

Peramalan Kelembapan Rata-Rata dan Temperatur Maksimum pada Provinsi Nusa Tenggara Timur Menggunakan *Box-Jenkins*

Naziehah Taibatunniswah¹, Irhamah², Kartika Fithriasari³, Santi Wulan Purnami⁴

¹Departemen Statistika, Institut Teknologi Sepuluh Nopember (ITS), Surabaya, Indonesia

²Departemen Statistika, Institut Teknologi Sepuluh Nopember (ITS), Surabaya, Indonesia

³Departemen Statistika, Institut Teknologi Sepuluh Nopember (ITS), Surabaya, Indonesia

⁴Departemen Statistika, Institut Teknologi Sepuluh Nopember (ITS), Surabaya, Indonesia

e-mail: ¹ziehah2307@gmail.com, ²irhamah@statistika.its.ac.id, ³kartika_f@statistika.its.ac.id, ⁴santi_wp@statistika.its.ac.id

Abstrak— Perbedaan cuaca dan iklim di berbagai daerah Indonesia dalam waktu yang bersamaan dapat terjadi karena lokasi negara Indonesia yang berada tepat di garis Ekuator. Perbedaan cuaca dan iklim ini mendorong pemerintah untuk memperhatikan kondisi geografis, meteorologi, maupun klimatologi secara seksama agar aktivitas masyarakat Indonesia tidak terganggu. Badan Meteorologi, Klimatologi, dan Geofisika (BMKG) merupakan Lembaga Pemerintah Non Departemen dengan status tetap yang melakukan pencatatan beberapa macam kondisi iklim dan cuaca. Salah satunya yakni temperatur maksimum dan kelembapan rata-rata. Daerah yang menjadi fokus penelitian adalah Provinsi Nusa Tenggara Timur (NTT) tepatnya di Stasiun Meteorologi Gewayantana, Flores Timur. Dalam analisis ini akan dilakukan peramalan temperatur maksimum dan kelembapan rata-rata untuk mengetahui kemungkinan terjadinya hujan di daerah Provinsi Nusa Tenggara Timur menggunakan *ARIMA Box-Jenkins*. Berdasarkan hasil perhitungan akurasi dan pengujian asumsi, model $MA[1,2]$ adalah model *ARIMA* paling baik yang dapat menggambarkan variabel kelembapan rata-rata dengan baik serta model $AR[1,2,3,4,5,6,13,22,185]$ merupakan model *ARIMA* terbaik yang dapat menggambarkan variabel temperatur maksimum dengan baik.

Kata Kunci—BMKG, *Box-Jenkins*, Kelembapan Rata-Rata, NTT, Temperatur Maksimum.

I. PENDAHULUAN

Indonesia merupakan negara yang berada di garis ekuator yang mana menjadi penyebab terjadinya perbedaan cuaca dan iklim di berbagai daerah dalam waktu yang bersamaan. Perbedaan cuaca dan iklim tersebut dapat dilihat dari jenis musim yang sedang melanda di sebagian wilayah tertentu dan jenis musim lainnya di sisa wilayah. Hal ini menjadi dorongan bagi pemerintah untuk memperhatikan kondisi geografis, meteorologi, maupun klimatologi yang ada pada setiap wilayah di Indonesia agar aktivitas masyarakat Indonesia tidak terganggu. Oleh karena itu pada tahun

1866, pemerintah Hindia Belanda meresmikan kegiatan pengamatan cuaca dan geofisika menjadi instansi pemerintah dengan nama Observatorium Magnetik dan Meteorologi. Seiring berjalannya waktu, instansi pemerintah tersebut berganti nama menjadi Badan Meteorologi, Klimatologi, dan Geofisika (BMKG) dengan status tetap sebagai Lembaga Pemerintah Non Departemen melalui Peraturan Presiden Nomor 61 Tahun 2008^[1].

Badan Meteorologi, Klimatologi, dan Geofisika (BMKG) melakukan pencatatan beberapa macam parameter pada kondisi iklim dan cuaca yakni temperatur, kelembapan, kecepatan angin, arah angin, curah hujan, dan intensitas radiasi matahari. Selain itu, BMKG juga melakukan pencatatan kualitas udara, gempa bumi, dan tsunami^[1]. Dikarenakan setiap wilayah di Indonesia dapat mempunyai cuaca dan iklim yang berbeda, maka setiap daerah di Indonesia memiliki stasiun-stasiun pencatatnya masing-masing.

Provinsi Nusa Tenggara Timur merupakan salah satu provinsi di Indonesia yang mana cuaca dan iklimnya dipengaruhi oleh faktor global, seperti El Nino – La Nina dan *Australian Monsoon Index (AUSMI)*. Pada daerah Provinsi Nusa Tenggara Timur terdapat 14 Unit Pelaksana Tugas (UPT) yang bertugas mencatat kondisi iklim dan cuaca. UPT tersebut terbagi menjadi tiga macam yaitu 10 stasiun meteorologi, 3 stasiun geofisika, dan 1 stasiun klimatologi^[2].

Salah satu stasiun di Provinsi Nusa Tenggara Timur yang digunakan dalam penelitian adalah Stasiun Meteorologi Gewayantana yang berada di Larantuka, Flores Timur. Pada stasiun ini, data iklim harian yang tersedia cukup lengkap, terutama pada parameter kelembapan rata-rata dan temperatur maksimum. Kedua parameter tersebut bisa menjadi perkiraan akan turunnya hujan di waktu mendatang selain menggunakan parameter curah hujan. Apabila temperatur maksimum rata-rata lebih rendah dari temperatur minimum biasanya dan kelembapan rata-rata relatif rendah, maka dapat diketahui bahwa akan terjadi hujan dengan intensitas sedang hingga lebat.

Oleh karena itu, penelitian ini melakukan analisis peramalan pada parameter kelembapan rata-rata dan temperatur maksimum untuk mendapatkan prediksi kemungkinan terjadinya hujan di wilayah tersebut menggunakan metode ARIMA *Box-Jenkins*.

II. METODOLOGI PENELITIAN

A. Sumber Data

Data yang digunakan dalam penelitian ini berasal dari data sekunder, yaitu data iklim harian pada bulan Januari 2019 hingga Januari 2020 di Stasiun Meteorologi Gewayantana, Flores Timur, Provinsi Nusa Tenggara Timur yang diperoleh dari Pusat Database *Online* Badan Meteorologi, Klimatologi, dan Geofisika (BMKG). Pengambilan data dilakukan pada:

Hari : Sabtu, 9 Januari 2021

Alamat web : <https://dataonline.bmkg.go.id/>

B. Variabel Penelitian

Variabel penelitian yang digunakan dalam penelitian ini yaitu sebagai berikut.

Tabel 1. Variabel Penelitian

Variabel	Keterangan
X ₁	Kelembapan Rata-Rata (%)
X ₂	Temperatur Maksimum (°C)

C. Langkah Analisis

Langkah analisis yang digunakan pada penelitian ini adalah sebagai berikut.

1. Mengumpulkan data.
2. Melakukan *preprocessing* data.
3. Melakukan analisis statistika deskriptif.
4. Melakukan analisis *time series plot*.
5. Melakukan pemeriksaan stasioneritas data.
6. Melakukan penentuan model.
7. Melakukan estimasi dan uji signifikansi parameter.
8. Melakukan pemeriksaan diagnostik (uji asumsi residual).
9. Melakukan pemilihan model terbaik.
10. Menarik kesimpulan dan saran.

III. ANALISIS DAN PEMBAHASAN

A. Preprocessing Data

Data iklim harian yang dicatat oleh Stasiun Meteorologi Gewayantana tidak sepenuhnya lengkap terdiri atas 396 hari dalam setahun lebih sebulan. Ada beberapa data yang tidak terukur maupun tidak dilakukan pengukuran sehingga ketersediaan data tiap bulannya pada masing-masing variabel yang diamati berbeda-beda. Berikut merupakan ketersediaan data yang tercantum pada situs BMKG serta jumlah *missing value* yang terdeteksi pada masing-masing variabel penelitian yang digunakan.

Tabel 2. Missing Value

Bulan	Ketersediaan data X ₁ (%)	Jumlah Missing Value	Ketersediaan data X ₂ (%)	Jumlah Missing Value
Jan-19	90.32258	3	100	0
Feb-19	92.857	2	100	0
Mar-19	96.774	1	96.774	1
Apr-19	96.67	1	100	0

Mei-19	93.548387	2	100	0
Jun-19	93.33	2	100	0
Jul-19	93.548387	2	100	0
Agu-19	93.548387	2	100	0
Sep-19	86.67	4	100	0
Okt-19	96.774	1	100	0
Nov-19	93.33	2	100	0
Des-19	87.096774	4	100	0
Jan-20	96.774	1	100	0
Total		27		1

Tabel 2 memberikan informasi bahwa data kelembapan rata-rata terdapat *missing value* sebanyak 27, sedangkan pada data temperatur maksimum ditemukan *missing value* sebanyak 1. Oleh karena itu, *missing value* tersebut diimputasi menggunakan metode *linear interpolation*. Imputasi dilakukan dengan bantuan *software Microsoft Excel*. Imputasi tersebut dilakukan dengan cara perbandingan antara data yang hilang (y) dengan data yang telah ada sebelumnya dan sesudahnya (y₀ dan y₁) pada setiap waktu data yang hilang (x) terhadap waktu data yang telah ada sebelumnya dan sesudahnya (x₀ dan x₁) sebagaimana yang dirangkum dalam rumus berikut.

$$\frac{y - y_0}{x - x_0} = \frac{y_1 - y_0}{x_1 - x_0} \quad (1)$$

B. Karakteristik Data

Salah satu cara untuk mengetahui karakteristik data dapat dilakukan dengan statistika deskriptif. Penyajian statistika deskriptif dari data iklim harian pada bulan Januari 2019 hingga Januari 2020 di Stasiun Meteorologi Gewayantana, Provinsi Nusa Tenggara Timur adalah sebagai berikut.

Tabel 3. Statistika Deskriptif

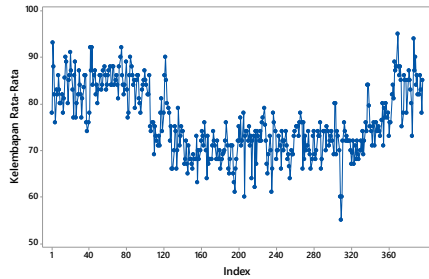
Variabel	Mean	Varians	Minimum	Maksimum
X ₁	76.128	54.374	55.0	95.0
X ₂	32.268	2.368	27.2	38.6

Berdasarkan Tabel 3 didapatkan informasi bahwa rata-rata dari kelembapan rata-rata di wilayah Provinsi Nusa Tenggara Timur pada bulan Januari 2019 hingga Januari 2020 sebesar 76,1% dengan varians perubahan kelembapan tiap harinya cukup besar. Kemudian untuk rata-rata temperatur maksimum yang dialami oleh wilayah Provinsi Nusa Tenggara Timur di bulan Januari 2019 hingga Januari 2020 adalah 32,3°C dengan keragaman perubahan suhu tiap harinya cukup kecil.

C. Analisis Time Series Plot

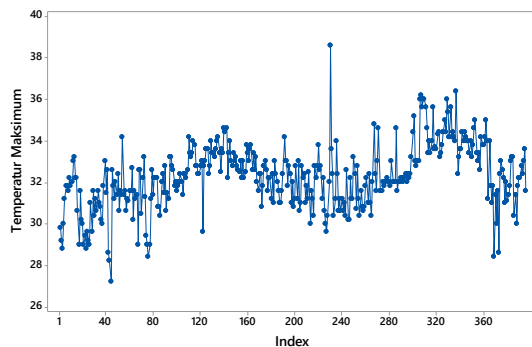
Data yang digunakan pada penelitian ini berupa data iklim harian di bulan Januari 2019 hingga Januari 2020. Berikut merupakan *time series plot* dari data variabel

kelembapan rata-rata (X_1) di bulan Januari 2019 hingga Januari 2020.



Gambar 1. Time Series Plot Kelembapan Rata-Rata

Gambar 1 menunjukkan bahwa pola data kelembapan rata-rata di Provinsi Nusa Tenggara Timur pada bulan Januari 2019 hingga Januari 2020 cenderung fluktuatif dan tidak membentuk pola *trend* maupun musiman. Selain itu, kelembapan rata-rata di wilayah ini berkisar 55 – 95 % dengan angka terendah terjadi pada tanggal 5 Mei 2019 dan angka tertinggi pada tanggal 4 Januari 2020. Kemudian dilakukan analisis *time series plot* pada data variabel temperatur maksimum (X_2) di bulan Januari 2019 hingga Januari 2020 yang ditampilkan pada Gambar 2.



Gambar 2. Time Series Plot Temperatur Maksimum

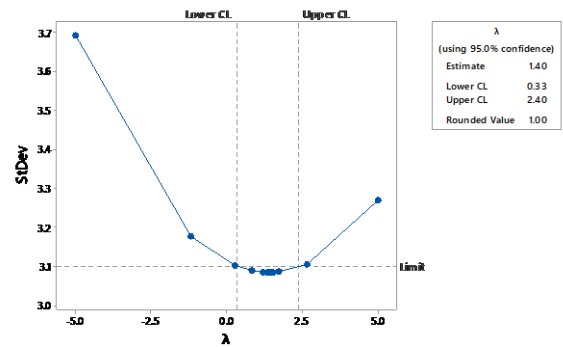
Gambar 2 memperlihatkan bahwa data temperatur maksimum di Provinsi Nusa Tenggara Timur pada bulan Januari 2019 hingga Januari 2020 membentuk pola yang cenderung fluktuatif dan tidak membentuk pola *trend* maupun musiman. Temperatur maksimum terendah yang pernah terjadi adalah 27,2°C pada tanggal 13 Februari 2019 dan temperatur maksimum tertinggi sebesar 38,6°C pada tanggal 18 Agustus 2019.

Berdasarkan *time series plot* yang telah ditampilkan, kedua variabel tersebut dibagi menjadi dua bagian data pada masing-masing variabel, yakni data *in-sample* sebanyak 365 data harian di bulan Januari hingga Desember 2019 dan data *out-sample* sebanyak 31 data harian di bulan Januari 2020. Alasan pembagiannya adalah data iklim harian pada tahun 2019 memiliki pola *time series* yang berfluktuatif dan mulai terjadi kenaikan di akhir bulan Desember 2019 pada data kelembapan rata-rata dan terjadi penurunan pada data temperatur maksimum. Oleh karena itu, untuk meramalkan iklim pada beberapa bulan kedepan, digunakan data bulan Januari 2020 yang mana telah mengikuti pola bulan Desember 2019 di masing-masing variabel.

D. Pemeriksaan Stasioneritas Data

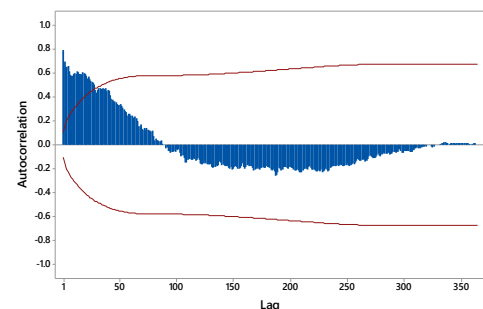
Sebelum membentuk model peramalan, maka perlu dilakukan pemeriksaan data *in-sample* untuk mengetahui apakah data sudah stasioner dalam *mean*

dan varians atau belum. Berikut adalah pemeriksaan stasioneritas menggunakan *Box-Cox plot* pada data *in-sample* kelembapan rata-rata (X_1).



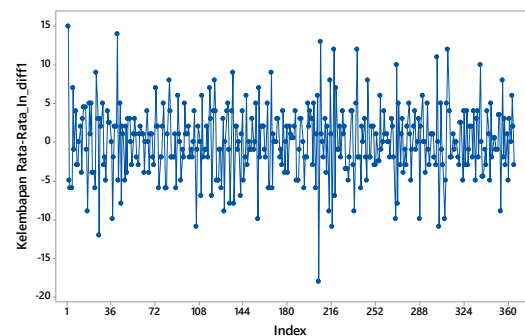
Gambar 3. Box-Cox Plot Kelembapan Rata-Rata

Informasi yang didapatkan pada Gambar 3 adalah *rounded value* yang bernilai 1 dan memiliki nilai 1 dalam rentang CL. Hal ini bermakna bahwa data *in-sample* kelembapan rata-rata telah stasioner dalam varians. Kemudian dilanjutkan dengan pemeriksaan stasioner dalam *mean* menggunakan *ACF plot* yang ditampilkan pada Gambar 4.



Gambar 4. ACF Plot Kelembapan Rata-Rata

Berdasarkan Gambar 4 diketahui bahwa ACF dari data *in-sample* kelembapan rata-rata membentuk pola *dies down* secara lambat dengan taraf signifikansi 5%. Hal ini mengindikasikan bahwa data *in-sample* kelembapan rata-rata belum stasioner dalam *mean* sehingga perlu dilakukan adanya *differencing lag 1* pada data tersebut. *Time series plot* hasil dari *differencing lag 1* pada data *in-sample* kelembapan rata-rata ditunjukkan sebagai berikut.

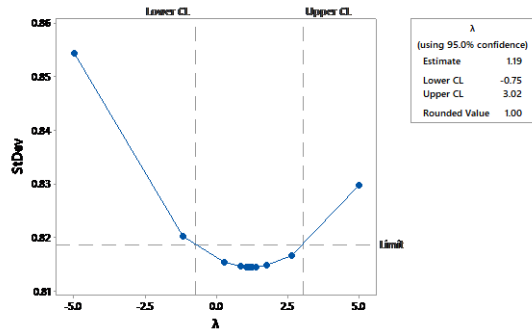


Gambar 5. Time Series Plot Kelembapan Rata-Rata Hasil Differencing Lag 1

Gambar 5 memberikan informasi yakni data sudah berfluktuasi di sekitar *mean*. Hal ini berarti data *in-sample* kelembapan rata-rata hasil *differencing lag 1* telah stasioner dalam *mean*.

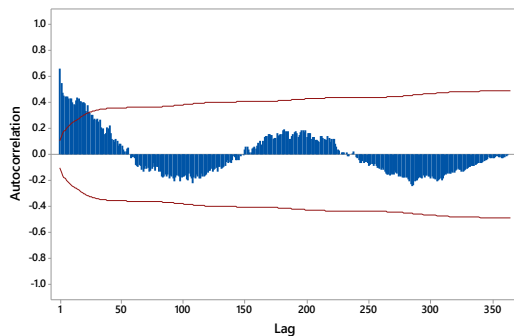
Selanjutnya pemeriksaan stasioner dilakukan pada variabel kedua yakni data *in-sample* temperatur maksimum (X_2). Pemeriksaan ini juga dilakukan

menggunakan *Box-Cox plot* sebagaimana yang ditampilkan pada Gambar 6.



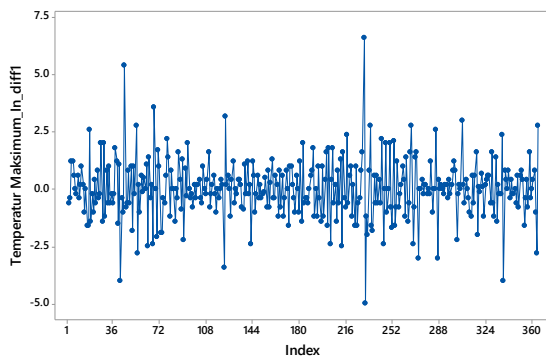
Gambar 6. Box-Cox Plot Temperatur Maksimum

Berdasarkan Gambar 6 diperoleh informasi yakni *rounded value* pada data *in-sample* temperatur maksimum bernilai 1 dengan rentang CL yang mengandung nilai 1. Hal ini mengindikasikan bahwa data *in-sample* temperatur maksimum sudah stasioner dalam varians. Dengan demikian pemeriksaan dapat dilanjutkan pada stasioner dalam *mean* menggunakan *ACF plot* sebagaimana yang ditunjukkan pada Gambar 7.



Gambar 7. ACF Plot Temperatur Maksimum

Informasi yang diperoleh dari Gambar 7 adalah *ACF plot* yang membentuk pola *dies down* secara lambat dengan taraf signifikan 5%. Hal ini mengindikasikan bahwa data *in-sample* temperatur maksimum belum stasioner dalam *mean*. Oleh karena itu, perlu dilakukan adanya *differencing lag 1* pada data *in-sample* temperatur maksimum. Berikut merupakan *time series plot* hasil dari *differencing lag 1* pada data *in-sample* temperatur maksimum.



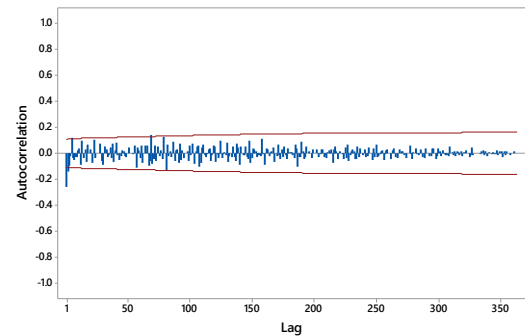
Gambar 8. Time Series Plot Temperatur Maksimum Hasil Differencing Lag 1

Gambar 8 menunjukkan bahwa data *in-sample* temperatur maksimum hasil dari *differencing lag 1* membentuk pola fluktuatif di sekitar *mean*. Hal ini menjelaskan yakni data *in-sample* temperatur maksimum hasil *differencing lag 1* telah stasioner dalam *mean*. Dengan demikian, kedua data variabel

dapat dilanjutkan ke tahap analisis berikutnya yakni penentuan model.

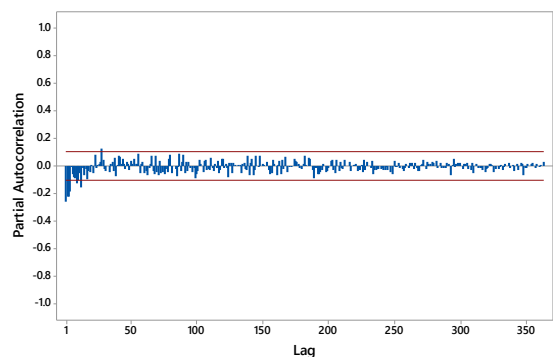
E. Penentuan Model

Model ARIMA dapat diperoleh dari visualisasi data *in-sample* masing-masing variabel hasil *differencing lag 1* menggunakan *ACF plot* dan *PACF plot*. Untuk *ACF plot* sendiri digunakan untuk menentukan model *Moving Average (MA)* dengan melihat *lag* berapa saja yang keluar dari batas signifikansi 5%. *Lag* tersebut yang nantinya akan menjadi indeks dari *q* pada model *MA(q)*. Sedangkan *PACF plot* digunakan untuk menentukan model *Autoregressive (AR)* dengan melihat *lag* berapa saja yang keluar dari batas signifikansi 5%. *Lag* tersebut yang nantinya akan menjadi indeks dari *p* pada model *AR(p)*. Berikut adalah *ACF plot* hasil dari *differencing lag 1* pada data *in-sample* kelembapan rata-rata.



Gambar 9. ACF Plot Kelembapan Rata-Rata Hasil Differencing Lag 1

Berdasarkan Gambar 9 diketahui bahwa terdapat beberapa *lag* yang keluar dari batas signifikansi 5%. *Lag* tersebut yakni 1, 2, 5, dan 69. Oleh karena itu, model yang terbentuk adalah *MA([1,2,5,69])* atau dapat ditulis juga dengan model *ARIMA(0,1,[1,2,5,69])*. Kemudian dilakukan *PACF plot* pada data yang sama yakni data *in-sample* kelembapan rata-rata hasil *differencing lag 1* untuk mendapatkan model *AR(p)* yang ditampilkan pada Gambar 10.

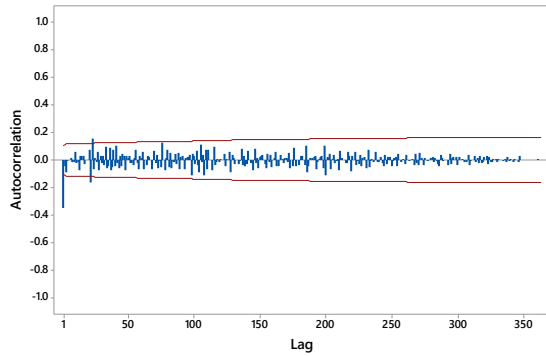


Gambar 10. PACF Plot Kelembapan Rata-Rata Hasil Differencing Lag 1

Informasi yang diperoleh dari Gambar 10 adalah terdapat beberapa *lag* yang keluar dari batas signifikansi 5%. *Lag* tersebut yakni 1, 2, 3, 4, 9, 12, dan 28 sehingga model yang terbentuk adalah *AR([1,2,3,4,9,12,28])* atau dapat ditulis juga dengan model *ARIMA([1,2,3,4,9,12,28],1,0)*.

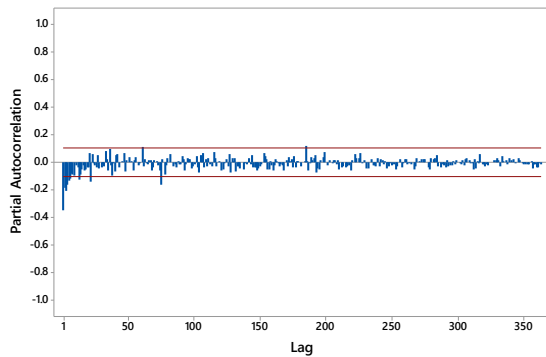
Selanjutnya penentuan model dilakukan pada variabel kedua yakni data *in-sample* temperatur maksimum (X_2) hasil *differencing lag 1*. Berikut

merupakan *ACF plot* pada data *in-sample* temperatur maksimum hasil *differencing lag 1*.



Gambar 11. *ACF Plot* Temperatur Maksimum Hasil *Differencing Lag 1*

Gambar 11 memberikan informasi bahwa data *in-sample* temperatur maksimum hasil *differencing lag 1* memiliki beberapa *lag* yang keluar dari batas signifikansi 5% yakni pada *lag* 1, 22, dan 23. Dari *lag-lag* tersebut didapatkan model yakni model $MA([1,22,23])$ atau model $ARIMA(0,1,[1,22,23])$. Kemudian dilihat pada *PACF plot* dengan data yang sama untuk mendapatkan model *AR*-nya yang ditampilkan pada Gambar 12.



Gambar 12. *PACF Plot* Temperatur Maksimum Hasil *Differencing Lag 1*

Berdasarkan Gambar 12 didapatkan informasi yakni data *in-sample* temperatur maksimum hasil *differencing lag 1* terdapat beberapa *lag* yang keluar dari batas signifikansi 5% yaitu pada *lag* 1, 2, 3, 4, 5, 6, 13, 22, 61, 75, dan 185. Dari 11 *lag* tersebut, model yang terbentuk yakni model $AR([1,2,3,4,5,6,13,22,61,75,185])$ atau dapat juga ditulis dengan model $ARIMA([1,2,3,4,5,6,13,22,61,75,185],1,0)$. Seluruh model yang terbentuk dari kedua data variabel adalah model *ARIMA subset* dengan ditandai adanya kurung siku dalam model.

F. Estimasi dan Uji Signifikansi Parameter

Setelah didapatkan model-model *ARIMA* di setiap variabel, langkah selanjutnya yakni menguji kelayakan tiap parameter yang ada pada masing-masing variabel. Hipotesis yang digunakan dalam pengujian ini adalah gagal tolak H_0 apabila *p-value* lebih besar dari α maka parameter dikatakan tidak signifikan dan tolak H_0 apabila *p-value* kurang dari α maka parameter dikatakan signifikan. Berikut merupakan hasil estimasi

dan pengujian signifikansi parameter dari data *in-sample* kelembapan rata-rata hasil *differencing lag 1*.

Tabel 4. Estimasi dan Uji Signifikansi Kelembapan Rata-Rata

Model ARIMA	Parameter	Estimasi	P-Value
([1,2,3,4,9,12,28],1,0)	AR[1]	-0.41029	< 0.0001
	AR[2]	-0.34927	< 0.0001
	AR[3]	-0.29725	< 0.0001
	AR[4]	-0.17833	0.0007
	AR[9]	-0.05948	0.2176
	AR[12]	-0.08712	0.0724
	AR[28]	0.07429	0.1275
(0,1,[1,2,5,69])	MA[1]	0.50089	< 0.0001
	MA[2]	0.28445	< 0.0001
	MA[5]	0.03201	0.4414
	MA[69]	-0.00752	0.8488

Tabel 4 memberikan informasi bahwa parameter AR[9], AR[12], dan AR[28] pada model *ARIMA* yang pertama memiliki *p-value* lebih besar dari $\alpha = 0.05$ sehingga keputusan gagal tolak H_0 . Hal ini berarti ketiga parameter tersebut tidak signifikan dan dikeluarkan dari model *ARIMA* pertama. Sedangkan untuk model *ARIMA* yang kedua, parameter MA[5] dan MA[69] memiliki *p-value* lebih besar daripada $\alpha = 0.05$ sehingga keputusan yang diambil adalah gagal tolak H_0 . Hal ini bermakna kedua parameter tersebut tidak signifikan dan dikeluarkan dari model *ARIMA* kedua.

Kemudian dilanjutkan pengujian signifikansi pada variabel yang kedua yaitu temperatur maksimum. Berikut merupakan hasil estimasi dan pengujian signifikansi parameter dari model yang didapatkan pada data *in-sample* temperatur maksimum hasil *differencing lag 1*.

Tabel 5. Estimasi dan Uji Signifikansi Temperatur Maksimum

Model ARIMA	Parameter	Estimasi	P-Value
([1,2,3,4,5,6,13,22,61,75,185],1,0)	AR[1]	-0.50902	< 0.0001
	AR[2]	-0.38462	< 0.0001
	AR[3]	-0.35306	< 0.0001
	AR[4]	-0.27854	< 0.0001
	AR[5]	-0.19759	0.0007
	AR[6]	-0.11614	0.0249
	AR[13]	-0.09394	0.0429
	AR[22]	-0.13680	0.0033
	AR[61]	0.08962	0.0675
	AR[75]	-0.07734	0.1185
	AR[185]	0.22147	0.0016
(0,1,[1,22,23])	MA[1]	0.67518	< 0.0001
	MA[22]	0.08748	0.1036
	MA[23]	-0.10026	0.0612

Berdasarkan Tabel 5 didapatkan informasi bahwa terdapat dua parameter pada model *ARIMA* pertama yang memiliki *p-value* lebih besar dari $\alpha = 0.05$ atau keputusan gagal tolak H_0 yakni AR[61] dan AR[75]. Hal ini mengindikasikan bahwa kedua parameter tersebut tidak signifikan sehingga dikeluarkan dari model pertama. Sedangkan pada model *ARIMA* kedua diperoleh dua parameter yang memiliki *p-value* lebih besar dari $\alpha = 0.05$ atau keputusan gagal tolak H_0 yakni MA[22] dan MA[23]. Kedua parameter tersebut berarti tidak signifikan sehingga dikeluarkan dari model *ARIMA* kedua.

Dengan demikian, pada variabel kelembapan rata-rata memiliki 4 parameter *AR* dan 2 parameter *MA* yang signifikan. Serta pada variabel temperatur maksimum memiliki 9 parameter *AR* dan 1 parameter

MA yang signifikan. Berikut merupakan rangkuman model yang terbentuk dari seluruh parameter yang signifikan.

Tabel 6. Model ARIMA yang Terbentuk		
Variabel	Model ARIMA	Persamaan Model
Kelembapan Rata-rata	[(1,2,3,4),1,0)	$Z_t = 1 + 0.41029B + 0.34927B^2 + 0.29725B^3 + 0.17833B^4$
	(0,1,[1,2])	$Z_t = 1 - 0.50089B - 0.28445B^2$
Temperatur Maksimum	[(1,2,3,4,5,6,13,22,185),1,0)	$Z_t = 1 + 0.50902B + 0.38462B^2 + 0.35306B^3 + 0.27854B^4 + 0.19759B^5 + 0.11613B^6 + 0.09394B^{13} + 0.1368B^{22} + 0.22147B^{85}$
	(0,1,[1])	$Z_t = 1 - 0.67518B$

Berdasarkan persamaan model yang telah dicantumkan pada Tabel 6, model yang terbentuk sudah berisi yang signifikan. Dengan demikian keempat model tersebut dapat dilanjutkan ke tahap berikutnya.

G. Pemeriksaan Diagnostik

Pemeriksaan diagnostik atau yang biasa disebut dengan pengujian asumsi residual dilakukan untuk mengetahui apakah model ARIMA yang terdiri atas parameter telah signifikan sudah memenuhi asumsi residual atau belum. Asumsi residual yang harus dipenuhi ada dua, yakni *white noise* dan berdistribusi normal.

1. White Noise

Pada penelitian ini, pemeriksaan diagnostik *white noise* dilakukan menggunakan uji *Ljung-Box*. Hipotesis yang digunakan yaitu hipotesis nol yakni residual sudah *white noise* dan hipotesis satu yakni residual tidak *white noise*. Apabila *p-value* yang didapatkan bernilai lebih besar dari α , maka keputusan yang diambil adalah gagal tolak H_0 sehingga asumsi residual *white noise* pada model ARIMA tersebut telah terpenuhi. Berikut merupakan hasil pengujian residual *white noise* pada variabel kelembapan rata-rata.

Tabel 7. Uji Asumsi Residual *White Noise* Kelembapan Rata-Rata

Model ARIMA	Lag	Statistik Uji	P-Value
[(1,2,3,4),1,0)	12	17.54	0.0036
	24	25.49	0.0842
	36	33.78	0.2475
	48	43.10	0.3816
(0,1,[1,2])	12	15.35	0.0527
	24	28.83	0.0912
	36	39.25	0.1769
	48	50.02	0.2468

Hasil pengujian asumsi residual *white noise* untuk model ARIMA([1,2,3,4],1,0) menunjukkan bahwa pada lag 24, 36, dan 48 memiliki *p-value* lebih besar daripada $\alpha = 0.05$, sedangkan pada lag 12 memiliki *p-value* kurang dari $\alpha = 0.05$. Hal ini mengindikasikan bahwa model ARIMA tersebut masih terdapat autokorelasi pada residual sehingga model tidak memenuhi asumsi residual *white noise*. Untuk model ARIMA(0,1,[1,2]) memperlihatkan bahwa pada lag 12, 24, 36, dan 48 seluruhnya memiliki *p-value* lebih besar dari $\alpha = 0.05$ yang mana berarti model ARIMA tersebut tidak terdapat autokorelasi. Dengan kata lain, model ARIMA(0,1,[1,2]) telah memenuhi asumsi residual *white noise*. Kemudian dilanjutkan dengan pengujian

asumsi residual *white noise* pada variabel temperatur maksimum yang hasilnya dirangkum dalam Tabel 8.

Tabel 8. Uji Asumsi Residual *White Noise* Temperatur Maksimum

Model ARIMA	Lag	Statistik Uji	P-Value
[(1,2,3,4,5,6,13,22,185),1,0)	12	12.28	0.0005
	24	19.21	0.1168
	36	30.15	0.2186
	48	45.54	0.1583
(0,1,[1])	12	29.51	0.0005
	24	39.82	0.0078
	36	46.21	0.0632
	48	60.59	0.0602

Hasil pengujian asumsi residual *white noise* pada variabel temperatur maksimum menunjukkan bahwa untuk model ARIMA([1,2,3,4,5,6,13,22,185],1,0) memiliki *p-value* kurang dari $\alpha = 0.05$ pada lag 12 dan *p-value* lebih dari $\alpha = 0.05$ pada lag 24, 36, dan 48. Hal ini mengindikasikan bahwa model ARIMA tersebut masih terdapat autokorelasi sehingga model belum memenuhi asumsi residual *white noise*. Sedangkan untuk model ARIMA(0,1,[1]) terdapat *p-value* kurang dari $\alpha = 0.05$ pada lag 12 dan 24 serta *p-value* lebih dari $\alpha = 0.05$ pada lag 36 dan 48. Hal ini berarti model ARIMA tersebut masih terdapat autokorelasi sehingga model belum memenuhi asumsi residual *white noise*. Namun apabila dibandingkan kedua model ARIMA tersebut, model ARIMA([1,2,3,4,5,6,13,22,185],1,0) masih lebih baik daripada model ARIMA(0,1,[1]) dikarenakan jumlah lag yang memenuhi asumsi lebih banyak.

2. Distribusi Normal

Asumsi residual yang harus terpenuhi selanjutnya adalah residual berdistribusi normal. Hipotesis yang digunakan yakni gagal tolak H_0 jika *p-value* lebih dari α maka residual berdistribusi normal dan tolak H_0 jika *p-value* kurang dari α maka residual tidak berdistribusi normal. Berikut adalah hasil pengujian asumsi distribusi normal menggunakan uji *Kolmogorov-Smirnov*.

Tabel 9. Uji Asumsi Residual Distribusi Normal

Variabel	Model ARIMA	Statistik Uji	P-Value
Kelembapan Rata-rata	[(1,2,3,4),1,0)	0.053189	0.017
	(0,1,[1,2])	0.051060	0.030
Temperatur Maksimum	[(1,2,3,4,5,6,13,22,185),1,0)	0.051153	0.029
	(0,1,[1])	0.056668	< 0.010

Berdasarkan informasi yang ditampilkan dalam Tabel 9 diketahui bahwa seluruh model ARIMA baik variabel kelembapan rata-rata maupun variabel temperatur maksimum memiliki *p-value* kurang dari $\alpha = 0.05$ sehingga keputusan yang diambil adalah tolak H_0 . Hal ini mengindikasikan bahwa seluruh residual dari keempat model yang ada tidak berdistribusi normal.

H. Pemilihan Model Terbaik

Pada tahap ini, seluruh residual model dihitung nilai akurasi yakni *AIC* (*Akaike Information Criterion*), *MSE* (*Mean Square Error*) dan *MAPE* (*Mean Absolute Percentage Error*) baik pada data *in-sample* maupun *out-sample* untuk menjadi bahan pertimbangan dalam pemilihan model terbaik. Hasil perhitungan akurasi dan

pengujian sebelumnya pada variabel kelembapan rata-rata dirangkum menjadi satu dalam Tabel 10.

Tabel 10. Pemilihan Model Terbaik Kelembapan Rata-Rata

Pengujian	Model ARIMA ([1,2,3,4],1,0)	Model ARIMA (0,1,[1,2])
Signifikansi Parameter	Signifikan	Signifikan
Asumsi <i>White Noise</i>	Tidak memenuhi	Memenuhi
Asumsi Distribusi Normal	Tidak memenuhi	Tidak memenuhi
<i>In-Sample</i>	<i>AIC</i> 2064.84 <i>MSE</i> 16.3837	2054.73 16.1997
<i>Out-Sample</i>	<i>MAPE</i> 6.5026 <i>MSE</i> 41.724	8.0393 63.277

Tabel 10 memberikan informasi yaitu model MA[1,2] atau ARIMA(0,1,[1,2]) memiliki nilai *AIC* dan *MSE* terkecil untuk data *in-sample*, sedangkan model AR atau ARIMA([1,2,3,4],1,0) memiliki nilai *MAPE* dan *MSE* terkecil untuk data *out-sample*. Akan tetapi pada penelitian ini, model terbaik akan mengutamakan nilai akurasi terkecil pada data *in-sample* dan asumsi yang terpenuhi lebih banyak sehingga model terbaik yang akan digunakan untuk variabel kelembapan rata-rata adalah model MA[1,2] atau ARIMA(0,1,[1,2]).

Selanjutnya dilakukan pemilihan model terbaik pada variabel kedua yakni temperatur maksimum. Berikut merupakan hasil rangkuman dari perhitungan akurasi dan pengujian yang dilakukan di tahap sebelumnya pada variabel temperatur maksimum.

Tabel 11. Pemilihan Model Terbaik Temperatur Maksimum

Pengujian	Model ARIMA ([1,2,3,4,5,6,13,22,185],1,0)	Model ARIMA (0,1,[1])
Signifikansi Parameter	Signifikan	Signifikan
Asumsi <i>White Noise</i>	Tidak Memenuhi	Tidak Memenuhi
Asumsi Distribusi Normal	Tidak memenuhi	Tidak memenuhi
<i>In-Sample</i>	<i>AIC</i> 1106.50 <i>MSE</i> 1.15203	1127.08 1.27381
<i>Out-Sample</i>	<i>MAPE</i> 6.57146 <i>MSE</i> 5.78833	5.70747 4.58966

Berdasarkan informasi yang diberikan pada Tabel 11 yaitu model ARIMA([1,2,3,4,5,6,13,22,185],1,0) atau AR[1,2,3,4,5,6,13,22,185] memiliki nilai *AIC* dan *MSE* terkecil pada data *in-sample*, sedangkan model ARIMA(0,1,[1]) atau MA[1] memiliki nilai *MAPE* dan *MSE* terkecil pada data *out-sample*. Namun pada penelitian ini, pemilihan model terbaik akan mengutamakan nilai akurasi terkecil pada data *in-sample* dengan jumlah asumsi terpenuhi yang terbanyak. Dikarenakan dari kedua model sama-sama tidak memiliki asumsi yang terpenuhi sehingga pemilihan model terbaik hanya berdasarkan nilai akurasi. Oleh karena itu, model terbaik yang dipilih pada variabel temperatur maksimum adalah model ARIMA([1,2,3,4,5,6,13,22,185],1,0) atau AR[1,2,3,4,5,6,13,22,185]. Dengan demikian, persamaan model

terbaik pada masing-masing variabel dicantumkan dalam Tabel 12.

Tabel 12. Persamaan Model ARIMA Terbaik

Variabel	Model ARIMA	Persamaan Model
Kelembapan Rata-Rata	(0,1,[1,2])	$Z_t = 1 - 0.50089B - 0.28445B^2$ $Z_t = 1 + 0.50902B + 0.38462B^2$ $+ 0.35306B^3 + 0.27854B^4$ $+ 0.19759B^5 + 0.11613B^6$ $+ 0.09394B^{13} + 0.1368B^{22}$ $+ 0.22147B^{185}$
Temperatur Maksimum	([1,2,3,4,5,6,13,22,185],1,0)	

IV. KESIMPULAN DAN SARAN

Kesimpulan dalam penelitian yaitu pada data variabel kelembapan rata-rata ditemukan *missing value* sebanyak 27 dan data variabel temperatur maksimum ditemukan *missing value* sebanyak 1. *Missing value* tersebut diimputasi menggunakan metode *linear interpolation* dengan bantuan *software Microsoft Excel*. Kedua variabel tersebut dibagi menjadi dua bagian data, yakni data *in-sample* sebanyak 365 data harian di bulan Januari hingga Desember 2019 dan data *out-sample* sebanyak 31 data harian di bulan Januari 2020. Dari hasil penentuan model dan pengujian signifikansi parameter didapatkan model ARIMA([1,2,3,4],1,0) atau AR[1,2,3,4], ARIMA(0,1,[1,2]) atau MA[1,2] untuk variabel kelembapan rata-rata dan model ARIMA([1,2,3,4,5,6,13,22,185],1,0) atau AR[1,2,3,4,5,6,13,22,185] dan model ARIMA(0,1,[1]) atau MA[1] untuk variabel temperatur maksimum. Kemudian pengujian asumsi residual *white noise* didapatkan hasil yakni hanya model MA[1,2] pada variabel kelembapan rata-rata yang memenuhi, sedangkan pengujian asumsi residual berdistribusi normal didapatkan hasil bahwa seluruh residual model tidak ada yang memenuhi. Selanjutnya model terbaik yang dilihat dari nilai akurasi terkecil pada data *in-sample* yakni model MA[1,2] atau ARIMA(0,1,[1,2]) untuk variabel kelembapan rata-rata dan model ARIMA([1,2,3,4,5,6,13,22,185],1,0) atau AR[1,2,3,4,5,6,13,22,185] untuk variabel temperatur maksimum.

Saran bagi pemerintah Provinsi Nusa Tenggara Timur yang mengelola Stasiun Meteorologi Gewayantana adalah dapat meningkatkan ketelitian dalam mencatat data iklim harian agar data yang ada tidak menimbulkan gangguan untuk masyarakat dalam beraktivitas sehari-hari. Saran bagi penulis untuk penelitian selanjutnya yakni meningkatkan ketelitian dalam perhitungan maupun pengujian agar tidak terjadi kesalahan dalam analisis.

DAFTAR PUSTAKA

- [1] BMKG. (2021). *Profil Sejarah Badan Meteorologi, Klimatologi, dan Geofisika* di <https://bmkg.go.id/> (di akses 13 Januari)
- [2] Firmansyah, H. (2020). *BMKG NTT* di <https://ntt.bmkg.go.id/> (diakses pada 13 Januari)