



**PENERAPAN MODEL STAR (*SPACE TIME AUTOREGRESSIVE*) DAN
ARIMA (*AUTOREGRESSIVE INTEGRATED MOVING AVERAGE*) UNTUK
PERAMALAN DATA CURAH HUJAN DI KABUPATEN JEMBER**

SKRIPSI

Oleh

Retnaningrum

NIM 091810101028

**JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS JEMBER**

2015



**PENERAPAN MODEL STAR (*SPACE TIME AUTOREGRESSIVE*) DAN
ARIMA (*AUTOREGRESSIVE INTEGRATED MOVING AVERAGE*) UNTUK
PERAMALAN DATA CURAH HUJAN DI KABUPATEN JEMBER**

SKRIPSI

diajukan guna melengkapi tugas akhir dan memenuhi salah satu syarat
untuk menyelesaikan Program Studi Matematika (S1)
dan mencapai gelar Sarjana Sains

Oleh

Retnaningrum

NIM 091810101028

**JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS JEMBER**

2015

PERSEMBAHAN

Skripsi ini saya persembahkan untuk:

1. Ibunda Suparmi dan Ayahanda Sutarno tercinta, yang dengan tulus ikhlas memberikan dukungan moral dan material serta kasih sayangnya yang tidak akan pernah pupus sampai kapanpun;
2. Ketiga adik-adikku tercinta Diana Cahya Ningrum, Yusniar Fidiya Ningrum, Dimas Wahyu Septian, yang selalu memberikan dorongan semangat serta menyayangi dan mendoakanku;
3. Dosen dan guru-guruku sejak taman kanak-kanak sampai dengan perguruan tinggi;
4. Almamater Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Jember.

MOTTO

Dan berdoalah: “Ya Tuhanku, tempatkanlah aku pada tempat yang diberkahi, dan Engkau adalah sebaik-baik yang memberi tempat.”

(Terjemahan Q.S *Al-Mu'minun*: 29)¹⁾

¹⁾ Departemen Agama Republik Indonesia. 2004. *Al Quran dan Terjemahannya*. Bandung: CV Jumanatul Ali-Art.

PERNYATAAN

Saya yang bertanda tangan dibawah ini:

Nama : Retnaningrum

NIM : 091810101028

menyatakan dengan sesungguhnya bahwa skripsi yang berjudul “Penerapan Model STAR (*Space Time Autoregressive*) dan ARIMA (*Autoregressive Integrated Moving Average*) untuk Peramalan Data Curah Hujan di Kabupaten Jember” adalah benar-benar hasil karya sendiri, kecuali kutipan yang sudah saya sebutkan sumbernya, belum pernah diajukan pada institusi manapun dan bukan karya jiplakan. Saya bertanggung jawab atas keabsahan dan kebenaran isinya sesuai dengan sikap ilmiah yang harus dijunjung tinggi.

Demikian pernyataan ini saya buat dengan sebenarnya, tanpa ada tekanan dan paksaan dari pihak manapun serta bersedia mendapat sanksi akademik jika ternyata dikemudian hari pernyataan ini tidak benar.

Jember, Desember 2015

Yang menyatakan,

Retnaningrum
NIM 091810101028

SKRIPSI

**PENERAPAN MODEL STAR (*SPACE TIME AUTOREGRESSIVE*) DAN
ARIMA (*AUTOREGRESSIVE INTEGRATED MOVING AVERAGE*) UNTUK
PERAMALAN DATA CURAH HUJAN DI KABUPATEN JEMBER**

Oleh

Retnaningrum

NIM 091810101028

Pembimbing

Dosen Pembimbing Utama : Dr. Alfian Futuhul Hadi, S.Si, M.Si.

Dosen Pembimbing Anggota : Dian Anggraeni, S.Si, M.Si.

PENGESAHAN

Skripsi yang berjudul “Penerapan Model STAR (*Space Time Autoregressive*) dan ARIMA (*Autoregressive Integrated Moving Average*) untuk Peramalan Data Curah Hujan di Kabupaten Jember” telah diuji dan disahkan pada:

hari, tanggal:

tempat : Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Jember

Tim Penguji:

Ketua,

Sekretaris,

Dr. Alfian Futuhul Hadi, S.Si, M.Si.
NIP 197407192000121001

Dian Anggraeni, S.Si, M.Si.
NIP 198202162006042002

Penguji I,

Penguji II,

Prof. Drs. I Made Tirta, M.Sc., Ph.D.
NIP 195912201985031002

Kosala Dwidja P, S.Si, M.Si.
NIP 196908281998021001

Mengesahkan

Dekan,

Prof. Drs. Kusno, DEA., Ph.D.
NIP 196101081986021001

RINGKASAN

Penerapan Model STAR (*Space Time Autoregressive*) dan ARIMA (*Autoregressive Integrated Moving Average*) untuk Peramalan Data Curah Hujan di Kabupaten Jember; Retnaningrum; 091810101028; 2015; 82 halaman; Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Jember.

Space Time Autoregressive (STAR) merupakan salah satu metode statistik dengan pendekatan *space-time* yang dapat digunakan untuk menganalisis data deret waktu (*time series*) dengan melibatkan faktor geografis (lokasi) dalam melakukan peramalan. Salah satu contoh data yang diduga mempunyai keterkaitan antar waktu dan lokasi adalah data curah hujan di kabupaten Jember, yang mana selain mempunyai keterkaitan dengan data pada waktu sebelum-sebelumnya juga mempunyai keterkaitan dengan data pada lokasi lain yang disebut dengan hubungan spasial. Dalam penelitian ini dilakukan penerapan model STAR dengan bobot lokasi seragam. Selain itu dilakukan pula pemodelan dengan model ARIMA (*Autoregressive Integrated Moving Average*) yang mana model ini berbasis data deret waktu tunggal (*univariate*). Pemodelan dengan ARIMA digunakan sebagai pembandingan apabila korelasi spasial tidak signifikan. Tujuan penelitian ini adalah untuk melakukan model peramalan yang sesuai serta mengembangkan model terbaik dalam meramalkan banyaknya curah hujan di kabupaten Jember berdasarkan pendekatan ARIMA dan mendeskripsikan pola curah hujan melalui keterkaitan pada setiap wilayah di kabupaten Jember dengan membentuk model STAR, serta membandingkan hasil peramalan model ARIMA dan STAR pada data curah hujan di kabupaten Jember.

Penelitian ini dilakukan pada data curah hujan bulanan kabupaten Jember pada periode bulan Januari 2005 sampai bulan Desember 2012 (data *training*) dan

data *testing* pada bulan Januari 2013 sampai bulan Juni 2013, yang terdiri dari empat bagian wilayah yakni Jember Barat, Jember Selatan, Jember Tengah, Jember Timur. Proses penelitian ini dilakukan dalam beberapa langkah. Langkah pertama melakukan peramalan data deret waktu ARIMA pada masing-masing wilayah (lokasi) yang melalui tahapan identifikasi model, estimasi parameter dan diagnosa model, memilih model ARIMA terbaik dan melakukan peramalan terhadap model ARIMA yang terbentuk. Langkah kedua melakukan peramalan data deret ruang-waktu dengan model STAR untuk semua lokasi dalam waktu yang bersamaan dengan langkah pemodelan *space* yang melalui tahapan yaitu identifikasi model STAR, estimasi parameter dan diagnosa model STAR, memilih model STAR terbaik dan dilakukan peramalan untuk melihat gambaran curah hujan untuk waktu mendatang. Langkah selanjutnya setelah masing-masing model dari ARIMA dan STAR didapatkan, dilakukan perbandingan ketepatan hasil ramalan berdasarkan nilai kesalahan ramalan terkecil.

Dari hasil analisis yang telah dilakukan, didapatkan model ramalan terbaik yang sesuai dengan data yaitu model ARIMA $(2,0,2)(1,0,1)^{12}$ untuk Jember Barat, ARIMA $(2,0,2)(1,0,1)^{12}$ untuk Jember Selatan, ARIMA $(1,0,0)(2,0,0)^{12}$ untuk Jember Tengah, ARIMA $(1,0,0)(1,0,1)^{12}$ untuk Jember Timur, dan model STAR (1_1) untuk keempat lokasi. Dari model ARIMA dan STAR yang terbentuk akan dipilih model terbaik yang menghasilkan kesalahan ramalan terkecil. Pemilihan model terbaik didasarkan pada nilai RMSE dari model. Berdasarkan nilai RMSE dari peramalan, didapatkan perbandingan hasil ketepatan ramalan untuk gambaran umum data curah hujan di kabupaten Jember yang diwakili dari empat lokasi menunjukkan bahwa model STAR (1_1) dengan nilai RMSE gabungan 72,69220849 menghasilkan ketepatan ramalan yang lebih baik dari pada model ARIMA yang memiliki nilai RMSE gabungan 99,23743772. Sehingga dapat dikatakan bahwa model yang paling sesuai dengan kondisi data adalah model STAR (1_1) .

PRAKATA

Segala Puji syukur penulis panjatkan kehadirat Allah SWT yang telah melimpahkan rahmat serta karunia-Nya sehingga penulis dapat menyelesaikan skripsi yang berjudul “Penerapan Model STAR (*Space Time Autoregressive*) dan ARIMA (*Autoregressive Integrated Moving Average*) untuk Peramalan Data Curah Hujan di Kabupaten Jember”. Penyusunan skripsi ini ditujukan untuk memenuhi salah satu syarat menyelesaikan pendidikan strata satu (S1) pada Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Jember.

Penulis menyadari bahwa sangatlah sulit untuk menyelesaikan skripsi ini tanpa bantuan dan bimbingan dari berbagai pihak baik secara materil maupun moril. Oleh karena itu, dengan segala hormat penulis menyampaikan banyak terimakasih kepada:

1. Bapak Dr. Alfian Futuhul Hadi, S.Si, M.Si. selaku Dosen Pembimbing Utama dan Ibu Dian Anggraeni, S.Si, M.Si. selaku Dosen Pembimbing Anggota yang telah menyediakan waktu, tenaga dan pikiran guna memberikan petunjuk dan bimbingannya dalam penulisan skripsi ini;
2. Bapak Prof. Drs. I Made Tirta, M.Sc.,Ph.D. selaku Dosen Penguji I dan Bapak Kosala Dwidja Purnomo, S.Si, M.Si. selaku Dosen Penguji II yang telah memberikan masukan dan saran yang membangun dalam penyusunan skripsi ini;
3. Seluruh staf pengajar Program Sarjana (S1) Jurusan Matematika Fakultas MIPA Universitas Jember, yang telah memberikan ilmunya selama perkuliahan;
4. Ibunda Suparmi dan Ayahanda Sutarno, yang selalu memberikan kasih sayang, semangat dan doa yang tiada pernah berhenti, serta ketiga adik-adikku tersayang yang telah memberikan perhatian dan dorongan demi terselesaikannya skripsi ini;

5. Teman-teman seperjuanganku di Jurusan Matematika angkatan 2009 atas kekeluargaan dan keceriaan yang selalu kita buat selama masa-masa indah menjadi mahasiswa di kampus tercinta;
6. Serta pihak-pihak lain yang tidak dapat disebutkan satu persatu.

Akhir kata, penulis berharap Tuhan Yang Maha Esa berkenan membalas segala kebaikan saudara-saudara semua. Semoga skripsi ini dapat membawa manfaat bagi pengembangan ilmu pengetahuan.

Jember, Desember 2015

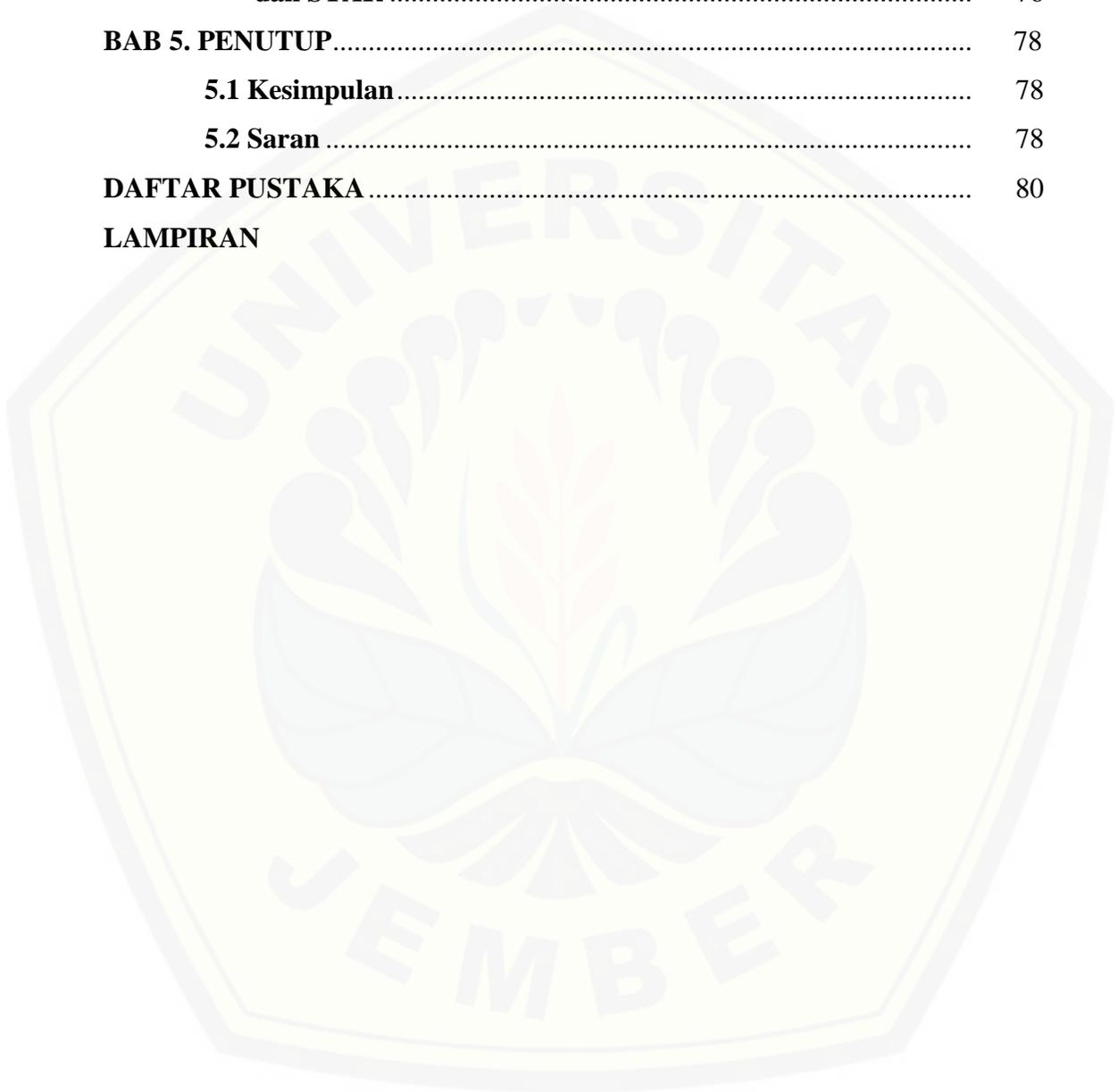
Penulis

DAFTAR ISI

	Halaman
HALAMAN JUDUL	i
HALAMAN PERSEMBAHAN	ii
HALAMAN MOTTO	iii
HALAMAN PERNYATAAN.....	iv
HALAMAN PEMBIMBINGAN.....	v
HALAMAN PENGESAHAN.....	vi
RINGKASAN	vii
PRAKATA.....	ix
DAFTAR ISI.....	xi
DAFTAR TABEL	xiv
DAFTAR GAMBAR.....	xvi
DAFTAR LAMPIRAN	xviii
BAB 1. PENDAHULUAN	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah	3
1.3 Tujuan.....	3
1.4 Manfaat.....	3
BAB 2. TINJAUAN PUSTAKA.....	4
2.1 Analisis Deret Waktu (<i>Time Series Analysis</i>).....	4
2.2 Metode Deret Waktu Box-Jenkins (ARIMA).....	6
2.2.1 Stasioneritas Model Deret Waktu <i>Univariate</i>	7
2.2.2 ACF (<i>Autocorrelation Function</i>) dan PACF (<i>Partial Autocorrelation Function</i>)	9
2.2.3 Model <i>Autoregressive-AR(p)</i>	13

2.2.4 Model <i>Moving Average</i> -MA(q)	13
2.2.5 Model Autoregressive Moving Average-ARMA(p,q).....	14
2.2.6 Model <i>Autoregressive Integrated Moving Average</i> (ARIMA)	14
2.3 Metode Deret Waktu <i>Multivariate Ruang-Waktu</i> (<i>Space Time</i>)	20
2.3.1 Stasioneritas Model Deret Waktu <i>Multivariate</i>	21
2.3.2 <i>Matrix Autocorrelation Function</i> (MACF) dan <i>Matrix Partial Autocorrelation Function</i> (MPACF).....	22
2.3.3 Model <i>Space Time Autoregressive</i> (STAR).....	24
2.3.4 Pemilihan Bobot Lokasi Pada Model STAR	24
2.3.5 Estimasi Parameter Model STAR.....	25
2.3.6 Pengujian Kecocokan Model STAR.....	27
2.3.7 Kriteria Pemilihan Model STAR Terbaik.....	28
BAB 3. METODE PENELITIAN	30
3.1 Data Penelitian	30
3.2 Metode Analisis	32
BAB 4. HASIL DAN PEMBAHASAN	39
4.1 Peramalan Data Deret Waktu Curah Hujan dengan Model ARIMA	39
4.1.1 Peramalan ARIMA Untuk Data Curah Hujan di Wilayah Jember Barat	39
4.1.2 Peramalan ARIMA Untuk Data Curah Hujan di Wilayah Jember Selatan	47
4.1.3 Peramalan ARIMA Untuk Data Curah Hujan di Wilayah Jember Tengah	53
4.1.4 Peramalan ARIMA Untuk Data Curah Hujan di Wilayah Jember Timur	59

4.2 Peramalan Data Curah Hujan dengan Model STAR.....	65
4.3 Perbandingan Ketepatan Hasil Ramalan Model ARIMA dan STAR	76
BAB 5. PENUTUP	78
5.1 Kesimpulan.....	78
5.2 Saran	78
DAFTAR PUSTAKA	80
LAMPIRAN	



DAFTAR TABEL

	Halaman
Tabel 2.1 Penentuan Nilai Ordo Pada Proses ARIMA Berdasarkan Plot ACF dan PACF	16
Tabel 4.1 Nilai Estimasi Parameter Model-Model Sementara Pemodelan Data Curah Hujan di Wilayah Jember Barat Tanpa Proses <i>Differencing</i>	42
Tabel 4.2 Ringkasan Uji Proses Ljung-Box Untuk Model ARIMA (2,0,1)(1,0,1) ¹² dan ARIMA (2,0,2)(1,0,1) ¹²	44
Tabel 4.3 Validasi Model Pemodelan Curah Hujan di Wilayah Jember Barat	45
Tabel 4.4 Perbandingan Data Prediksi Curah Hujan Per Bulan dengan Data Aktual di Wilayah Jember Barat	47
Tabel 4.5 Nilai Estimasi Parameter Model-Model Sementara Pemodelan Data Curah Hujan di Wilayah Jember Selatan Tanpa Proses <i>Differencing</i>	50
Tabel 4.6 Ringkasan Uji Proses Ljung-Box Untuk Model ARIMA (2,0,2)(1,0,1) ¹²	51
Tabel 4.7 Perbandingan Data Prediksi Curah Hujan Per Bulan dengan Data Aktual di Wilayah Jember Selatan	53
Tabel 4.8 Nilai Estimasi Parameter Model-Model Sementara Pemodelan Data Curah Hujan di Wilayah Jember Tengah Tanpa Proses <i>Differencing</i>	56
Tabel 4.9 Ringkasan Uji Proses Ljung-Box Untuk Model ARIMA (1,0,0)(2,0,0) ¹²	58

Tabel 4.10 Perbandingan Data Prediksi Curah Hujan Per Bulan dengan Data Aktual di Wilayah Jember Tengah.....	59
Tabel 4.11 Nilai Estimasi Parameter Model-Model Sementara Pemodelan Data Curah Hujan di Wilayah Jember Timur Tanpa Proses <i>Differencing</i>	62
Tabel 4.12 Ringkasan Uji Proses Ljung-Box Untuk Model ARIMA $(1,0,0)(1,0,1)^{12}$	63
Tabel 4.13 Perbandingan Data Prediksi Curah Hujan Per Bulan dengan Data Aktual di Wilayah Jember Timur.....	65
Tabel 4.14 Matriks Korelasi Data Curah Hujan Antar Wilayah	65
Tabel 4.15 Hasil Pengujian Stasioneritas dengan Menggunakan Uji ADF	68
Tabel 4.16 Representasi Skematik (a) MACF dan (b) MPACF Untuk Data Curah Hujan di Empat Bagian Wilayah kabupaten Jember	69
Tabel 4.17 Ringkasan Nilai AIC dari Model <i>Autoregressive</i>	70
Tabel 4.18 Estimasi Parameter Model STAR(1 ₁)	71
Tabel 4.19 Hasil Uji <i>Portmanteau</i> Model STAR(1 ₁).....	73
Tabel 4.20 Perbandingan Data Prediksi Curah Hujan Per Bulan dengan Data Aktual di Wilayah Kabupaten Jember	76
Tabel 4.21 Perbandingan Hasil Ketepatan Ramalan Antara Model ARIMA dan STAR.....	77

DAFTAR GAMBAR

	Halaman
Gambar 3.1 Peta Persebaran Wilayah di Kabupaten Jember (Sumber: DPU Kab. Jember).....	31
Gambar 3.2 Skema Proses Penelitian.....	36
Gambar 3.3 Tahapan ARIMA	37
Gambar 3.4 Tahapan STAR	38
Gambar 4.1 Plot Data Curah Hujan Bulanan di Wilayah Jember Barat Tahun 2005-2012.....	40
Gambar 4.2 Fungsi Autokorelasi (ACF) dan Fungsi Autokorelasi Parsial (PACF) Data Curah Hujan di Wilayah Jember Barat	41
Gambar 4.3 Plot Peramalan Curah Hujan Data Curah Hujan di Wilayah Jember Barat.....	46
Gambar 4.4 Plot Data Curah Hujan Bulanan di Wilayah Jember Selatan Tahun 2005-2012.....	48
Gambar 4.5 Fungsi Autokorelasi (ACF) dan Fungsi Autokorelasi Parsial (PACF) Data Curah Hujan di Wilayah Jember Selatan	49
Gambar 4.6 Plot Peramalan Curah Hujan Data Curah Hujan di Wilayah Jember Selatan.....	52
Gambar 4.7 Plot Data Curah Hujan Bulanan di Wilayah Jember Tengah Tahun 2005-2012.....	54
Gambar 4.8 Fungsi Autokorelasi (ACF) dan Fungsi Autokorelasi Parsial (PACF) Data Curah Hujan di Wilayah Jember Tengah	55
Gambar 4.9 Plot Peramalan Curah Hujan Data Curah Hujan di Wilayah Jember Tengah.....	58
Gambar 4.10 Plot Data Curah Hujan Bulanan di Wilayah Jember Timur	

Tahun 2005-2012.....	60
Gambar 4.11 Fungsi Autokorelasi (ACF) dan Fungsi Autokorelasi Parsial (PACF) Data Curah Hujan di Wilayah Jember Timur	61
Gambar 4.12 Plot Peramalan Curah Hujan Data Curah Hujan di Wilayah Jember Timur.....	64
Gambar 4.13 Plot Data Curah Hujan Bulanan di Empat Bagian Wilayah Kabupaten Jember Tahun 2005-2012.....	67
Gambar 4.14 Fungsi Autokorelasi (ACF) Data Curah Hujan (a) Jember Barat (b) Jember Selatan (c) Jember Tengah (d) Jember Timur.....	68
Gambar 4.15 Plot Deret Waktu Hasil Peramalan Model STAR(1 ₁) Untuk Data Curah Hujan (a) Jember Barat (b) Jember Selatan (c) Jember Tengah (d) Jember Timur.....	75

DAFTAR LAMPIRAN

	Halaman
A. Data Curah Hujan (mm) di Kabupaten Jember	82
B. Program Analisis Data Curah Hujan Wilayah Jember Barat dengan Menggunakan ARIMA	108
B1. Plot Data Curah Hujan (mm) di Wilayah Jember Barat.....	108
B2. Uji Kehomogenan Nilai Tengah Data Curah Hujan di Wilayah Jember Barat	108
B3. Skrip Hasil Estimasi Parameter Untuk Model-Model Sementara Data Curah Hujan di Wilayah Jember Barat Tanpa Proses <i>Differencing</i>	109
B4. Ljung-Box Test Sisaan Model ARIMA $(2,0,1)(1,0,1)^{12}$ dan ARIMA $(2,0,2)(1,0,1)^{12}$	110
B5. Skrip Validasi Model Untuk Pemodelan Curah Hujan di Wilayah Jember Barat.....	111
B6. Skrip Peramalan Curah Hujan di Wilayah Jember Barat	111
C. Program Analisis Data Curah Hujan Wilayah Jember Selatan dengan Menggunakan ARIMA	113
C1. Plot Data Curah Hujan (mm) di Wilayah Jember Selatan.....	113
C2. Uji Kehomogenan Nilai Tengah Data Curah Hujan di Wilayah Jember Selatan	113
C3. Skrip Hasil Estimasi Parameter Untuk Model-Model Sementara Data Curah Hujan di Wilayah Jember Selatan Tanpa Proses <i>Differencing</i> ..	113
C4. Ljung-Box Test Sisaan Model ARIMA $(2,0,2)(1,0,1)^{12}$	115
C5. Skrip Peramalan Curah Hujan di Wilayah Jember Selatan.....	115

D. Program Analisis Data Curah Hujan Wilayah Jember Tengah dengan Menggunakan ARIMA	117
D1. Plot Data Curah Hujan (mm) di Wilayah Jember Tengah	117
D2. Uji Kehomogenan Nilai Tengah Data Curah Hujan di Wilayah Jember Tengah.....	117
D3. Skrip Hasil Estimasi Parameter Untuk Model-Model Sementara Data Curah Hujan di Wilayah Jember Tengah Tanpa Proses <i>Differencing</i> ..	117
D4. Ljung-Box Test Sisaan Model ARIMA (1,0,0)(2,0,0) ¹²	118
D5. Skrip Peramalan Curah Hujan di Wilayah Jember Tengah.....	119
E. Program Analisis Data Curah Hujan Wilayah Jember Timur dengan Menggunakan ARIMA	120
E1. Plot Data Curah Hujan (mm) di Wilayah Jember Timur	120
E2. Uji Kehomogenan Nilai Tengah Data Curah Hujan di Wilayah Jember Timur.....	120
E3. Skrip Hasil Estimasi Parameter Untuk Model-Model Sementara Data Curah Hujan di Wilayah Jember Timur Tanpa Proses <i>Differencing</i> ...	120
E4. Ljung-Box Test Sisaan Model ARIMA (1,0,0)(1,0,1) ¹²	121
E5. Skrip Peramalan Curah Hujan di Wilayah Jember Timur.....	122
F. Program Analisis Data Curah Hujan di Kabupaten Jember dengan Menggunakan STAR.....	123
F1. Plot Data Curah Hujan (mm) di Kabupaten Jember	123
F2. Uji Kehomogenan Nilai Tengah Data Curah Hujan	123
F3. Skrip dan Output Model STAR.....	124
F4. Skrip dan Output Estimasi Parameter Model STAR.....	124
F5. Pengolahan Model STAR(1 ₁) Untuk Pembentukan Matriks	125
F6. Tabel Matriks Untuk Pembentukan model STAR(1 ₁) dengan Bobot Seragam.....	131

BAB 1. PENDAHULUAN

1.1 Latar Belakang

Peramalan (*forecasting*) merupakan prediksi nilai-nilai sebuah variabel kepada nilai yang diketahui dari variabel tersebut atau variabel yang berhubungan. Dalam melakukan peramalan senantiasa berupaya menyelesaikan dengan model pendekatan-pendekatan yang sesuai perilaku aktual data dan pengalaman, peramalan diperlukan untuk memberikan informasi sebagai dasar untuk membuat suatu keputusan dalam berbagai kegiatan, sehingga ramalan yang tepat akan memberikan manfaat lebih bagi yang memerlukan hasil ramalan (Montgomery: 2008). Seiring berkembangnya waktu, peramalan telah banyak dikembangkan dalam berbagai bidang untuk melakukan perencanaan dan pengambilan keputusan. Hal ini didukung oleh sebuah metode peramalan yang dikembangkan oleh George Box dan Gwilym Jenkins yaitu metode peramalan Box-Jenkins.

Metode peramalan Box-Jenkins pada dasarnya merupakan sebuah metode yang menggunakan pendekatan iteratif dalam mengidentifikasi suatu model yang paling tepat dari semua kemungkinan model yang ada. Metode peramalan Box-Jenkins lebih dikenal dengan model *Autoregressive Integrated Moving Average* (ARIMA). Model ARIMA adalah model yang secara penuh mengabaikan variabel penjelas dalam membuat peramalan. ARIMA menggunakan nilai masa lalu dan sekarang dari variabel respon untuk menghasilkan peramalan jangka pendek yang akurat (Pankratz: 1983). Oleh karena itu, dalam penelitian ini penulis menggunakan data banyaknya curah hujan di kabupaten Jember untuk dilakukan peramalan dengan ARIMA, hal ini didasarkan karena data curah hujan yang diambil melalui ijin Dinas Perairan kabupaten Jember menurut waktu pengumpulannya merupakan data deret waktu (*time series*) karena diurut dalam rentang waktu tertentu, selain itu data yang

diperoleh sifatnya dinamis karena telah memperhitungkan perubahan berdasarkan waktu secara kontinu.

Curah hujan merupakan salah satu fenomena alam yang terdapat dalam siklus hidrologi dan sangat dipengaruhi iklim, yang keberadaannya sangat penting dalam kehidupan karena dapat mencukupi kebutuhan air yang sangat dibutuhkan oleh semua makhluk hidup. Saat ini kondisi perubahan curah hujan yang tidak tetap atau berubah setiap saat, selain membawa keuntungan juga dapat mengakibatkan kerugian. Hal ini berkaitan dengan kabupaten Jember yang sebagian besar wilayahnya merupakan lahan pertanian dan perkebunan, sehingga informasi mengenai banyaknya curah hujan sangat berguna bagi para petani dalam mengantisipasi kemungkinan terjadinya peristiwa-peristiwa ekstrim yang tidak diinginkan seperti banjir dan kekeringan yang berakibat kegagalan dalam proses produksinya, dengan demikian perlu diadakannya kajian secara menyeluruh tentang peramalan curah hujan.

Peramalan mengenai curah hujan di kabupaten Jember telah dilakukan oleh beberapa peneliti diantaranya Lestari (2000) telah mengkaji pemodelan dan peramalan banyaknya hari hujan di Jember dengan proses ARMA, dan Imami (2008) melakukan aplikasi model *time series* semiparametrik ketika ARIMA gagal memenuhi *white noise* dengan studi kasus peramalan curah hujan di Jember. Penelitian-penelitian ini terbatas pada data acuan yang digunakan dalam proses peramalan yang hanya berdasar pada waktu sebelumnya dan belum memperhatikan pengaruh faktor geografis. Variabilitas curah hujan yang besar bergantung dalam ruang dan waktu. Menurut skala ruang, variabilitas curah hujan banyak dipengaruhi letak geografi, ketinggian tempat, letak lintang, topografi dan arah angin umum (Aldrian: 2003). Dengan demikian perlu adanya pengembangan metode lain untuk meningkatkan tingkat akurasi peramalan serta mengantisipasi keragaman perubahan curah hujan dalam ruang dan waktu. Metode statistika yang dapat digunakan untuk menganalisis data deret waktu dengan melibatkan faktor geografis adalah *Space Time Autoregressive* (STAR), yang mana model ini mampu menggabungkan unsur dependensi waktu dan lokasi pada suatu deret waktu *multivariate* (Pfeifer: 1980).

1.2 Rumusan Masalah

Pada penelitian ini dilakukan peramalan terhadap data curah hujan di kabupaten Jember dengan menggunakan model ARIMA dan model STAR. Data curah hujan diduga selain mempunyai keterkaitan dengan data pada waktu sebelum-sebelumnya juga mempunyai keterkaitan dengan data pada lokasi lain yang disebut dengan hubungan spasial, sehingga diperlukan model STAR dalam menggambarkan keterkaitan waktu dan lokasi. Model ARIMA digunakan sebagai pembanding apabila korelasi spasial tidak signifikan. Permasalahan yang akan dibahas dalam skripsi ini adalah bagaimana melakukan metode peramalan terhadap data curah hujan di kabupaten Jember dengan menggunakan ARIMA dan STAR, serta mendapatkan model ramalan terbaik dengan membandingkan dan mengevaluasi ketepatan ramalan berdasarkan nilai kesalahan ramalan terkecil antara kedua model tersebut.

1.3 Tujuan

Tujuan yang ingin dicapai dari penulisan skripsi ini adalah:

1. melakukan model peramalan yang sesuai serta mengembangkan model terbaik dalam meramalkan banyaknya curah hujan di kabupaten Jember berdasarkan pendekatan ARIMA dan mendeskripsikan pola curah hujan melalui keterkaitan pada setiap wilayah di kabupaten Jember dengan membentuk model STAR.
2. membandingkan hasil peramalan model ARIMA dan STAR pada data curah hujan di kabupaten Jember.

1.4 Manfaat

Manfaat yang akan didapat dari penulisan skripsi ini adalah mengetahui model peramalan dan prediksi yang sesuai untuk banyaknya curah hujan, serta mengetahui keterkaitan banyaknya curah hujan pada setiap wilayah di kabupaten Jember sehingga dapat digunakan sebagai bentuk upaya mengantisipasi berbagai dampak yang ditimbulkan.

BAB 2. TINJAUAN PUSTAKA

Pada bab ini akan dijelaskan mengenai teori dasar dan literatur yang menjadi dasar dalam penyelesaian masalah penelitian ini. Berbagai sumber yang digunakan, baik berupa buku, artikel, jurnal digunakan untuk mendukung teori penyelesaian skripsi ini. Adapun pembahasan teori mencakup analisis deret waktu, teori dan metode peramalan deret waktu Box-Jenkins yaitu ARIMA (*Autoregressive Integrated Moving Average*), serta metode deret ruang-waktu yaitu model STAR (*Space Time Autoregressive*).

2.1 Analisis Deret Waktu (*Time Series Analysis*)

Setiap urutan hasil pengukuran dari respon yang selalu berubah (variabel) menurut waktu disebut deret waktu (*time series*). *Time series* biasanya dinyatakan dengan persamaan matematika yang berisi nilai-nilai respon sebagai fungsi dari waktu atau yang ekuivalen dengan itu, sebagai angka pada sebuah grafik dengan koordinat vertikal menunjukkan nilai respon acak dan dipetakan terhadap waktu pada sumbu horisontal. Hasil pengukuran dari respon ini nantinya akan dapat diproses melalui analisa sehingga didapatkan hasil perkiraan untuk masa depan. Pada proses analisa ini sangat beragam namun intinya tetap menggunakan pola data deret waktu (*time series*) untuk memproyeksikan masa depan melalui mekanisme tertentu dan proses inilah yang disebut sebagai analisis deret waktu (*time series analysis*) (Pankratz: 1983).

Peramalan dengan *time series* pada dasarnya merupakan suatu himpunan pengamatan yang dibangun secara berurutan dalam waktu. *Time series* (deret waktu) merupakan proses stokastik $\{Y_t, t \in T\}$, dengan $T = \{\pm 1, \pm 2, \pm 3, \dots\}$ yang merupakan indeks parameter waktu, dimana waktu bisa saja merupakan parameter

yang kontinu atau pun parameter diskret. Biasanya waktu yang digunakan merupakan indeks parameter diskret dengan unit dari waktu yang bisa saja meliputi tahun, bulan, hari, atau detik, bergantung pada situasi yang dimodelkan. Waktu atau periode yang dibutuhkan untuk melakukan suatu peramalan disebut sebagai *lead time*, yang bervariasi dalam setiap persoalan. Dalam penggunaan waktu t pada pengamatan deret waktu untuk meramalkan nilai di masa depan telah menjadi dasar bagi perencanaan ekonomi, bisnis, produksi serta optimalisasi proses industri. Dengan nilai Y_t maka data deret waktu sebelumnya (Y_{t-1}, Y_{t-2}, \dots) dapat digunakan untuk meramalkan nilai pada beberapa periode ke depan beserta *lead time* (k) yang menyatakan periode peramalan di masa mendatang. Fungsi dari $Y_t(k)$ merupakan peramalan yang dibuat dari awal pengamatan t dari objek, Y_{t+1} yang terjadi pada *lead time* (k) sehingga mendapatkan nilai *mean square deviations* $Y_{t+1} - Y_t(k)$ sekecil mungkin diantara nilai aktual dengan peramalan untuk setiap *lead time* ($k = 1, 2, \dots$) (Shumway: 2011).

2.1.1 Pola Data *Time Series*

Salah satu hal yang terpenting dalam metode peramalan dengan data deret waktu (*time series*) yang dimilikinya adalah memperhatikan atau mengidentifikasi jenis pola data. Identifikasi pola data dilakukan untuk mengetahui unsur pola yang terkandung pada suatu deret data, sehingga deret data tersebut dapat disesuaikan dengan metode peramalan *time series* yang digunakan. Aspek penting dari pemilihan teknik peramalan yang sesuai dari data *time series* adalah dengan memperhatikan jenis pola data yang berbeda. Ada empat jenis yang umum, yaitu: Tren, Musiman, siklis, dan acak.

1. Pola Tren (*Trend*)

Pola *trend* terbentuk ketika data observasi terlihat meningkat atau menurun dalam periode waktu yang lebih panjang. *Trend* merupakan komponen jangka panjang yang mendasari pertumbuhan atau penurunan data *time series*. *Trend* dapat

disebabkan oleh misalnya pertumbuhan populasi, inflasi, perubahan teknologi, dan peningkatan produktivitas.

2. Pola Musiman (*Seasonality*)

Pola ini terjadi ketika data observasi dipengaruhi oleh faktor musiman. Komponen musiman mengacu pada suatu pola perubahan yang berulang dengan sendirinya dari tahun ke tahun. Fluktuasi musiman ini umumnya terjadi setiap mingguan, bulanan, atau triwulanan. Variasi musiman mencerminkan kondisi cuaca, liburan atau panjangnya hari bulan kalender.

3. Pola Siklis (*Cycle*)

Pola ini hampir memiliki karakteristik yang sama dengan pola musiman, terjadi ketika data observasi terlihat naik turun dalam periode waktu yang lebih panjang.

4. Pola Acak (*Random*)

Pola ini terjadi karena data observasi tidak dipengaruhi faktor-faktor khusus sehingga pola menjadi tidak menentu dan tidak dapat diperkirakan secara biasa (Cryer: 2008).

Langkah yang harus dilakukan untuk menganalisis data deret waktu adalah dengan memplotkan data tersebut secara grafis. Dari hasil plot data tersebut dapat diketahui apakah pola data trend, musiman, siklik atau acak. Dengan mengetahui secara jelas pola data dari suatu data deret waktu maka dapat dipilih teknik-teknik peramalan yang mampu secara efektif mengekstrapolasi pola data.

2.2 Metode Deret Waktu *Univariate* Box-Jenkins (ARIMA)

Penerapan metode Box-Jenkins digunakan pada kajian terhadap data deret waktu *univariat* (tunggal). Pada metode Box-Jenkins dikenal model deret waktu yang secara umum terdiri dari model regresi diri (*Autoregressive*), model rata-rata bergerak (*Moving Average*), dan model gabungan keduanya (*Autoregressive-Moving Average*) serta model *Autoregressive Integrated Moving Average* (ARIMA). Model ARIMA telah dipelajari secara mendalam oleh George Box dan Gwilym Jenkins (1976) dan nama mereka sering disinonimkan dengan model ARIMA yang diterapkan untuk

analisis dan peramalan data deret waktu (Pankratz: 1983). Oleh karena itu diperlukan pemahaman atas konsep kestasioneran data, autokovarians, autokorelasi, serta autokorelasi parsial dalam menentukan dan mengidentifikasi model-model dengan metode Box-Jenkins yang merupakan model dengan menggambarkan *time series* yang stasioner.

2.2.1 Stasioneritas Model Deret Waktu *Univariate*

Data *time series* pada umumnya tidak stasioner, oleh karena itu harus dilakukan uji stasioneritas data terlebih dahulu. Uji stasioneritas data merupakan langkah awal yang dilakukan untuk mengestimasi sebuah model yang akan digunakan. Suatu proses yang penting dalam model-model stokastik (suatu famili dari peubah acak $\{Y_t, t \in T\}$, $T = \{\pm 1, \pm 2, \pm 3, \dots\}$ yang didefinisikan pada suatu ruang probabilistik) untuk menggambarkan suatu *time series* adalah apa yang disebut sebagai model – model stasioner yang mengasumsikan bahwa proses tetap berada dalam keseimbangan disekitar konstan mean level. Menurut Shumway (2011) deret waktu (*time series*) dikatakan stasioner kuat jika fungsi sebaran bersama dari Y_{t_1}, \dots, Y_{t_n} sama dengan fungsi sebaran bersama dari $Y_{t_1+k}, \dots, Y_{t_n+k}$. Dengan demikian $P\{Y_{t_n} \leq c_n\} = P\{Y_{t_n+k} \leq c_n\}$ untuk $n = 1, 2, \dots, n$ dan t_n yang menyatakan indeks waktu serta lag antar waktu yang dapat bernilai $k = 0, \pm 1, \pm 2, \dots$. Di sisi lain deret waktu dikatakan stasioner lemah, jika nilai tengah deret waktu konstan sepanjang waktu ($\mu_t = \mu$) dan nilai autokovarians konstan ($\gamma_{(t+k,t)} = \gamma_{(k,0)}$) untuk semua waktu t dan lag $-k$. Pada kenyataannya definisi kestasioneran yang digunakan adalah stasioner lemah. Dengan demikian, data deret waktu dikatakan stasioner jika perilaku data tersebut berfluktuasi di sekitar nilai tengah dan variansi yang relatif konstan sepanjang periode waktu.

Pada pemeriksaan kestasioneran deret waktu dapat dilakukan dengan pendekatan informal dan formal. Pendekatan informal dapat dilakukan dengan

eksplorasi plot deret waktu yang konsep stasioneritas ini dapat digambarkan secara praktis yaitu sebagai berikut:

- a. Apabila suatu deret waktu diplot kemudian tidak terbukti adanya perubahan nilai tengah dari waktu ke waktu maka dapat dikatakan bahwa data deret waktu tersebut stasioner pada nilai tengahnya.
- b. Apabila plot data deret waktu tidak memperlihatkan adanya perubahan variansi yang jelas dari waktu ke waktu maka dapat dikatakan data deret waktu tersebut adalah stasioner pada variansinya.
- c. Apabila plot data deret waktu memperlihatkan nilai tengahnya menyimpang (dengan beberapa plot *trend cycle*) dari waktu ke waktu maka dapat dikatakan data deret waktu tersebut tidak stasioner pada nilai tengahnya.
- d. Apabila data deret waktu memperlihatkan nilai tengahnya menyimpang (berubah setiap waktu) dan variansi tidak konstan setiap waktu maka dapat dikatakan deret waktu tersebut tidak stasioner pada nilai tengahnya dan variansinya.

Sedangkan untuk pendekatan formal dapat dilakukan dengan menggunakan uji akar unit. Salah satu uji akar unit untuk kestasioneran dalam nilai tengah adalah uji *Augmented Dickey-Fuller* (ADF) (Cryer: 2008). Uji ADF pertama kali diperkenalkan oleh David Dickey dan Wayne Fuller. Model sederhana yang digunakan pada uji ADF adalah $\Delta Y_t = \beta_1 + \delta Y_{t-1} + e_t$ dengan $\delta = \rho - 1$. Hipotesis yang diuji dari persamaan ini adalah

$H_0: \delta = 1$ (Y_t tidak stasioner)

$H_1: \delta < 1$ (Y_t stasioner)

Uji signifikansi untuk hipotesis tersebut menggunakan uji τ (tau), karena α berdistribusi τ . Statistik ujinya adalah $\tau = \frac{\hat{\rho}}{SE(\hat{\rho})}$ (Harris: 2003). Dickey dan Fuller sudah menyusun tabel untuk uji ini. Kaidah keputusan yang digunakan ialah tolak H_0 jika τ lebih kecil dari nilai τ pada ADF dengan taraf nyata tertentu atau p -value $< 5\%$. Ketika dalam pemeriksaan kestasioneran deret waktu, diketahui bahwa data tidak stasioner dalam nilai tengah, maka diperlukan proses pembedaan (*differencing*).

Menurut Pankratz (1983) pembedaan adalah operasi sederhana yang melibatkan perubahan sukuensial dalam menghitung nilai suatu deret waktu. Pembedaan dilakukan ketika data tidak stasioner dalam nilai tengah (nilai tengah berubah setiap waktu). Persamaan proses pembedaan adalah $w_t = Y_t - Y_{t-1}$ dimana $t = 2, 3, 4, \dots, n$ dengan w_t merupakan variabel yang menyatakan selisih antara pengamatan dalam deret waktu. Adapun transformasi *Box-Cox* dilakukan jika kestasioneran dalam variansi tidak dapat diselesaikan dengan pembedaan. Transformasi ini didefinisikan sebagai berikut (Cryer: 2008) :

$$T(Y_t) = \begin{cases} Y_t^\lambda - 1 & ; \text{ untuk } \lambda \neq 0 \\ \log(Y_t) & ; \text{ untuk } \lambda = 0 \end{cases}$$

Sehingga dapat dikatakan bahwa diperlukan pemenuhan asumsi tentang kestasioneritas data pada pemodelan deret waktu yang berarti bahwa tidak terdapat pertumbuhan atau penurunan data. Bila tidak terdapat perubahan pada *trend* waktu maka dapat disebut stasioner dimana rata-rata deret pengamatan di sepanjang waktu selalu konstan. Apabila suatu data tidak stasioner maka diperlukan diferensiasi pada data tersebut. Bila data masih belum stasioner maka perlu dideferensiasi lagi hingga stasioner dan jika suatu deret data bukan merupakan data yang stasioner, maka sebelum melakukan pembuatan model deret waktu, maka perlu dilakukan pembedaan atau transformasi (Cowpertwait: 2008).

2.2.2 ACF (*Autocorrelation Function*) dan PACF (*Partial Autocorrelation Function*)

Untuk melihat adanya ketergantungan antara pengamatan, maka diperlukan proses uji korelasi antar pengamatan. Untuk proses Y_t yang stasioner mempunyai rata-rata (mean) $E(Y_t) = \mu_t = \mu$ dan variansi $var(Y_t) = E(Y_t - \mu_t)^2 = \sigma^2$ adalah konstan dan kovariansi $cov(Y_t, Y_s) = E(Y_t - \mu_t)(Y_s - \mu_s) = \gamma_{t,s}$ adalah fungsi dari selisih waktu $|t - s|$. Estimator untuk mean adalah rata-rata sampel yaitu

$$\bar{Y} = \frac{1}{n} \sum_{t=1}^n Y_t \quad (2.1)$$

dimana n merupakan banyaknya pengamatan deret waktu. Kemudian untuk kovariansi antara Y_t dan Y_{t+k} yang dipisahkan oleh k interval waktu disebut sebagai autokovarians pada lag k dan didefinisikan sebagai berikut:

$$\gamma_k = \text{cov}(Y_t, Y_{t+k}) = E(Y_t - \mu, Y_{t+k} - \mu) \quad (2.2)$$

sehingga penaksirnya adalah kovariansi sampel antara Y_t dan Y_{t+k} diberikan oleh

$$\hat{\gamma}_k = \frac{1}{n} \sum_{t=1}^{n-k} (Y_t - \bar{Y})(Y_{t+k} - \bar{Y}) \text{ atau } \hat{\gamma}_k = \frac{1}{n} \sum_{t=k+1}^n (Y_t - \bar{Y})(Y_{t-k} - \bar{Y}) \quad (2.3)$$

karena $\gamma_k = \gamma_{-k}$ secara umum autokorelasi pada lag k didefinisikan sebagai berikut:

$$\begin{aligned} \rho_k = \text{korr}(Y_t, Y_{t+k}) &= \frac{\text{cov}(Y_t, Y_{t+k})}{\sqrt{\text{var}(Y_t)}\sqrt{\text{var}(Y_{t+k})}} \\ &= \frac{E(Y_t - \mu, Y_{t+k} - \mu)}{\sqrt{E[(Y_t - \mu)^2]}E[(Y_{t+k} - \mu)^2]} \\ &= \frac{E(Y_t - \mu, Y_{t+k} - \mu)}{\sigma^2} = \frac{\gamma_k}{\gamma_0} \end{aligned} \quad (2.4)$$

berdasarkan persamaan (2.2) dan karena proses stasioner, maka variansi $\text{var}(Y_t) = \text{var}(Y_{t+k}) = \sigma^2 = \gamma_0$ pada waktu t adalah sama untuk waktu $t + k$. Berdasarkan persamaan (2.3) maka penaksir untuk γ_0 adalah

$$\hat{\gamma}_0 = \frac{1}{n} \sum_{t=1}^n (Y_t - \bar{Y})^2 \quad (2.5)$$

maka autokorelasi pada lag k adalah $\rho_k = \frac{\gamma_k}{\gamma_0}$ dimana γ_k dinamakan fungsi autokovarians dan ρ_k dinamakan fungsi autokorelasi pada analisis deret waktu, karena masing-masing menyatakan kovariansi dan korelasi antara Y_t dan Y_{t-k} dari proses yang sama, hanya dipisahkan oleh jarak waktu k atau lag k . Karena $\text{cov}(Y_t, Y_{t+k}) = \text{cov}(Y_{t+k}, Y_t) = \text{cov}(Y_t, Y_{t-k})$ maka yang perlu ditentukan adalah γ_k untuk $k \geq 0$. Himpunan $\{\rho_k; k = 0, 1, 2, \dots\}$ dinamakan fungsi autokorelasi (*Autocorrelation Function*) disingkat dengan ACF. Koefisien korelasi (fungsi autokorelasi) merupakan statistik kunci dalam analisis deret waktu, yaitu menyatakan ukuran korelasi (hubungan linier) deret waktu itu dengan dirinya sendiri dengan selisih waktu (lag) 0,1,2 periode atau lebih. Untuk suatu pengamatan deret waktu

Y_1, Y_2, \dots, Y_n , maka nilai autokorelasi antara Y_t dan Y_{t+k} dinamakan nilai autokorelasi lag k sampel atau penaksir (estimator) yang diberikan oleh

$$\hat{\rho}_k = \frac{\hat{\gamma}_k}{\hat{\gamma}_0} = \frac{\frac{1}{n} \sum_{t=1}^{n-k} (Y_t - \bar{Y})(Y_{t+k} - \bar{Y})}{\frac{1}{n} \sum_{t=1}^n (Y_t - \bar{Y})^2} \quad (2.6)$$

Perilaku fungsi korelasi diri (ACF/Autocorrelation Function) yaitu fungsi yang menunjukkan besarnya korelasi antara pengamatan pada waktu ke- t dengan pengamatan pada waktu sebelum-sebelumnya $t - 1$ juga dapat digunakan sebagai dasar penentu kestasioneran deret waktu. Deret waktu yang stasioner dapat dilihat dari pola fungsi korelasi diri yang menunjukkan penurunan nilai-nilai korelasi diri yang cepat mendekati nol. Sebaliknya, deret waktu yang tidak stasioner terjadi jika penurunan nilai-nilai korelasi diri yang lambat.

Autokorelasi parsial merupakan ukuran yang lain pada analisa deret waktu. Autokorelasi parsial digunakan untuk mengukur tingkat korelasi (keeratan) antara Y_t dan Y_{t-k} , apabila pengaruh dari lag waktu $1, 2, \dots, k - 1$ dianggap terpisah. Fungsi autokorelasi parsial adalah suatu fungsi yang menunjukkan besarnya korelasi antara pengamatan ke t yaitu Y_t dengan pengamatan waktu-waktu sebelumnya $Y_{t-1}, Y_{t-2}, \dots, Y_{t-k}$. Rumus autokorelasi parsial sebagai berikut:

$$\phi_{kk} = \text{korr}(Y_t, Y_{t-k} | Y_{t-1}, Y_{t-2}, \dots, Y_{t-k}) \quad (2.7)$$

Nilai ϕ_{kk} dapat ditentukan melalui persamaan Yule-Walker

$$\rho_j = \phi_{k1}\rho_{j-1} + \phi_{k2}\rho_{j-2} + \dots + \phi_{k1}\rho_{j-k} \quad (2.8)$$

Untuk $j = 1, 2, 3, \dots, k$ didapat persamaan linier

$$\begin{bmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{k-1} \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{k-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_{k-1} & \rho_{k-1} & \rho_{k-1} & \dots & 1 \end{bmatrix} \begin{bmatrix} \phi_{k1} \\ \phi_{k2} \\ \vdots \\ \phi_{kk} \end{bmatrix} = \begin{bmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_k \end{bmatrix}$$

Sehingga untuk $k = 1, 2, 3, \dots$ didapat nilai autokorelasi parsial lag $1, 2, 3, \dots$ adalah

$$\phi_{11} = \rho_1 ;$$

$$\phi_{22} = \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}} = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2} ;$$

$$\phi_{33} = \frac{\begin{vmatrix} 1 & \rho_1 & \rho_1 \\ \rho_1 & 1 & \rho_2 \\ \rho_2 & \rho_1 & \rho_3 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_2 \\ \rho_2 & \rho_1 & 1 \end{vmatrix}} = \frac{\rho_1^3 - 2\rho_1\rho_2 + \rho_1\rho_2^2 - \rho_1^2\rho_3 + \rho_3}{1 - \rho_1^2 + 2\rho_1^2\rho_2 - \rho_2^2};$$

⋮

$$\phi_{kk} = \frac{\begin{vmatrix} 1 & \rho_1 & \rho_1 & \dots & \rho_{k-2} & \rho_1 \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{k-3} & \rho_1 \\ \vdots & \vdots & \vdots & \dots & \vdots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \dots & \rho_1 & \rho_k \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_1 & \dots & \rho_{k-2} & \rho_{k-1} \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{k-3} & \rho_{k-2} \\ \vdots & \vdots & \vdots & \dots & \vdots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \dots & \rho_1 & 1 \end{vmatrix}} \tag{2.9}$$

atau dengan menggunakan metode yang lebih efisien untuk menyelesaikan persamaan

Yule-Walker (2.8) yaitu rumus Durbin $\phi_{kk} = \frac{\rho_k - \sum_{j=1}^{k-1} \phi_{k-1, j} \rho_{k-j}}{1 - \sum_{j=1}^{k-1} \phi_{k-1, j} \rho_j}$ (2.10)

dimana $\phi_{kj} = \phi_{k-1, j} - \phi_{kk} \phi_{k-1, k-j}; j = 1, 2, \dots, k - 1$. Estimator untuk ϕ_{kk} adalah $\hat{\phi}_{kk}$ yang diperoleh dengan mengganti ρ_k pada persamaan (2.9) dengan penaksirnya yaitu $\hat{\rho}_k$, dengan demikian fungsi dari k , ϕ_{kk} dinamakan fungsi autokorelasi parsial (*Partial Autocorrelation Function*) disingkat dengan PACF (Montgomery: 2008).

Untuk proses $\{a_t\}$ yang dinamakan *white noise* apabila bentuk peubah acak yang berurutan tidak saling berkorelasi dan mengikuti distribusi tertentu yang identik. Proses *white noise* mempunyai fungsi mean konstan $E(a_t) = \mu_t$ untuk $t = 1, 2, \dots$ dan $var(a_t) = \sigma_a^2$, serta kovariansi untuk proses ini adalah $\gamma_k = cov(a_t, a_{t+k}) = 0$ untuk $k \neq 0$. Berdasarkan definisi ini, dapat diketahui bahwa proses *white noise* $\{a_t\}$ adalah stasioner dan mempunyai sifat:

fungsi autokovariansinya konstan: $\gamma_k = \begin{cases} var(a_t) = \sigma_a^2 & \text{untuk } k = 0 \\ cov(a_t, a_{t+k}) = 0 & \text{untuk } k \neq 0 \end{cases}$

sehingga fungsi autokorelasinya yaitu $\rho_k = \begin{cases} 1 & \text{untuk } k = 0 \\ 0 & \text{untuk } k \neq 0 \end{cases}$

dan $\phi_{kk} = \begin{cases} 1 & \text{untuk } k = 0 \\ 0 & \text{untuk } k \neq 0 \end{cases}$. Pada umumnya proses *white noise* diasumsikan berdistribusi normal dengan mean nol dan variansi konstan σ_a^2 sehingga dapat ditulis $a_t \sim NID(0, \sigma_a^2)$ (Cowpertwait: 2008).

2.2.3 Model *Autoregressive*-AR(p)

Secara umum model *Autoregressive* atau AR(p) adalah bentuk regresi tetapi bukan yang menghubungkan variabel respon dengan variabel penjelas, melainkan menghubungkan nilai-nilai sebelumnya diri sendiri pada time lag (selang waktu). yang bermacam-macam. Menurut Montgomery (2008) model regresi diri berordo p atau AR(p) mempunyai bentuk sebagai berikut:

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + a_t \quad (2.11)$$

dimana:

Y_t = deret waktu stasioner

ϕ_p = parameter *autoregressive* ke- p

$Y_{t-1}, Y_{t-2}, \dots, Y_{t-p}$ = nilai masa lalu pada saat $t - p$

a_t = nilai kesalahan pada saat t .

Apabila digunakan bentuk operator shift mundur untuk *autoregressive* order p adalah sebagai berikut

$$\Phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$$

dimana $B := BY_t = Y_{t-1}$ dan $B^p Y_t = Y_{t-p}$

sehingga persamaan (2.11) dapat dituliskan menjadi

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) y_t = \delta + a_t$$

$$\text{atau } \Phi(B) y_t = \delta + a_t \quad (2.12)$$

2.2.4 Model *Moving Average*-MA(q)

Model rata-rata bergerak atau *Moving Average* merupakan salah satu metode analisis teknikal sederhana. Dilakukan dengan cara mencari rata-rata bergerak dari

suatu variabel selama beberapa periode yang dipengaruhi oleh unsur kesalahan pada saat ini dan masa lalu. Bentuk umum model *Moving Average* dengan orde q atau MA(q) dinyatakan sebagai berikut:

$$Y_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q} \quad (2.13)$$

dimana:

Y_t = deret waktu stasioner

θ_q = merupakan parameter *Moving Average*

a_{t-q} = unit residual (*white noise*) atau nilai galat waktu $t - q$.

Jika (2.13) ditulis ulang dalam bentuk operator shift mundur, maka persamaan ini dapat ditulis:

$$\begin{aligned} Y_t &= a_t (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) \\ &= a_t (1 - \sum_{i=1}^q \theta_i B^i) \end{aligned} \quad (2.14)$$

atau $Y_t = \Theta(B)a_t$ dimana $\Theta(B) = (1 - \sum_{i=1}^q \theta_i B^i)$ (Montgomery: 2008).

2.2.5 Model *Autoregressive Moving Average*-ARMA(p,q)

Selain model regresi diri atau AR(p) dan rata-rata bergerak atau MA(q), model deret waktu juga dapat dijelaskan oleh model regresi diri dan rata-rata bergerak secara bersamaan. Model yang dapat memuat kedua proses tersebut adalah model ARMA (*Autoregressive Moving Average*) didefinisikan sebagai berikut:

$$\begin{aligned} Y_t &= \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q} \\ &= \sum_{i=1}^p \phi_i Y_{t-i} + a_t - \sum_{i=1}^q \theta_i a_{t-i} \end{aligned} \quad (2.15)$$

atau $\Phi(B)Y_t = \Theta(B)a_t$ (Montgomery: 2008).

2.2.6 Model *Autoregressive Integrated Moving Average* (ARIMA)

Model ARIMA merupakan gabungan antara model regresi diri (*Autoregressive*) dan model rata-rata bergerak (*Moving Average*) dengan data yang

telah mengalami proses *differencing* (pembedaan) sebanyak d kali. Menurut Montgomery (2008) secara umum model ARIMA(p,d,q) dinyatakan dalam bentuk persamaan berikut:

$$\Phi(B)(1 - B)^d Y_t = \Theta(B)a_t \quad (2.16)$$

dimana:

Y_t = data deret waktu yang mengalami *differencing*

$$\Phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$$

$$\Theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q$$

$(1 - B)^d$ = *differencing* non-musiman pada orde ke- d

B = *backshift operator*; dan a_t = nilai galat waktu t .

Generalisasi dari model ARIMA untuk data yang memiliki pola musiman, yaitu model ARIMA musiman. Dimana model ARIMA musiman ini merupakan model ARIMA yang digunakan untuk menyelesaikan *time series* musiman yang terdiri dari dua bagian yaitu bagian non musiman dan bagian musiman. Bagian non musiman dari model ini adalah model ARIMA. Dalam model ini pola musiman dalam satu periode diasumsikan berulang sendiri pada periode berikutnya, hanya dengan gangguan acak tambahan. Menurut Montgomery (2008) secara umum bentuk model ARIMA musiman atau ARIMA(p,d,q) (P,D,Q)^S adalah sebagai berikut:

$$\Phi_p(B^S)\Phi_p(B)(1 - B)^d(1 - B^S)^D Y_t = \Theta_Q(B^S)\Theta_q(B)a_t \quad (2.17)$$

dengan

p, d, q = order AR, MA dan *differencing* non musiman

P, D, Q = order AR, MA dan *differencing* musiman

S = jumlah periode setiap musim

$$\Phi_p(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$$

$$\Phi_p(B^S) = 1 - \phi_1 B^S - \phi_2 B^{2S} - \dots - \phi_p B^{pS}$$

$(1 - B)^d$ = *differencing* non-musiman

$(1 - B^S)^D$ = *differencing* musiman

$$\Theta_q(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q$$

$$\Theta_Q(B^S) = 1 - \theta_1 B^S - \theta_2 B^{2S} - \dots - \theta_Q B^{QS}$$

B = *backshift operator*; dan a_t = nilai galat waktu t .

Salah satu metode yang bisa digunakan untuk menduga model ARIMA adalah metode Box-Jenkins. Selain itu metode ini dapat digunakan hanya pada data deret waktu yang stationer (Pankratz: 1983). Metode ini terdiri dari tiga langkah yaitu identifikasi model, pendugaan parameter dan diagnostik model (Pankratz: 1983).

Identifikasi model merupakan tahap pertama untuk menentukan model-model sementara terhadap deret waktu yang telah stasioner, yaitu dengan menentukan nilai p , q dan d pada ARIMA non musiman sedangkan nilai P , Q dan D pada ARIMA musiman. Penentuan nilai-nilai tersebut dilakukan dengan mengamati grafik fungsi ACF (korelogram) dan PACF (korelogram parsial). Nilai p (ordo proses AR) dapat ditentukan dengan melihat nilai pada grafik fungsi PACF dan nilai q (ordo proses MA) dapat ditentukan dengan melihat nilai pada grafik fungsi ACF. Secara umum bentuk ACF dan PACF model ARIMA musiman stasioner menyerupai ACF dan PACF pada ARIMA non musiman. Hanya nilai yang keluar terlihat seperti berkelipatan atau berulang di setiap periodenya. Adapun cara identifikasi nilai ordo tersebut secara lengkap dapat dilihat pada tabel 2.1 berikut.

Tabel 2.1 Penentuan Nilai Ordo Pada Proses ARIMA Berdasarkan Plot ACF dan PACF

No	Kemungkinan plot ACF dan PACF	Model ARIMA
1	ACF nyata pada lag ke-1,2,3,...,q dan terpotong setelah lag q (<i>cuts off</i>), PACF menurun cepat membentuk pola eksponensial atau sinus (<i>tails off</i>)	MA(q)
2	ACF <i>tails off</i> , PACF nyata pada lag ke-1,2,...,p dan <i>cuts off</i> setelah lag ke-p	AR(p)
3	ACF nyata pada lag ke-1,2,...,q lalu <i>cuts off</i> , PACF nyata pada lag ke-1,2,3,...,p dan <i>cuts off</i>	MA(q) jika ACF <i>cuts off</i> lebih tajam, AR(p) jika PACF <i>cuts off</i> lebih tajam

No	Kemungkinan plot ACF dan PACF	Model ARIMA
4	Tidak ada autokorelasi yang nyata pada plot ACF dan PACF	ARMA(0,0)
5	ACF <i>tail off</i> , PACF <i>tail off</i>	ARMA(p,q)
Pola ACF dan PACF Musiman dengan S Periode Per Musiman		
6	ACF menurun secara eksponensial (<i>tails off</i>) pada lag musiman, PACF <i>cuts off</i> setelah lag PS	AR(P)
7	ACF <i>cuts off</i> setelah lag QS, PACF <i>tails off</i> pada lag musiman	MA(Q)
8	ACF <i>tails off</i> pada lag musiman, PACF <i>tails off</i> pada lag musiman	ARMA(P,Q)

Tahap kedua adalah pendugaan parameter. Pendugaan parameter bertujuan untuk menentukan apakah parameter sudah layak digunakan dalam model. Pendugaan parameter dapat dilakukan dengan menggunakan metode *Maximum Likelihood Estimation* (MLE). Asumsi yang diperlukan pada metode MLE adalah error a_t (nilai galat waktu t) berdistribusi normal (Box: 2008). Dengan demikian, fungsi kepadatan peluangnya adalah

$$f(a_t | \sigma_a^2) = (2\pi\sigma_a^2)^{-\frac{1}{2}} \exp\left(-\frac{a_t^2}{2\sigma_a^2}\right) \quad (2.18)$$

dengan mengingat bahwa error bersifat independen, maka distribusi bersama dari $a_1, a_2, a_3, \dots, a_n$ adalah

$$f(a_1, a_2, a_3, \dots, a_n | \sigma_a^2) = (2\pi\sigma_a^2)^{-\frac{n}{2}} \exp\left(-\frac{1}{2\sigma_a^2} \sum_{i=1}^n a_i^2\right) \quad (2.19)$$

error a_t dapat dinyatakan dalam fungsi Y_t dan parameter-parameter ϕ, θ, σ_a^2 serta error-error sebelumnya. Secara umum bentuk a_t adalah

$$a_t = Y_t - \phi_1 Y_{t-1} - \phi_p Y_{t-p} + \theta_1 a_{t-1} + \dots + \theta_q a_{t-q} \quad (2.20)$$

fungsi *likelihood* untuk parameter model jika diketahui data pengamatan adalah

$$L(\phi, \theta, \sigma_a^2 | Y) = (2\pi\sigma_a^2)^{-\frac{n}{2}} \exp\left(-\frac{1}{2\sigma_a^2} S(\phi, \theta)\right) \quad (2.21)$$

dengan $S(\phi, \theta) = \sum_{t=1}^n -[(Y_t - \phi_1 Y_{t-1} - \phi_p Y_{t-p}) + (\theta_1 a_{t-1} + \dots + \theta_q a_{t-q})]^2$

maka dari persamaan 2.21 diperoleh fungsi *log likelihood*nya adalah

$$l = \ln L(\phi, \theta, \sigma_a^2 | Y) = -\frac{n}{2} \log 2\pi - \frac{n}{2} \log \sigma_a^2 - \frac{1}{2\sigma_a^2} S(\phi, \theta) \quad (2.22)$$

maksimum fungsi *log likelihood* dihitung dengan mencari turunan pertama persamaan (2.22) terhadap masing-masing parameter dan menetapkan turunan parsial yang dihasilkan sama dengan nol.

$$\frac{\partial(\phi, \theta, \sigma_a^2 | Y)}{\partial \phi} = 0; \quad \frac{\partial(\phi, \theta, \sigma_a^2 | Y)}{\partial \theta} = 0; \quad \frac{\partial(\phi, \theta, \sigma_a^2 | Y)}{\partial \sigma_a^2} = 0$$

untuk mendapatkan standart baku parameter diestimasi dengan metode MLE, digunakan matriks informasi yang dinotasikan dengan $I(\phi, \theta)$. Matriks ini diperoleh dengan menghitung dahulu derivatif kedua terhadap setiap parameternya, yang dinotasikan dengan l_{ij} .

$$l_{ij} = \frac{\partial^2 l(\beta, \sigma_a^2 | Y)}{\partial \beta_i \partial \beta_j} \text{ dan } I(\beta) = -E(l_{ij}) \quad (2.23)$$

untuk varians parameter dinotasikan dengan $V(\hat{\beta})$ dan standart baku parameternya adalah $SE(\hat{\beta})$.

$$V(\hat{\beta}) = [I(\beta)]^{-1} \text{ dan } SE(\hat{\beta}) = [V(\hat{\beta})]^{1/2} \quad (2.24)$$

Parameter yang telah diperoleh selanjutnya diuji signifikansinya. Adapun perumusan hipotesis yang dilakukan adalah sebagai berikut:

$H_0: \hat{\beta} = 0$ (parameter tidak signifikan terhadap model)

$H_1: \hat{\beta} \neq 0$ (parameter signifikan terhadap model)

Statistik uji yang digunakan adalah

$$|t - \text{hitung}| = \frac{\hat{\beta}}{SE(\hat{\beta})}$$

dengan $\hat{\beta}$ adalah parameter dugaan dan $SE(\hat{\beta})$ adalah standar error dari setiap parameter dugaan.

Pendugaan parameter untuk suatu model dikatakan berpengaruh signifikan atau H_0 ditolak jika nilai $|t - \text{hitung}| > t - \text{tabel} (t_{(\alpha/2); df=n-1})$ atau $p\text{-value} < \alpha$. Untuk

α adalah taraf nyata (*level of significance*) yang dalam bernilai 0.05 (5%). *Freedom of degree (df)* adalah tingkat kepercayaan yang didapat dari operasi pengurangan antara jumlah data dengan jumlah perkiraan parameter, dan n jumlah pengamatan.

Tahap selanjutnya yang dilakukan adalah diagnosa model untuk melihat model yang relevan dengan data. Pada tahap ini model harus dicek kelayakannya dengan melihat sifat sisaan dari sisi kebebasannya (*white noise*) dan kenormalan. Secara umum pengecekan kebebasan sisaan model dapat dilakukan dengan menggunakan uji Q_* modifikasi Box-Pierce (Ljung-Box). Berikut persamaan uji Q_* (Cryer: 2008) adalah

$$Q_* = n(n+2) \left(\sum_{k=1}^k \frac{r_k^2}{n-k} \right) \quad (2.25)$$

dengan r_k adalah nilai korelasi diri sisaan pada lag ke- k , n banyaknya data yang diamati, dan k adalah lag maksimum. Statistik uji Q_* Ljung-Box menyebar mengikuti sebaran $\chi^2_{(k-p-q)}$ dengan p adalah ordo AR dan q adalah ordo MA. Berikut ini proses pengujian *white noise* dengan menggunakan uji Ljung-box, adapun perumusan hipotesis yang dilakukan adalah sebagai berikut:

$H_0: \rho_1 = \rho_2 = \dots = \rho_k = 0$ (Residual/sisaan memenuhi syarat *white noise*)

H_1 : minimal ada satu $\rho_i \neq 0$ untuk $i = 1, 2, \dots, k$ (Residual belum memenuhi syarat *white noise*)

Dengan daerah kritis: jika nilai Q_* lebih besar dari nilai $\chi^2_{(k-p-q)}$ untuk tingkat kepercayaan tertentu ($df = k - p - q$) atau nilai peluang statistik Q_* Ljung-Box lebih kecil dari taraf nyata (α) atau p -value $> \alpha$, maka dapat dikatakan bahwa sisaan tidak saling bebas (H_0 ditolak). Selain pengecekan kebebasan pada sisaan, kenormalan pada sisaan dapat dilihat dari p -value hasil uji *shapiro-wilk normality*. Jika p -value yang dihasilkan $> \alpha$, maka dapat disimpulkan bahwa sisaan telah memenuhi asumsi kenormalan sisaan, dengan $\alpha = 5\%$.

Setelah semua proses dalam metode Box-Jenkins dilakukan tahap berikutnya adalah melakukan validasi model terhadap model-model sementara yang telah

diperoleh dan menetapkan model terbaik yang dapat digunakan untuk prediksi beberapa waktu kedepan dengan membandingkan nilai *Akaike's Information Criterion* (AIC) dan *Root Mean Square Error* (RMSE). Untuk penentuan model dengan menggunakan kriteria pemilihan model terbaik melalui pendekatan nilai AIC, dimana nilai AIC semakin kecil maka model yang didapatkan semakin baik dengan mempertimbangkan banyaknya parameter dalam model, persamaan AIC sebagai berikut:

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

dengan n banyaknya pengamatan, M banyaknya parameter dalam model, serta $\hat{\sigma}_a^2$ merupakan estimasi varians residual. RMSE diperoleh dari akar *Mean Square Error* (MSE) dengan n periode peramalan dirumuskan sebagai berikut:

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

model terbaik dipilih berdasarkan nilai RMSE terkecil. Jika dalam proses ini didapatkan model yang relevan dengan data, maka langkah terakhir adalah proses peramalan. Peramalan merupakan proses untuk menentukan data beberapa periode waktu kedepan dari titik waktu ke $-t$ (Wei: 2006).

2.3 Metode Deret Waktu *Multivariate Ruang-Waktu (Space Time)*

Nilai dari suatu objek sering kali berubah seiring berjalannya waktu. Namun, ada pula objek yang berubah tidak hanya dipengaruhi oleh waktu, tapi juga dipengaruhi kondisi wilayah di sekitarnya. Hal yang demikianlah memberikan efek bahwa dalam melakukan prediksi nilai objek yang dipengaruhi oleh waktu dan ruang tidak dapat diselesaikan dengan hanya mengandalkan model deret waktu yang merupakan fungsi dari waktu saja, namun diperlukan sebuah model yang mampu merepresentasikan faktor waktu dan ruang (*space*) yang biasa disebut sebagai *spatio temporal* atau *space time models*, dengan demikian penerapan metode deret ruang-

waktu digunakan pada kajian data deret waktu *multivariate*. Metode ruang waktu erat kaitannya dengan model *Space Time Autoregressive* (STAR) yang pertama kali diperkenalkan oleh Cliff pada tahun 1975 (Giacomini: 2004). Dalam hal ini model ruang waktu (*space time*) merupakan salah satu model yang dapat menggambarkan dan meramalkan secara bersamaan suatu N observasi deret waktu, dimana N merupakan jumlah bagian wilayah observasi dan hubungan antar wilayah yang berbeda diwujudkan dalam bentuk matriks pembobot.

2.3.1 Stasioneritas Model Deret Waktu *Multivariate*

Analisis deret waktu *multivariate* pada umumnya digunakan untuk memodelkan dan menjelaskan interaksi serta pergerakan diantara sejumlah variabel deret waktu. Variabel atau peubah tersebut dinotasikan $Z_t = (Z_{1t}, \dots, Z_{mt})'$ untuk $t = \{0, \pm 1, \pm 2, \dots\}$ dan m menunjukkan suatu dimensi deret waktu pada vektor variabel acak. Pilihan komponen deret waktu *univariate* Z_{it} yang termasuk dalam Z_t akan tergantung pada daerah dan pemahaman tentang sistem yang diteliti, tetapi secara implisit bahwa komponen deret akan saling baik serentak di *time lags*. Representasi dan pemodelan ini saling berkaitan yang akan menjadi kepentingan utama analisis deret waktu *multivariate*. Seperti dengan kasus *univariate*, yang penting dalam konsep model representasi dan analisis deret waktu, yang memungkinkan berguna. Hasil pemodelan yang akan diperoleh dari sampel realisasi terbatas deret, adalah stasioneritas. Menurut Box (2008) serupa dengan kasus *univariate*, proses vektor Z_t dikatakan stasioner kuat jika distribusi peluang vektor acak $Z_{t_1}, Z_{t_2}, \dots, Z_{t_m}$ dan $Z_{t_1+l}, Z_{t_2+l}, \dots, Z_{t_m+l}$ adalah sama untuk deret t_1, t_2, \dots, t_m dan semua lag $l = \{0, \pm 1, \pm 2, \dots\}$. Dengan demikian, distribusi probabilitas pengamatan dari proses vektor stasioner adalah invarian terhadap pergeseran waktu.

Pada model *multivariate*, peramalan data dilakukan dengan menambahkan variabel lain yang mempunyai hubungan jangka panjang untuk mendapatkan keakuratan peramalan. Uji yang sangat sederhana untuk melihat kestasioneran data

adalah dengan analisis grafik, yang dilakukan dengan membuat plot antara nilai observasi dan waktu. Berdasarkan plot tersebut dapat dilihat pola data. Jika diperkirakan mempunyai mean dan variansi yang konstan, maka data tersebut dapat disimpulkan stasioner. Dalam menentukan kestasioneran sebaran data dengan menggunakan grafik tidak mudah, untuk itu dibutuhkan uji formal dalam menentukan kestasioneran data. Ada dua macam pengujian formal yang dapat dilakukan yaitu korelogram dan *unit root test*. Serupa dengan model deret waktu *univariate* secara visual kestasioneran data pada model deret waktu *multivariate* juga dapat dilihat dari plot MACF dan MPACF. Plot MACF dan MPACF yang turun secara lambat mengindikasikan bahwa data belum stasioner dalam rata-rata sehingga perlu dilakukan pembedaan untuk menstasionerkan data. Selain korelogram, kestasioneran juga dapat dilihat dengan menggunakan sebuah uji formal yaitu uji *Augmented Dickey-Fuller* (ADF). Kestasioneran dalam varians, dikatakan belum stasioner jika dalam plot data menunjukkan data mengalami fluktuasi yang berubah-ubah dan tidak konstan dari waktu ke waktu. Agar data stasioner dalam varians, maka transformasi perlu dilakukan (Wei: 2006).

2.3.2 Matrix Autocorrelation Function (MACF) dan Matrix Partial Autocorrelation Function (MPACF)

Suatu vektor deret waktu sebanyak n pengamatan Z_1, Z_2, \dots, Z_n matriks korelasi sampel dinyatakan sebagai

$$\hat{\rho}(k) = [\hat{\rho}_{ij}(k)] \quad (2.26)$$

dimana $\hat{\rho}_{ij}(k)$ merupakan korelasi silang sampel dari komponen deret ke- i dan ke- j yaitu

$$\hat{\rho}_{ij}(k) = \frac{\sum_{t=1}^{n-k} (Z_{i,t} - \bar{Z}_i)(Z_{j,t+k} - \bar{Z}_j)}{\left[\sum_{t=1}^n (Z_{i,t} - \bar{Z}_i)^2 \sum_{t=1}^n (Z_{j,t} - \bar{Z}_j)^2 \right]^{1/2}} \quad (2.27)$$

dengan \bar{Z}_i dan \bar{Z}_j adalah mean sampel dari komponen deret yang bersesuaian. Fungsi matriks korelasi (*Matrix Autocorrelation Function*) sampel sangat diperlukan dalam

model MA, bila matriks korelasinya bernilai nol setelah lag ke- q maka model yang bersesuaian adalah MA(q). Untuk bentuk matriks dan grafik semakin kompleks apabila dimensi dan vektornya semakin besar, sehingga menyulitkan dalam hal pengidentifikasian. Sehingga untuk mempermudah metode yang digunakan adalah dengan menggunakan simbol yang dinotasikan dengan (+), (-) dan (.) pada matriks korelasi sampel ke (i, j) . Simbol (+) diartikan sebagai $\hat{\rho}_{ij}(k)$ lebih besar dari 2 kali standart error dan menunjukkan hubungan memiliki korelasi positif. Simbol (-) menyatakan suatu nilai $\hat{\rho}_{ij}(k)$ kurang dari -2 kali standart error dan menunjukkan hubungan yang memiliki korelasi negatif. Sedangkan symbol (.) menotasikan $\hat{\rho}_{ij}(k)$ berada diantara ± 2 kali standart error dan menunjukkan tidak adanya korelasi (Wei: 2006).

Fungsi matriks parsial korelasi (*Matrix Partial Autocorrelation Function*) sampel sangat diperlukan dalam model AR. Korelasi antara Z_t dan Z_{t+k} bisa diketahui setelah ketergantungan linear pada peubah $Z_{t+1}, Z_{t+2}, \dots, Z_{t+k-1}$ dihilangkan. Persamaan matriks fungsi korelasi diri parsial (MPACF) dirumuskan sebagai berikut (Wei: 2006):

$$\phi_{kk} = \frac{\text{cov}[(Z_t - \hat{Z}_t), (Z_{t+k} - \hat{Z}_{t+k})]}{\sqrt{\text{var}(Z_t - \hat{Z}_t)} \sqrt{\text{var}(Z_{t+k} - \hat{Z}_{t+k})}} \quad (2.28)$$

matriks fungsi korelasi parsial pada lag ke- k dinotasikan dengan $\wp(k)$ sebagai koefisien matriks terakhir jika data diterapkan untuk suatu proses vektor regresi diri (*autoregressive*) pada orde ke- k . Hal ini merupakan pengembangan definisi fungsi parsial sampel untuk deret waktu *univariate* yang dikemukakan oleh Box-Jenkins pada tahun 1976. Sehingga $\wp(k)$ sama dengan Φ_{ss} dalam regresi linier peubah ganda (*multivariate*). Seperti fungsi korelasi parsial (PACF) untuk kasus deret waktu peubah tunggal (*univariate*), MACF juga bersifat terputus setelah lag p pada model VAR(p). Sama halnya dengan deret waktu *univariate*, untuk identifikasi pada model

deret waktu *multivariate* juga dapat dilihat dari pola atau matriks fungsi korelasi (MACF) dan matriks fungsi korelasi parsial (MPACF) (Wei: 2006).

2.3.3 Model *Space Time Autoregressive* (STAR)

Model *Space Time Autoregressive* (STAR) atau regresi diri ruang waktu adalah model yang dikategorikan berdasarkan lag yang berpengaruh secara linier baik dalam lokasi maupun waktu (Pfeifer: 1980). Seperti metode deret waktu pada umumnya, model ruang-waktu ini juga membutuhkan suatu sistem yang tidak berubah. Menurut Pfeifer (1980) model regresi diri ruang waktu dengan orde *autoregressive* p dan orde spasial λ_p atau dapat dituliskan STAR($p\lambda_p$) yang dirumuskan sebagai berikut:

$$Z(t) = \sum_{k=1}^p \sum_{l=0}^{\lambda_k} \Phi_{kl} W^{(l)} Z(t-k) + e(t) \quad (2.29)$$

dengan

Φ_{kl} = parameter *autoregressive* pada lag waktu ke $-k$ dan lag spasial ke $-l$

λ_k = spasial lag dari bentuk *autoregressive* orde p

$W^{(l)}$ = matriks bobot ukuran $(N \times N)$ pada lag spasial $l = 0, 1, 2, \dots, N$

$e(t)$ = vektor residual ukuran $(N \times 1)$ berdistribusi normal *multivariate* dengan mean 0 dan matriks varians-kovarians $\sigma^2 I_N$

$Z(t)$ = vektor acak ukuran $(N \times 1)$ pada waktu t

N = jumlah lokasi.

2.3.4 Pemilihan Bobot Lokasi Pada Model STAR

Untuk pemilihan atau penentuan bobot lokasi pada model STAR, dilakukan dengan menggunakan matriks bobot seragam. Matriks bobot merupakan matriks bujur sangkar yang memiliki entri-entri berupa bobot lokasi yang bersesuaian. Bobot untuk entri matriks pada model STAR biasanya ditentukan dengan memperhatikan sifat-sifat fisik atau karakteristik misalnya luas wilayah, kepadatan penduduk, batas

antara dua lokasi, atau sarana transportasi, dimana setiap bobot tersebut tidak tergantung pada waktu. Asumsi untuk kajian bobot ini adalah bahwa bobot suatu lokasi terhadap dirinya sendiri adalah nol.

Bobot w_{ij} pada lag spasial 1 dinyatakan oleh W berupa matriks bujur sangkar $N \times N$ sebagai berikut:

$$W = \begin{bmatrix} 0 & w_{12} & \cdots & w_{1N} \\ w_{21} & 0 & \cdots & w_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ w_{N1} & w_{N2} & \cdots & 0 \end{bmatrix}$$

Salah satu bobot yang sering digunakan adalah bobot seragam. Bobot seragam ditentukan berdasarkan banyaknya tetangga terdekat dalam suatu kelompok lokasi tertentu (lag spasial). Secara matematis bobot seragam ditentukan sebagai berikut:

$$W_{ij} = \begin{cases} \frac{1}{N} & , \text{ jika } i \text{ dan } j \text{ merupakan tetangga terdekat} \\ 0 & , \text{ jika lainnya} \end{cases} \quad (2.30)$$

dengan N menyatakan banyaknya tetangga terdekat dari lokasi i pada lag spasial 1. Sifat-sifat matriks bobot seragam diantaranya:

$$W_{ij} > 0$$

$$\sum_{j=1}^N W_{ij} = 1, \text{ untuk setiap } i \text{ dan } \sum_{i=1}^N \sum_{j=1}^N W_{ij} = N$$

diagonal matriks bobot W adalah nol (Ruchjana: 2001).

2.3.5 Estimasi Parameter Model STAR

Dalam melakukan estimasi parameter model STAR ditentukan dengan menggunakan metode kuadrat terkecil (MKT), estimasi parameter $\hat{\phi}_{kl}$ yang memiliki k sebagai order dan l sebagai spasial order dari model STAR, diperoleh dengan menentukan nilai $\hat{\phi}_{kl}$ yang menghasilkan jumlah kuadrat residu minimum. Apabila diambil sebuah model STAR(1₁) yaitu

$$Z(t) = \phi_{10}Z(t-1) + \phi_{11}WZ(t-1) + e(t)$$

mempunyai bentuk model linier yang dapat ditulis sebagai berikut:

$$\begin{bmatrix} Z(2) \\ Z(3) \\ \dots \\ Z(T) \end{bmatrix} = \begin{bmatrix} Z(1) \\ Z(2) \\ \dots \\ Z(T-1) \end{bmatrix} \begin{bmatrix} W^{(1)}Z(1) \\ W^{(1)}Z(2) \\ \dots \\ W^{(1)}Z(T-1) \end{bmatrix} \begin{bmatrix} \phi_{10} \\ \phi_{11} \end{bmatrix} + \begin{bmatrix} e(2) \\ e(3) \\ \dots \\ e(T) \end{bmatrix} \quad (2.31)$$

Persamaan (2.31) dapat dituliskan menjadi $Z(t) = X\vec{\phi} + e(t)$

dimana $X = [Z(t-1) \quad WZ(t-1)]$ (Ruchjana: 2001a).

Galat model STAR(1₁) adalah sebagai berikut:

$$e(t) = Z(t) - \phi_{10}Z(t-1) - \phi_{11}WZ(t-1)$$

Jumlah kuadrat residu STAR(1₁) dinyatakan dengan:

$$S = (e(t))^2 = (Z(t) - \phi_{10}Z(t-1) - \phi_{11}WZ(t-1))^2$$

Turunan terhadap ϕ_{10} dan ϕ_{11} adalah

$$\frac{\partial S}{\partial \phi_{10}} = -2[Z(t) - \phi_{10}Z(t-1) - \phi_{11}WZ(t-1)]Z(t-1) = 0$$

$$Z(t)Z(t-1) = \hat{\phi}_{10}Z(t-1)Z(t-1) + \hat{\phi}_{11}WZ(t-1)Z(t-1)$$

$$\frac{\partial S}{\partial \phi_{11}} = -2[Z(t) - \phi_{10}Z(t-1) - \phi_{11}WZ(t-1)]WZ(t-1) = 0$$

$$Z(t)WZ(t-1) = \hat{\phi}_{10}Z(t-1)WZ(t-1) + \hat{\phi}_{11}WZ(t-1)WZ(t-1)$$

Dalam bentuk matriks menjadi:

$$\begin{bmatrix} Z(t)Z(t-1) \\ Z(t)WZ(t-1) \end{bmatrix} = \begin{bmatrix} Z(t-1)Z(t-1) \\ Z(t-1)WZ(t-1) \end{bmatrix} \begin{bmatrix} \hat{\phi}_{10} \\ \hat{\phi}_{11} \end{bmatrix}$$

$$\begin{bmatrix} \hat{\phi}_{10} \\ \hat{\phi}_{11} \end{bmatrix} = \begin{bmatrix} Z(t-1)Z(t-1) \\ Z(t-1)WZ(t-1) \end{bmatrix}^{-1} \begin{bmatrix} Z(t)Z(t-1) \\ Z(t)WZ(t-1) \end{bmatrix}$$

Atau

$$\hat{\phi} = (X^t X)^{-1} X^t Z \text{ dengan } \hat{\phi} = \begin{bmatrix} \hat{\phi}_{10} \\ \hat{\phi}_{11} \end{bmatrix}, X = [Z(t-1) \quad WZ(t-1)] \text{ dan } X^t X \text{ matriks}$$

non singular (Suryamah: 2013). Untuk proses ini dapat dilanjutkan menggunakan model linier dengan bantuan program pengolahan data statistika R. Estimasi parameter STAR(1₁) mengharuskan memenuhi kondisi stasioner sebagai berikut (Pfeifer : 1980):

$$-1 < \hat{\phi}_{10} < 1 \text{ dan } -1 < \hat{\phi}_{11} < 1 \quad (2.32)$$

2.3.6 Pengujian Kecocokan Model STAR

Untuk mencegah galat peramalan yang terlalu besar, harus dilakukan pengecekan terlebih dahulu terhadap model tersebut apakah sudah cocok atau belum untuk data tersebut. Pengujian dilakukan secara serentak dan secara individu. Pengujian parameter secara serentak merupakan uji serentak yang dilakukan untuk memeriksa keberartian koefisien ϕ secara keseluruhan atau serentak. Perumusan hipotesis adalah sebagai berikut:

$$H_0 : \phi_{10}^{(i)} = \phi_{11}^{(i)} = 0$$

$$H_1 : \text{minimal ada satu } \phi_{kl}^{(i)} \neq 0 \text{ dengan } k = 1, l = 0,1 \text{ dan } i = 1,2,\dots,N$$

Statistik uji yang digunakan adalah F –hitung = $\frac{MSR}{MSE}$

$$MSR \text{ (Mean Square Regression)} = \frac{\sum_{i=1}^n (\hat{Z}_i(T) - \bar{Z}(T))^2}{n} \text{ dan}$$

$$MSE \text{ (Mean Square Error)} = \frac{1}{n} \sum_{l=1}^n (Z_{l+1}(T) - \hat{Z}_l(T))^2$$

dengan n merupakan banyak ramalan yang dilakukan. Kriteria pengujian dengan mengambil taraf nyata $\alpha = 6\%$, $df_1 = N \times N$, $df_2 = n$ maka dari tabel distribusi F diperoleh $F_{(\frac{\alpha}{2}, (N \times N), n)}$. Jika F –hitung $> F$ –tabel maka H_0 ditolak. Selain dari daerah kritik tersebut, dapat juga digunakan daerah kritik yang lain yaitu H_0 ditolak apabila p –value $< \alpha = 6\%$ yang artinya $\phi_{kl}^{(i)}$ secara statistik mempunyai arti dalam model (sesuai sebagai model peramalan).

Pengujian parameter secara parsial (individu) dilakukan untuk menguji keberartian koefisin ϕ secara parsial dengan membandingkan dugaan ϕ dengan penduga standart errornya. Dengan hipotesisnya dalah sebagai berikut:

$$H_0 : \phi_{1l}^{(i)} = 0 \text{ (parameter tidak signifikan)}$$

$$H_1 : \phi_{1l}^{(i)} \neq 0 \text{ (parameter signifikan)}$$

dengan $l = 0,1$ dan $i = 1,2,\dots,N$

$$\text{Statistika uji yang digunakan adalah } t \text{ –hitung} = \frac{|\phi_{1l}^{(i)}|}{SE(\phi_{1l}^{(i)})}$$

dimana $SE(\phi_{10}^{(i)}) = \left\{ \frac{\sum_{i=1}^n (Z_i(t))^2}{n \sum_{i=1}^n (Z_i(t) - \bar{Z}_i(t))^2} \right\}^{\frac{1}{2}} \times \sigma$ dan $SE(\phi_{11}^{(i)}) = \left\{ \frac{\sigma^2}{n \sum_{i=1}^n (Z_i(t) - \bar{Z}_i(t))^2} \right\}^{\frac{1}{2}}$

serta $\phi_{1l}^{(i)}$ adalah parameter dengan $l = 0,1$ dan $i = 1,2,\dots,N$; $SE(\phi_{1l}^{(i)})$ merupakan standart error untuk parameter dengan $l = 0,1$ dan $i = 1,2,\dots,N$.

Kriteria pengujian dengan mengambil taraf nyata $\alpha = 6\%$, $df = n$ maka dari tabel distribusi t diperoleh $t_{(\frac{\alpha}{2};n)}$. Jika t -hitung $>$ t -tabel maka H_0 ditolak. Selain dari daerah kritik tersebut, dapat juga digunakan daerah kritik yang lain yaitu H_0 ditolak bila p -value $<$ $\alpha = 6\%$ yang artinya $\phi_{1l}^{(i)}$ secara statistik mempunyai arti dalam model (sesuai sebagai model peramalan).

2.3.7 Kriteria Pemilihan Model STAR Terbaik

Akaike's Information Criterion (AIC) dan nilai *Root Mean Square Error* (RMSE) akan digunakan dalam proses pemilihan model terbaik. Berikut ini dijelaskan masing-masing kriteria pemilihan model terbaik. Salah satu kriteria pemilihan dalam penentuan model terbaik pada *training* adalah *Akaike's Information Criterion* (AIC). Model terbaik adalah model dengan nilai AIC paling kecil. Berikut cara perhitungan nilai AIC:

$$AIC(p) = \ln|\tilde{\Sigma}_u(p)| + \frac{2p}{T} K^2$$

dimana $\tilde{\Sigma}_u p = T^{-1} \sum_{t=1}^T \hat{u}_t \hat{u}_t'$ adalah matriks taksiran kovarian residual dari model VAR(p), T merupakan banyak observasi dan K merupakan jumlah parameter dalam model.

Root Mean Square Error (RMSE) adalah ukuran perbedaan antara nilai prediksi dari model atau penaksir dengan nilai sebenarnya dari observasi. RMSE digunakan untuk memperoleh gambaran keseluruhan standar deviasi yang muncul saat menunjukkan perbedaan antara model, atau hubungan yang dimiliki. Untuk mengetahui besarnya nilai RMSE, dapat digunakan rumus sebagai berikut (Wei: 2006):

$$\text{RMSE} = \sqrt{\text{MSE}} = \sqrt{\frac{1}{M} \sum_{t=1}^M (Z_t - \hat{Z}_t)^2}$$

dengan M merupakan banyak data ramalan yang digunakan, Z_t data sebenarnya, \hat{Z}_t data hasil ramalan. Nilai RMSE berkisar antara 0 sampai ∞ . Semakin kecil nilai RMSE maka model semakin bagus.



BAB 3. METODE PENELITIAN

Pada bab ini akan dipaparkan beberapa uraian mengenai metode penelitian sebagai konsep untuk melakukan penelitian ini. Adapun metode penelitian ini meliputi data penelitian dan metode analisis dengan menggunakan ARIMA dan STAR.

3.1 Data Penelitian

Data yang digunakan adalah data sekunder yang diperoleh dari Dinas Pengairan kabupaten Jember yaitu data rata-rata curah hujan bulanan periode bulan Januari 2005 sampai bulan Juni 2013 yang dibagi menjadi empat bagian wilayah yakni Jember Barat, Jember Selatan, Jember Tengah, Jember Timur (Gambar 3.1). Untuk masing-masing wilayah terdiri dari 102 data dan dibagi menjadi 2 bagian.

- a. Data *training* (*in-sample*/pemodelan) : Januari 2005–Desember 2012 (96 data)
- b. Data *testing* (*out-sample*/validasi) : Januari 2013–Juni 2013 (6 data)

Untuk variabel-variabel yang digunakan dalam penelitian ini adalah sebagai berikut:

- a. Variabel yang digunakan untuk model ARIMA

$Y_1(t)$: Rata-rata curah hujan di wilayah Jember Barat

$Y_2(t)$: Rata-rata curah hujan di wilayah Jember Selatan

$Y_3(t)$: Rata-rata curah hujan di wilayah Jember Tengah

$Y_4(t)$: Rata-rata curah hujan di wilayah Jember Timur

- b. Variabel yang digunakan untuk model STAR

$Z_1(t)$: Rata-rata curah hujan di wilayah Jember Barat

$Z_2(t)$: Rata-rata curah hujan di wilayah Jember Selatan

$Z_3(t)$: Rata-rata curah hujan di wilayah Jember Tengah

$Z_4(t)$: Rata-rata curah hujan di wilayah Jember Timur



Gambar 3.1 Peta Persebaran Wilayah di Kabupaten Jember (Sumber: DPU Pengairan Kab.Jember)

3.2 Metode Analisis

Dari informasi data yang telah diperoleh, selanjutnya adalah melakukan pengolahan data melalui metode analisis berikut ini dengan menggunakan metode statistika yang dibantu dengan program pengolahan data statistika R 3.0.3 dan SAS. Berikut tahapan analisis yang digunakan untuk penelitian ini.

a. Tahap peramalan data deret waktu

Langkah-langkah yang terdapat dalam tahapan peramalan data deret waktu ini dilakukan dengan menggunakan pemodelan ARIMA Box-Jenkins antara lain berikut ini.

- 1) Identifikasi model, merupakan tahap penting yang merupakan proses pemilihan model. Berikut tahapan yang harus dilakukan dalam identifikasi model.
 - a) mengeksplorasi data dengan cara membuat plot data *training* secara grafis dan ACF untuk mengetahui kestasioneran data.
 - b) melakukan uji untuk mengetahui kestasioneran data terhadap mean yang dapat dilakukan dengan menggunakan uji *Augmented Dickey-Fuller* (ADF). Jika data belum stasioner dalam mean, baik mean non-musiman maupun mean musiman maka perlu dilakukan proses *differencing* (pembedaan) non-musiman maupun *differencing* musiman sampai data menjadi stasioner. Jika data belum stasioner dalam varians maka dilakukan transformasi *Box-Cox*.
 - c) apabila data telah stasioner maka dengan melihat pola ACF dan PACF akan dapat mengidentifikasi bentuk model-model sementara ARIMA.

2) Estimasi parameter

Estimasi parameter pada tahapan ini dapat dilakukan untuk menentukan apakah parameter dari setiap model-model sementara ARIMA tersebut sudah layak untuk digunakan, dengan melihat hasil estimasi parameter yang merupakan proses penentuan nilai parameter-parameter pada model yang dihasilkan melalui penggunaan metode kemungkinan maksimum dan kemudian dilakukan pengujian parameter-parameternya menggunakan statistik uji t

(*t*-test). Model dengan parameter-parameter yang $|t - hitung| > t_{(\alpha/2); df=n-1}$ atau *p*-value $> \alpha$ untuk $\alpha = 0,05$, akan diikuti sertakan pada proses berikutnya.

3) Diagnosa model

Diagnosa model merupakan proses untuk memeriksa ketepatan model yang dihasilkan dan memberikan petunjuk kearah perbaikan model dengan melakukan pemeriksaan sisaan (residual) yang meliputi asumsi *white noise* dengan menggunakan uji Ljung-Box (Persamaan 2.25). Model-model yang memenuhi asumsi sisaan (residual) dinamakan model yang layak dan diikutsertakan pada proses selanjutnya.

4) Pemilihan model ARIMA terbaik

Untuk menetapkan model ARIMA terbaik yang dapat digunakan untuk prediksi beberapa waktu kedepan berdasarkan hasil evaluasi atau kriteria pada data *training* yakni membandingkan nilai AIC dari model-model yang ada. Model ramalan yang baik adalah jika memiliki nilai AIC yang lebih kecil serta pada data *testing* dengan melihat nilai RMSE yang terkecil.

5) Melakukan peramalan data *testing* dengan model ARIMA yang terbentuk untuk beberapa waktu kedepan pada data curah hujan di masing-masing lokasi.

b. Tahap peramalan data deret ruang-waktu (*Space Time*)

Langkah-langkah yang terdapat dalam tahapan peramalan data deret ruang-waktu (*Space Time*) ini dilakukan dengan menggunakan pemodelan STAR, antara lain berikut ini.

1) Identifikasi model STAR

Berikut ini proses identifikasi model STAR yang harus dilakukan untuk melakukan pemilihan model.

a) Mendeskripsikan pola data dengan mencari nilai korelasi antar lokasi menggunakan matriks korelasi. Mengeksplorasi data dengan cara membuat plot data *training* secara grafis dan melakukan pengujian stasioneritas data

secara *multivariate* dengan membentuk plot MACF dan MPACF, serta secara *univariate* melalui plot ACF.

- b) Jika data tidak stasioner dalam mean maka dilakukan proses *differencing* (pembedaan) sampai data menjadi stasioner. Jika data tidak stasioner dalam variansi maka dilakukan transformasi *Box-Cox*.
- c) Apabila data telah stasioner maka dilakukan perhitungan dan analisis bentuk dari MACF dan MPACF, serta nilai AIC pada beberapa orde AR. Ketiga besaran ini digunakan sebagai dasar untuk penentuan orde model STAR, khususnya pada nilai AIC yang tekecil.
- d) Membentuk model $STAR(p_{(\lambda_p)})$ dengan menetapkan nilai matriks pembobot seragam.

2) Estimasi parameter

Setelah kandidat model STAR sudah dipilih dalam tahap identifikasi, tahap selanjutnya melakukan estimasi parameter model STAR yang ditentukan dengan menggunakan metode kuadrat terkecil.

3) Diagnosa pengecekan model

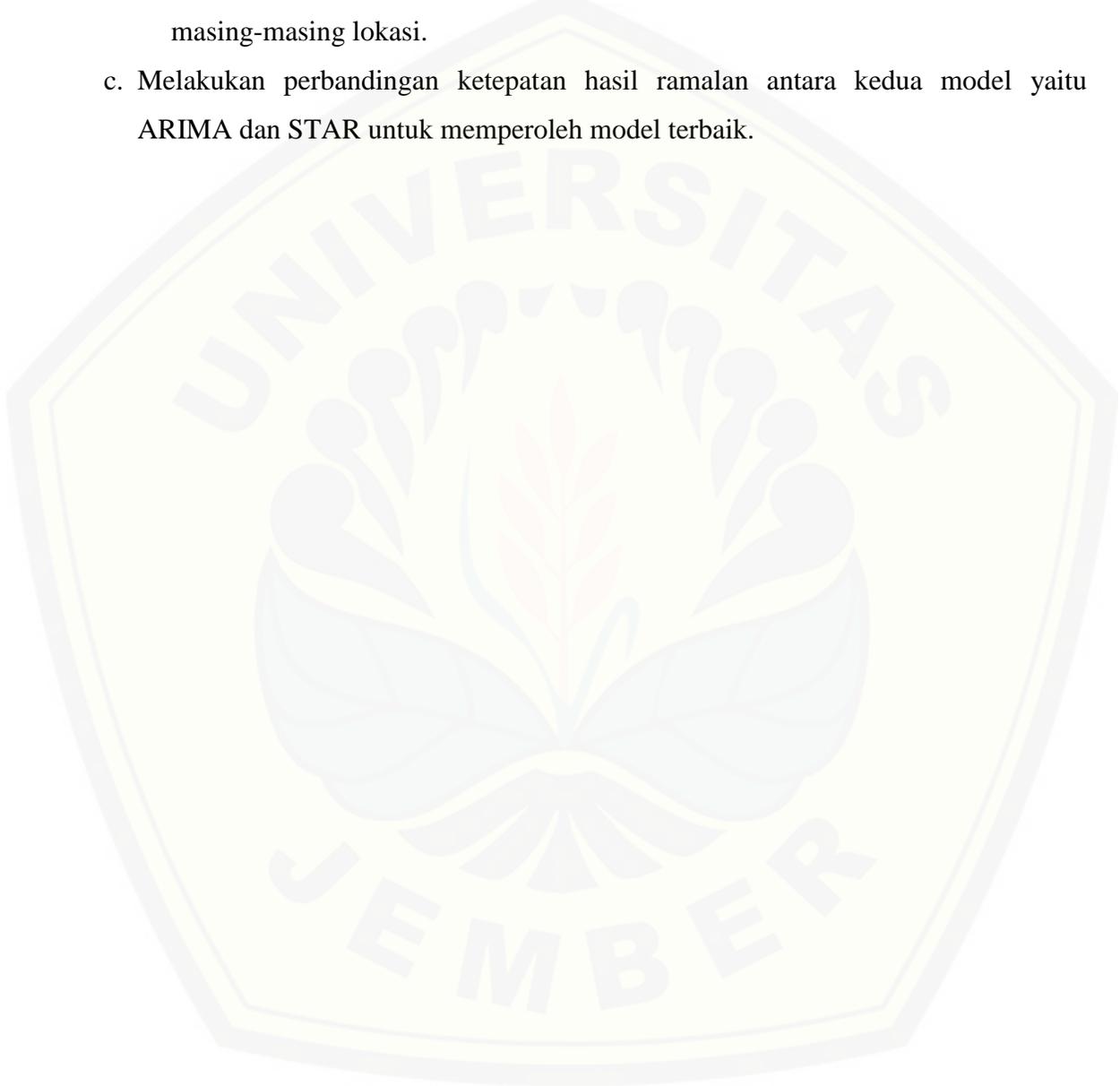
Setelah sebuah kandidat model dipilih dan parameter-parameternya diestimasi, model harus masuk dalam tahap uji diagnosa model untuk menentukan apakah model cukup untuk mewakili data. Untuk pengujian diagnosa model dilakukan dengan uji *white noise* residual. Pengujian asumsi *white noise* menggunakan uji *Portmanteau* dengan hipotesis:

$H_0: \rho_1 = \rho_2 = \dots = \rho_k = 0$ (Residual memenuhi syarat *white noise*)

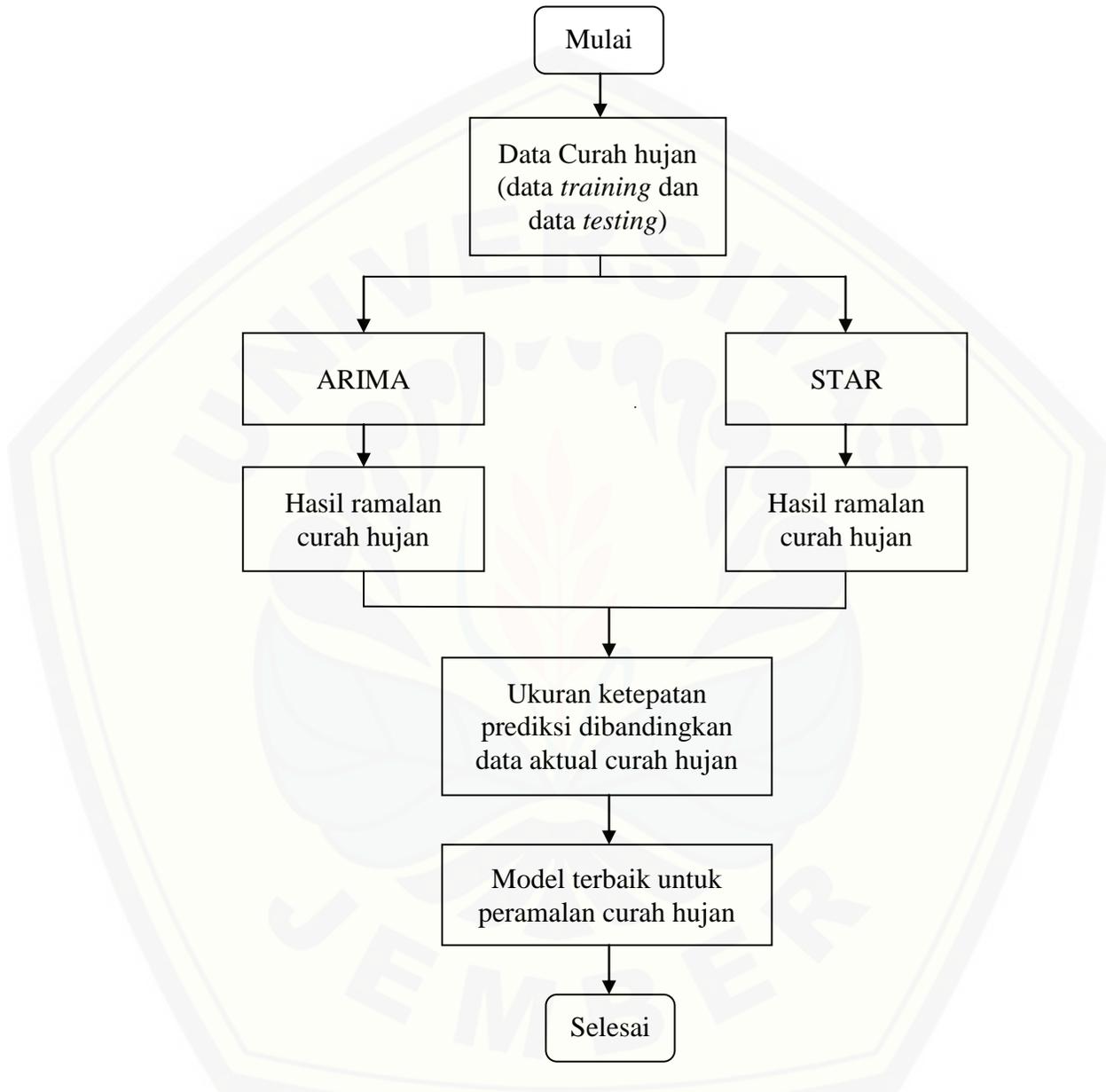
H_1 : minimal ada satu $\rho_i \neq 0$ untuk $i = 1, 2, \dots, k$ (Residual belum memenuhi syarat *white noise*)

dengan statistik uji $Q_{LB} = n(n+2) \left(\sum_{k=1}^m \frac{\hat{r}_k^2}{n-k} \right)$ dengan daerah penolakan yaitu tolak H_0 jika $Q_{LB} > \chi_{(k-p-q)}^2$ atau p -value $< \alpha = 5\%$ yang berarti bahwa data belum memenuhi syarat cukup (residual belum memenuhi syarat *white noise* (Box: 2004).

- 4) Menentukan model STAR terbaik berdasarkan nilai AIC dan RMSE terkecil yang dihasilkan dari model-model yang ada, selanjutnya dilakukan peramalan data *testing* untuk melihat gambaran curah hujan pada waktu mendatang di masing-masing lokasi.
- c. Melakukan perbandingan ketepatan hasil ramalan antara kedua model yaitu ARIMA dan STAR untuk memperoleh model terbaik.



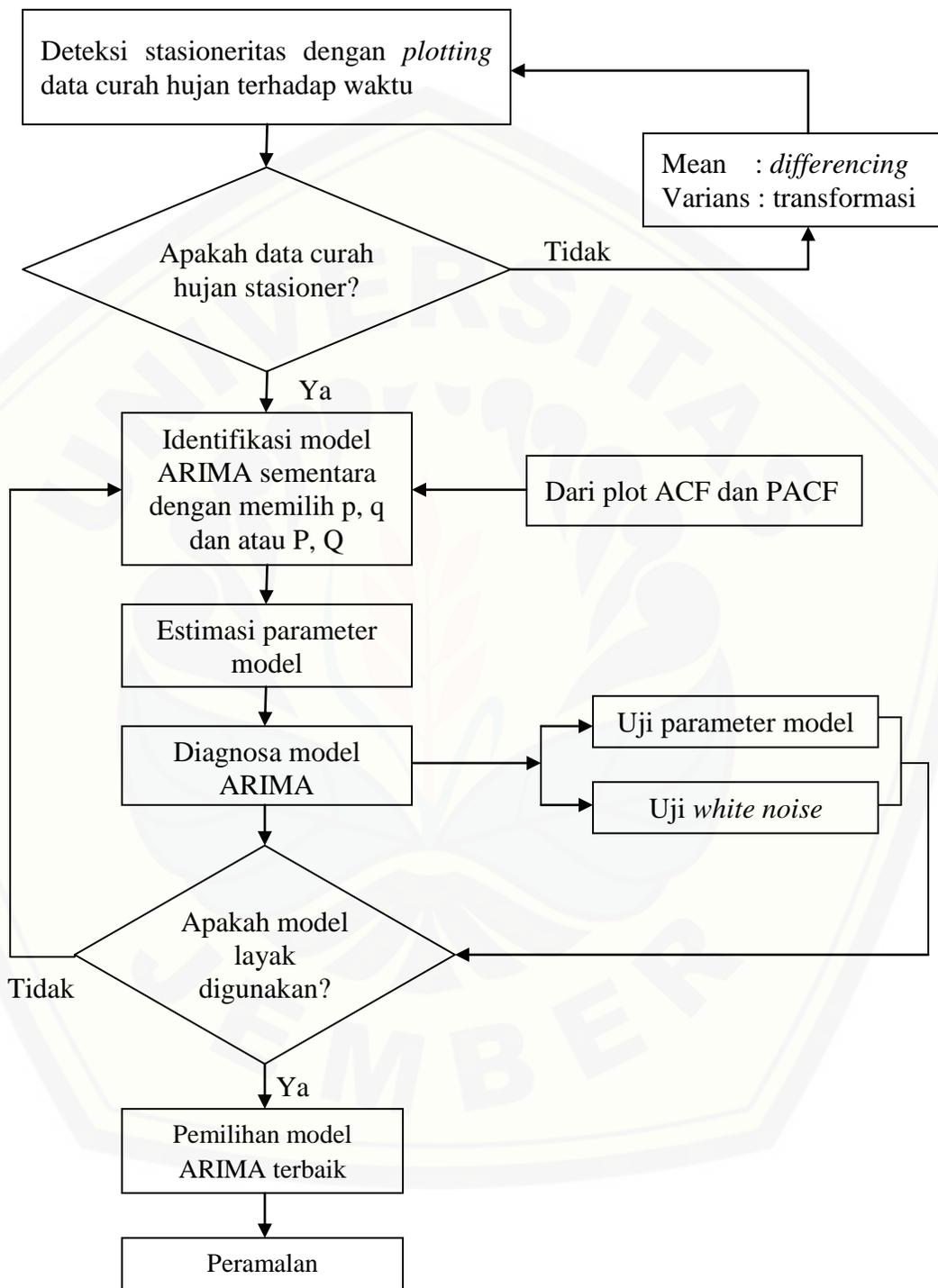
Skema dari tahapan analisis dapat dilihat pada Gambar 3.2 dibawah ini.



Gambar 3.2 Skema Proses Penelitian

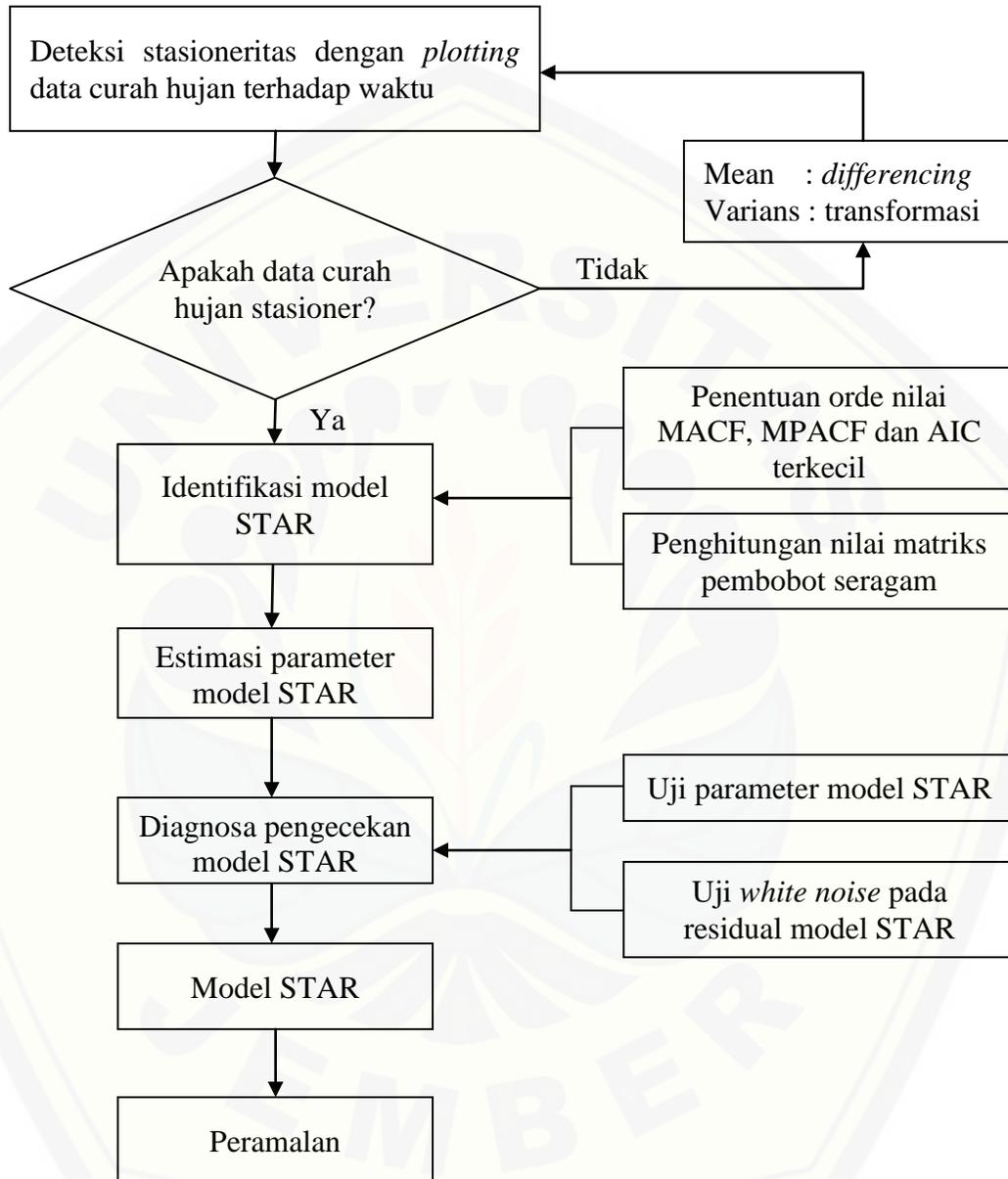
Selanjutnya akan dijelaskan langkah untuk tahapan ARIMA dan STAR pada proses penelitian tersebut yang dapat dilihat pada Gambar 3.3 dan Gambar 3.4 berikut.

Tahapan ARIMA



Gambar 3.3 Tahapan ARIMA

Tahapan STAR



Gambar 3.4 Tahapan STAR