

II. TINJAUAN PUSTAKA

2.1 Curah Hujan

Hujan adalah proses kondensasi uap air di atmosfer menjadi butir air yang cukup berat untuk jatuh dan tiba didarat. Dua proses yang mungkin terjadi bersamaan dapat mendorong udara menjadi semakin jenuh menjelang hujan, yaitu pendinginan udara atau penambahan uap air ke udara. Butiran hujan mempunyai ukuran yang berbeda-beda mulai dari ukuran besar, hingga butiran yang kecil (Afrila, 2018).

Curah hujan atau yang sering disebut presipitasi adalah jumlah air hujan yang turun di daerah dalam satuan waktu tertentu. Jumlah curah hujan merupakan volume air yang terkumpul di permukaan bidang datar dalam suatu periode tertentu (harian, mingguan, bulanan, atau tahunan). Curah hujan merupakan salah satu unsur cuaca yang unsur datanya diperoleh dengan cara mengukurnya menggunakan alat penakar hujan sehingga dapat diketahui jumlahnya dalam satuan milimeter (mm). Curah hujan dibatasi sebagai tinggi air hujan yang diterima di permukaan sebelum mengalami aliran permukaan, evaporasi dan peresapan ke dalam tanah (Chandra dan Suprpto, 2016).

2.2 Suhu Udara

Suhu udara adalah ukuran energi kinetik rata-rata dari pergerakan molekul-molekul. Suhu suatu benda ialah keadaan yang menentukan kemampuan benda tersebut untuk memindahkan panas ke benda-benda lain atau menerima panas dari benda-benda lain. Dalam sistem dua benda, benda yang kehilangan panas dikatakan benda yang bersuhu lebih tinggi. Suhu dapat didefinisikan secara mikroskopik berkaitan dengan gerakan molekul sedemikian rupa sehingga semakin besar kecepatan molekul maka semakin tinggi suhunya. Secara mikroskopik suhu suatu benda dapat didefinisikan sebagai tingkat atau derajat kepanasan benda tersebut. Pengukuran suhu udara pada beberapa negara dalam meteorologi dinyatakan dengan satuan derajat *Celsius* yang lambangnya °C. Sedangkan untuk keperluan *meteorologist* satuan derajat *Fahrenheit* dengan lambang °F masih tetap digunakan. Secara resmi telah disepakati untuk pelaporan internasional digunakan skala *Celsius*. Skala suhu °C dan °F masing-masing didefinisikan dengan menggunakan skala suhu Kelvin yang merupakan skala suhu dasar ilmu pengetahuan (Seojitno, 1973).

Menurut Ance (1986), Suhu dikatakan sebagai derajat panas atau dingin yang diukur berdasarkan skala tertentu dengan menggunakan thermometer. Suhu udara permukaan bumi merupakan salah satu unsur penting yang diamati oleh pengamat cuaca (*Meteorological Station maupun Climatological Station*). Pengukuran suhu udara dilakukan setiap hari selama 24 jam dicatat pada tiap

jamnya. Hasil pengukuran ini digunakan untuk menentukan suhu rata-rata harian. Suhu rata-rata harian digunakan untuk menentukan suhu rata-rata bulanan dan suhu rata-rata bulanan digunakan untuk menentukan suhu rata-rata tahunan. Ketinggian suatu tempat yang ada dipermukaan bumi berpengaruh terhadap tekanan udara dan suhu udara. Semakin tinggi tempat di permukaan bumi, semakin rendah suhu udaranya (Purwantara, 2015).

2.3 Model ARIMA

ARIMA adalah gabungan model $AR(p)$ dan $MA(q)$ melalui proses *differencing*, ARIMA memiliki kelambanan waktu (penentuan orde). Kelambanan waktu 1 periode pada proses *autoregressive* disebut *autoregressive* orde pertama atau disingkat $AR(1)$ dan simbol untuk menyatakan banyaknya kelambanan waktu pada proses *autoregressive* adalah p . Kelambanan waktu 1 periode pada proses *moving average* disebut *moving average* orde pertama atau disingkat $MA(1)$. Simbol untuk banyaknya kelambanan waktu pada proses *moving average* adalah q . Nilai p dan q dapat lebih dari 1. Proses *differencing* dapat dilakukan sekali atau dapat dilakukan lebih dari sekali sampai data bersifat stasioner. Biasanya proses *differencing* ini tidak lebih dari 2 kali. Simbol proses diferensi data adalah d .

Model ARIMA terdiri dari dua aspek yaitu *autoregressive* dan *moving average*. Penulisan model ARIMA untuk $AR(p)$, $MA(q)$ dan *differencing* sebanyak d kali ditulis dengan ARIMA (p, d, q) . Misalnya dalam suatu proses ARIMA menggunakan *autoregressive* orde pertama. *Moving average* orde pertama dan dilakukan *differencing* sekali untuk memperoleh data yang stasioner, maka penulisannya adalah ARIMA $(1, 1, 1)$. Proses ARIMA yang tidak mempunyai bagian MA ditulis sebagai $ARI(p, d)$ atau ARIMA $(p, d, 0)$. Proses ARIMA yang tidak mempunyai bagian AR ditulis sebagai $IMA(d, q)$ atau ARIMA $(0, d, q)$.

Bentuk umum model ARIMA (p, d, q) adalah:

$$(1 - \phi_1 B - \dots - \phi_p B^p)(1 - B)^d Z_t = (1 - \theta_1 B - \dots - \theta_q B^q) a_t \quad (1)$$

dengan Z_t adalah data (observasi). a_t adalah residual model ARIMA, ϕ_p adalah parameter model *autoregressive* ke- p , θ_q adalah parameter model *moving average* ke- q (Suseno, 2017).

2.3.1 Pengujian Stasioneritas

Sebelum melakukan analisis lanjutan terhadap data deret waktu, hal yang penting dilakukan adalah mengidentifikasi karakteristik data. Penetapan karakteristik seperti stasioner, musiman, non-musiman dan memerlukan suatu pendekatan yang sistematis dan ini akan menolong untuk mendapatkan gambaran yang jelas mengenai model-model yang akan digunakan (Aswi dan Sukarna, 2006)

Hal yang perlu diperhatikan adalah bahwa kebanyakan deret berkala bersifat nonstasioner dan bahwa aspek-aspek AR dan MA dari model ARIMA hanya berkenaan dengan deret berkala yang stasioner. Stasioneritas berarti tidak terdapat pertumbuhan atau penurunan pada data. Data secara kasarnya harus horizontal sepanjang sumbu waktu. Dengan kata lain, fluktuasi data berada di sekitar suatu nilai rata-rata yang konstan, tidak tergantung pada waktu dan variansi dari fluktuasi tersebut pada pokoknya tetap konstan setiap waktu.

Jenis deret waktu yang sangat penting adalah deret waktu stasioner. Rangkaian deret waktu dikatakan stasioner jika sifat-sifatnya tidak terpengaruh oleh perubahan asal waktu. Artinya, jika distribusi probabilitas gabungan dari observasi $y_t, y_{t+1}, \dots, y_{t+n}$ sama persis dengan distribusi probabilitas gabungan dari pengamatan $y_{t+k}, y_{t+k+1}, \dots, y_{t+k+n}$ maka deret waktu tersebut benar-benar stasioner (*strictly stationary*). Ketika $n = 0$ asumsi stasioneritas berarti distribusi probabilitas y_t adalah sama untuk semua periode waktu dan dapat ditulis sebagai $f(y)$.

Stasioner menyiratkan jenis keseimbangan atau stabilitas statistik dalam data. Akibatnya, deret waktu mempunyai mean konstan yang biasanya dimodelkan sebagai berikut:

$$\mu_y = E(y) = \int_{-\infty}^{\infty} yf(y)dy \quad (2)$$

(Montgomery, 2008)

a. Stasioner dalam variansi

Ketika data tidak stasioner dalam variansi, dilakukan transformasi untuk menstabilkan variansi atau membuat variansi menjadi homogen. Salah satu transformasi yang bisa digunakan adalah *Power Transformation*. Secara umum, *Power Transformation* diberikan sebagai berikut.

$$T(Z_t) = \begin{cases} \frac{Z_t^\lambda - 1}{\lambda} & ; \lambda \neq 0 \\ \ln Z_t & ; \lambda = 0 \end{cases} \quad (3)$$

Jika λ yang bernilai nol, maka dilakukan Transformasi Logaritma sebagai berikut.

$$\lim_{\lambda \rightarrow 0} T(Z_t) = \lim_{\lambda \rightarrow 0} \frac{Z_t^\lambda - 1}{\lambda} = \ln(Z_t) \quad (4)$$

Dengan Z_t adalah pengamatan pada waktu ke- t dan λ adalah parameter yang digunakan untuk transformasi. Metode ini diperkenalkan oleh Box dan Cox (1964) di dalam Wei (2006). Berikut adalah nilai λ yang sering digunakan dan bentuk transformasinya.

Tabel 1. Tabel Transformasi

Nilai λ	Transformasi
-1,0	$\frac{1}{Z_t}$
-0.5	$\frac{1}{\sqrt{Z_t}}$
0,0	$\ln Z_t$
0,5	$\sqrt{Z_t}$
1.0	Z_t (tidak perlu ditransformasi)

Data dikatakan stasioner dalam variansi apabila batas bawah dan batas atas dari transformasi *Box-Cox* memuat nilai $\lambda = 1$, sebaliknya dikatakan tidak stasioner dalam variansi apabila batas bawah dan batas atas tidak memuat angka 1. Stasioneritas data runtun waktu harus terpenuhi karena digunakan untuk membuat kesimpulan statistik tentang struktur proses stokastik berdasarkan catatan yang diamati dari proses tersebut (Cryer & Chan, 2008).

b. Stasioner dalam rata-rata

Seringkali *time series* tidak stasioner dalam rata-rata akibat adanya pola seiring dengan perubahan waktu. Suatu deret yang nonstasioner dapat direduksi menjadi deret yang stasioner melalui *differencing* menggunakan operator *shift* mundur (*backward shift*) yang dinotasikan ∇^d atau dapat ditulis $(1 - B)^d$ dengan :

$$\nabla Z_t = Z_t - Z_{t-1} \quad (5)$$

Penulisan lain untuk *differencing* disebut operator *backsift* yang didefinisikan dengan $B^i Z_t = Z_{t-i}$ sehingga menjadi:

$$\begin{aligned} (1 - B)Z_t &= \nabla Z_t \\ Z_t - B Z_t &= (1 - B)Z_t \end{aligned} \quad (6)$$

dengan $\nabla = (1 - B)$. Jika *differencing* pertama tidak menghasilkan *time series* yang stasioner maka dapat dilakukan *differencing* kedua yaitu:

$$y_t = \nabla^2 Z_t = \nabla(\nabla Z_t) = (1 - B)^2 Z_t = (1 - 2B + B^2)Z_t = Z_t - 2Z_{t-1} + Z_{t-2} \quad (7)$$

(Montgomery, 2008)

Suatu proses stasioner dalam rata-rata jika $E(Z_t) = \mu_t = \mu$ adalah konstan untuk setiap t . Proses pemeriksaan kestasioneran ini dapat digunakan diagram deret waktu (*time series plot*) yaitu diagram pencar antara nilai peubah Z_t dengan waktu t . Demi mengetahui apakah data mengandung unit root atau tidak, maka perlu digunakan uji *unit root*. Salah satu dari uji *unit root* ini yang digunakan adalah *Augmented Dickey Fuller* (*ADF-test*) dimana filosofi dari uji ADF ini adalah dengan mengikuti proses

autoregressive orde pertama atau AR(1). Uji ADF mempunyai persamaan sebagai berikut:

$$\Delta Z_t = \beta_1 + \beta_2 + \delta Z_{t-1} + \phi_i \sum_{i=1}^k \Delta Z_{t-1} + \varepsilon_t \quad (8)$$

dengan :

- ΔZ_t : *first different* dari Z
- β_1 : nilai konstan atau *intercept*
- β_2 : koefisien untuk trend
- δ : koefisien untuk lag Z
- ϕ : koefisien untuk *difference* lag Z
- ε : error
- k : lag
- t : periode waktu $t = \{1, 2, \dots, n\}$

a) Hipotesis :

H_0 : terdapat unit *root*, data stasioner

H_1 : tidak terdapat unit *root*, data tidak stasioner

b) Statistik Uji :

$$T_{hitung} = \frac{\hat{\delta}}{sd(\hat{\delta})} \quad (9)$$

c) Kriteria Pengujian:

Jika nilai $|T_{hitung}| > |T_{(a,n-1)}|$ (dengan $a = 0,05$), maka H_0 ditolak yang berarti data sudah stasioner terhadap rata-rata.

2.3.2 Identifikasi Model

Identifikasi model ARIMA yang sesuai yaitu merubah data, apabila diperlukan dapat menjadi bentuk yang stasioner dan memutuskan model sementara dengan menganalisis ACF (*Autocorrelation Function*) dan PACF (*Partial Autocorrelation Function*).

Menurut Makridakis dkk (1999), data stasioner adalah data yang tidak memiliki pertumbuhan atau penurunan data, artinya fluktuasi data berada disekitar suatu nilai rata-rata konstan, tidak tergantung pada waktu dan variansi dari fluktuasi tersebut. Hal ini dapat dilihat melalui nilai autokorelasi (plot ACF) atau dengan kata lain suatu kumpulan nilai yang tidak mempunyai hubungan dengan nilai lainnya. Nilai terdahulu tidak dapat memprediksi keluarnya hubungan dengan nilai lainnya. Nilai terdahulu tidak dapat memprediksi keluarnya nilai yang berikutnya. Hal ini menyampaikan bahwa data stasioner adalah data acak.

Menurut Wei (2006) ACF (*Autocorrelation Function*) adalah korelasi diantara variabel itu sendiri dengan selang satu atau beberapa periode

kebelakang. Autokorelasi digunakan pada data *time series* untuk mengukur bagaimana nilai saling berhubungan dengan nilai masa depan ($Z_{t+1}, Z_{t+2}, \dots, Z_{t+k}$) atau sama untuk nilai masa lalu ($Z_{t-1}, Z_{t-2}, Z_{t-k}$). Bentuk autokorelasi pada *time series* dapat digunakan untuk mengidentifikasi model ARIMA. Proses stasioner Z_t mempunyai mean $E(Z_t) = \mu$ dan varians $Var(Z_t) = E(Z_t - \mu)^2$, yang konstan dan kovarians $Cov(Z_t, Z_s)$, yaitu hanya berfungsi selisih waktu $|t - s|$. Oleh karena itu, dalam kasus ini, kovarians antara Z_t dan Z_{t-k} sebagai berikut :

$$\gamma_k = Cov(Z_t, Z_{t+k}) = E(Z_t - \mu)(Z_{t+k} - \mu), \quad (10)$$

kemudian korelasi antara Z_t dan Z_{t+k} adalah

$$\rho_k = \frac{Cov(Z_t, Z_{t+k})}{\sqrt{Var(Z_t)}\sqrt{Var(Z_{t+k})}} = \frac{\gamma_k}{\gamma_0} \quad (11)$$

PACF (Partial Autocorrelation Function) adalah suatu ukuran dari korelasi dua variabel *time series* stasioner setelah pengaruh dari variabel lainnya dihilangkan. Penentuan nilai PACF dapat menggunakan Cramer's rules untuk $k = 1, 2, \dots$, sehingga didapatkan persamaan umum PACF

$$\begin{aligned} \phi_{11} &= \rho_1 \\ \phi_{22} &= \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}} \\ \phi_{33} &= \frac{\begin{vmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_2 \\ \rho_2 & \rho_1 & \rho_3 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{vmatrix}} \\ \phi_{kk} &= \frac{\begin{vmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{k-2} & \rho_1 \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{k-3} & \rho_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \dots & \rho_1 & \rho_k \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{k-2} & \rho_1 \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{k-3} & \rho_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \dots & \rho_1 & 1 \end{vmatrix}} \quad (12) \end{aligned}$$

Data yang tidak stasioner dalam variansi biasanya akan ditransformasi dengan transformasi box-cox dan data yang tidak stasioner dalam rata-rata akan dilakukan *differencing* (menghitung perubahan atau selisih nilai observasi), sehingga menjadi data yang stasioner. Nilai selisih yang diperoleh dicek kembali apakah sudah stasioner atau belum. Jika belum stasioner maka dilakukan *differencing* kembali (Wei, 2006).

2.3.3 Estimasi Parameter

Metode pendugaan parameter dapat dilakukan dengan metode *maximum likelihood estimation*. Dalam ARIMA biasanya digunakan untuk memperkirakan parameter dalam model sementara yang diidentifikasi. Sebagian besar model

ARIMA adalah model *nonlinier* dan memerlukan penggunaan prosedur model *nonlinier* (Montgomery, 2008).

Metode *maximum likelihood* dapat digunakan untuk memperoleh nilai taksiran parameter dari model ARIMA. Misalkan parameter model ARIMA adalah AR(1), proses pendugaan atau estimasi parameter menggunakan metode *maximum likelihood* untuk mendapatkan nilai taksiran parameter model AR(1) yaitu sebagai berikut (Aswi & Sukarna, 2006):

$$\begin{aligned}
 Z_t &= \phi Z_{t-1} + a_t \\
 Z_t - \mu &= \phi_1 (Z_{t-1} - \mu) + a_t \\
 Z_t &= \mu + \phi_1 (Z_{t-1} - \mu) + a_t \\
 Z_t &= \mu + \phi_1 Z_{t-1} - \phi_1 \mu + a_t \\
 Z_t &= \mu(1 - \phi_1) + \phi_1 Z_{t-1} + a_t \\
 &\text{dengan, } \theta_0 = \mu(1 - \phi_1) \\
 a_t &\sim WN(0, \sigma_a^2) \text{ atau } a_t \sim i.i.d. N(0, \sigma_a^2)
 \end{aligned} \tag{13}$$

Fungsi kepadatan peluang dari Z_t untuk AR(1) berdistribusi normal. Penjabaran fungsi kepadatan peluang untuk data pertama atau Z_1 dengan rata-rata $\mu = \frac{\theta_0}{1 - \phi_1}$

dan variansi $\frac{\sigma_a^2}{1 - \phi_1^2}$ adalah:

$$f(Z_1; \theta_0, \phi_1, \sigma_a^2) = \frac{1}{\sqrt{\frac{2\pi\sigma_a^2}{1 - \phi_1^2}}} \exp \left(-\frac{\left(Z_1 - \left(\frac{\theta_0}{1 - \phi_1} \right) \right)^2}{\frac{2\sigma_a^2}{1 - \phi_1^2}} \right) \tag{14}$$

Untuk data kedua :

$$Z_2 = \theta_0 + \phi_1 Z_1 + a_2$$

Dengan $\theta_0 + \phi_1 Z_1$ sebagai konstanta, kemudian $a_2 \sim N(0, \sigma_a^2)$ dapat ditulis menjadi:

$$Z_2 \sim N(\theta_0 + \phi_1 Z_1, \sigma_a^2)$$

Fungsi kepadatan peluang untuk data yang kedua ini dapat ditulis sebagai berikut:

$$f(Z_2; \theta_0, \phi_1, \sigma_a^2) = \frac{1}{\sqrt{\frac{2\pi\sigma_a^2}{1 - \phi_1^2}}} \exp \left(-\frac{1}{2\sigma_a^2} (Z_2 - \theta_0 - \phi_1 Z_1)^2 \right)$$

Fungsi kepadatan peluang untuk data ketiga, keempat dan seterusnya, dapat diperoleh dengan cara yang sama. Apabila data yang digunakan sampai pada $t = T$, fungsi kepadatan peluang untuk waktu $t = T$ dapat dituliskan sebagai berikut:

$$f(Z_T; \theta_0, \phi_1, \sigma_a^2) = \frac{1}{\sqrt{2\pi\sigma_a^2}} \exp \left(-\frac{1}{2\sigma_a^2} (Z_T - \theta_0 - \phi_1 Z_{T-1})^2 \right) \tag{15}$$

Fungsi *likelihood* untuk model AR(1) dapat diperoleh dengan mengalikan seluruh fungsi kepadatan peluang sebagai berikut:

$$\begin{aligned}
L(Z_1, Z_2, \dots, Z_T; \theta_0, \phi_1, \sigma_a^2) &= f(Z_1; \theta_0, \phi_1, \sigma_a^2) f(Z_2; \theta_0, \phi_1, \sigma_a^2) \dots f(Z_T; \theta_0, \phi_1, \sigma_a^2) \\
L(Z_1, Z_2, \dots, Z_T; \theta_0, \phi_1, \sigma_a^2) &= f(Z_1; \theta_0, \phi_1, \sigma_a^2) \cdot \prod_{t=2}^T f(Z_t; \theta_0, \phi_1, \sigma_a^2)
\end{aligned} \tag{16}$$

Berdasarkan dari bentuk persamaan (14) dan (15) dapat disisalkan

$$\begin{aligned}
&\frac{1}{\frac{\sqrt{2\pi\sigma_a^2}}{1-\phi_1^2}} \exp\left(-\frac{\left(Z_1 - \left(\frac{\theta_0}{1-\phi_1}\right)\right)^2}{\frac{2\sigma_a^2}{1-\phi_1^2}}\right) = A \text{ dan} \\
&\left(\frac{1}{\sqrt{2\pi\sigma_a^2}}\right)^{T-1} \exp\left(-\frac{1}{2\sigma_a^2} \sum_{t=2}^T (Z_t - \theta_0 - \phi_1 Z_{t-1})^2\right) = B, \text{ maka} \\
&L(Z_1, Z_2, \dots, Z_T; \theta_0, \phi_1, \sigma_a^2) = A \cdot B.
\end{aligned}$$

Penyederhanaan bentuk di atas dapat dilakukan dengan mentransformasikan dalam bentuk persamaan logaritma yang dinotasikan dengan $(\theta_0, \phi_1, \sigma_a^2)$ menjadi:

$$\begin{aligned}
L(\theta_0, \phi_1, \sigma_a^2) &= \ln f(Z_1; \theta_0, \phi_1, \sigma_a^2) + \ln \sum_{t=2}^T f(Z_t; \theta_0, \phi_1, \sigma_a^2) \\
&= -\frac{1}{2} \ln(2\pi) - \frac{1}{2} \ln\left(\frac{\sigma_a^2}{1-\phi_1^2}\right) - \frac{\left(Z_1 - \left(\frac{\theta_0}{1-\phi_1}\right)\right)^2}{\frac{2\sigma_a^2}{1-\phi_1^2}} - \frac{T-1}{2} \ln(2\pi) \\
&\quad - \frac{T-1}{2} \ln(\sigma_a^2) - \frac{1}{2\sigma_a^2} \sum_{t=2}^T (Z_t - \theta_0 - \phi_1 Z_{t-1})^2
\end{aligned} \tag{17}$$

Penaksir *maximum likelihood* untuk $(\theta_0, \phi_1, \sigma_a^2)$ adalah nilai-nilai yang dapat $L(\theta_0, \phi_1, \sigma_a^2)$. Masalah yang dihadapi selanjutnya adalah menentukan titik-titik untuk setiap parameter yang memaksimumkan fungsi logaritma *likelihood*-nya dengan mengambil turunan pertamanya dan menyamakannya dengan nol. Langkah selanjutnya adalah menganggap bahwa nilai Z_1 sebagai peubah deterministic karena untuk $t = 1$ nilai $Z_1 = \theta_0 + a_t$ sehingga kita cukup memaksimumkan fungsi

$$\ln \sum_{t=2}^T f(Z_t; \theta_0, \phi_1, \sigma_a^2) = -\frac{T-1}{2} \ln(2\pi) - \frac{1}{2\sigma_a^2} \sum_{t=2}^T (Z_t - \theta_0 - \phi_1 Z_{t-1})^2 \tag{18}$$

Ada tiga parameter yang akan ditaksir pada model ini yaitu $(\theta_0, \phi_1, \sigma_a^2)$. Dengan mengambil turunan terhadap ketiga parameter tersebut dan menyamakannya dengan nol diperoleh

$$\begin{aligned}
\frac{\partial \ln L}{\partial \theta_0} &= 0 \\
\frac{\partial \ln L}{\partial \phi_1} &= 0, \text{ dan} \\
\frac{\partial \ln L}{\partial \sigma_a^2} &= 0
\end{aligned}$$

Mendifferensialkan fungsi $\ln \sum_{t=2}^T f(Z_t; \theta_0, \phi_1, \sigma_a^2)$ terhadap θ_0 dan ϕ_1 ekuivalen dengan meminimumkan

$$\sum_{t=2}^T (Z_t - \theta_0 - \phi_1 Z_{t-1})^2 = \sum_{t=2}^T (a_t)^2 \quad (19)$$

Sehingga kita dapat menyelesaikannya dengan cara penaksiran kuadrat kecil sederhana.

$$\frac{\partial (\sum_{t=2}^T (Z_t - \theta_0 - \phi_1 Z_{t-1})^2)}{\partial \theta_0} = -2 \sum_{t=2}^T (Z_t - \theta_0 - \phi_1 Z_{t-1}) = 0$$

dan

$$\frac{\partial (\sum_{t=2}^T (Z_t - \theta_0 - \phi_1 Z_{t-1})^2)}{\partial \phi_1} = -2 \sum_{t=2}^T (Z_t - \theta_0 - \phi_1 Z_{t-1})(Z_{t-1}) = 0$$

Hasil penjabaran persamaan di atas adalah

$$-\sum_{t=2}^T Z_t + (T-1)\widehat{\theta}_0 + \widehat{\phi}_1 \sum_{t=2}^T Z_{t-1} = 0 \quad (20)$$

dan

$$-\sum_{t=2}^T Z_t \cdot Z_{t-1} + \widehat{\theta}_0 \sum_{t=2}^T Z_{t-1} + \widehat{\phi}_1 \sum_{t=2}^T (Z_{t-1})^2 = 0 \quad (21)$$

Kedua persamaan (21) dan (22) diselesaikan secara simultan untuk memperoleh nilai $\widehat{\theta}_0$ dan $\widehat{\phi}_1$. Penyelesaian $-\sum_{t=2}^T Z_t + (T-1)\widehat{\theta}_0 + \widehat{\phi}_1 \sum_{t=2}^T Z_{t-1} = 0$ untuk memperoleh nilai $\widehat{\theta}_0$ memberikan

$$\widehat{\theta}_0 = \frac{\sum_{t=2}^T Z_t}{(T-1)} - \widehat{\phi}_1 \frac{\sum_{t=2}^T Z_{t-1}}{(T-1)}$$

Kemudian disubstitusikan kepersamaan (21) menghasilkan

$$\widehat{\phi}_1 = \frac{(T-1) \sum_{t=2}^T Z_t \cdot Z_{t-1} - \sum_{t=2}^T Z_t \sum_{t=2}^T Z_{t-1}}{(T-1) \sum_{t=2}^T (Z_{t-1})^2 - (\sum_{t=2}^T Z_{t-1})^2}$$

Nilai $\widehat{\phi}_1$ sudah diperoleh dan selanjutnya kita dapat mendistribusikan pada persamaan (21) untuk memperoleh nilai $\widehat{\theta}_0$, yaitu:

$$\widehat{\theta}_0 = \frac{\sum_{t=2}^T (Z_{t-1})^2 \sum_{t=2}^T Z_t - \sum_{t=2}^T Z_{t-1} \sum_{t=2}^T Z_t \cdot Z_{t-1}}{(T-1) \sum_{t=2}^T (Z_{t-1})^2 - (\sum_{t=2}^T Z_{t-1})^2}$$

Titik penyelesaian untuk $\widehat{\theta}_0$ dan $\widehat{\phi}_1$ sesungguhnya terjadi dengan titik-titik dimana turunan pertama persamaan (19) dan (20) adalah nol, yaitu pada saat SSE minimum. Hal ini dapat diverifikasi dengan menghitung turunan kedua yang menunjukkan bahwa

$$\frac{\partial^2 \sum_{t=2}^T (a_t)^2}{\partial \theta_0^2} > 0 \text{ dan } \partial^2 \sum_{t=2}^T (a_t)^2 > 0$$

Selanjutnya kita akan mencari nilai

$$\frac{\partial \ln \sum_{t=2}^T f(Z_t; \theta_0, \phi_1, \sigma_a^2)}{\partial \sigma_a^2} = 0$$

Dari hasil penjabaran dengan cara yang sama seperti di atas kita peroleh

$$\frac{\partial \ln \sum_{t=2}^T f(Z_t; \theta_0, \phi_1, \sigma_a^2)}{\partial \sigma_a^2} = \frac{T-1}{2\sigma_a^2} + \frac{1}{2\sigma_a^4} \sum_{t=2}^T (Z_t - \theta_0 - \phi_1 Z_{t-1})^2 = 0$$

$$\hat{\sigma}_a^2 = \frac{\sum_{t=2}^T (Z_t - \hat{\theta}_0 - \hat{\phi}_1 Z_{t-1})^2}{T-1} \quad (22)$$

dengan cara yang sama kita dapat memperoleh taksiran parameter dari model AR(2), MA(1), MA(2) serta gabungan dari AR dan MA.

2.3.4 Pemeriksaan Diagnostik

Pemeriksaan diagnostik model ARIMA dapat dilakukan dengan uji signifikansi parameter dan uji kesesuaian model (meliputi uji *White noise* dan uji normalitas residual). Uji *White noise* bertujuan untuk melihat apakah residual berkorelasi dengan yang lainnya. Uji asumsi normalitas bertujuan untuk mengetahui apakah residual sudah memenuhi asumsi kenormalan atau belum (Aswi dan Sukarna, 2006)

1. Uji Signifikansi Parameter

Model ARIMA yang baik dapat menggambarkan suatu kejadian adalah model yang salah satunya menunjukkan bahwa penaksiran parameternya signifikan berbeda dengan nol. Secara umum, misalkan θ adalah suatu parameter pada model ARIMA Box-Jenkins dan $\hat{\theta}$ adalah nilai taksiran dari parameter tersebut, serta $SE(\hat{\theta})$ adalah *standar error* dari nilai taksiran $\hat{\theta}$, maka uji signifikansi parameter dapat dilakukan dengan tahapan sebagai berikut.

a) Hipotesis

$$H_0: \hat{\theta} = 0$$

$$H_1: \hat{\theta} \neq 0$$

b) Statistik Uji

$$t = \frac{\hat{\theta}}{SE(\hat{\theta})} \quad (23)$$

c) Daerah Penolakan

Tolak H_0 jika $|t| > t_{\alpha/2; df} = n - n_p$,

dengan n = jumlah data dan n_p = banyaknya parameter

atau dengan menggunakan nilai *p-value*, yakni tolak H_0 jika nilai $p > 0,05$ dengan kata lain parameter θ signifikan.

2. Uji Kesesuaian Model

Uji kesesuaian model meliputi kecukupan model (uji apakah sisanya *White noise*) dan uji asumsi distribusi normal.

i. Uji *White noise*

a) Hipotesis

$$H_0 : \rho_1 = \rho_2 = \dots = \rho_K = 0$$

$$H_1 : \text{Minimal ada satu } \rho_j \neq 0, j = 0, 1, 2, \dots, k$$

Statistik uji, yaitu uji Ljung-Box:

$$Q = n(n+2) \sum_{k=1}^m \frac{\rho_k^2}{(n-k)} \quad (24)$$

dengan:

m : banyaknya lag

n : banyaknya pengamatan asli

ρ_k : koefisien autokorelasi

b) Kriteria pengambilan keputusan

Tolak H_0 jika $Q > \chi_{\alpha; df=K-p-q}^2$

dengan nilai p dan q adalah orde dari ARMA (p, q)

atau dengan menggunakan nilai p -value, yakni tolak H_0 jika nilai $p < 0,05$

dengan kata lain residual telah *White noise* (Aswi & Sukarna, 2006).

ii. Uji Residual Berdistribusi Normal

Uji distribusi normal bertujuan untuk mengetahui apakah residual telah berdistribusi normal atau belum. Salah satu cara yang dapat dilakukan dengan menggunakan uji *Kolmogorov-Smirnov*.

a) Hipotesis:

$$H_0 : F(a_t) = F_0(a_t), \forall a_t$$

$$H_1 : F(a_t) \neq F_0(a_t), \forall a_t$$

b) Statistik uji, yaitu :

$$D = \sup |S(a_t) - F_0(a_t)| \quad (25)$$

dengan $F(a_t)$ adalah fungsi distribusi kumulatif dari residual

c) Kriteria pengambilan keputusan ($\alpha = 0,05$)

H_0 ditolak apabila $D > D_{1-\frac{\alpha}{2}}$ atau H_0 ditolak jika p -value $< 0,05$ dengan

kata lain residual sudah berdistribusi normal (Aswi & Sukarna, 2006).

2.3.5 Pemilihan Model Terbaik

Pada pemodelan data deret waktu, ada kemungkinan terdapat beberapa model yang sesuai yaitu semua parameternya signifikan, sisia memenuhi asumsi *White noise* serta berdistribusi normal. Untuk menentukan model yang terbaik dari beberapa model yang memenuhi syaraat tersebut dapat digunakan Kriteria *Akaike's Information Criterion* (AIC). Model terbaik dipilih yang nilai kriterianya terkecil. (Hanke et al, 2001)

Akaike's Information Criterion (AIC)

paling kecil diantara model-model lainnya (Suhartono, 2011). Kriteria AIC dapat dirumuskan sebagai berikut:

$$AIC = n \times \ln\left(\frac{SSE}{n}\right) + 2p + n + n \times \ln(2\pi) \quad (26)$$

dengan :

n : banyaknya pengamatan

SSE : *Sum Square Error*

p : Banyaknya parameter dalam model

Semakin kecil nilai AIC model yang diperoleh berarti semakin baik untuk digunakan (Aswi & Sukarna, 2006).

2.4 Data Split

Data *Split* adalah proses membagi *dataset* menjadi dua yaitu data *train* dan data *test*. Proses ini merupakan bagian penting dalam model *time series*, untuk mencari rasio data *split* terbaik untuk menghasilkan model dengan keakuratan paling baik. Data *time series* adalah data yang bentuknya berurutan dan memiliki rentang waktu perekaman yang konsisten. Beberapa *rule of thumb* atau referensi dalam membagi data pertama berdasarkan *professor* Andrew Ng mengatakan *rule of thumb* dalam pembagian data *split* yaitu 70:30 menjadi mungkin pada data 100, 1000 atau 10000, namun pada jumlah data 1000000 menjadi tidak sesuai dengan rasio 70:30 maka dipakailah *split data* 99:1. Selain itu terdapat *rule of thumb* lain dalam pembagian data yaitu 80:20 atau dikenal dengan *Pareto Principle* yang sering dipakai dalam matematika, ekonomi dan komputer (Adinugroho, 2022).

2.5 Model Fungsi Transfer

Model fungsi transfer adalah suatu model yang menggambarkan nilai prediksi dari suatu *time series* (deret *output* atau Y_t) berdasarkan pada nilai-nilai dari deret itu sendiri (Y_t) dan berdasarkan pula pada data *time series* yang mempunyai hubungan (deret *Input* atau X_t) dengan deret *output*.

Diasumsikan bahwa x_t dan y_t merupakan deret yang telah ditransformasikan dan kedua deret tersebut stasioner. Bentuk umum model transfer single *Input* ditunjukkan pada persamaan berikut:

$$Y_t = v(B)X_t + \eta_t \quad (27)$$

dengan,

$$v(B) = (v_0B + v_1B + v_2B^2 + \dots + v_kB^k), \quad (28)$$

Keterangan :

Y_t = representasi dari deret *output* pada waktu ke- t

X_t = representasi dari deret *Input* pada waktu ke- t

n_t = deret *Noise*

$v(B)$ = bobot *respon impulse*, koefisien pada model fungsi transfer

k = orde fungsi transfer

Pola data ditemukan adanya kemungkinan data yang tidak stasioner, maka deret *Input* dan deret *output* harus ditransformasikan dengan tepat (untuk mengatasi ragam yang non stasioner), dibedakan (untuk mengatakan nilai tengah yang nonstasioner) dan mungkin perlu dihilangkan unsur musimannya (*deseasonalized*) (untuk menyederhanakan model fungsi transfer)(Markidakis *et al*, 1999) sehingga model fungsi transfer juga ditulis sebagai berikut:

$$y_t = \frac{\omega(B)}{\delta(B)} x_{t-b} + n_t \text{ atau } y_t = \frac{\omega(B)}{\delta(B)} x_{t-b} + \frac{\theta(B)}{\phi(B)} a_t \quad (29)$$

dengan,

$$\omega(B) = \omega_0 - \omega_1 B - \omega_2 B^2 - \dots - \omega_s B^s$$

$$\delta(B) = 1 - \delta_1 B - \delta_2 B^2 - \dots - \delta_r B^r$$

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

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

keterangan :

y_t = nilai Y_t yang telah ditransformasikan dan dibedakan

x_t = nilai X_t yang telah ditransformasikan dan dibedakan

$\omega(B)$ = operator orde s yang menyatakan banyaknya pengamatan masa lampau deret *Input* yang berpengaruh terhadap deret *output*

$\delta(B)$ = operator orde r yang menyatakan banyaknya pengamatan masa lalu deret *output*

$\theta(B)$ = operator MA orde ke- q

$\phi(B)$ = operator AR orde ke- p

n_t = deret *Noise*

a_t = gangguan(*Noise*) acak r, s, p, q , dan b konstanta.

Parameter $\theta(B)$ dan $\phi(B)$ menunjukkan operator *moving average* dan *auto regressive* dengan gangguan(*Noise*) n_t . Sedangkan $\omega(B)$ dan $\delta(B)$ menggantikan $v(B)$ yang merupakan konstanta fungsi transfer.

2.5.1 Identifikasi Model Fungsi Transfer

Adapun tahap-tahapan dalam pemodelan fungsi transfer single *Input* adalah sebagai berikut :

1) *Prewhitening* deret *Input*

Prewhitening deret *Input* bertujuan untuk menjadikan deret *Input* menjadi lebih sederhana dengan menghilangkan seluruh pola yang diketahui supaya tertinggal hanya *White noise*. *Prewhitening* deret *Input* x_t dengan proses ARIMA $(p, 0, q)$ adalah:

$$\phi(B)x_t = \theta_x(B)\alpha_t \quad (30)$$

Mengubah deret *Input* x_t menjadi deret α_t sebagai berikut:

$$\frac{\phi_x(B)}{\theta_x(B)}x_t = \alpha_t \quad (31)$$

2) *Prewhitening* deret *output*

Apabila *Prewhitening* dilakukan untuk x_t maka *Prewhitening* juga diterapkan terhadap y_t supaya fungsi transfer dapat memetakan x_t kedalam y_t . Transformasi pada y_t tidak harus mengubah y_t menjadi *White noise*. berikut merupakan *prewhitening* deret y_t :

$$\frac{\phi_x(B)}{\theta_x(B)}y_t = B_t \quad (32)$$

3) *Cross correlation function* (CCF)

Setelah deret *Input* dan deret *output* yang telah melalui proses *Prewhitening*, maka selanjutnya adalah menghitung *cross correlation* antara kedua deret tersebut. *Cross Correlation Function* (CCF) digunakan untuk mengukur kekuatan hubungan antara dua variabel random pada selisih waktu k . Bentuk fungsi kovarian antara x_t dan y_{t+k} (Wei, 2006).

$$\gamma_{xy}(k) = E\{(x_t - \mu_x)(y_{t+k} - \mu_y)\} \quad (33)$$

dengan $k = 0, \pm 1, \pm 2, \pm 3, \dots$, Dengan $\mu_x = E(x_t)$ dan $\mu_y = E(y_t)$. Bentuk *cross correlation function* ditulis sebagai berikut :

$$\hat{p}_{xy}(k) = \frac{\hat{\gamma}_{xy}(k)}{\sigma_x \sigma_y} \quad (34)$$

dengan σ_x dan σ_y adalah standar deviasi dari x_t dan y_t .

Fungsi sampel *cross correlation function* ditulis sebagai berikut :

$$\hat{p}_{xy}(k) = \frac{\hat{\gamma}_{xy}(k)}{s_x s_y} \quad (35)$$

untuk $k = 0, \pm 1, \pm 2, \pm 3, \dots$, dengan :

$$\hat{\gamma}_{xy}(k) = \begin{cases} \frac{1}{n} \sum_{t=1}^{n-k} (x_t - \bar{x})(y_{t+k} - \bar{y}) & k \geq 0 \\ \frac{1}{n} \sum_{t=1}^n (x_t - \bar{x})(y_{t+k} - \bar{y}) & k < 0 \end{cases} \quad (36)$$

$$S_x = \sqrt{\hat{\gamma}_{xx}(0)} \text{ dan } S_y = \sqrt{\hat{\gamma}_{yy}(0)}$$

4) Penetapan orde b, r, s untuk model fungsi transfer yang menghubungkan deret *Input* dan deret *output* (Makridakis *et al*, 1999)

- a. Nilai b menyatakan bahwa y_t mulai dipengaruhi oleh x_t pada periode $t + b$
- b. Nilai s menyatakan seberapa lama deret y_t terus dipengaruhi oleh nilai-nilai baru dari deret *Input* x_t atau y_t dipengaruhi oleh $x_{t-b-1}, x_{t-b-2}, \dots, x_{t-b-s}$

c. Nilai r menyatakan bahwa y_t dipengaruhi oleh nilai masa lalunya $y_{t-1}, \dots, y_{t-r}$.

setelah menetapkan (b, r, s) kemudian dilakukan penaksir bobot *respon impulse*.

$$v_k = \frac{\sigma_\beta}{\sigma_a} \rho_{a\beta}(k) \quad (37)$$

5) Penaksir Awal Deret *Noise* (η_t)

Bobot respon implus diukur secara langsung dan ini memungkinkan dilakukannya perhitungan nilai taksiran dari deret *Noise* n_t dengan

$$\begin{aligned} \hat{\eta}_t &= y_t - \hat{y}_t \\ \hat{\eta}_t &= y_t - \frac{\hat{\omega}(B)}{\hat{\delta}(B)} x_t \\ \hat{\eta}_t &= y_t - \hat{v}(B)x_t \\ \hat{\eta}_t &= y_t - v_0 x_t - v_1 x_{t-1} - v_2 x_{t-2} - \dots - v_n x_{t-n} \end{aligned} \quad (38)$$

6) Penetapan orde p, q untuk model ARIMA $(p, 0, q)$ dari deret *Noise* (n_t)

Setelah didapat persamaan, maka nilai nilai ini dimodelkan dengan pendekatan ARIMA sehingga diperoleh orde p dan q . Model deret *Noise* n_t dapat dinyatakan dengan :

$$\phi_n(B)\eta_t = \theta_n(B)a_t \quad (39)$$

dengan:

- $\phi(B)$ = Polinomial *autoregressive* orde ke- p dari n_t
- $\theta(B)$ = Polinomial *moving average* orde ke- q dari n_t
- a_t = Residual dari deret n_t

Setelah deret n_t diperoleh, langkah selanjutnya adalah memperoleh nilai dari ramalan dengan menggunakan model fungsi transfer yang terbentuk, dalam membentuk model fungsi transfer dapat mengikuti langkah berikut:

Contoh dugaan orde untuk model fungsi transfer: $b = 2, r = 2, s = 2$. Berdasarkan model umum fungsi transfer sebagai berikut, maka:

$$y_t = \frac{\omega(B)}{\delta(B)} x_{t-b} + \eta_t \quad (40)$$

dengan,

$$\begin{aligned} \omega(B) &= \omega_0 - \omega_1 B - \omega_2 B^2 - \dots - \omega_s B^s \\ \delta(B) &= 1 - \delta_1 B - \delta_2 B^2 - \dots - \delta_r B^r \\ \theta(B) &= 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q \\ \phi(B) &= 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p \end{aligned}$$

Dugaan sementara model fungsi transfer untuk $b = 2, r = 2, s = 2$, yaitu

$$y_t = \mu + \frac{C(1 - \omega_1 B - \omega_2 B^2)}{(1 - \delta_1 B - \delta_2 B^2)} x_{t-2} + \eta_t$$

atau, jika parameter μ dikeluarkan dari model, diperoleh

$$y_t = \frac{C(1 - \omega_1 B - \omega_2 B^2)}{(1 - \delta_1 B - \delta_2 B^2)} x_{t-2} + \eta_t$$

Dengan $x_t = X_t - X_{t-1}$ dan $y_t = Y_t - Y_{t-1}$

2.5.2 Penaksiran Parameter Model Fungsi Transfer

Penaksiran parameter model fungsi transfer menggunakan metode *conditional least square* (CLS), dengan melibatkan parameter ω, δ, ϕ , dan θ . Metode CLS merupakan suatu metode yang dilakukan untuk mencari nilai parameter dengan meminimumkan jumlah kuadrat kesalahan. Setelah mengidentifikasi model fungsi transfer pada persamaan 42, selanjutnya parameter $\delta = (\delta_1, \dots, \delta_r)'$, $\omega = (\omega_0, \omega_1, \dots, \omega_s)'$, $\phi = (\phi_1, \dots, \phi_p)'$, $\theta = (\theta_1, \dots, \theta_q)'$, dan σ_a^2 akan diestimasi. Kemudian persamaan (42) dapat ditulis dalam bentuk berikut:

$$\delta_r(B)\phi(B)y_t = \phi(B)\omega_s(B)x_{t-b} + \delta_r(B)\theta(B)a_t \quad (41)$$

atau dapat ditulis dalam bentuk :

$$c(B)y_t = d(B)x_{t-b} + e(B)a_t \quad (42)$$

dengan

$$\begin{aligned} c(B) &= \delta(B)\phi(B) = (1 - \delta_1 B - \dots - \delta_r B^r)(1 - \phi_1 B - \dots - \phi_p B^p) \\ &= (1 - c_1 B - c_2 B^2 - \dots - c_{p+r} B^{p+r}) \\ d(B) &= \phi(B)\omega(B) = (1 - \phi_1 B - \dots - \phi_r B^r)(\omega_0 - \omega_1 B - \dots - \omega_s B^s) \\ &= (1 - c_1 B - c_2 B^2 - \dots - c_{p+r} B^{p+r}), \end{aligned}$$

dan

$$\begin{aligned} e(B) &= \delta(B)\theta(B) = (1 - \delta_1 B - \dots - \delta_r B^r)(1 - \theta_1 B - \dots - \theta_q B^q) \\ &= (1 - e_1 B - e_2 B^2 - \dots - e_{r+q} B^{r+q}), \end{aligned} \quad (43)$$

maka,

$$a_t = y_t - c_1 y_{t-1} - \dots - c_{p+r} y_{t-p-r} - d_0 x_{t-b} - d_1 x_{t-b-1} + \dots + d_{p+s} x_{t-b-p-s} + e_{r+q} a_{t-r-q}$$

dengan c_i, d_j , dan e_k adalah fungsi dari $\delta_i, \omega_j, \phi_k$, dan θ_i . Sedangkan asumsi a_t adalah *White noise* $N(0, \sigma_a^2)$, sehingga fungsi *conditional likelihood* :

$$L(\delta, \omega, \theta, \sigma_a^2 | b, x, y, x_0, y_0, a_0) = (2\pi\sigma_a^2)^{-\frac{n}{2}} \exp \left[-\frac{1}{2\sigma_a^2} \sum_{t=1}^n a_t^2 \right], \quad (44)$$

dengan x_0, y_0, a_0 adalah beberapa nilai awal yang sesuai untuk menghitung a_t dari Persamaan (45) sama dengan nilai awal yang diperlukan dalam pendugaan model ARIMA *univariat*.

Secara umum, metode penaksiran *maximum likelihood* juga dapat digunakan untuk menduga parameter $\omega, \delta, \phi, \theta$, dan σ_a^2 . Sebagai contoh dengan mengatur nilai a sama dengan 0 sebagai nilai ekspektasi kondisional, pendugaan terkecil nonlinier parameter tersebut diperoleh dengan nilai SSE, yaitu :

$$S(\delta, \omega, \phi, \theta | b) = \sum_{t=t_0}^n a_t^2 \quad (45)$$

dengan $t_0 = \max \{p + r + 1, b + p + s + 1\}$ (Wei, 2006).

2.5.3 Uji Diagnosis Model Fungsi Transfer

Setelah model diidentifikasi dan parameternya diestimasi, maka perlu dilakukan pemeriksaan diagnostik model sebelum melakukan peramalan. Berikut pemeriksaan diagnostic model fungsi transfer (Wei, 2006).

a. Pengujian *cross correlation*

Menurut Makridakis dkk (1999) dalam proses perkiraan bobot fungsi transfer terdapat asumsi bahwa deret *Input* yang telah diprewhitening (α_t) adalah bebas dari komponen *Noise* (a_t).

Wei (1990) menjelaskan bahwa untuk sebuah model fungsi transfer yang layak maka koefisien korelasi silang antara a_t dengan α_t seharusnya tidak menunjukkan suatu pola tertentu dan berada diantara dua kesalahan standar $2(n - k)^{-1/2}$.

Sebuah uji Ljung-Box dapat digunakan untuk mengetahui kelayakan model fungsi transfer di bawah hipotesis nol yaitu korelasi silang antara α_t dan a_t tidak berbeda nyata dari nol, dengan uji statistik sebagai berikut :

$$Q = m(m + 2) \sum_{k=0}^k \frac{r_{\alpha_t, a_t}^2(k)}{m - k} \quad (46)$$

dengan :

$$m = n - t_0 + 1$$

n = banyaknya pengamatan

$$t_0 = \max \{p + r + 1, b + p + s + 1\}$$

k = lag maksimum

$r_{\alpha_t, a_t}(k)$ = Koefisien korelasi silang antara deret α_t dan a_t pada lag k

Statistik Q menyebar mengikuti sebaran chi-kuadrat dengan derajat bebas $(k + 1 - r - s)$. Jika Q lebih kecil dari $\chi^2(k + 1 - r - s)$ maka dapat dikatakan bahwa model fungsi transfer sudah layak.

b. Pengujian *Autocorrelation*

Fungsi autokorelasi nilai sisa menunjukkan suatu pola dapat dikatakan model fungsi transfer tidak layak. Dan uji Ljung-Box digunakan untuk menguji kelayakan *Noise*.

Wei (1990), menyatakan bahwa untuk menguji kelayakan model digunakan uji Ljung-Box dengan hipotesis nol yaitu autokorelasi nilai sisaan tidak berbeda nyata dari nol, akan ditunjukkan bahwa statistik Q berdistribusi $\chi^2(k - m)$.

$$Q = m(m + 2) \sum_{k=1}^k \frac{r_{a_t}^2(k)}{m - k} \quad (47)$$

dengan:

$$m = n - t_0 + 1$$

$r_{a_t}(k)$ = Koefisien autokorelasi a_t pada lag k

$$t_0 = \max(p + r + 1, p + s + 1)$$

k = lag maksimum

m = jumlah parameter yang diduga dalam model.

Jika nilai statistik Q lebih kecil daripada $\chi^2(k - p - q)$ maka dapat dikatakan model untuk deret *Noise* η_t sudah layak untuk digunakan.

2.5.4 Pemilihan Model Terbaik

Pada pemodelan data deret waktu, ada kemungkinan terdapat beberapa model yang sesuai yaitu semua parameternya signifikan, sisia memenuhi asumsi *White noise* serta berdistribusi normal. Penentuan model yang terbaik dari beberapa model yang memenuhi syarat tersebut dapat digunakan beberapa kriteria antara lain: Kriteria Akaike's Information Criterion(AIC), *Root Mean Square Error* (RMSE), dan (*Mean Absolute Percentage Error*) MAPE . Model terbaik dipilih yang nilai kriterianya terkecil (Hanke et al, 2001).

1. Akaike's Information Criterion (AIC)

Paling kecil diantara model-model lainnya (Suhartono, 2011). Kriteria AIC dapat dirumuskan sebagai berikut:

$$AIC = n \times \ln\left(\frac{SSE}{n}\right) + 2p + n \times \ln(2\pi) \quad (48)$$

dengan :

n : banyaknya pengamatan

SSE : Sum Square Error

p : banyaknya parameter dalam model

Semakin kecil nilai AIC model yang diperoleh berarti semakin baik untuk digunakan (Aswi & Sukarna, 2006).

2. Root Mean Square Error (RMSE)

Root Mean Square Error (RMSE) adalah suatu kriteria pemilihan model terbaik berdasarkan pada hasil sisa peramalannya. RMSE dirumuskan sebagai berikut:

$$RMSE = \sqrt{\frac{\sum_{t=1}^n (x_t - y_t)^2}{n}} \quad (49)$$

dengan,

n = banyaknya data

Semakin kecil nilai RMSE berarti nilai taksiran semakin mendekati nilai sebenarnya, atau model yang dipilih merupakan model terbaik (Aswi & Sukarna, 2006).

3. *Mean Absolute Percentage Error* (MAPE)

Mean Absolute Percentage Error (MAPE) adalah suatu kriteria pemilihan model terbaik berdasarkan pada hasil sisa peramalannya. Kriteria MAPE dirumuskan sebagai berikut:

$$MAPE = \sum_{t=1}^n \frac{(y_t - \hat{y}_t)}{\hat{y}_t} \times 100\% \quad (50)$$

Keterangan:

n = banyaknya data

y = data aktual

$\hat{y}_t$ = data ramalan

Tabel 2. Interpretasi Nilai MAPE

MAPE	Interpretasi
<10%	Prediksi sangat akurat
10%-20%	Prediksi yang baik
20%-50%	Prediksi yang layak
>50%	Prediksi yang tidak akurat

(Sumari,2020)

Semakin kecil nilai MAPE berarti nilai taksiran semakin mendekati nilai sebenarnya, atau model yang dipilih merupakan model terbaik (Aswi & Sukarna, 2006).