ANALISIS DATA PADA KASUS POSITIF COVID-19 DI PROVINSI LAMPUNG MENGGUNAKAN METODE *AUTOREGRESSIVE INTEGRATED MOVING AVERAGE* (ARIMA) DAN METODE PEMULUSAN EKSPONENSIAL GANDA HOLT

(Skripsi)

Oleh

DINDHA AGUSTINA
NPM 1717031077

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG
2021

ABSTRAK

ANALISIS DATA PADA KASUS POSITIF COVID-19 DI PROVINSI LAMPUNG MENGGUNAKAN METODE *AUTOREGRESSIVE INTEGRATED MOVING AVERAGE* (ARIMA) DAN METODE PEMULUSAN EKSPONENSIAL GANDA HOLT

Oleh

DINDHA AGUSTINA

Metode ARIMA merupakan metode peramalan data deret waktu yang meliputi proses stasioner, autokorelasi, autokorelasi parsial dan lain-lain, sedangkan metode pemulusan eksponensial ganda holt adalah salah satu metode peramalan yang digunakan pada data deret waktu yang mengandung pola *trend*. Tujuan dari penelitian ini adalah menerapkan metode ARIMA dan metode pemulusan eksponensial ganda holt pada data deret waktu. Pemilihan model yang terbaik dalam peramalan adalah model yang memiliki nilai MSE (*Mean Square Error*) terkecil. Data yang dianalisis merupakan data harian pasien positif COVID-19 di Provinsi Lampung pada periode 1 September 2020 hingga 30 Juni 2021. Data dibagi menjadi dua bagian, yaitu data *training* dan data *testing*.

Hasil penelitian menunjukkan bahwa berdasarkan MSE, pengujian data *testing* harian pasien positif COVID-19 di Provinsi Lampung dengan menggunakan metode ARIMA memberikan hasil yang lebih baik dibandingkan dengan metode pemulusan eksponensial ganda Holt.

Kata kunci : Peramalan, *Trend*, Data *Training*, Data *Testing*, Metode ARIMA, Metode Pemulusan Eksponensial Ganda Holt.

ABSTRACT

DATA ANALYSIS OF POSITIVE CASES OF COVID-19 IN LAMPUNG PROVINCE USING AUTOREGRESSIVE INTEGRATED MOVING AVERAGE (ARIMA) METHODS AND DOUBLE HOLT EXPONENTIAL SMOOTHING METHODS

By

DINDHA AGUSTINA

ARIMA method is a time series data forecasting method which includes stationary processes, autocorrelation, partial autocorrelation and others, while the Holt double exponential smoothing method is one of the forecasting methods used on time series data containing trend patterns. The purpose of this study is to apply the ARIMA method and the Holt double exponential smoothing method on time series data. The best model selection in forecasting is the model that has the smallest MSE (Mean Square Error) value. The data analyzed is data on the number of positive COVID-19 patients in Lampung Province in the period September 1, 2020 to June 30, 2021. The data is divided into two parts, namely *training* data and *testing* data.

The results showed that based on MSE, testing the data testing the number of positive COVID-19 patients in Lampung Province using the ARIMA method gave better results than the Holt double exponential smoothing method.

Key words : *Forecasting*, *Trend*, *Training* Data, *Testing* Data, ARIMA Method, Double Exponential Smoothing.

ANALISIS DATA PADA KASUS POSITIF COVID-19 DI PROVINSI LAMPUNG MENGGUNAKAN METODE *AUTOREGRESSIVE INTEGRATED MOVING AVERAGE* (ARIMA) DAN METODE PEMULUSAN EKSPONENSIAL GANDA HOLT

Oleh
DINDHA AGUSTINA

Skripsi

Sebagai Salah Satu Syarat untuk Mencapai Gelar
SARJANA MATEMATIKA

Pada

Jurusan Matematika
Fakultas Matematika dan Ilmu Pengetahuan Alam

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSTAS LAMPUNG
BANDAR LAMPUNG
2021

Judul Skripsi : **ANALISIS DATA PADA KASUS POSITIF COVID-19 DI PROVINSI LAMPUNG MENGGUNAKAN METODE *AUTOREGRESSIVE INTEGRATED MOVING AVERAGE* (ARIMA) DAN METODE PEMULUSAN EKSPONENSIAL GANDA HOLT**

Nama Mahasiswa : **Dindha Agustina**

Nomor Pokok Mahasiswa : 1717031077

Jurusan : Matematika

Fakultas : Matematika dan Ilmu Pengetahuan Alam

MENYETUJUI

1. Komisi Pembimbing

Drs. Nusyirwan, M.Si.
NIP 19661010 199205 1 001

Dr. Fitriani, S.Si., M.Sc.
NIP 19840627 200604 2 001

2. Ketua Jurusan Matematika

Dr. Aang Nuryaman, S.Si., M.Si.
NIP 19740316 200501 1 001

MENGESAHKAN

1. Tim Penguji

Ketua	: **Drs. Nusyirwan, M.Si**	....
Sekretaris	: **Dr. Fitriani, S.Si., M.Sc.**	....
Penguji Bukan Pembimbing	: **Widiarti, S.Si, M.Si**	....

2. Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam

Dr. Eng. Suripto Dwi Yuwono, M.T.
NIP 19740705 200003 1 001

Tanggal Lulus Ujian Skripsi : **23 Agustus 2021**

PERNYATAAN SKRIPSI MAHASISWA

Saya yang bertanda tangan di bawah ini:

Nama	: **Dindha Agustina**
Nomor Pokok Mahasiswa	: **1717031077**
Jurusan	: **Matematika**
Judul	: **ANALISIS DATA PADA KASUS POSITIF COVID-19 DI PROVINSI LAMPUNG MENGGUNAKAN METODE *AUTOREGRESSIVE INTEGRATED MOVING AVERAGE* (ARIMA) DAN METODE PEMULUSAN EKSPONENSIAL GANDA HOLT**

Dengan ini menyatakan bahwa skripsi ini adalah hasil pekerjaan saya sendiri dan sepanjang pengetahuan saya tidak berisi materi yang telah dipublikasikan atau ditulis orang lain atau telah dipergunakan dan diterima sebagai persyaratan penyelesaian studi pada universitas atau institut lain.

Bandar Lampung, Agustus 2021
Yang Menyatakan,

SEPULUH RIBU RUPIAH 10000 TGL. 20 METERAI TEMPEL 08056AJX385196066

Dindha Agustina
1717031077

RIWAYAT HIDUP

Penulis bernama Dindha Agustina, dilahirkan di Dipasena pada tanggal 28 Agustus 1999, sebagai anak kedua dari tiga bersaudara dari Bapak Kadir dan Ibu Parida.

Menempuh pendidikan di Taman Kanak-kanak (TK) Darma Wanita Bd Mulya pada tahun 2004-2005, Sekolah Dasar (SD) di SDN 01 Bd Mulya pada tahun 2005-2011, Sekolah Menengah Pertama (SMP) di SMPN 01 Rawajitu Timur pada tahun 2011-2014, dan Sekolah Menengah Atas (SMA) di SMK Kesehatan Bhakti Nusantara pada tahun 2014-2017.

Pada tahun 2017 penulis terdaftar sebagai mahasiswa S1 Matematika di Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung melalui jalur SBMPTN. Pada tanggal 6 Januari 2020 sampai dengan 7 Februari 2020, penulis melakukan Praktek Kerja Lapangan (PKL) di Kantor Badan Pusat Statistik Provinsi Lampung. Pada tanggal 1 Juli 2020 sampai dengan 10 Agustus 2020, penulis melakukan Kuliah Kerja Nyata (KKN) di Desa Bumi Dipasena Jaya, Rawajitu Timur.

PERSEMBAHAN

Alhamdulillah, puji syukur kehadirat Allah SWT yang maha pengasih lagi Maha Penyayang. Berkat rahmat serta hidayah-Nya penulis dapat menyelesaikan skripsi ini. Tak lupa shalawat serta salam kepada junjungan nabi besar Muhammad SAW yang kita nantikan syafaatnya di yaumil akhir kelak, Aamiin

Dengan kerendahan hati dan rasa syukur, kupersembahkan sebuah karya kecil ini sebagai tanda cinta dan sayangku kepada Abah Bunda, serta kakak dan adik tercinta. Terimakasih atas setiap tetes keringat dan doa dari Abah dan Bunda untuk kebahagiaan dan keberhasilan.

Bapak/Ibu dosen, Bapak/Ibu guru, Sahabat, Teman-temanku yang telah banyak membantu dalam perjalananku hingga saat ini. Terimakasih sudah membuat hari-hari penulis berasa begitu berwarna.

KATA INSPIRASI

“Allah tidak akan membebani seseorang melainkan sesuai dengan kesanggupannya”

(Q.S Al-Baqarah:185)

“Berlelah-lelah lah, manisnya hidup terasa setelah berjuang”

(Imam Syafi’i)

“Wahai orang-orang yang Beriman! Mohonlah pertolongan (kepada Allah) dengan sabar dan sholat. Sungguh, Allah beserta orang-orang yang sabar”

(H.R Tarmidzi)

“Tidak peduli seberapa sulit atau tidak mungkin untuk dicapai, kamu tidak boleh kehilangan pandangan terhadap tujuanmu”

(Monkey D. Luffy (One Piece))

“Jalani saja, lalui saja rintangan yang ada dihadapanmu, kamu pasti bisa melaluinya dengan yakin dan sungguh-sungguh”

(Dindha Agustina)

SANWACANA

Puji syukur penulis ucapkan kehadirat Allah SWT, karena atas rahmat dan hidayah-Nya penulis dapat menyelesaikan skripsi yang berjudul " Analisis Data Pada Kasus Positif COVID-19 di Provinsi Lampung Menggunakan Metode *Autoregressive Integrated Moving Average* (ARIMA) dan Metode Pemulusan Eksponensial Ganda Holt". Skripsi ini disusun sebagai salah satu syarat untuk memperoleh gelar Sarjana Matematika di Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung

Dalam kesempatan ini penulis mengucapkan terima kasih kepada:

1. Bapak Drs. Nusyirwan, M.Si., selaku dosen Pembimbing pertama atas kesediaannya untuk memberikan bimbingan, kritik dan saran dalam proses penyelesaian skripsi ini.
2. Ibu Dr. Fitriani, S.Si., M.Sc., selaku dosen Pembimbing kedua atas kesediaannya memberikan bimbingan, kritik dan saran dalam proses penyelesaian skripsi ini.
3. Ibu Widiarti, S.Si., M.Si., selaku dosen Penguji yang telah memberi pengarahan, kritik dan saran dalam proses penyelesaian skripsi ini.

4. Bapak Dr. Aang Nuryaman, S.Si., M.Si., selaku ketua Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung dan selaku dosen Pembimbing Akademik.
5. Bapak Dr. Eng. Suripto Dwi Yuwono, M.T., selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.
6. Seluruh Dosen dan Staff Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.
7. Keluarga tercinta yang selalu memberikan semangat, motivasi, dan doa yang tak terhingga kepada penulis.
8. Irvan Dwi Saputra, S.Kom., yang telah memberikan semangat dan motivasi bagi penulis.
9. Sahabat-sahabat terbaikku Anggit Khusnul Khotimah, Della Egidia Goeba, Maria Ulfa, Nadhira Dewiantari, Shintia Anjar Wati, Vina Nurmadani, yang selalu memberi semangat dan keceriaan kepada penulis.
10. Seluruh pihak yang telah membantu serta mendoakan penulis dalam menyelesaikan skripsi ini.

Semoga Allah SWT membalas kebaikan bagi semua pihak yang telah berjasa selama menyelesaikan skripsi ini, penulis menyadari bahwa skripsi ini masih jauh dari kesempurnaan, akan tetapi sedikit harapan semoga skripsi yang sederhana ini dapat berguna dan bermanfaat bagi kita semua. Aamiin.

Bandar Lampung, 2021
Penulis

Dindha Agustina

DAFTAR ISI

Halaman

IV. HASIL DAN PEMBAHASAN

V. KESIMPULAN

DAFTAR PUSTAKA

LAMPIRAN

DAFTAR GAMBAR

Gambar Halaman

DAFTAR TABEL

I. PENDAHULUAN

1.1 Latar Belakang

Di awal tahun 2020, dunia dikagetkan dengan kejadian infeksi berat yang dikenal sebagai COVID-19 (*Corona Virus Disease* 2019) yang disebabkan oleh virus *Severe Acute Respiratory Syndrome Coronavirus-2* (SARS-Cov-2). Pandemi COVID-19 ini berawal dari laporan Cina kepada WHO terdapat 44 pasien pneumonia yang berat di suatu wilayah yaitu Kota Wuhan, Provinsi Hubei, Cina dan telah menyebar ke 213 negara di seluruh dunia, salah satunya Indonesia.

Secara umum gejala yang paling umum ditemukan adalah demam dan batuk tidak berdahak. Hampir 90% kasus menunjukkan gejala demam dan 67% menunjukkan gejala batuk tidak berdahak. Kemudian disusul dengan 40% pasien mengeluhkan gejala *fatigue* (tidak enak badan/pegal-pegal) dan 33% pasien melaporkan adanya batuk berdahak. Dari seluruh gejala, hanya 18.6% pasien yang melaporkan adanya gejala kesulitan bernapas (*dyspnea*). Banyak dari gejala yang dilaporkan oleh pasien COVID-19 hampir serupa dengan gejala flu. Namun, pasien COVID-19 jarang mengeluhkan adanya gejala hidung tersumbat atau pilek dibandingkan dengan flu pada umumnya.

Hingga saat ini jumlah kasus positif COVID-19 di Provinsi Lampung bertambah setiap harinya. Untuk melihat penyebaran kasus positif COVID-19 diperlukan adanya peramalan pada kasus tersebut, sehingga tindakan yang tepat dapat dilakukan. Peramalan adalah suatu kegiatan dalam memperkirakan atau kegiatan yang meliputi pembuatan perencanaan di masa yang akan datang dengan menggunakan data masa lalu dan masa sekarang, sehingga dapat membuat prediksi di masa yang akan datang. Ramalan yang dilakukan pada umumnya akan berdasarkan pada data masa lampau. Data masa lampau dikumpulkan, dipelajari, dan dihubungkan dengan perjalanan waktu.

Untuk meminimalisir terjadinya kesalahan pada hasil peramalan yang diinginkan maka perlu digunakan metode yang sesuai. Pemulusan eksponensial ganda merupakan salah satu model peramalan yang terbaik untuk melakukan peramalan pada data deret waktu yang memuat *trend*. Sedangkan metode ARIMA merupakan metode yang dikembangkan oleh George Box dan Gwilym Jenkins yaitu metode yang sering digunakan dalam peramalan data deret waktu (Wulandari, dkk., 2019). Metode ARIMA mencangkup proses-proses pemeriksaan pola data, stasioner dan nonstasioner, autokorelasi, autokorelasi parsial, dan lain-lain.

Pada penelitian ini, penulis ingin membandingkan metode ARIMA, metode pemulusan eksponensial ganda dua parameter dari Holt. Pembandingan tersebut digunakan untuk mencari metode yang terbaik dengan nilai *error* terkecil yang diukur melalui nilai MSE (*Mean Square Error*)

1.2 Tujuan Penelitian

Adapun tujuan yang ingin dicapai dalam tugas akhir ini adalah:

1. menentukan pendugaan metode ARIMA, dan metode pemulusan eksponensial ganda Holt pada peramalan data deret waktu,
2. membandingkan metode peramalan terbaik dari metode ARIMA dan metode pemulusan eksponensial ganda *Holt* yang dapat digunakan untuk menguji data *testing* kasus positif COVID-19 di Provinsi Lampung pada periode 1 Mei 2021 hingga 30 Juni 2021.

1.3 Manfaat Penelitian

Adapun manfaat yang dapat diambil dari penulisan tugas akhir ini adalah:

1. mahasiswa mampu mengaplikasikan ilmu matematika dan statistik yang telah diperoleh ke dalam permasalahan sehari-hari,
2. dapat menambah pengetahuan dan sebagai referensi untuk pembaca apabila ingin melakukan penelitian mengenai peramalan dengan data yang memuat *trend*,
3. sebagai bahan perbandingan untuk penelitian selanjutnya.

II. TINJAUAN PUSTAKA

2.1 *Training* dan *Testing Set*

Training set adalah himpunan data yang digunakan untuk melatih atau membangun model. Sedangkan *testing set* adalah himpunan data yang digunakan untuk menguji model setelah proses latihan selesai. Pada umumnya, rasio pembagian dataset menjadi *training* dan *testing set* adalah (90% ∶ 10%), (80% ∶ 20%), (70% ∶ 30%) atau 50% ∶ 50%).

2.2 Peramalan

Peramalan adalah proses menduga sesuatu yang akan terjadi dimasa yang akan datang. Berdasarkan teori, peramalan adalah perkiraan terjadinya sebuah kejadian di masa depan, berdasarkan data yang ada di masa lampau. Peramalan bertujuan memperoleh ramalan yang dapat mengurangi kesalahan meramal biasanya diukur dengan menggunakan *Mean Square Error* (MSE), *Mean Absolute Error* (MAE), dan *Mean Absolute Precentage Error* (MAPE) (Subagyo, 1984).

Terdapat dua langkah dasar yang harus dilakukan dalam membuat atau menghasilkan suatu peramalan yang akurat dan berguna. Langkah dasar yang pertama adalah pengumpulan data yang relevan dengan tujuan peramalan yang dimaksud dan menurut informasi-informasi yang dapat menghasilkan peramalan yang akurat. Langkah dasar kedua adalah memilih metode peramalan yang tepat yang akan digunakan dalam mengolah informasi yang terkandung dalam data yang telah dikumpulkan (Said, 2011).

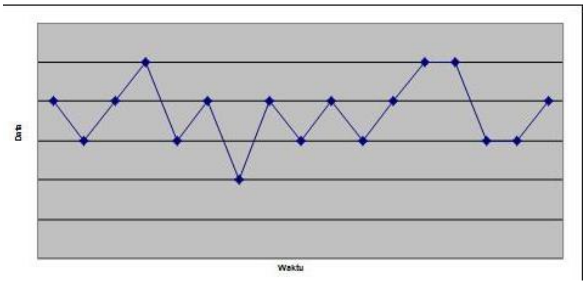
2.3 Analisis Deret Waktu (*Time Series Analysis*)

Analisis deret waktu merupakan serangkaian observasi terhadap suatu variabel yang diambil secara beruntun berdasarkan interval waktu yang tetap (Wei, 2006). Biasanya jarak atau interval dari waktu ke waktu sama. Pada umumnya pengamatan dan pencatatan itu dilakukan dalam jangka waktu tertentu, misalnya akhir tahun, tiap permulaan tahun, tiap sepuluh tahun, dan sebagainya. Deret waktu (*Time Series*) merupakan serangkaian data pengamatan yang terjadi berdasarkan indeks waktu secara berurutan dengan interval waktu tetap. Analisis deret waktu adalah salah satu prosedur statistika yang diterapkan untuk meramalkan struktur variabel statistik keadaan yang akan terjadi di masa yang akan datang dalam rangka pengambilan keputusan (Aswi dan Sukarna, 2006).

Menurut Makridakis, dkk. (2003), pola data deret waktu dapat dibedakan menjadi empat jenis, yaitu:

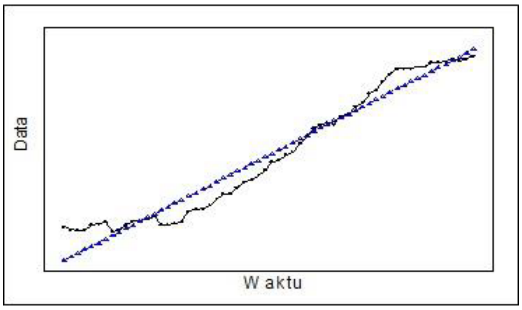
1. Pola acak (*random*) atau pola horizontal, dihasilkan oleh banyak pengaruh independen yang menghasilkan pola non-sistematik dan tidak berulang dari beberapa nilai rataan. Pola acak terjadi karena data yang diambil tidak dipengaruhi oleh faktor-faktor khusus sehingga pola menjadi tidak menentu

dan tidak dapat diperkirakan secara biasa. Struktur datanya dapat digambarkan sebagai berikut:



Gambar 1. Pola horizontal

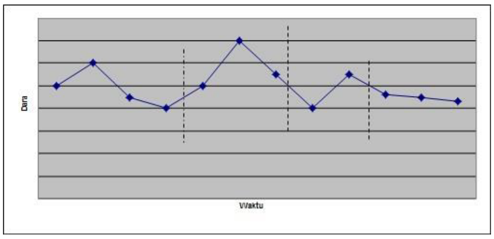
2. Pola tren (*trend*), peningkatan atau penurunan secara umum dari deret waktu yang terjadi selama beberapa periode tertentu. Trend disebabkan oleh perubahan jangka panjang yang terjadi disekitar faktor-faktor yang mempengaruhi data deret waktu. Pola perkembangan data ini membentuk karakteristik yang mendekati garis linier. Gradien yang naik atau turun menunjukkan peningkatan atau pengurangan nilai data sesuai dengan waktu. Struktur datanya dapat digambarkan sebagai berikut:



Gambar 2. Pola tren

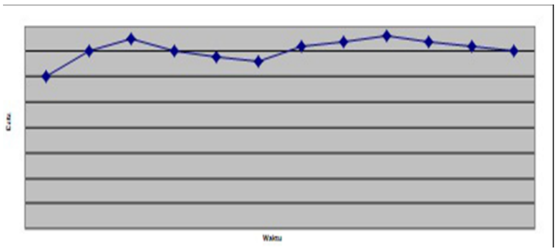
3. Pola musiman (*seasonal*), dihasilkan oleh kejadian yang terjadi secara musiman atau periodik (contoh: iklim, liburan, kebiasaan manusia). Suatu periode

musim dapat terjadi tahunan, bulanan, harian dan untuk beberapa aktivitas bahkan setiap jam. Pola ini terbentuk karena adanya pola kebiasaan dari data dalam suatu periode kecil sehingga grafik yang dihasilkan akan serupa jangka waktu tertentu berulang-ulang. Struktur datanya dapat digambarkan sebagai berikut:



Gambar 3. Pola musiman

4. Pola siklis, biasanya dihasilkan oleh pengaruh ekspansi ekonomi dan bisnis dan kontraksi (resesi dan depresi). Pengaruh siklis ini sulit diprakirakan karena pengaruhnya berulang tetapi tidak periodik. Pola ini masih terus dikembangkan dan diteliti lebih lanjut pemodelannya sehingga dapat diperoleh hasil yang tepat. Struktur datanya dapat digambarkan sebagai berikut:



Gambar 4. Pola siklis

2.4 Stasioneritas

Stasioner berarti bahwa tidak terdapat perubahan drastis pada data. Fluktuasi data berada disekitar suatu nilai rata-rata yang konstan, tidak tergantung pada waktu dan variansi dari fluktuasi tersebut. Stasioner dibagi menjadi dua yaitu :

1. Stasioner dalam rata-rata

   Stasioner dalam rata-rata adalah fluktuasi data berada di sekitar suatu nilai rata-rata yang konstan, tidak tergantung pada waktu dan variansi dari fluktuasi tersebut. Dari bentuk plot data seringkali dapat diketahui bahwa data tersebut stasioner atau tidak stasioner. Apabila dilihat dari plot ACF, maka nilai-nilai autokorelasi dari data stasioner akan turun menjadi nol sesudah time *lag* (selisih waktu) kelima atau keenam.
2. Stasioner dalam ragam

   Sebuah data *time series* dikatakan stasioner dalam ragam apabila struktur dari waktu ke waktu mempunyai fluktuasi data yang tetap atau konstan dan tidak berubah-ubah. Secara visual untuk melihat hal tersebut dapat dibantu dengan menggunakan plot *time series*, yaitu dengan melihat fluktuasi data dari waktu ke waktu. Apabila tidak stasioner dalam ragam maka perlu dilakukan perhitungan dengan metode *box-cox transformation* sehingga data tersebut stasioner dalam ragam (Wei, 2006).

2.5 Pembedaan (*Differencing*)

Menurut Makridakis, dkk. (2003) pembedaan (*differencing*) digunakan untuk mengatasi data yang tidak stasioner dalam rata-rata. Proses pembedaan dilakukan setelah data stasioner dalam ragam. Yang dimaksud diferensiasi adalah menghitung perubahan atau selisih nilai observasi. Untuk mengatasi masalah data yang tidak stasioner pada ragamnya, umumnya dilakukan transformasi data ke bentuk ln (logaritma natural) atau akar kuadrat. Data yang tidak stasioner terhadap ragam juga dapat disebabkan oleh pengaruh musiman (*seasonal*), sehingga setelah dihilangkan pengaruh musimannya data dapat menjadi data yang stasioner. Selanjutnya, jika data tidak stasioner baik pada nilai tengahnya maupun ragamnya dilakukan proses diferensiasi atau transformasi ln atau akar kuadrat.

2.6 Transformasi *Box-Cox*

Untuk menstabilkan ragam dalam suatu data seri digunakan transformasi *Box-Cox*. Transformasi log dan akar kuadrat merupakan anggota dari keluarga *power transformation* yang disebut *Box-Cox Transformation*. Dengan transformasi ini kita mendefinisikan seri baru $T(x_i)$ sebagai

$$T(x_i) = \frac{(x_i^{\lambda} - 1)}{\lambda}, \tag{2.1}$$

Dimana λ adalah bilangan *real*. Jika nilai $\lambda = \frac{1}{2}$, maka disebut transformasi akar karena $T(x_i)$ adalah akar dari x_i (Rofi, 2019).

2.7 Uji *Augmented Dickey-Fuller* (ADF)

Uji ADF merupakan pengujian yang sangat populer dan dikenalkan oleh David Dickey dan Whyne Fuller. Dalam uji ini dibentuk persamaan regresi dari data aktual pada periode ke-t dan ke $t-1$. Dalam uji akar unit digunakan model berikut:

$$Y_t = \rho Y_{t-1} + u_t. \tag{2.2}$$

Jika koefisien regresi dari Y_{t-1} $\rho = 1$, maka disimpulkan bahwa Y_t tidak stasioner. Dengan demikian Y_t dapat disebut mempunyai “*unit root*” atau berarti data tidak stasioner. Bila Persamaan (2.2) dikurangi sisi kanan dan kiri maka persamaannya menjadi:

$$Y_t - Y_{t-1} = \rho Y_{t-1} - Y_{t-1} + u_t$$

$$\Delta Y_t = (\rho - 1)Y_{t-1} + u_t, \tag{2.3}$$

atau dapat ditulis dengan:

$$\Delta Y_t = \delta Y_t + u_t, \tag{2.4}$$

dengan:

ΔY_t = hasil *difference* data pada periode ke-t,

Y_t = data aktual pada periode ke- t,

Y_{t-1} = data aktual pada periode ke- t-1,

δ = koefisien regresi,

u_t = *error* yang *white noise* dengan rata-rata = 0 dan ragam = δ^2.

Pada tahap ini sudah dilakukan pembedaan sebagai metode untuk menanggulangi masalah ketidakstasioneran data dan kemudian data akan diuji kembali. Dari persamaan dapat dibuat hipotesis:

$$H_0 : \delta = 0$$

$$H_0 : \delta \neq 0$$

Jika hipotesis $\delta = 0$ ditolak dengan derajat kepercayaan α maka $\rho = 1$ artinya terdapat akar unit, sehingga data deret waktu Y_t tidak stasioner. Dengan membentuk persamaan regresi ΔY_t dan Y_{t-1} akan diperoleh koefisien regresinya, yaitu δ. Hipotesis dalam uji akar unit menjelaskan bahwa apabila hasil uji menyatakan nilai *Augmented Dickey-Fuller test statistic* lebih kecil nilai kritis pada derajat kepercayaan tertentu atau nilai tingkat signifikansinya lebih kecil dari derajat kepercayaan $\alpha = 0{,}05$, maka hipotesis nol yang menyatakan bahwa data tersebut tidak stasioner ditolak dan demikian sebaliknya (Wei, 2006).

2.8 *Auto Correlation* (AC) dan *Partial Auto Correlation* (PAC)

Korelogram memberikan nilai *Auto Correlation* (AC) dan *Partial Auto Correlation* (PAC). *Auto Correlation* (AC) sama halnya dengan koefisien korelasi, hanya saja koefisien ini menunjukkan keeratan hubungan antara nilai variabel yang sama namun pada periode waktu yang berbeda, sedangkan *Partial Auto Correlation* (PAC) mengukur hubungan antara nilai yang sekarang dengan nilai yang sebelumnya (untuk *lag* tertentu), sedangkan pengaruh *lag* lainnya dianggap tetap.

Adapun nilai autokorelasi untuk *lag* $1,2,3,\ldots,k$ dapat dicari dengan persamaan berikut:

$$r_k = \frac{\sum_{t=k+1}^{n}(Y_t-\bar{Y})(Y_{t-k}-\bar{Y})}{\sum_{t=1}^{n}(Y_t-\bar{Y})^2}, \qquad (2.5)$$

dengan:

r_k = autokorelasi pada lag ke-k,

Y_t = data pengamatan ke-t,

$\bar{Y}$ = rata-rata data,

Y_{t-k} = data pengamatan ke-t-k.

Nilai autokorelasi parsial *lag* ke-k digunakan persamaan berikut

$$\rho_k = \frac{r_k}{r_0}, \qquad (2.6)$$

dengan:

r_k = autokorelasi populasi k,

r_0 = autokorelasi populasi 0.

2.9 *Autoregressive Integrated Moving Average* (ARIMA)

ARIMA sering juga disebut metode runtun waktu Box-Jenkins. ARIMA sangat baik ketepatannya untuk prakiraan jangka pendek, sedangkan untuk prakiraan jangka panjang ketepatan prakiraannya kurang baik. Biasanya akan cenderung mendatar/konstan untuk periode yang cukup panjang. ARIMA dapat diartikan sebagai gabungan dari dua model, yaitu model *autoregressive* (AR) yang di integrasikan dengan model *Moving Average* (MA). Model ARIMA umumnya dituliskan dengan notasi ARIMA (p,d,q). Notasi p adalah derajat proses AR, notasi *d* adalah orde pembedaan dan notasi *q* adalah derajat proses MA (Nachrowi dan Usman, 2006).

Model ARIMA adalah model yang secara penuh mengabaikan independen variabel dalam membuat prakiraan. ARIMA menggunakan nilai masa lalu dan sekarang dari variabel dependen untuk menghasilkan prakiraan jangka pendek yang akurat. ARIMA cocok jika observasi deret waktu (*time series*) secara statistik berhubungan satu sama lain (*dependent*).

2.9.1 Model *Autoregressive* (AR)

Model AR adalah model yang menggambarkan bahwa nilai masa sekarang dipengaruhi oleh nilai masa lampau. Model AR dengan orde *p* dinotasikan dengan AR (berorde *p*). Bentuk umum model AR(berorde *p*) adalah:

$$Y_t = \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \cdots + \alpha_p Y_{t-p} + \varepsilon_t, \tag{2.7}$$

dengan:

Y_t = nilai variabel pada waktu ke-t,

$Y_{t-1}, Y_{t-2}, \cdots Y_{t-p}$ = variabel yang merupakan *lag* dari Y_t,

$\alpha_1, \alpha_2, \cdots \alpha_p$ = koefisien *autoregressive* (AR),

ε_t = nilai galat pada waktu ke- t,

p = orde AR.

2.9.2 Model *Moving Average* (MA)

Model *Moving Average* (MA) adalah model perataan nilai dengan mengambil sekelompok nilai pengamatan yang kemudian dicari rata-rata nya, Model *moving average* dinotasikan dengan MA(berorde q), persamaannya adalah:

$$Y_t = \varepsilon_t - \beta_1 \varepsilon_{t-1} - \beta_2 \varepsilon_{t-2} - \cdots - \beta_q \varepsilon_{t-q}, \qquad (2.8)$$

dengan:

Y_t	: nilai variabel pada waktu ke-t,
$\varepsilon_t, \varepsilon_{t-1}, \dots, \varepsilon_{t-q}$	: nilai residu pada waktu *t*, $t-1, \dots, t-q$,
$\beta_1, \beta_2 \dots, \beta_q$	: parameter model *Moving Average* (MA),
q	: orde MA.

2.9.3 Model *Autoregressive Moving Average* (ARMA)

Model ARMA (berode *p*,*q*) merupakan kombinasi dari model AR (berorde *p*) dan MA(berorde *q*) yang memiliki asumsi bahwa data periode sekarang dipengaruhi oleh data pada periode sebelumnya dan nilai residual pada periode sebelumnya (Assauri, 1984). Model ARMA (berorde *p*,*q*) dinyatakan sebagai berikut:

$$Y_t = \alpha_1 Y_{t-1} + \cdots + \alpha_p Y_{t-p} + \varepsilon_t - \beta_1 \varepsilon_{t-1} - \cdots - \beta_q \varepsilon_{t-q}. \qquad (2.9)$$

2.10 *Diagnostic Checking*

Diagnostic Checking yang akan dilakukan meliputi uji signifikansi parameter dan kesesuaian model. Uji kesesuaian model terdiri dari asumsi *white noise* dan berdistribusi normal. Pengujian ini dilakukan untuk memeriksa kesesuaian antara hasil estimasi model dengan data yang ada.

2.10.1 Uji Signifikansi Parameter

Uji signifikansi parameter dilakukan setelah diperoleh nilai estimasi dari parameter-parameter yang terdapat dalam model. Uji signifikansi parameter dilakukan dengan menggunakan signifikansi level 0.05. Uji signifikansi parameter dapat dilakukan dengan tahap sebagai berikut (Wei, 2006):

Tabel 1. Hipotesis Parameter Model AR dan MA

Keterangan	Parameter AR	Parameter MA
Hipotesis	H_0: parameter tidak signifikan	H_0: parameter tidak signifikan
	H_1: parameter signifikan	H_1: parameter signifikan
Statistik uji	$t = \frac{\alpha}{SE(\alpha)}$	$t = \frac{\beta}{SE(\alpha)}$

Keterangan :

α = nilai taksiran dari parameter AR,

β = nilai taksiran dari parameter MA,

SE = *Standar Error*.

Jika diterapkan taraf signifikan 5%, maka daerah penolakannya adalah tH_0 jika $|t| > {}^{t\alpha}/_{2.\,df}$ atau *p-value* < 5%.

2.10.2 Uji *White Noise*

Uji asumsi *white noise* merupakan tidak adannya autokorelasi residual dengan residual data sebelumnya dan mengikuti distribusi normal. Statistik uji yang digunakan dalam pengujian asumsi residual *white noise* adalah statistik uji *Box-Pierce*. Berikut merupakan pengujian *white noise* untuk melihat adanya autokorelasi dengan menggunakan statistic uji *Box-Pierce* (Wei, 2006).

Hipotesis:

H_0 = residual bersifat acak.

H_1 = residual tidak bersifat acak.

Hipotesis diatas diuji menggunakan rumus sebagai berikut:

$$Q = n \sum_{k=1}^{K} r_k^2, \tag{2.10}$$

dengan:

n = banyaknya data,

r_k = nilai autokorelasi *lag* ke-k,

K = nilai maksimum *lag*,

k = nilai *lag*.

Jika ditetapkan taraf signifikan $\alpha = 0.05$, maka daerah penolakannya adalah terima H_0 jika $Q > \chi^2_{\alpha,df}$ atau *p-value* >α.

Setelah melakukan uji *Box-Pierce*, pengujian dilanjutkan dengan uji asumsi residual berdistribusi normal. Pada pengujian asumsi residual berdistribusi normal ini statistik uji yang digunakan adalah statistik uji *Kolmogorov-Smirnov*. Berikut merupakan bentuk pengujian asumsi residual berdistribusi normal (Daniel, 1989).

Hipotesis:

H_0 = residual berdistribusi normal

H_1 = residual tidak berdistribusi normal

Hipotesis diatas diuji menggunakan rumus sebagai berikut:

$$D = |FT - FS| \quad (2.11)$$

dimana:

FT = frekuensi kumulatif teoritis

FS = frekuensi kumulatif observasi

Jika ditetapkan taraf signifikan $\alpha = 0.05$, maka daerah penolakannya adalah terima H_0 jika $D_{hit} < D_{tabel}$ atau *p-value* > α.

2.11 *Exponential Smoothing*

Menurut Handoko (1984), prinsip dari metode *exponential smoothing* adalah menggunakan nilai penghalusan secara eksponensial sebagai ramalan nilai di masa mendatang. *Exponential smoothing* adalah suatu tipe teknik peramalan rata-rata bergerak yang melakukan penimbangan terhadap data di masa lalu dengan cara eksponensial sebagai data paling akhir mempunyai bobot lebih besar dalam rata-rata bergerak. *Smothing* adalah mengambil rata-rata dari nilai pada beberapa periode untuk menaksir nilai pada suatu periode.

Menurut Makridakis, dkk. (2003), *exponential smoothing* adalah suatu metode peramalan rata-rata bergerak yang melakukan pembobotan menurun secara eksponensial terhadap nilai-nilai observasi yang lebih tua. Metode *exponential smoothing* merupakan pengembangan dari metode *moving average*. Dalam metode ini peramalan dilakukan dengan mengulang perhitungan secara terus menerus dengan menggunakan data terbaru.

2.11.1 Metode Pemulusan Eksponensial Ganda Holt

Menurut Makridakis, dkk. (2003), metode eksponensial dari Holt dalam prinsipnya serupa dengan Brown namun Holt tidak menggunakan rumus pemulusan berganda secara langsung. Sebagai gantinya, Holt memuluskan nilai *trend* dengan parameter yang berbeda dari parameter yang digunakan pada deret asli. Rumus untuk pemulusan eksponensial ganda Holt adalah sebagai berikut:

$$
\begin{aligned}
S_t &= \alpha(X_t - S_{t-1} - b_{t-1}) + S_{t-1} + b_{t-1} \\
&= (\alpha X_t - \alpha S_{t-1} - \alpha b_{t-1}) + S_{t-1} + b_{t-1} \\
&= \alpha X_t - \alpha S_{t-1} + S_{t-1} - \alpha b_{t-1} + b_{t-1} \\
&= \alpha X_t + (1-\alpha)S_{t-1} + (1-\alpha)b_{t-1}
\end{aligned}
$$

$$S_t = \alpha X_t + (1-\alpha)(S_{t-1} + b_{t-1}), \tag{2.12}$$

dengan:

S_t = pemulusan eksponensial pada periode ke-t,

X_t = nilai aktual pada periode ke-t,

α = parameter pemulusan eksponensial, $0 < \alpha < 1$,

b_{t-1} = nilai pemulusan unsur *trend* pada periode ke- t-1.

Untuk menghitung pemulusan unsur *trend* digunakan persamaan sebagai berikut:

$$b_t = \gamma(S_t - S_{t-1}) + (1-\gamma)b_{t-1}, \tag{2.13}$$

dengan:

γ = konstanta pembobot pemulusan unsur *trend* $(0 < \gamma < 1)$,

S_t = pemulusan eksponensial pada periode ke-t,

S_{t-1} = pemulusan eksponensial pada periode ke-t-1,

b_t = pemulusan unsur *trend* pada periode ke-t,

b_{t-1} = pemulusan unsur *trend* pada periode ke-t-1.

Peramalan menggunakan metode pemulusan eksponensial ganda dari Holt yaitu dengan menghitung pemulusan eksponensial dan pemulusan unsur *trend*. Setelah kedua faktor ditemukan nilai pemulusannya, langkah terakhir adalah peramalan data pada periode *m* yang akan datang dengan rumus:

$$F_{t+1} = S_t + b_t.1$$

$$F_{t+2} = S_t + b_t.2$$

$$F_{t+3} = S_t + b_t.3$$

$$F_{t+m} = S_t + b_t.m, \tag{2.14}$$

dengan:

F_t = nilai yang ingin diramalkan,

S_t = pemulusan eksponensial pada periode ke-t,

b_t = pemulusan unsur *trend* pada periode ke-t,

m = periode waktu yang akan diramalkan.

2.12 Proses Inisialisasi

Proses inisialisasi atau penentuan nilai awal memiliki peranan cukup penting dalam melakukan peramalan dengan metode pemulusan eksponensial. Misalkan untuk menghitung pemulusan periode kedepan maka:

$$S_{t+1} = \alpha X_t + (1-\alpha)S_t. \qquad (2.15)$$

Bila $t = 1$, Persamaan (2.14) menjadi:

$$S_2 = \alpha X_1 + (1-\alpha)S_1 \qquad (2.16)$$

Untuk memperoleh nilai S_2, S_1 harus diketahui terlebih dahulu. nilai S_1 tersebut adalah:

$$S_1 = \alpha X_0 + (1-\alpha)S_0 \qquad (2.17)$$

Akan tetapi, nilai X_0 dan S_0 tidak ada. Begitu juga dengan nilai S_1 yang harus diketahui untuk meghitung nilai S_2, akan tetapi nilai tersebut tidak diperoleh dari data yang ada. Oleh karena itu, diperlukan suatu cara untuk menentukan nilai S_1 tersebut. Berikut adalah cara penentuan nilai awal untuk metode pemulusan eksponensial (Pindyck, 1998):

$$S_1 = X_1 \qquad (2.18)$$

$$b_1 = X_2 - X_1 \qquad (2.19)$$

dengan:

S_1 = pemulusan eksponensial periode ke-1,

X_1, X_2 = nilai aktual pada periode ke-1, ke-2,

b_1 = pemulusan unsur *trend* periode ke-1.

2.13 Kesalahan Peramalan (*Forecast Error*)

Ukuran kesalahan peramalan digunakan untuk mengevaluasi nilai parameter peramalan. Nilai parameter peramalan yang terbaik adalah nilai yang memberikan kesalahan peramalan yang terkecil, kesalahan peramalan pada metode pemulusan eksponensial ganda Holt dihitung menggunakan:

1. *Mean Absolute Error* (MAE)

$$MAE = \frac{1}{n}\sum_{t=1}^{n}\left|Y_t - \hat{Y}_t\right| \tag{2.21}$$

2. *Mean Square Error* (MSE)

$$MSE = \frac{1}{n}\sum_{t=1}^{n}(Y_t - \hat{Y}_t)^2 \tag{2.22}$$

3. *Mean Absolute Percentage Error* (MAPE)

$$MAPE = \frac{1}{n}\left(\sum_{t=1}^{n}\left|\frac{Y_t - \hat{Y}_t}{Y_t}\right|\right) \times 100\% \tag{2.23}$$

dengan:

n = banyaknya data yang diamati,

$\hat{Y}_t$ = peramalan ke t,

Y_t = data ke t.

III. METODOLOGI PENELITIAN

3.1 Waktu dan Tempat Penelitian

Penelitian ini dilakukan pada semester genap tahun akademik 2020/2021 di Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.

3.2 Data Penelitian

Data yang digunakan dalam penulisan ini adalah data sekunder data harian penderita positif COVID-19 di Provinsi Lampung yang diperoleh dari Dinas Kesehatan Provinsi Lampung dari 1 September 2020 hingga 30 Juni 2021.

3.3 Metode Penelitian

Langkah-langkah yang digunakan dalam penelitian ini, yaitu:

1. Membagi data menjadi dua bagian data *training* dan data *testing.* Data *training* sebesar 80% dan data *testing* sebesar 20%.
2. *Autoregressive Integrated Moving Average* (ARIMA).

a. Membuat plot data harian kasus positif COVID-19 di Provinsi Lampung.
b. Memeriksa kestasioneran data terhadap rata-rata dan ragam.
c. Mengidentifikasi model ARIMA dengan menggunakan metode pemilihan model melalui ACF dan PACF.
d. Menentukan parameter p, d, dan q dalam ARIMA.
e. Mengestimasi parameter dan mengevaluasi model ARIMA.
f. Menentukan model terbaik.

3. Pemulusan Eksponensial Ganda Holt.
   a. Memeriksa kecenderungan data.
   b. Pendugaan parameter α dan γ untuk metode Pemulusan Eksponensial Ganda Holt. Kemudian menentukan parameter terbaik dengan mempertimbangkan nilai MSE.
   c. Perhitungan nilai pemulusan eksponensial ganda dua parameter dari Holt.
4. Hasil pemilihan model yang diperoleh dari metode ARIMA kemudian dibandingkan dengan hasil yang diperoleh dari metode Pemulusan Eksponensial Ganda Holt. Pembandingan ini dilakukan dengan mempertimbangkan nilai MSE untuk dua model tersebut.
5. Melakukan peramalan kasus positif COVID-19 dengan menggunakan model terpilih dan diuji menggunakan data *testing*.

V. KESIMPULAN

Kesimpulan dari penelitian ini adalah metode ARIMA, metode Pemulusan Eksponensial Ganda Holt dapat digunakan untuk menguji data *testing* jumlah pasien positif COVID-19 di Provinsi Lampung. Model ARIMA yang paling tepat dengan data adalah ARIMA(2,1,1) dengan nilai MSE sebesar 375.56. Model Pemulusan Eksponensial Ganda Holt yang paling tepat adalah dengan parameter pemulusan $\alpha = 0.3$ dan $\gamma = 0.001$ dengan nilai MSE sebesar 413. Pengujian data *training* pasien positif COVID-19 di Provinsi Lampung lebih tepat menggunakan metode ARIMA(2,1,1) karena menghasilkan MSE yang lebih kecil dibandingkan metode Pemulusan Eksponensial Ganda Holt.

DAFTAR PUSTAKA

Aswidan Sukarna. 2006. *Analisis Deret Waktu dan Aplikasi*. Andira, Makasar.

Box, G.E.P. dan Jenkins, G.M., 1976. *Time Series Analysis: Forecasting and. Control*. Holden-Day, San Fransisco.

Makridakis, S., Wheelwright, S.C., dan McGee, V.E. 2003. *Metode dan Aplikasi Peramalan*. Edisi Revisi. Binarupa Aksara, Jakarta.

Montgomery, D., Jennings, C., dan Kulachi, M. 2008. *Introduction to Time Series Anaysis and Forecasting.* John Wiley and Sons, Inc., New York.

Pankratz, A. 1991. *Forecasting with Dynamic Regression Models. John Wiley and Sons, Inc., New York.*

Suhartono. 2007. Teori dan aplikasi model intervensi fungsi pulse. *Jurnal Ilmiah Matstat.* **7**(2):191-214.

Suhartono. 2008. *Analisis Data Statistik Dengan R.* Surabaya: Jurusan Statistika, ITS.

Sumintro. 1999. *Metode dan Aplikasi Peramalan.* Binarupa Aksara, Jakarta.

Supranto, J. 1983. *Statistika Teori dan Aplikasi*. Erlangga, Jakarta.

Sutaryo, Natasha, Y., Lintang, S., dan Dea, S.S. 2020. *Buku Praktis Penyakit Virus Corona 19 (COVID-19).* Gadjah Mada University Press, Yogyakarta.

Wei, W.W.S. 2006. *Time Series Analysis: Univariate and Multivariate Methods*. 2nd Edition. Pearson Education Hall, New Jersey.

Wulandari, R.A. dan Gernowo, R. 2019. Metode *Autoregressive Integrated Moving Average* (ARIMA) dan Meode *Adaptive Neuro Fuzzy Inference System* (ANFIS) Dalam Analisis Curah Hujan. *Jurnal Berkala Fisika*. **22**(1): 41.