



TESIS - ST 2309

**PEMODELAN *STATISTICAL DOWNSCALLING* DENGAN  
*PROJECTION PURSUIT REGRESSION* DAN *ADAPTIVE SPLINE  
THRESHOLD AUTOREGRESSION* UNTUK PERAMALAN  
CURAH HUJAN**

NOVIYANTI SANTOSO  
NRP. 1312201027

DOSEN PEMBIMBING:  
Dr. Sutikno, M.Si.  
Dr. I Nyoman Latra, M.S.

PROGRAM MAGISTER  
JURUSAN STATISTIKA  
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
INSTITUT TEKNOLOGI SEPULUH NOPEMBER  
SURABAYA  
2014



TESIS - ST 2309

**STATISTICAL DOWNSCALLING MODELLING WITH  
PROJECTION PURSUIT REGRESSION AND ADAPTIVE SPLINE  
THRESHOLD AUTOREGRESSION TO FORECASTING  
RAINFALL**

**NOVIYANTI SANTOSO  
NRP. 1312201027**

**SUPERVISOR:  
Dr. Sutikno, M.Si  
Dr. I Nyoman Latra, MS**

**MASTER DEGREE PROGRAM  
DEPARTEMENT OF STATISTICS  
FACULTY OF MATHEMATICS AND NATURAL SCIENCE  
SEPULUH NOPEMBER INSTITUTE OF TECHNOLOGY  
SURABAYA  
2014**

# **PEMODELAN *STATISTICAL DOWNSCALING* DENGAN *PROJECTION PURSUIT REGRESSION* DAN *ADAPTIVE SPLINE THRESHOLD AUTOREGRESSION* UNTUK PERAMALAN CURAH HUJAN**

**Nama Mahasiswa** : Noviyanti Santoso  
**NRP** : 1312201027  
**Dosen Pembimbing** : Dr. Sutikno, M.Si  
**Co-Pembimbing** : Dr. I Nyoman Latra, MS

## **ABSTRAK**

Penyediaan data iklim merupakan hal yang penting sebagai upaya untuk mengetahui perubahan iklim di suatu wilayah. Sampai saat ini, *general circulation model* (GCM) diakui banyak pihak sebagai alat untuk upaya memahami sistem iklim. GCM merupakan alat prediksi utama iklim dan cuaca secara numerik serta sebagai informasi primer untuk menilai pengaruh perubahan iklim. Namun informasi GCM yang dihasilkan masih berskala global (ratusan km), sehingga sulit untuk mendapatkan informasi skala regional. Salah satu upaya untuk mendapatkan informasi skala lokal perlu digunakan *statistical downscaling* (SD). Metode SD adalah memprediksi curah hujan lokal beresolusi tinggi berdasarkan data *General Circulation Model* (GCM) berskala global melalui model regresi. Data GCM merupakan data spasial dan temporal yang memungkinkan terjadi korelasi spasial antar *grid* yang berbeda dalam satu domain dan autokorelasi dalam data deret waktu. Selain itu bentuk fungsi regresi yang tidak diketahui juga menjadi permasalahan tersendiri. Pada penelitian ini, metode SD yang digunakan adalah *Projection Pursuit Regression* (PPR) dan *Adaptive Spline Threshold Autoregression* (ASTAR). Model dengan metode ASTAR digunakan untuk menyusun model error dari hasil model PPR. Model gabungan antara PPR dan ASTAR selanjutnya disebut model hybrid. Oleh karena itu dua pendekatan model yang digunakan, yaitu model hybrid dan non hybrid. Kriteria kebaikan model untuk validasi menggunakan nilai *Root Mean Square Error Prediction* (RMSEP) dan nilai  $R_{pred}^2$  atau disebut dengan koefisien determinasi. Lokasi penelitian di sentra produksi padi, yaitu Kabupaten Banyuwangi, Jember, Bojonegoro, Ngawi, dan Lamongan. Hasil penelitian menunjukkan bahwa model hybrid PPR dan ASTAR mempunyai kinerja yang lebih baik daripada non hybrid, jika terjadi autokorelasi. Hal ini ditunjukkan pada nilai RMSEP model *hybrid* lebih rendah daripada model *non hybrid* baik untuk data *in-sample* maupun *out-sample*.

**Kata Kunci** : ASTAR, GCM, *Statistical Downscaling*, PPR, Reduksi dimensi

# STATISTICAL DOWNSCALING MODELLING WITH PROJECTION PURSUIT REGRESSION AND ADAPTIVE SPLINE THRESHOLD AUTOREGRESSION TO FORECASTING RAINFALL

**Name** : Noviyanti Santoso  
**NRP** : 1312201027  
**Supervisor** : Dr. Sutikno, M.Si.  
**Co-Supervisor** : Dr. I Nyoman Latra, M.S.

## ABSTRACT

*Climate data supplying is one of an important things as an effort to detect climate change in some certain areas. Till now days, general circulation model (GCM) is admitted as a device to understand the climate system. GCM is a major prediction device numerically for weather and climate also as primary information to verify the effect of climate change. Yet, the result of GCM information is still in global scale (hundred kilometers), so it is difficult to gain information in regional scale. One of the efforts to gain the local scale information is needed the usage of statistical downscaling (SD). SD method is able to predict local fall of rain with high resolution based on GCM global scale's data through regression model. GCM data is temporal and spatial data which enable a spatial correlation between different grid in one domain and autocorrelation in time line data to happen. Besides, the unknown form of regression function is also becoming a problem. In this research, SD method which is used are Projection Pursuit Regression (PPR) for dimension reduction and Adaptive Spline Threshold Autoregression (ASTAR). Model with ASTAR method is used to arrange error model from PPR model result. Combination model between PPR and ASTAR is known as hybrid model. Hence, there are two approaching model which is used, hybrid and non hybrid model. The criteria of a good model for validation uses Root Mean Square Error Prediction (RMSEP) and  $R_{pred}^2$  grades or known as determination coefficient. This research takes place at rice production center, in Banyuwangi, Jember, Bojonegoro, Ngawi, and Lamongan regency. Research result shows that hybrid model of PPR and ASTAR have a better work than non hybrid, if there is autocorrelation. This is shown from RMSEP's grade with hybrid model lower than non hybrid, as in-sample or out-samples data.*

**Keywords** : ASTAR, GCM, Statistical Downscaling, PPR, Dimension reduction.

**PEMODELAN STATISTICAL DOWNSCALLING DENGAN PROJECTION  
PURSUIT REGRESSION DAN ADAPTIVE SPLINE THRESHOLD  
AUTOREGRESSION UNTUK PERAMALAN CURAH HUJAN**

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

**di  
Institut Teknologi Sepuluh Nopember  
oleh:**

**NOVIYANTI SANTOSO  
NRP 1312 201 027**


**Tanggal Ujian : 21 Juli 2014**

**Periode Wisuda : September 2014**

**Disetujui oleh:**

  
**1. Dr. Sutikno, M.Si.  
NIP 19710313 199702 1 001**

**(Pembimbing)**

  
**2. Dr. I Nyoman Latra, M.S.  
NIP 19511130 197901 1 001**

**(Pembimbing)**

  
**3. Dr. Purhadi, M.Sc.  
NIP 19620204 198701 1 001**

**(Penguji)**

  
**4. Dr. Bambang Widjanarko Otok, M.Si.  
NIP 19681124 199412 1 001**

**(Penguji)**

  
**5. Dr. Santi Wulan Purnami, M.Si.  
NIP 19720923 199803 2 001**

**(Penguji)**

**Direktur Program Pascasarjana,**

  
**Prof. Dr. Ir. Adi Soeprijanto, M.T.  
NIP 19640405 199002 1 001**

## KATA PENGANTAR

Puji syukur penulis panjatkan kehadiran Allah SWT, yang atas rahmat-Nya maka penulis dapat menyelesaikan penyusunan laporan Tesis yang berjudul **PEMODELAN *STATISTICAL DOWNSCALLING* DENGAN *PROJECTION PURSUIT REGRESSION* DAN *ADAPTIVE SPLINE THRESHOLD AUTOREGRESSION* UNTUK MERAMALKAN CURAH HUJAN.**

Dalam penulisan laporan ini penulis ucapkan terima kasih kepada pihak-pihak yang membantu dalam menyelesaikan rang-kaiannya dalam proses Tesis ini, khususnya kepada :

1. Bapak Dr. Muhammad Mashuri, M.T. selaku Ketua Jurusan Statistika ITS.
2. Bapak Dr. Sutikno, M.Si dan Bapak Dr. I Nyoman Latra, M.S selaku dosen pembimbing Tesis yang telah membantu proses pelaksanaan Tesis dari awal hingga akhir.
3. Bapak, ibu, mama, suami, dan anak yang telah mendoakan dan selalu memotivasi untuk menyelesaikan Tesis.
4. Dek Iis, dek Elvira, dek Novianti selaku teman seperjuangan di bawah bimbingan Bapak Sutikno yang selalu membantu dalam segala hal.
5. Ayu, Mbak Erna, Gusmi, dek Rina yang selalu memberi informasi dan semangat selama kuliah hingga penyelesaian Tesis.
6. Teman-teman S2 Jurusan Statistika yang banyak membantu proses pelaksanaan Tesis ini.
7. Dan semua pihak yang tidak dapat disebutkan satu-persatu yang telah membantu hingga pelaksanaan Tesis ini dapat terselesaikan dengan baik.

Dalam Penulisan laporan ini penulis merasa masih banyak kekurangan-kekurangan baik pada teknis penulisan maupun materi, mengingat akan kemampuan yang dimiliki penulis. Untuk itu kritik dan saran dari semua pihak sangat penulis harapkan demi penyempurnaan pembuatan laporan ini.

Akhir kata, penulis berharap semoga Allah memberikan imbalan yang setimpal pada mereka yang telah memberikan bantuan, dan dapat menjadikan semua bantuan ini sebagai ibadah, Amiin Yaa Robbal ‘Alamiin.

Surabaya, Juni 2014

Penulis

## DAFTAR ISI

	halaman
<b>HALAMAN JUDUL</b> .....	i
<b>LEMBAR PENGESAHAN</b> .....	iii
<b>ABSTRAK</b> .....	v
<b>ABSTRACT</b> .....	vii
<b>KATA PENGANTAR</b> .....	ix
<b>DAFTAR ISI</b> .....	xi
<b>DAFTAR GAMBAR</b> .....	xiii
<b>DAFTAR TABEL</b> .....	xv
<b>DAFTAR LAMPIRAN</b> .....	xvii
<b>BAB I PENDAHULUAN</b>	
1.1 Latar Belakang .....	1
1.2 Permasalahan Penelitian .....	4
1.3 Tujuan Penelitian .....	4
1.4 Manfaat Penelitian .....	4
1.5 Batasan Masalah .....	5
<b>BAB II TINJAUAN PUSTAKA</b>	
2.1 <i>Projection Pursuit Regression</i> (PPR) .....	7
2.2 <i>Adaptive Spline Threshold Autoregression</i> (ASTAR) .....	14
2.3 Model <i>Hybrid</i> PPR dan ASTAR .....	16
2.4 Validasi Model .....	16
2.5 Metode Deteksi Autokorelasi Residual .....	17
2.6 <i>General Circulation Models</i> (GCM) .....	18
2.7 Teknik <i>Statistical Downscaling</i> .....	19
2.8 Curah Hujan .....	20
<b>BAB III METODOLOGI PENELITIAN</b>	
3.1 Sumber Data .....	23
3.2 Variabel Penelitian .....	23



3.3 Metode Analisis .....	24
<b>BAB IV ANALISIS DATA DAN PEMBAHASAN</b>	
4.1 Deskripsi Curah Hujan di Lima Kabupaten .....	29
4.2 Pemodelan SD dengan <i>Projection Pursuit Regression</i> (PPR) .....	33
4.3 Penyusunan Model <i>Hybrid</i> .....	37
4.3.1 Identifikasi Autokorelasi Residual .....	37
4.3.2 Penyusunan Model Residual dengan ASTAR .....	39
4.3.3 Validasi Model <i>Hybrid</i> PPR dan ASTAR.....	41
<b>BAB V KESIMPULAN DAN SARAN</b>	
5.1 Kesimpulan.....	45
5.2 Saran.....	46
<b>DAFTAR PUSTAKA</b> .....	47
<b>LAMPIRAN</b> .....	51

## DAFTAR GAMBAR

	halaman
Gambar 2.1 Ilustrasi Geometrik Proyeksi $X$ Terhadap $Z$ .....	9
Gambar 2.2 Diagram Alir Algoritma PPR .....	12
Gambar 2.3 Ilustrasi Teknik <i>Statistical Downscaling</i> .....	20
Gambar 2.4 Pola Curah Hujan di Indonesia.....	22
Gambar 3.1 Penentuan Pemilihan Grid di Setiap Lokasi Pengamatan .....	24
Gambar 3.2 Diagram Alir Penelitian .....	26
Gambar 4.1 <i>Boxplot</i> Curah Hujan di Kabupaten (a)Banyuwangi, (b)Jember, (c)Bojonegoro, (d)Lamongan, dan (e)Ngawi.....	30
Gambar 4.2 <i>Time Series Plot</i> Curah Hujan di Kabupaten (a)Banyuwangi, (b)Jember, (c)Bojonegoro, (d)Lamongan, dan (e)Ngawi.....	31
Gambar 4.3 <i>Plot</i> Antara Variabel Respon dan Variabel Prediktor di Kabupaten Banyuwangi .....	32
Gambar 4.4 <i>Plot</i> Antara Variabel Respon dan Variabel Prediktor di Kabupaten Bojonegoro.....	32
Gambar 4.5 <i>Plot</i> Antara Variabel Respon dan Variabel Prediktor di Kabupaten Jember .....	33
Gambar 4.6 <i>Plot</i> Antara Observasi dan Prediksi Curah Hujan di Kabupaten Banyuwangi dengan PPR.....	36
Gambar 4.7 <i>Plot</i> Antara Observasi dan Prediksi Curah Hujan di Kabupaten Jember dengan PPR.....	36
Gambar 4.8 <i>Plot</i> Antara Observasi dan Prediksi Curah Hujan di Kabupaten Bojonegoro dengan PPR .....	36
Gambar 4.9 <i>Plot</i> Antara Observasi dan Prediksi Curah Hujan di Kabupaten Lamongan dengan PPR .....	37
Gambar 4.10 <i>Plot</i> Antara Observasi dan Prediksi Curah Hujan di Kabupaten Ngawi dengan PPR .....	37
Gambar 4.11 <i>Plot</i> ACF Residual Curah Hujan di Kabupaten (a)Banyuwangi, (b)Jember, (c)Bojonegoro, (d)Lamongan, dan (e)Ngawi.....	38

Gambar 4.12	<i>Plot Antara Observasi dan Prediksi Curah Hujan di Kabupaten Banyuwangi dengan PPR dan Hybrid ASTAR</i> .....	42
Gambar 4.13	<i>Plot Antara Observasi dan Prediksi Curah Hujan di Kabupaten Jember dengan PPR dan Hybrid ASTAR</i> .....	43
Gambar 4.14	<i>Plot Antara Observasi dan Prediksi Curah Hujan di Kabupaten Bojonegoro dengan PPR dan Hybrid ASTAR</i> .....	43
Gambar 4.15	<i>Plot Antara Observasi dan Prediksi Curah Hujan di Kabupaten Lamongan dengan PPR dan Hybrid ASTAR</i> .....	44
Gambar 4.16	<i>Plot Antara Observasi dan Prediksi Curah Hujan di Kabupaten Ngawi dengan PPR dan Hybrid ASTAR</i> .....	44

## DAFTAR TABEL

	halaman
Tabel 3.1	Lokasi Stasiun Pengamatan.....23
Tabel 4.1	Nilai RMSEP dan $R^2_{prediction}$ Menurut Banyaknya Fungsi Di Lima Kabupaten .....34
Tabel 4.2	Model PPR Data Curah Hujan di Lima Kabupaten.....35
Tabel 4.3	Nilai ACF, Ljung-Box, dan <i>P-Value</i> .....39
Tabel 4.4	Model Residual dengan ASTAR di Lima Kabupaten .....40
Tabel 4.5	Perbandingan Nilai RMSEP dan $R^2_{prediction}$ Model <i>Non Hybrid</i> dan <i>Hybrid</i> ASTAR.....42



# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Penyediaan data iklim merupakan hal yang penting sebagai upaya untuk mengetahui kondisi iklim di suatu wilayah, diantaranya dapat memberikan informasi untuk menyusun kalender tanam bagi para petani. Bagi daerah yang merupakan lumbung padi seperti Kabupaten Jember, Lamongan, Bojonegoro, Banyuwangi dan Ngawi adanya informasi iklim menjadi krusial karena berguna untuk meningkatkan produksi dan kualitas padi sertaantisipasi kegagalan panen akibat pergeseran awal musim dan periode masa tanam (Runtunuwu dan Syahbuddin, 2007).

Sampai saat ini, *general circulation model* (GCM) diakui banyak pihak sebagai alat untuk upaya memahami sistem iklim. GCM merupakan alat prediksi utama iklim dan cuaca secara numerik serta sebagai informasi primer untuk menilai pengaruh perubahan iklim (Wigena, 2006). Model proyeksi perubahan iklim di masa depan seringkali menggunakan data luaran GCM yang telah diproduksi beberapa negara maju yang terletak di lintang tinggi. Namun informasi GCM yang dihasilkan masih berskala global (ratusan km), sehingga sulit untuk mendapatkan informasi skala regional. Untuk daerah-daerah dengan topografi yang kompleks, di sepanjang garis pantai, dan daerah-daerah dengan tutupan lahan yang sangat heterogen, seperti halnya di Indonesia keluaran model GCM hasilnya kurang sensitif (Wilby, Charles, Zorita, Timbal, Whetton, dan Mearns. 2004).

Menurut Fernandez (2005), GCM masih bisa digunakan untuk memperoleh informasi skala lokal atau regional bila teknik *downscaling* digunakan. Teknik *downscaling* adalah suatu proses transformasi data dari suatu *grid* dengan skala besar menjadi data pada *grid-grid* dengan skala yang lebih kecil. Wilby dan Wigley (1997) mengatakan bahwa *downscaling* merupakan suatu cara menginterpolasi variabel-variabel prediktor atmosfer berskala regional

terhadap variabel-variabel berskala lebih kecil. Terdapat dua jenis *downscaling*, yaitu *dynamical downscaling* dan *statistical downscaling* (SD).

*Statistical downscaling* (SD) adalah proses *downscaling* yang berbasis regresi dimana data pada *grid-grid* berskala besar dalam periode dan jangka waktu tertentu digunakan sebagai dasar untuk menentukan data pada *grid* yang berskala kecil (Wigena, 2006). Data *grid* berskala besar atau GCM digunakan sebagai variabel prediktor dan observasi permukaan atau *grid* skala kecil sebagai variabel respon. Teknik SD digunakan untuk pemanfaatan kajian iklim, salah satu kegunaannya yaitu untuk memprediksi curah hujan pada skala lokal berdasarkan data GCM berskala global. Salah satu keuntungan utama teknik ini adalah komputasinya lebih murah dan dapat dengan mudah diaplikasikan pada luaran berbagai simulasi dan eksperimen berbasis GCM (Sutikno, 2008).

Data GCM merupakan data spasial dan temporal yang sangat memungkinkan terjadi korelasi spasial antar *grid* yang berbeda dalam satu domain dan autokorelasi dalam data deret waktu. Selain itu, permasalahan pada teknik SD yang muncul adalah bentuk fungsi yang tidak diketahui. Hal ini tentu saja melanggar asumsi dasar metode regresi baku. Selama ini sudah banyak metode yang digunakan untuk mengatasi masalah tersebut. Regresi komponen utama dapat mengatasi masalah korelasi spasial atau multikolinearitas antar *grid*, sementara kombinasi analisis komponen utama dan ARIMA dapat mengatasi masalah autokorelasi (Notodiputro, Wigena, dan Fitriadi, 2004). Solusi untuk menangani bentuk fungsi yang tidak diketahui digunakan regresi nonparametrik. Beberapa metode regresi nonparametrik SD yang telah digunakan antara lain, jaringan syaraf tiruan (Crane dan Hewitson, 1998; Cavaros dan Hewitson 2002; Sutikno dan Boer 2005) regresi bayes; regresi kontinum, *multivariate adaptive regression spline* (MARS) (Corte-Real, Zhang, & Wang, 1995; Fischer, Dewitte, & Maitrepirre, 2004), regresi *fourier*, dan regresi komponen utama. Haryoko (2004) menggunakan metode regresi *fourier* dan regresi komponen utama untuk memprediksi curah hujan di Indramayu. Disamping itu, Haryoko (2004) menyimpulkan bahwa model dengan reduksi dimensi transformasi *fourier* lebih baik daripada regresi komponen utama. Namun dalam penelitian tersebut belum mempertimbangkan kasus autokorelasi dan korelasi spasial. Selain itu beberapa

metode SD yang telah digunakan tersebut belum dapat menghasilkan akurasi nilai prediksi yang tinggi. Hal ini dapat disebabkan karena metode SD yang digunakan tidak sesuai dengan pola data GCM yang bersifat nonlinear. Metode nonlinear lain yang berkembang adalah *projection pursuit regression* (PPR) dan *adaptive spline threshold autoregression* (ASTAR).

Menurut Friedman dan Stuetzle (1989), metode PPR termasuk metode regresi nonlinier, nonparametrik, berbasis proyeksi, dan tidak mempunyai bentuk model baku serta tidak memerlukan asumsi yang ketat seperti model regresi parametrik. Penggunaan metode PPR dapat mengatasi masalah pada rata-rata lokal, fungsi polinomial, dan *recursive partitioning*, yaitu dengan cara menggunakan sejumlah fungsi pemulus (*smoothing*) dari hasil proyeksi atau reduksi dimensi. Wigena (2006) menggunakan metode PPR untuk peramalan curah hujan di Kabupaten Indramayu. Penelitian tersebut menyimpulkan bahwa pendugaan curah hujan bulanan dengan metode PPR lebih akurat dan pola nilai dugaan lebih mendekati pola data aktualnya daripada model PCR. Namun model PPR masih belum mengatasi adanya kasus autokorelasi. Oleh karena untuk mengatasi kasus tersebut dilakukan pemodelan residual dari model utama yang seringkali disebut model gabungan (*hybrid*). Beberapa penelitian yang menggabungkan dua metode antara lain: PCR dan ARIMA oleh Notodiputro *et al* (2004), Wigena dan Aunuddin (2004) menggunakan model *hybrid* PP dan ANN, sementara Sutikno (2008) melakukan penelitian dengan menggunakan model *hybrid* MARS dan ASTAR. Penelitian dengan model *hybrid* PPR dan ASTAR masih terbatas. Ketiga penelitian tersebut menghasilkan bahwa model *hybrid* memberikan hasil yang lebih baik dibandingkan dengan model *non hybrid*.

Sama halnya dengan PPR, metode ASTAR juga merupakan pendekatan metode nonparametrik. Metode ASTAR merupakan pengembangan dari MARS tetapi variabel prediktornya berupa nilai lag deret waktu (Lewis dan Steven, 1991). Metode tersebut merupakan pemodelan data dengan atau tidak melibatkan variabel prediktor. Istriana (2009) melakukan peramalan curah hujan dengan membandingkan metode ARIMA dan ASTAR, hasil yang diperoleh adalah metode ASTAR memberi hasil ramalan yang lebih baik.



Pada penelitian ini dilakukan pemodelan SD dengan menggunakan metode *hybrid* PPR dan ASTAR untuk meramalkan curah hujan pada lima Kabupaten sentra lumbung padi di Jawa Timur. Gabungan kedua model tersebut berpotensi mengatasi masalah autokorelasi dan korelasi spasial serta menghasilkan akurasi nilai prediksi yang tinggi.

## **1.2 Permasalahan Penelitian**

Berdasarkan latar belakang yang diurai di atas, maka permasalahan pada penelitian ini adalah sebagai berikut:

1. Bagaimana model terbaik SD antara data luaran GCM dengan curah hujan di Kabupaten Jember, Lamongan, Bojonegoro, Banyuwangi dan Ngawi menggunakan metode PPR dan *hybrid* PPR - ASTAR?
2. Bagaimana perbandingan kinerja antara model curah hujan dengan metode PPR dengan metode *hybrid* PPR - ASTAR?

## **1.3 Tujuan Penelitian**

Tujuan penelitian ini adalah sebagai berikut.

1. Mendapatkan model terbaik SD antara data luaran GCM dengan curah hujan di Kabupaten Jember, Lamongan, Bojonegoro, Banyuwangi dan Ngawi dengan menggunakan PPR dan ASTAR.
2. Membandingkan model *hybrid* PPR – ASTAR dan *non hybrid* dengan kriteria RMSEP

## **1.4 Manfaat Penelitian**

Manfaat yang diharapkan pada penelitian ini adalah sebagai berikut.

1. Membantu Badan Meteorologi Klimatologi (BMKG) dalam mendapatkan model curah hujan sehingga dapat digunakan untuk meramalkan curah hujan.
2. Adanya informasi peramalan curah hujan dapat digunakan untuk membuat kalender tanam, sehingga membantu petani untuk meminimalisir kerugian panen.

3. Mengembangkan khasanah ilmu Statistika khususnya bidang klimatologi.

### **1.5 Batasan Masalah**

Batasan masalah pada penelitian ini adalah data luaran yang digunakan dari aplikasi model CSIRO (*Commonwealth Scientific and Industrial Research Organization*) 20C3M yang diukur pada *grid* 3x3.

*(Halaman ini sengaja dikosongkan)*

## BAB II

### TINJAUAN PUSTAKA

#### 2.1 *Projection Pursuit Regression (PPR)*

*Projection Pursuit Regression (PPR)* adalah suatu metode regresi nonparametrik dan non linier. PPR digunakan untuk memproses data yang berdimensi besar yang dapat menggambarkan informasi dalam dimensi kecil melalui proses proyeksi. Oleh karena itu metode ini seringkali digunakan untuk mengatasi masalah rata-rata lokal, fungsi *polynomial*, dan *recursive partitioning* (Friedman & Tukey, 1974; Wigena, 2006).

Selama ini metode SD menggunakan *Principal Component Regression (PCR)* dan *Principal Component Analysis (PCA)* untuk *pre-processing* (Wigena, 2006). PCA adalah suatu prosedur untuk mereduksi dimensi data melalui transformasi variabel-variabel asal yang berkorelasi menjadi sekumpulan variabel baru yang tidak berkorelasi. Prosedur reduksi dimensi dengan PCA berdasarkan keragaman terbesar, hasil dari reduksi dimensi PCA kemudian digunakan untuk PCR (Safitri, 2012). Jika  $\mathbf{X}$  adalah matriks variabel prediktor berukuran  $t \times n$  dan  $\mathbf{Z}$  adalah matriks variabel baru hasil *pre-processing* berukuran  $t \times g$ , maka:

$$\mathbf{Z}^T = \mathbf{A} \mathbf{X} \quad (2.1)$$

dimana  $\mathbf{A}$  adalah matriks pemetaan atau proyeksi berukuran  $g \times n$ . Matriks  $\mathbf{A}$  bersifat ortogonal. Dalam metode PP, matriks  $\mathbf{X}$  ditransformasi dengan persamaan (2.1), namun prosedur untuk mendapatkan matriks  $\mathbf{A}$  berbeda dengan metode PCA, yaitu dengan cara memaksimalkan indeks proyeksi  $I(\boldsymbol{\alpha})$  seperti pada persamaan (2.6). Matriks  $\mathbf{A}$  disebut matriks koefisien proyeksi dan modelnya adalah:

$$y_i = \sum_{m=1}^M \mathbf{S}_{\boldsymbol{\alpha}_m}(\mathbf{z}) \quad (2.2)$$

dimana  $\mathbf{S}$  disebut fungsi pemulus dan  $\mathbf{z}$  yaitu *inner product* antara  $\boldsymbol{\alpha}_m$  dan  $\mathbf{x}$ , yaitu

$$\mathbf{z} = \boldsymbol{\alpha}_m \bullet \mathbf{x} \quad (2.3)$$

Besaran  $\boldsymbol{\alpha}_m$  disebut *loading* sedangkan  $\mathbf{z}$  disebut skor variabel prediktor.

Pemodelan PPR diawali dengan memaksimumkan indeks proyeksi, menentukan fungsi-fungsi variabel tunggal secara empirik berdasarkan proyeksi-proyeksi optimum, serta menjumlahkan fungsi-fungsi tersebut. (Jones & Sibson 1987). Berdasarkan Friedman & Stuetzle 1981, algoritma PPR adalah:

1. Menstandarisasi variabel respon. Proses standarisasi  $y_i$  adalah sebagai berikut:

$$w_i = \left( \frac{y_i - \mu}{\sigma} \right)$$

2. Penentuan nilai awal residual ( $e$ ) dan nilai  $M$  (banyaknya fungsi)

$$\begin{aligned} e_i &\leftarrow y_i, & i = 1, 2, \dots, t \\ M &\leftarrow 0 \end{aligned} \tag{2.4}$$

dimana  $\sum_{i=1}^t y_i = 0$  (variabel respon yang distandarkan). Setelah  $y_i$  distandarkan, langkah selanjutnya adalah meregresikan dengan metode *Ordinary Least Square* (OLS) antara  $w_i$  dengan  $\mathbf{X}$  (variabel prediktor).

$$w_i = \beta_0 + \sum_{k=1}^n \beta_k X_{ik} + \varepsilon_i \tag{2.5}$$

keterangan:

$\mathbf{y}$  = vektor variabel respon

$\mathbf{X}$  = matriks variabel prediktor

$\mu$  = rata-rata variabel respon

$\sigma$  = standar deviasi variabel respon

$i$  = pengamatan ke- $i$ ,  $i = 1, 2, \dots, t$

$k$  = variabel prediktor ke- $k$ ,  $k = 1, 2, \dots, n$

Residual ( $e_i$ ) yang hasil regresi antara  $w_i$  dan  $\mathbf{X}$  akan digunakan untuk menentukan indeks proyeksi minimum dalam model PPR.

Penentuan banyaknya fungsi didapatkan dari *scatterplot* antara variabel respon dan variabel prediktor. Penentuan banyaknya fungsi berdasarkan optimalisasi banyak fungsi  $m = 1, 2, 3, 4, \dots, 5$ . Safitri (2012) telah melakukan simulasi dalam menentukan optimalisasi banyak fungsi hingga  $m = 10$ , hasil yang diperoleh adalah nilai validasi RMSE semakin besar seiring

dengan bertambahnya fungsi. Oleh sebab itu dibatasi banyak fungsi  $m = 5$ . Banyak fungsi yang dipilih berdasarkan hasil validasi terbaik.

### 3. Penentuan $\alpha$ dan $S_\alpha$ dalam model

Untuk kombinasi linier persamaan (2.3), ditentukan fungsi pemulus  $S_\alpha(z)$  sesuai dengan nilai-nilai  $z$  dengan menggunakan indeks proyeksi  $I(\alpha)$  sebagai berikut:

$$I(\alpha) = 1 - \frac{\sum_{i=1}^t (e_i - S_\alpha(\alpha \cdot x_i))^2}{\sum_{i=1}^t e_i^2} \quad (2.6)$$

keterangan:

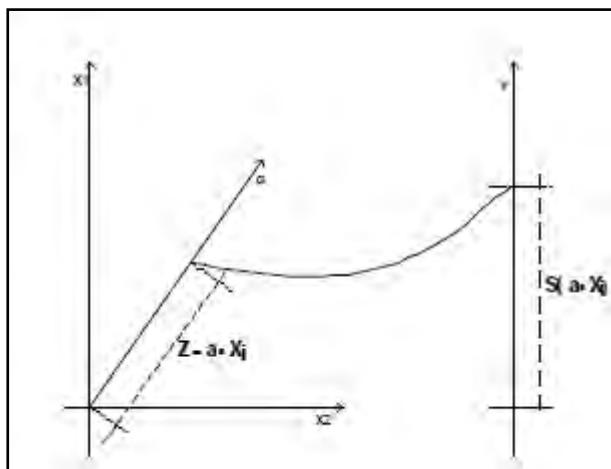
$e$  = residual

$S$  = fungsi pemulus

$x$  = variabel prediktor

$i$  = pengamatan ke- $i$ ,  $i = 1, 2, \dots, t$

Menentukan vektor koefisien  $\alpha_{M+1}$  yang memaksimalkan  $I(\alpha)$  atau  $\alpha_{M+1} = [\max_\alpha I(\alpha)]^{-1}$  dan fungsi pemulusnya adalah  $S_{\alpha_{M+1}}(z)$ . Ilustrasi geometrik proyeksi  $X$  terhadap  $Z$  dengan menggunakan nilai indeks proyeksi  $I(\alpha)$  ditunjukkan pada Gambar 2.1.



Sumber: Wigena, 2006

**Gambar 2.1** Ilustrasi Geometrik Proyeksi  $X$  terhadap  $Z$

4. Akhir algoritma

Jika  $I(\alpha)$  lebih kecil dari nilai *threshold*, maka stop: jika tidak, menentukan nilai residual dan nilai M sebagai berikut:

$$\begin{aligned} e_i &\leftarrow e_i - S_\alpha(\mathbf{z}), & i = 1, 2, \dots, t \\ M &\leftarrow M+1 \end{aligned} \quad (2.6)$$

dan kembali ke langkah 2. Nilai *threshold* didapatkan berdasarkan batas antara kombinasi linier dalam *scatterplot* antara variabel prediktor dan variabel respon.

Fungsi pemulus  $S_\alpha(\mathbf{z})$  ditentukan secara nonparametrik. Bentuk umum hubungan antara variabel respon dan  $\mathbf{z}$  dengan fungsi pemulusnya dapat dituliskan sebagai berikut:

$$y_i = S_\alpha(\mathbf{z}_i) + \varepsilon_i \quad (2.7)$$

Dalam regresi nonparametrik fungsi  $f(x)$  diduga dengan  $S_\alpha(\mathbf{z})$ , yang ditentukan berdasarkan rata-rata lokal, yaitu:

$$S(y_i) = AVE(y_j)_{i-a \leq j \leq i+a} \quad (2.8)$$

untuk lebar jendela (*bandwidth*)  $a$  tertentu dan dengan formulasi AVE seperti median atau rata-rata. Pemilihan nilai  $a$  sangat menentukan keragaman penduga dan besarnya bias. Nilai  $a$  yang terlalu kecil akan memperbesar keragaman ragam penduga, sedangkan nilai  $a$  yang terlalu besar akan memperbesar bias. Penentuan fungsi pemulus  $S_\alpha(\mathbf{z})$  menurut Friedman dan Stuetzle (1981):

- 1) Menentukan median untuk setiap tiga respon secara sekuensial untuk menghilangkan pengaruh data pencilan
- 2) Menentukan penduga ragam respon pada setiap titik dengan residual kuadrat rata-rata (*average squared residual*) dari penduga linier lokal ke- $a$  tertentu.
- 3) Pemulusan penduga ragam dengan rata-rata bergerak dan  $a$  tetap untuk menghindari perhitungan lebih dari satu kombinasi linier pada persamaan (2.3)
- 4) Pemulusan berurut dari tahap 1) dengan pengepasan (*fitting*) linier lokal dengan nilai  $k$  yang diperoleh pada tahap 3).

Menurut Hall (1989) menguraikan metode PPP secara matematik berdasarkan fungsi kernel (*kernel-based PPR*) dan sifat penduga PP. Pada dasarnya bahwa solusi PPR *invariant* terhadap setiap transformasi baik rotasi maupun penskalaan variabel prediktor. Berikut adalah penduga PP untuk mendapatkan proyeksi pertama. Berdasarkan persamaan (2.8),  $E(Y_i|X_i = x) = G(x)$ , dimana  $S(\bullet)$  disebut fungsi target (Hall, 1989). Jika  $G(\bullet)$  adalah fungsi pemetaan  $\mathfrak{R}^P \rightarrow \mathfrak{R}$ ,  $f(\bullet)$  adalah fungsi densitas yang memetakan variabel pada  $\mathfrak{R}^P$ , dan  $\mathbf{X}$  adalah variabel acak berdimensi  $p$  dengan densitas  $f(\bullet)$ , maka untuk skalar  $z$ ,

$$S_{\alpha}(z) = E\{G(x) | \alpha \bullet \mathbf{x} = z\} \quad (2.10)$$

Proyeksi pertama terhadap  $G(x)$  adalah fungsi  $G_1(x) = S_{\alpha_1}(z)$  di mana  $\alpha_1$  meminimumkan  $L(\alpha)$  berikut:

$$L(\alpha) = E\left[\{G(x) - S_{\alpha}(z)\}^2\right] \quad (2.11)$$

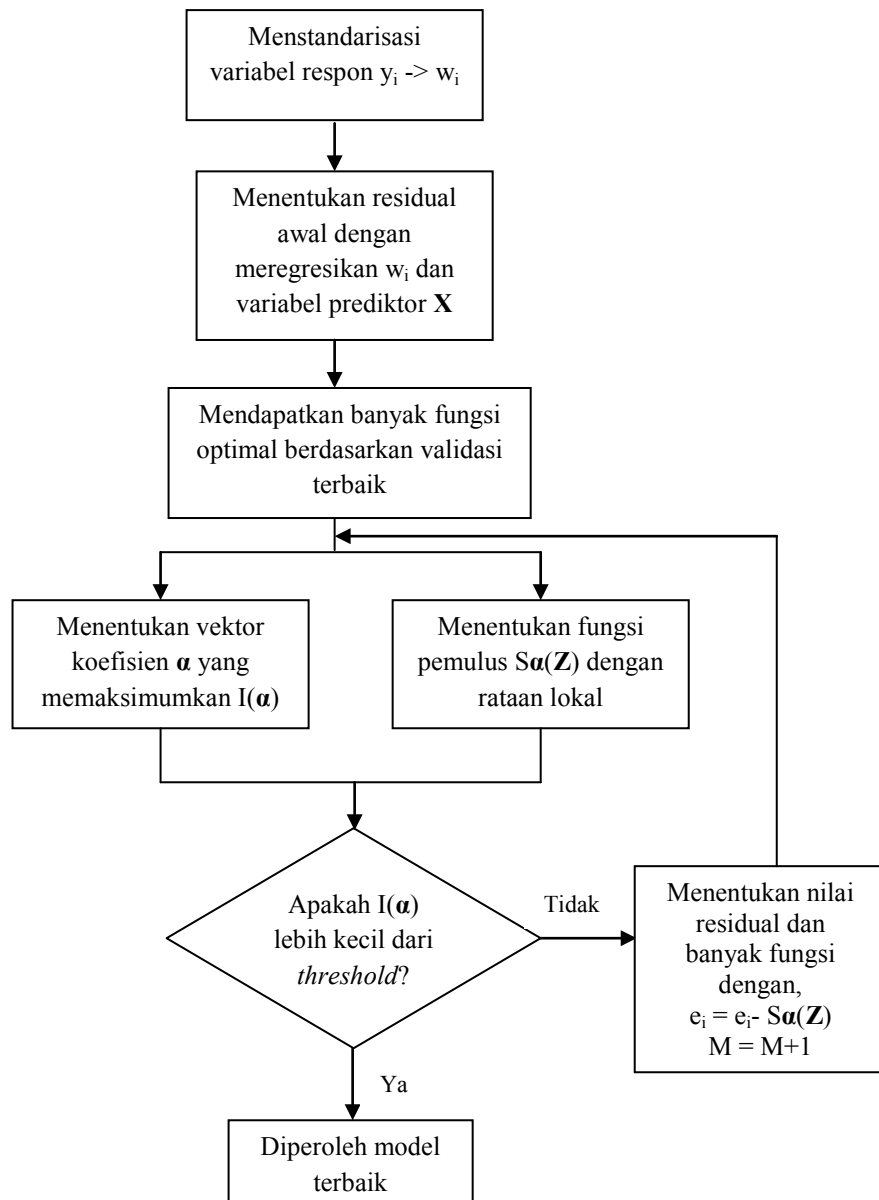
Sehingga penduga  $\alpha_1$  akan meminimumkan penduga  $L(\alpha)$ , yaitu:

$$\hat{L}(\alpha) = \frac{1}{t} \sum_{i=1}^t \{y_i - \hat{S}_{\alpha(i)}(\alpha \bullet \mathbf{x}_i)\}^2 \quad (2.12)$$

dan penduga proyeksi pertamanya adalah

$$\hat{G}_{1(i)}(x) = \hat{S}_{\alpha_1(i)}(\hat{\alpha}_1 \bullet \mathbf{x}) \quad (2.13)$$





**Gambar 2.2** Diagram Alir Algoritma PPR

Pada Gambar 2.2 menunjukkan tahapan atau algoritma pembentukan model pada PPR. Proses proyeksi ke- $m$  memiliki tahapan yang sama seperti langkah-langkah mendapatkan proyeksi kesatu. Penduga  $\hat{S}_{\alpha_1(i)}(\mathbf{z})$  akan konvergen terhadap  $S_{\alpha}(\mathbf{z})$  dan konsisten, di mana  $\hat{\alpha}_1$  juga konvergen terhadap  $\alpha$  sehingga didapatkan model PPR adalah sebagai berikut:

$$y_i = \sum_{m=1}^M S_{\alpha_m}(\alpha_m \bullet \mathbf{x}) \quad (2.14)$$

$$= \bar{y}_i + \sum_{m=1}^M \delta_m \left( \sum_{k=1}^n \alpha_{km} x_{ik} \right) + \varepsilon_i$$

dengan:

$S_{\alpha_m}(\alpha_m \bullet \mathbf{x})$  = suatu fungsi yang tidak diketahui

$\alpha_m$  =  $(\alpha_{1m}, \alpha_{2m}, \dots, \alpha_{km})$  = vektor satuan (arah *projection pursuit*)

$\mathbf{x}_{ik}$  =  $(x_{i1}, x_{i2}, \dots, x_{ik})$  = variabel prediktor

$y_i$  = variabel respon

$\bar{y}_i$  = rata-rata variabel respon

$m$  = banyak fungsi

$k$  = banyak variabel prediktor

$\varepsilon_i$  = faktor acak dengan asumsi adanya autokorelasi;  $E(\varepsilon_i) = 0$  dan

$$\text{var}(\varepsilon_i) = \sigma^2$$

Didefinisikan bahwa:

1.  $G(\mathbf{x}) = E(Y_i | X_i = \mathbf{x})$
2.  $\Omega$  = gugus vektor satuan berdimensi  $p$  ( $\alpha_m$ )
3.  $S_{\alpha}(\alpha \bullet \mathbf{x}) = E(Y_i | \alpha \bullet \mathbf{x}_i = \alpha \bullet \mathbf{x})$ , dimana  $\alpha \in \Omega$

Komponen proyeksi pertama dari  $G(\mathbf{x})$  adalah  $S_{\alpha_1}(\alpha_1 \bullet \mathbf{x})$  dimana  $\alpha_1 \in \Omega$

meminimumkan  $L(\alpha)$  (mengacu pada persamaan 2.11) berikut:

$$L(\alpha) = E \left[ \left\{ G(x_i) - S_{\alpha}(\alpha \bullet \mathbf{x}_i) \right\}^2 \right]$$

Diasumsikan bahwa nilai minimum ini bersifat unik. Misalkan bahwa:

$$S_{\alpha}^{(2)}(\alpha \bullet \mathbf{x}) = E(y_t - S_{\alpha_1}(\alpha_1 \bullet \mathbf{x}_i) | \alpha \bullet \mathbf{x}_i = \alpha \bullet \mathbf{x})$$

Maka komponen proyeksi keduanya adalah  $S_{\alpha_2}^{(2)}(\alpha_2 \bullet \mathbf{x})$  dimana  $\alpha_2 \in \Omega$

meminimumkan  $L^{(2)}(\alpha_1, \alpha)$  berikut:

$$L^{(2)}(\alpha_1, \alpha) = E \left[ G(x_i) - S_{\alpha_1}(\alpha_1 \bullet \mathbf{x}_i) - S_{\alpha}(\alpha \bullet \mathbf{x}_i) \right]^2$$

Diasumsikan juga bahwa nilai minimum ini bersifat unik. Dengan cara yang sama dapat diperoleh komponen-komponen proyeksi berikutnya dari  $G(\mathbf{x})$ .  $L(\alpha)$  dapat digunakan untuk mengukur pendekatan  $S_\alpha(\alpha \bullet \mathbf{x})$  terhadap  $G(\mathbf{x})$ ,  $L^{(2)}(\alpha_1, \alpha)$  untuk mengukur pendekatan  $S_{\alpha_1}(\alpha_1 \bullet \mathbf{x}) + S_\alpha^{(2)}(\alpha \bullet \mathbf{x})$  terhadap  $G(\mathbf{x})$ .

## 2.2 Adaptive Spline Threshold Autoregression (ASTAR)

Metode ASTAR adalah pemodelan deret waktu *nonlinear threshold* sebagai pengembangan metode *Multivariate Adaptive Regression Spline* (MARS) dengan  $\{V_i\}$  sebagai variabel respon dan  $\{V_{i-j}\}$  sebagai variabel prediktor (Lewis & Steven, 1991). Dalam penerapannya, metode ini lebih banyak digunakan untuk pemodelan data iklim atau curah hujan yang merupakan data deret waktu namun tidak linier.

Pada ASTAR, titik knot disebut *threshold* dimana pembentukannya sama seperti knot pada MARS. Misalkan  $V_i$  dengan  $i = 1, 2, \dots, t$  adalah variabel respon pada model ASTAR dan variabel prediktor lag  $p = 3$  yaitu  $V_{i-1}, V_{i-2}$ , dan  $V_{i-3}$ . Pada setiap langkah *forward* dalam algoritma ASTAR memilih satu dari satu himpunan bentuk baru untuk model ASTAR. Kandidat himpunan untuk inisial langkah *forward* pada algoritma ASTAR adalah sebagai berikut:

$$\begin{aligned} &(V_{i-1} - t^*)_+ \text{ dan } (t^* - V_{i-1})_+ \text{ atau} & (2.15) \\ &(V_{i-2} - t^*)_+ \text{ dan } (t^* - V_{i-2})_+ \text{ atau} \\ &(V_{i-3} - t^*)_+ \text{ dan } (t^* - V_{i-3})_+ \end{aligned}$$

Dengan beberapa *threshold*  $t^*$  pada masing-masing domain variabel prediktor. Sebagai contoh langkah pertama diasumsikan bahwa ASTAR memilih variabel rediktor lagnya adalah  $V_{i-2}$  dengan  $t^* = t_1$ , berarti bahwa  $(V_{i-2} - t_1)_+$  dan  $(t_1 - V_{i-2})_+$  adalah kurva awal dari model ASTAR selain konstanta. Himpunan kandidat pada langkah kedua dari algoritma *forward* termasuk seluruh himpunan kandidat pada (2.15) dan himpunan kandidat yang baru sebagai berikut.

$$\begin{aligned} &(V_{i-2} - t_1)_+(V_{i-1} - t^*)_+ \text{ dan } (V_{i-2} - t_1)_+(t^* - V_{i-1})_+ \text{ atau} \\ &(V_{i-2} - t_1)_+(V_{i-3} - t^*)_+ \text{ dan } (V_{i-2} - t_1)_+(t^* - V_{i-3})_+ \text{ atau} \end{aligned}$$

$$(t_1 - V_{i-2})_+(V_{i-1} - t^*)_+ \text{ dan } (t_1 - V_{i-2})_+(t^* - V_{i-1})_+ \text{ atau} \\ (t_1 - V_{i-2})_+(V_{i-3} - t^*)_+ \text{ dan } (t_1 - V_{i-2})_+(t^* - V_{i-3})_+ \quad (2.16)$$

Karena pada langkah pertama terpilih  $(V_{i-2} - t_1)_+$  dan  $(t_1 - V_{i-2})_+$ , maka pada langkah kedua, variabel tersebut menjadi kandidat dalam perkalian *threshold* satu variabel  $V_{i-j}$  dengan beberapa *threshold* yang lain ( $t^* \neq t_1$ ). Pemilihan jumlah dan lokasi knot secara otomatis berdasarkan nilai ASR (*Average Sum of Square Residual*) yang minimum. Jika algoritma *forward stepwise* sudah selesai, maka dilanjutkan dengan *backward stepwise*. Algoritma *backward* digunakan untuk mendapatkan jumlah basis fungsi berdasarkan nilai GCV (*General Cross Validation*) yang minimum (Friedman, 1991). Contoh model ASTAR yang sederhana adalah

$$V_i = c + \phi_1(V_{i-d_1} - t_1) + \phi_2(V_{i-d_2} - t_2) + \phi_3(V_{i-d_1} - t_1)(V_{i-d_2} - t_2) + \varepsilon_i \quad (2.17)$$

dengan:

$c$  : konstanta

$t_1, t_2$  : masing-masing nilai knot variabel  $V_{t-d_1}$  dan  $V_{t-d_2}$

$d_1, d_2$  : merupakan nilai lag 1 dan 2

Pemilihan model akhir menggunakan *backward stepwise* berdasarkan satu dari empat kriteria (Lewis dan Ray, 1997) yaitu GCV, AIC (*Akaike Information Criterion*), *Amemyia's prediction criterion*, dan SC (*Schwarz-Rissanen Criterion*).

Persamaan model ASTAR dalam bentuk matrik dapat ditulis menjadi:

$$\mathbf{V} = \mathbf{B}^* \boldsymbol{\gamma} + \boldsymbol{\varepsilon} \quad (2.18)$$

dengan,

$$\mathbf{V} = (v_1, \dots, v_n)^T, \quad \boldsymbol{\gamma} = (\gamma_0, \dots, \gamma_m)^T, \quad \boldsymbol{\varepsilon} = (\varepsilon_1, \dots, \varepsilon_n)^T \\ \mathbf{B}^* = \begin{bmatrix} 1 & s_{1m}(V_{t-1(1,m)} - t_{1m})_+ & \cdots & s_{1m}(V_{t-1(M,m)} - t_{Mm})_+ \\ 1 & s_{1m}(V_{t-1(1,m)} - t_{1m})_+ & \cdots & s_{1m}(V_{t-1(M,m)} - t_{Mm})_+ \\ \vdots & \vdots & \vdots & \vdots \\ 1 & s_{1m}(V_{t-1(1,m)} - t_{1m})_+ & \cdots & s_{1m}(V_{t-1(M,m)} - t_{Mm})_+ \end{bmatrix}$$

### 2.3 Model *Hybrid* PPR dan ASTAR

Model *hybrid* yang dimaksud adalah gabungan antara model utama dengan model pada residualnya. Modifikasi model ini diperoleh dengan cara memodelkan lagi residual dari model yang telah terbentuk (Sutikno, 2003).

Pada penelitian ini model yang digabungkan menggunakan metode PPR pada model awal dan ASTAR pada model residualnya. Secara umum model regresi adalah sebagai berikut:

$$y_i = f(X_i) + \varepsilon_i$$

dengan,

$$f(X) \text{ adalah model utama yaitu } y_i = \bar{y}_i + \sum_{m=1}^M \delta_m \left( \sum_{k=1}^n \alpha_{km} X_{ik} \right)$$

$\varepsilon_i$  adalah model residualnya yaitu:

$$\hat{e}_i = c + \phi_1(e_{i-d_1} - t_1) + \phi_2(e_{i-d_2} - t_2) + \phi_3(e_{i-d_1} - t_1)(e_{i-d_2} - t_2)$$

### 2.4 Validasi Model

Validasi adalah proses terakhir dalam pemodelan untuk menentukan apakah model yang dihasilkan merupakan representasi berarti dan akurat dari sistem nyata. Beberapa metode validasi diantaranya dengan cara membagi data menjadi dua bagian (*in sample* dan *out sample*). Bagian pertama adalah data yang dinamakan model *building himpunan* digunakan untuk membentuk model. Data yang kedua digunakan untuk menguji (validasi) model yang dinamakan *prediction himpunan*.

Kriteria kebaikan model untuk validasi dapat menggunakan nilai *Root Mean Square Error Prediction* (RMSEP) adalah nilai varians dari residual yang dapat menunjukkan keakuratan suatu model. RMSEP mempunyai nilai minimal 0, semakin kecil nilai RMSEP menunjukkan bahwa perbedaan antara nilai dugaan hasil pemodelan dengan data aktualnya semakin kecil pula. Sehingga model yang terbaik adalah model dengan nilai RMSEP terkecil. Nilai RMSEP diperoleh dengan menggunakan rumus:

$$RMSEP = \sqrt{MS_{Error}} = \sqrt{\frac{\sum_{i=1}^t (Y_i - \hat{Y}_i)^2}{t}} \quad (2.19)$$

Selain itu, kriteria kebaikan model lainnya yang juga dapat digunakan adalah nilai  $R^2$  atau disebut dengan koefisien determinasi. Koefisien determinasi menunjukkan proporsi keragaman total nilai-nilai variabel respon yang dapat diterangkan oleh variabel prediktor dalam model yang digunakan. Nilai  $R^2$  dan berkisar antara 0 sampai 1, atau 0% sampai 100%. Semakin besar nilai  $R^2$ , maka semakin baik model yang didapatkan. Nilai  $R^2$  diperoleh dengan menggunakan rumus:

$$R^2 = \frac{SS_{Regression}}{SS_{Total}} = \frac{\sum_{i=1}^t (\hat{Y}_i - \bar{Y})^2}{\sum_{i=1}^t (Y_i - \bar{Y})^2} \quad (2.20)$$

dengan :

$Y_i$  = nilai variabel respon pada data validasi ke-  $i$

$\hat{Y}_i$  = nilai dugaan pada data validasi ke-  $i$

$\bar{Y}$  = nilai rata-rata variabel respon

$t$  = banyaknya pengamatan *out sample*

## 2.5 Metode Deteksi Autokorelasi Residual

Pada model *time series* salah satu syarat yang harus dipenuhi adalah residual bersifat *white noise* yaitu tidak ada korelasi antar residual dengan  $\mu = 0$  dan  $\sigma =$  konstan (Wei, 2006). Oleh sebab itu perlu dilakukan pendeteksian terhadap residual. Metode yang digunakan ada dua, yaitu secara visual dan statistik. Secara visual, autokorelasi dapat dilihat pada *plot* ACF residual. Apabila terdapat lag yang keluar batas ACF, maka dapat dikatakan bahwa terdapat korelasi pada residualnya.

Selain secara visual, pendeteksian autokorelasi residual dapat dilakukan dengan menggunakan uji Ljung-Box. Uji Ljung-Box merupakan modifikasi dari uji statistik Q original yang ditemukan oleh Box dan Pierce pada tahun 1970.

Kemudian pada tahun 1978 Ljung dan Box memodifikasi karena menemukan hasil bahwa pada pendekatan sampel  $n = 100$  statistik uji Q tidak memberikan kinerja yang baik (Cryer dan Chan, 2008). Hipotesis untuk uji Ljung-Box adalah:

$H_0$  : Tidak terdapat autokorelasi pada residual

$H_1$  : Terdapat autokorelasi pada residual

Statistik uji Ljung-Box adalah sebagai berikut:

$$Q = t(t+2) \sum_{d=1}^D \frac{\rho_d^2}{(t-d)}$$

dimana  $\rho_d = \frac{\sum_{i=1}^{t-d} (e_i - \bar{e})(e_{i+d} - \bar{e})}{\sum_{i=1}^t (e_i - \bar{e})^2}$   $\chi^2 > \chi^2$

$$p\text{-value} = P(\bar{e} > Q | \mu = 0)$$

$$= P(Z > \frac{Q - \mu}{\sigma} | \mu = 0)$$

Tolak  $H_0$  jika  $p\text{-value} < \alpha (0,05)$  atau  $\chi_{hit}^2 > \chi_{D-m}^2$

Keterangan:

$t$  = banyaknya pengamatan

$d$  = banyaknya lag

$e$  = residual

## 2.6 General Circulation Model (GCM)

*General Circulation Model (GCM)* merupakan model numerik, deterministik, dan simulasi komputer yang kompleks tentang kondisi iklim dengan berbagai komponennya yang berubah sepanjang waktu. GCM menjadi alat yang penting dalam memahami sistem iklim karena dapat memberikan informasi mengenai pergeseran iklim di masa yang akan datang. Model ini menggambarkan proses fisika pada atmosfer, daratan, lautan, dan mampu mensimulasikan kondisi iklim dalam skala besar. GCM menghasilkan data dalam bentuk *grid* atau petak

wilayah dengan resolusi rendah yang merepresentasikan keadaan iklim global. Meskipun GCM menghasilkan informasi iklim yang baik dalam skala besar, namun GCM tidak dirancang untuk informasi dalam skala kecil (Huth dan Kysley, 2000). Fernandez (2005) mengatakan untuk mengatasi hal tersebut maka digunakan teknik *statistical downscaling*.

Menurut Wigena (2006), beberapa model GCM antara lain GISS (*Goddard Institute for Space Studies*), GFDL (*Geophysical Fluid Dynamic Laboratory*), UKMO (*United Kingdom Meteorological Office*), CSIRO (*Commonwealth Scientific and Industrial Research Organization*), dan NCEP (*National Centers for Environmental Prediction*). Pada setiap model GCM memiliki perbedaan resolusi spasial, jumlah lapisan, dan persamaan dalam membangkitkan parameter atmosfer. Selama ini GCM telah dikembangkan dan digunakan di Indonesia untuk simulasi, prediksi, dan pembuatan skenario iklim.

## **2.7 Teknik *Statistical Downscaling***

*Statistical downscaling* adalah proses *downscaling* yang disusun berdasarkan adanya hubungan *grid* skala besar atau GCM sebagai variabel prediktor dan *grid* skala lokal sebagai variabel respon kemudian dinyatakan dalam model statistik (Wigena, 2006). Pendekatan pada teknik ini menggunakan model regresi sebagai berikut:

$$y = f(x) + \varepsilon$$

dimana:

y : variabel respon (curah hujan observasi, beresolusi tinggi: titik/wilayah)

x : variabel prediktor (GCM beresolusi rendah)

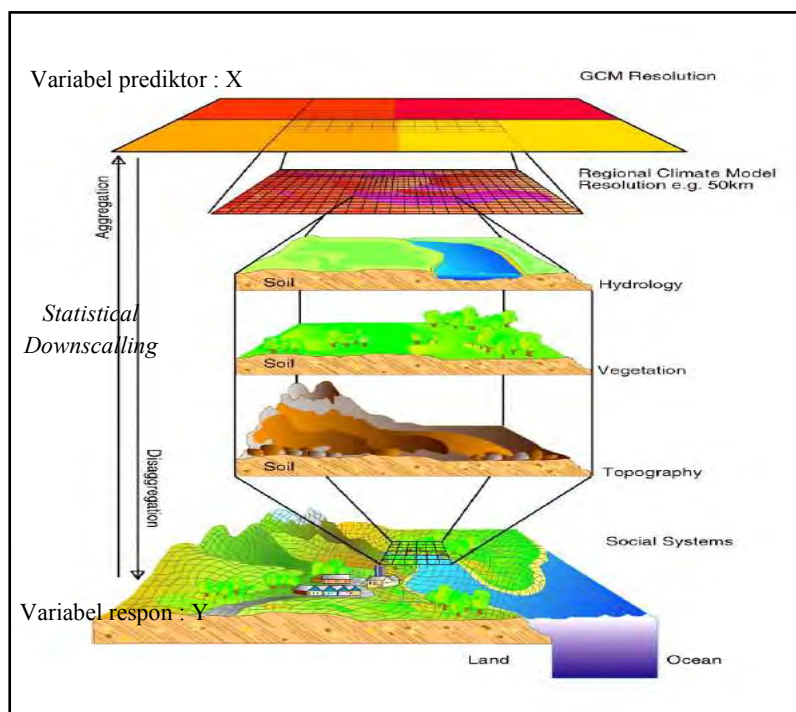
$\varepsilon$  : residual

Pada pendekatan SD data GCM digunakan untuk meramal iklim skala lokal. Dua hal yang sangat penting dalam pendekatan ini adalah penentuan variabel prediktor dan penentuan jumlah *grid* (Wilby & Wigley, 2000). Menurut Busuioc, Chen, dan Hellstrom (2001) model SD akan memberikan hasil yang baik, apabila beberapa syarat berikut dipenuhi, yaitu (1) Antara variabel respon dan prediktor memiliki korelasi yang kuat sehingga menjelaskan keragaman iklim



yang baik, (2) Variabel prediktor disimulasi baik oleh GCM, dan (3) Hubungan variabel respon dan prediktor tidak dipengaruhi waktu dan perubahan iklim.

Pada umumnya model SD menggunakan data deret waktu (t) dan data spasial GCM (s). Banyaknya variabel respon, prediktor, adanya autokorelasi dan kolinieritas pada variabel tersebut menjadikan model SD sangat kompleks. Oleh sebab itu dilakukan pengembangan dan modifikasi pada penerapannya sebagai solusi kompleksitas model SD tersebut. Ilustrasi teknik *statistical downscaling* disajikan pada Gambar 2.3 berikut.



Sumber: [www.consulclima.co.uk](http://www.consulclima.co.uk)

**Gambar 2.3** Ilustrasi Teknik *Statistical Downscaling*

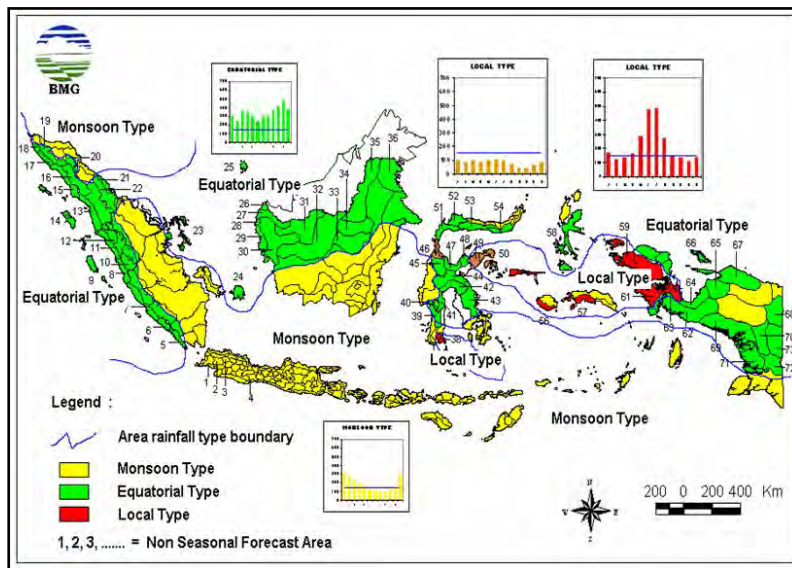
## 2.8 Curah Hujan

Hujan adalah sebuah presipitasi berwujud cairan yang terjadi karena proses kondensasi uap air di atmosfer. Jumlah curah hujan secara aktif diukur oleh radar, di Indonesia rata-rata curah hujan berubah-ubah setiap tahunnya, namun tergolong banyak yaitu 2000-3000mm per tahun.

Menurut BMKG, berdasarkan distribusi data rata-rata curah hujan bulanan, umumnya wilayah Indonesia dibagi menjadi 3 (tiga) pola hujan, yaitu :

1. Pola hujan *moonson*, yang wilayahnya memiliki perbedaan yang jelas antara periode musim hujan dan periode musim kemarau kemudian dikelompokkan dalam Zona Musim (ZOM), tipe curah hujan yang bersifat unimodial (satu puncak musim hujan, DJF musim hujan, JJA musim kemarau).
2. Pola hujan ekuatorial, yang wilayahnya memiliki distribusi hujan bulanan bimodial dengan dua puncak musim hujan maksimum dan hampir sepanjang tahun masuk dalam kriteria musim hujan. Pola ekuatorial dicirikan oleh tipe curah hujan dengan bentuk bimodial (dua puncak hujan) yang biasanya terjadi sekitar bulan Maret dan Oktober atau pada saat terjadi ekinoks (matahari berada dekat ekuator).
3. Pola hujan lokal, yang wilayahnya memiliki distribusi hujan bulanan kebalikan dengan pola *moonson*. Pola lokal dicirikan oleh bentuk pola hujan unimodial (satu puncak hujan), tetapi bentuknya berlawanan dengan tipe hujan *moonson*.

Wilayah sebaran pola hujan di Indonesia dapat dilihat pada Gambar 2.4. Tipe ekuatorial proses terjadinya berhubungan dengan pergerakan zona konvergensi ke utara dan selatan, wilayah sebarannya adalah Sumatera dan Kalimantan. Tipe *monsoon* dipengaruhi oleh angin laut dalam skala yang sangat luas, wilayah sebarannya adalah di pulau Jawa, Bali dan Nusa Tenggara dan sebagian Sumatera. Tipe lokal dicirikan dengan besarnya pengaruh kondisi lingkungan fisis setempat, seperti bentang perairan atau lautan, pegunungan yang tinggi, serta pemanasan lokal yang intensif, wilayah sebarannya meliputi Papua, Maluku dan sebagian Sulawesi. Jumlah curah hujan juga dipengaruhi oleh arah datang angin, pada sisi pegunungan atau gunung yang menghadap arah datang angin lembab (*windward side*) curah hujannya tinggi dan pada sisi sebaliknya (*leeward side*) curah hujannya sangat rendah atau rendah.



Sumber: BMG, 2008

**Gambar 2.4** Pola Curah Hujan di Indonesia

## BAB III

### METODOLOGI PENELITIAN

#### 1.1 Sumber Data

Data yang digunakan pada penelitian ini adalah data sekunder curah hujan yang diperoleh dari Badan Meteorologi, Klimatologi, dan Geofisika (BMKG) periode Januari 1984 sampai dengan Desember 2000 dan data GCM diperoleh dari website: <http://climexp.knmi.nl>. Jumlah data pada setiap lokasi penelitian yaitu 204. Wilayah yang menjadi lokasi penelitian adalah lima kabupaten dengan masing-masing dua stasiun pengamatan seperti yang disajikan pada Tabel 3.1 berikut.

**Tabel 3.1.** Lokasi Stasiun Pengamatan

No.	Kabupaten	Nama Stasiun Pengamatan	Lintang Selatan	Bujur Timur
1.	Ngawi	Karangjati dan Walikukun	7.35 - 7.52	110.167 - 111.67
2.	Lamongan	Lamongan dan Babat	6.86 - 7.38	112.56 - 112.56
3.	Bojonegoro	Jatiblimbing dan Kedungadem	6.98 - 7.62	112.42 - 112.15
4.	Jember	Ledokombo dan Pakusari	8 - 8.5	113.5 - 113.75
5.	Banyuwangi	Kawah ijen dan Kabat	7.43 - 8.46	113.53 - 114.38

#### 1.2 Variabel Penelitian

Variabel respon yang digunakan pada penelitian ini adalah curah hujan. Curah hujan tersebut diukur secara langsung di dua stasiun pada lima kabupaten. Kemudian data hasil pengukuran curah hujan di dua stasiun pengamatan tersebut, dihitung nilai rata-ratanya. Rata-rata curah hujan tersebut yang menjadi variabel respon.

Variabel prediktor adalah presipitasi merupakan luaran dari aplikasi model *Commonwealth Scientific and Industrial Research Organization (CSIRO) 20C3M* yang diambil pada *grid* 3x3. Penentuan titik *grid* ditunjukkan pada Gambar 3.1 berikut ini.

1	2	3
4	5	6
7	8	9

Sumber: BMKG, 2011

**Gambar 3.1** Penentuan Pemilihan Grid di Setiap Lokasi Stasiun Pengamatan

Bagian yang berwarna merah merupakan *grid* GCM CSIRO yang terdekat dengan pengamatan BMKG. Karena curah hujan diukur pada *grid* 3x3 maka terdapat 9 variabel prediktor pada pemodelan menggunakan PPR.

### 1.3 Metode Analisis Data

Langkah-langkah analisis pada penelitian ini adalah sebagai berikut:

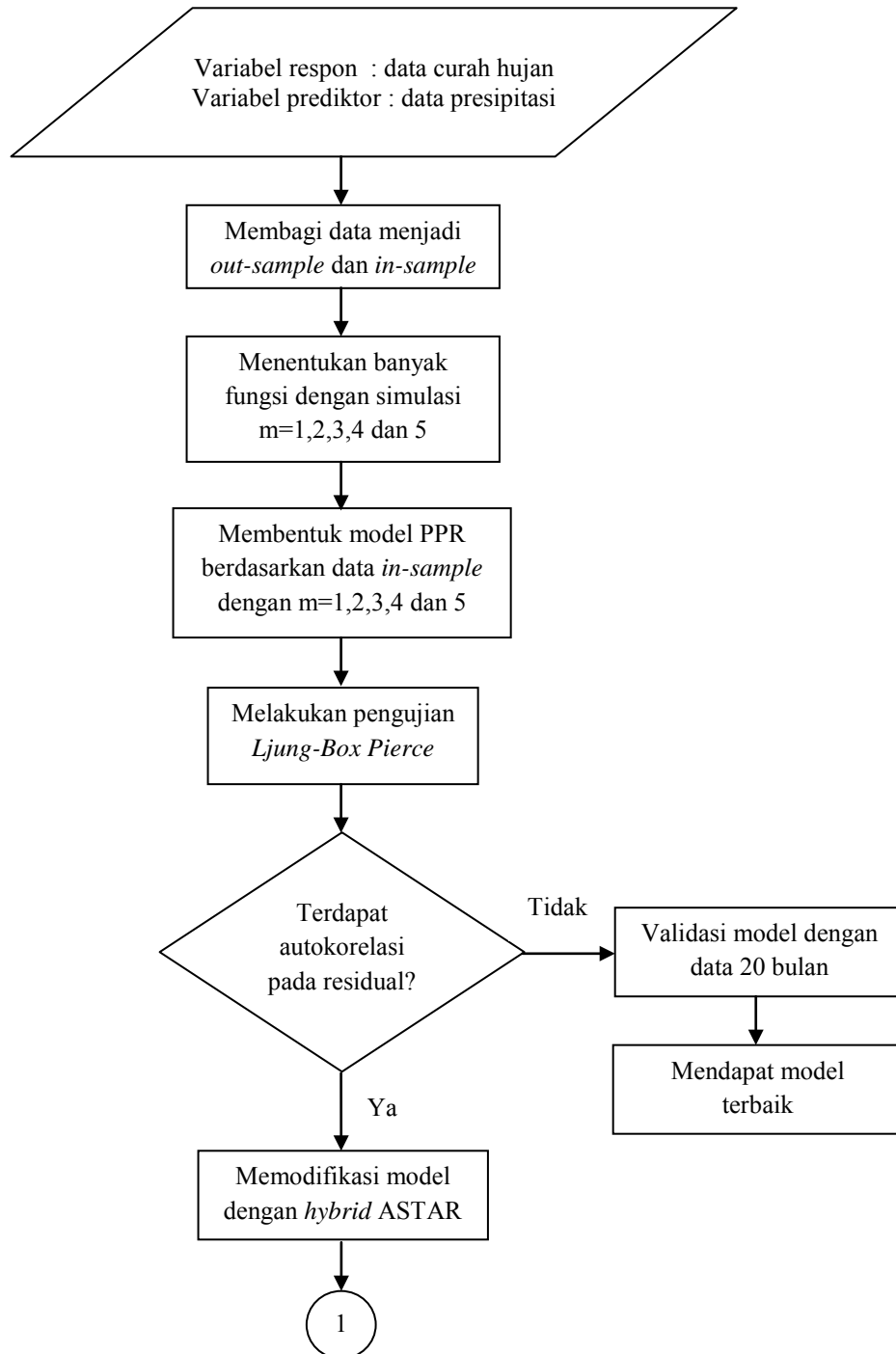
1. Pemodelan dengan menggunakan pendekatan PPR melalui tahapan sebagai berikut:
  - a. Membagi data curah hujan dan *output* CSIRO 20C3M menjadi *in-sample* dan *out-sample*. Banyaknya *in-sample* yaitu 90% dari data yaitu tahun 1984-1998, sedangkan *out-sample* sebanyak 10% dari data yaitu data tahun 1999-2000.
  - b. Menentukan banyaknya fungsi dalam model PPR. Banyaknya fungsi ditentukan dari simulasi  $m = 1, 2, 3, 4,$  dan  $5$ .
  - c. Membangun model PPR berdasarkan data *in-sample* dengan simulasi banyak fungsi  $m = 1, 2, 3, 4,$  dan  $n = 5$  dengan bentuk fungsi sebagai berikut:

$$y_i = \sum_{m=1}^M S_{am} (\alpha_m \bullet x)$$

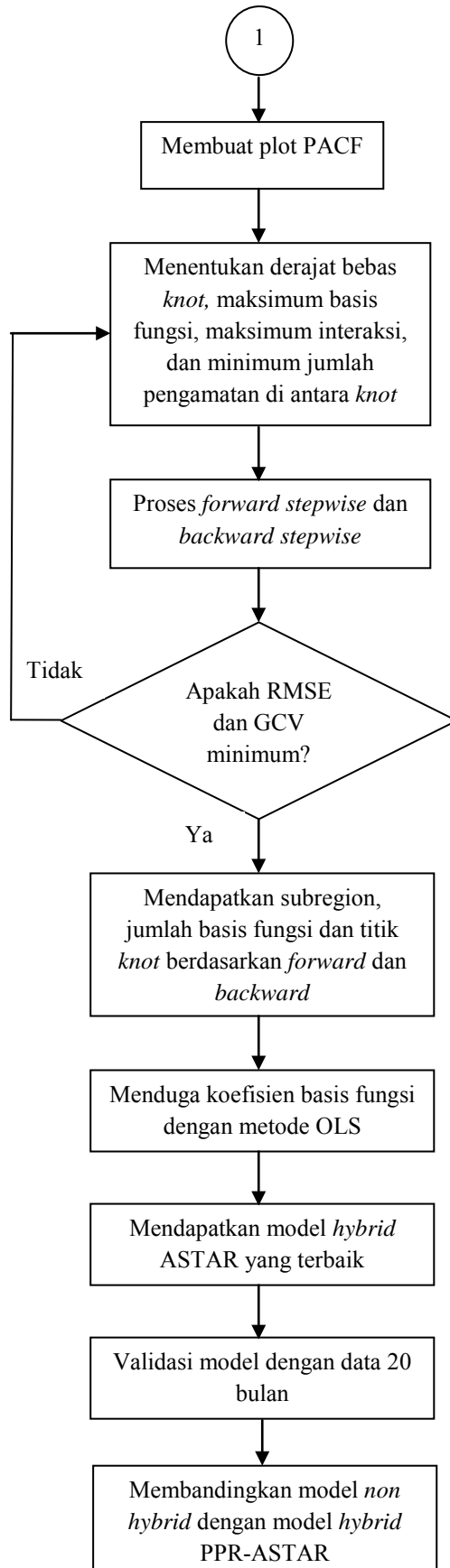
Sebagai variabel respon yaitu curah hujan. Variabel prediktor yaitu presipitasi luaran CSIRO 20C3M dengan ukuran *grid* 3x3.

- d. Melakukan validasi model berdasarkan data *out-sample* dengan menggunakan RMSE dan  $R^2$  sebagai kriteria kebaikan model.
2. Melakukan pendeteksian autokorelasi dengan membuat plot antara nilai residual dan observasi serta pengujian statistik *Ljung-Box Pierce*.
3. Pembentukan model ASTAR dengan tahapan sebagai berikut:
  - a. Menentukan variabel prediktor berdasarkan nilai  $PACF = \hat{\phi}_k$  yang berada di luar batas selangnya.
  - b. Menentukan maksimum derajat bebas *knot*, maksimum basis fungsi, maksimum jumlah interaksi, jumlah pengamatan di antara *knot*.
  - c. Proses *forward stepwise* untuk mendapatkan subregion melalui kriteria ASR terkecil.
  - d. Proses *backward stepwise* untuk mendapatkan jumlah basis fungsi melalui metode GCV minimum.
  - e. Melakukan pemilihan titik-titik *knot* dengan menggunakan algoritma *forward* dan *backward*.
  - f. Menduga koefisien dari basis fungsi yang telah terpilih dengan menggunakan *Ordinary Least Square* (OLS).
4. Melakukan validasi dengan menggunakan data 20 bulan yaitu tahun 1999-2000 dengan model terbaik yang telah terpilih.

Diagram alir tahapan penyusunan model peramalan curah hujan dengan metode PPR dan ASTAR adalah sebagai berikut:



**Gambar 3.2** Diagram Alir Penelitian



**Gambar 3.2** (Lanjutan)



*(Halaman ini sengaja dikosongkan)*

## BAB IV

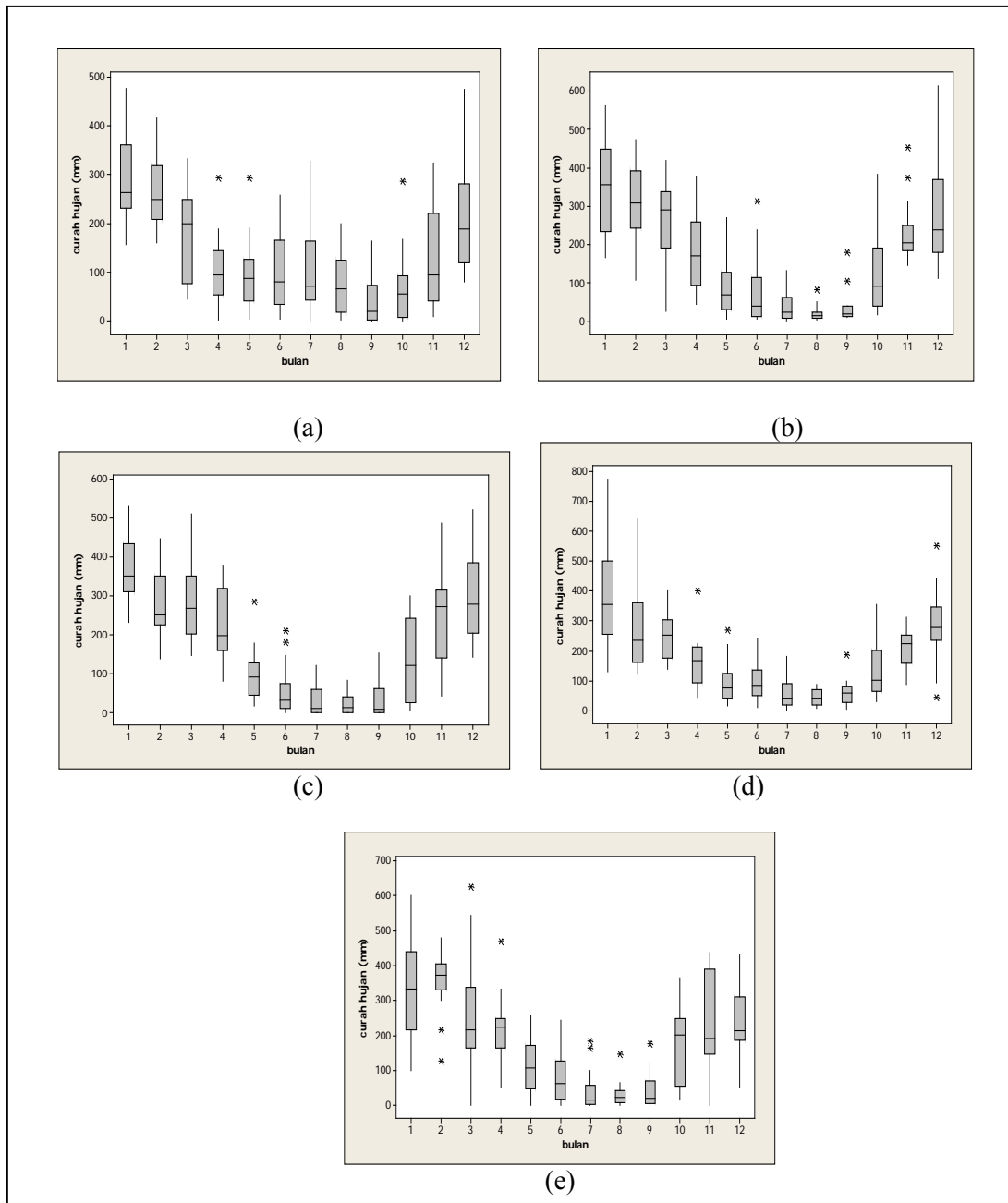
### ANALISIS DATA DAN PEMBAHASAN

Pada bagian ini membahas pemodelan SD dengan PPR dan *hybrid* PPR dan ASTAR . Bagian awal mendeskripsikan curah hujan di wilayah penelitian dan bagian akhir dilakukan validasi model baik *hybrid* maupun *non hybrid* di masing-masing Kabupaten.

#### 4.1 Deskripsi Curah Hujan di Lima Kabupaten

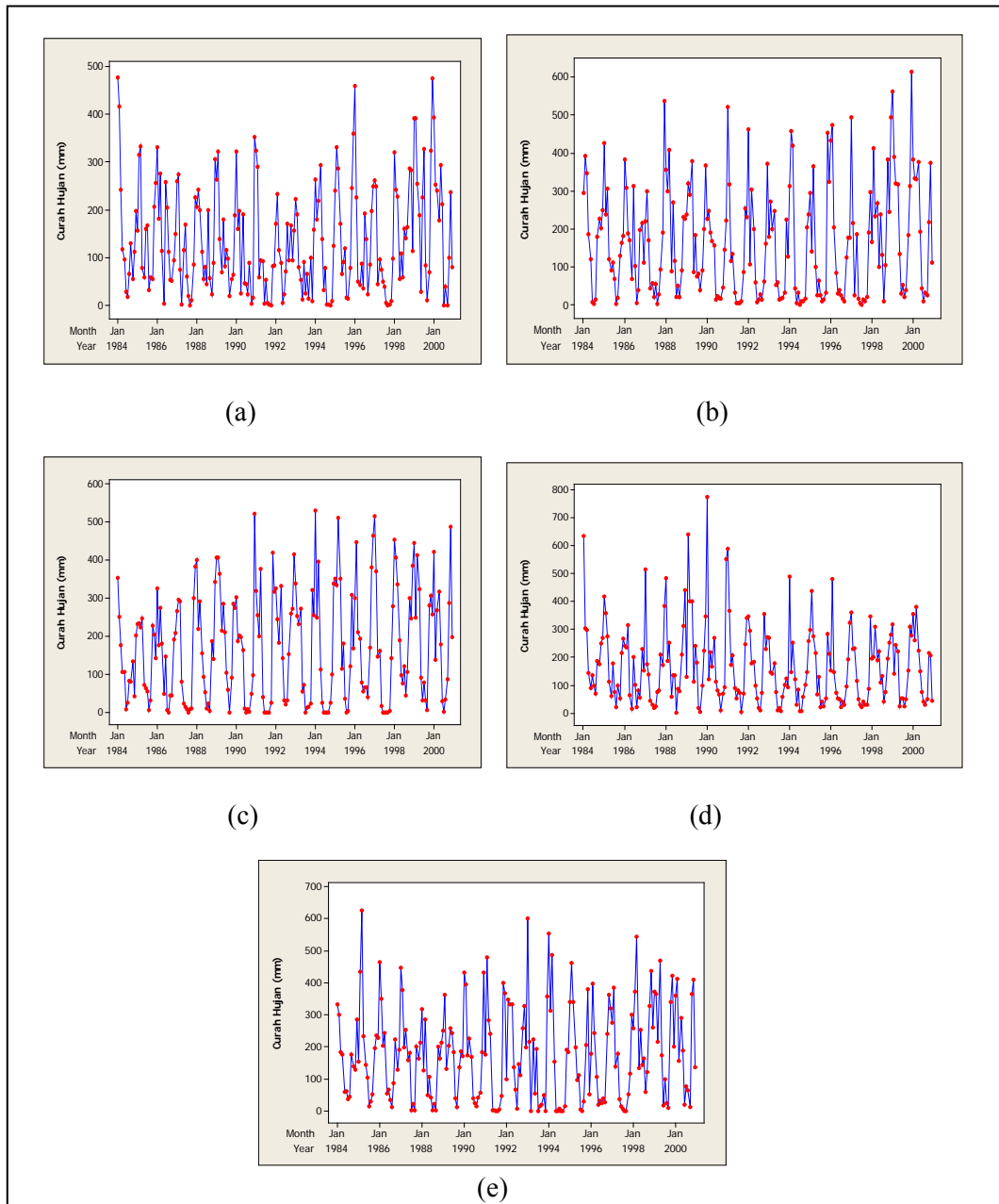
Curah hujan di lima kabupaten wilayah penelitian memiliki karakteristik yang hampir sama. Berdasarkan Gambar 4.1 *boxplot* curah hujan pada tiap kabupaten menunjukkan bahwa terdapat data *outlier* pada seluruh wilayah pengamatan. Rataan curah hujan tertinggi pada wilayah penelitian terjadi pada bulan Januari, kecuali Kabupaten Ngawi pada bulan Februari. Hal ini sesuai dengan tipe hujan di wilayah Jawa yang memiliki tipe hujan *moonson* (satu puncak musim hujan biasanya terjadi pada bulan Desember, Januari, atau Februari). Rataan curah hujan tertinggi di Kabupaten Banyuwangi, Jember, Bojonegoro, Lamongan dan Ngawi masing-masing adalah 298 mm; 350,7 mm; 369,6 mm; 385,9 mm dan 358,2 mm.

Bulan Agustus merupakan bulan yang memiliki rataaan curah hujan terendah untuk Kabupaten Bojonegoro, Jember, Lamongan, dan Ngawi, sedangkan untuk Banyuwangi rataaan curah hujan terendah terjadi pada bulan September. Rataan curah hujan terendah untuk masing-masing kabupaten adalah 43mm; 19,7mm; 23,9mm; 39,5mm dan 31,4mm.



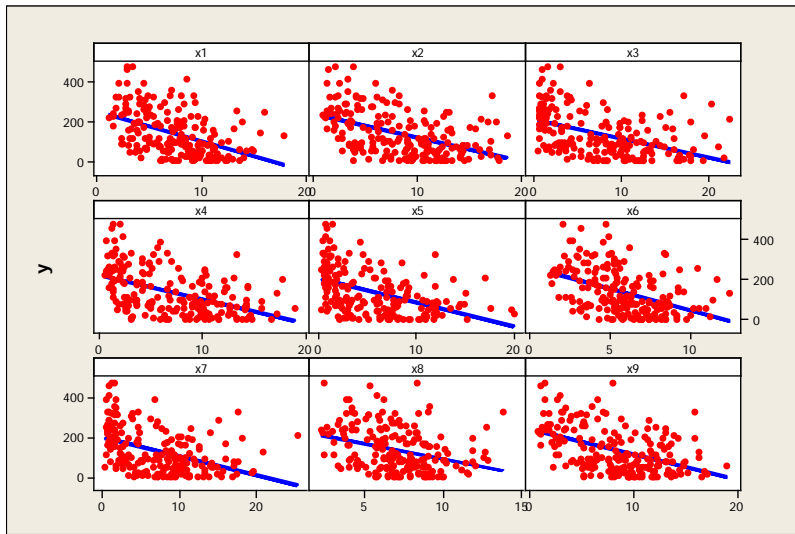
**Gambar 4.1** *Boxplot* Curah Hujan di Kabupaten (a) Banyuwangi, (b) Jember, (c) Bojonegoro, (d) Lamongan, dan (e) Ngawi

*Time series plot* curah hujan di lima Kabupaten disajikan pada Gambar 4.2. Curah hujan di lima kabupaten tampak fluktuatif. Dengan standar deviasi yang besar, nampak bahwa curah hujan di lima Kabupaten mengalami kenaikan dan penurunan yang signifikan selama periode 17 tahun dari Januari 1984 sampai Desember 2000.

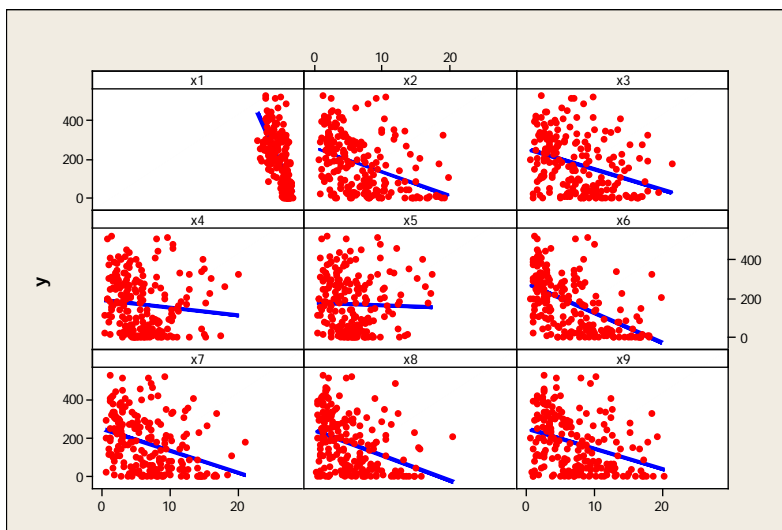


**Gambar 4.2** Time Series Plot Curah Hujan di Kabupaten (a) Banyuwangi, (b) Jember, (c) Bojonegoro, (d) Lamongan, dan (e) Ngawi

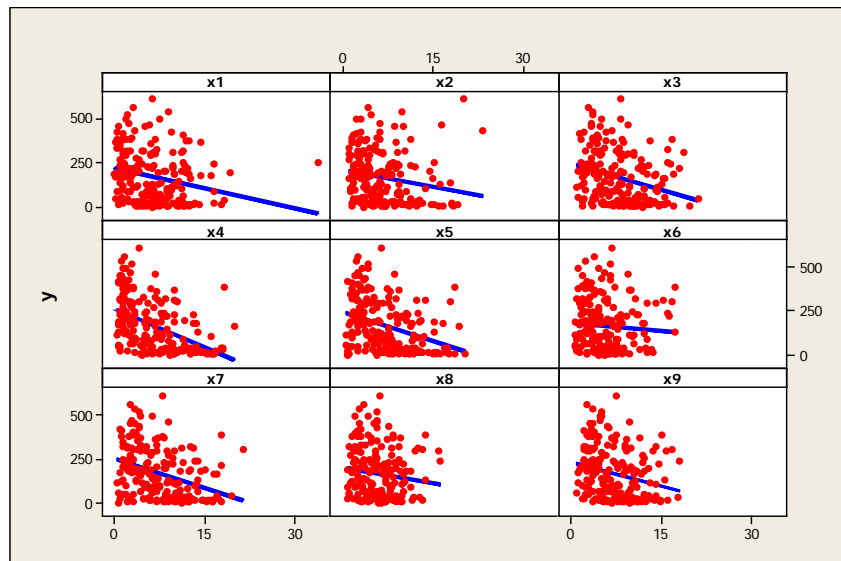
Setelah dilakukan identifikasi data curah hujan (variabel respon) melalui *time series plot*, kemudian dilakukan identifikasi pola hubungan antara variabel respon dan variabel prediktor. Secara visual, pola hubungan tersebut dapat ditunjukkan sebagai berikut.



**Gambar 4.3** *Plot* Antara Variabel Respon dan Variabel Prediktor di Kabupaten Banyuwangi



**Gambar 4.4** *Plot* Antara Variabel Respon dan Variabel Prediktor di Kabupaten Bojonegoro



**Gambar 4.5** *Plot* Antara Variabel Respon dan Variabel Prediktor di Kabupaten Jember

Pada Gambar 4.3 sampai Gambar 4.5 secara jelas nampak bahwa tidak terdapat pola hubungan linier antar variabel respon dan variabel prediktor di Kabupaten Banyuwangi, Bojonegoro, dan Jember. Hal ini berarti sesuai dengan karakteristik data curah hujan yang tidak linier. Oleh sebab itu dalam pembentukan model sebaiknya menggunakan metode non linier agar diperoleh hasil yang sesuai. Secara lengkap untuk *plot* Kabupaten Lamongan dan Ngawi disajikan pada Lampiran 2.

#### **4.2** Pemodelan SD dengan *Projection Pursuit Regression (PPR)*

Pemodelan SD dengan PPR dilakukan dengan cara meregresikan data luaran GCM sebagai variabel prediktor dan data rata-rata curah hujan sebagai variabel respon. Data curah hujan yang digunakan adalah rata-rata dari dua stasiun di Kabupaten Banyuwangi, Jember, Bojonegoro, Lamongan, dan Ngawi.

Tahap awal dalam proses pembentukan model PPR adalah menentukan proyeksi atau banyak fungsi yang akan disertakan dalam model. Proses optimalisasi proyeksi dilakukan dengan cara simulasi banyak fungsi yang akan digunakan dalam model  $m = 1, 2, 3, 4, \text{ dan } n = 5$ . Kemudian membangun model antara observasi curah hujan dan luaran GCM, dan terakhir adalah melakukan validasi model dengan data *out-sample*. Hasil validasi model kemudian

membandingkan nilai RMSEP dan  $R^2_{pred}$  pada masing-masing fungsi, nilai RMSEP yang terkecil dan atau  $R^2_{pred}$  yang terbesar adalah yang dipilih sebagai model terbaik. Berikut adalah hasil perbandingan nilai RMSEP dan  $R^2_{pred}$  pada masing-masing banyaknya fungsi.

**Tabel 4.1** Nilai RMSEP dan  $R^2_{pred}$  Menurut Banyaknya Fungsi di Lima Kabupaten

Kabupaten	m = 1		m = 2		m = 3		m = 4	
	RMSEP	$R^2_{pred}$	RMSEP	$R^2_{pred}$	RMSEP	$R^2_{pred}$	RMSEP	$R^2_{pred}$
Banyuwangi	142,79	34,0%	133,32	21,0%	144,88	22,7%	159,76	39,0%
Jember	139,23	42,0%	156,53	65,0%	150,47	49,0%	149,02	51,0%
Bojonegoro	121,31	70,0%	130,01	75,5%	142,33	59,3%	130,24	66,9%
Lamongan	120,79	70,3%	122,41	64,8%	122,20	56,0%	125,07	72,6%
Ngawi	122,04	72,2%	123,97	65,9%	150,13	65,8%	160,24	77,3%

Berdasarkan Tabel 4.1 di atas dapat dilihat bahwa nilai RMSEP dan  $R^2_{pred}$  tidak berpengaruh pada banyaknya fungsi (m) artinya semakin banyak fungsi tidak selalu semakin kecil RMSEP dan juga tidak selalu semakin besar nilai  $R^2_{pred}$ . Pada stasiun di Kabupaten Banyuwangi model optimal pada  $m = 2$  dengan nilai RMSEP = 133,32 dan  $R^2_{pred} = 21,0\%$ . Pada stasiun di Kabupaten Jember, Bojonegoro, Lamongan dan Ngawi model optimal pada  $m = 1$  dengan RMSEP = 139,23 dan  $R^2_{pred} = 42,0\%$  untuk Jember, RMSEP = 121,31 dan  $R^2_{pred} = 70,0\%$  untuk Bojonegoro. Sedangkan untuk Kabupaten Lamongan RMSEP = 120,79 dan  $R^2_{pred} = 70,3\%$  dan pada Kabupaten Ngawi yaitu RMSEP = 122,04 dan  $R^2_{pred} = 72,2\%$ . Dari uraian tersebut secara umum nilai RMSEP pada lima Kabupaten yaitu antara 120 – 143 sedangkan  $R^2_{pred}$  antara 21,0% - 73%. Model PPR data curah hujan di kelima Kabupaten dari nilai RMSEP dan  $R^2_{pred}$  optimal disajikan pada Tabel 4.2 berikut.

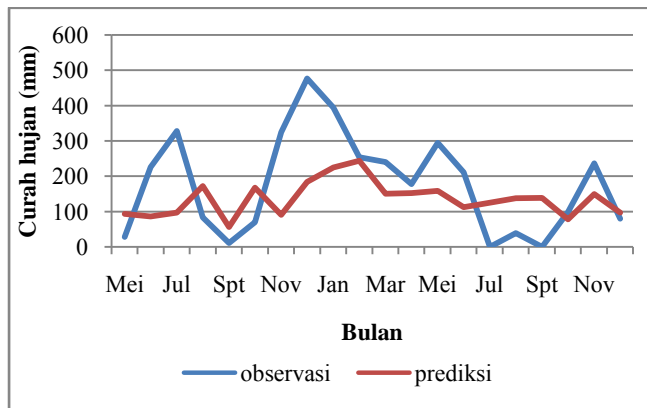
**Tabel 4.2** Model PPR Data Curah Hujan di Lima Kabupaten

Kabupaten	Model
Banyuwangi	$\hat{y} = 138,17 + 78,28(0,144X_1 - 0,109X_2 - 0,144X_3 - 0,119X_4 - 0,594X_5 - 0,425X_6 + 0,432X_7 + 0,438X_8 - 0,138X_9) + 43,66(-0,241X_1 - 0,353X_2 - 0,413X_3 + 0,198X_4 - 0,108X_5 + 0,410X_6 - 0,202X_8 + 0,488X_9)$
Jember	$\hat{y} = 166,40 + 94,92(0,08X_1 - 0,007X_2 + 0,028X_3 - 0,571X_4 + 0,267X_5 + 0,472X_6 - 0,317X_7 - 0,438X_8 + 0,288X_9)$
Bojonegoro	$\hat{y} = 173,20 + 112,63(-0,921X_1 + 0,181X_2 + 0,003X_3 + 0,128X_4 + 0,077X_5 - 0,245X_6 - 0,187X_7 + 0,037X_8 + 0,02X_9)$
Lamongan	$\hat{y} = 170,44 + 93,52(0,0002X_1 - 0,4X_2 + +0,004X_3 + 0,861X_4 - 0,031X_5 - 0,13X_6 - 0,093X_7 - 0,0377X_8 - 0,262X_9)$
Ngawi	$\hat{y} = 179,00 + 112,11(-0,877X_1 + 0,228X_2 + 0,238X_3 + 0,143X_4 + 0,083X_5 - 0,207X_6 - 0,323X_7 + 0,051X_8 + 0,035X_9)$

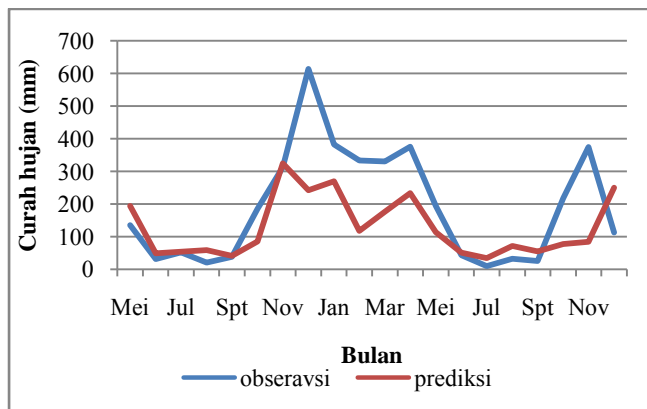
Koefisien proyeksi  $\alpha_{ik}$  dengan  $i = 1$  dan  $2$  yang memproyeksikan  $X_k$  pada persamaan PPR optimal untuk model curah hujan Kabupaten Banyuwangi karena banyak fungsi yang optimal adalah  $m=2$ , sedangkan untuk kabupaten lainnya koefisien proyeksi hanya  $\alpha_{1k}$ .

Secara visual berikut disajikan plot hasil ramalan dengan kriteria RMSEP terkecil menggunakan data *out-sample*. Pada Gambar 4.6 sampai Gambar 4.10 menunjukkan bahwa hasil prediksi curah hujan di Kabupaten Banyuwangi, Jember, Bojonegoro, Lamongan dan Ngawi mendekati nilai observasi. Hal ini dapat dikatakan bahwa model PPR dengan banyak fungsi  $m=2$  baik untuk meramalkan curah hujan di Kabupaten Banyuwangi sedangkan banyak fungsi  $m=1$  baik untuk meramalkan curah hujan di Kabupaten Jember, Bojonegoro, Lamongan dan Ngawi.

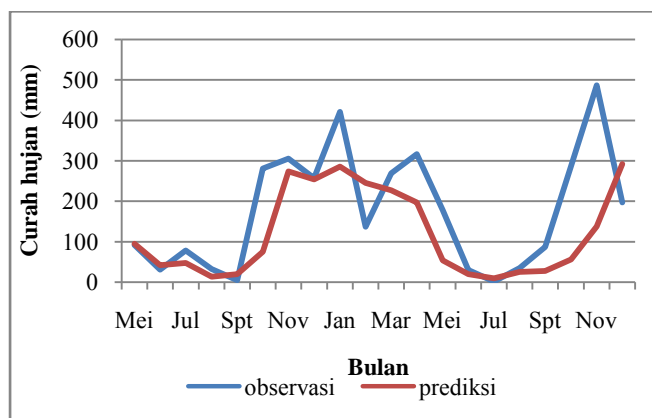




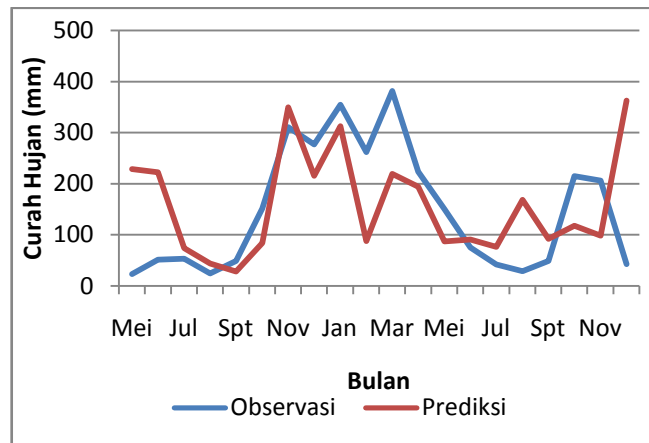
**Gambar 4.6** Plot antara Observasi dan Prediksi Curah Hujan di Kabupaten Banyuwangi dengan PPR



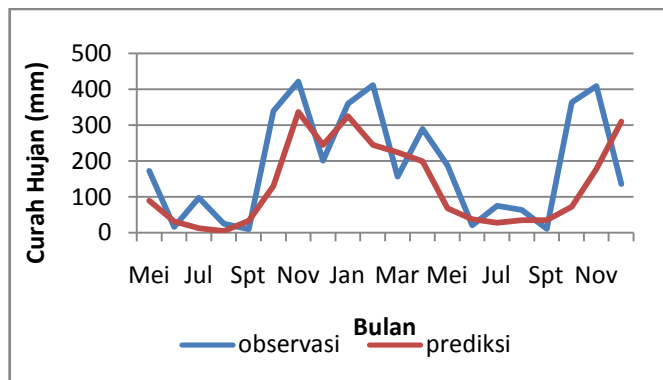
**Gambar 4.7** Plot antara Observasi dan Prediksi Curah Hujan di Kabupaten Jember dengan PPR



**Gambar 4.8** Plot antara Observasi dan Prediksi Curah Hujan di Kabupaten Bojonegoro dengan PPR



**Gambar 4.9** Plot antara Observasi dan Prediksi Curah Hujan di Kabupaten Lamongan dengan PPR



**Gambar 4.10** Plot antara Observasi dan Prediksi Curah Hujan di Kabupaten Ngawi dengan PPR

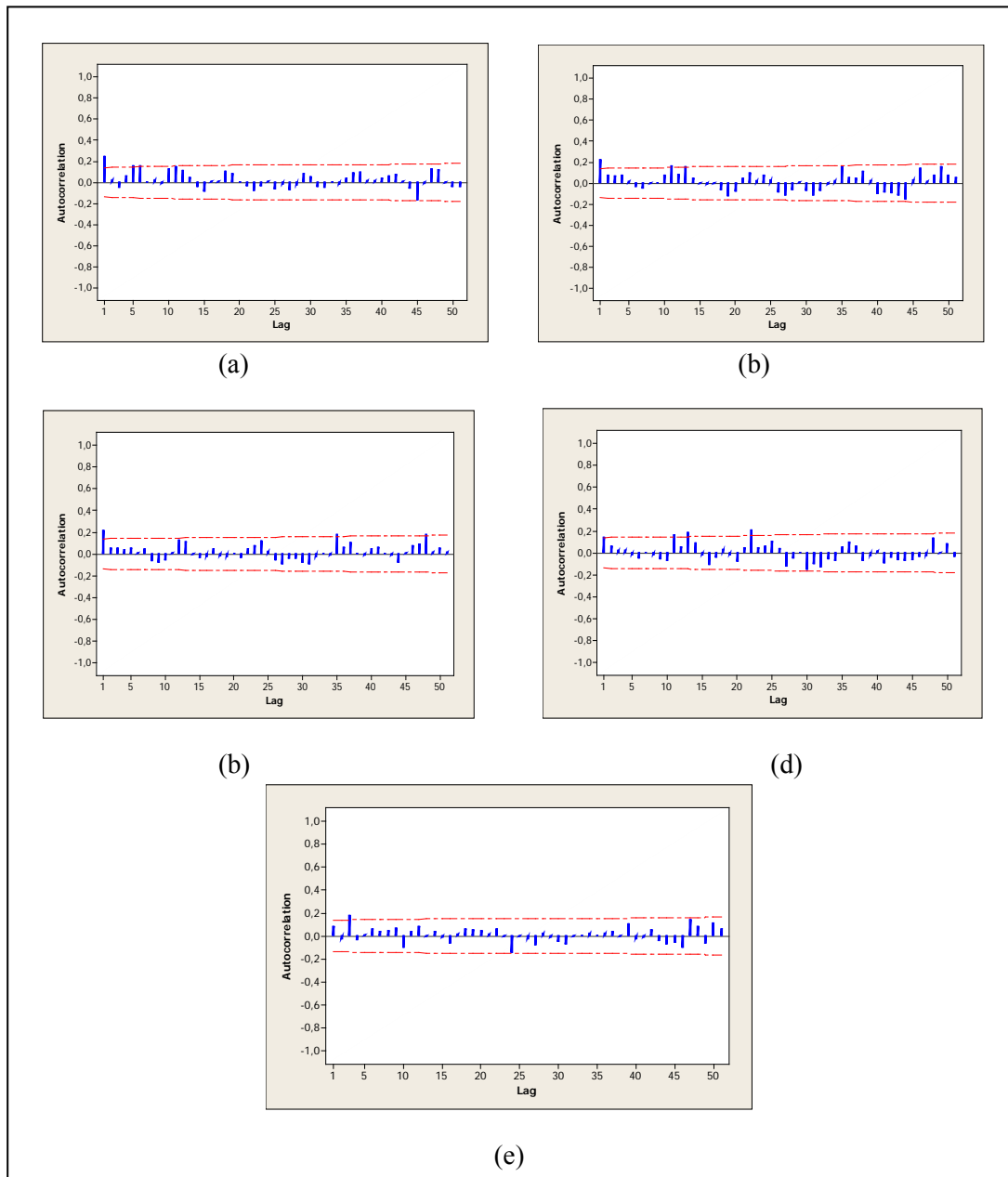
### 4.3 Penyusunan Model *Hybrid*

Modifikasi model adalah memodelkan residual dari model PPR yang telah dibangun. Sehingga model akhir yang diperoleh adalah gabungan (*hybrid*) dari model utama PPR dan model dari residualnya.

#### 1) Identifikasi Autokorelasi Residual

Pada penyusunan model hybrid, langkah awal yang perlu dilakukan adalah identifikasi terhadap residual apakah terdapat korelasi antar series atau biasa disebut autokorelasi. Terdapat dua cara untuk menguji autokorelasi yaitu secara visual dengan plot ACF dan pengujian statistik L-jung Box. Pada Gambar 4.11 nampak bahwa terdapat lag yang keluar batas ACF sehingga dapat dikatakan

bahwa terdapat autokorelasi pada residual curah hujan di seluruh daerah. Pada Kabupaten Banyuwangi nampak lag 1, 5, dan 6 keluar batas ACF, sedangkan pada Kabupaten Jember lag 1, 11, dan 13 yang keluar batas ACF.



**Gambar 4.11** Plot ACF Residual Curah Hujan di Kabupaten (a) Banyuwangi, (b) Jember, (c) Bojonegoro, (d) Lamongan, dan (e) Ngawi

Pada Tabel 4.3 ditunjukkan bahwa seluruh lag yang keluar batas signifikan atau dapat disimpulkan secara statistik terbukti adanya kasus autokorelasi pada

residual curah hujan di Kabupaten Banyuwangi, Jember, Bojonegoro, Lamongan dan Ngawi, sehingga untuk mengatasi hal ini dilakukan modifikasi model *hybrid* dengan menggunakan ASTAR.

**Tabel 4.3** Nilai ACF, L-jung Box, dan *P-Value*

Kabupaten	Lag	ACF	L-jung Box	<i>p-value</i>
Banyuwangi	1	0,250	12,94	0,0003*
	5	0,157	19,56	0,0015*
	6	0,162	25,12	0,0003*
Jember	1	0,226	10,55	0,0011*
	11	0,169	22,48	0,0209*
	13	0,161	29,62	0,0053*
Bojonegoro	1	0,216	9,64	0,0019*
Lamongan	1	0,144	4,27	0,0388*
	13	0,190	22,83	0,0437*
	22	0,212	40,48	0,0095*
Ngawi	3	0,179	8,38	0,0387*

Ket: (\*) signifikan pada  $\alpha = 5\%$

## 2) Penyusunan Model Residual Dengan ASTAR

Seperti halnya pada PPR, langkah awal pada pembentukan model ASTAR yaitu dilakukan simulasi masukan agar diperoleh model yang optimal. Beberapa parameter ASTAR yang perlu dikombinasikan yaitu maksimum interaksi (MI) yaitu 1, 2, dan 3, maksimum basis fungsi (BF) 24, 36, dan 48, minimum span (MS) 0, 5, dan 10, sedangkan derajat bebas (DF) 1, 2, dan 3. Setelah dikombinasikan keempat parameter ASTAR tersebut maka akan dipilih nilai GCV minimum untuk memenuhi konsep parsemoni, yaitu membuang fungsi basis yang memiliki kontribusi kecil terhadap respon dengan *forward stepwise*.

Pada Kabupaten Banyuwangi model akhir curah hujan dengan ASTAR diperoleh nilai GCV sebesar 6063. Kombinasi yang menghasilkan model terbaik adalah MI = 1, BF = 24, MS = 0, dan DF = 2 dengan RMSEP minimum 117,06. Sedangkan pada Kabupaten Jember, kombinasi masukan yaitu MI = 1, BF = 36, MS = 10, dan DF = 1 dengan nilai GCV = 10582 dan RMEP minimum 111,49.

Secara lengkap kombinasi masukan ASTAR optimal untuk Kabupaten Bojonegoro, Lamongan, dan Ngawi disajikan pada Lampiran 3, 4, dan 5.

**Tabel 4.4** Model Residual dengan ASTAR di Lima Kabupaten

Kabupaten	Model
Banyuwangi	$\hat{e}_t = -33,41 + 0,81 (e_{t-1} - 75)_+ - 1,4 (e_{t-4} - 71)_+ + 0,7 (e_{t-4} + 1)_+ + 0,17 (e_{t-6} + 172)_+ - 1,1 (-82 - e_{t-8})_+ - 20,43 (e_{t-11} - 144)_+ - 4,59 (e_{t-11} - 98)_+ + 26,2 (e_{t-11} - 135)_+ - 0,58 (e_{t-12} + 51)_+ + 1,12 (e_{t-12} - 14)_+$
Jember	$\hat{e}_t = -1,03 - 2,78 (e_{t-1} + 70)_+ - 0,87 (e_{t-1} - 122)_+ + 2,67 (e_{t-1} + 35)_+ + 0,74 (e_{t-1} + 220)_+ - 0,71 (-35 - e_{t-4})_+ - 4,58 (e_{t-9} + 20)_+ - 9,91 (e_{t-9} + 49)_+ + 14,36 (e_{t-9} + 40)_+ - 0,29 (116 - e_{t-10})_+ - 0,72 (e_{t-10} - 116)_+ + 5,31 (e_{t-11} + 20)_+ - 4,06 (e_{t-11} - 0)_+ - 1,26 (e_{t-11} + 70)_+$
Bojonegoro	$\hat{e}_t = 32,89 + 0,22 (e_{t-1} + 62)_+ - 0,47 (-25 - e_{t-3})_+ - 1,29 (e_{t-4} - 47)_+ + 1,58 (e_{t-4} - 100)_+ - 0,48 (20 - e_{t-4})_+ + 0,3 (16 - e_{t-8})_+ - 3,18 (e_{t-11} - 69)_+ + 4,38 (e_{t-11} - 99)_+ - 0,26 (14 - e_{t-12})_+$
Lamongan	$\hat{e}_t = -441,75 + 1,84 (e_{t-1} - 228)_+ - 1,69 (e_{t-1} - 135)_+ + 3,56 (-163 - e_{t-1})_+ + 0,33 (e_{t-1} + 163)_+ + 0,61 (-87 - e_{t-2})_+ + 0,3 (65 - e_{t-4})_+ + 0,4 (e_{t-4} - 65)_+ + 1,06 (e_{t-11} + 133)_+ + 1,05 (127 - e_{t-11})_+ + 2,09 (e_{t-12} - 293)_+ - 1,02 (e_{t-12} + 66)_+ + 3,4 (-163 - e_{t-12})_+ + 0,85 (e_{t-12} + 163)_+$
Ngawi	$\hat{e}_t = -48,86 + 0,35 (e_{t-9} + 93)_+ - (e_{t-3} - 54)_+ (84 - e_{t-12})_+ (e_{t-3} + 314)_+ + (84 - e_{t-12})_+ (-76 - e_{t-7})_+ (e_{t-9} + 93)_+ + 0,01 (e_{t-9} + 93)_+ (-54 - e_{t-10})_+ (e_{t-9} + 314)_+ - (84 - e_{t-12})_+$

Pada Tabel 4.5 dapat diketahui bahwa pada setiap model curah hujan yang terbentuk, lag yang relatif banyak mempengaruhi model berbeda-beda. Namun secara keseluruhan terlihat bahwa lag 1, lag 4, dan lag 11 relatif banyak mempengaruhi model. Hal ini berarti bahwa dugaan  $t$  ( $e_t$ ) dipengaruhi oleh nilai pengamatan pada  $t = 1$  ( $e_{t-1}$ ),  $t = 4$  ( $e_{t-4}$ ), dan  $t = 11$  ( $e_{t-11}$ ).

Model residual yang terbentuk merupakan model yang kompleks karena terdiri dari beberapa basis fungsi. Pada Kabupaten Banyuwangi, model residual tidak terdapat interaksi.

$$BF1 = (e_{t-1} - 75)_+ = \begin{cases} (e_{t-1} - 75) & , \text{jika } e_{t-1} > 75 \\ 0 & , \text{lainnya} \end{cases}$$

Model tersebut menunjukkan bahwa, setiap kenaikan satu satuan BF1 akan menyebabkan kenaikan residual sebesar 0,81 mm. Hal ini terjadi apabila residual satu bulan sebelumnya lebih besar dari 75 mm.

### 3) Validasi Model *Hybrid* PPR dan ASTAR

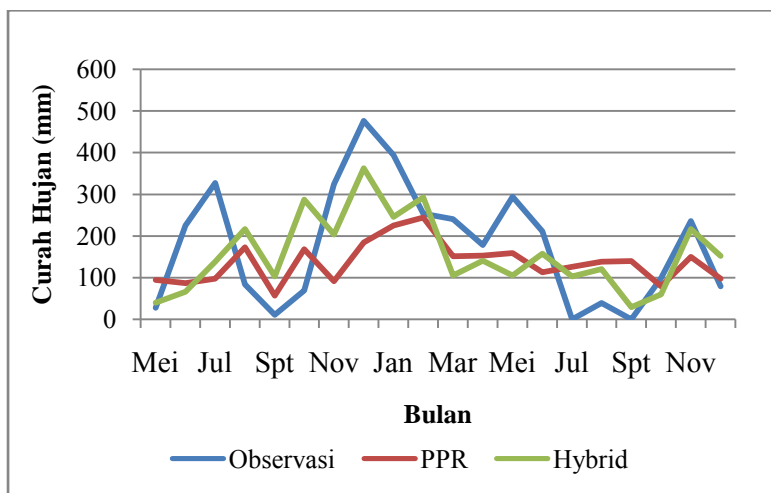
Model *hybrid* adalah gabungan dari model PPR dan ASTAR. Residual dari model PPR karena dideteksi adanya autokorelasi maka diatasi dengan memodelkan residual tersebut dengan ASTAR. Validasi model *hybrid* menunjukkan bahwa terjadi perbaikan nilai RMSEP jika dibandingkan dengan model *non-hybrid* baik untuk data *in-sample* maupun *out-sample*. Untuk data *out-sample* model curah hujan di Kabupaten Banyuwangi, RMSEP model *non-hybrid* adalah 133,32 dengan  $R^2_{\text{pred}}$  sebesar 21,0% namun dengan menggunakan model *hybrid* nilai RMSEP semakin kecil yaitu 117,19 dengan nilai korelasi 62,8%. Sedangkan untuk data *in-sample* nilai RMSEP pada model *non-hybrid* adalah 79,3 dan semakin kecil menjadi 68,5 pada model *hybrid*. Hal ini berarti bahwa tingkat ketepatan prediksi pada model *hybrid* PPR dan ASTAR lebih baik dibanding dengan model PPR saja.

Pada Kabupaten Lamongan, nilai RMSEP pada model *hybrid* lebih kecil dari model *non-hybrid* namun  $R^2_{\text{pred}}$  pada model *hybrid* tidak lebih baik. Sedangkan pada Kabupaten Ngawi, RMSEP pada model *hybrid* sedikit lebih besar dibandingkan model *non-hybrid*, namun  $R^2_{\text{pred}}$  lebih baik yaitu dari 72,2% menjadi 80,1%. Hal ini dimungkinkan terjadi karena kasus autokorelasi residual pada model curah hujan tidak terlalu besar.

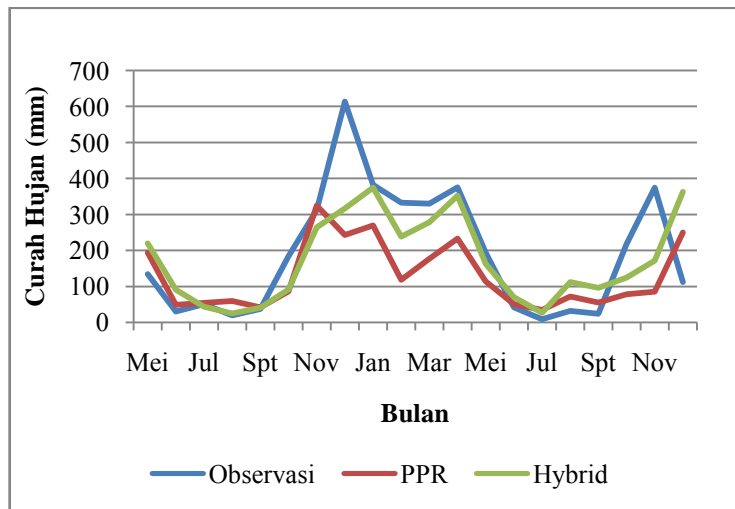
**Tabel 4.5** Perbandingan RMSEP dan  $R^2_{pred}$  Model *Non-Hybrid* dan *Hybrid* ASTAR

Kabupaten	Out-Sample				In-Sample			
	Non-Hybrid		Hybrid		Non-Hybrid		Model Hybrid	
	RMSEP	$R^2_{pred}$	RMSEP	$R^2_{pred}$	RMSEP	$R^2_{pred}$	RMSEP	$R^2_{pred}$
Banyuwangi	133,32	21,0%	117,19	62,8%	79,3	47,4%	68,5	61,5%
Jember	139,23	42,0%	111,49	49,9%	105,9	44,9%	90	74,3%
Bojonegoro	121,31	70,0%	97,75	52,1%	89,7	60,4%	78,1	78,6%
Lamongan	120,79	70,3%	118,25	66,7%	105,6	44,2%	79,8	73,6%
Ngawi	122,04	72,2%	123,32	80,1%	98,3	58,7%	88,5	70,0%

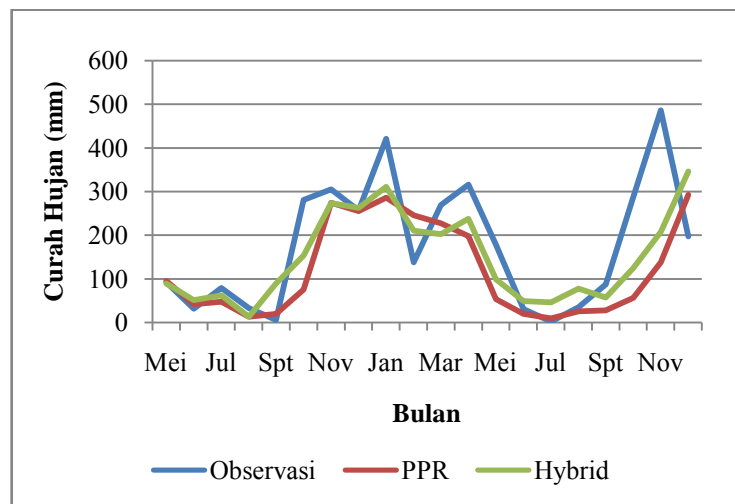
Secara visual tampak pada Gambar 4.12 s ampai Gambar 4.16 bahwa prediksi curah hujan pada model PPR atau *non-hybrid* berada di bawah *plot* model *hybrid*. Hal ini berarti bahwa ketepatan prediksi curah hujan pada model *hybrid* lebih mendekati nilai observasi dibandingkan dengan prediksi pada model *non hybrid*.



**Gambar 4.12** Plot antara Observasi dan Prediksi Curah Hujan di Kabupaten Banyuwangi dengan PPR dan *Hybrid* ASTAR

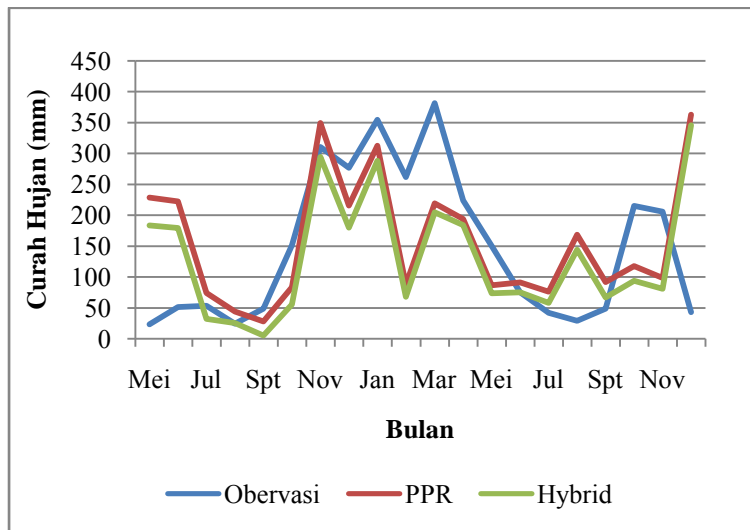


**Gambar 4.13** Plot antara Obervasi dan Prediksi Curah Hujan di Kabupaten Jember dengan PPR dan *Hybrid* ASTAR

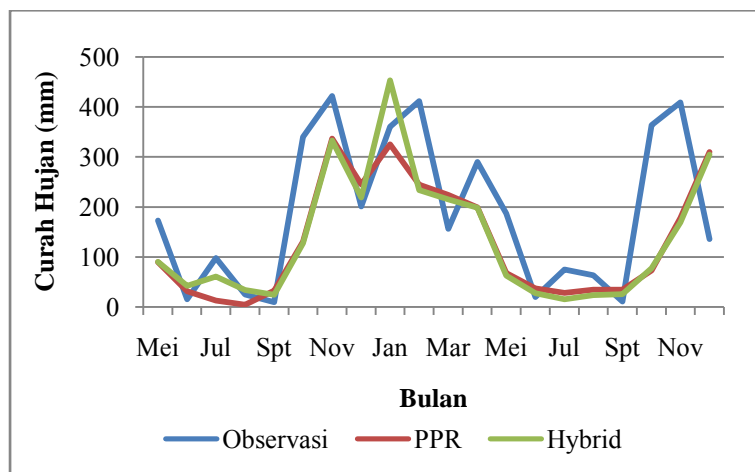


**Gambar 4.14** Plot antara Obervasi dan Prediksi Curah Hujan di Kabupaten Bojonegoro dengan PPR dan *Hybrid* ASTAR





**Gambar 4.15** Plot antara Observasi dan Prediksi Curah Hujan di Kabupaten Lamongan dengan PPR dan *Hybrid* ASTAR



**Gambar 4.16** Plot antara Observasi dan Prediksi Curah Hujan di Kabupaten Ngawi dengan PPR dan *Hybrid* ASTAR

## BAB V

### KESIMPULAN DAN SARAN

#### 5.1 Kesimpulan

Berdasarkan analisis dan pembahasan diperoleh beberapa kesimpulan sebagai berikut.

1. Model curah hujan dengan menggunakan metode *hybrid* PPR dan ASTAR pada Kabupaten Banyuwangi, Bojonegoro, Jember, Lamongan dan Ngawi.

Banyuwangi  $\hat{y} = 138,17 + 78,28(0,144X_1 - 0,109X_2 - 0,144X_3 - 0,119X_4 - 0,594X_5 - 0,425X_6 + 0,432X_7 + 0,438X_8 - 0,138X_9) + 43,66(-0,241X_1 - 0,353X_2 - 0,413X_3 + 0,198X_4 - 0,108X_5 + 0,410X_6 - 0,202X_8 + 0,488X_9) + \hat{e}_t$

dimana  $\hat{e}_t = -33,41 + 0,81(e_{t-1} - 75)_+ - 1,4(e_{t-4} - 71)_+ + 0,7(e_{t-4} + 1)_+ + 0,17(e_{t-6} + 172)_+ - 1,1(-82 - e_{t-8})_+ - 20,43(e_{t-11} - 144)_+ - 4,59(e_{t-11} - 98)_+ + 26,2(e_{t-11} - 135)_+ - 0,58(e_{t-12} + 51)_+ + 1,12(e_{t-12} - 14)_+$

Jember  $\hat{y} = 166,40 + 94,92(0,08X_1 - 0,007X_2 + 0,028X_3 - 0,571X_4 + 0,267X_5 + 0,472X_6 - 0,317X_7 - 0,438X_8 + 0,288X_9) + \hat{e}_t$

dimana  $\hat{e}_t = -1,03 - 2,78(e_{t-1} + 70)_+ - 0,87(e_{t-1} - 122)_+ + 2,67(e_{t-1} + 35)_+ + 0,74(e_{t-1} + 220)_+ - 0,71(-35 - e_{t-4})_+ - 4,58(e_{t-9} + 20)_+ - 9,91(e_{t-9} + 49)_+ + 14,36(e_{t-9} + 40)_+ - 0,29(116 - e_{t-10})_+ - 0,72(e_{t-10} - 116)_+ + 5,31(e_{t-11} + 20)_+ - 4,06(e_{t-11} - 0)_+ - 1,26(e_{t-11} + 70)_+$

Bojonegoro  $\hat{y} = 173,20 + 112,63(-0,921X_1 + 0,181X_2 + 0,003X_3 + 0,128X_4 + 0,077X_5 - 0,245X_6 - 0,187X_7 + 0,037X_8 + 0,02X_9) + \hat{e}_t$

dimana  $\hat{e}_t = 32,89 + 0,22(e_{t-1} + 62)_+ - 0,47(-25 - e_{t-3})_+ - 1,29(e_{t-4} - 47)_+ + 1,58(e_{t-4} - 100)_+ - 0,48(20 - e_{t-4})_+ + 0,3(16 - e_{t-8})_+ - 3,18(e_{t-11} - 69)_+ + 4,38(e_{t-11} - 99)_+ - 0,26(14 - e_{t-12})_+$

Lamongan  $\hat{y} = 170,44 + 93,52(0,0002X_1 - 0,4X_2 + 0,004X_3 + 0,861X_4 - 0,031X_5 - 0,13X_6 - 0,093X_7 - 0,0377X_8 - 0,262X_9) + \hat{e}_t$

dimana  $\hat{e}_t = -441,75 + 1,84(e_{t-1} - 228)_+ - 1,69(e_{t-1} - 135)_+ + 3,56(-163 - e_{t-1})_+ + 0,33(e_{t-1} + 163)_+ + 0,61(-87 - e_{t-2})_+ + 0,3(65 - e_{t-4})_+ + 0,4(e_{t-4} - 65)_+ + 1,06(e_{t-11} + 133)_+ + 1,05(127 - e_{t-11})_+ + 2,09(e_{t-12} - 293)_+ - 1,02(e_{t-12} + 66)_+ + 3,4(-163 - e_{t-12})_+ + 0,85(e_{t-12} + 163)_+$

Ngawi  $\hat{y} = 179,00 + 112,11(-0,877X_1 + 0,228X_2 + 0,238X_3 + 0,143X_4 + 0,083X_5 - 0,207X_6 - 0,323X_7 + 0,051X_8 + 0,035X_9) + \hat{e}_t$

dimana  $\hat{e}_t = -48,86 + 0,35 (e_{t-9} + 93)_+ - (e_{t-3} - 54)_+ (84 - e_{t-12})_+ (e_{t-3} + 314)_+ + (84 - e_{t-12})_+ (-76 - e_{t-7})_+ (e_{t-9} + 93)_+ + 0,01 (e_{t-9} + 93)_+ (-54 - e_{t-10})_+ (e_{t-9} + 314)_+ - (84 - e_{t-12})_+$

2. Hasil validasi model *hybrid* PPR dan ASTAR dengan kriteria RMSEP menunjukkan kinerja yang lebih baik jika dibandingkan dengan model *non hybrid*.

## 5.2 Saran

Beberapa saran yang direkomendasikan untuk pengembangan penelitian selanjutnya yaitu:

1. Kriteria kebaikan model sebaiknya tidak hanya menggunakan RMSE dan  $R^2$  tetapi menggunakan kriteria lain misalnya MAPE, karena hasil dari keduanya tidak berbanding lurus.
2. Membedakan antara pemodelan curah hujan untuk musim kemarau dan musim hujan. Hal ini dimungkinkan memberi hasil ramalan yang lebih baik.

## DAFTAR PUSTAKA

- Aldrin, M., Bolviken, E., dan Schweder, Tore. 1991. Projection Pursuit Regression for Moderate Non-Linearities. *Computational Statistics & Data Analysis* 16: 379-403.
- BMG, 2008. *Pola Curah Hujan*. <<http://www.bmg.go.id>>
- Boer, R. dan I. Las. 2003. Sistem Produksi Padi Nasional Dalam Perspektif Kebijakan Iklim Global. Hlm 215-234. *Dalam Bambang S. et al.* (Ed). Kebijakan Perberasan dan Inovasi Teknologi Padi: Buku Dua. Prosiding Pertemuan Ilmiah Bagian Proyek Litbang Tanaman Padi, 2003. Sukamandi.
- Busuioc A, Chen D, & Hellström C. 2001. Performance of Statistical Downscaling Models in GCM Validation and Regional Climate Change: Application for Swedish Precipitation. *Int. J. Climatology* 21: 557-578.
- Cavaeos T, Hewitson B. 2002. Relative Performance of Empirical Predictors of Daily Precipitation. *Proc. of the 1<sup>st</sup> Biennial Meeting of the IEMSS*, Lugano, Switeerland, 2:349-354.
- Cheng, Ping. & Li, Guo-Ying. 1993. Some Recent Development in Projection Pursuit in China. *Statistics Sinica* 3: 35-51.
- Crane RG, Hewitson BC. 1998. Doubled CO<sub>2</sub> Precipitation Changes for The Susquehanna Basin: Downscaling from GENESIS General Circulation Model. *Int. J. Climatology* 18:65-76.
- Corte-Real J, Ehang X, Wang X. 1995. Downscaling GCM Information to Regional Scales: A Non-parametric Multivariate Regression Approach. *Climate Dynamics* 11 (7): 413-424.
- Cryer, JD. & Chan Kung-Sik. 2008. *Time Series Analysis With Application in R Second Edition*. USA: Springer.
- Fernandez, E. 2005. On The Influence of Predictors Area in Statistical Downscaling of Daily Parameters. *Report Climate* no.09/2005. Norwegian Meteorological Institute. Oslo.
- Fischer M, Dewitte B, & Maîtrepierre L. 2004., A non-linear statistical downscaling model: El Niño/Southern Oscillation impact on precipitation over New Caledonia, *Geophys. Res. Lett.* 31, L16204, doi:10.1029/2004GL020112.
- Friedman, JH. 1991. Multivariate Adaptive Regression Splines (with Discussion). *The Annals of Statistics*, 19:1-141.
- Friedman, JH & Stuetele, Werner. 1981. Projection Pursuit Regression. *Journal of The American Statistical Association* 376: 817-823.

- Hall, P. 1989. On Projection Pursuit Regression. *The Annals of Statistics* 17(2): 573-588.
- Haryoko U.2004. Pendekatan Reduksi Dimensi Luaran GCM untuk Penyusunan Model *Statistical Downscaling* [Tesis]. Bogor: Sekolah Pascasarjana, Institut Pertanian Bogor.
- Istriana. 2009. Peramalan Curah Hujan dengan Menggunakan Metode ARIMA dan ASTAR. Tugas Akhir. Institut Teknologi Sepuluh Nopember Surabaya.
- Jones MC & Sibson R. 1987. What is Projection Pursuit?. *J R Statistic Social A.* 150:1-36.
- Lewis, PAW & Ray, BK. 1997. Modeling Long-Range Dependence, Nonlinearity, and Periodic Phenomena in Sea Surface Temperatures Using TSMARS. *J. Amer. Statist. Assoc.* Vol. 92. No. 439. pp. 881-893.
- Lewis, PAW & Stevens, JG. 1991. Nonlinear Modelling of Time Series Using Multivariate Adaptive Regression Splines. *J. Amer. Statist. Assoc* 86: 864 – 877.
- Nason, Guy Philip Nason. 1992. Design and Choice of Projection Indices. Thesis for the Degree Ph.D. University of Bath.
- Notodiputro KA, Wigena AH, & Fitriadi. 2004. Pendekatan Regresi Komponen Utama dan ARIMA dalam Statistical Downscaling. Forum Statistika dan Komputasi. Edisi Khusus Seminar Nasional Statistika.
- Runtuuwu, E. dan H. Syahbuddin. 2007. Perubahan Pola Curah Hujan dan Dampaknya Terhadap Periode Masa Tanam. *Jurnal Tanah dan Iklim.* 26: 1–12.
- Safitri, Rina. 2012. Pemodelan MOS dengan Pendekatan PPR untuk Meramalkan Suhu Minimum, Suhu Maksimum, dan Kelembaban. Tugas Akhir. Institut Teknologi Sepuluh Nopember Surabaya.
- Sutikno. 2008. *Statistical Downscaling* Luaran GCM dan Pemanfaatannya untuk Peramalan Produksi Padi. Disertasi. Bogor: Institut Pertanian Bogor.
- Sutikno, Bakti RD, Susanti P. & Istriana. 2004. Prakiraan Cuaca Dengan Metode ARIMA, NN, dan ASTAR di Stasiun Juanda Surabaya. *Jurnal Sains Dirgantara* 8: 43-61.
- Sutikno dan Boer R. 2005. Penggunaan Regresi *Splines* Adaptif Berganda, Regresi Komponen Utama, dan Jaringan Syaraf Tiruan untuk Statistical Downscaling GCM. *Di Dalam Prosiding Seminar Nasional Statistika VII ITS Surabaya.* 26 Nopember 2005.
- Wei, WWS. 2006. *Time Series Analysis: Univariate and Multivariate Methods Second Edition.* USA : Addison - Wesley Publishing Co.

- Wigena, AH. 2006. *Pemodelan Statistical Downscaling dengan Regresi Projection Pursuit* untuk Peramalan Curah Hujan Bulanan. Disertasi. Bogor: Institut Pertanian Bogor.
- Wilby RL, Charles SP, Eorita E, Timbal B, Whetton P, Mearns LO. 2004. Guidelines for Use of Climate Scenarios Developed from Statistical Downscaling Methods. *Intergovernmental Panel on Climate Change*. Australia.
- Wilby RL & Wigley TML. 1997. Downscaling General Circulation Model Output: A review of methods and limitations. *Progress in Physical Geoghrapy*. 4:530-548.
- Ehang, X., Xiaou Tang, L., dan Yeung Shum, Heung.2008.  $L_1$  Regularieed Projection Pursuit for Additive Model Learning. Beijing-China.

## BIODATA PENULIS



Penulis yang bernama Noviyanti Santoso ini dilahirkan di Surabaya, 30 Nopember 1987. Penulis merupakan anak kedua dari tiga bersaudara. Penulis telah menikah dan memiliki satu putra. Pendidikan formal yang pernah ditempuh antara lain : TK Dharma Wanita Sampang, SDN Pranti Sedari, SMPN 1 Waru, SMAN 16 Surabaya, dan S1 Statistika ITS.

Penulis diterima di Pascasarjana Statistika FMIPA-ITS tahun 2012 melalui beasiswa unggulan calon dosen DIKTI. Percaya dan yakin bahwa Allah tidak akan membebani makhlukNYA diluar batas kemampuan kita menghadapinya. Saran dan kritik dapat dikirim melalui email : [nanami.santo@gmail.com](mailto:nanami.santo@gmail.com).

## DAFTAR LAMPIRAN

	halaman
Lampiran 1. Statistika Deskriptif Curah Hujan Tiap Bulan.....	51
Lampiran 2. <i>Plot</i> Antara Variabel Respon dan Variabel Prediktor di Kabupaten Lamongan dan Ngawi .....	54
Lampiran 3. Kombinasi Masukan ASTAR Optimal untuk Kabupaten Bojonegoro .....	55
Lampiran 4. Kombinasi Masukan ASTAR Optimal untuk Kabupaten Lamongan .....	56
Lampiran 5. Kombinasi Masukan ASTAR Optimal untuk Kabupaten Ngawi .....	57
Lampiran 6. Syntax PPR.....	58
Lampiran 7. Syntax Uji Ljung-Box .....	59
Lampiran 8. Data Curah Hujan dan Presipitasi Kabupaten Bojonegoro .....	60





## LAMPIRAN

### Lampiran 1. Statistika Deskriptif Curah Hujan Tiap Bulan

#### Banyuwangi

Bulan	Rataan	Standar Deviasi	Koefisien Variasi
Januari	298,1	92,9	31,18
Februari	264,9	72,7	27,43
Maret	178,7	93,2	52,16
April	104,5	71,2	68,1
Mei	94,5	75,2	79,58
Juni	96,8	78,6	81,15
Juli	101,4	85,6	84,42
Agustus	71,6	62,4	87,12
September	43	49,8	115,69
Oktober	67,1	72,5	108,07
Nopember	126,9	102	80,41
Desember	210,6	111,6	53,01

#### Bojonegoro

Bulan	Rataan	Standar Deviasi	Koefisien Variasi
Januari	369,6	83,7	22,66
Februari	277,2	87	31,38
Maret	277,8	98,7	35,53
April	224,3	91	40,57
Mei	96,4	68	70,52
Juni	58,3	64,1	109,89
Juli	31,59	39,38	124,66
Agustus	23,97	26,78	111,71
September	34,3	45,9	133,66
Oktober	132	107,9	81,74
Nopember	249,9	121,1	48,44
Desember	302,9	106,8	35,26

## Jember

Bulan	Rataan	Standar Deviasi	Koefisien Variasi
Januari	350,7	121,9	34,75
Februari	307,3	100,7	32,77
Maret	260,1	106,2	40,82
April	183	105,2	57,48
Mei	88,3	70,1	79,39
Juni	76	90,3	118,91
Juli	36,88	38,27	103,76
Agustus	19,74	20,12	101,97
September	35,2	43,5	123,64
Oktober	120	96,5	80,39
Nopember	232,2	80,4	34,63
Desember	287,6	148,6	51,66

## Lamongan

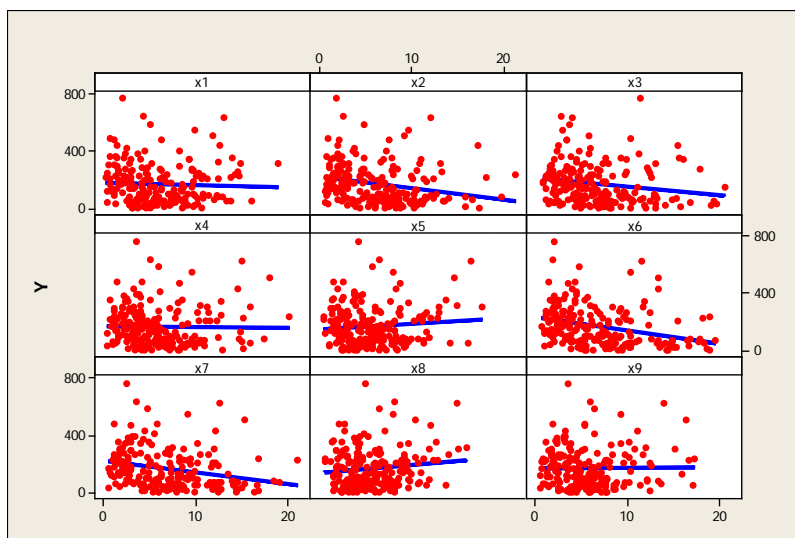
Bulan	Rataan	Standar Deviasi	Koefisien Variasi
Januari	385,9	176,5	45,72
Februari	278	141,4	50,86
Maret	248,4	78,6	31,65
April	160,1	86	53,73
Mei	92,4	68,9	74,6
Juni	97,6	65,6	67,22
Juli	57,7	50,1	86,93
Agustus	39,5	26,05	65,95
September	58,5	44,3	75,77
Oktober	131,7	84,9	64,46
Nopember	210,1	67,8	32,25
Desember	285,2	121,4	42,55

## Ngawi

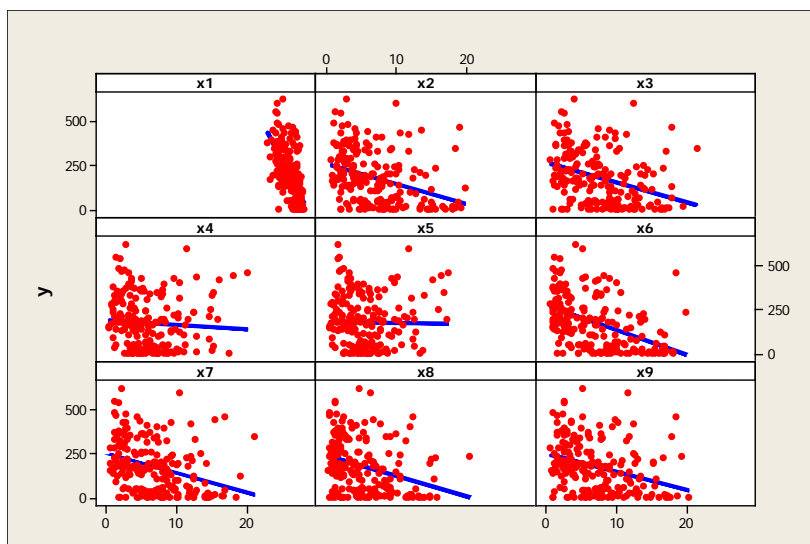
Bulan	Rataan	Standar Deviasi	Koefisien Variasi
Januari	330,1	140,2	42,46
Februari	358,2	85,6	23,9
Maret	267,1	160,3	60
April	218,3	93,5	42,85
Mei	112,3	81,4	72,48
Juni	78,6	73,2	93,15
Juli	40	57	142,42
Agustus	31,44	35,75	113,69
September	41,6	52	124,96
Oktober	179,4	109,5	61,03
Nopember	248,9	133,5	53,64
Desember	241,7	93,6	38,7

**Lampiran 2. Plot Antara Variabel Respon dan Variabel Prediktor di Kabupaten Lamongan dan Ngawi**

**Lamongan**



**Banyuwangi**



**Lampiran 3. Kombinasi Masukan ASTAR Optimal untuk Kabupaten  
Bojonegoro**

MI	BF	MS	DF	GCV	RMSEP
1	24	0	1	10582	114,96
2	24	0	1	6386	122,66
3	24	0	1	6116	109,5
1	24	5	1	7173	109,91
2	24	5	1	6461	115,85
3	24	5	1	6364	116,94
1	24	10	1	7202	101,58
2	24	10	1	6460	110,03
3	24	10	1	6919	114,22
1	24	0	2	7587	125,47
2	24	0	2	7085	119,65
3	24	0	2	6801	106,21
1	24	5	2	7469	109,91
2	24	5	2	7309	113,77
3	24	5	2	7203	113,5
1	24	10	2	7591	98,99
2	24	10	2	7174	110
3	24	10	2	7299	111
1	24	0	3	7662	125,47
2	24	0	3	7701	119,65
3	24	0	3	7307	114,5
1	24	5	3	7654	125,16
2	24	5	3	7517	113,77
3	24	5	3	7625	113,5
1	24	10	3	7793	125,82
2	24	10	3	7673	113,71
3	24	10	3	7509	111,21
1	36	0	1	7130	119,62
2	36	0	1	6324	120,63
3	36	0	1	5672	110,86
1	36	5	1	7345	107,65
2	36	5	1	5963	113,56
3	36	5	1	5954	122,89
1	36	10	1	7211	103,6
2	36	10	1	6359	108,94
3	36	10	1	5716	115,16
1	36	0	2	7550	117,61

**Lampiran 4. Kombinasi Masukan ASTAR Optimal untuk Kabupaten Lamongan**

MI	BF	MS	DF	GCV	RMSEP
1	24	0	1	8778	119,81
2	24	0	1	6344	120
3	24	0	1	6344	120
1	24	5	1	9068	122
2	24	5	1	7740	125
3	24	5	1	7436	122
1	24	10	1	9499	123
2	24	10	1	8057	133
3	24	10	1	7613	123
1	24	0	2	8777	118,25
2	24	0	2	7206	120
3	24	0	2	7206	120
1	24	5	2	9354	122
2	24	5	2	8472	126
3	24	5	2	8366	123
1	24	10	2	9771	123
2	24	10	2	8873	127
3	24	10	2	8250	122
1	24	0	3		
2	24	0	3	8102	124
3	24	0	3	8102	124
1	24	5	3		
2	24	5	3		
3	24	5	3		
1	24	10	3		
2	24	10	3	9390	124
3	24	10	3	8646	122
1	36	0	1	8024	125
2	36	0	1	5887	128
3	36	0	1	5392	123

**Lampiran 5. Kombinasi Masukan ASTAR Optimal untuk Kabupaten Ngawi**

MI	BF	MS	DF	GCV	RMSEP
1	24	0	1	8619	148,62
2	24	0	1	7480	172,38
3	24	0	1	7774	153,95
1	24	5	1	9473	128,25
2	24	5	1	7916	165
3	24	5	1	7988	142,77
1	24	10	1	9322	128,94
2	24	10	1	8787	130,1
3	24	10	1	7731	188,82
1	24	0	2	9119	150,79
2	24	0	2	8784	158,97
3	24	0	2	8369	161,12
1	24	5	2	9675	125,17
2	24	5	2	8646	160,65
3	24	5	2	8843	126,86
1	24	10	2	9578	128,94
2	24	10	2	9352	123,38
3	24	10	2	8784	141,01
1	24	0	3	9554	129,22
2	24	0	3	9213	137,41
3	24	0	3	8859	161,11
1	24	5	3	9754	122,46
2	24	5	3	9248	152,16
3	24	5	3	9268	128,53
1	24	10	3		
2	24	10	3	9603	124,95
3	24	10	3	9739	132,31
1	36	0	1	8658	150,79
2	36	0	1	7291	221,62
3	36	0	1	7200	204,43
1	36	5	1	9495	122,46
2	36	5	1	7776	195,37
3	36	5	1	7251	145,14
1	36	10	1	9583	122,38
2	36	10	1	8841	130,1
3	36	10	1	8185	246
1	36	0	2	9184	150,79
2	36	0	2	8875	152,11



## Lampiran 6. Syntax PPR

```
> Prediktor.Outsample<-matrix(c(Prediktor.Outsample),a,b)
> Judul<-ppreg(prediktorinsample,responisample,c,.)
> Hasil.ramalan<-ppreg(prediktorinsample,responisample,c,xpred=
Prediktor.Outsample)
```

Keterangan:

a = banyaknya baris dalam matriks

b = banyaknya kolom dalam matriks

c = banyak fungsi

## Lampiran 7. Syntax Uji Ljung-Box

```
> x=bwi
> Box.test(x,lag=1,type="Ljung")
Box-Ljung test
data: x
X-squared = 12.9439, df = 1, p-value = 0.000321
> Box.test(x,lag=5,type="Ljung")
Box-Ljung test
data: x
X-squared = 19.5601, df = 5, p-value = 0.001511
> Box.test(x,lag=6,type="Ljung")
Box-Ljung test
data: x
X-squared = 25.1172, df = 6, p-value = 0.0003248
```

```
> x=bjn
> Box.test(x,lag=1,type="Ljung")
Box-Ljung test
data: x
X-squared = 9.6377, df = 1, p-value = 0.001906
```

### Lampiran 8. Data Curah Hujan dan Presipitasi Kabupaten Bojonegoro

tahun	x1	x2	x3	x4	x5	x6	x7	x8	x9	y
1984	2,836	3,853	2,874	1,331	0,649	4,705	1,604	8,293	7,954	477,5
1984	8,620	3,219	0,806	2,282	0,574	4,909	0,775	5,720	2,808	416,5
1984	3,775	2,964	0,988	2,018	0,730	2,780	1,232	3,460	3,265	241,5
1984	5,952	8,176	6,937	6,733	4,095	5,588	5,593	6,679	8,502	117
1984	7,305	8,988	8,109	8,113	7,631	4,764	8,447	6,129	8,747	95,5
1984	8,114	11,839	16,576	10,694	11,774	6,153	15,660	7,436	11,848	28,5
1984	11,251	12,295	11,214	9,883	6,120	9,175	9,234	8,774	13,232	18
1984	11,237	11,066	8,543	13,346	11,556	8,583	10,547	7,840	9,986	65,5
1984	17,951	18,560	19,472	17,225	15,050	12,468	20,875	12,640	16,298	129,5
1984	9,181	10,902	13,171	8,168	17,635	5,413	14,961	6,727	10,600	54,5
1984	3,401	6,613	15,939	2,628	3,952	2,548	7,060	4,297	13,502	111,5
1984	3,902	7,188	8,097	1,302	1,794	5,906	3,311	7,373	12,842	198,5
1985	3,039	5,191	2,911	1,319	0,714	5,226	1,564	9,126	6,732	156,5
1985	2,701	4,670	1,284	2,357	1,041	3,025	1,606	5,148	5,374	315,5
1985	4,867	4,010	0,467	3,091	0,355	3,488	0,613	4,104	3,191	332,5
1985	2,987	3,359	2,168	1,916	2,123	5,368	2,285	7,320	4,952	78,5
1985	4,739	8,855	8,596	6,734	3,936	4,314	7,535	6,794	9,630	58,5
1985	11,120	12,850	12,992	12,398	9,210	7,631	12,612	9,029	13,369	159,5
1985	7,924	9,116	9,644	8,994	7,264	6,397	9,804	6,281	8,175	167,5
1985	10,829	10,514	9,882	13,117	10,855	8,421	7,889	9,288	11,973	31,5
1985	14,839	14,120	10,474	15,716	11,731	10,151	10,184	10,833	12,342	58
1985	9,976	6,979	6,966	6,044	10,502	8,181	6,629	8,690	7,811	55
1985	6,785	9,299	10,229	7,979	5,805	4,745	10,278	9,820	8,467	206
1985	13,458	10,665	5,058	5,132	2,510	10,450	4,571	12,720	12,136	257
1986	8,036	8,228	3,637	6,659	3,041	8,372	4,612	13,775	7,017	331
1986	5,483	3,325	0,464	2,403	0,792	3,948	0,847	5,698	3,244	182,5
1986	5,047	5,601	1,818	4,839	1,897	4,816	2,678	6,668	5,180	276,5
1986	7,006	10,125	6,307	8,706	6,043	5,948	7,610	6,544	9,471	113,5
1986	7,816	10,450	10,301	8,808	7,119	6,420	8,767	7,824	10,143	4
1986	7,839	8,990	8,297	7,682	4,645	5,587	6,971	7,586	9,239	258,5
1986	7,618	10,630	12,318	10,039	5,655	5,421	9,789	6,674	11,068	205,5
1986	10,850	11,489	11,055	12,923	6,877	8,241	9,230	9,655	12,105	111,5
1986	14,940	14,385	9,914	18,953	11,924	10,948	8,295	8,799	11,046	54
1986	14,282	11,882	12,564	14,101	19,710	7,328	14,109	7,703	9,699	51,5
1986	4,934	5,719	7,113	3,575	3,309	5,122	5,811	6,444	4,919	95
1986	7,213	6,192	3,055	4,155	1,597	5,112	2,280	5,714	5,128	150
1987	4,906	4,199	1,498	2,203	1,022	4,776	1,106	8,150	5,339	259,5
1987	6,095	5,585	1,835	2,852	0,603	5,234	1,556	6,900	5,188	275
1987	6,019	5,659	1,461	4,668	1,636	7,801	2,089	11,436	6,032	74

tahun	x1	x2	x3	x4	x5	x6	x7	x8	x9	y
1987	7,263	9,332	7,819	6,444	3,975	5,782	6,739	5,646	9,948	2
1987	7,979	10,593	7,928	9,338	8,972	6,258	9,403	7,020	10,100	116
1987	9,441	14,968	14,986	14,464	11,827	7,518	16,053	8,842	13,635	170
1987	11,070	15,917	12,999	14,278	14,250	8,510	14,105	8,331	14,852	61
1987	11,495	10,759	10,337	10,596	6,761	9,165	8,915	8,685	10,845	20
1987	9,255	11,781	12,366	10,599	9,293	7,869	11,203	9,887	13,674	0
1987	9,466	9,558	6,684	10,983	2,493	5,841	4,411	6,404	6,830	11
1987	8,072	7,547	6,913	5,134	4,623	7,455	6,700	12,798	10,465	86
1987	9,005	6,236	3,967	4,367	4,395	8,239	4,741	8,169	8,073	226,5
1988	8,348	3,578	1,082	2,897	0,862	4,564	1,336	5,454	2,799	207
1988	1,533	1,327	0,617	0,733	0,688	2,939	0,621	6,034	1,533	241,5
1988	4,015	3,647	1,122	1,883	0,662	4,066	1,079	6,295	4,697	199
1988	4,539	6,225	5,102	5,626	2,789	4,529	3,373	5,653	6,352	111,5
1988	6,273	9,625	8,895	7,544	6,648	5,556	8,476	6,310	8,999	55
1988	8,577	14,750	10,468	14,277	7,948	6,175	10,358	7,779	13,282	80,5
1988	6,824	7,847	7,966	7,199	5,837	5,371	9,184	6,354	8,365	44,5
1988	13,933	16,806	10,597	17,768	13,691	11,656	11,200	11,875	14,337	200
1988	13,715	16,831	18,173	12,660	13,488	10,573	13,005	12,009	19,028	57
1988	8,428	6,862	9,296	4,474	2,850	5,774	4,563	5,647	8,168	23
1988	2,339	3,355	2,325	1,660	1,186	3,049	2,073	4,457	4,012	89
1988	4,662	6,610	4,133	4,228	1,952	5,217	3,644	5,817	3,919	306,5
1989	2,736	2,271	0,395	1,522	0,550	1,828	0,626	3,221	1,906	264,5
1989	2,037	1,487	0,980	0,848	1,168	2,152	1,567	3,807	1,233	322
1989	3,936	2,707	2,099	1,331	0,750	4,201	1,086	6,088	5,931	139
1989	4,543	4,487	2,782	3,645	1,975	3,415	2,567	4,405	5,129	68,5
1989	8,275	9,660	7,618	9,995	8,639	6,590	8,470	8,076	8,143	179
1989	8,400	14,279	13,324	10,020	14,834	6,533	17,044	8,185	14,999	81
1989	8,527	10,784	10,093	10,720	8,959	6,023	8,643	7,816	10,832	116
1989	12,540	17,053	12,179	14,330	10,615	11,409	10,047	12,535	16,368	98
1989	9,834	9,106	10,193	9,834	10,733	7,358	10,133	8,810	8,463	20
1989	9,620	9,652	8,940	9,591	7,600	7,183	10,361	8,985	8,997	55
1989	4,676	4,071	4,491	5,030	4,192	3,468	3,451	4,739	5,625	64
1989	5,764	6,066	2,225	3,615	1,398	5,461	1,886	6,048	5,913	189
∴	∴	∴	∴	∴	∴	∴	∴	∴	∴	∴
2000	11,554	12,800	13,402	11,872	7,441	8,985	11,032	11,114	13,311	39
2000	10,754	12,466	12,200	11,538	8,666	8,188	12,338	9,136	10,567	0
2000	7,217	8,048	9,761	7,880	8,745	5,559	9,281	6,419	6,911	99
2000	5,104	6,892	12,336	4,204	2,678	3,990	8,250	5,266	7,140	236
2000	5,949	9,327	9,127	3,656	1,976	5,859	5,164	6,606	13,073	79,5