_{1t}, x_{2t}, \dots, x_{kt}]$, maka diperoleh (Cameron dan Trivedi, 1998: 9):

$$\log(\lambda_t) = \mathbf{x}_t^T \quad \text{atau} \quad \lambda_t = \exp(\mathbf{x}_t^T)$$

B. Prediksi (*Forecast*)

Prediksi (*forecast*) adalah teknik untuk memprediksi keadaan di masa yang akan datang dengan tetap memperhatikan keadaan di masa lalu. Berdasarkan periode waktunya, prediksi dibagi menjadi 3 bentuk (Montgomery *et al.*, 2008: 1):

1. Jangka Pendek (*Short Term*)

Jangka pendek meliputi beberapa kurun waktu (hari, minggu, bulan). Data historis terdahulu masih relevan digunakan karena prediksinya untuk jangka waktu singkat. Contoh prediksi dalam jangka pendek adalah memprediksi hasil penjualan produksi.

2. Jangka Menengah (*Medium Term*)

Jangka menengah meliputi kurun waktu satu musim (triwulan atau kuartal) sampai dua tahun. Data historis masa lalu masih dianggap relevan untuk memprediksi masa depan. Contoh prediksi jangka menengah adalah memprediksi anggaran penjualan atau produksi.

3. Jangka Panjang (*Long Term*)

Jangka panjang meliputi kurun waktu lebih dari dua tahun. Prediksi untuk jangka panjang pada umumnya dilakukan berdasarkan intuisi dan pengalaman seseorang, tapi banyak perusahaan juga menggunakan data historis.

C. Analisis Deret Waktu (*Time Series Analysis*)

Data deret waktu (*time series*) adalah serangkaian data yang dikumpulkan, direkam atau diamati terhadap suatu peristiwa, kejadian, gejala atau perubahan yang diamati dari waktu ke waktu. Deret waktu digunakan untuk memperoleh gambaran dari keadaan atau sifat variabel di masa lalu untuk memprediksi keadaan atau sifat variabel tersebut pada periode yang akan datang. Beberapa komponen deret waktu yang perlu diperhatikan adalah sebagai berikut (Makridakis *et al.*, 1997: 25):

1. Gerakan Horizontal

Gerakan horizontal adalah suatu pergerakan data yang berfluktuasi di sekitaran nilai konstan atau rata-rata. Pola ini juga disebut stasioner terhadap rata-rata (*mean*).

2. Gerakan Musiman

Gerakan musiman (*seasonal*) merupakan gerakan berulang-ulang secara teratur selama kurang lebih satu tahun, misalnya pola yang berulang setiap minggu, bulan atau kuartalan (per 4 bulan).

3. Gerakan Siklis

Pola gerakan siklis merupakan gerakan naik atau turun pada periode yang tidak tetap. Data yang sering menunjukkan pola gerakan siklis adalah data perdagangan, industri dan keuangan.

4. Gerakan *Trend*

Pola gerakan *trend* adalah gerakan data yang cenderung menuju ke satu arah baik naik maupun turun.

D. Stasioneritas

Stasioner adalah keadaan dimana nilai rata-rata (*mean*) dan varians data observasi tidak mengalami perubahan atau konstan dari waktu ke waktu. Secara teoritis dapat dituliskan sebagai berikut (Palit dan Papovic, 2005: 18):

1. $E(Y_t) = \mu$

2. $Var(Y_t) = E\{(Y_t - \mu)^2\} = \sigma_0$

3. $Cov(Y_t, Y_{t+k}) = E\{(Y_t - \mu)(Y_{t+k} - \mu)\} = \sigma_k$

dengan

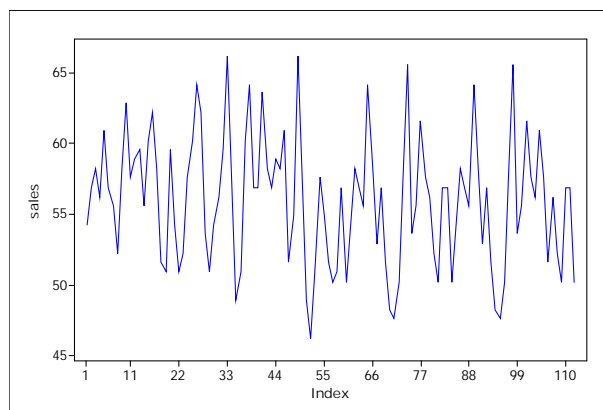
$$t = 1, 2, \dots, n$$

$$k = \dots, -2, -1, 0, 1, 2, \dots$$

n = banyaknya observasi

μ, σ_0 , dan σ_k adalah konstan.

Penentuan suatu data stasioner atau nonstasioner dapat dilihat pada pola *plot* data. Jika *plot* data cenderung konstan dimana tidak terjadi kenaikan atau penurunan dan atau sebaran data berada di sekitar garis rata-rata (*mean*), maka dapat dikatakan bahwa data sudah stasioner. Contoh data stasioner dapat dilihat pada Gambar 2.1 berikut:



(Sumber: Hanke dan Wichern, 2005: 434)

Gambar 2.1. Pola Data Stasioner

Selain *plot* data, *plot* autokorelasi juga dapat dijadikan alternatif untuk melihat kestasioneran data. Data dikatakan stasioner, jika *lag* setelah *lag* kedua atau ketiga menurun dengan cepat menuju nol. Sedangkan jika data tidak stasioner, *lag-lag* cenderung lambat menuju nol dalam kurun waktu tertentu. Uji Statistik yang digunakan untuk melihat kestasioneran data dalam *mean* adalah Uji Dickey-Fuller (*Dickey-Fuller Test*). Hipotesis Uji Dickey-Fuller adalah (Wei, 2006: 191):

$$H_0 : w_1 = 1 \text{ (Data tidak stasioner)}$$

$$H_1 : w_1 < 1 \text{ (Data stasioner)}$$

dengan statistik uji:

$$T = \frac{\hat{w} - 1}{SE(\hat{w})} \quad (2.2)$$

dimana

$$\hat{W} = \frac{\sum_{t=1}^n Z_{t-1} Z_t}{\sum_{t=1}^n Z_{t-1}^2} \quad (2.3)$$

Data yang tidak stasioner dapat diatasi dengan beberapa cara, misalnya tidak stasioner dalam rata-rata (*mean*) diatasi dengan melakukan *differencing*. Jika data tidak stasioner dalam varians dapat diatasi dengan melakukan Transformasi *Box-Cox* dengan persamaan sebagai berikut (Wei, 2006: 85):

$$T(Z_t) = \begin{cases} \frac{Z_t^\lambda - 1}{\lambda}, & \lambda \neq 0 \\ \ln Z_t, & \lambda = 0 \end{cases} \quad (2.4)$$

E. Kriteria Keباikan Model

Pada dasarnya, nilai prediksi tidak sama dengan nilai sebenarnya, sehingga memungkinkan adanya suatu kesalahan. Pada proses prediksi itu sendiri diupayakan untuk menghasilkan nilai prediksi yang mendekati nilai sebenarnya sehingga meminimumkan kesalahan (*error*). Ada beberapa kriteria yang dapat digunakan untuk mengukur kesalahan prediksi, antara lain adalah *Mean Square Error (MSE)* dan *Root Mean Square Error (RMSE)* yang dirumuskan sebagai berikut (Hanke dan Wichern, 2005: 80):

1. Mean Square Error (MSE)

Nilai *MSE* digunakan untuk mengukur ketepatan nilai dugaan model yang dinyatakan dalam rata-rata kuadrat kesalahan. Rumus *MSE* adalah:

$$MSE = \frac{1}{n} \sum_{t=1}^n (Y_t - \hat{Y}_t)^2 \quad (2.5)$$

2. Root Mean Square Error (RMSE)

Nilai *RMSE* adalah nilai akar kuadrat dari *MSE*. Rumus *RMSE* adalah:

$$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{n} \sum_{t=1}^n (Y_t - \hat{Y}_t)^2} \quad (2.6)$$

dengan

Y_t = Nilai observasi ke- t

\hat{Y}_t = Nilai prediksi observasi ke- t

n = banyak observasi

F. Autokorelasi

Autokorelasi (*Autocorrelation*) digunakan untuk menentukan koefisien korelasi pada data deret waktu dan untuk mempelajari pola data seperti *trend* atau musiman. Autokorelasi dibagi menjadi dua fungsi yaitu fungsi autokorelasi dan fungsi autokorelasi parsial.

1. Fungsi autokorelasi

Fungsi autokorelasi atau *Autocorrelation Function (ACF)* adalah semua himpunan autokorelasi untuk semua *lag-k* yang disimbolkan dengan \dots_k dengan $k = 1, 2, 3, \dots$ dan $\dots_0 = 0$.

Diketahui suatu deret $\{Y_t\}$ yaitu $Y_1, Y_2, Y_3, \dots, Y_N$ adalah stasioner, maka korelasi antara observasi Y_t dengan Y_{t+k} adalah (Wei, 2006: 10):

$$\dots_k = \frac{\text{Cov}(Y_t, Y_{t+k})}{\sqrt{\text{Var}(Y_t)\text{Var}(Y_{t+k})}} = \frac{\text{Cov}(Y_t, Y_{t+k})}{\text{Var}(Y_t)} = \frac{\chi_k}{\chi_0} \quad (2.7)$$

dimana $\text{Var}(Y_{t+k}) = \text{Var}(Y_t) = \chi_0$

dengan

\dots_k = autokorelasi pada *lag-k*

\sim = rata-rata (*mean*)

χ_k = autokovarians pada *lag-k*

t = waktu observasi, $t = 1, 2, 3, \dots, n$

n = banyaknya observasi

χ_k disebut fungsi autokovarians dan \dots_k disebut fungsi autokorelasi atau *Autocorrelation Function (ACF)* karena keduanya merepresentasikan kovarians dan korelasi antara Y_t dan Y_{t+k} pada *lag-k*. Jadi, rumus *ACF* sampel pada *lag-k* dituliskan sebagai berikut (Wei, 2006: 20):

$$\hat{\dots}_k = r_k = \frac{\sum_{t=1}^{n-k} (Y_t - \bar{Y})(Y_{t+k} - \bar{Y})}{\sum_{t=1}^n (Y_t - \bar{Y})^2}, \quad k = 0, 1, 2, \dots \quad (2.8)$$

dengan

r_k = koefisien autokorelasi sampel pada *lag-k*

\bar{Y} = rata-rata observasi $\{Y_t\}$

Y_t = observasi pada waktu ke- t

Y_{t+k} = observasi pada waktu ke- $t+k$

Menurut Hanke dan Wichern (2005), pengujian signifikansi autokorelasi bertujuan untuk mengetahui apakah autokorelasi berbeda signifikan dengan nol. Hipotesis yang digunakan adalah:

$H_0 : \dots_k = 0$ (autokorelasi pada *lag-k* tidak berbeda signifikan dengan nol)

$H_1 : \dots_k \neq 0$ (autokorelasi pada *lag-k* berbeda signifikan dengan nol)

dan statistik uji menggunakan uji t yang dirumuskan sebagai berikut:

$$t_{hitung} = \frac{r_k}{SE(r_k)} \quad (2.9)$$

$$SE(r_k) = \sqrt{\frac{1 + 2 \sum_{i=1}^{k-1} r_i^2}{n}} \quad (2.10)$$

dengan

$SE(r_k)$ = *standar error* autokorelasi sampel pada *lag-k*

r_i = koefisien autokorelasi sampel pada *lag-i*

n = banyaknya observasi

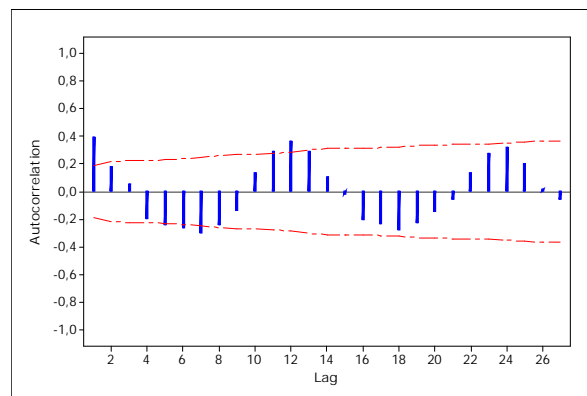
Kriteria keputusan yang digunakan adalah H_0 ditolak jika:

$$t_{hitung} < -t_{n-k; \left(\frac{\alpha}{2}\right)} \text{ atau } t_{hitung} > t_{n-k; \left(\frac{\alpha}{2}\right)}$$

Signifikansi autokorelasi sampel juga dapat dilihat dari *correlogram*. *Correlogram* adalah *plot lag-k* dengan r_k yang memiliki dua komponen utama, yaitu selang kepercayaan dan *lag* data. Selang kepercayaan direpresentasikan dengan garis putus-putus berwarna merah, sedangkan *lag* data direpresentasikan dengan garis tegak berwarna biru. Selang kepercayaan dapat ditentukan dengan rumus (Hanke dan Wichern, 2005: 64):

$$r_k \pm t_{n-k; \left(\frac{\alpha}{2}\right)} \times SE(r_k) \quad (2.11)$$

Contoh *correlogram ACF* sampel dapat dilihat pada Gambar 2.2 berikut.



Gambar 2.2. *Correlogram* untuk ACF

2. Fungsi Autokorelasi Parsial Sampel

Fungsi autokorelasi parsial sampel atau *Partial Autocorrelation Function (PACF)* sampel adalah korelasi antara Y_t dengan Y_{t+k} setelah pengaruh $Y_{t+1}, Y_{t+2}, \dots, Y_{t+k-1}$ dihilangkan terhadap Y_{t+k} . Autokorelasi parsial sampel dinotasikan dengan \hat{w}_{kk} dengan rumus (Wei., 2006: 14):

$$\hat{W}_{kk} = \frac{\begin{vmatrix} 1 & \dots_1 & \dots_2 & \dots & \dots_1 \\ \dots_1 & 1 & \dots_1 & \dots & \dots_2 \\ \dots_2 & \dots_1 & 1 & \dots & \dots_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \dots_{k-1} & \dots_{k-2} & \dots_{k-3} & \dots & \dots_k \end{vmatrix}}{\begin{vmatrix} 1 & \dots_1 & \dots_2 & \dots & \dots_{k-1} \\ \dots_1 & 1 & \dots_1 & \dots & \dots_{k-2} \\ \dots_2 & \dots_1 & 1 & \dots & \dots_{k-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \dots_{k-1} & \dots_{k-2} & \dots_{k-3} & \dots & 1 \end{vmatrix}} \quad (2.12)$$

untuk $k = 1, 2, \dots$ diperoleh:

$$\hat{W}_{11} = \frac{|\dots_1|}{|P_1|} = \frac{|\dots_1|}{1} = \dots_1$$

$$\hat{W}_{22} = \frac{|\dots_2|}{|P_2|} = \frac{\begin{vmatrix} 1 & \dots_1 \\ \dots_1 & \dots_2 \end{vmatrix}}{\begin{vmatrix} 1 & \dots_1 \\ \dots_1 & 1 \end{vmatrix}} = \frac{\dots_2 - \dots_1^2}{1 - \dots_1^2}$$

dan seterusnya sampai \hat{W}_{kk} .

Pengujian signifikansi autokorelasi parsial menggunakan hipotesis:

$H_0 : w_{kk} = 0$ (autokorelasi parsial pada $lag-k$ tidak berbeda signifikan dengan nol)

$H_1 : w_{kk} \neq 0$ (autokorelasi parsial pada $lag-k$ berbeda signifikan dengan nol)

dan statistik uji menggunakan uji t yang dirumuskan sebagai berikut:

$$t_{hitung} = \frac{\hat{W}_{kk}}{SE(\hat{W}_{kk})} \quad (2.13)$$

$$SE(\hat{W}_{kk}) = \sqrt{\frac{1}{n}} \quad (2.14)$$

dengan

$SE(\hat{w}_{kk})$ = standar error autokorelasi parsial sampel pada lag-k

\hat{w}_{kk} = koefisien autokorelasi parsial sampel pada lag-k

n = banyaknya observasi

Kriteria keputusan yang digunakan adalah H_0 ditolak jika:

$$t_{hitung} < -t_{n-k; \left(\frac{\alpha}{2}\right)} \text{ atau } t_{hitung} > t_{n-k; \left(\frac{\alpha}{2}\right)}$$

Selang kepercayaan pada *correlogram PACF* sampel dapat ditentukan dengan rumus:

$$\hat{w}_{kk} \pm t_{n-k; \left(\frac{\alpha}{2}\right)} \times SE(\hat{w}_{kk}) \quad (2.15)$$

G. Model Hybrid

Pada keadaan sebenarnya, sering ditemukan data deret waktu yang tidak hanya memiliki struktur autokorelasi linier, namun juga memiliki struktur autokorelasi nonlinier. Berdasarkan hal tersebut, Zhang (2003) merekomendasikan model *Hybrid* yaitu kombinasi dua atau lebih sistem dalam satu fungsi yang mempertimbangkan struktur autokorelasi linier dan nonlinier. Pada penelitian ini, digunakan model *Hybrid INARIMA-RBFNN* yang merupakan gabungan Model *INARIMA* dan *RBFNN*. Secara umum, kombinasi dari model *time series* yang memiliki struktur autokorelasi linier dan nonlinier dapat dituliskan sebagai berikut (Zhang, 2003: 165):

$$y_t = L_t + N_t \quad (2.16)$$

dimana L_t adalah komponen linier dan N_t adalah komponen nonlinier. *INARIMA* digunakan untuk memodelkan komponen linier yang akan menghasilkan *residual* yang masih mengandung struktur nonlinier. Nilai *residual* dari model linier dirumuskan sebagai berikut (Zhang, 2003: 165):

$$e_t = y_t - \hat{L}_t \quad (2.17)$$

dengan

$e_t = residual$ pada waktu ke- t

$y_t = data$ observasi pada waktu ke- t

$\hat{L}_t = nilai$ prediksi data observasi pada waktu ke- t

Suatu model belum dapat dikatakan linier jika masih terdapat struktur nonlinier yang dapat dilihat dari nilai *residual* (Zhang, 2003). Nilai *residual* ini kemudian akan di analisis menggunakan *RBFNN*. Diketahui n *neuron input* dari *residual* dan dituliskan dengan:

$$e_t = f(e_{t-1}, e_{t-2}, \dots, e_{t-n}) + v_t \quad (2.18)$$

dengan

$f = fungsi$ nonlinier *neural network*

$e_t = residual$ pada waktu ke- t

$e_{t-1} = residual$ pada waktu ke- $t-1$

Hasil prediksi pada persamaan (2.16) adalah \hat{N}_t , maka hasil prediksi *Hybrid* dapat dituliskan sebagai berikut (Zhang, 2003: 165):

$$\hat{y}_t = \hat{L}_t + \hat{N}_t \quad (2.19)$$

H. Model *Autoregressive Integrated Moving Average (ARIMA)*

Pada model umum *ARIMA* nilai variabel yang akan datang diasumsikan merupakan fungsi linier dari beberapa observasi sebelumnya. Model *ARIMA* adalah (Wei, 2006: 72):

$$w_p(B)(1-B)^d Z_t = {}_n q(B)a_t \quad (2.20)$$

dengan $w_p(B) = (1 - w_1 B - w_2 B^2 - \dots - w_p B^p)$ merupakan model stasioner *Autoregressive (AR)* dan ${}_n q(B) = (1 - {}_n 1 B - {}_n 2 B^2 - \dots - {}_n q B^q)$ merupakan model *invertible Moving Average*.

Menurut Montgomery *et al.* (2008) terdapat 3 tahapan untuk menentukan model *ARIMA*, yaitu:

1. Identifikasi model *ARIMA* menggunakan data periode sebelumnya.
2. Estimasi parameter yang nilainya belum diketahui.
3. Melakukan uji diagnostik *residual* yang dihasilkan untuk memenuhi kriteria kebaikan model.

I. Model *Integer-valued Autoregressive Integrated Moving Average (INARIMA)*

Menurut Cameron dan Trivedi (1998), pada model *Integer-valued Autoregressive Integrated Moving Average (INARIMA)*, nilai Y_t merupakan variabel acak hasil penjumlahan atau data *count* yang dependen dengan Y_t pada waktu sebelumnya (Y_{t-k}). Pada proses *INARIMA*, terdapat *binomial thinning operation* yang didefinisikan sebagai jumlahan dari variabel acak *Bernoulli* yang dirumuskan sebagai berikut:

$$r \circ Y = \sum_{j=1}^Y U_j \quad (2.21)$$

dimana Y adalah variabel acak bilangan bulat positif dan $r \in (0,1)$. $\{U_j\} = U_1, U_2, \dots, U_j$ dengan $j=1, 2, \dots, Y$ adalah variabel acak berdistribusi *Bernoulli*, $P(U_j = 1) = r, P(U_j = 0) = 1 - r$ yang memenuhi *iid (identically independent distributed)*. $\{U_j\}$ independen dengan Y dan mempunyai *mean* r dan *varians* $r \dagger^2$.

1. Model *Integer-valued Autoregressive INAR(p)*

Diberikan proses stokastik diskrit bilangan bulat positif $\{Y_t\}$ mengikuti model *INAR(p)*, maka proses tersebut akan memenuhi persamaan (Silva *et al.*, 2005: 519):

$$Y_t = r_1 \circ Y_{t-1} + r_2 \circ Y_{t-2} + \dots + r_p \circ Y_{t-p} + Z_t \quad (2.22)$$

dengan $\{Z_t\}$ adalah barisan variabel acak bilangan bulat positif yang memenuhi *iid* dan $E(Z_t) = \lambda_t$. Bentuk $r_i \circ Y_{t-i}$ menjelaskan bahwa Y_{t-i} adalah banyaknya percobaan dengan peluang sukses sebesar r_i untuk setiap percobaan.

Model *INAR* yang melibatkan variabel prediktor dan mengasumsikan $\{Y_t\}$ sebagai variabel respon. Diberikan r variabel prediktor pada waktu ke- t adalah $\mathbf{x}_t = (x_{t,0}, x_{t,1}, \dots, x_{t,r})$, dengan $x_{t,0} = 1$. Nilai p adalah order model *INAR*. Jika diketahui $\mathbf{x}_i = (x_{i,0}, x_{i,1}, \dots, x_{i,r})$ dan $\mathbf{u} = (u_0, u_1, \dots, u_r)$, maka model *INAR*(p) dengan variabel prediktor adalah sebagai berikut:

$$Y_t = \sum_{i=1}^p r_i \circ Y_{t-i} + Z_t \quad (2.23)$$

dengan $E(Z_t) = \lambda_t$ dan $\lambda_t = \exp(\mathbf{x}_t^T \mathbf{w}_t)$. Nilai $0 < r_i < 1$, sehingga dapat dituliskan dalam bentuk fungsi logistik sebagai berikut (Cameron dan Trivedi, 1998: 237):

$$r_i = \frac{1}{1 + \exp(-\mathbf{w}_t^T \mathbf{x}_i)} \quad (2.24)$$

dengan $\sum_{i=1}^p r_i < 1$.

2. Model *Integer-valued Moving Average INMA*(q)

Persamaan *INMA*(q) didefinisikan sebagai berikut:

$$Y_t = s_0 \circ Z_t + s_1 \circ Z_{t-1} + \dots + s_q \circ Z_{t-q} \quad (2.25)$$

dimana Z_t adalah variabel acak bilangan bulat positif yang memenuhi *iid* dengan $s_0 = 1$, dan $s_j = (0,1)$, $j = 1, 2, \dots, q$.

3. Model *Integer-valued Autoregressive Moving Average INARMA(p,q)*

Persamaan umum *INARMA(p,q)* adalah sebagai berikut:

$$Y_t = \sum_{i=1}^p r_i \circ Y_{t-i} + Z_t + \sum_{j=1}^q s_j \circ Z_{t-j} \quad (2.26)$$

dimana $r_1, r_2, \dots, r_{(p)} \in (0,1)$ dan $s_1, s_2, \dots, s_{(q)} \in (0,1)$, sedangkan $\{Z_t\}$ adalah variabel acak bilangan bulat positif yang memenuhi *iid* yang independen dengan Y_t . Mekanisme *binomial thinning* untuk *INAR(p)* dan *INMA(q)* adalah berurutan, sehingga diasumsikan bahwa operasi *thinning* pada $r_i \circ Y_{t-1}$ untuk $i=1,2,\dots,p$ dan $s_j \circ Z_{t-j}$ untuk $j=1,2,\dots,q$ adalah saling independen.

J. Conditional Least Square (CLS)

Conditional Least Square (CLS) adalah salah satu metode estimasi parameter yang dikenalkan pertama kali oleh Klimko dan Nelson (1987). Diberikan ekspektasi bersyarat untuk Y_t dengan syarat Y_1, Y_2, \dots, Y_{t-1} yang dapat dituliskan sebagai berikut (Turkman *et al.*, 2014: 218):

$$E(Y_t | Y_1, Y_2, \dots, Y_{t-1}) = E(Y_t | \mathcal{F}_{t-1}) \quad (2.27)$$

Prinsip dasar metode *CLS* adalah untuk meminimumkan *conditional least of squares*:

$$Q(\) = \sum_{t=1}^n (Y_t - E(Y_t | \mathcal{F}_{t-1}))^2 \quad (2.28)$$

dengan

$$\begin{aligned} Q(\) &= \text{conditional least of squares} \\ &= \text{Vektor parameter yang belum diketahui} \end{aligned}$$

Sebagai ilustrasi, diketahui model *INAR(1)* adalah sebagai berikut:

$$Y_t = r_1 \circ Y_{t-1} + Z_t$$

dimana $\{Z_t\}$ mempunyai *mean* β_t , sehingga diperoleh:

$$E(Y_t) = r_1 E(Y_{t-1}) + \beta_t \quad (2.29)$$

Kemudian, ekspektasi bersyarat dari Y_t dengan syarat Y_{t-1}, Z_t dituliskan sebagai berikut:

$$E(Y_t | Y_{t-1}, Z_t) = r_1 Y_{t-1} + \beta_t \quad (2.30)$$

Berdasarkan persamaan (2.28), metode *CLS* akan meminimumkan *conditional least of squares*:

$$\begin{aligned} Q(r_1, \beta) &= \sum_{t=1}^n (Y_t - E(Y_t | Y_{t-1}, Z_t))^2 \\ &= \sum_{t=1}^n (Y_t - (r_1 Y_{t-1} + \beta_t))^2 \end{aligned}$$

K. *Neural Network (NN)*

Neural Network (NN) atau Jaringan Saraf Tiruan (JST) adalah sistem pemrosesan informasi yang memiliki karakteristik mirip jaringan saraf biologi, khususnya otak manusia. *NN* dibentuk sebagai generalisasi model matematika dari jaringan saraf biologi dengan asumsi (Fausett, 1994: 3):

- Pemrosesan informasi terjadi pada banyak elemen sederhana (*neuron*).
- Sinyal dikirim diantara *neuron-neuron* melalui penghubung-penghubung.
- Penghubung antara *neuron* memiliki bobot yang akan memperkuat atau memperlemah sinyal.
- Untuk menentukan *output*, setiap *neuron* menggunakan fungsi aktivasi yang dikenakan pada banyaknya *input* yang diterima.

1. Kerangka *Neural Network*

Pada *Neural Network*, *neuron-neuron* akan dikumpulkan dalam lapisan-lapisan (*layer*) yang disebut lapisan *neuron (neuron layer)*.

Pada umumnya, *NN* memiliki tiga lapisan, yaitu (Fausett, 1994: 12):

a. Lapisan *Input* (*Input Layer*)

Pada lapisan ini, terdapat node-node yang disebut *neuron input*. *Neuron-neuron input* menerima *input* berupa informasi dari luar.

b. Lapisan Tersembunyi (*Hidden Layer*)

Lapisan tersembunyi merupakan lapisan yang terletak antara lapisan *input* dan lapisan *output*. Node-node pada lapisan tersembunyi disebut *neuron* tersembunyi.

c. Lapisan *Output* (*Output Layer*)

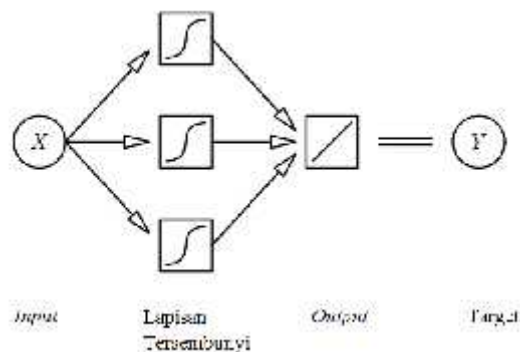
Node-node pada lapisan *output* disebut *neuron output*. Keluaran dari lapisan ini merupakan hasil dari *NN* terhadap suatu informasi.

2. Arsitektur *Neural Network*

Arsitektur jaringan saraf adalah pengaturan *neuron* dalam setiap lapisan dan pola hubungan antar lapisan. Terdapat tiga arsitektur *NN*, yaitu Jaringan Lapisan Tunggal, Jaringan Banyak Lapisan, dan Jaringan Kompetitif (Fausett, 1994: 12-14). Pada penelitian ini digunakan arsitektur jaringan banyak lapisan dan jaringan kompetitif.

a. Jaringan Banyak Lapisan (*Multilayer Net*)

Jaringan dengan banyak lapisan memiliki lapisan *input*, satu atau lebih lapisan tersembunyi dan lapisan *output*. Gambar 2.3 berikut adalah contoh jaringan banyak lapisan dengan fungsi aktivasi *Sigmoid* pada lapisan tersembunyi dan fungsi aktivasi Linier pada lapisan *output*.

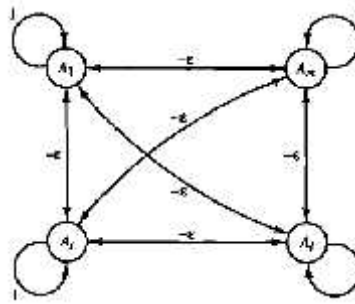


(Sumber: Sarle (1994))

Gambar 2.3. Jaringan Banyak Lapisan

b. Jaringan Lapisan Kompetitif (*Competitive Layer Net*)

Pada jaringan ini, setiap *neuron* dapat dihubungkan dengan dirinya sendiri dan *neuron* yang lain. Contoh lapisan kompetitif dapat dilihat pada Gambar 2.4 berikut.



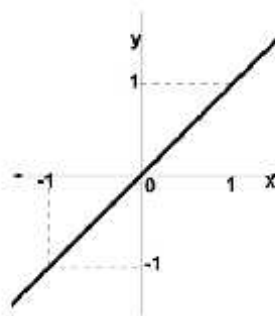
(Sumber: Fausett, 1994: 14)

Gambar 2.4. Jaringan Lapisan Kompetitif

3. Fungsi Aktivasi

Fungsi aktivasi adalah fungsi yang digunakan untuk menentukan keluaran (*output*) suatu *neuron*. Fungsi aktivasi yang digunakan pada penelitian ini adalah fungsi aktivasi Linier (Identitas) atau disebut juga *Purelin*. Fungsi Linier memiliki nilai *output* yang sama dengan nilai *input*-nya. Fungsi Linier dirumuskan sebagai berikut:

$$y = x \text{ untuk semua } x$$



(Sumber: Sri Kusumadewi, 2004, 54)

Gambar 2.5. Fungsi Aktivasi Linier

4. Metode Pembelajaran dan Pelatihan

Tujuan utama proses pembelajaran adalah untuk mengatur bobot-bobot yang ada pada jaringan saraf, sehingga diperoleh bobot akhir yang tepat sesuai dengan pola data yang dikenai perlakuan atau pelatihan. Pada *NN* terdapat dua metode pembelajaran, yaitu (Yeung *et al.*, 1998: 5):

a. Pembelajaran Terawasi (*Supervised Learning*)

Pada proses pembelajaran ini, satu informasi yang diberikan pada *neuron input* akan dijalankan sampai *neuron output*. Hasil *output* akan dicocokkan dengan target, jika terdapat perbedaan, maka akan muncul *error*. Jika nilai *error* cukup besar, akan dilakukan pembelajaran yang lebih banyak lagi.

b. Pembelajaran Tidak Terawasi (*Unsupervised Learning*)

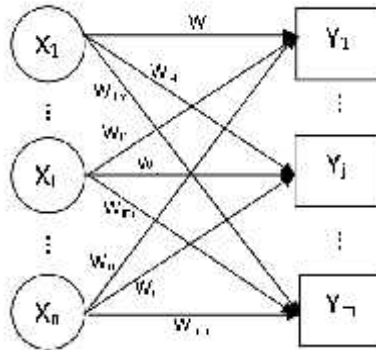
Pada proses pembelajaran ini, nilai bobot disusun dalam satu interval atau *range* tertentu tergantung dari nilai *input* yang diberikan. Pembelajaran ini bertujuan mengelompokkan unit-unit yang hampir sama dalam satu area tertentu.

L. *Self Organizing Map (SOM)*

Pada jaringan *SOM*, suatu lapisan yang berisi *neuron-neuron* akan menyusun dirinya sendiri dalam suatu kelompok. Pada proses pengelompokkan, vektor *input* yang memiliki jarak paling pendek dengan bobot akan menjadi pemenang dan diperbaharui bobotnya. Metode pembelajaran pada *SOM* adalah *unsupervised learning*, dimana tidak ada target kelompok yang ditetapkan untuk setiap *input*.

Menurut Kusumadewi (2004), salah satu algoritma pembelajaran untuk *SOM* adalah algoritma pembelajaran *Kohonen*. Pembelajaran kompetitif dengan Metode *Kohonen* diawali dengan memilih secara acak suatu vektor *input*. Jaringan *SOM Kohonen* terdiri dari dua lapisan (*layer*), yaitu lapisan *input* dan lapisan *output*. Setiap *neuron output* dihubungkan dengan *neuron input* dengan bobot, sehingga vektor bobot memiliki dimensi yang sama dengan vektor *input*.

Arsitektur *SOM* dapat dilihat pada Gambar 2.6 berikut:



(Sumber: Sarle(1994))

Gambar 2.6. Arsitektur SOM

Berdasarkan Gambar 2.6, arsitektur *SOM* memiliki n *neuron input* dan m *neuron output*. Bobot w_{ij} adalah bobot yang menghubungkan *neuron input* ke- i dengan *neuron output* ke- j . Pada proses pembelajaran, terdapat laju pembelajaran yang disebut *learning rate* (τ) yang ditentukan oleh peneliti. Jika *learning rate* ditentukan terlalu besar, maka algoritma akan tidak stabil. Sebaliknya, jika *learning rate* ditentukan terlalu kecil, maka algoritma akan konvergen dalam jangka waktu yang lama. Algoritma *SOM* adalah sebagai berikut (Fausett, 1994: 170):

Langkah 0. Inisialisasi bobot w_{ij} .

Tentukan parameter dan *learning rate* (τ).

Langkah 1. Jika kondisi berhenti salah, lakukan langkah 2-8

Langkah 2. Untuk setiap *input* x_i lakukan langkah 3-5

Langkah 3. Untuk setiap j (*output*), akan dihitung:

$$D(j) = \sum_{i=1}^n (w_{ij} - x_i)^2 \quad (2.31)$$

Langkah 4. Tentukan indeks J sedemikian sehingga $D(J)$ minimum.

Langkah 5. Update bobot pemenang (J)

$$w_{ij}(\text{baru}) = w_{ij}(\text{lama}) + \tau [x_i - w_{ij}(\text{lama})] \quad (2.32)$$

Langkah 6. Update *learning rate*.

Langkah 7. Mengurangi (*reduce*) jarak antar *neuron input*.

Langkah 8. Uji kondisi berhenti.

M. *Radial Basis Function Neural Network (RBFNN)*

Model *RBFNN* terdiri dari 3 lapisan, yaitu lapisan *input (input layer)*, lapisan tersembunyi (*hidden layer*) dan lapisan *output (output layer)*. Pada lapisan *input* terdapat *input* yang kemudian dibawa ke lapisan tersembunyi yang akan memproses data *input* secara nonlinier dengan fungsi aktivasi. *Output* dari lapisan tersembunyi selanjutnya diproses di lapisan *output* secara linier.

Model *RBFNN* menggunakan fungsi basis sebagai fungsi aktivasi untuk setiap *neuron* pada lapisan tersembunyi. Beberapa fungsi basis adalah sebagai berikut (Hanrahan, 2011: 10):

- Fungsi *Gaussian*

$$w(x) = \exp\left(\frac{-x^2}{2\tau_{ij}^2}\right) \quad (2.33)$$

- Fungsi *Multikuadratik*

$$w(x) = \left(x^2 + \tau_{ij}^2\right)^{\frac{1}{2}} \quad (2.34)$$

- Fungsi *Invers Multikuadratik*

$$w(x) = \frac{1}{\left(x^2 + \tau_{ij}^2\right)^{\frac{1}{2}}} \quad (2.35)$$

dengan x adalah jarak (*norm*) *Euclidean* antara *input* $x_i, i = 1, 2, \dots, p$ dengan pusat *neuron* tersembunyi $\sim_{ij}, j = 1, 2, \dots, m$ yang dirumuskan dengan:

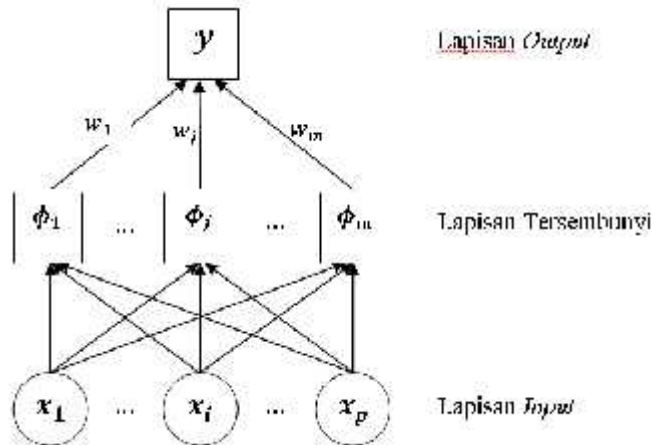
$$x = \|x_i - \sim_{ij}\|$$

$\tau_{ij}^2 = \text{varians neuron tersembunyi ke-}j$

Fungsi aktivasi yang sering digunakan dalam *RBFNN* adalah fungsi *Gaussian* pada lapisan tersembunyi dan fungsi aktivasi Linier pada lapisan *output*.

1. Arsitektur *RBFNN*

Arsitektur *RBFNN* dapat dilihat pada Gambar 2.7 berikut:



(Sumber: Orr, 1996: 10)

Gambar 2.7. Arsitektur *Radial Basis Function Neural Network*

Pada arsitektur *RBFNN* di atas, terdapat p input pada lapisan *input*, m fungsi basis sebagai fungsi aktivasi *neuron* pada lapisan tersembunyi dan 1 *neuron* pada lapisan *output*. *Output* y yang diperoleh dari model *RBFNN* merupakan hasil kombinasi linier dari bobot $\{w_j\}_{j=1}^m$ dengan fungsi aktivasi $w_j(x)$ dan dirumuskan sebagai berikut (Orr, 1996: 11):

$$y = \sum_{j=1}^m w_j w_j(x_i) \tag{2.36}$$

dengan

m = banyaknya fungsi aktivasi *neuron* tersembunyi

w_j = bobot *output* ke- j

$w_j(x_i)$ = fungsi aktivasi *neuron* pada lapisan tersembunyi ke- j

x_i = x_1, x_2, \dots, x_p yang merupakan *neuron input*

Berdasarkan fungsi basis *Gaussian*, diperoleh persamaan sebagai berikut:

$$\begin{aligned}
 w_j(x) &= \exp\left(-\sum_{i=1}^p \left(\frac{x_i - \tilde{ij}}{2\uparrow_{ij}}\right)^2\right) \\
 &= \exp\left(-\left\{\left(\frac{x_1 - \tilde{1,1}}{2\uparrow_{1,1}}\right)^2 + \dots + \left(\frac{x_i - \tilde{ij}}{2\uparrow_{ij}}\right)^2 + \dots + \left(\frac{x_p - \tilde{pm}}{2\uparrow_{pm}}\right)^2\right\}\right) \quad (2.37)
 \end{aligned}$$

dengan

x_i = variabel *input* ke- i , $i = 1, 2, \dots, p$

\tilde{ij} =pusat/rata-rata *neuron* pada lapisan tersembunyi ke- j *input* ke- i

\uparrow_{ij} =simpangan baku *neuron* pada lapisan tersembunyi ke- j *input* ke- i

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

2. Metode *Global Ridge Regression*

Metode *Global Ridge Regression* mengestimasi bobot dengan menambahkan parameter regulasi yang bernilai positif pada *Sum Square Error (SSE)* sehingga diperoleh persamaan (Orr, 1996: 24):

$$C = \sum_{i=1}^n (y_i - \hat{y}_i)^2 + \check{S} \sum_{j=1}^m w_j^2 \quad (2.38)$$

dengan

y_i = nilai observasi ke- i

\hat{y}_i = prediksi nilai observasi ke- i

S = parameter regulasi

w_j = bobot ke- j

n = banyaknya observasi

Nilai bobot optimum ditentukan dengan mendiferensialkan persamaan (2.38) menjadi:

$$\begin{aligned}
 \frac{\partial C}{\partial w_j} &= 2 \sum_{i=1}^n (y_i - \hat{y}_i) \frac{\partial y}{\partial w_j} + 2\tilde{S}w_j \\
 &= 2 \left(\sum_{i=1}^n y_i \frac{\partial y}{\partial w_j} - \sum_{i=1}^n \hat{y}_i \frac{\partial y}{\partial w_j} + \tilde{S}w_j \right) \\
 \sum_{i=1}^n y_i \frac{\partial y}{\partial w_j} - \sum_{i=1}^n \hat{y}_i \frac{\partial y}{\partial w_j} + \tilde{S}w_j &= 0 \\
 \sum_{i=1}^n y_i \frac{\partial y}{\partial w_j} + \tilde{S}w_j &= \sum_{i=1}^n \hat{y}_i \frac{\partial y}{\partial w_j} \tag{2.39}
 \end{aligned}$$

Persamaan (2.36) yang merupakan persamaan *output* dan didiferensialkan menjadi:

$$\frac{\partial y}{\partial w_j} = w_j(\mathbf{x})$$

Sehingga persamaan menjadi:

$$\sum_{i=1}^n y_i w_j(\mathbf{x}) + \tilde{S}w_j = \sum_{i=1}^n \hat{y}_i w_j(\mathbf{x}) \tag{2.40}$$

dan dalam notasi vektor menjadi:

$$\begin{aligned}
 \begin{bmatrix} w_1^T \mathbf{y} \\ w_2^T \mathbf{y} \\ \vdots \\ w_m^T \mathbf{y} \end{bmatrix} + \tilde{S} \begin{bmatrix} \hat{w}_1 \\ \hat{w}_2 \\ \vdots \\ \hat{w}_m \end{bmatrix} &= \begin{bmatrix} w_1^T \hat{\mathbf{y}} \\ w_2^T \hat{\mathbf{y}} \\ \vdots \\ w_m^T \hat{\mathbf{y}} \end{bmatrix} \\
 \Phi^T \mathbf{y} + \tilde{S} \hat{\mathbf{w}}_j &= \Phi^T \hat{\mathbf{y}} \tag{2.41}
 \end{aligned}$$

dengan

\mathbf{y} = vektor nilai *output*

$\hat{\mathbf{y}}$ = vektor prediksi nilai *output*

$\hat{\mathbf{w}}$ = vektor prediksi nilai bobot

S = parameter regulasi

Berdasarkan uraian optimasi bobot di atas, diperoleh persamaan:

$$\begin{aligned}\Phi^T \mathbf{y} &= \Phi^T \hat{\mathbf{y}} + \check{\mathbf{S}} \hat{\mathbf{w}} \\ &= \Phi^T \Phi \hat{\mathbf{w}} + \check{\mathbf{S}} \hat{\mathbf{w}} \\ &= (\Phi^T \Phi + \check{\mathbf{S}} \mathbf{I}_m) \hat{\mathbf{w}}\end{aligned}$$

dimana \mathbf{I}_m adalah matriks identitas berukuran $m \times n$. Jadi diperoleh persamaan untuk memprediksi bobot adalah sebagai berikut:

$$\hat{\mathbf{w}} = (\Phi^T \Phi + \check{\mathbf{S}} \mathbf{I}_m)^{-1} \Phi^T \hat{\mathbf{y}} \quad (2.42)$$

N. Demam Berdarah (DB)

1. Pengertian Demam Berdarah (DB)

Demam Berdarah (DB) adalah penyakit demam akut yang disebabkan oleh virus *Dengue* yang masuk ke peredaran darah manusia melalui gigitan nyamuk dari genus *Aedes aegypti* dan *Aedes albopictus*. *Aedes aegypti* adalah nyamuk pembawa virus *Dengue* yang paling banyak ditemukan menyebabkan penyakit demam berdarah (Kemenkes RI, 2014: 153). Penyakit DB dapat muncul sepanjang tahun dan dapat menyerang semua kelompok umur. Penyakit DB berkaitan dengan kondisi lingkungan dan perilaku masyarakat. Proses penularan DB adalah sebagai berikut:

- a. Demam berdarah ditularkan oleh nyamuk *Aedes aegypti* dan *Aedes albopictus* betina.
- b. Nyamuk ini mendapatkan virus *Dengue* sewaktu menggigit/menghisap darah penderita DB atau di dalam darahnya terdapat virus *Dengue* namun tidak menunjukkan gejala sakit.
- c. Virus *Dengue* yang terhisap akan berkembangbiak dan menyebar ke seluruh tubuh nyamuk, termasuk kelenjar liur.
- d. Jika nyamuk tersebut menggigit/menghisap darah orang lain, virus yang ada di dalam tubuh nyamuk akan dipindahkan bersama air liur nyamuk.

- e. Virus *Dengue* akan menyerang sel pembeku darah dan merusak dinding pembuluh darah kecil (*kapiler*), akibatnya terjadi pendarahan dan kekurangan cairan.

2. Faktor yang Mempengaruhi Penyebaran Demam Berdarah

Menurut Supartha (2008), beberapa faktor yang mempengaruhi penyebaran penyakit DB adalah faktor inang (*host*), lingkungan (*environment*) dan faktor patogen (virus). Faktor inang menyangkut kerentanan tubuh manusia dan imunitas terhadap penyakit, sedangkan faktor lingkungan menyangkut kondisi geografis (ketinggian dari permukaan laut, curah hujan, angin, kelembaban, musim), kondisi demografi (kepadatan, mobilitas, perilaku, adat istiadat, sosial ekonomi penduduk). Faktor patogen menyangkut jenis dan kepadatan nyamuk sebagai vektor penular penyakit.

Pada penelitian ini, faktor yang menjadi perhatian adalah faktor lingkungan secara geografis yang meliputi kelembaban, tekanan udara, temperatur, durasi penyinaran matahari, curah hujan, dan kecepatan angin.

a. Kelembaban Udara

Kelembaban udara menggambarkan kandungan uap air di udara. Kelembaban udara dinyatakan dalam persen (%). Kelembaban udara mempengaruhi umur nyamuk. Pada suhu 20°C kelembaban 27%, umur nyamuk betina adalah 101 hari dan umur nyamuk jantan adalah 35 hari. Pada kelembaban 55% umur nyamuk betina 88 hari dan nyamuk jantan 50 hari. Tingkat kelembaban 60% merupakan batas paling rendah untuk memungkinkan hidup nyamuk. Pada kelembaban kurang dari 60% umur nyamuk akan menjadi pendek, nyamuk tidak dapat menjadi vektor karena tidak cukup waktu untuk perpindahan virus dari lambung ke kelenjar liur. Pada kelembaban tinggi, nyamuk menjadi lebih aktif dan sering menggigit sehingga meningkatkan penularan demam berdarah.

b. Tekanan Udara

Tekanan udara adalah tenaga yang bekerja untuk menggerakkan massa udara dalam setiap satuan luas tertentu. Tekanan udara dipengaruhi oleh ketinggian suatu tempat. Tekanan udara akan berkurang dengan bertambahnya ketinggian suatu tempat. Alat yang digunakan untuk mengukur tekanan udara adalah *Barometer*. Satuan tekanan udara adalah *milibar (mb)*.

Tekanan udara erat kaitannya dengan temperatur, semakin rendah temperatur suatu tempat, maka tekanan udara akan semakin besar. Hal ini juga berpengaruh terhadap kelangsungan hidup nyamuk. Nyamuk tidak akan mampu bertahan di tempat dengan tekanan udara tinggi.

c. Temperatur

Temperatur atau suhu merupakan karakteristik yang dimiliki oleh suatu benda yang berhubungan dengan panas dan energi. Jika suatu benda dialiri panas, maka suhu benda tersebut akan meningkat, sebaliknya suhu benda tersebut akan menurun jika benda tersebut kehilangan panas. *Thermometer* adalah alat pengukur suhu, baik suhu udara maupun suhu air. Satuan umum yang digunakan adalah Derajat *Celsius (°C)*.

Nyamuk dapat bertahan hidup pada suhu rendah, namun metabolismenya menurun. Suhu rata-rata optimum untuk pertumbuhan nyamuk berkisar antara 25°C - 27°C . Pertumbuhan nyamuk akan terhenti sama sekali pada suhu kurang dari 10°C atau lebih dari 40°C . Sementara untuk pertumbuhan jentik memerlukan suhu antara 25°C - 30°C .

d. Durasi Penyinaran Matahari

Penyinaran matahari ke Bumi dipengaruhi oleh kondisi awan dan perbedaan sudut datang sinar matahari. Durasi penyinaran matahari berbeda-beda di setiap daerah. Perbedaan ini dipengaruhi oleh perbedaan luas permukaan yang mendapat penyinaran matahari, perbedaan panjang *atmosfer* yang dilalui sinar matahari, dan sebagainya.

Penyinaran matahari berpengaruh terhadap keseluruhan komponen iklim seperti temperatur, curah hujan, dan tekanan udara. Semakin lama durasi penyinaran matahari, maka temperatur akan naik, sedangkan curah hujan dan tekanan udara akan menurun. Nyamuk lebih menyukai tempat yang rindang dan tidak terkena sinar matahari langsung.

e. Curah Hujan

Curah hujan merupakan ketinggian air hujan yang terkumpul dalam tempat yang datar, tidak menguap, tidak meresap, dan tidak mengalir. Jumlah hujan dicatat dalam *inci* atau *milimeter (mm)*. Hujan dapat mempengaruhi kehidupan nyamuk dengan dua cara yaitu menyebabkan naiknya kelembaban udara dan menambah tempat perindukan. Setiap 1 mm curah hujan menambah kepadatan nyamuk 1 ekor, akan tetapi jika curah hujan dalam 1 minggu sebesar 140 mm, maka larva akan hanyut dan mati, sehingga kepadatan nyamuk berkurang.

Curah hujan yang tinggi akan menambah jumlah tempat perkembangbiakan nyamuk alami. Sampah-sampah kering seperti botol bekas, kaleng-kaleng, dan daun-daunan yang memungkinkan menampung air hujan merupakan tempat perindukan yang disukai untuk bertelurnya nyamuk *Aedes aegypti*.

f. Kecepatan Angin

Kecepatan angin adalah kecepatan udara yang bergerak secara horizontal pada ketinggian dua meter di atas permukaan tanah. Kecepatan angin akan berbeda pada permukaan yang tertutup oleh vegetasi pada ketinggian tertentu, misalnya tanaman padi, jagung dan kedelai. Kecepatan angin dapat diukur dengan alat yang disebut dengan *Anemometer*. Kecepatan angin diukur dalam satuan *knot* (*mm* laut per jam) atau dinyatakan dalam satuan meter per detik.

Angin dapat berpengaruh pada penerbangan dan penyebaran nyamuk. Jika kecepatan angin 11-14 m/detik atau 25-31 mil/jam, akan menghambat penerbangan nyamuk. Kecepatan angin pada saat matahari terbit dan tenggelam merupakan saat terbang nyamuk ke dalam atau keluar rumah dimana hal ini mempengaruhi banyaknya kontak antara manusia dengan nyamuk.

BAB 3

METODOLOGI PENELITIAN

A. Analisis Variabel Penelitian

Data banyak kasus DB dan data iklim adalah data sekunder yang masing-masing diperoleh dari Dinas Kesehatan dan Badan Pusat Statistik (BPS) Provinsi Jawa Timur. Data iklim merupakan hasil observasi Badan Meteorologi, Klimatologi dan Geofisika (BMKG) stasiun Juanda. Data banyak kasus DB dan data iklim yang digunakan adalah data per bulan dari Januari 2006 sampai Desember 2014 (Lampiran 1 dan 2).

Berdasarkan data yang diperoleh, terdapat dua variabel yang digunakan, yaitu variabel respon dan variabel prediktor. Banyak kasus DB di Kota Surabaya sebagai variabel respon, sedangkan iklim sebagai variabel prediktor. Keterangan variabel dapat dilihat pada Tabel 3.1 berikut ini.

Tabel 3.1. Variabel-variabel

Variabel Respon	Y	Banyak Kasus DB di Kota Surabaya
Variabel Prediktor	X_1	Kelembaban (%)
	X_2	Tekanan Udara (Mbs)
	X_3	Temperatur (0° C)
	X_4	Penyinaran Matahari (%)
	X_5	Curah Hujan (mm)
	X_6	Kecepatan Angin (knot)

Data yang digunakan adalah data bulanan selama 9 tahun yaitu sebanyak 108 data. Data dibagi menjadi dua, yaitu data *training* dan data *testing*. Data yang digunakan sebagai data *training* adalah data dari Januari 2006 sampai Desember 2013, yaitu sebanyak 96 data. Sedangkan data *testing* adalah data dari Januari sampai Desember 2014 yaitu sebanyak 12 data.

B. Metode Analisis

Tahap-tahap yang akan dilakukan pada penelitian ini adalah:

1. Eksplorasi Data

Data banyak kasus DB dianalisis untuk mengetahui pola data seperti pola musiman (*seasonal*) atau *trend*. Hal ini dapat dilakukan dengan melihat *time series plot* data banyak kasus DB.

2. Uji Stasioneritas

Kestasioneritasan data dapat dilihat dari *time series plot* dan hasil pengujian dengan Uji Dickey Fuller dan Transformasi Box-Cox. Data dikatakan stasioner jika pola data pada *time series plot* cenderung konstan dan tersebar di sekitaran garis rata-rata (*mean*). Pada Box-Cox *plot*, data yang sudah stasioner dalam varians akan menghasilkan nilai $\lambda = 0$, kemudian dilanjutkan dengan uji Dickey Fuller untuk menguji kestasioneritasan data dalam *mean*.

3. Identifikasi Model *INARIMA*

Identifikasi model bertujuan untuk menentukan orde p dan q pada proses *Integer-valued Autoregressive (INAR)* dan *Integer-valued Moving Average (INMA)*. Penentuan orde dilakukan dengan melihat *lag-lag* yang signifikan pada plot *ACF* dan *PACF*.

4. Pembagian Data

Data dibagi menjadi dua, yaitu data *training* dan data *testing*. Beberapa komposisi data *training* dan data *testing* yang dapat digunakan masing-masing 80% dan 20%, 75% dan 25% dan sebagainya tergantung dari tujuan dan kebutuhan peneliti. Pada penelitian ini, komposisi pembagian data adalah dari 108 data, digunakan 96 data untuk data *training* dan 12 data untuk data *testing*.

5. Estimasi Parameter

Estimasi parameter dilakukan untuk menentukan nilai parameter-parameter pada model yang belum diketahui sehingga dapat dihitung nilai prediksi untuk model tersebut. Metode estimasi parameter yang digunakan adalah Metode *Conditional Least Square (CLS)*.

6. Peramalan model *INARIMA*

Prediksi dilakukan dengan menggunakan model yang telah terbentuk dan nilai-nilai parameter yang telah diestimasi. Hasil yang diperoleh dari proses ini adalah nilai-nilai prediksi dari data observasi dan nilai *residual*.

7. Identifikasi Model *RBFNN*

Pada proses ini, data *input* yang digunakan adalah *residual* dari hasil prediksi proses *INARIMA*. Identifikasi model bertujuan menentukan banyak *neuron input* pada struktur jaringan *RBFNN*. Penentuan *input* dilakukan dengan melihat *lag-lag* signifikan pada *plot ACF* dan *PACF*.

8. Menentukan Nilai Pusat dan Varians

Data *input* dikelompokkan menjadi beberapa kelompok menggunakan *Self Organizing Map (SOM)* sehingga nilai pusat dan varians setiap kelompok dapat ditentukan. Nilai pusat diperoleh dari bobot-bobot hasil pembelajaran *SOM*, sedangkan nilai varians dihitung dari data masing-masing kelompok yang terbentuk. Banyaknya kelompok menentukan banyaknya *neuron* pada lapisan tersembunyi, sehingga ditentukan banyaknya kelompok yang meminimumkan nilai *MSE* dan *RMSE*.

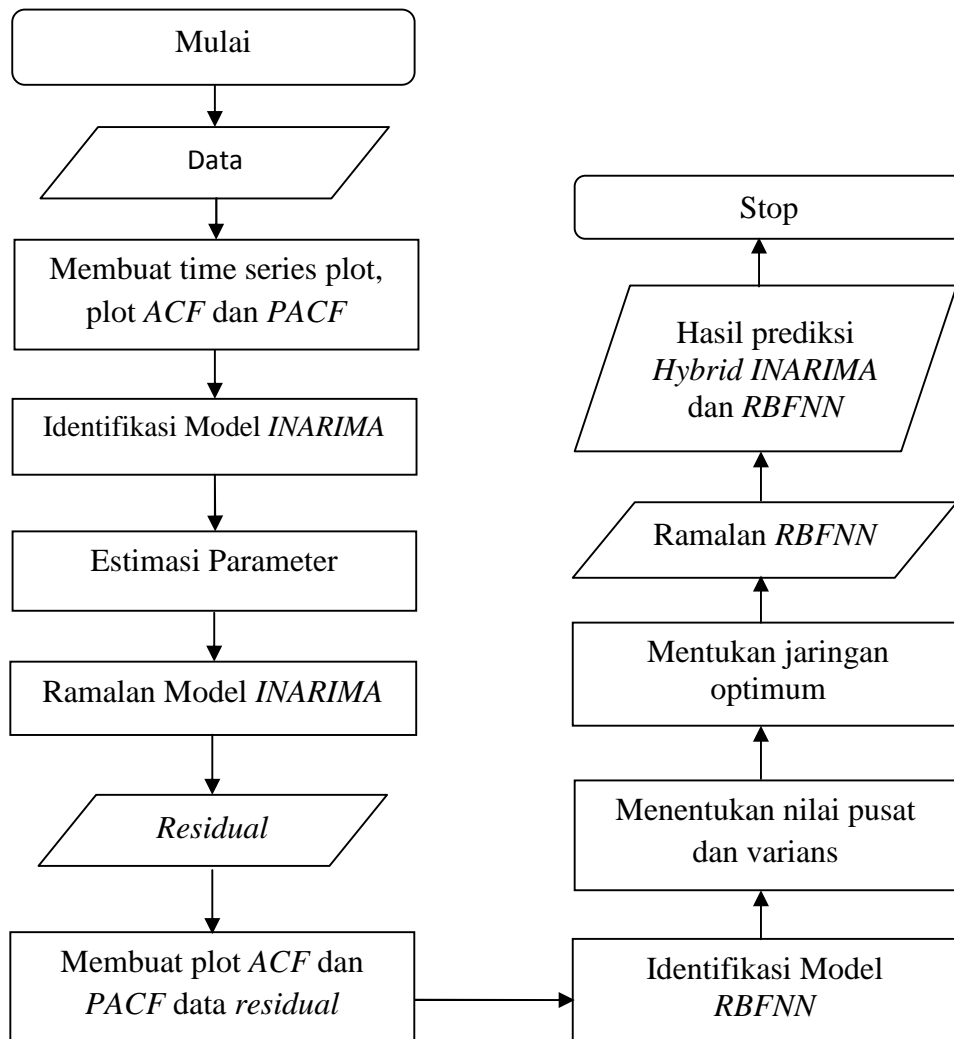
9. Ramalan model *RBFNN*

Langkah selanjutnya yaitu menghitung nilai ramalan dari model *RBFNN* untuk *residual*.

10. Ramalan model Hybrid *INARIMA-RBFNN*

Langkah terakhir yaitu melakukan peramalan dengan model *Hybrid INARIMA-RBFNN* yang dilakukan dengan menjumlahkan hasil ramalan model *INARIMA* dengan hasil ramalan model *RBFNN*.

Uraian di atas dapat dibentuk dalam diagram alur berikut:



Gambar 3.1. Diagram Alur

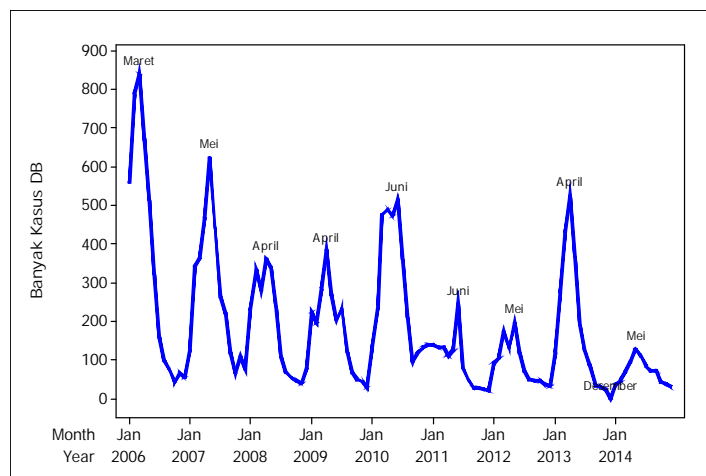
BAB 4

HASIL DAN PEMBAHASAN

Pada bab ini, akan diuraikan proses dan hasil penelitian yang merupakan penyelesaian dari rumusan masalah. Beberapa proses yang akan dibahas pada bab ini adalah deskripsi data banyak kasus demam berdarah dan data iklim, identifikasi model dan estimasi parameter untuk model *INARIMA*, pemodelan dengan *RBFNN*, dan peramalan menggunakan metode *Hybrid INARIMA-RBFNN*.

A. Deskripsi Data Banyak Kasus Demam Berdarah

Data yang digunakan pada penelitian ini adalah data sekunder banyak kasus demam berdarah (DB) di Kota Surabaya dari Januari 2006 sampai Desember 2014. Data ini diperoleh dari Dinas Kesehatan Provinsi Jawa Timur. Data dapat dilihat pada Lampiran 1. Berikut adalah *plot* sebaran data banyak kasus demam berdarah di Kota Surabaya:



Gambar 4.1. *Plot* Banyak Kasus Demam Berdarah di Kota Surabaya

Berdasarkan Gambar 4.1 dan Lampiran 1, dapat dilihat bahwa banyak kasus DB tertinggi terjadi pada bulan Maret 2006 sebanyak 841 kasus. Selanjutnya data mengalami penurunan. Setiap tahun mulai tahun 2006 sampai 2014, terdapat 1 bulan dengan angka tertinggi. Misal seperti pada

tahun 2007, bulan dengan banyak kasus DB tertinggi ada pada bulan Mei sebanyak 624 kasus. Tahun 2008 ada pada bulan April sebanyak 361 kasus, sampai tahun 2014, bulan Mei menjadi bulan dengan kasus DB terbanyak. Dalam kurun waktu 9 tahun, terdapat 4 bulan dimana banyak kasus DB tertinggi sering terjadi, yaitu bulan Maret, April, Mei, dan Juni. Pada bulan Desember 2013, terjadi penurunan hingga mencapai angka 0. Nilai ini merupakan nilai terendah sepanjang tahun 2006 sampai 2014. Selanjutnya banyak kasus DB kembali mengalami peningkatan meskipun tidak drastis. Jika dibandingkan dengan tahun-tahun sebelumnya, banyak kasus DB sepanjang tahun 2014 adalah yang paling rendah. Rata-rata banyak kasus demam berdarah mengalami peningkatan pada pertengahan tahun, berkisar antara bulan April sampai Juli.

B. Deskripsi Data Iklim

Selain data DB, pada penelitian ini juga digunakan data iklim. Komponen iklim yang menjadi salah satu faktor penyebaran DB yang digunakan pada penelitian ini adalah kelembaban udara, tekanan udara, temperatur, penyinaran matahari, curah hujan, dan kecepatan angin. Data iklim ini merupakan data sekunder yang diperoleh dari Badan Pusat Statistik (BPS) Provinsi Jawa Timur. Data merupakan hasil observasi Badan Meteorologi, Klimatologi dan Geofisika (BMKG) stasiun Juanda. Data yang digunakan adalah data per bulan dari Januari 2006 sampai Desember 2014 (Lampiran 2).

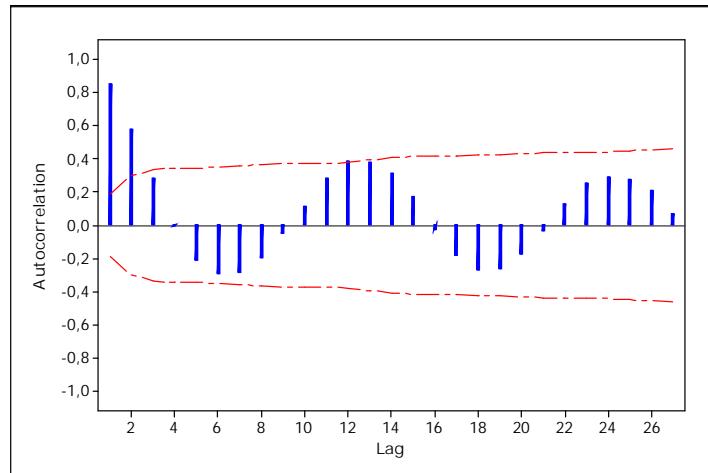
Berdasarkan informasi data iklim pada Lampiran 2, pada bulan Februari 2006, curah hujan mencapai nilai maksimum yaitu 886 mm dan kecepatan angin mencapai nilai minimum yaitu 0 knot. Tingginya curah hujan akan menambah jumlah tempat perkembangbiakan nyamuk alami. Pada saat yang sama, kecepatan angin yang minimum mengakibatkan penerbangan nyamuk tidak terhambat sehingga mempengaruhi banyaknya kontak antara manusia dengan nyamuk. Kondisi ini dapat dijadikan indikasi sebagai penyebab tingginya kasus DB pada bulan Maret 2006 yang mencapai 841 kasus.

Pada bulan Agustus sampai Nopember 2014, penyinaran matahari mencapai maksimum sehingga temperatur dan tekanan udara juga mencapai nilai maksimum. Kondisi ini tidak disukai nyamuk karena rata-rata temperatur optimum nyamuk berkisar antara 25°C - 27°C , sedangkan pada bulan-bulan tersebut temperatur maksimum mencapai $29,9^{\circ}\text{C}$. Pengaruh kondisi ini terhadap banyaknya kasus DB adalah pada bulan Agustus sampai Desember 2014 kasus DB menurun menjadi sebanyak 29 kasus yang sebelumnya mencapai angka diatas 100 kasus.

C. Identifikasi Model *INARIMA*

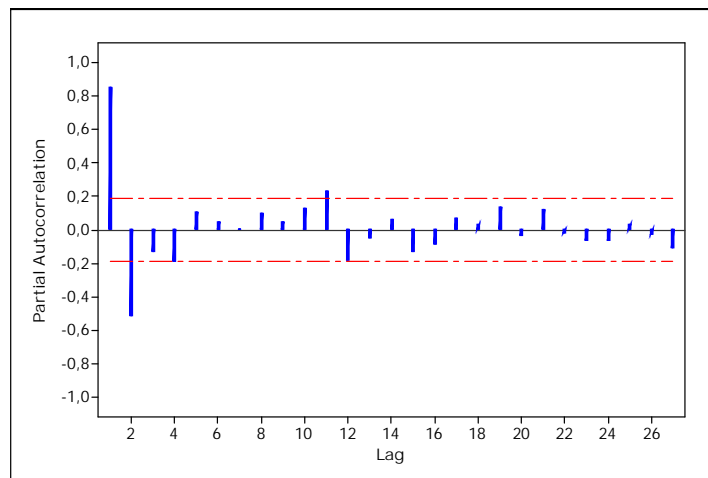
Data banyak kasus demam berdarah di Kota Surabaya merupakan data yang stasioner. Hal ini ditunjukkan dengan hasil pengujian dengan uji statistik *Augmented Dickey Fuller* (Lampiran 7). Hipotesis alternatif pada uji ini adalah data stasioner. Pengujian ini menghasilkan nilai $p\text{-value} = 0,01$ sehingga nilai $p\text{-value} < \alpha = 0,05$. Maka diperoleh keputusan tolak hipotesis awal dan menerima hipotesis alternatif. Jadi dapat disimpulkan bahwa data banyak kasus demam berdarah di Kota Surabaya adalah stasioner.

Identifikasi model bertujuan untuk menentukan order p dan q pada proses *Integer-valued Autoregressive (INAR)* dan *Integer-valued Moving Average (INMA)*. Identifikasi model dilakukan pada data yang sudah stasioner. Penentuan order dilakukan dengan melihat *lag-lag* yang signifikan pada *plot Autocorrelation Function (ACF)* dan *Partial Autocorrelation Function (PACF)*. Gambar 4.2 berikut merupakan *plot ACF* dari data banyak kasus DB yang menunjukkan bahwa *lag-lag* awal turun secara cepat menuju 0 (*dies down*).



Gambar 4.2. *Plot ACF* Data Banyak Kasus DB di Kota Surabaya

Plot PACF dari data banyak kasus DB pada Gambar 4.3, menunjukkan *lag* 1 dan 2 signifikan, sedangkan *lag-lag* berikutnya tidak ada yang signifikan. Oleh karena itu, *plot PACF* dikatakan *cut off after lag 2*.



Gambar 4.3. *Plot PACF* Data Banyak Kasus DB di Kota Surabaya

Berdasarkan analisis *plot ACF* dan *PACF*, dapat diduga bahwa model yang terbentuk untuk data banyak kasus DB di Kota Surabaya adalah *Integer-valued Autoregressive (INAR)* dengan order 2 atau *INAR(2)*. Pada penelitian ini, digunakan 6 variabel prediktor yang merupakan faktor-faktor iklim yang mempengaruhi penyebaran banyak kasus DB di Kota Surabaya.

Model $INAR(2)$ dengan 6 variabel prediktor dituliskan sebagai berikut:

$$Y_t = r_1 \circ Y_{t-1} + r_2 \circ Y_{t-2} + Z_t$$

dengan $Z_t \sim POI(\lambda_t)$, $\lambda_t = \exp(\mathbf{x}_t^T \boldsymbol{\beta})$ dan $r_i = \frac{1}{1 + \exp(-\mathbf{w}_i^T \boldsymbol{\beta}_i)}$

dimana

$$\mathbf{x}_t = (x_{t,0}, x_{t,1}, \dots, x_{t,6}), \text{ dengan } x_{t,0} = 1$$

$$\mathbf{w}_1 = (w_{1,0}, w_{1,1}, \dots, w_{1,6})$$

$$\mathbf{w}_2 = (w_{2,0}, w_{2,1}, \dots, w_{2,6})$$

$$\boldsymbol{\beta} = (\beta_0, \beta_1, \dots, \beta_6)$$

$\mathbf{w}_2 = (w_{2,0}, w_{2,1}, \dots, w_{2,6})$ dengan $w_{2,0} = 1$ untuk r_1 , $w_{2,1} = 1$ untuk r_2 dan 0

untuk $w_{2,r}$ yang lain, sehingga

$$r_1 = \frac{1}{1 + \exp(-(0, 1, 0, \dots, 0)(1, x_{1,1}, \dots, x_{1,6}))}$$

$$r_1 = \frac{1}{1 + \exp(-x_1)} \text{ dan}$$

$$r_2 = \frac{1}{1 + \exp(-(0, 0, 1, \dots, 0)(1, x_{2,1}, x_{2,2}, \dots, x_{1,6}))}$$

$$r_2 = \frac{1}{1 + \exp(-x_2)}$$

Persamaan r_1 dan r_2 disubstitusikan pada model $INAR(2)$, sehingga model

$INAR(2)$ dapat ditulis sebagai berikut:

$$Y_t = \frac{1}{1 + \exp(-x_1)} \circ Y_{t-1} + \frac{1}{1 + \exp(-x_2)} \circ Y_{t-2} + Z_t$$

dengan

$$\begin{aligned} E(Y_t | \mathbf{x}_t, Y_{t-1}, Y_{t-2}) &= \frac{1}{1 + \exp(-x_1)} Y_{t-1} + \frac{1}{1 + \exp(-x_2)} Y_{t-2} + \lambda_t \\ &= \frac{1}{1 + \exp(-x_1)} Y_{t-1} + \frac{1}{1 + \exp(-x_2)} Y_{t-2} + \exp(\beta_0 + \beta_1 x_1 + \dots + \beta_6 x_6) \end{aligned}$$

D. Estimasi Parameter

Data banyak kasus demam berdarah di Kota Surabaya dari Januari 2006 sampai Desember 2016 adalah sebanyak 108 data. Berdasarkan hasil identifikasi model, *input* yang digunakan pada model *INAR(2)* adalah Y_{t-1} dan Y_{t-2} , sehingga dua data awal tidak digunakan. Data ini kemudian dibagi menjadi data *training* sebanyak 94 data dan data *testing* sebanyak 12 data. Data *training* digunakan untuk proses estimasi parameter.

Metode estimasi parameter yang digunakan adalah metode *Conditional Least Square (CLS)*. Berdasarkan prinsip dasar metode *CLS* pada persamaan (2.26), maka diminimumkan *conditional sum of squares*:

$$Q(\beta, \gamma) = \sum_{t=1}^{94} (Y_t - E(Y_t | \mathbf{x}_t, Y_{t-1}, Y_{t-2}))^2$$

$$Q(\beta, \gamma) = \sum_{t=1}^{94} \left(Y_t - \left(\frac{1}{1 + \exp(-\mathbf{w}_t^T \beta_1)} Y_{t-1} + \frac{1}{1 + \exp(-\mathbf{w}_t^T \beta_2)} Y_{t-2} + \exp(\mathbf{x}_t^T \gamma) \right) \right)^2$$

Hasil estimasi parameter dapat di lihat pada Tabel 4.1 berikut.

Tabel 4.1. Hasil Estimasi *INAR(2)* dengan Metode *CLS*

Parameter	Nilai Estimasi	Standard Error	t Stat	p-value
\hat{u}_0	-510,39	271,06	-1,883	0,063
\hat{u}_1	0,185	0,0577	3,203	0,0019
\hat{u}_2	0,463	0,265	1,747	0,084
\hat{u}_3	1,163	0,479	2,427	0,017
\hat{u}_4	-0,0057	0,0158	-0,358	0,720
\hat{u}_5	0,00029	0,00099	0,292	0,771
\hat{u}_6	-0,119	0,153	-0,779	0,438
$\hat{\chi}_1$	1,226	0,264	4,642	$1,228 \times 10^{-5}$
$\hat{\chi}_2$	-12905	0		0

Berdasarkan hasil estimasi parameter pada Tabel 4.1, terdapat beberapa parameter yang tidak signifikan karena nilai $p\text{-value} > \alpha = 0,05$. Parameter yang signifikan adalah \hat{u}_1 , \hat{u}_3 , \hat{x}_1 , dan \hat{x}_2 . Selanjutnya, dilakukan proses estimasi kembali menggunakan parameter-parameter yang signifikan dan diperoleh dugaan model sebagai berikut:

1. Model *INAR(2)* dengan x_1 dan x_3

Model umum *INAR(2)* dengan variabel prediktor yang signifikan adalah x_1 dan x_3 dapat dituliskan sebagai berikut:

$$Y_t = \frac{1}{1 + \exp(-\hat{x}_1)} \circ Y_{t-1} + \frac{1}{1 + \exp(-\hat{x}_2)} \circ Y_{t-2} + Z_t$$

dimana Z_t adalah $\exp(\hat{u}_1 x_1 + \hat{u}_3 x_3)$. Parameter-parameter pada model kemudian diestimasi menggunakan *CLS* yang ditampilkan pada Tabel 4.2.

Tabel 4.2. Hasil Estimasi Model *INAR(2)* dengan x_1 dan x_3

Parameter	Nilai Estimasi
\hat{u}_1	-0,358
\hat{u}_3	-1,59
\hat{x}_1	1,924
\hat{x}_2	$-2,35 \times 10^9$

Berdasarkan model dan hasil estimasi yang diperoleh, dapat ditentukan

nilai $\hat{r}_1 = \frac{1}{1 + \exp(-1,924)} = 0,872$, sehingga model *INAR(2)* dengan x_1

dan x_3 menjadi

$$Y_t = 0,872 \circ Y_{t-1} + \frac{1}{1 + \exp(-2,35 \times 10^9)} \circ Y_{t-2} + Z_t$$

dimana Z_t adalah $\exp(-0,358x_1 - 1,59x_3)$.

2. Model *INAR(1)* dengan x_1

Model umum *INAR(1)* dengan variabel prediktor yang signifikan adalah x_1 dapat dituliskan sebagai berikut:

$$Y_t = \frac{1}{1 + \exp(-\hat{x}_1)} \circ Y_{t-1} + Z_t$$

dimana Z_t adalah $\exp(\hat{u}_1 x_1)$. Parameter \hat{x}_1 dan \hat{u}_1 kemudian diestimasi menggunakan *CLS* yang menghasilkan nilai $\hat{x}_1 = 1,139$ dan $\hat{u}_1 = 0,049$.

Nilai estimator yang diperoleh disubstitusikan pada model umum sehingga diperoleh $\hat{r}_1 = \frac{1}{1 + \exp(-1,139)} = 0,757$ dan model dengan hasil

estimasi sebagai berikut:

$$Y_t = 0,757 \circ Y_{t-1} + Z_t$$

dimana Z_t adalah $\exp(-0,049x_1)$.

3. Model *INAR(2)*

Model umum *INAR(2)* tanpa variabel prediktor adalah sebagai berikut:

$$Y_t = r_1 \circ Y_{t-1} + r_2 \circ Y_{t-2} + Z_t$$

Parameter r_1 dan r_2 kemudian diestimasi menggunakan *CLS* yang menghasilkan nilai $\hat{r}_1 = 0,904$ dan $\hat{r}_2 = 1,04 \times 10^{-9}$. Nilai estimator yang diperoleh disubstitusikan pada model umum sehingga diperoleh model dengan hasil estimasi sebagai berikut:

$$Y_t = 0,904 \circ Y_{t-1} + 1,04 \times 10^{-9} \circ Y_{t-2} + Z_t$$

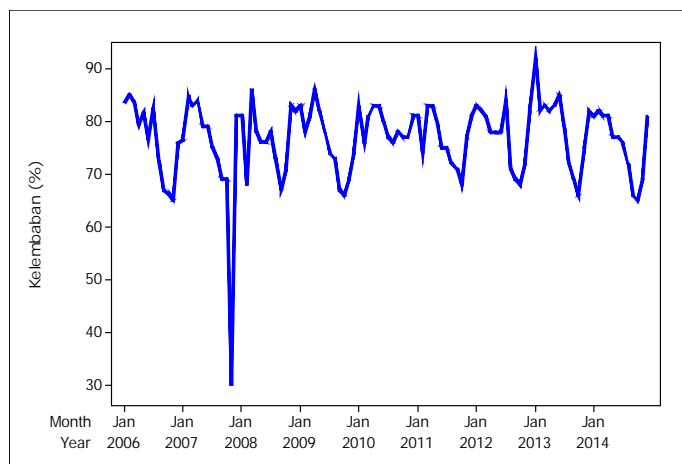
E. Ramalan Model *INAR*

Model-model *INAR* yang diperoleh kemudian digunakan untuk menghitung nilai ramalan banyak kasus demam berdarah di Kota Surabaya 12 periode ke depan. Data aktual terakhir yang digunakan pada penelitian ini adalah bulan Desember 2014, sehingga periode ramalan dilakukan untuk bulan Januari sampai Desember 2015. Pada model *INAR* dengan variabel

prediktor terdapat nilai variabel prediktor pada waktu ke- t sehingga dibutuhkan nilai x_1 dan x_3 untuk periode Januari sampai Desember 2015. Masing-masing nilai ramalan variabel prediktor ini akan ditentukan menggunakan metode *Autoregressive Integrated Moving Average (ARIMA)*.

1. Pemodelan *ARIMA* Kelembaban Udara (x_1)

Pola data kelembaban udara di Kota Surabaya pada Gambar 4.4 menunjukkan bahwa data berada di sekitar rata-rata yaitu pada kelembaban 76,75%. Pada Nopember 2007, kelembaban udara berada pada nilai terendah yaitu 29,9%, sedangkan nilai kelembaban tertinggi terjadi pada bulan Januari 2013.

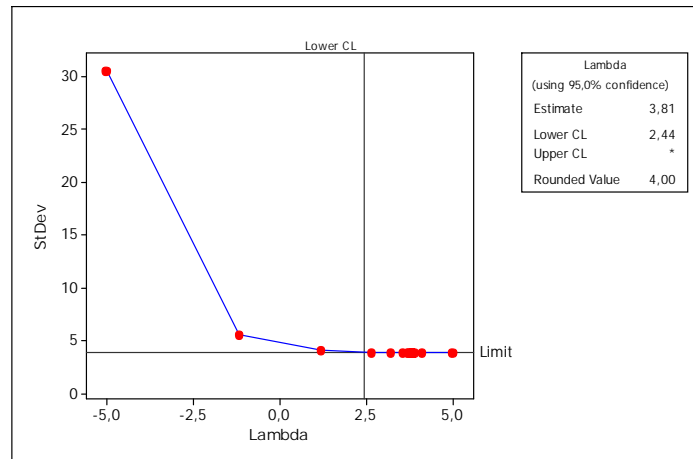


Gambar 4.4. *Time Series Plot* Data Kelembaban Udara

a. Identifikasi

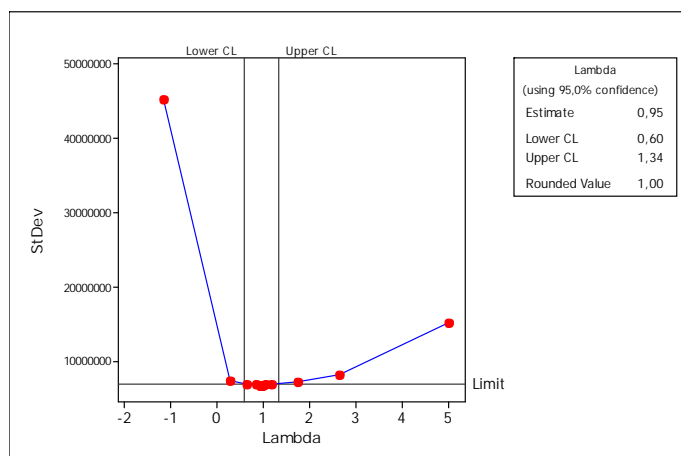
Pada tahap identifikasi, akan diuji apakah data deret waktu sudah stasioner dalam *mean* dan varians. Pertama, akan diuji kestasioneran data dalam varians menggunakan Transformasi *Box-Cox*. Pada proses ini, terdapat nilai $\lambda(\lambda)$ yang akan mempengaruhi transformasi apa yang digunakan. Jika nilai $\lambda = 1$, maka tidak perlu dilakukan transformasi dan menunjukkan bahwa data sudah stasioner dalam

varians. Gambar 4.5 berikut merupakan *Box-Cox plot* data kelembaban udara di Kota Surabaya sebelum transformasi.



Gambar 4.5. *Box-Cox Plot* Kelembaban Udara Sebelum Transformasi

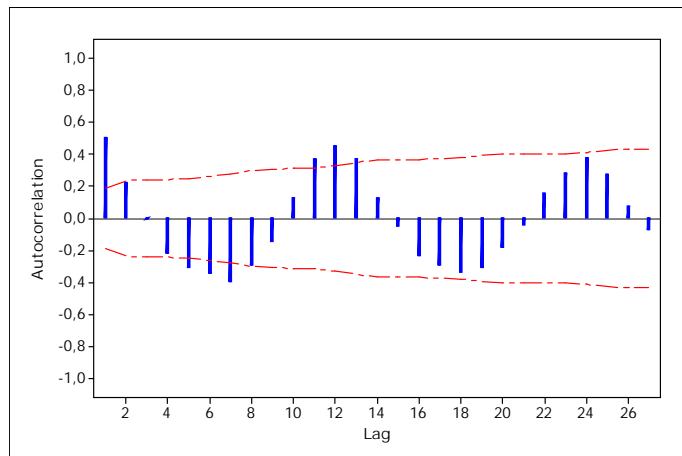
Berdasarkan Gambar 4.5, nilai λ yang diperoleh adalah 4 sehingga data perlu ditransformasi. Data yang telah ditransformasi kemudian diuji kembali dengan *Box-Cox plot* yang ditunjukkan pada Gambar 4.6. Pada Gambar 4.6, dapat dilihat bahwa nilai $\lambda = 1$. Maka data kelembaban udara sudah stasioner dalam varians.



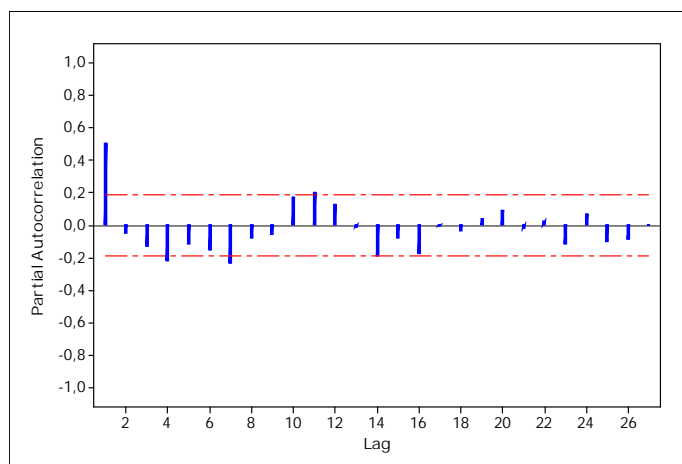
Gambar 4.6. *Box-Cox Plot* Data Kelembaban Udara Setelah Transformasi

Setelah data stasioner dalam varians, selanjutnya diuji apakah data stasioner dalam *mean* dengan uji statistik *Dickey-Fuller*. Pengujian ini menghasilkan nilai $p\text{-value} = 0,01 < \alpha = 0,05$, sehingga tolak hipotesis awal (data tidak stasioner dalam *mean*). Jadi, data kelembaban udara di Kota Surabaya sudah stasioner dalam *mean*.

Data yang sudah stasioner kemudian digunakan untuk menduga model yang sesuai. Pendugaan model dilakukan dengan menentukan order *ARIMA* dengan melihat *plot ACF* dan *PACF*. *Plot ACF* dan *PACF* data kelembaban udara di Kota Surabaya ditunjukkan pada Gambar 4.7 dan 4.8.



Gambar 4.7. *Plot ACF* Data Kelembaban Udara



Gambar 4.8. *Plot PACF* Data Kelembaban Udara

Berdasarkan *plot ACF* dan *PACF* data kelembaban udara pada Gambar 4.7 dan 4.8, model dugaan awal adalah *AR(1)* atau *ARIMA* (1,0,0) yaitu:

$$\dot{x}_{1t} = w_1 \dot{x}_{1t-1} + a_{1t}$$

b. Estimasi

Estimasi parameter dilakukan dengan metode *Conditional Least Square (CLS)* dan diperoleh hasil estimasi parameter untuk model data kelembaban udara adalah sebagai berikut:

Tabel 4.3. Hasil Estimasi Parameter untuk Kelembaban Udara

Parameter	Nilai Estimasi	Standard Error	t Stat	p-value
w_1	0,856	0,05048	16,97	0,0001
\check{S}_{AO}	38984571	6200174,3	6,29	0,0001
\check{S}_{LS}	-27374596	7759492,4	-3,53	0,0006

Maka model *ARIMA* untuk kelembaban udara di Kota Surabaya adalah

$$\dot{x}_{1t} = 0,856 \dot{x}_{1t-1} + 38984571 I_{st}^{34} - 27374596 I_{at}^{36} a_{1t}$$

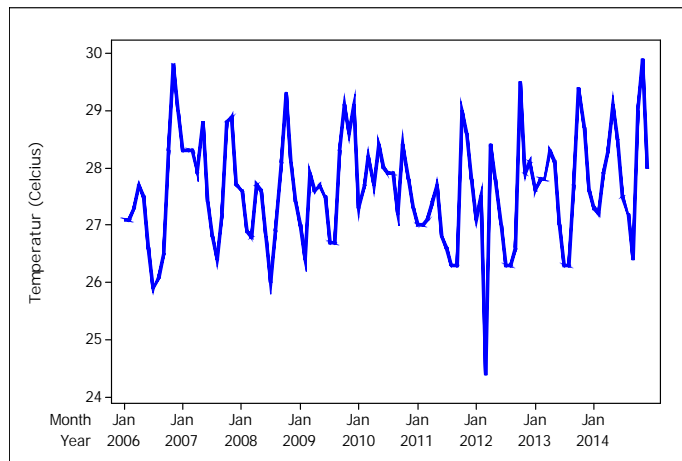
dengan $\dot{x}_{1t} = (x_{1t})^{\frac{1}{4}}$

c. Ramalan Kelembaban Udara

Berdasarkan model dan hasil estimasi yang diperoleh, kemudian dilakukan ramalan 12 periode ke depan untuk kelembaban udara di Kota Surabaya. Hasil ramalan kelembaban udara untuk bulan Januari sampai Desember 2015 dapat dilihat pada Lampiran 8.

2. Ramalan *ARIMA* Temperatur (x_3)

Data temperatur di Kota Surabaya memiliki nilai maksimum dan minimum masing-masing adalah 29,9°C dan 24,4°C. Nilai temperatur yang minimum terjadi pada bulan Maret 2012, sedangkan nilai maksimum terjadi pada bulan Nopember 2014. Gambar 4.9 berikut merupakan *time series plot* data temperatur di Kota Surabaya dari Januari 2006 sampai Desember 2014.

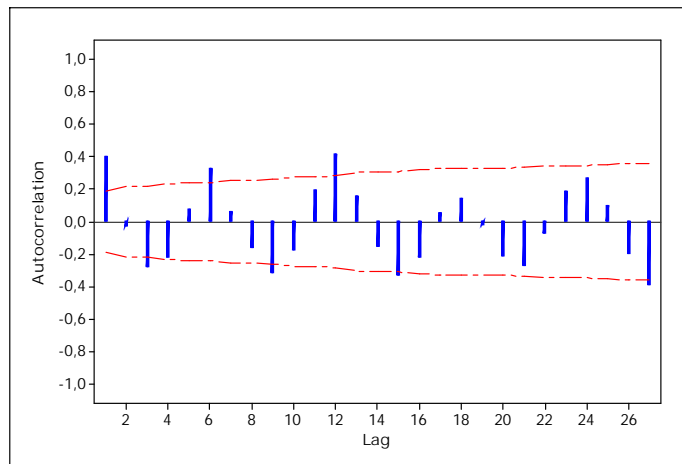


Gambar 4.9. *Time Series Plot* Data Temperatur

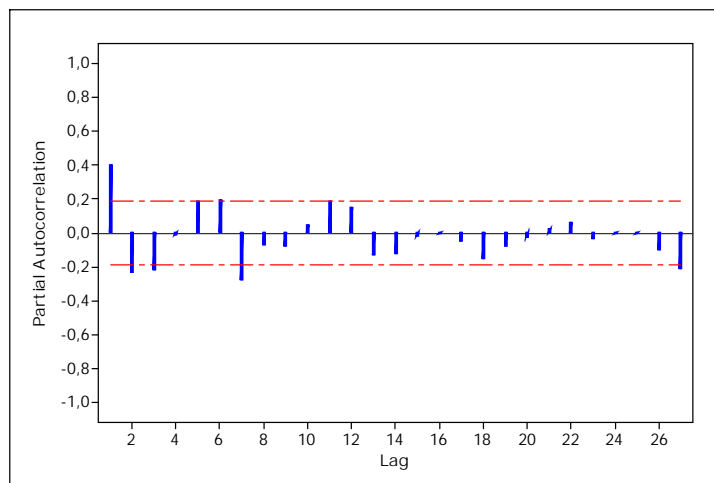
a. Identifikasi

Pertama, akan diuji kestasioneran data dalam varians menggunakan Transformasi *Box-Cox*. Pada proses Transformasi *Box-Cox* untuk data temperatur di Kota Surabaya menghasilkan nilai $\lambda = -0,5$ sehingga data perlu ditransformasi. Data yang telah ditransformasi kemudian diuji kembali dengan *Box-Cox plot* sampai diperoleh nilai $\lambda = 1$. Maka data temperatur sudah stasioner dalam varians.

Setelah data stasioner dalam varians, selanjutnya diuji apakah data stasioner dalam *mean* dengan uji statistik *Dickey-Fuller*. Pengujian ini menghasilkan nilai *p-value* = 0,01 yang berarti kurang dari $\alpha = 0,05$ dengan *differencing* 12 sehingga tolak hipotesis awal (data tidak stasioner dalam *mean*). Jadi dapat disimpulkan bahwa data temperatur di Kota Surabaya sudah stasioner dalam *mean*. Data yang sudah stasioner kemudian digunakan untuk menduga model yang terbentuk. Pendugaan model dilakukan dengan menentukan order *ARIMA* dengan melihat *plot ACF* dan *PACF*. *Plot ACF* dan *PACF* data temperatur di Kota Surabaya ditunjukkan pada Gambar 4.10 dan 4.11.



Gambar 4.10. *Plot ACF* Data Temperatur



Gambar 4.11. *Plot PACF* Data Temperatur

Berdasarkan *plot ACF* dan *PACF* data temperatur pada Gambar 4.10 dan 4.11, model dugaan awal adalah *ARIMA* $(1,0,0)(1,1,0)^{12}$, yaitu:

$$\dot{x}_{3t} = w_1 \dot{x}_{3t-1} + (1 + \Phi_1) \dot{x}_{3t-12} + (w_1 + w_1 \Phi_1) \dot{x}_{3t-13} + \Phi_1 \dot{x}_{3t-24} - w_1 \Phi_1 \dot{x}_{3t-25} + a_{3t}$$

b. Estimasi

Estimasi parameter dilakukan dengan metode *CLS* dan diperoleh hasil estimasi pada Tabel 4.4 berikut.

Tabel 4.4. Hasil Estimasi Parameter untuk Temperatur

Parameter	Nilai Estimasi	Standard Error	t Stat	p-value
w_1	0,31016	0,09812	3,16	0,0021
Φ_1	-0,34892	0,10282	-3,39	0,0010

Maka model *ARIMA* untuk temperatur di Kota Surabaya adalah

$$\dot{x}_{3t} = 0,31016\dot{x}_{3t-1} + 0,65108\dot{x}_{3t-12} + 0,20193\dot{x}_{3t-13} - 0,34892\dot{x}_{3t-24} + 0,10822\dot{x}_{3t-25} + a_{3t}$$

dengan $\dot{x}_{3t} = \left(\frac{1}{(x_{3t})^{1/1,14}} \right)^2$

c. Ramalan Temperatur

Berdasarkan model dan hasil estimasi yang di peroleh, kemudian dilakukan ramalan 12 periode ke depan untuk temperatur di Kota Surabaya. Hasil ramalan temperatur untuk bulan Januari sampai Desember 2015 dapat dilihat pada Lampiran 8.

Berdasarkan model dan hasil peramalan variabel prediktor yang telah diperoleh, maka dapat ditentukan nilai ramalan untuk data *training* dan data *testing* untuk masing-masing dugaan model. Proses pada setiap model juga menghasilkan nilai akurasi *MSE* dan *RMSE* yang ditampilkan pada Tabel 4.5.

Tabel 4.5. Nilai *MSE* dan *RMSE* untuk Model dugaan *INAR*

Model dugaan	<i>MSE</i>	<i>RMSE</i>
<i>INAR</i> (2) dengan x_1 dan x_3	516,491	22,726
<i>INAR</i> (1) dengan x_1	440,458	20,987
<i>INAR</i> (2)	491,145	22,162

Berdasarkan Tabel 4.6 diperoleh dugaan model yang meminimumkan nilai *MSE* dan *RMSE* adalah model *INAR(1)* dengan variabel prediktor x_1 . Pada penelitian ini akan dilakukan analisis lebih lanjut dengan *Hybrid RBFNN* menggunakan model *INAR(1)* dengan x_1 dan model *INAR(2)*.

F. Ramalan *Hybrid* untuk Model *INAR(1)* dengan x_1

Berdasarkan model *INAR(1)* dengan variabel prediktor x_1 dan hasil estimasi parameternya, perhitungan nilai ramalan untuk periode Januari 2015 atau untuk $t = 109$ adalah sebagai berikut:

$$\hat{Y}_{109} = 0,757 \circ Y_{108} + \exp(-0,049x_{1,109})$$

$$\hat{Y}_{109} = (0,757) \circ (29) + \exp((-0,049)(80,7))$$

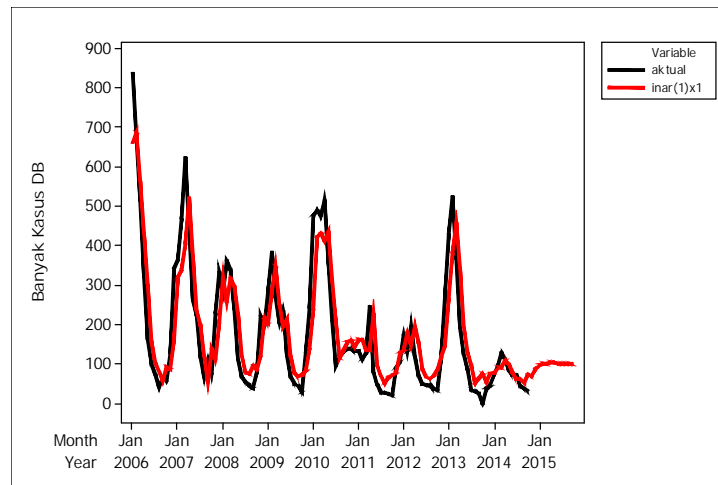
$$\hat{Y}_{109} = 68,025$$

Perhitungan nilai ramalan untuk periode Februari sampai Desember 2015 dapat dilakukan dengan cara yang sama. Hasil ramalan model *INAR(1)* dengan x_1 ditunjukkan pada Tabel 4.6.

Tabel 4.6. Hasil Ramalan *INAR(1)* dengan x_1

Bulan	Banyak Kasus DB	Pembulatan
Januari 2015	68,025	69
Februari	86,913	87
Maret	95,824	96
April	99,823	100
Mei	101,433	102
Juni	101,905	102
Juli	101,855	102
Agustus	101,586	102
September	101,242	101
Oktober	100,892	101
Nopember	100,563	101
Desember	100,268	101

Plot data aktual dengan hasil ramalan data *training* dan data *testing* serta hasil ramalan periode Januari sampai Desember 2015 dapat dilihat pada Gambar 4.12. Gambar 4.12 menunjukkan nilai hasil ramalan mendekati dan mengikuti pola data aktual.



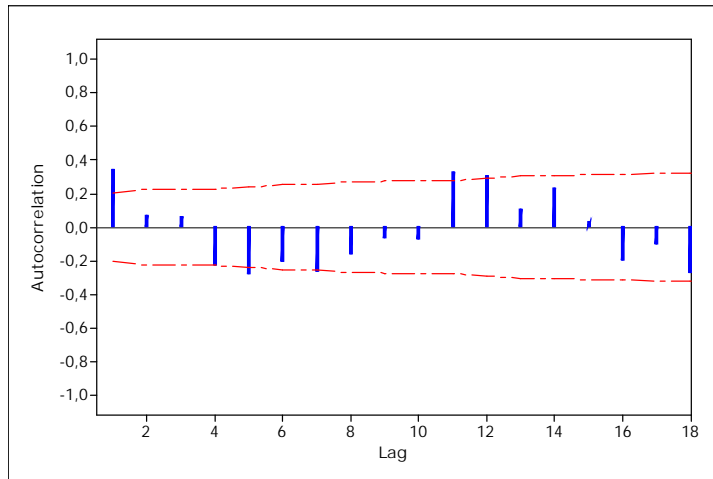
Gambar 4.12. *Plot* Data Aktual dengan Hasil Ramalan Model $INAR(1)$ dengan x_1

Hasil ramalan untuk data *training* dan data *testing* yang diperoleh kemudian digunakan untuk menghitung nilai *residual*. Nilai *residual* merupakan nilai kesalahan hasil prediksi atau selisih antara nilai aktual dengan hasil peramalan yang diperoleh. Nilai *residual* dari hasil peramalan model $INAR(1)$ dengan x_1 ini yang akan menjadi *input* untuk *Hybrid Radial Basis Function Neural Network (RBFNN)*.

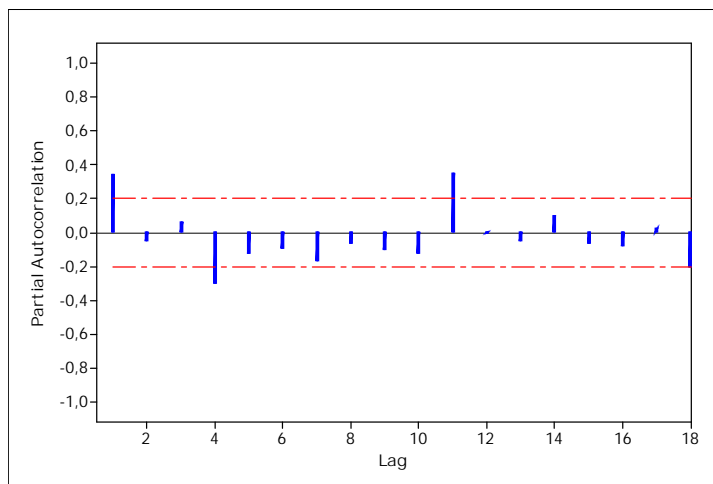
1. Identifikasi Model *RBFNN*

Setelah proses pemodelan data banyak kasus demam berdarah dilakukan menggunakan metode $INAR(1)$ dengan x_1 , kemudian dilakukan pemodelan menggunakan *RBFNN*. Pada proses ini, data *input* yang digunakan adalah *residual* dari hasil prediksi proses $INAR(1)$ dengan x_1 . Data *residual* dinotasikan dengan a_t . Tahap pertama yang dilakukan adalah identifikasi model yang bertujuan untuk menentukan banyak *neuron input*

pada struktur jaringan *RBNN*. Penentuan *input* dilakukan dengan melihat *lag-lag* signifikan pada *plot ACF* dan *PACF*. Berbeda dengan cara identifikasi model pada model *INAR*, pada proses *neural network* tidak mempertimbangan *lag-lag* turun cepat (*dies down*) atau *lag-lag* yang terputus setelah *lag* tertentu (*cut off*).



Gambar 4.13. *Plot ACF* a_t



Gambar 4.14. *Plot PACF* a_t

Plot ACF dan *PACF* pada Gambar 4.13 dan 4.14 menunjukkan *lag-lag* yang signifikan adalah *lag 1* dan *lag 4*. Oleh karena itu, *input* pada model *RBNN* adalah a_{t-1} dan a_{t-4} , sehingga terdapat 2 *neuron* pada lapisan *input*.

2. Menentukan Nilai Pusat dan Varians

Pada proses ini, data *input* dikelompokkan menjadi beberapa kelompok menggunakan *Self Organizing Map (SOM)* sehingga nilai pusat dan varians setiap kelompok dapat ditentukan. Nilai pusat diperoleh dari bobot-bobot hasil pembelajaran *SOM*. Kemudian dari setiap kelompok yang terbentuk, ditentukan nilai varians masing-masing kelompok. Banyaknya kelompok menentukan banyaknya *neuron* pada lapisan tersembunyi. Pembentukan model optimum dilakukan dengan menentukan banyak *neuron* pada lapisan tersembunyi dengan nilai kriteria kebaikan model yang minimum. Tabel 4.7 berikut adalah hasil pembelajaran *SOM* untuk 2 sampai 7 kelompok.

Tabel 4.7. Nilai *MSE* dan *RMSE* untuk 4 sampai 7 Kelompok

Banyaknya Kelompok	<i>MSE</i>	<i>RMSE</i>
2	371,494	19,274
3	396,015	19,900
4	1593,500	39,919
5	519,264	22,787
6	655,698	25,607
7	522,768	22,864

Berdasarkan Tabel 4.8, nilai *MSE* dan *RMSE* yang minimum ada pada kelompok 2 yaitu 371,494 untuk *MSE* dan 19,274 untuk *RMSE*, sehingga ditentukan banyak *neuron* pada lapisan tersembunyi adalah 2 *neuron*. Nilai pusat dan varians dari 2 kelompok dapat dilihat pada Lampiran 6. Jadi, model terbaik yang diperoleh pada penelitian ini adalah model *RBFNN* dengan 2 *neuron* pada lapisan *input*, 2 *neuron* pada lapisan tersembunyi dan 1 *neuron* pada lapisan *output*.

3. Prediksi *RBFNN*

Setelah model optimum terbentuk, langkah selanjutnya adalah melakukan peramalan. Arsitektur *RBFNN* yang telah terbentuk adalah arsitektur dengan 2 *neuron* pada lapisan *input*, 2 *neuron* pada lapisan tersembunyi dan 1 *neuron* pada lapisan *output* dapat dituliskan dengan

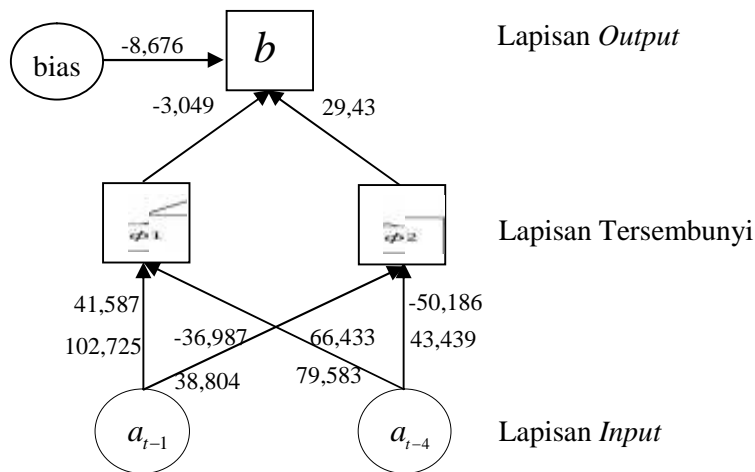
$$b_i = \sum_{j=1}^2 w_j w_j(a)$$

Pada proses pembelajaran *RBFNN*, diperoleh nilai-nilai bobot pada setiap penghubung *neuron* pada lapisan tersembunyi dengan *neuron* pada lapisan *output*. Banyak *neuron* pada lapisan tersembunyi yang terbentuk sebanyak 2 *neuron*, sehingga terdapat 2 bobot *neuron* dan 1 bobot bias yang ditunjukkan pada Tabel 4.8 berikut.

Tabel 4.8. Bobot Pembelajaran *RBFNN*

Bobot	Nilai
w_{bias}	-8,6762
w_1	-3,0499
w_2	29,4333

Arsitektur model terbaik *RBFNN* dapat dilihat pada Gambar 4.15 berikut:



Gambar 4.15. Arsitektur Model terbaik *RBFNN*

Model dan bobot-bobot yang diperoleh dari pembelajaran *RBFNN* kemudian digunakan untuk perhitungan nilai ramalan periode ke depan. Data banyak kasus demam berdarah dari Januari 2006 sampai Desember 2014 yang digunakan ada sebanyak 108 data, sehingga periode Januari 2015 adalah

periode ke 109. Perhitungan nilai ramalan untuk periode Januari 2015 (\hat{b}_{109})

adalah sebagai berikut:

$$\begin{aligned}\hat{b}_{109} &= \sum_{j=1}^2 w_j w_j(a) \\ &= -8,6762 - 3,0499w_1(a) + 29,4333w_2(a)\end{aligned}$$

Input yang digunakan pada model *RBFFNN* adalah a_{t-1} dan a_{t-4} atau a_{108} dan a_{105} yang masing-masing mempunyai nilai $-43,75$ dan $9,60$. Nilai *input*, nilai pusat (\sim_{ij}) dan varians (\dagger_{ij}^2) yang telah diperoleh digunakan untuk menentukan nilai $w_j(a)$, $i = t-1, t-4$ dan $j = 1, 2$ dengan perhitungan sebagai berikut:

$$\begin{aligned}w_j(a) &= \exp\left(-\sum_{i=t-1}^{t-4} \left(\frac{a_i - \sim_{ij}}{2\dagger_{ij}}\right)^2\right) \\ w_1(a) &= \exp\left[-\left\{\left(\frac{a_{t-1} - \sim_{t-1,1}}{2\dagger_{t-1,1}}\right)^2 + \left(\frac{a_{t-4} - \sim_{t-4,1}}{2\dagger_{t-4,1}}\right)^2\right\}\right] \\ &= \exp\left[-\left\{\left(\frac{-43,75 - 41,59}{(2)(102,72)}\right)^2 + \left(\frac{9,60 - 66,43}{(2)(79,58)}\right)^2\right\}\right] \\ &= 0,549 \\ w_2(a) &= \exp\left[-\left\{\left(\frac{a_{t-1} - \sim_{t-1,2}}{2\dagger_{t-1,2}}\right)^2 + \left(\frac{a_{t-4} - \sim_{t-4,2}}{2\dagger_{t-4,2}}\right)^2\right\}\right] \\ &= \exp\left[-\left\{\left(\frac{-57,68 + 36,99}{(2)(38,80)}\right)^2 + \left(\frac{9,60 + 50,19}{(2)(43,44)}\right)^2\right\}\right] \\ &= 0,382\end{aligned}$$

Nilai $w_j(a)$ yang diperoleh kemudian disubstitusikan pada persamaan *output*, sehingga diperoleh hasil peramalan *residual* untuk periode Januari 2015 adalah

$$\begin{aligned}\hat{b}_{109} &= (-8,6762) - (3,0499)(0,549) + (29,433)(0,382) \\ &= 1,326\end{aligned}$$

Nilai ramalan *residual* untuk periode Februari sampai Desember 2015 dapat ditentukan dengan cara yang sama. Hasil ramalan dapat dilihat pada Tabel 4.9 berikut ini.

Tabel. 4.9. Hasil Ramalan *RBFNN*

Periode		Hasil Ramalan
Bulan	t	<i>RBFNN</i>
Januari 2015	109	1,327
Februari	110	4,084
Maret	111	1,973
April	112	8,668
Mei	113	-3,279
Juni	114	-1,169
Juli	115	-1,036
Agustus	116	-0,829
September	117	0,449
Oktober	118	-0,526
Nopember	119	-0,307
Desember	120	-0,426

4. Prediksi *Hybrid*

Berdasarkan persamaan (2.17), nilai ramalan banyak kasus demam berdarah di Kota Surabaya untuk periode Januari sampai Desember 2015 ditentukan dengan menjumlahkan nilai ramalan model *INAR(1)* dengan x_1 (\hat{L}_{109}) dengan nilai ramalan model *RBFNN* (\hat{N}_{109}). Nilai ramalan untuk periode Januari 2015 adalah sebagai berikut:

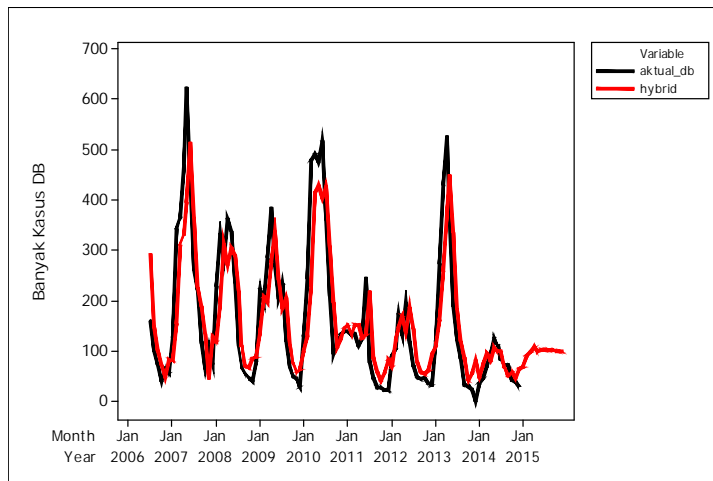
$$\begin{aligned}\hat{y}_{109} &= \hat{L}_{109} + \hat{N}_{109} \\ \hat{y}_{109} &= 68,024 + 1,327 \\ &= 69,351 \approx 70\end{aligned}$$

Jadi, banyak kasus demam berdarah di Kota Surabaya pada Januari 2015 diperkirakan sebanyak 70 kasus. Hasil peramalan banyak kasus demam berdarah di Kota Surabaya untuk periode Februari sampai Desember 2015 dapat dilihat pada Tabel 4.10 berikut.

Tabel 4.10. Hasil Ramalan *Hybrid INAR(1)* dengan x_t dan *RBFNN*

Periode Bulan	t	Hasil Peramalan		Hybrid = $\hat{L}_{109} + \hat{N}_{109}$	Pembulatan
		<i>INAR(2)</i>	<i>RBFNN</i>		
Januari	109	68	1,327	69,351	70
Februari	110	87	4,084	90,997	91
Maret	111	96	1,973	97,796	98
April	112	100	8,668	108,491	109
Mei	113	101	-3,279	98,154	99
Juni	114	102	-1,169	100,736	101
Juli	115	102	-1,036	100,819	101
Agustus	116	102	-0,829	100,757	101
September	117	101	0,449	101,691	102
Oktober	118	101	-0,526	100,366	101
Nopember	119	101	-0,307	100,257	101
Desember	120	100	-0,426	99,842	100

Gambar 4.16 adalah *plot* data aktual banyak kasus demam berdarah di Kota Surabaya dengan hasil ramalan data *training*, data *testing* dan periode Januari sampai Desember 2015 menggunakan model *Hybrid INAR(1)* dengan x_t dengan dan *RBFNN*. Hasil ramalan memberikan nilai *MSE* sebesar 371,494 dan *RMSE* 19,274.



Gambar 4.16. *Plot Data Aktual dengan Data Hasil Ramalan Hybrid*

Gambar 4.16 menunjukkan bahwa nilai hasil ramalan mendekati dan mengikuti pola data aktual. Jika diperhatikan, Gambar 4.16 tidak jauh berbeda dengan Gambar 4.12. Hal ini dapat diidentifikasi bahwa komponen linier pada proses linier lebih dominan dibanding dengan komponen nonliniernya. Pada titik-titik tertentu, hasil yang diberikan pada Gambar 4.16 memberikan nilai ramalan yang lebih mendekati nilai aktual.

G. Ramalan Hybrid untuk Model *INAR(2)*

Model *INAR(2)* tanpa variabel prediktor dan hasil estimasi parameternya adalah sebagai berikut:

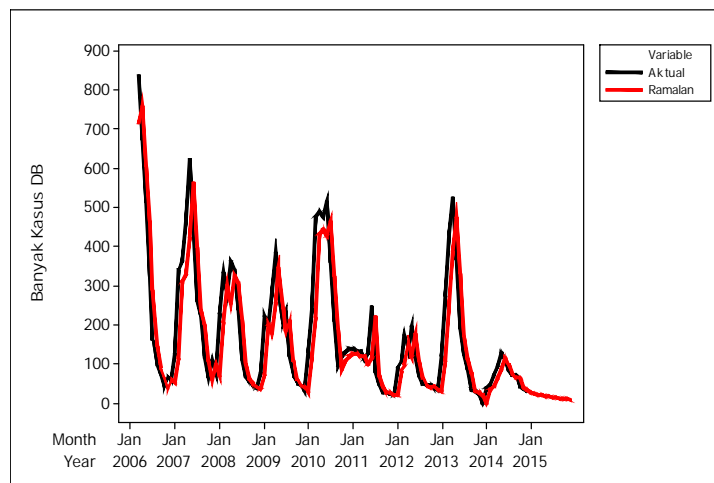
$$Y_t = 0,904 \circ Y_{t-1} + 1,04 \times 10^{-9} \circ Y_{t-2} + Z_t$$

Berdasarkan model tersebut, nilai ramalan untuk periode Januari sampai Desember 2015 dapat ditentukan dan ditunjukkan pada Tabel 4.11.

Tabel 4.11. Hasil Ramalan *INAR(2)*

Bulan	Banyak Kasus DB	Pembulatan
Januari 2015	26,216	27
Februari	23,699	24
Maret	21,424	22
April	19,367	20
Mei	17,508	18
Juni	15,827	16
Juli	14,308	15
Agustus	12,934	13
September	11,693	12
Oktober	10,570	11
Nopember	9,555	10
Desember	8,638	9

Plot data aktual dengan hasil ramalan data *training* dan data *testing* serta hasil ramalan periode Januari sampai Desember 2015 dapat dilihat pada Gambar 4.17.

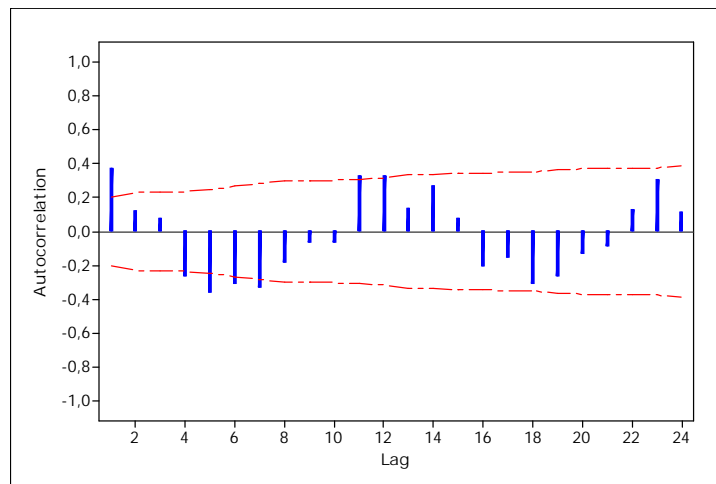


Gambar 4.17. *Plot* Data Aktual dengan Hasil Ramalan Model *INAR(2)*

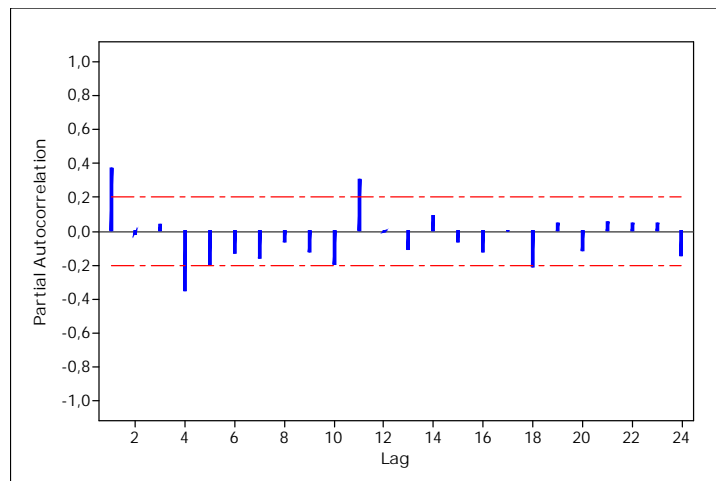
Hasil ramalan untuk data *training* dan data *testing* yang diperoleh kemudian digunakan untuk menghitung nilai *residual*. Nilai *residual* dari hasil peramalan model *INAR(2)* ini akan menjadi *input* untuk *Hybrid Radial Basis Function Neural Network (RBFNN)*.

1. Identifikasi Model *RFNN*

Setelah proses pemodelan data banyak kasus demam berdarah dilakukan menggunakan metode *INAR(2)*, kemudian dilakukan pemodelan menggunakan *RFNN*. Pada proses ini, data *input* yang digunakan adalah *residual* dari hasil prediksi proses *INAR(2)* dan dinotasikan dengan a_t . Tahap pertama yang dilakukan adalah identifikasi model dengan melihat *lag-lag* signifikan pada plot *ACF* dan *PACF* a_t .



Gambar 4.18. *Plot ACF* a_t



Gambar 4.19. *Plot PACF* a_t

Berdasarkan Gambar 4.18 dan 4.19, *lag-lag* yang signifikan pada plot *ACF* dan *PACF* a_t adalah *lag 1* dan *lag 4*. Oleh karena itu, *input* yang digunakan adalah a_{t-1} dan a_{t-4} , sehingga terdapat 2 *neuron* pada lapisan *input*.

2. Menentukan Nilai Pusat dan Varians

Pada proses ini, data *input* dikelompokkan menjadi beberapa kelompok menggunakan *Self Organizing Map (SOM)* sehingga nilai pusat dan varians setiap kelompok dapat ditentukan. Tabel 4.12 berikut adalah hasil pembelajaran *SOM* untuk 2 sampai 5 kelompok.

Tabel 4.12. Nilai *MSE* dan *RMSE* untuk 2 sampai 5 Kelompok

Banyaknya Kelompok	<i>MSE</i>	<i>RMSE</i>
2	417,742	20,4387
3	460,087	21,4497
4	1274,8	35,704
5	1003,7	31,6812

Berdasarkan Tabel 4.12, nilai *MSE* dan *RMSE* yang minimum ada pada kelompok 2 yaitu 417,6425 untuk *MSE* dan 20,4363 untuk *RMSE*, sehingga ditentukan banyak *neuron* pada lapisan tersembunyi adalah 2 *neuron*. Nilai pusat dan varians dari 2 kelompok dapat dilihat pada Lampiran 6. Jadi, model terbaik yang diperoleh pada penelitian ini adalah model *RBFNN* dengan 2 *neuron* pada lapisan *input*, 2 *neuron* pada lapisan tersembunyi dan 1 *neuron* pada lapisan *output*.

3. Prediksi *RBFNN*

Setelah model optimum terbentuk, langkah selanjutnya adalah melakukan peramalan. Arsitektur *RBFNN* yang telah terbentuk adalah arsitektur dengan 2 *neuron* pada lapisan *input*, 2 *neuron* pada lapisan tersembunyi dan 1 *neuron* pada lapisan *output* dapat dituliskan dengan

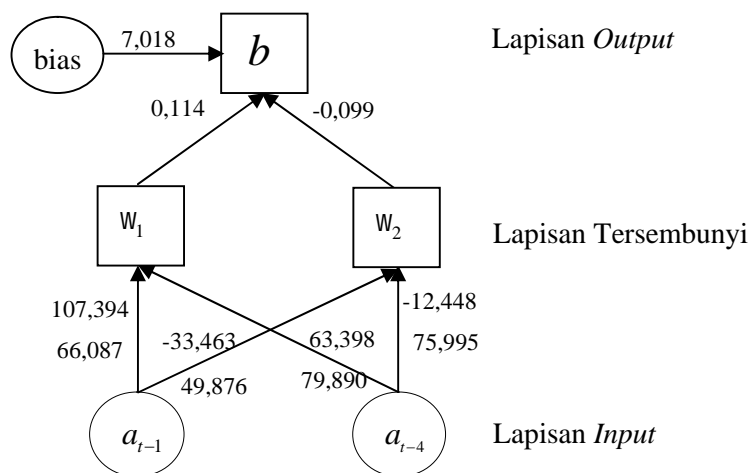
$$b_i = \sum_{j=1}^2 w_j w_j(a)$$

Pada proses pembelajaran *RBFNN*, diperoleh nilai-nilai bobot pada setiap penghubung *neuron* pada lapisan tersembunyi dengan *neuron* pada lapisan *output*. Banyak *neuron* pada lapisan tersembunyi yang terbentuk sebanyak 2 *neuron*, sehingga terdapat 2 bobot *neuron* dan 1 bobot bias yang ditunjukkan pada Tabel 4.13 berikut.

Tabel 4.13. Bobot Pembelajaran *RBFNN*

Bobot	Nilai
w_{bias}	7,0181
w_1	0,1144
w_2	-0,0993

Arsitektur model terbaik *RBFNN* dapat dilihat pada Gambar 4.20 berikut:



Gambar 4.20. Arsitektur Model terbaik *RBFNN*

Model dan bobot-bobot yang diperoleh dari pembelajaran *RBFNN* kemudian digunakan untuk perhitungan nilai ramalan periode Januari sampai Desember 2015. Perhitungan nilai ramalan untuk periode Januari 2015 (\hat{b}_{109}) adalah sebagai berikut:

$$\begin{aligned}\hat{b}_{109} &= \sum_{j=1}^2 w_j w_j(a) \\ &= 7,0181 + 0,1144w_1(a) - 0,0993w_2(a)\end{aligned}$$

Input yang digunakan pada model *RBFNN* adalah a_{t-1} dan a_{t-4} atau a_{108} dan a_{105} yang masing-masing mempunyai nilai -4,448 dan 6,816. Nilai *input*, nilai pusat (\sim_{ij}) dan varians (\dagger_{ij}^2) yang telah diperoleh digunakan untuk menentukan nilai $w_j(a)$, $i = t-1, t-4$ dan $j = 1, 2$ dengan perhitungan sebagai berikut:

$$\begin{aligned}w_j(a) &= \exp\left(-\sum_{i=t-1}^{t-4} \left(\frac{a_i - \sim_{ij}}{2\dagger_{ij}}\right)^2\right) \\ w_1(a) &= \exp\left[-\left\{\left(\frac{a_{t-1} - \sim_{t-1,1}}{2\dagger_{t-1,1}}\right)^2 + \left(\frac{a_{t-2} - \sim_{t-4,1}}{2\dagger_{t-4,1}}\right)^2\right\}\right] \\ &= \exp\left[-\left\{\left(\frac{-4,448 - 107,394}{(2)(66,086)}\right)^2 + \left(\frac{6,816 - 63,398}{(2)(79,890)}\right)^2\right\}\right] \\ &= 0,185\end{aligned}$$

$$\begin{aligned}w_2(a) &= \exp\left[-\left\{\left(\frac{a_{t-1} - \sim_{t-1,2}}{2\dagger_{t-1,2}}\right)^2 + \left(\frac{a_{t-2} - \sim_{t-4,2}}{2\dagger_{t-4,2}}\right)^2\right\}\right] \\ &= \exp\left[-\left\{\left(\frac{4,448 + 33,463}{(2)(49,875)}\right)^2 + \left(\frac{6,816 + 12,447}{(2)(75,995)}\right)^2\right\}\right] \\ &= 0,817\end{aligned}$$

Nilai $w_j(a)$ yang diperoleh kemudian disubstitusikan pada persamaan *output*, sehingga diperoleh hasil peramalan *residual* untuk periode Januari 2015 adalah

$$\begin{aligned}\hat{b}_{109} &= 7,0181 + (0,1144)(0,185) + (-0,0993)(0,817) \\ &= 6,9581\end{aligned}$$

Nilai ramalan *residual* untuk periode Februari sampai Desember 2015 dapat ditentukan dengan cara yang sama. Hasil ramalan dapat dilihat pada Tabel 4.14 berikut ini.

Tabel. 4.14. Hasil Ramalan *RBFNN*

Periode		Hasil Ramalan
Bulan	t	<i>RBFNN</i>
Januari 2015	109	6,958
Februari	110	6,967
Maret	111	6,973
April	112	6,972
Mei	113	6,977
Juni	114	6,977
Juli	115	6,977
Agustus	116	6,977
September	117	6,977
Oktober	118	6,977
Nopember	119	6,977
Desember	120	6,977

4. Prediksi *Hybrid*

Berdasarkan persamaan (2.17), nilai ramalan banyak kasus demam berdarah di Kota Surabaya untuk periode Januari sampai Desember 2015 ditentukan dengan menjumlahkan nilai ramalan model *INAR(2)* (\hat{L}_{109}) dengan nilai ramalan model *RBFNN* (\hat{N}_{109}). Nilai ramalan untuk periode Januari 2015 adalah sebagai berikut:

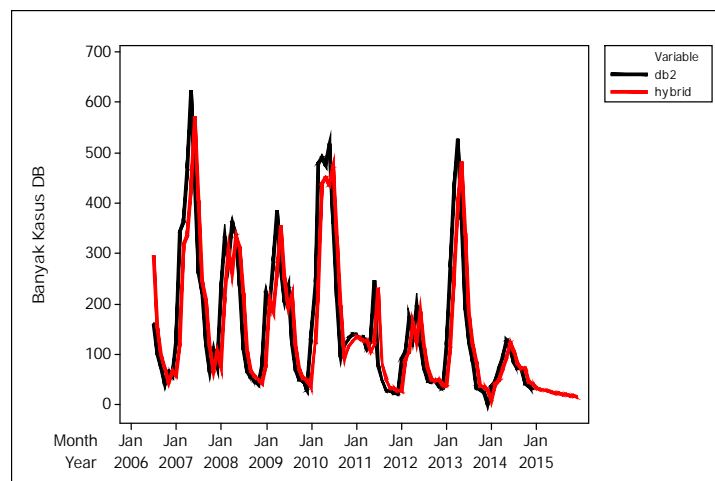
$$\begin{aligned}\hat{y}_{109} &= \hat{L}_{109} + \hat{N}_{109} \\ \hat{y}_{109} &= 26,22 + 6,958 \\ &= 33,174 \approx 34\end{aligned}$$

Jadi, banyak kasus demam berdarah di Kota Surabaya pada Januari 2015 diperkirakan sebanyak 34 kasus. Hasil peramalan banyak kasus demam berdarah di Kota Surabaya untuk periode Februari sampai Desember 2015 dapat dilihat pada Tabel 4.15 berikut.

Tabel 4.15. Hasil Ramalan *Hybrid INAR(2)* dan *RBFNN*

Periode Bulan	t	Hasil Peramalan		Hybrid = $\hat{L}_{109} + \hat{N}_{109}$	Pembulatan
		<i>INAR(2)</i>	<i>RBFNN</i>		
Januari	109	26,216	6,9582	33,174	34
Februari	110	23,699	6,9675	30,667	31
Maret	111	21,424	6,9735	28,398	29
April	112	19,367	6,9722	26,340	27
Mei	113	17,508	6,9770	24,485	25
Juni	114	15,827	6,9770	22,804	23
Juli	115	14,308	6,9770	21,285	22
Agustus	116	12,934	6,9770	19,911	20
September	117	11,693	6,9770	18,670	19
Oktober	118	10,570	6,9770	17,547	18
Nopember	119	9,555	6,9770	16,532	17
Desember	120	8,638	6,9770	15,615	16

Gambar 4.21 adalah *plot* data aktual banyak kasus demam berdarah di Kota Surabaya dengan hasil ramalan data *training*, data *testing* dan hasil ramalan periode Januari sampai Desember 2015 menggunakan model *Hybrid INAR(2)* dan *RBFNN*. Hasil ramalan memberikan nilai *MSE* sebesar 417,7425 dan *RMSE* 20,4387.



Gambar 4.21. *Plot* Data Aktual dengan Data Hasil Ramalan *Hybrid*

Gambar 4.21 menunjukkan bahwa nilai hasil ramalan mendekati dan mengikuti pola data aktual.

Berdasarkan hasil analisis peramalan *Hybrid INAR(1)* dengan x_1 -*RBFFNN* dan *Hybrid INAR(2)-RBFFNN*, dapat disimpulkan bahwa *Hybrid INAR(1)* dengan x_1 -*RBFFNN* memberikan hasil yang lebih baik. Hal ini ditunjukkan dengan nilai akurasi *MSE* dan *RMSE* yang lebih kecil dibanding *Hybrid INAR(2)-RBFFNN*. Pada penelitian ini, juga dilakukan peramalan banyak kasus demam berdarah di Kota Surabaya dengan model *RBFFNN*. *Input* yang digunakan pada proses ini adalah data banyak kasus demam berdarah di Kota Surabaya. Berdasarkan proses identifikasi model pada *INAR*, dapat dilihat bahwa *lag-lag* yang signifikan adalah *lag 1* dan *2*. Oleh karena itu, *input* yang digunakan adalah y_{t-1}, y_{t-2} . Pada proses ini, digunakan juga 6 variabel prediktor yang kemudian menjadi *input* pada pembelajaran *RBFFNN*. Jadi, pada model *RBFFNN* terdapat 8 *neuron* pada lapisan *input*.

Data *input* kemudian dikelompokkan menjadi beberapa kelompok dengan metode *Self Organizing Map (SOM)*. Banyak kelompok yang terbentuk merupakan banyak *neuron* pada kapisan tersembunyi. Pemilihan banyaknya kelompok yang akan digunakan dalam pembentukan model *RBFFNN* dilihat dari nilai *MSE* dan *RMSE* yang minimum. Pada Tabel 4.16 ditunjukkan banyaknya kelompok dengan nilai kebaikan model *MSE* dan *RMSE*nya.

Tabel 4.16. Nilai *MSE* dan *RMSE* untuk 2 sampai 12 kelompok

Banyaknya Kelompok	<i>MSE</i>	<i>RMSE</i>
2	15.351	123,90
3	14.753	121,46
4	15.155	123,10
5	15.928	126,21
6	16.808	129,64
7	16.823	129,70
8	16.009	126,53
9	16.127	126,99
10	16.944	130,17
11	13.428	115,88
12	15.449,0	124,30

Tabel 4.16 menunjukkan bahwa nilai *MSE* dan *RMSE* minimum terdapat pada 11 kelompok. Jadi, dapat ditentukan banyak *neuron* pada lapisan tersembunyi adalah 11 *neuron*. Berdasarkan hal ini, model *RBFNN* pada proses ini mempunyai model optimum dengan 8 *neuron* pada lapisan *input*, 11 *neuron* pada lapisan tersembunyi dan 1 *neuron* pada lapisan *output*. Nilai *MSE* dan *RMSE* yang dihasilkan relatif besar yaitu 13.428 untuk *MSE* dan 115,88 untuk *RMSE*.

Jika dibandingkan hasil peramalan data banyak kasus demam berdarah di Kota Surabaya menggunakan model *INAR(1)* dengan x_1 , *INAR(2)*, *RBFNN*, *Hybrid INAR(1)* dengan x_1 -*RBFNN*, dan *Hybrid INAR(2)-RBFNN* nilai kesalahan *MSE* dan *RMSE* terkecil dihasilkan oleh model *Hybrid INAR(1)* dengan x_1 -*RBFNN*. Sedangkan nilai *MSE* dan *RMSE* terbesar dihasilkan oleh model *RBFNN*. Berdasarkan hal tersebut, dapat dikatakan bahwa penyelesaian linier untuk data banyak kasus demam berdarah di Kota Surabaya dari Januari 2006 sampai Desember 2014 ini lebih baik dibandingkan dengan penyelesaian nonlinier. Pada penyelesaian linier dengan model *Hybrid INAR(1)* dengan x_1 , komponen *residual* yang dihasilkan dimodelkan menggunakan *RBFNN* yang kemudian model *Hybrid Hybrid INAR(1)* dengan x_1 -*RBFNN* memberikan hasil yang lebih baik.

Halaman ini sengaja dikosongkan.

Lampiran 1

Data Banyak Kasus Demam Berdarah di Kota Surabaya dari bulan Januari 2006 sampai Desember 2014

Bulan	Tahun								
	2006	2007	2008	2009	2010	2011	2012	2013	2014
Januari	557	126	228	222	128	138	90	118	36
Februari	790	342	333	196	242	132	105	273	46
Maret	841	364	280	284	477	133	173	430	71
April	672	468	361	384	490	110	131	526	94
Mei	513	624	338	268	475	126	196	367	127
Juni	320	441	237	204	515	246	118	189	110
Juli	160	261	109	234	360	78	70	123	82
Agustus	99	221	66	119	209	46	48	85	71
September	72	118	54	67	94	28	45	32	71
Oktober	40	65	45	49	119	28	46	30	42
Nopember	65	109	39	44	131	22	36	23	37
Desember	56	75	79	29	139	21	33	0	29

Lampiran 2

Data Kelembaban, Tekana Udara, Temperatur, Penyinaran Matahari, Curah Hujan, dan Kecepatan Angin per bulan di Kota Surabaya pada bulan Januari 2006 sampai Desember 2014.

Bulan	Kelembaban (%)	Tekanan Udara (Mbs)	Temperatur (0° C)	Penyinaran Matahari (%)	Curah Hujan (mm)	Kecepatan Angin (knot)
Januari (2006)	83,6	1008,3	27,1	47,4	377,8	7,0
Februari	85,1	1009,4	27,1	65,0	886,0	0,0
Maret	83,8	1009,0	27,3	50,1	288,2	7,1
April	79,3	1009,4	27,7	67,1	262,8	5,5
Mei	81,7	1010,7	27,5	71,3	228,5	5,1
Juni	76,5	1011,3	26,6	85,8	0,0	5,4
Juli	82,5	1012,9	25,9	98,7	0,0	8,0
Agustus	72,3	1012,9	26,1	99,0	0,0	7,4
September	67,0	1013,1	26,5	98,1	0,0	7,1
Oktober	66,3	1013,2	28,3	97,3	0,0	8,1
Nopember	65,2	1011,0	29,8	90,0	31,1	7,3
Desember	75,9	1009,5	29,0	63,6	147,9	4,9
...
Januari (2014)	81,0	1010,3	27,3	42,0	259,0	8,8
Februari	82,0	1009,8	27,2	40,0	247,0	8,5
Maret	81,0	1010,8	27,9	74,0	455,0	6,0
April	81,0	1010,4	28,3	54,0	273,0	5,8
Mei	77,0	1010,6	29,1	86,0	105,0	7,0
Juni	77,0	1010,3	28,5	89,0	202,0	7,0
Juli	76,0	1012,1	27,5	90,0	48,0	7,0
Agustus	72,0	1012,7	27,2	99,0	0,0	8,2
September	66,0	1013,3	26,4	67,0	0,0	7,6
Oktober	65,0	1012,3	29,0	98,0	0,0	8,0
Nopember	69,0	1010,6	29,9	81,0	72,0	6,9
Desember	81,0	1009,4	28,0	33,0	320,0	6,9

Lampiran 3

Sintaks Estimasi Parameter untuk model *INAR(2)* data banyak kasus DB di Kota Surabaya

```
function nonlinier
clc;

%Data Training n=94
data = dlmread('training.txt');
[n,p] = size(data);

y = data(:,1);
x = data(:,2:9);

%Model Nonlinier
modelfun = @(b,x)(x(:,1)/(1+exp(-b(8)))+(x(:,2)/(1+exp(-
b(9))))+(exp(b(1)+b(2)*x(:,3)+b(3)*x(:,4)+b(4)*x(:,5)+b(5)*x(:,6)+
b(6)*x(:,7)+b(7)*x(:,8))));

%Initial Value
beta0 = [0 0 0 0 0 0 0 0 0];

mdl = fitnlm(x,y,modelfun,beta0)

%Prediksi
xnew = x;
ypred = predict(mdl,xnew);

%Plot y vs ypred
dataplot = [y,ypred];
ts1 = timeseries(dataplot,1:n);
plot(ts1);

end
```

Lampiran 4

Sintaks *Self Organizing Maps (SOM)*

```
function somhybrid
clc;
rng(1);
data = dlmread('inputrbf.txt');
[n,p] = size(data);

%Data Training n=83
data1 = data(1:83,2:3);
[n1,p1] = size(data1);

p = transpose(data1);

%Membangun Jaringan
JumlahKlas = 13;

%Proses Pembelajaran
net = newc([-1000 1000; -1000 1000], JumlahKlas);

%Menentukan Bobot Awal
Bobot_Input = net.IW{1,1};
Bobot_Bias_Input = net.b{1,1};

%Pelatihan
net.trainParam.epochs = 1000;
net.trainParam.goal = 0.001;
net.trainParam.lr = 0.01;
net = train(net,p);

%Menentukan Bobot Akhir
BobotAkhir_Input = net.IW{1,1};
BobotAkhir_Input = transpose(BobotAkhir_Input)
BobotAkhir_Bias_Input = net.b{1,1};

%Simulasi
yhat1 = sim(net,p);
yhat1 = vec2ind(yhat1);
yhat1 = transpose(yhat1)

%Data Testing n=12
data2 = data(84:n,2:3);
[n2,p2] = size(data2);

p2 = transpose(data2);

%Simulasi
yhat2 = sim(net,p2);
yhat2 = vec2ind(yhat2);
yhat2 = transpose(yhat2)

end
```

Lampiran 5

Sintaks *Radial Basis Function Neural Network (RBFNN)* dengan 2 neuron input, 2 neuron tersembunyi dan 1 output.

```
function input_RBF

clc;

data=dlmread('inputrbf.txt');
[n,p]=size(data);
kluster=dlmread('klas2.txt');

%nilai mean dan standar deviasi
mean=kluster(1:2,:);
SD=kluster(3:4,:);

%Data Training n=84
Y=data(1:83,1);
[n2,p2]=size(Y);
X=data(1:83,2:3);
X=transpose(X);

%Data Testing n=12
Yt=data(84:n,1);
[n1,p1]=size(Yt);
Xt=data(84:n,2:3);
Xt=transpose(Xt);

%Prediksi Data Training
H=rbfDesign(X,mean,SD,'b');
lambda=globalRidge(H,Y,0.05);
w=inv(transpose(H)*H+lambda*eye(3))*transpose(H)*Y;
Yhat=H*w;

%Prediksi Data Testing
Ht=rbfDesign(Xt,mean,SD,'b');
Yhat2=Ht*w;

%Kriteria Model Terbaik
for i = 1:n1
e(i) = Yt(i)-Yhat2(i);
end

sse = sum(e.^2);
mse = sse/n1
rmse= sqrt(mse)

end
```

Lampiran 6

Nilai Pusat dan Varians model *RBFFNN*

Arsitektur model *RBFFNN* yang terbentuk adalah dengan 2 *neuron input*, 2 *neuron* tersembunyi dan 1 *output* untuk *Hybrid INAR(1)* dengan x_1 -*RBFFNN*.

	<i>Input</i>	kelompok 1	kelompok 2
Pusat/	t-1	41,5866	-36,9875
Rata-rata	t-4	66,4331	-50,1864
Standar	t-1	102,7254	38,80422
deviasi	t-4	79,58322	43,43944

Arsitektur model *RBFFNN* yang terbentuk adalah dengan 2 *neuron input*, 2 *neuron* tersembunyi dan 1 *output* untuk *Hybrid INAR(2)*-*RBFFNN*.

	<i>Input</i>	kelompok 1	kelompok 2
Pusat/	t-1	107,394	-33,4632
Rata-rata	t-4	63,3981	-12,4477
Standar	t-1	66,08683	49,87598
deviasi	t-4	79,89031	75,99549

Lampiran 7

Uji *Dickey Fuller* untuk Stasioneritas Data Banyak Kasus DB di Surabaya.

```
> yt=scan("datadb.txt")
Read 108 items
> adf.test(yt)

      Augmented Dickey-Fuller Test

data:  yt
Dickey-Fuller = -5.8805, Lag order = 4, p-value = 0.01
alternative hypothesis: stationary

Warning message:
In adf.test(yt) : p-value smaller than printed p-value
```

Lampiran 8

Hasil Ramalan Variabel Prediktor untuk Tahun 2015

Bulan	Kelembaban (%)	Temperatur (C)
Januari	80,7	27,4
Februari	80,5	27,4
Maret	80,3	27,9
April	80,1	28,3
Mei	80,0	28,7
Juni	79,8	27,9
Juli	79,7	27,1
Agustus	79,6	26,8
September	79,5	26,8
Oktober	79,5	29,1
Nopember	79,4	29,5
Desember	79,3	27,8

Halaman ini sengaja dikosongkan.

BAB 5

KESIMPULAN DAN SARAN

A. Kesimpulan

Berdasarkan pembahasan dan proses penelitian untuk meramalkan data banyak kasus demam berdarah di Kota Surabaya dengan 6 variabel prediktor untuk 12 periode ke depan, diperoleh kesimpulan sebagai berikut:

1. Model *INARIMA* terbaik yang terbentuk dari data banyak kasus demam berdarah di Kota Surabaya adalah *INAR(1)* dengan variabel prediktor x_1 . Hasil ramalan model *INAR(1)* dengan x_1 menghasilkan nilai *MSE* 440,458 dan *RMSE* 20,987.
2. Model *RBFNN* yang terbentuk dari data banyak kasus demam berdarah di Kota Surabaya adalah model dengan 8 *neuron input*, 11 *neuron tersembunyi* dan 1 *output*. Nilai *MSE* dan *RMSE* yang dihasilkan masing-masing yaitu 13.428 dan 115,88.
3. Model *Hybrid INAR(1)* dengan x_1 dan *RBFNN* untuk data banyak kasus demam berdarah di Kota Surabaya memberikan hasil ramalan dengan nilai *MSE* sebesar 371,494 dan *RMSE* 19,274.

Berdasarkan 3 kesimpulan yang diperoleh, untuk peramalan banyak kasus demam berdarah di Kota Surabaya, nilai kesalahan *MSE* dan *RMSE* terkecil adalah untuk model *Hybrid INAR(1)* dengan x_1 dan *RBFNN*. Sedangkan nilai *MSE* dan *RMSE* terbesar ada pada model *RBFNN*. Berdasarkan hal tersebut, dapat dikatakan bahwa penyelesaian linier untuk data banyak kasus demam berdarah di Kota Surabaya dari Januari 2006 sampai Desember 2014 ini lebih baik dibandingkan dengan penyelesaian nonlinier. Pada penyelesaian linier dengan model *Hybrid INAR(1)* dengan x_1 , nilai *residual* yang dihasilkan dimodelkan menggunakan *RBFNN* yang kemudian

hasil peramalan menggunakan model *Hybrid INAR(1)* dengan x_1 dan *RBFNN* memberikan hasil yang lebih baik.

B. Saran

Pada penelitian ini, model *Hybrid* yang digunakan adalah model yang direkomendasikan oleh Zhang (2003), pada penelitian selanjutnya, dapat dikembangkan menggabungkan 2 model dengan model *Hybrid* yang lain. Estimasi parameter pada model *INAR(2)* dilakukan menggunakan metode *Conditional Least Square (CLS)*, untuk penelitian selanjutnya dapat digunakan metode lain seperti metode Bayes, *Marcov Chain Monte Carlo (MCMC)*.

DAFTAR PUSTAKA

- Badan Pusat Statistika Provinsi Jawa Timur (BPS Jatim). 2007. *Surabaya dalam Angka 2007*. Surabaya: BPS.
- _____. 2008. *Surabaya dalam Angka 2008*. Surabaya: BPS.
- _____. 2009. *Surabaya dalam Angka 2009*. Surabaya: BPS.
- _____. 2010. *Surabaya dalam Angka 2010*. Surabaya: BPS.
- _____. 2011. *Surabaya dalam Angka 2011*. Surabaya: BPS.
- _____. 2012. *Surabaya dalam Angka 2012*. Surabaya: BPS.
- _____. 2013. *Surabaya dalam Angka 2013*. Surabaya: BPS.
- _____. 2014. *Surabaya dalam Angka 2014*. Surabaya: BPS.
- _____. 2015. *Surabaya dalam Angka 2015*. Surabaya: BPS.
- Bartoszynski, Robert and Bugaj, Magdalena N.. 2007. *Probability and Statistical Inference, 2nd Edition*. Hoboken, New Jersey: Wiley.
- Cameron, A. Colon and Trivedi, Pravin K. 1998. *Regression Analysis of Count Data*. Cambridge: Cambridge University Press.
- Dinas Kesehatan Provinsi Jawa Timur (Dinkes Jatim). 2015. *Daftar Penderita dan Kematian DBD Provinsi Jawa Timur Tahun 2014*. Surabaya.
- Fausett, Laurence. 1994. *Fundamentals of Neural Networks (Architectures, Algorithms and Applications)*. Upper Saddle River, New Jersey: Prentice Hall.
- Fithriasari, K., Iriawan, N., Ulama, B.S.S. 2013. On The Multivariate Time Series Rainfall Modeling Using Time Delay Neural Network. *International Journal of Applied Mathematics and Statistics*TM. **44**, 193-201.
- Hanke, J.E. and Wichern, D.W.. 2005. *Business Forecasting, 8th Edition*. Upper Saddle River, New Jersey: Prentice Hall.

- Hanrahan, Grady. 2011. *Artificial Neural Networks in Biological and Environmental Analysis*. London: CRC Press.
- Kedem, Benjamin and Fakianos, Konstantinos. 2002. *Regression Models for Time Series Analysis*. Canada: John Wiley and Son, Inc.
- Kementrian Kesehatan Republik Indonesia (Kemenkes RI). 2010. *Buletin Jendela Epidemiologi*. Jakarta.
- _____. 2014. *Petunjuk Teknis Jumantik – DSN Anak sekolah*. Jakarta.
- _____. 2015. *Profil Kesehatan Indonesia 2014*. Jakarta.
- Khandewal, Ina., Adhikari, Ratnadip and Verma, Ghanshyam. 2005. Time Series Forecasting using Hybrid ARIMA and ANN Models based on DWT Decomposition. *Procedia Computer Science*. **48**, 173-179.
- Kim, Hee-Young and Park, Yousung. 2010. Markov Chain Approach to Forecast in the Binomial Autoregressive Models. *Communication of the Korean Statistical Society*. **17**, 441-450.
- Kiusalaas, Joan. 2005. *Numerical Methode in Engineering with Matlab*. Cambridge: Cambridge University Press.
- Kusumadewi, Sri. 2004. *Membangun jaringan Saraf Tiruan Menggunakan Matlab dan Excel Link*. Yogyakarta: Graha Ilmu.
- Mafrida, Arma Fauziyatul. 2012. Pemodelan Prediksi Jumlah Kasus Penderita Demam Berdarah Dengue (DBD) di Kota Surabaya dengan Metode Integer-valued Autoregressive Moving Average (INARMA). *Skripsi*. ITS.
- Makridakis, Spyros G., Wheelwright, Steven C. and Hyndman, Rob J.. 1997. *Forecasting: Methods and Application, 3th Edition*. New York: Wiley.
- Montgomery, Douglas C., Jennings, Cheryl and Kulachi, Murat. 2008. *Introduction to Time Series Analysis and Forecasting*. Hoboken, New Jersey: Willey.
- Orr, Mark J.L.. 1996. *Introduction to Radial Basis Function Networks*. Edinburg, Scotland: University of Edinburg.
- Palit, A.K. and Papovic, D.. 2005. *Computational Intellegence in Time Series Forecasting*. Glasgow: Springer.
- Park, J., Sandberg, I. (1991). Universal Approximation using Radial Basis Function Networks. *Neural Computation*. **3**, 246-257.

- Pedeli, Xanthi and Karlis, Dimitris. 2013. Some Properties of Multivariate INAR(1) Processes. *Computational Statistic and Data Analysis*. **67**, 213-225.
- Riyadi, M.A.A., Fithriasari, K., Widodo, D.A. 2016. Data Mining Peramalan Konsumsi Listrik dengan Pendekatan Cluster Time Series sebagai Preprocessing. *Jurnal Sains dan Seni ITS*. **5**, D121-D126.
- Shukur, Osamah B. and Lee, Muhammad H.. 2015. Daily Wind Speed Forecasting Through Hybrid KF-NN Model based on ARIMA. *Renewable Energy*. **76**. 637-647.
- Silva, Isabel., Silva, M. Eduarda., Pereira, Isabel., and Silva, Nelia. 2005. Replicated INAR(1) Process. *Methodology and Computing in Applied Probability*. **7**, 517-542.
- Silva, Nelia., Pereira, Isabel and Silva, M. Eduarda. 2009. Forecasting in INAR(1) Model. *Revstat – statistical Journal*. **7**, 829-851.
- Sugiartawan, Putu dan Mauko, Arfan. 2015. Optimasi Radial Basis Function Neural Network dengan Growing Hierarchical Self Organizing Map pada Data Time Series. *Jurnal Keamanan Jaringan*. Vol 1, hlm. 32-37.
- Supartha, I.W.. 2008. *Pengendalian Terpadu Vektor Virus Demam Berdarah Dengue*. Universitas Udayana.
- Tatar, Afshin., Naseri, Saeid., Sirach, Nick., Lee, Moonyang., Bahadori, Alireza .2015. Prediction of Reservoir Brine Properties using RBFNN. *Petroleum*. **1**, 1-9.
- Turkman, Kamil Feridun., Scotto, Manuel Gonzales., and Bermudez, Patricia de Zea. 2014. *Non-Linear Time Series Extreme Event and Integer Value Problems*. New York: Springer.
- Wei, W.W.S.. 2006. *Time Series Analysis Univariate and Multivariate Method, 2nd Edition*. New York: Pearson Education.
- Weib, Christian H.. 2015. A Poisson INAR(1) Model with Serially Dependent Innovations. *Metrika*. **78**, 829-851.
- Wiyanti, D.T. dan Pulungan, R.. 2012. Peramalan Deret Waktu Menggunakan Model Fungsi Basis Radial (RBF) dan Autoregressive Moving Average (ARIMA). *Jurnal MIPA*. Vol. 35, hlm. 175-182.

- Yeung, Daniel S., Cloete, Ian., Shi, Daming., and Ng, Wing W.Y.. 1998. *Sensitivity Analysis for Neural Network*. London: Springer.
- Zahrati, Z. dan Fithriasari, K. 2016. Multi-output Neural Network for The Temperature Forecasting in Semarang. *The 2016 Conference on Fundamental and Applied Science for Advanced Technology (CONFAST 2016)*. **1746**, 020040.
- Zhang, G.P.. 2003. Time Series Forecasting using a Hybrid ARIMA and Neural Networks Model. *Neurocomputing*. **50**, 159-175.
- Zhu, Jia Zeng., Cao, Jin Xin and Zhu, Yuan. 2014. Traffic Volume Forecasting based on Radial Basis Function Neural Network with the Consideration of traffic Flows at Adjacent Intersection. *Transportasi Research Part C*. **47**, 139-154.

BIOGRAFI PENULIS



Fajarani Juliaristi, penulis yang disapa dengan nama Risti ini dilahirkan di Gerung, Lombok Barat pada tanggal 9 Juli 1993 dari ayah yang bernama Khotib Anwari dan ibu bernama Sulistiyasih. Penulis merupakan anak pertama dari dua bersaudara. Penulis menyelesaikan Sekolah Dasar di SD Negeri 6 Lembar pada tahun 1999-2005. Kemudian penulis melanjutkan pendidikan di SMP Negeri 1 Gerung dan lulus pada tahun 2008. Penulis melanjutkan pendidikannya di SMA Negeri 5 Mataram pada tahun 2008-2010. Setelah lulus SMA penulis diterima di Program Studi Matematika, Universitas Negeri Yogyakarta (UNY) dan lulus sebagai seorang sarjana pada tahun 2014. Kemudian penulis melanjutkan pendidikan Pasca Sarjana (S2) di Institut Teknologi Sepuluh Nopember Program Studi Statistika. Penulis dapat dihubungi melalui email: fajaranijuliaristi@gmail.com.