PENERAPAN MODEL *AUTOREGRESSIVE INTEGRATED MOVING AVERAGE WITH EXOGENUS (*ARIMAX) DENGAN VARIASI KALENDER

(Skripsi)

Oleh

Wilda Maryati

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG
2019

ABSTRACT

THE APPLICATION OF AUTOREGRESSIVE INTEGRATED MOVING AVERAGE WITH EXOGENUS (ARIMAX) MODEL WITH CALENDAR VARIATION

By

Wilda Maryati

Models that commonly used for time series data is autoregressive integrated moving average (ARIMA) model. However, the time series data that are affected by the effects of calendar variations, the ARIMA model is not good enough in modeling. One way to capture the effects of variation of the calendar is the Autoregressive Integrated Moving Average Exogenous (ARIMAX) model. In this research, ARIMAX model is applied on the monthly data of Train Passenger at Tanjung Karang Station from January 2000 to December 2018. From the evaluation results, Eid Fitr is an influential variable. The best ARIMAX model gained is ARIMAX model (1,1,12) with MAPE value of 5.83%.

Kata Kunci: ARIMA, ARIMAX, *calendar variation*

ABSTRAK

PENERAPAN MODEL *AUTOREGRESSIVE INTEGRATED MOVING AVERAGE WITH EXOGENUS (*ARIMAX) DENGAN VARIASI KALENDER

Oleh

Wilda Maryati

Model yang umum digunakan untuk data deret waktu adalah model *autoregressive integrated moving average* (ARIMA). Namun pada data deret waktu yang dipengaruhi efek variasi kalender, model ARIMA belum cukup baik dalam pemodelan. Salah satu cara menangkap efek variasi kalender tersebut adalah dengan model *Autoregressive Integrated Moving Average Exogenous* (ARIMAX). Pada penelitian ini, model ARIMAX diterapkan pada data bulanan Penumpang Kereta Api di Stasiun Tanjung Karang periode Januari 2000 sampai Desember 2018. Dari hasil evaluasi, Idul Fitri merupakan peubah yang berpengaruh. Model ARIMAX terbaik yang didapatkan adalah
model ARIMAX(1,1,12) dengan nilai MAPE sebesar 5.83%.

Kata Kunci: ARIMA, ARIMAX, variasi kalender

PENERAPAN MODEL *AUTOREGRESSIVE INTEGRATED MOVING AVERAGE WITH EXOGENUS (*ARIMAX) DENGAN VARIASI KALENDER

Oleh

Wilda Maryati

Skripsi

Sebagai Salah Satu Syarat untuk Mencapai Gelar
SARJANA SAINS

Pada

Jurusan Matematika
Fakultas Matematika dan Ilmu Pengetahuan Alam

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG
2019

Judul Skripsi : **PENERAPAN MODEL *AUTOREGRESSIVE INTEGRATED MOVING AVERAGE WITH EXOGENUS* (ARIMAX) DENGAN VARIASI KALENDER**

Nama Mahasiswa : Wilda Maryati

Nomor Pokok Mahasiswa : 1517031174

Program Studi : Matematika

Fakultas : Matematika dan Ilmu Pengetahuan Alam

MENYETUJUI

1. Komisi Pembimbing

Widiarti, S.Si., M.Si.
NIP 19800502 200501 2 003

Dian Kurniasari, S.Si., M.Sc.
NIP 19690305 199603 2 001

2. Ketua Jurusan Matematika

Prof. Dra. Wamiliana, MA., Ph.D.
NIP 19631108 198902 2 001

MENGESAHKAN

1. Tim Penguji

Ketua : **Widiarti, S.Si., M.Si.** ....

Sekretaris : **Dian Kurniasari, S.Si., M.Sc.** ....

Penguji
Bukan Pembimbing : **Dr. Khoirin Nisa, S.Si., M.Si.** ....

2. Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam

Drs. Suratman, M.Sc.
NIP 19640604 199003 1 002

Tanggal Lulus Ujian Skripsi : **24 Mei 2019**

PERNYATAAN

Yang bertanda tangan di bawah ini :

Nama	: Wilda Maryati
Nomor Pokok Mahasiswa	: 1517031174
Jurusan	: Matematika

Dengan ini menyatakan bahwa skripsi saya yang berjudul **"PENERAPAN MODEL *AUTOREGRESSIVE MOVING AVERAGE WITH EXOGENUS* (ARIMAX) DENGAN VARIASI KALENDER"** adalah hasil pekerjaan saya sendiri. Semua hasil tulisan dalam skripsi ini telah mengikuti kaidah penulisan karya ilmiah Universitas Lampung. Apabila kemudian hari terbukti bahwa skripsi ini merupakan hasil salinan atau telah dibuat orang lain, maka saya bersedia menerima sanksi sesuai dengan ketentuan akademik yang berlaku.

Bandar Lampung, 17 Juni 2019
Penulis

METERAI TEMPEL
TGL. 20
6160DAFF751452141
6000
ENAM RIBU RUPIAH

Wilda Maryati
NPM.1517031174

RIWAYAT HIDUP

Penulis dilahirkan di Bandar Lampung pada tanggal 25 Maret 1998, sebagai anak kedua dari tiga bersaudara dari pasangan Bapak M. Sidik dan Ibu Ade Masitoh.

Penulis telah menempuh Pendidikan di Sekolah Dasar Negeri (SDN) 1 Kebun Jeruk Kota Bandar Lampung tahun 2003-2009, Sekolah Menengah Pertama Negeri (SMPN) 5 Kota Bandar Lampung tahun 2009-2012, dan Sekolah Menengah Atas Negeri (SMAN) 1 Kota Bandar Lampung pada 2012-2015.

Pada tahun 2015 penulis terdaftar sebagai Mahasiswi Program Studi S1 Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung melalui jalur SBMPTN dan menerima Beasiswa Bidikmisi. Selama menjadi mahasiswi, penulis bergabung di Generasi Muda Matematika (GEMATIKA) periode 2015-2016, pengurus Himpunan Mahasiswa Jurusan Matematika (HIMATIKA) sebagai anggota Bidang Kaderisasi periode 2016 dan 2017, anggota Biro Dana dan Usaha ROIS FMIPA Unila 2017, anggota Departemen Pengembangan Sumber Daya Mahasiswa BEM FMIPA Unila 2017, dan sekretaris Departemen Pengembangan Sumber Daya Mahasiswa BEM FMIPA Unila 2018.

Pada bulan Januari sampai dengan Februari 2018 penulis melaksanakan Kerja Praktik (KP) di Badan Pusat Statistik Kota Bandar Lampung guna menerapkan ilmu yang telah diperoleh sewaktu kuliah. Pada bulan Juli sampai dengan Agustus 2018 penulis melaksanakan Kuliah Kerja Nyata (KKN) Kebangsaan di Desa Maringgai, Kecamatan Labuhan Maringgai, Kabupaten Lampung Timur.

KATA INSPIRASI

Ingatlah Sesungguhnya Pertolongan Allah itu Dekat.

(QS:Al-Baqarah, 214)

Dia yang Menaruh Kepercayaan pada Dunia maka Dunia akan Mengkhianatinya.

(Ali bin Abi Thalib)

Manusia yang Paling Lemah Adalah Dia yang Paling Malas Berdoa Kepada Allah.

(H.R. Abu Ya'la)

PERSEMBAHAN

Dengan mengucap Alhamdulillah, puji syukur kehadirat Allah SWT
Ku persembahkan karya kecil sangat sederhana ini kepada:

Abahku M. Sidik & Emakku Ade Masitoh

Terimakasih Abah, Emak yang telah mengasihi dan menyayangiku dengan penuh rasa tulus dan telah berjuang dengan ikhlas, tak kenal lelah dan waktu. Senantiasa berdoa di setiap jejak langkah kakiku dan memberiku semangat ketika diriku mulai merasa gentar melewati ini semua.

Kakakku Peni Yulianti & Adikku Bunga Ayu Lestari

Terimakasih telah memberi semangat dan doa yang tulus. Mungkin karya ini tak sebanding dengan pengorbanan yang telah kalian lakukan. Kalian adalah motivasi terbesar dalam hidupku, karena aku ingin baiknya diriku
maka lebih baik lagi kalian.

SANWACANA

Alhamdulillah puji syukur kepada Allah SWT yang telah melimpahkan segala rahmat dan hidayah-Nya sehingga skripsi ini dapat terselesaikan. Shalawat serta salam semoga selalu tercurah kepada junjungan alam Nabi Muhammad SAW, penuntun jalan bagi seluruh umat manusia. Skripsi yang berjudul "Penerapan Model *Autoregressive Integrated Moving Average With Exogenus (*ARIMAX) dengan Variasi Kalender" adalah salah satu syarat untuk memperoleh gelar Sarjana Sains di Universitas Lampung.

Penulis menyadari bahwa terselesainya skripsi ini tidak akan terwujud tanpa bantuan dan doa dari mereka yang senantiasa mendukung penulis. Oleh karena itu penulis ingin mengucapkan terimakasih kepada:

1. Ibu Widiarti, S.Si., M.Si., selaku dosen pembimbing satu yang telah membimbing, mengarahkan, dan memotivasi penulis.
2. Ibu Dian Kurniasari, S.Si., M.Sc., selaku dosen pembimbing kedua yang telah memberikan pengarahan dalam proses penyusunan skripsi ini.
3. Ibu Dr. Khoirin Nisa, S.Si., M.Si., selaku penguji atas saran dan kritik yang diberikan untuk skripsi ini.
4. Ibu Dr. Asmiati, S.Si., M.Si., selaku dosen pembimbing akademik yang telah membimbing penulis selama mengikuti perkuliahan.

5. Ibu Prof. Dra. Wamiliana, MA., Ph.D., selaku Ketua Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.
6. Bapak Drs. Suratman, M.Sc., selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.
7. Abah dan Emak tercinta yang selalu mendoakan penulis, selalu memberi dukungan, dan kasih sayang yang tulus kepada penulis.
8. Kakak dan Adik yang tersayang Peni Yulianti dan Bunga Ayu Lestari atas dukungannya selama ini, memberi semangat untuk menyelesaikan skripsi ini.
9. Sahabat-sahabat penulis (Anggun, Liza, Dhenty, Ulfa, Pipin, Riza, Silvi, Ratri) atas dukungan dan persahabatan yang indah.
10. Departemen PSDM berserta pimpinan BEM FMIPA Unila 2018.
11. Teman-teman di Jurusan Matematika angkatan 2015, serta Keluarga HIMATIKA.
12. Teman-teman KKN Kebangsaan (Ongka, Akbar, Bima, Ardina, Kurnia, dan Habsari) yang selalu saling mendukung untuk menyelesaikan skripsi.
13. Seluruh pihak yang telah memotivasi, membantu, dan mendoakan penulis yang tidak dapat disebutkan satu per satu.

Penulis menyadari bahwa skripsi ini memiliki ketidaksempurnaan. Semoga skripsi ini dapat memberikan manfaat bagi kita semua.

Bandar Lampung, 24 Mei 2019
Penulis,

Wilda Maryati
NPM. 1517031174

DAFTAR ISI

V. KESIMPULAN

DAFTAR PUSTAKA

LAMPIRAN

DAFTAR TABEL

DAFTAR GAMBAR

I. PENDAHULUAN

1.1 Latar Belakang dan Masalah

Data *time series* merupakan serangkaian pengamatan yang terurut berdasarkan waktu dengan jarak yang sama (Wei, 2006). Model yang sering digunakan dalam menganalisis data *time series* adalah *Autoregressive Integrated Moving Average* (ARIMA). Model ARIMA sering digunakan karena memiliki fleksibilitas yang tinggi dalam menganalisis berbagai data *time series* dan nilai ramalan yang dihasilkan lebih akurat. Salah satu pengembangan model ARIMA adalah *Seasonal Autoregressive Integrated Moving Average* (SARIMA). Model SARIMA digunakan saat data *time series* memperlihatkan pola periodik yang jelas pada interval waktu tertentu, misalnya harian, mingguan, atau bulanan (Montgomery, *et al*., 2008).

Model SARIMA dapat mengatasi pola musiman dari satu periode waktu. Pada beberapa kasus, pola data juga dapat dipengaruhi oleh variasi kalender. Variasi kalender merupakan pola berulang dengan panjang periode yang bervariasi akibat pengaruh penanggalan kalender yang berbeda-beda setiap tahunnya. Menurut Suhartono, *et.a*l (2010), model yang dapat mengatasi pengaruh variasi kalender adalah model *Autoregressive Integrated Moving Average Exogenus* (ARIMAX). Model ARIMAX merupakan perluasan dari model ARIMA. Pada model ini, faktor yang mempengaruhi variabel respon *y* pada waktu *t* tidak hanya dipengaruhi oleh fungsi variabel *y* dalam waktu *t*, tetapi juga dipengaruhi oleh variabel-variabel *predictor* lain pada waktu ke-*t* (Rosadi, 2011).

Pada umumnya masyarakat memanfaatkan hari-hari libur atau hari besar untuk pulang ke kampung halaman ataupun untuk pergi ke tempat-tempat wisata umum. Hal ini mengakibatkan jumlah penumpang transportasi semakin meningkat, salah satunya adalah jumlah penumpang kereta api. Lonjakan penumpang di hari-hari besar tertentu menjadi kendala yang dihadapi oleh PT. Kereta Api (Persero). Hal tersebut terjadi karena kapasitas angkut yang disediakan tidak seimbang, yakni jumlah permintaan angkutan penumpang kereta api jauh lebih besar dibandingkan dengan kapasitas tempat duduk yang disediakan.

Stasiun Kereta Api Tanjung Karang adalah salah satu *unit* yang mengalami lonjakan efek hari-hari besar. Tingginya lonjakan jumlah penumpang kereta api pada hari-hari besar tersebut menyebabkan pola musiman yang bergerak. Pergeseran hari-hari besar tersebut merupakan salah satu efek variasi kalender. Oleh karena itu dalam proses memodelkan dan meramalkan jumlah penumpang kereta api Stasiun Tanjung Karang, akan digunakan metode peramalan analisis variasi kalender dengan model ARIMAX dimana efek variasi kalender digunakan sebagai variabel *predictor* yang berupa variabel dummy. Pendugaan parameter pada model ARIMAX akan diduga dengan menggunakan metode *Ordinary Least Square* (OLS).

1.2 Tujuan Penelitian

Adapun tujuan dari penelitian ini adalah:

1. Mengkaji model ARIMAX dengan metode pendugaan parameter *Ordinary Least Square* (OLS).
2. Memodelkan dan meramalkan jumlah penumpang kereta api di Stasiun Tanjung Karang dengan metode ARIMAX.

1.3 Manfaat Penelitian

Manfaat dari penelitian ini adalah menambah wawasan terhadap model ARIMAX dengan efek variasi kalender serta memperoleh model untuk meramalkan jumlah penumpang kereta api Stasiun Tanjung Karang pada periode berikutnya.

II. LANDASAN TEORI

2.1 Data *Time Series*

Data *time series* merupakan serangkaian pengamatan yang terurut berdasarkan waktu dengan jarak yang sama (Wei, 2006). Data *time series* sering kali ditemukan dalam berbagai bidang disiplin ilmu seperti ekonomi, pertanian, meteorologi, biologi, serta disiplin ilmu lainnya. Data bentuk *time series* dapat dicatat berdasarkan periode waktu harian, mingguan, bulanan, tahunan, ataupun periode waktu tertentu lainnya dalam rentang waktu yang sama (Cryer dan Chan, 2008). Pada data time series nilai pengamatan suatu periode waktu diasumsikan dipengaruhi oleh nilai pengamatan pada periode waktu sebelumnya, sehingga analisis data *time series* memungkinan untuk melakukan peramalan (*forecasting*) di masa mendatang.

2.2 Kestasioneran

Sekumpulan data dinyatakan stasioner jika nilai rata-rata dan varian dari data *time series* tersebut tidak mengalami perubahan secara sistematik sepanjang waktu, atau sebagian ahli menyatakan rata-rata dan variannya konstan (Nachrowi dan Usman, 2006).

Kestasioneran dalam analisis *time series* diperlukan untuk memperkecil kekeliruan model (Mulyana, 2004). Kestasioneran dibedakan menjadi dua, yakni stasioner dalam rataan dan stasioner dalam ragam. Data yang tidak stasioner dalam rataan, dapat diatasi menggunakan metode *differencing* atau pembedaan pada orde tertentu. Uji formal untuk mengetahui kestasioneran dalam rataan dapat menggunakan uji *Augmented Dickey-Fuller* (ADF) atau dikenal juga dengan uji akar unit.

Data *time series* dikatakan stasioner dalam ragam jika fluktuasi datanya tetap atau konstan, tidak ada perubahan ragam pada fluktuasi. Uji untuk mengetahui kestasioneran dalam ragam salah satunya adalah diagram Box Cox. Data yang tidak stasioner dalam ragam dapat ditangani dengan transformasi data, dilakukan dengan cara memangkatkan data *time series* dengan λ. Nilai λ dapat ditentukan berdasarkan nilai *rounded value*.

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

Pada metode *time series*, alat utama untuk mengidentifikasi model dari data yang akan diramalkan adalah dengan menggunakan *Autocorrelation Function* (ACF) dan *Partial Autocorrelation Function* (PACF) (Wei, 2006).

2.3.1 *Autocorrelation Function* (ACF)

Hipotesis untuk menguji koefisien ACF adalah sebagai berikut:

$H_0 = \rho_k = 0$ (Koefisien autokorelasi tidak berbeda secara signifikan)

$H_1 = \rho_k \neq 0$ (Koefisien autokorelasi berbeda secara signifikan)

Taraf signifikansi : $\alpha = 5\%$

Stasistik uji : $t = \frac{\rho_k}{SE\rho_k}$

dengan:

$\rho_k = \frac{cov(x_t, x_{t+k})}{\sqrt{var(x_t)var(x_{t+k})}}$ dan $SE(\rho_k) = \sqrt{\frac{1+2\sum_{j=1}^{k-1} r_j^2}{T}}$

$SE(\rho_k)$: *standard error* autokorelasi pada *lag* ke *k*

ρ_k : autokorelasi pada saat *lag* ke *k*

k : time *lag*

T : banyaknya observasi dalam data deret waktu

Kriteria keputusan :

Tolak H_0 jika nilai $|t_{hitung}| > t_{\alpha/2,df}$ dengan derajat bebas $df = T - 1, T$ merupakan banyaknya data dan *k* adalah *lag* koefisien ACF yang diuji.

2.3.2 *Partial Autocorrelation Function* (PACF)

Hipotesis untuk menguji koefisien PACF adalah sebagai berikut:

$H_0 = \rho_k = 0$

$H_1 = \rho_k \neq 0$

Taraf signifikansi : $\alpha = 5\%$

Stasistik uji : $t = \frac{\emptyset_{kk}}{SE(\emptyset_{kk})}$

dengan:

$SE(\emptyset_{kk}) = \frac{1}{T}$ dan $\phi_{kk} = Corr(x_t, x_{t+k} | x_{t+1}, x_{t+2}, \ldots, x_{t+k-1})$

Kriteria keputusan :

Tolak H_0 jika nilai $\left|t_{hitung}\right| > t_{\alpha/2,df}$dengan derajat bebas $df = T - 1, T$ merupakan banyaknya data dan *k* adalah *lag* PACF yang diuji.

2.4 Uji Signifikansi

Menurut Gujarati (2004) uji hipotesis (uji signifikansi) dilakukan untuk membuktikan kebenaran dari hipotesa yang diajukan dalam penelitian. Uji hipotesis yang dilakukan pada penelitian ini yaitu uji *t* (Uji Parsial). Uji *t* digunakan untuk menguji hubungan regresi secara parsial/individu. Pengujian ini dilakukan untuk mengukur tingkat signifikan setiap variabel *predictor* terhadap variabel respon dalam suatu model regresi.

Kriteria yang digunakan dalam uji *t* adalah sebagai berikut :

H_0 = variabel independen tidak signifikan mempengaruhi variabel dependen

H_1 = variabel independen signifikan mempengaruhi variabel dependen

Taraf signifikasi :$\alpha = 0{,}05$

Kriteria pengujiannya: nilai *prob.* t-hitung < α, maka tolak H_0

2.5 Analisis Regresi *Time Series* untuk Variasi Kalender

Regresi dalam konteks *time series* memiliki bentuk yang sama dengan regresi linier umum. Dengan mengasumsikan output atau bentuk respon y_t, yang dipengaruhi oleh kemungkinan data input atau *predictor*, dimana inputnya merupakan *fix* dan diketahui, hubungan ini dapat ditunjukkan dengan model regresi linier (Shumway dan Stoffer, 2006).

Model regresi linier untuk data yang memiliki variasi kalender adalah

$$y_t = \beta_0 + \beta_1 D_{1,t} + \beta_2 D_{2,t} + \cdots + \beta_n D_{n,t} + \varepsilon_t \qquad (2.1)$$

Dimana $D_{n,t}$ adalah variabel *dummy* untuk efek variasi kalender ke-*n*.

Metode Pendugaan parameter model regresi *dummy* menggunakan metode *Ordinary Least Square* (OLS). Prinsip dari OLS adalah menentukan nilai parameter-parameter yang tidak diketahui sehingga menghasilkan jumlah kuadrat sisaan yang bernilai seminimal mungkin (Gujarati, 2004).

2.6 Model *Time Series* Box Jenkins

Menurut Box dan Jenkins (1976), macam-macam model *time series* diantaranya adalah model *Autoregressive* (AR), *Moving-Average* (MA), *Autoregressive Moving-Average* (ARMA), dan Model ARMA yang melalui proses integrasi, dinyatakan sebagai *Autoregressive Integrated Moving Average* (ARIMA).

2.6.1 *Autoregressive* (AR)

Proses *Autoregressive* (AR) merupakan proses regresi diri menggunakan setiap nilai dalam seri fungsi linier dari nilai sebelumnya. Fungsi AR pertama menjelaskan bahwa hanya satu nilai sebelumnya yang digunakan sebagai fungsi dari nilai saat ini. Model AR dapat berjenjang 0, 1, 2, ..., sampai dengan *p*. Bentuk umum model AR dengan ordo *p* yaitu AR(*p*) dinyatakan sebagai berikut:

$$Y_t = \delta + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + \varepsilon_t \qquad (2.2)$$

dengan:

Y_t = nilai variabel pada waktu ke-*t*

δ = suatu konstanta

ϕ_p = parameter AR ordo ke *p*

ε_t = nilai kesalahan pada saat *t*

2.6.2 *Moving Average* (MA)

Moving Average (MA) disebut juga sebagai proses rata-rata bergerak. MA menjelaskan bahwa kesalahan pada nilai saat ini dipengaruhi oleh kesalahan pada periode sebelumnya. $\mathrm{Y_t}$ diperoleh dengan menerapkan bobot 1, $-\theta_1, -\theta_2, \ldots, -\theta_q$ untuk $\varepsilon_t, \varepsilon_{t-1}, \varepsilon_{t-2}, \ldots, \varepsilon_{t-q}$ sehingga untuk mendapatkan Y_{t-1} penerapannya menjadi $\varepsilon_{t+1}, \varepsilon_t, \varepsilon_{t-1}, \ldots, \varepsilon_{t-q+1}$. Model MA berjenjang 0,1,..., sampai dengan jenjang *q*. Bentuk umum model MA dengan ordo *q* yaitu MA(*q*) dinyatakan sebagai berikut:

$$Y_t = \delta + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q} \qquad (2.3)$$

dengan:

Y_t = nilai variabel pada waktu ke-*t*

δ = suatu konstanta

θ_q = parameter MA ordo ke-*q*

ε_{t-q} = nilai kesalahan pada saat *t*-*q*

2.6.3 *Autoregressive Moving Average* (ARMA)

Gabungan antara model AR(p) dan model MA(q) disebut dengan model *Autoregressive Moving Average* (ARMA). Data *time series* stasioner Y_t dikatakan mengikuti model ARMA (p,q) apabila:

$$Y_t = \delta + \phi_1 Y_{t-1} + \cdots + \phi_p Y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \cdots - \theta_q \varepsilon_{t-q} \qquad (2.4)$$

Parameter ϕ dan θ merupakan parameter proses regresi diri dan rataan bergerak yang stasioner dan *invertible* serta ε_t adalah *white noise*. Model tersebut juga dapat dituliskan secara singkat sebagai berikut:

$$\phi_p(B) Y_t = \delta + \theta_q(B) \varepsilon_t \qquad (2.5)$$

2.6.4 *Autoregressive Integrated Moving Average* (ARIMA)

Data *time series* dikatakan mengikuti model ARIMA jika dilakukan penanganan dengan pembedaan orde ke-d, sehingga $W_t = (1 - \mathrm{B})^d Y_t$ adalah proses yang stasioner. Menurut Cryer dan Chan (2008), jika W_t mengikuti model ARMA (p,q), dapat dikatakan bahwa Y_t adalah proses ARIMA (p,d,q). Bentuk umum model ARIMA (p,d,q) adalah:

$$\phi_p(B)(1 - B)^d Y_t = \delta + \theta_q(B) \varepsilon_t \qquad (2.6)$$

2.7 Proses *White Noise*

Suatu proses $\{\varepsilon_t\}$ disebut proses *white noise* jika data terdiri dari variabel acak yang tidak berkorelasi dan berdistribusi normal. Proses *white noise* dapat dideteksi menggunakan uji autokorelasi residual pada analisis *error*-nya. Uji korelasi residual digunakan untuk mendektesi ada tidaknya korelasi residual antar *lag*. Uji statistik yang dapat digunakan untuk mendeteksi adanya autokorelasi adalah uji Ljung *Box-Pierce*.

Langkah-langkah pengujian korelasi residual yaitu :

H_0 = Tidak ada autokorelasi antar sisaan

H_1 = Terdapat autokorelasi antar sisaan

Statistik uji Ljung-Box adalah:

$$Q_{LB} = n(n+2)\sum_{k=1}^{K}\frac{\hat{r}_1^2}{n-1} \sim \chi^2_{(K-p-q)} \tag{2.7}$$

dengan:

n : banyaknya amatan

$\hat{r}_t^2$: korelasi diri sisaan ke t dengan t+k

K : besarnya lag pada pengujian

Hipotesis nol ditolak ketika $Q_{LB} > \chi^2_{(K-p-q)}$ atau nilai-p $Q_{LB} < \alpha$. (Wei, 2006)

2.8 *Autoregressive Integrated Moving Average with Exogenous* (ARIMAX)

Model *Autoregressive Integrated Moving Average with Exogenous* (ARIMAX) adalah model ARIMA dengan peubah tambahan (Cryer dan Chan, 2008). Peubah tambahan yang digunakan untuk data deret waktu dengan variasi kalender berupa variabel *dummy*. Pada pemodelan ini, variabel *dummy* bernilai 1 untuk waktu-waktu terjadinya hari khusus dan bernilai 0 untuk waktu-waktu selainnya.

Model awal dibangun dengan pemodelan regresi linier dengan variabel *dummy* sebagai peubah prediktor yang akan menghasilkan sisaan. Menurut Box *et al*. (2015), sisaan yang dihasilkan dari model regresi deret waktu dapat dipastikan saling berkorelasi, sehingga penanganan yang tepat yaitu melalui pemodelan ARIMA pada sisaan model regresi. Model ARIMAX untuk variasi kalender dapat dituliskan sebagai berikut:

$$Y_t = \beta_0 + \beta_1 D_{1,t} + \beta_2 D_{2,t} + \cdots + \beta_n D_{n,t} + \frac{\theta_q(B)\varepsilon_t}{\phi_p(B)(1-B)^d} \quad (2.8)$$

dengan:

Y_t	: kombinasi linier dari gabungan pengamatan dan sisaan pada waktu-waktu sebelumnya
$D_{1,t}, D_{2,t}, \dots, D_{n,t}$	: variabel *dummy* hari-hari khusus
$\beta_1, \beta_2, \dots, \beta_n$	: parameter variabel *dummy* hari-hari khusus
ϕ_p	: parameter regresi diri ordo ke-p
θ_q	: parameter rataan bergerak ordo ke-q
$\phi_p(B)$	: $\left(1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p\right)$
$\theta_q(B)$	: $\left(1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q\right)$

Model-model ARIMA yang sudah diidentifikasi orde AR dan MA selanjutnya dibentuk model ARIMA. Kemudian model-model ARIMA tersebut akan dipilih untuk mendapatkan model ARIMAX terbaik. Pendugaan parameter model menggunakan metode *Ordinary Least Square* (OLS) dan uji signifikansi menggunakan uji-*t*. Selanjutnya dilakukan pemeriksaan asumsi *white noise* menggunakan uji Ljung-*Box Pierce*.

2.9 Peramalan

Peramalan pada dasarnya merupakan proses menyusun informasi tentang kejadian masa lampau yang berurutan untuk menduga kejadian di masa depan. Peramalan pada umumnya digunakan untuk memprediksi sesuatu yang kemungkinan besar akan terjadi misalnya kondisi permintaan, banyaknya curah hujan, kondisi ekonomi, dan lain-lain. Atas dasar logika, langkah dalam metode peramalan secara umum adalah mengumpulkan data, menyeleksi dan memilih data, memilih model peramalan, menggunakan model terpilih untuk melakukan peramalan, evaluasi hasil akhir (Frechtling, 2001).

2.10 Kriteria Kebaikan Model

2.10.1 Kriteria Kebaikan Model pada Data Pemodelan

Menurut Pesaran dan Shin (1997), kriteria kebaikan model dengan pendekatan data pemodelan terdiri dari *Akaike's Information Criterion* (AIC) dan *Schwartz's Bayesian Criterion* (SBC) yang mengikuti persamaan berikut:

$$AIC(M) = n\ln\left(\hat{\sigma}_a{}^2\right) + 2M \quad (2.9)$$

$$SBC(M) = n\ln\left(\hat{\sigma}_a{}^2\right) + M\ln(n) \quad (2.10)$$

dengan:

$\hat{\sigma}_a{}^2$ = Estimasi Maksimum Likelihood untuk σ

M = Jumlah parameter yang ditaksir

n = Jumlah pengamatan data pemodelan.

2.10.2 Kriteria Kebaikan Model pada Data Validasi

Menurut Makridakis dan Hibon (2000), kriteria kebaikan model pada data validasi terdiri dari kriteria *Root Mean Square Error* (RMSE) dan *Mean Absolute Percentage Error* (MAPE) yang mengikuti persamaan berikut

$$RMSE = \sqrt{\frac{1}{L}\sum_{i=1}^{L} e_l{}^2} \tag{2.11}$$

$$MAPE = \sqrt{\frac{1}{L}\sum_{i=1}^{L} \left|\frac{e_l}{y_{n+l}}\right|} \tag{2.12}$$

dengan:

e_l = $y_{n+l} - \hat{y}_n(l)$

L = Jumlah pengamatan data validasi

y_{n+l} = Nilai data aktual pada ramalan ke-*l*

III. METODOLOGI PENELITIAN

3.1 Waktu dan Tempat Penelitian

Penelitian ini dilakukan pada semester ganjil tahun ajaran 2018/2019, yang bertempat di Jurusan Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Lampung.

3.2 Data Penelitian

Data yang digunakan dalam penelitian ini adalah data sekunder yang bersumber dari buku Lampung dalam Angka tahun 2000 sampai dengan 2019 berupa jumlah penumpang kereta api yang berangkat dari Stasiun Tanjung Karang, Bandar Lampung pada bulan Januari tahun 2000 sampai dengan bulan Maret 2019, yang diperoleh dari Badan Pusat Statistika Provinsi Lampung Tahun 2018. Data yang dilibatkan dalam penelitian ini terbagi menjadi dua, yaitu data pemodelan dan data validasi. Data pemodelan yang digunakan sebanyak 228 observasi yaitu data pada bulan Januari 2000 hingga Desember 2018, dan data validasi yang digunakan sebanyak 3 observasi yaitu data dari bulan Januari 2019 hingga Maret 2019. Perangkat lunak yang digunakan untuk membantu proses komputasi dalam penelitian ini adalah program Minitab dan Eviews.

3.3 Metode Penelitian

Langkah penyelesaian analisis dengan menggunakan model ARIMAX adalah sebagai berikut:

1. Mengkaji model ARIMAX

Parameter-parameter model ARIMAX akan dikaji dengan menggunakan metode pendugaan parameter *Ordinary Least Square (OLS)*.

2. Menentukan variabel *dummy* untuk model regresi.

Variabel *dummy* untuk variasi kalender dalam penelitian ini ditentukan dengan melihat pola data dari plot jumlah penumpang kereta api Stasiun Tanjung Karang

3. Melakukan pengujian signifikansi parameter model regresi *dummy*.

Signifikansi parameter dilakukan menggunakan statistik uji t. Jika ada parameter yang tidak signifikan, maka dilakukan estimasi ulang dengan tidak menyertakan parameter yang tidak signifikan. Sehingga diperoleh model regresi *dummy* dengan semua parameter yang signifikan.

4. Melakukan *diganostic white noise* pada residual regresi variabel *dummy*.

Jika sudah memenuhi asumsi *white noise*, maka dilanjutkan dengan peramalan. Namun jika tidak *white noise* maka dilanjutkan dengan pendugaan model ARIMA pada residual.

5. Melakukan uji stasioneritas

Uji stasioneritas dilakukan untuk mengetahui apakah data sudah stasioner baik dalam varians ataupun mean, untuk menguji secara varians maka kita menggunakan metode transformasi Box-cox, sedangkan untuk menguji secara mean dapat dilakukan dengan uji stasioneritas menggunakan uji akar unit *Augmented Dickey Fuller* (ADF). Jika tidak stasioner maka lakukan *differencing.*

6. Pendugaan model ARIMA

Identifikasi model ARIMA dapat dilakukan dengan melakukan plot fungsi ACF dan plot fungsi PACF pada data yang telah stasioner. Identifikasi model ini akan memperoleh model ARIMA (p,d,q) yang memungkinkan.

7. Melakukan pengujian signifikansi parameter model ARIMA

Melakukan uji signifikansi parameter-parameter dari model ARIMA menggunakan uji t.

8. Pemilihan model ARIMA terbaik

Pemilihan model terbaik dengan AIC dan SBC berdasarkan nilai yang paling kecil.

9. Pemodelan ARIMAX

Model ARIMA yang terbaik yang diperoleh dari langkah 8, diestimasi secara simultan dengan model regresi *dummy* yang telah diperoleh sebelumnya.

10. Melakukan pengujian signifikansi parameter model ARIMAX

Melakukan uji signifikansi parameter-parameter dari model ARIMAX menggunakan uji t.

11. Melakukan *diganostic white noise* pada residual model ARIMAX.

Jika sudah memenuhi asumsi *white noise*, maka dilanjutkan dengan peramalan.

12. Peramalan data menggunakan model ARIMAX

Model ARIMAX yang telah memenuhi asumsi *white noise* selanjutnya akan dilakukan peramalan pada data tersebut dengan model ARIMAX.

13. Pengepasan data validasi dengan data hasil peramalan

Data validasi digunakan untuk mengetahui seberapa baiknya hasil peramalan model ARIMAX dibandingkan dengan data sebenarnya.

14. Mengukur kebaikan peramalan dengan RMSE dan MAPE

Setelah didapatkan hasil peramalannya, langkah terakhir adalah melihat apakah peramalan itu cukup baik untuk digunakan, yaitu dengan melihat nilai dari RMSE dan MAPE.

V. KESIMPULAN

Berdasarkan hasil dan pembahasan pada penelitian ini dapat diambil kesimpulan yaitu:

1. Penduga parameter untuk ϕ_1 dari model variasi kalender AR (1) adalah:

$$\hat{\phi}_1 = \frac{\sum_{t=2}^{n}(Y_t - D_t)(Y_{t-1} + D_{t-1})}{\sum_{t=2}^{n} 2\,(Y_{t-1} + D_{t-1})^2}$$

   Penduga parameter untuk (θ_1) dari model variasi kalender MA(1) yaitu :

$$\hat{\theta}_1 = \frac{\sum_{t=2}^{n}(D_t)(\varepsilon_{t-1}) - \sum_{t=2}^{n}(Y_t)(\varepsilon_{t-1})}{\sum_{t=2}^{n}(\varepsilon_{t-1})^2}$$

2. Model terbaik dari data jumlah penumpang kereta api Stasiun Tanjung Karang dengan penambahan variabel *dummy* adalah ARIMAX (1,1,12) dengan bentuk matematis sebagai berikut:

$$Y_t = 4712.060D_{1,t} + \frac{0.507675_1(B)\varepsilon_t}{-0.513009_{12}(B)(1-B)^1}$$

3. Nilai peramalan dari data jumlah penumpang kereta api stasiun tanjung karang yang diperoleh untuk 12 bulan selama tahun 2019 berturut-turut adalah: 81525, 73090, 83094, 86243, 89675, 76717, 96458, 93736, 88019, 86639, 88249 dan 96126.

DAFTAR PUSTAKA

Badan Pusat Statistik. 2018. *Provinsi Lampung dalam Angka Tahun 2000-2018.* Badan Pusat Statistik Provinsi Lampung, Bandar Lampung,

Box, G. E. P., dan Jenkins, G. M. 1976. *Time series analysis: Forecasting and control*. San Holden-Day, Francisco.

Box, G.E.P., Jenkins, G.M., Reinsel, G.C.,dan Ljung, G.M. 2015. *Time Series Analysis: Forecasting and Control*. J Wiley, New Jersey (US).

Cryer, J. D. dan Chan, K. S. 2008. *Time Series Analysis: With Applications in R* (2nd edition.). Springer Science, New York.

Frechtling, D. C. 2001. *Forecasting Tourism Demand: Methods and Strategies*, Butterworth-Heinemann, Oxford.

Gujarati, D. N. 2004. *Basic Econometrics*. The-McGraw Hill Company.

Makridakis, S. dan Hibon, M. 2000. The M3-Competition: Result, Conclusion and Implications. *International Journal of Forecasting* 16(1):451-476.

Montgomery, D.C., Jennings, C.L.dan Kulahci, M. 2008. *Introduction to Time Series Analysis and Forecasting*. J Wiley , New Jersey (US).

Mulyana. 2004. *Buku Ajar Analisis Deret Waktu*. Universitas Padjajaran FMIPA, Bandung.

Nachrowi, D.N. dan Usman H. 2006. *Pendekatan Populer dan Praktis Ekonometrika untuk Analisis Ekonomi dan Keuangan.* Universitas Indonesia, Jakarta.

Pesaran, M. H. dan Shin, Y. 1997. An Autoregressive Distributed Lag Modelling Approach to Cointegration Analysis. *Symposium at the Centennial of Ragnar Frisch, The Norwegian Academy of Science and Letters,* Oslo, March 3-5.

Shumway, R. H.dan Stoffer, D. S. 2006. *Time Series Analysis and Its Applications (With R Examples, Second Edition)* . Springer, New York.

Rosadi, D. 2011. *Ekonometrika & Analisis Runtun Waktu Terapan dengan Eviews*. Andi, Yogyakarta.

Suhartono, Lee, M. H., dan Hamzah, N. A. 2010. Calendar Variation Model Based on Time Series Regression for Sales Forecast : The Ramadhan Effects. *Proceedings of the Regional Conference on Statistical Sciences*, 30-41.

Wei, W. W. S. 2006. *Time series Analysis Univariate and Multivariate Methods.* Pearson education, Inc, New York.