



***PREDICTION INTERVAL PADA MODEL SEASONAL
AUTOREGRESSIVE INTEGRATED MOVING AVERAGE
(SARIMA) UNTUK PERAMALAN CURAH HUJAN DAN
INFORMASI KEKERINGAN***

PROPOSAL SKRIPSI

diajukan guna memenuhi persyaratan kelulusan S1

Oleh

**Vita Mami Nikmatillah
141810101013**

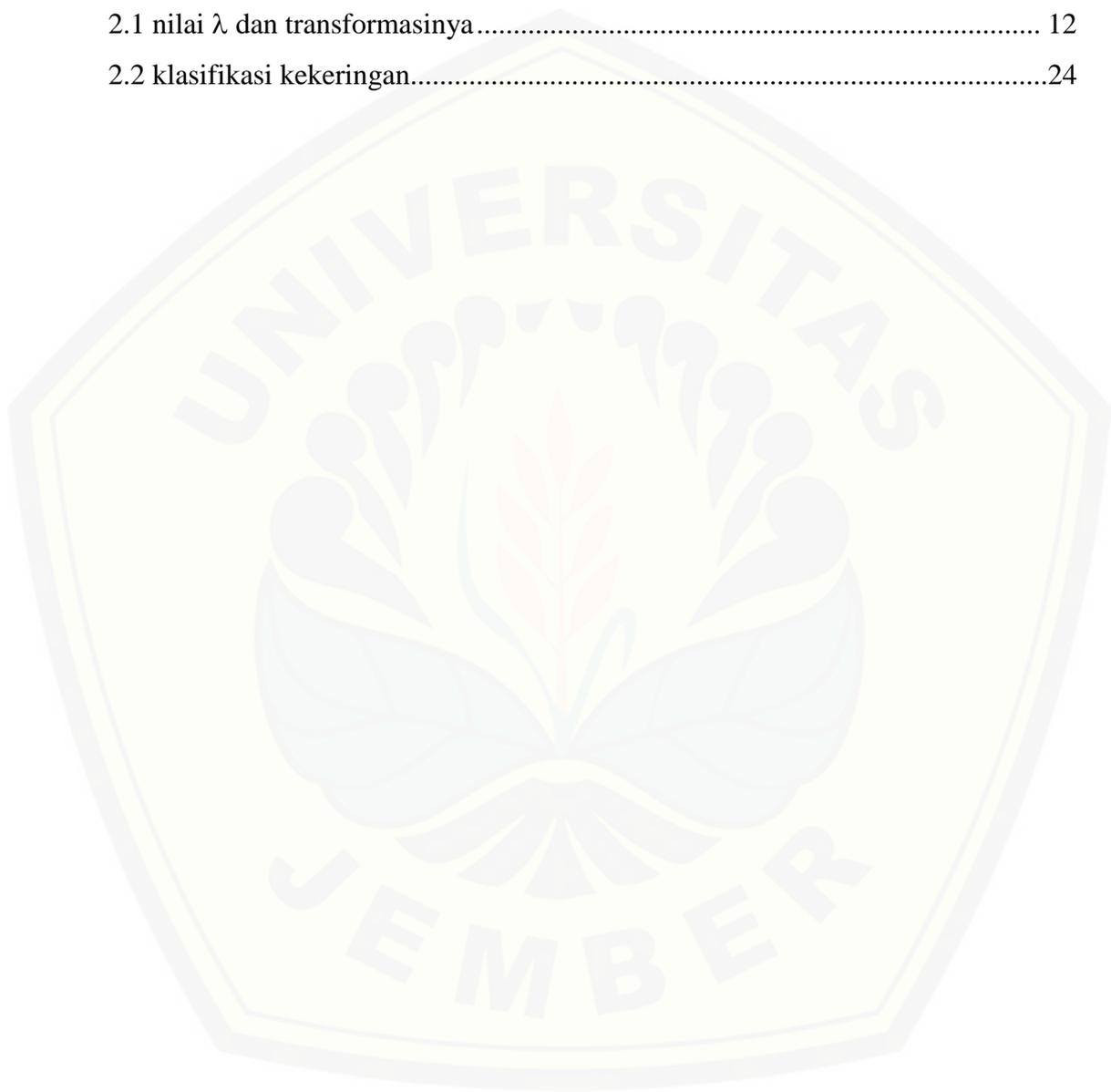
**JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS JEMBER
2018**

DAFTAR ISI

	Halaman
HALAMAN JUDUL	i
DAFTAR ISI	ii
DAFTAR TABEL	iii
DAFTAR GAMBAR	iv
BAB 1. PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Tujuan	4
1.4 Manfaat	4
BAB 2. TINJAUAN PUSTAKA	5
2.1 Peramalan	5
2.1.1 Metode Peramalan	5
2.2 Analisis Deret Waktu (<i>Time Series</i>)	6
2.3 Kestasioneran Data	8
2.3.1 Pembedaan Data (<i>Differencing</i>).....	10
2.3.2 Transformasi <i>Box-Cox</i>	12
2.3.3 Fungsi Autokorelasi dan Fungsi Autokorelasi Parsial	12
2.4 Model ARIMA (<i>Autoregressive Integrated Moving Average</i>)	16
2.5 Estimasi Parameter	19
2.6 Uji <i>White Noise</i>	20
2.7 Pemilihan Model Terbaik	21
2.8 Menghitung <i>Prediction Interval</i>	22
2.8 Metode SPI (<i>Standardized Precipitation Index</i>)	23
BAB 3. METODE PENELITIAN	25
3.1 Data	25
3.2 Langkah-Langkah Penelitian	25
DAFTAR PUSTAKA	28

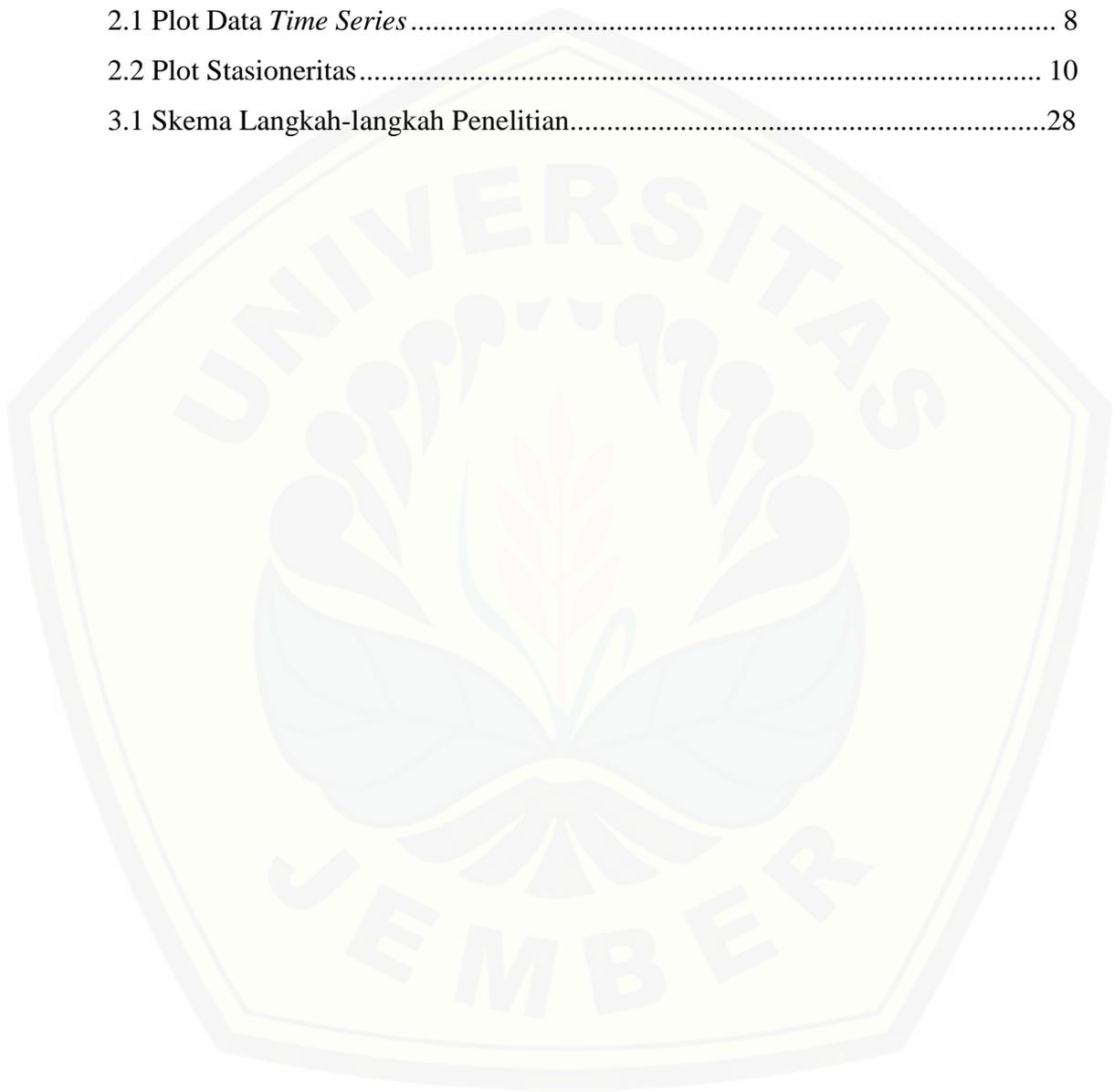
DAFTAR TABEL

	Halaman
2.1 nilai λ dan transformasinya.....	12
2.2 klasifikasi kekeringan.....	24



DAFTAR GAMBAR

	Halaman
2.1 Plot Data <i>Time Series</i>	8
2.2 Plot Stasioneritas	10
3.1 Skema Langkah-langkah Penelitian.....	28



BAB 1. PENDAHULUAN

1.1 Latar Belakang

Peramalan (*forecasting*) adalah analisis data *time series* yang menggunakan kejadian di masa lalu untuk mengetahui perkembangan kejadian di masa yang akan datang. Peramalan banyak digunakan dalam kehidupan sehari-hari sebagai modal perencanaan di masa yang akan datang, baik perencanaan kegiatan, perencanaan keuangan dan beberapa perencanaan lain yang membutuhkan prediksi. Prediksi adalah bilangan yang disebut sebagai titik ramal, dimana nilai yang dihasilkan belum tentu benar. Interval prediksi atau *prediction interval* merupakan interval dari nilai-nilai peramalan yang berjalan. Nilai ramal yang muncul cukup mendekati nilai yang akan dicapai. Perhitungan *prediction interval* dalam proses peramalan merupakan bagian terpenting untuk mengetahui indikasi nilai ramalan yang tak pasti. Nilai ketidakpastian inilah yang menjadi faktor interval prediksi dilakukan. Hal ini bertujuan untuk mengetahui informasi ketidakpastian yang diperlukan untuk kedepannya.

Badan Meteorologi dan Geofisika (BMKG) adalah salah satu instansi yang menerapkan ilmu peramalan dalam meramalkan jumlah curah hujan di suatu daerah. Curah hujan adalah jumlah air hujan yang turun pada suatu daerah dalam waktu tertentu. Letak topografi suatu daerah mempengaruhi curah hujan yang akan terjadi. Model peramalan yang banyak digunakan dalam meramalkan curah hujan adalah model ARIMA (*Autoregressive Integrated Moving Average*). Model ARIMA atau *Box-Jenkins* merupakan gabungan dari beberapa model seperti model *Autoregressive* (AR), model *Moving Average* (MA) dan model *Autoregressive Moving Average* (ARMA). ARIMA adalah penggabungan model (AR) dan model (MA) yang telah mengalami *differencing*. Model ARIMA banyak digunakan dalam peramalan curah hujan karena memiliki beberapa karakter yang sangat cocok untuk kasus curah hujan, terutama ARIMA musiman atau *Seasonal Autoregressive Integrated Moving Average* (SARIMA). Model SARIMA merupakan hasil modifikasi dari model ARIMA. SARIMA banyak digunakan pada data musiman.

Kekeringan adalah kejadian berkurangnya curah hujan dari kondisi normalnya dalam kurun waktu yang lama. Pada sektor pertanian, kekeringan merupakan hal yang sangat di takuti para petani karena dapat mempengaruhi produksi yang kemudian mengakibatkan kerugian. Fenomena kekeringan yang terjadi secara rutin ini maka perlu diadakan analisa kekeringan guna mengetahui tingkat kekeringan yang terjadi di suatu daerah. Metode yang digunakan dalam menganalisa tingkat kekeringan dengan menggunakan data curah hujan adalah *Standardized Precipitation Index* (SPI). Oleh karena itu perlu dilakukan peramalan curah hujan sebagai informasi dasar dalam menentukan indeks kekeringan di masa yang akan datang.

Beberapa penelitian tentang peramalan curah hujan telah dilakukan oleh beberapa peneliti diantaranya Lusiani (2011) pemodelan ARIMA curah hujan di kota Bandung, Ukhra (2014) pemodelan dan peramalan data deret waktu dengan metode SARIMA dan Retnaningrum (2015) penerapan model STAR (*Space Time Autoregressive*) dan ARIMA untuk peramalan data curah hujan di kabupaten Jember. Penelitian-penelitian ini terbatas pada satu titik data hasil peramalan tanpa mempertimbangkan suatu interval probabilitas tertentu. *Prediction interval* merupakan bagian penting dari proses peramalan yang dimaksud untuk menunjukkan kemungkinan ketidakpastian dalam perkiraan titik. Penelitian tentang *Prediction interval* telah dilakukan diantaranya YAR and Chatfield (1990) *Prediction interval* untuk Prosedur Peramalan Holt-Winters, Chatfield (1993) Menghitung *Prediction interval*, dan Safitri (2008) Interval Prediksi untuk model *Time Series*. Selain menentukan ramalan curah hujan yang akan datang dibutuhkan penelitian tentang tingkat kekeringan yang terjadi disuatu daerah. Mutjahiddin (2014) menyimpulkan bahwa Kekeringan terjadi akibat adanya penyimpangan kondisi cuaca dari kondisi normal yang terjadi di suatu wilayah. Penyimpangan tersebut dapat berupa berkurangnya curah hujan dibanding dengan kondisi normal. Kurniawan (2016) meneliti tentang kombinasi ARIMA dan *Standardized Precipitation Index* (SPI) untuk menentukan indeks kekeringan di kabupaten Boyolali.

Letak topografi yang beragam dan kondisi alam yang mengalami perubahan suhu yang sangat dinamis ini menyebabkan kabupaten Jember melakukan beberapa upaya untuk mengurangi dampak yang terjadi. Berdasarkan penelitian diatas peneliti ingin memberikan penelitian baru. Peneliti ingin memberikan bentuk interval peramalan untuk curah hujan dengan model *time series* yang mengandung parameter musiman dan menentukan tingkat kekeringan dari hasil ramalan curah hujan. Data penelitian ini menggunakan data curah hujan 77 stasiun hujan di kabupaten Jember yang tersebar berdasarkan letak topografi. Terakhir, menghitung interval prediksi nilai peramalan dengan menggunakan model SARIMA serta menganalisa tingkat kekeringan yang akan terjadi dengan metode SPI.

1.2 Rumusan Masalah

Pada penelitian ini dilakukan peramalan pada data curah hujan di kabupaten Jember dengan menggunakan model SARIMA. Data hasil ramalan curah hujan akan dihitung nilai *Prediction interval*. Perhitungan *Prediction interval* ini sangat penting untuk memberikan ketidakpastian pada nilai ramal. Selain memperoleh nilai *Prediction interval*, nilai peramalan ini digunakan untuk mengklasifikasi tingkat kekeringan. Nilai ramal yang diperoleh ditransformasi menggunakan metode SPI untuk mendapat informasi tingkat kekeringan. Berdasarkan hal tersebut, rumusan masalah yang akan dibahas dalam penelitian ini adalah:

1. Bagaimana memperoleh nilai ketidakpastian dengan membangun *Prediction Interval* (PI) pada nilai ramal curah hujan yang diperoleh dari model SARIMA?
2. Bagaimana cara menghitung indeks kekeringan dengan metode *Standardized Precipitation Index* (SPI) pada nilai ramalan yang diperoleh dari model SARIMA?
3. Bagaimana indeks kekeringan yang terjadi di kabupaten Jember yang diperoleh dari nilai *Prediction Interval* (PI) dan *Standardized Precipitation Index* (SPI) ?

1.3 Tujuan

Adapun tujuan dari penelitian ini adalah :

1. Memperoleh informasi nilai ketidakpastian dengan membangun *Prediction Interval* (PI) pada nilai ramal yang diperoleh dari model SARIMA.
2. Memperoleh nilai *Standardized Precipitation Index* (SPI) pada nilai ramalan yang diperoleh dari model SARIMA?
3. Menyajikan informasi tingkat kekeringan (bulan basah/kering) di kabupaten Jember yang diperoleh dari nilai *Prediction Interval* (PI) dan *Standardized Precipitation Index* (SPI) ?

1.4 Manfaat

Manfaat dari penelitian ini diharapkan dapat memberikan informasi yang seluas-luasnya tentang curah hujan, ketidakpastian curah hujan yang terjadi serta memberikan informasi tentang fase kekeringan di kabupaten Jember, sehingga dapat membantu pemerintah kabupaten Jember dalam upaya mengantisipasi dampak yang ditimbulkan dari fenomena alam ini.

BAB 2. TINJAUAN PUSTAKA

2.1 Peramalan

Peramalan adalah suatu kegiatan yang dapat memprediksi apa yang akan terjadi di masa depan dengan menggunakan kondisi atau data di masa lalu. Peramalan merupakan kegiatan dalam memperkirakan apa yang akan terjadi pada masa yang akan datang, atau lebih tepatnya peramalan adalah kegiatan mencoba menduga perubahan yang akan terjadi (Assauri, 1984). Penggunaan data masa lalu dari sebuah variabel untuk mengestimasi nilai dimasa depan. Peramalan merupakan alat bantu yang sangat penting dalam perencanaan yang efektif dan efisien (Subagyo, 1986).

Berikut adalah prinsip-prinsip dalam sebuah peramalan:

1. Peramalan melibatkan kesalahan (*error*).
2. Peramalan menggunakan tolak ukur kesalahan peramalan.
3. Permalan jangka pendek lebih akurat dibanding peramalan.
4. Peramalan kelompok lebih akurat dibanding peramalan individu

Adapun prosedur yang digunakan dalam peramalan adalah sebagai berikut :

1. Mendefinisikan tujuan peramalan.
2. Membuat diagram pencar (plot data).
3. Memilih jenis model peramalan yang tepat.
4. Melakukan peramalan.
5. Hitung kesalahan ramalan(*forecast error*).
6. Memilih metode peramalan dengan kesalahan yang kecil.
7. Uji validalitas keakuratan data yang akan kita lakukan peraman.

2.1.1 Metode Peramalan

Metode peramalan dapat diklasifikasikan ke dalam 2 kelompok yaitu metode kualitatif dan metode kuantitatif.

a. Metode Kualitatif

Metode ini digunakan dimana tidak ada model matematik atau lebih mendasarkan kualitatif dimasa lalu yaitu berdasarkan pemikiran yang bersifat

intuisi, pendapat dan pengetahuan serta pengalaman. Metode ini banyak digunakan dalam banyak pengambilan keputusan sehari-hari. Dalam hal ini ramalan dikatakan baik atau tidak bergantung dari banyak hal antara lain pengalaman, perkiraan dan pengetahuan yang didapat.

b. Metode Kuantitatif

Metode yang penggunaannya didasari data mentah yang disertai model matematik. Menurut (Makridakis, 1999) peramalan dengan menggunakan metode kuantitatif dapat diterapkan apabila terdapat tiga kondisi berikut yaitu: tersedia informasi tentang masa lalu, informasi tersebut dapat dikuantitatifkan dalam bentuk data numerik, dan dapat diasumsikan bahwa beberapa aspek pola masa lalu akan terus berlanjut dimasa mendatang. Metode Kuantitatif dibagi menjadi 3 macam :

- 1) Model Analisis Deret Waktu (*Time Series*).
- 2) Model Regresi.
- 3) Model Ekonometrik.

Sedangkan untuk kategori peramalan yang akan digunakan berdasar horizon waktu terbagi menjadi 3 bagian yaitu pertama Peramalan Jangka Pendek yang merupakan peramalan dengan jangka waktu 1 tahun. Kedua Peramalan Jangka Menengah yang merupakan peramalan dengan jangka waktu bulanan hingga 3 tahun. Ketiga Peramalan Jangka Panjang yang merupakan peramalan dengan jangka waktu 3 tahun atau lebih.

2.2 Analisis Deret Waktu (*Time Series*)

Menurut *Australia Bureau of Statistics*, data *time series* adalah sekumpulan data pengamatan yang diperoleh dari perhitungan dari waktu ke waktu. Pada umumnya pengumpulan dan pencatatan itu dilakukan dalam jangka waktu tertentu misalnya tiap bulan, tiap akhir tahun, sepuluh tahun dan sebagainya. Dalam teknik peramalan dengan *time series* ada 2 kategori utama yang perlu dilakukan pengujian yaitu pemulusan (*smoothing*) dan dekomposisi (*decomposition*). Dalam teknik pemulusan peramalan berdasarkan prinsip rata-rata dari kesalahan masa lalu (*Averaging Smoothing Past Errors*) dengan

menambahkan nilai ramalan sebelumnya dengan presentase kesalahan (*Percentage of The Errors*) antara nilai sebenarnya dan nilai ramalan (*forecasting value*). Pada teknik dekomposisi bergantung pada pola data yang ada (Hanke dan Wichern, 2005).

Berdasar variasi data *time series* terbagi menjadi 4 macam pola/komponen yaitu :

1. Pola Data Horizontal

Pola data horizontal terjadi saat data observasi berfluktuasi di sekitaran suatu nilai konstan atau *mean* yang membentuk garis horizontal. Data ini disebut juga dengan data stasioner. Contoh plot data horizontal adalah pada gambar a yaitu berupa plot data penjualan. Jumlah penjualan selalu meningkat atau menurun pada suatu nilai konstan secara konsisten dari waktu ke waktu.

2. Pola Data *Trend*

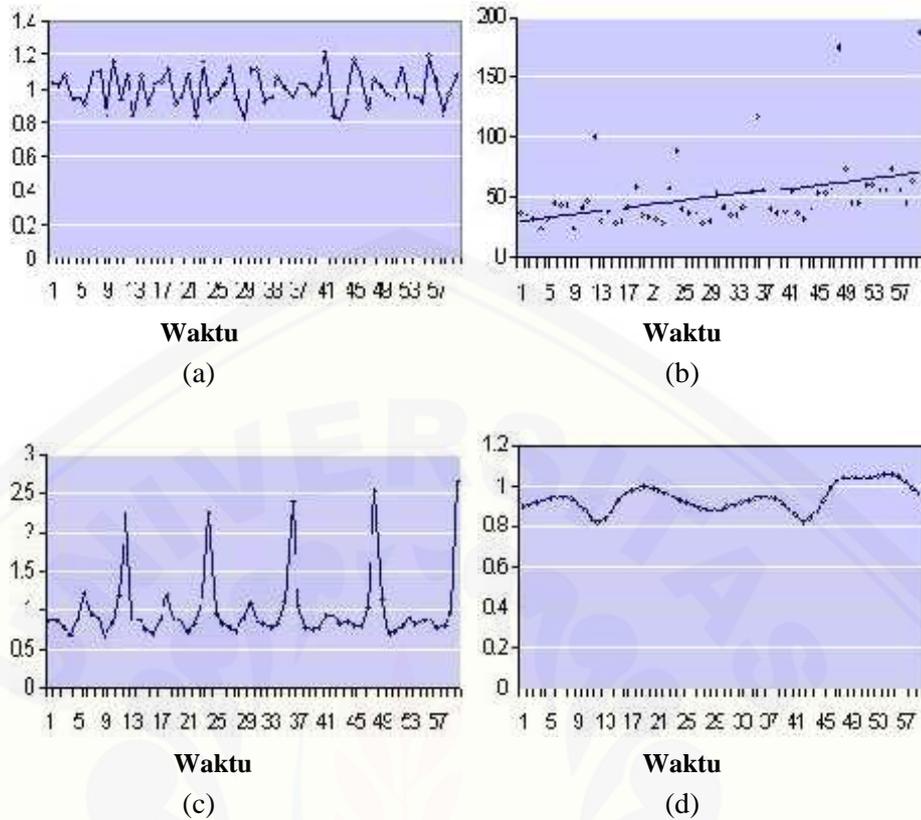
Pola data *trend* terjadi bilamana data pengamatan mengalami kenaikan atau penurunan selama periode jangka panjang. Suatu data pengamatan yang mempunyai *trend* disebut data *nonstasioner*. Plot data trend dicontohkan pada gambar b yaitu berupa data harga suatu produk yang meningkat dari tahun ke tahun.

3. Pola Data Musiman

Pola data musiman terjadi bilamana suatu deret dipengaruhi oleh faktor musiman. Pola data musiman dapat mempunyai pola musim yang berulang dari periode ke periode berikutnya. Misalnya pola yang berulang setiap bulan tertentu, tahun tertentu atau pada minggu tertentu. Contoh dari data musiman ada pada gambar c yaitu plot suplai bahan makanan tiap bulan. Dari plot tersebut terlihat bahwa terjadi pola yang berulang setiap periode dua belas bulan, sehingga bisa disimpulkan bahwa data tersebut merupakan pola data musiman.

4. Pola Data Siklis

Pola data siklis terjadi apabila deret data dipengaruhi oleh fluktuasi ekonomi jangka panjang seperti yang berhubungan dengan siklus bisnis. Gambar d di bawah ini adalah contoh plot pola data siklis. Keempat pola data *time series* tersebut disajikan pada gambar 2.1 berikut ini.



(a) Pola Horizontal; (b) Pola Trend; (c) Pola Musiman; (d) Pola Siklis

Gambar 2.1 Plot Pola Data *Time Series*

Berdasarkan Jenis Runtun waktunya, dibedakan menjadi dua yaitu:

1. Model-model linear untuk deret yang tidak statis (*non stationary series*)

Menggunakan model ARIMA (*Autoregressive Integrated Moving Average*). Model ini menggunakan teknik-teknik korelasi.

2. Model-model linear untuk deret yang statis (*stationary series*)

Menggunakan teknik penyaringan atau *filtering* untuk deret waktu, yaitu yang disebut dengan model ARMA (*Autoregressive Moving Average*) untuk suatu kumpulan data.

2.3 Kestasioneran Data

Stasioner berarti bahwa tidak terjadinya pertumbuhan dan penurunan data. Suatu data dapat dikatakan stasioner apabila pola data tersebut berada pada

kesetimbangan disekitar nilai rata-rata (*mean*) yang konstan dan variansi (*varians*) disekitar rata-rata tersebut konstan selama waktu tertentu (Makridakis, 1999). Time series dikatakan stasioner apabila tidak ada unsur *trend* dalam data dan tidak ada unsur musiman atau rata-rata dan variannya tetap.

Selain dari plot *time series*, stasioner dapat dilihat dari plot *Autocorrelation Function (ACF)* data tersebut. Apabila plot data *Autocorrelation Function (ACF)* turun mendekati nol secara cepat, pada umumnya setelah *lag* kedua atau ketiga maka dapat dikatakan stasioner (Hanke & Winchern, 2005).

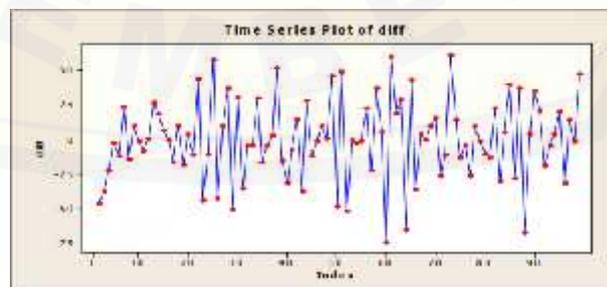
Uji akar unit untuk kestasioneran data *mean* adalah uji *Augmented Dickey-Fuller (ADF)*. Model sederhana yang digunakan dalam uji ADF adalah $\Delta X_t = \beta_1 + \delta X_{t-1} + \varepsilon_t$ dengan $\delta = \rho - 1$. Hipotesis yang diuji adalah

1. $H_0: \delta = 0$ (tidak stasioner)
 $H_0: \delta \neq 0$ (stasioner)
2. Uji statistik yang digunakan adalah

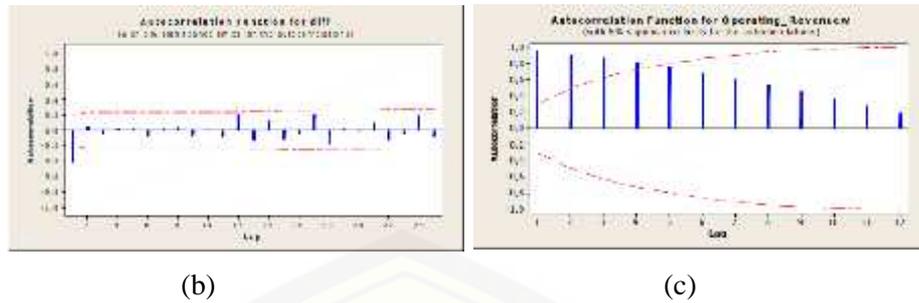
$$\tau = \frac{\rho}{SE\rho}$$

3. Kriteria keputusan H_0 ditolak jika $\tau_{ADF} > \tau$, $p\text{-value} < \alpha$

Data nonstasioner apabila terdapat unsur *trend* dalam data, yaitu mengalami kenaikan dan penurunan seiring bertambahnya periode waktu. Pada data nonstasioner yang memiliki *trend* akan memiliki nilai *Autocorrelation Function (ACF)* yang signifikan pada *lag-lag* awal kemudian mengecil secara bertahap, seperti Gambar 2.1 dibawah ini.



(a)



(a) Plot *time series* data Stasioner dalam mean dan varians; (b) Plot ACF data stasioner; (c) Plot ACF data tidak stasioner

Gambar 2.2 Plot Stationeritas

Selanjutnya stasioneritas dibagi menjadi 2 (Wei, 2006), yaitu:

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

Stasioner dalam mean adalah fluktuasi data berada di sekitar suatu nilai rata-rata yang konstan, tidak tergantung pada waktu dan variansi dari fluktuasi tersebut. Suatu proses stasioner dalam rata-rata jika $E(X_t) = \mu_t = \mu$ adalah konstan untuk setiap t . Bentuk plot data seringkali dapat diketahui bahwa data tersebut stasioner atau tidak stasioner. Apabila dilihat dari plot ACF, maka nilai-nilai utokorelasi dari data stasioner akan turun menuju nol sesudah time lag (selisih waktu) kedua atau ketiga.

b. Stasioner dalam varians

Suatu data *time series* dikatakan stasioner dalam varians apabila struktur data dari waktu ke waktu mempunyai fluktuasi data yang tetap atau konstan dan tidak berubah-ubah. Suatu proses stasioner pada varians jika $Var(X_t) = E(X_t - \mu_t)^2 = \sigma^2$ Secara visual untuk melihat hal tersebut dapat dibantu dengan menggunakan plot *time series*, yaitu dengan melihat fluktuasi data dari waktu ke waktu.

2.3.1 Pembedaan Data (*Differencing*)

Data yang tidak stasioner dalam *mean* dapat diatasi dengan proses *differencing* dilakukan untuk menstasionerkan data nonstasioner. Operator *shift*

mundur (*backward shift*) sangat tepat untuk menggambarkan proses *differencing* (Makridakis, 1999). Penggunaan *backward shift* adalah sebagai berikut

$$BX_t = X_{t-1} \quad (2.1)$$

dengan,

X_t : nilai variabel pada waktu t

X_{t-1} : nilai variabel pada waktu $t-1$

B : backward shift

Notasi B yang dipasang pada X mempunyai pengaruh menggeser data satu waktu kebelakang. Sebagai contoh, jika suatu data *time series* nonstasioner maka data tersebut dapat dibuat mendekati stasioner dengan melakukan *differencing* orde pertama dari data. Rumus untuk *differencing* orde pertama, yaitu

$$X_t' = X_t - X_{t-1} \quad (2.2)$$

dengan X_t' = nilai variabel X pada waktu t setelah *differencing*.

dengan menggunakan *backward shift*, persamaan (2.2) dapat ditulis menjadi

$$X_t' = X_t - BX_t \quad (2.3)$$

atau

$$X_t' = (1 - B)X_t \quad (2.4)$$

Sama halnya apabila perbedaan orde kedua (yaitu perbedaan pertama dari perbedaan pertama sebelumnya) harus dihitung, maka:

$$\begin{aligned} X_t'' &= X_t' - BX_{t-1}' \\ &= X_t - X_{t-1} - (X_{t-1} - X_{t-2}) \\ &= X_t - 2X_{t-1} + X_{t-2} \\ &= (1 - 2B + B^2)X_t \\ &= (1 - B)^2 X_t \end{aligned} \quad (2.5)$$

Pembedaan orde kedua diberi notasi $(1 - B)^2$, sedangkan pembeda pertama $(1 - B)$.

$$\begin{aligned} X_t^2 &= X_t - X_{t-2} \\ &= (1 - B^2)X_t \end{aligned} \quad (2.6)$$

Tujuan dari menghitung perbedaan adalah untuk mencapai stasioneritas dan secara umum apabila terdapat perbedaan orde ke-d untuk mencapai stasioneritas, ditulis sebagai berikut:

$$(1 - B^d)X_t \tag{2.7}$$

2.3.2 Transformasi Box-Cox

Data yang tidak stasioner dalam varian dapat diatasi dengan transformasi Box-Cox. Menurut (Cryer dan Chan, 2008) transformasi data digunakan untuk menstabilkan atau mendapatkan varian yang konstan. Transformasi Box-Cox yang didefinisikan oleh :

$$T(X_t) \begin{cases} X_t^\lambda & ; \lambda \neq 0 \\ \log(X_t) & ; \lambda = 0 \end{cases} \tag{2.8}$$

Berdasar rumus diatas λ adalah parameter transformasi Box-Cox dan X_t adalah nilai time series pada waktu ke-t. Berikut adalah tabel (2.1) beberapa nilai λ dengan transformasinya:

Tabel 2.1 beberapa nilai λ dan transformasinya

Nilai λ	Transformasi
-1	$\frac{1}{X_t}$
-0,5	$\frac{1}{\sqrt{X_t}}$
0	$\ln X_t$
0,5	$\sqrt{X_t}$
1	X_t

2.3.3 Fungsi Autokorelasi dan Fungsi Autokorelasi Parcial

Mengidentifikasi model yang akan digunakan untuk meramalkan adalah dengan menggunakan ACF dan PCAF. Menurut (Wei, 2006) dari proses stasioner suatu data *time series* X_t diperoleh $E(X_t) = \mu_t = \mu$ dan variansi $Var(X_t) = E(X_t - \mu_t)^2 = \sigma^2$, yang konstan dan kovarian $Cov(X_t, X_{t+k})$, yang fungsinya

pada perbedaan waktu $|t - (t + k)|$. Maka hasil dapat ditulis sebagai kovariansi antara X_t dan X_{t+k} sebagai berikut:

$$\gamma_k = \text{Cov}(X_t, X_{t+k}) = E(X_t - \mu_t)(X_{t+k} - \mu_t) \quad (2.9)$$

Korelasi antara X_t dan X_{t+k} sebagai berikut:

$$\rho_k = \frac{\text{Cov}(X_t, X_{t+k})}{\sqrt{\text{Var}(X_t)\text{Var}(X_{t+k})}} = \frac{\gamma_k}{\gamma_0} \quad (2.10)$$

Sebagai fungsi dari k , γ_k disebut fungsi autokovariansi dan ρ_k disebut fungsi autokorelasi (ACF), dalam analisis *time series* γ_k dan ρ_k menggambarkan kovarian dan korelasi antara X_t dan X_{t+k} dari proses yang sama, hanya dipisahkan oleh lag ke- k .

Fungsi autokovariansi sampel dan fungsi autokorelasi sampel dapat ditulis sebagai berikut:

$$\gamma_k = \sum_{t=1}^{n-k} (X_t - \bar{X}_t)(X_{t+k} - \bar{X}_t) \quad (2.11)$$

dan

$$r_k = \rho_k = \frac{\sum_{t=1}^{n-k} (X_t - \bar{X}_t)(X_{t+k} - \bar{X}_t)}{\sqrt{\sum_{t=1}^n (X_t - \bar{X}_t)^2}} \quad (2.12)$$

dengan,

r_k : koefisien autokorelasi

X_t : nilai variabel X pada waktu t

X_{t+k} : nilai variabel X pada waktu t+k

\bar{X}_t : nilai rata-rata variabel X_t

karena r_k merupakan fungsi atas k, maka hubungan koefisien autokorelasi dengan lagnya disebut dengan fungsi autokorelasi dan dinotasikan dengan ρ_k .

a. Autocorrelation Function (ACF)

Autokorelasi merupakan korelasi atau hubungan antar data pengamatan suatu data *time series*. Menurut (Wei, 2006) untuk mengetahui apakah koefisien autokorelasi signifikan atau tidak, perlu dilakukan uji. Pengujian dapat dilakukan hipotesis sebagai berikut:

4. $H_0: \rho_k = 0$ (koefisien autokorelasi tidak berbeda signifikan)

$H_0: \rho_k \neq 0$ (koefisien autokorelasi berbeda signifikan)

5. Uji statistik yang digunakan adalah

$$t = \frac{r_k}{SEr_k}$$

dengan,

$$SE(r_k) = \sqrt{\frac{1+2\sum_{j=1}^{k-1} r_j^2}{T}}$$

dimana,

$SE(r_k)$: standart error autokorelasi pada lag k

r_j autokorelasi pada saat lag j

k time lag

T banyaknya observasi dalam deret waktu

6. Kriteria keputusan H_0 ditolak jika $t_{hit} > t_{\frac{\alpha}{2}, n-1}$

b. *Partial Autocorrelation Function (PACF)*

Autokorelasi parsial merupakan korelasi antara X_t dan X_{t+k} dengan mengabaikan ketidakebebasan $X_{t+1}, X_{t+2}, \dots, X_{t+k-1}$. Menurut (Wei, 2006), autokorelasi parcial X_t dan X_{t+k} dapat diturunkan dari model regresi linear, dengan variabel dependent X_{t+k} dan variabel independent $X_{t+k-1}, X_{t+k-2}, \dots, X_t$ yaitu

$$X_{t+k} = \alpha_{k1}X_{t+k-1} + \alpha_{k2}X_{t+k-2} + \dots + \alpha_{kk}X_t + \epsilon_{t+k} \tag{2.13}$$

dengan α_{ki} merupakan parameter regresi ke- i untuk $i=1,2,\dots,k$ dan ϵ_{t+k} merupakan residu atau *error* dengan rata-rata nol dan tidak berkorelasi dengan X_{t+k-j} untuk $j=1,2,\dots,k$. Mengalikan X_{t+k-j} pada kedua ruas persamaan (diatas) dan menghitung nilai harapan (expected value), diperoleh

$$\begin{aligned} E(X_{t+k-j}X_{t+k}) &= \alpha_{k1}E(X_{t+k-j}X_{t+k}) + \alpha_{k2}E(X_{t+k-j}X_{t+k-1}) + \dots \\ &+ \alpha_{kk}E(X_{t+k-j}X_{t+k-2}) + E(X_{t+k-j}\epsilon_{t+k}) \end{aligned}$$

$$Y_j = \alpha_{k1}Y_{j-1} + \alpha_{k2}Y_{j-2} + \dots + \alpha_{kk}Y_{j-k} \tag{2.14}$$

dan

$$\rho_j = \alpha_{k1}\rho_{j-1} + \alpha_{k2}\rho_{j-2} + \dots + \alpha_{kk}\rho_{j-k} \tag{2.15}$$

Untuk $j=1,2,\dots,k$ diperoleh sistem persamaan berikut

$$\begin{aligned}
 \rho_1 &= k_1\rho_0 + k_2\rho_1 + \dots + k_k\rho_{k-1} \\
 \rho_2 &= k_1\rho_1 + k_2\rho_0 + \dots + k_k\rho_{k-1} \\
 &\vdots \\
 \rho_k &= k_1\rho_{k-1} + k_2\rho_{k-2} + \dots + k_k\rho_0
 \end{aligned} \tag{2.16}$$

dengan menggunakan aturan cramer, berturut-turut $k=1,2,\dots$ diperoleh

$$\begin{aligned}
 {}_{11} &= \rho_1 \\
 {}_{22} &= \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}} \\
 {}_{33} &= \frac{\begin{vmatrix} 1 & \rho_1 & \rho_1 \\ \rho_1 & 1 & \rho_2 \\ \rho_2 & \rho_1 & \rho_3 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{vmatrix}} \\
 {}_{kk} &= \frac{\begin{vmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{k-2} & \rho_1 \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{k-3} & \rho_2 \\ \vdots & \vdots & \vdots & \dots & \vdots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \dots & \rho_1 & \rho_k \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{k-2} & \rho_{k-1} \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{k-3} & \rho_{k-2} \\ \vdots & \vdots & \vdots & \dots & \vdots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \dots & \rho_1 & 1 \end{vmatrix}}
 \end{aligned} \tag{2.17}$$

Karena ${}_{kk}$ merupakan fungsi atas k , maka ${}_{kk}$ disebut fungsi autokorelasi parsial. hipotesis untuk menguji koefisien autokorelasi parsial berikut:

1. $H_0: \phi_{kk} = 0$ (koefisien autokorelasi tidak signifikan)

$H_0: \phi_{kk} \neq 0$ (koefisien autokorelasi signifikan)

2. Statistik uji yang digunakan: $t = \frac{\phi_{kk}}{SE\phi_{kk}}$ dengan

$$SE(\phi_{kk}) = \frac{1}{T}$$

3. Kriteria keputusan tolak H_0 jika $|t_{hit}| > t_{\frac{\alpha}{2}}$

2.4 Model *Autoregressive Integrated Moving Average* ARIMA

Persamaan model ARIMA menyatakan bagaimana sebuah nilai pada time series adalah linier terhadap nilai masa lalu. Peramalan dapat dihasilkan dari model ARIMA yang layak untuk data time series.

2.4.1 Model *autoregressive* (AR)

Model *Autoregressive* merupakan bentuk regresi yang menyatakan suatu ramalan sebagai fungsi nilai-nilai sebenarnya dari time series tertentu (Makridakis, 1999).

Model *Autoregressive* orde p dinotasikan dengan AR (p). Bentuk umum AR (p) adalah:

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + \varepsilon_t \quad (2.18)$$

dengan,

X_t : nilai variabel pada waktu ke- t

$X_{t-1}, X_{t-2}, \dots, X_{t-p}$: nilai masa lalu dari time series pada waktu $t-1, t-2, \dots, t-p$

ϕ_i : koefisien regresi, $i: 1, 2, 3, \dots, p$

ε_t : nilai *error* pada waktu ke- t

p : order *AR*

Persamaan (2.18) dapat ditulis menggunakan operator B (*backshift*):

$$X_t = \phi_1 B X_t + \phi_2 B^2 X_t + \dots + \phi_p B^p X_t + \varepsilon_t$$

atau

$$X_t = \phi(B) X_t \quad (2.19)$$

dimana $\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$ disebut operator (*AR*).

2.4.2 Model *Moving Average* (MA)

Menurut (Wei, 2006) Model *Moving Average* orde q dinotasikan MA (q) didefinisikan sebagai

$$X_t = \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_p \varepsilon_{t-q} + \varepsilon_t, \varepsilon_t \sim N(0, \sigma_t^2) \quad (2.20)$$

dengan,

X_t : nilai variabel pada waktu ke- t

$\varepsilon_t, \varepsilon_{t-1}, \varepsilon_{t-2}, \dots, \varepsilon_{t-q}$: nilai-nilai *error* pada waktu $t, t-1, t-2, \dots, t-q$ dan ε_t diasumsikan white noise dan normal

- θ_i : koefisien regresi, $i: 1,2,3,\dots,q$
- ε_t : nilai *error* pada waktu ke- t
- q : orde *MA*

Persamaan (2.20) dapat ditulis menggunakan operator B (*backshift*):

$$X_t = \theta(B)\varepsilon_t \tag{2.21}$$

dimana $\theta(B) = 1 + \theta_1B + \theta_2B^2 + \dots + \theta_qB^q$ disebut operator (*MA*).

2.4.3 Model *Autoregressive Moving Average* (ARMA)

Model *Autoregressive Moving Average* (ARMA) adalah hasil kombinasi dari model (AR) dan (MA) yang memiliki bentuk umum yaitu:

$$\begin{aligned} X_t &= \phi_1X_{t-1} + \phi_2X_{t-2} + \dots + \phi_pX_{t-p} + \varepsilon_t \\ &= \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2} + \dots + \theta_q\varepsilon_{t-q} + \varepsilon_t \end{aligned} \tag{2.22}$$

Persamaan (2.22) dapat ditulis menggunakan operator B (*backshift*):

$$(1 - \phi_1B - \phi_2B^2 - \dots - \phi_pB^p)X_t = (1 + \theta_1B + \theta_2B^2 + \dots + \theta_qB^q)\varepsilon_t \tag{2.23}$$

Sehingga diperoleh,

$$\theta(B)\varepsilon_t = \phi(B)X_t \tag{2.24}$$

dengan,

- X_t : nilai variabel pada waktu ke- t
- $\varepsilon_t, \varepsilon_{t-1}, \varepsilon_{t-2}, \dots, \varepsilon_{t-q}$: nilai-nilai *error* pada waktu $t, t-1, t-2, \dots, t-q$ dan ε_t diasumsikan white noise dan normal
- $X_{t-1}, X_{t-2}, \dots, X_{t-p}$ nilai masa lalu dari time series pada waktu $t-1, t-2, \dots, t-p$
- ϕ_i : koefisien regresi, $i: 1,2,3,\dots,p$
- θ_i : koefisien regresi, $i: 1,2,3,\dots,q$
- ε_t : nilai *error* pada waktu ke- t
- q : orde *MA*
- p : order *AR*

2.4.4 ARIMA (*Autoregressive Integrated Moving Average*)

Model ARIMA telah dipelajari secara mendalam oleh George Box dan Gwilym Jenkins (1976). Model ARIMA merupakan penggabungan antara model

Autoregressive (AR) dan *Moving Average* (MA) serta proses *differencing*. Secara umum, model ARIMA (p, d, q) dapat ditulis sebagai berikut:

$$\phi_p(B)(1 - B)^d X_t = \theta_q(B)\varepsilon_t \quad (2.25)$$

dimana,

$\phi_p(B)$ komponen AR non musiman dengan orde (p)

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

X_t besarnya pengamatan (kejadian) pada waktu ke t

ε_t nilai *error* pada saat t

$\theta_q(B)$ komponen MA non musiman dengan orde (q)

2.4.5 SARIMA (*Seasonal Autoregressive Integrated Moving Average*)

SARIMA atau ARIMA musiman adalah *time series* yang mempunyai sifat “berulang” setelah sekian periode waktu tertentu, misalnya satu tahun, satu bulan, triwulan dan seterusnya. Model ARIMA Musiman merupakan model ARIMA yang digunakan untuk menyelesaikan *time series* musiman yang terdiri dari dua bagian, yaitu bagian tidak musiman (non-musiman) dan bagian musiman. Bagian non-musiman dari metode ini adalah model ARIMA.

Secara umum bentuk model SARIMA (p, d, q)(P, D, Q)^s adalah:

$$\Phi_P(B^s)\phi_p(B)(1 - B)^d(1 - B)^D X_t = \theta_q(B)\Theta_Q(B^s)\varepsilon_t \quad (2.26)$$

dimana,

(p, d, q) : orde AR (p), orde *differencing* (d), orde MA (q) untuk data non musiman.

$\phi_p(B)$ komponen AR non musiman dengan orde (p)

$\Phi_P(B)$: komponen AR musiman dengan orde (P)

$\theta_q(B)$: komponen MA non musiman dengan orde (q)

$\Theta_Q(B)$: komponen MA musiman dengan orde (Q)

$(1 - B)^d$: *differencing* non musiman dengan orde (d)

$(1 - B)^D$: *differencing* musiman dengan orde (D)

X_t besarnya pengamatan (kejadian) pada waktu ke t

ε_t nilai residual pada saat t

s jumlah periode musiman

2.5 Estimasi Parameter

Setelah model dugaan dari data *time series* diperoleh, langkah selanjutnya ialah mengestimasi nilai parameter – parameter pada model tersebut. Salah satu metode estimasi yang dapat dipergunakan ialah *Maximum Likelihood Estimation* (MLE). Asumsi yang diperlukan pada metode MLE adalah error ε_t (nilai galat waktu t) berdistribusi normal (Box dan Jenkins, 1976). Sehingga fungsi kepadatan peluangnya adalah

$$f(\varepsilon_t | \sigma_\varepsilon^2) = (2\pi\sigma_\varepsilon^2)^{-\frac{1}{2}} \exp -\frac{\varepsilon_t^2}{2\sigma_\varepsilon^2} \tag{2.27}$$

dengan mengingat error bersifat independen, maka distribusi bersama dari $\varepsilon_1, \varepsilon_2, \varepsilon_3, \dots, \varepsilon_n$ adalah

$$f(\varepsilon_1, \varepsilon_2, \varepsilon_3, \dots, \varepsilon_n | \sigma_\varepsilon^2) = (2\pi\sigma_\varepsilon^2)^{-\frac{n}{2}} \exp\left(-\frac{1}{2\sigma_\varepsilon^2} \sum_{i=1}^n \varepsilon_i^2\right) \tag{2.28}$$

error ε_t dapat dinyatakan dalam fungsi X_t dan parameter-parameter $\phi, \theta, \sigma_\varepsilon^2$ serta error-error sebelumnya. Secara umum bentuk ε_t adalah

$$\varepsilon_t = X_t - \phi_1 X_{t-1} - \dots - \phi_p X_{t-p} + \theta_1 \varepsilon_{t-1} + \dots + \theta_p \varepsilon_{t-p} \tag{2.29}$$

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

$$L(\phi, \theta, \sigma_\varepsilon^2 | X) = (2\pi\sigma_\varepsilon^2)^{-\frac{n}{2}} \exp\left(-\frac{1}{2\sigma_\varepsilon^2} S(\phi, \theta)\right) \tag{2.30}$$

dengan $S(\phi, \theta) = \sum_{t=1}^n (X_t - \phi_1 X_{t-1} - \dots - \phi_p X_{t-p} + \theta_1 \varepsilon_{t-1} + \dots + \theta_p \varepsilon_{t-p})^2$

maka dari persamaan (2.30) diperoleh fungsi *likelihood*nya adalah

$$l = \ln L(\phi, \theta, \sigma_\varepsilon^2 | X) = -\frac{n}{2} \log 2\pi - \frac{n}{2} \log -\frac{1}{2\sigma_\varepsilon^2} S(\phi, \theta) \tag{2.31}$$

Maksimum fungsi *log likelihood* dihitung dengan mencari turunan pertama persamaan (2.31) terhadap masing-masing parameter dan menetapkan turunan parsial yang dihasilkan sama dengan nol.

$$\frac{\partial l(\phi, \theta, \sigma_\varepsilon^2 | X)}{\partial \phi} = 0, \frac{\partial l(\phi, \theta, \sigma_\varepsilon^2 | X)}{\partial \theta} = 0, \frac{\partial l(\phi, \theta, \sigma_\varepsilon^2 | X)}{\partial \sigma_\varepsilon^2} = 0$$

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

$$l_{ij} = \frac{\partial^2 l(\theta, \sigma^2 | x)}{\partial \beta_i \partial \beta_j} \text{ dan } I(\beta) = -E(l_{ij}) \quad (2.32)$$

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

$$V(\hat{\beta}) = [I(\beta)]^{-1} \text{ dan } SE(\hat{\beta}) = [V(\hat{\beta})]^{1/2} \quad (2.33)$$

Parameter yang diperoleh diuji signifikasinya. Pengujian dapat dilakukan hipotesis sebagai berikut:

1. $H_0: \hat{\beta} = 0$ (parameter tidak signifikan pada model)
 $H_0: \hat{\beta} \neq 0$ (parameter signifikan pada model)
2. Uji statistik yang digunakan adalah

$$|t_{hitung}| = \frac{\hat{\beta}}{SE(\hat{\beta})}$$

dimana,

$SE(\hat{\beta})$: standar error dari setiap parameter dugaan

$\hat{\beta}$ parameter dugaan

3. Kriteria keputusan H_0 ditolak jika $t_{hitung} > t_{\frac{\alpha}{2}, n-1}$ atau $p\text{-value} < \alpha$

2.6 Uji White Noise

Menurut (Wei, 2006), Suatu proses $\{\varepsilon_t\}$ disebut proses *white noise* jika *series*-nya terdiri dari variabel random yang tidak berkorelasi dan berdistribusi normal dengan rata-rata konstan $E(\varepsilon_t) = 0$, variansi konstan $Var(\varepsilon_t) = \sigma^2$ dan $Cov(\varepsilon_t, \varepsilon_{t+k}) = 0$ untuk $k \neq 0$. Dengan demikian proses *white noise* stasioner dengan fungsi autokovariansi

$$\gamma_k = \begin{cases} \sigma^2, & \text{jika } k = 0 \\ 0, & \text{jika } k \neq 0 \end{cases}$$

Fungsi autokorelasi

$$\rho_k = \begin{cases} 1, & \text{jika } k = 0 \\ 0, & \text{jika } k \neq 0 \end{cases}$$

Fungsi autokorelasi parsial

$$\alpha_k = \begin{cases} 1, & \text{jika } k = 0 \\ 0, & \text{jika } k \neq 0 \end{cases}$$

Pemeriksaan yang bersifat *white noise* dengan menggunakan uji *Ljung and Box (LB)* adalah sebagai berikut:

1. Hipotesis

H_0 : $\rho_k = 0$ sisaan *white noise*

H_1 : $\rho_k \neq 0$ sisaan tidak *white noise*

2. Taraf signifikansi $\alpha = 0,05$

3. Statistik uji

$$LB = n(n+2) \sum_{k=1}^n \frac{\widehat{\rho}_k^2}{n-k}$$

dengan,

n : banyaknya pengamatan

k : banyaknya lag

$\widehat{\rho}_k$: estimator korelasi pada lag ke k

4. Daerah Kritis (DK) : $\{LB \mid LB > X^2_{1-\alpha;k}\}$

5. Keputusan uji : H_0 di terima jika $LB \in DK$

2.7 Pemilihan Model Terbaik

Sedangkan untuk kriteria pemilihan model terbaik didasarkan pada nilai residual untuk nilai ramalan data *out sampel*. Penentuan model dengan menggunakan kriteria pemilihan model terbaik melalui pendekatan AIC. Semakin kecil nilai kriteria pemilihan model maka semakin besar kemungkinan suatu model tersebut dipilih karena tingkat akurasinya semakin tinggi. Persamaan AIC sebagai berikut:

$$AIC(M) = n \ln(\hat{\sigma}_a^2 + 2M) \quad (2.34)$$

Dengan n banyaknya pengamatan, M banyaknya parameter dalam model, serta $\hat{\sigma}_a^2$ merupakan estimasi varians residual. kriteria *RMSE (Root Mean Squared Error)* dengan n periode peramalan dirumuskan sebagai berikut.

$$RMSE = \sqrt{\frac{1}{M} \sum_{i=1}^M (X_i - \hat{X}_i)^2} \quad (2.35)$$

dimana M merupakan banyaknya peramalan yang dilakukan X_i adalah data yang sebenarnya dan \hat{X}_i adalah data hasil ramalan. Nilai *RMSE* berkisar antara 0 sampai . Semakin kecil nilai *RMSE* maka model yang digunakan semakin bagus (Wei, 2006).

2.8 Menghitung *Prediction Interval*

Prediction interval adalah suatu interval yang mencakup probabilitas tertentu, mencakup observasi masa depan dari populasi. Serangkaian waktu yang diamati mengandung t observasi yang dilambangkan (X_1, X_2, \dots, X_t) . misalkan kita ingin meramalkan nilai dari k langkah kedepan $(t+k)$. Titik peramalan pada waktu $(t+k)$ dibuat menggunakan data waktu n dilambangkan dengan $\hat{X}_t(k)$. Oleh karena itu penting untuk menentukan waktu perkiraan. Ketika nilai yang diamati kemudian menjadi tersedia, kita dapat menghitung kesalahan peramalan yang sesuai yang dilambangkan oleh $e_t(k)$ yang diperoleh dari

$$e_t(k) = X_{t+k} - \hat{X}_t(k) \quad (2.36)$$

Notasi untuk kesalahan peramalan, sama dalam titik peramalan untuk menentukan periode waktu kapan peramalan dibuat.

Menurut (Chatfield, 1993) dalam menghitung nilai *Prediction interval* bersumber dari bentuk persamaan umum $100(1 - \alpha)\%$ untuk nilai k langkah ke depan adalah sebagai berikut:

$$\hat{X}_t(k) \pm Z_{\frac{\alpha}{2}} \sqrt{Var(e_t(k))} \quad (2.37)$$

dimana formula yang tepat untuk $\hat{X}_t(k)$ dan untuk $Var(e_t(k))$ untuk model yang dianggap tepat dan menunjukkan titik presentase terdistribusi normal.

Interval simetris $\hat{X}_t(k)$ menghasilkan nilai asumsi perkiraan titik yang tak bias. Statistik biasa yang dapat menghitung nilai ketidakpastian dalam perkiraan

suatu variabel tunggal adalah *mean square prediction error* (PMSE) yaitu $E(e_t(k)^2)$. Skala statistik independen seperti mean absolute prediction error (MAPE) digunakan membandingkan keakuratan perkiraan untuk variabel yang berbeda. Untuk perkiraan tidak bias kita menggunakan PMSE dimana $E(e_t(k)^2) = Var(e_t(k))$ dengan menghitung persamaan (2.37) yang menggunakan $Var(e_t(k))$ sebagai PMSE untuk mendapat nilai *Prediction interval*.

2.9 Metode SPI (*Standardized Precipitation Index*)

Metode SPI dikembangkan (McKee, 1993) digunakan untuk menghitung indeks kekeringan. Metode ini mengukur kekurangan curah hujan pada berbagai periode berdasarkan kondisi normal. Perhitungan nilai SPI berdasarkan jumlah sebaran gamma yang didefinisikan sebagai fungsi frekuensi atau peluang kejadian sebagai berikut:

$$G(x) = \int_0^x g(x) dx = \frac{1}{\beta^\alpha \Gamma(\alpha)} \int_0^x t^{\alpha-1} e^{-t/\beta} dt \quad (2.38)$$

Nilai α dan β di estimasi untuk setiap stasiun hujan dengan menggunakan rumus sebagai berikut:

$$\alpha = \frac{1}{4A} \left(1 + \sqrt{1 + \frac{4A}{3}} \right) \quad (2.39)$$

$$A = \ln(\bar{x}) - \frac{\ln(x)}{n} \quad (2.40)$$

atau

$$\alpha = \frac{\bar{x}^2}{s^2} \quad (2.41)$$

$$\beta = \frac{\bar{x}}{\alpha} \quad (2.42)$$

untuk $x = 0$ maka nilai $G(x)$ menjadi:

$$H(x) = q + (1 - q) \cdot G(x) \quad (2.43)$$

dimana q = jumlah kejadian hujan (0 (m)) / jumlah data (n)

Nilai SPI merupakan transformasi dari distribusi gamma($G(x)$) menjadi standar normal dengan rata-rata (*mean*) 0 dan perbedaan 1, atau menggunakan rumusan dibawah ini:

Perhitungan Spi untuk $0 < H(x) \leq 0,5$

$$Z=SPI=-\left(t - \frac{C_0+C_1t+C_2t^2}{1+d_1t+d_2t^2+d_3t^3}\right) \tag{2.43}$$

$$t = \sqrt{\ln \frac{1}{H(x)^2}} \tag{2.44}$$

dimana,

$$C_0 = 2.515517 \qquad d_1 = 1.432788$$

$$C_1 = 0.802853 \qquad d_2 = 0.189269$$

$$C_2 = 0.010328 \qquad d_3 = 0.001308$$

Kriteria nilai indeks kekeringan metode SPI diklasifikasikan dalam tabel (2.2) dibawah ini:

Tabel 2.2 Klasifikasi Kekeringan

Klasifikasi	Nilai
Amat sangat basah	> 2,00
Sangat basah	1,50 - 1,99
Cukup Basah	1,00 - 1,49
Mendekati Normal	(- 0,99) - 0,99
Cukup Kering	(- 1,00) - (- 1,49)
Sangat kering	(- 1,50) - (- 1,99)
Amat sangat kering	< (- 2,00)

Sumber: Mc Kee, 1993

DAFTAR PUSTAKA

- Assauri, Sofyan. 1984. *Teknik dan Metode Peramalan*. Jakarta : Penerbit Fakultas Ekonomi Universitas Indonesia.
- Box, G.E.P. and G.M. Jenkins, 1976. *Time Series Analysis: Forecasting and Control*. 1st Edn., Holden-Day, San Fransisco.
- Chatfield, C. 1993. *Calculating Interval Forecasts*. *Journal of business and economic statistics*, April 1993, Vol. 11, No. 2. American Statistical Association.
- Cryer, J.D., dan Chan, KS. (2008). *Time Series Analysis: With Apllication in R: Second Edition*. USA: Spinger Science dan Busiiness Media, LLC.
- Hanke, J.E. & Wichern, D.W. 2005. *Business Forecasting Eight Edition*. New Jersey: Pearson Prentice Hall.
- Kurniawan, R. Kombinasi *Autoregressive Integrated Moving Average (ARIMA)* dan *Standardized Precipitation Index (SPI)* untuk Penentuan Indeks Kekeringan dengan Teknologi *Android*. Universitas Kristen Satya Wacana.
- Lusiani, A. 2011. *Pemodelan Autoregressive Integrated Moving Average (ARIMA) Curah Hujan di kota Bandung*. Sigma-Mu Vol.3 No.2 – September 2011. Politeknik Negeri Bandung.
- Makridakis, S., Wheelwright, S.C., & McGee, V.E. 1999. *Metode dan Aplikasi Peramalan Jilid 1* (Ir. Untung Sus Ardiyanto, M.Sc. & Ir. Abdul Basith, M.Sc. Terjemahan). Edisi Kedua. Jakarta: Penerbit Erlangga.
- Mckee, T.B., Doesken, N.J. & Kleist, J.1993. *The Relationship of Drought Frequency And Duration To Time Scales*. Colorado: Department of Atmospheric Science.
- Mujtahiddin, M. I. 2014. *Analisis Spasial Indeks Kekeringan Kabupaten Indramayu*. *Jurnal Meteorologi dan Geofisika* VOL. 15 No. 2 Tahun 2014: 99-107. Bandung.
- Retnaningrum. 2015. Penerapan model STAR dan ARIMA untuk peramalan Curah Hujan di kabupaten Jember. *UNEJ JURNAL XXXXXXXXXX* 2015, 1(1):1-5. Universitas Jember.
- Sarle, W. 1994. Neural network and Statistical Models. *In Proceeding 19th A SAS Users Group Int. Conf.*, pp. 1538-1550. Cary: SAS Institute.

Subagyo, Pangestu. 1986. *Forecasting Konsep dan aplikasi*. Yogyakarta: BPPE UGM.

Ukhra, A. 2014. pemodelan dan peramalan data deret waktu dengan metode *Seasonal ARIMA*. Jurnal Matematika UNAND Vol. 3 No. 3 Hal. 59 – 67. Universitas UNAND.

Yar, M and Chatfield, C. 1990. *Prediction Interval of the Holt-Winters forecasting procedure*. *Internasional Journal of Forecasting* 6 (1990) 127-137. North-Holland.

Wei, W.S. 2006. *Time Series Analysis, Univariate and Multivariate Method Second Edition*. New York : Pearson Education.

