

## *Jurnal Ilmiah Teori dan Aplikasi Statistika*

- ❑ **Peramalan Data Ekspor Non Migas Provinsi Kalimantan Timur Menggunakan Univariate Time Series**  
Farida Nur Hayati<sup>1</sup>, Diana Nurlaily<sup>2</sup>, Elly Pusporani<sup>3</sup>  
Institut Teknologi Kalimantan
- ❑ **Pemanfaatan Satelit Himawari-8 Untuk Estimasi Curah Hujan Dengan Metode Convective Stratiform Technique (CST) Dan Modified Convective Stratiform Technique (Mcsst) Di Wilayah Ekuatorial Dan Monsunal (Studi Kasus Sulawesi 2020)**  
Nur Habib Muzaki<sup>1</sup>, Eriska Febriati<sup>2</sup>, Yosafat Donni Haryanto<sup>3</sup>  
Sekolah Tinggi Meteorologi Klimatologi dan Geofisika
- ❑ **Pemodelan Kasus Kumulatif Covid-19 di Pulau Jawa dan Bali Dengan Pendekatan Multiple Classification Analysis (MCA)**  
Rezky Yayang Yakhamid<sup>1</sup>, Amelia Tri Wahyuni<sup>2</sup>, Nadidah Pengestika<sup>3</sup>, Hanifah<sup>4</sup>, Putu Adi Myarsithawan<sup>5</sup>,  
Risni Julaeli Yuhan<sup>6</sup>  
Politeknik Statistika STIS
- ❑ **Geographically Weighted Spline Nonparametric Regression dengan Fungsi Pembobot Bisquare dan Gaussian Pada Tingkat Pengangguran Terbuka Di Pulau Kalimantan**  
Hillidatul Ilmi<sup>1</sup>, Sifriyani<sup>2</sup>, Surya Prangga<sup>3</sup>  
Universitas Mulawarman
- ❑ **Analisis Multilevel Kemiskinan Rumah Tangga Pertanian di Provinsi Sulawesi Selatan Tahun 2020**  
Vollandio Ardhian Rastantra  
Badan Pusat Statistik Kabupaten Bulukumba
- ❑ **Analisis Volatilitas Harga Saham Sekor Minyak dan Gas di Indonesia pada Masa Pandemi Covid-19 dengan Metode ARIMA-GARCH**  
Nanda Septiana<sup>1</sup>, Primadina Hasanah<sup>2</sup>, Annisa Rahmita Soemarsono<sup>3</sup>  
Institut Teknologi Kalimantan
- ❑ **Analisis Kalman filter berbasis Google Trends untuk Prediksi Kedatangan Wisatawan Mancanegara di Indonesia Pasca Pandemi**  
Evita Purnaningrum<sup>1</sup>, Hanief Khoyyir Nafah<sup>2</sup>  
Universitas PGRI Adi Buana Surabaya

<b>Penanggung Jawab</b>	
Alfisyahrina Hapsery, S.Si., M.Si	(Universitas PGRI Adi Buana Surabaya)
<b>Pimpinan Redaksi</b>	
Muhammad Athoillah, S.Si., M.Si	(Universitas PGRI Adi Buana Surabaya)
<b>Editor Pelaksana</b>	
Sari Cahyaningtias, S.Si., M.Si	(Universitas PGRI Adi Buana Surabaya)
Fenny Fitriani, S.Si., M.Si	(Universitas PGRI Adi Buana Surabaya)
Nur Silviyah Rahmi, S.Si., M.Si	(Universitas PGRI Adi Buana Surabaya)
Sekar Utami Wijaya S.Stat., M.Si	(Universitas PGRI Adi Buana Surabaya)
<b>Mitra Bestari (Reviewer)</b>	
Dr.rer.pol. Dedy Dwi Prastyo, M.Si	(Institut Teknologi Sepuluh Nopember)
Dr. Drs. Agus Suharsono, M.S	(Institut Teknologi Sepuluh Nopember)
Dr. Bambang Widjanarko Otok	(Institut Teknologi Sepuluh Nopember)
Novri Suhermi, S.Si., M.Si., M.Sc	(Institut Teknologi Sepuluh Nopember)
Shofi Andari, S.Stat., M.Si	(Institut Teknologi Sepuluh Nopember)
Dr. RB Fajriya Hakim, S.Si., M.Si	(Universitas Islam Indonesia)
A'yunin Sofro, S.Si., M.Si., Ph.D.	(Universitas Negeri Surabaya)
Arief Rachman Hakim, S.Si., M.Si	(Universitas Diponegoro)
Dani Al Mahkya, S.Si., M.Si	(Sains Aktuaria Institut Teknologi Sumatera)
Dr. Sri Harini	(Universitas Islam Negeri Maulana Malik Ibrahim)
Dr. Faula Arina, M.Si	(Universitas Sultan Agung Tirtayasa)
Artanti Indrasetimingsih, S.Si., M.Si	(Universitas PGRI Adi Buana Surabaya)
Gangga Anuraga, S.Si., M.Si	(Universitas PGRI Adi Buana Surabaya)
Winda Aprianti, S.Si., M.Si	(Politeknik Negeri Tanah Laut)

J Statistika: **Jurnal Ilmiah Teori dan Aplikasi Statistika** merupakan media penerbitan dan pembahasan karya ilmiah dalam bidang ilmu statistika beserta aplikasinya, baik berupa hasil penelitian, bahasan tentang teori, metodologi, komputasi maupun aplikasi statistika dalam bidang lain. Jurnal ini akan terbit dua kali dalam setahun setiap bulan Juli dan Desember.

Redaksi mengundang para pakar dan praktisi dari dalam dan luar lingkungan Universitas PGRI Adi Buana Surabaya untuk menuliskan karya ilmiahnya yang relevan dengan bidang ilmu statistika. Naskah dikirim pada website Jstatistika <http://jurnal.unipasby.ac.id/> dengan format A4 margin 4-3-3-3 times new roman 11, spasi 1, disertai biodata penulis. Email redaksi: [jstat@unipasby.ac.id](mailto:jstat@unipasby.ac.id). Cp: 085645358346

## Pengantar Redaksi

Syukur Alhamdulillah bahwa Jurnal Ilmiah: JStatistika, Program Studi Statistika Universitas PGRI Adi Buana Surabaya volume 14 nomor 2 dapat terbit bulan Desember 2021 ini. Berbagai hambatan dapat kita atasi, semoga hambatan-hambatan tersebut tidak akan terjadi lagi pada penerbitan selanjutnya. Jurnal Ilmiah JStatistika terbitan kali ini memuat artikel yang membahas tentang **“Peramalan Data Ekspor Non Migas Provinsi Kalimantan Timur Menggunakan Univariate Time Series; Peramalan Data Ekspor Non Migas Provinsi Kalimantan Timur Menggunakan Univariate Time Series; Pemodelan Kasus Kumulatif Covid-19 di Pulau Jawa dan Bali Dengan Pendekatan Multiple Classification Analysis (MCA); Geographically Weighted Spline Nonparametric Regression dengan Fungsi Pembobot Bisquare dan Gaussian Pada Tingkat Pengangguran Terbuka Di Pulau Kalimantan; Analisis Multilevel Kemiskinan Rumah Tangga Pertanian di Provinsi Sulawesi Selatan Tahun 2020; Analisis Volatilitas Harga Saham Sekor Minyak dan Gas di Indonesia pada Masa Pandemi Covid-19 dengan Metode ARIMA-GARCH; Analisis Kalman filter berbasis Google Trends untuk Prediksi Kedatangan Wisatawan Mancanegara di Indonesia Pasca Pandemi”**.

Jurnal Ilmiah Statistika menerima artikel ilmiah dari hasil penelitian, laporan / studi kasus, kajian/ tinjauan pustaka, maupun penyegar ilmu statistika, yang berorientasi pada kemuthakhiran ilmu pengetahuan statistika agar dapat menjadi sumber informasi ilmiah yang mampu memberikan kontribusi dalam mengatasi permasalahan yang semakin kompleks. Redaksi mengundang berbagai ilmuwan dari berbagai lembaga pendidikan tinggi maupun penelitian untuk memberikan sumbangan ilmiahnya, baik berupa hasil penelitian maupun kajian ilmiah dalam bidang statistika dan aplikasinya.

Redaksi sangat mengharapkan masukan-masukan dari para pembaca, professional bidang statistika, atau yang terkait dengan penerbitan, demi semakin meningkatnya kualitas jurnal sebagaimana harapan kita bersama. Semoga artikel ilmiah yang termuat dalam Jurnal Ilmiah JStatistika bermanfaat bagi para akademisi dan professional yang berkecimpung dalam dunia statistika dan aplikasinya.

Surabaya, Desember 2021

**Pimpinan Redaksi**

## DAFTAR ISI

HALAMAN JUDUL

PENGANTAR REDAKSI

DAFTAR ISI

**Peramalan Data Ekspor Non Migas Provinsi Kalimantan Timur Menggunakan Univariate Time Series**

Farida Nur Hayati<sup>1</sup>, Diana Nurlaily<sup>2</sup>, Elly Pusporani<sup>3</sup>  
Institut Teknologi Kalimantan

59-66

**Pemanfaatan Satelit Himawari-8 Untuk Estimasi Curah Hujan Dengan Metode Convective Stratiform Technique (CST) Dan Modified Convective Stratiform Technique (McsT) Di Wilayah Ekuatorial Dan Monsunal (Studi Kasus Sulawesi 2020)**

Nur Habib Muzaki<sup>1</sup>, Eriska Febriati<sup>2</sup>, Yosafat Donni Haryanto<sup>3</sup>  
Sekolah Tinggi Meteorologi Klimatologi dan Geofisika

67-73

**Pemodelan Kasus Kumulatif Covid-19 di Pulau Jawa dan Bali Dengan Pendekatan Multiple Classification Analysis (MCA)**

Rezky Yayang Yakhamid<sup>1</sup>, Amelia Tri Wahyuni<sup>2</sup>, Nadidah Pengestika<sup>3</sup>, Hanifah<sup>4</sup>, Putu Adi Myarsithawan<sup>5</sup>, Risni Julaeni Yuhan<sup>6</sup>  
Politeknik Statistika STIS

74-83

**Geographically Weighted Spline Nonparametric Regression dengan Fungsi Pembobot Bisquare dan Gaussian Pada Tingkat Pengangguran Terbuka Di Pulau Kalimantan**

Hillidatul Ilmi<sup>1</sup>, Sifriyani<sup>2</sup>, Surya Prangga<sup>3</sup>  
Universitas Mulawarman

84-92

**Analisis Multilevel Kemiskinan Rumah Tangga Pertanian di Provinsi Sulawesi Selatan Tahun 2020**

Volandio Ardhian Rastantra  
Badan Pusat Statistik Kabupaten Bulukumba

93-98

**Analisis Volatilitas Harga Saham Sekor Minyak dan Gas di Indonesia pada Masa Pandemi Covid-19 dengan Metode ARIMA-GARCH**

Nanda Septiana<sup>1</sup>, Primadina Hasanah<sup>2</sup>, Annisa Rahmita Soemarsono<sup>3</sup>  
Institut Teknologi Kalimantan

99-109

**Analisis Kalman filter berbasis Google Trends untuk Prediksi Kedatangan Wisatawan Mancanegara di Indonesia Pasca Pandemi**

Evita Purnaningrum<sup>1</sup>, Hanief Khooyir Nafah<sup>2</sup>  
Universitas PGRI Adi Buana Surabaya

110-115

---

## Peramalan Data Ekspor Non Migas Provinsi Kalimantan Timur Menggunakan Univariate Time Series

Farida Nur Hayati<sup>(1)</sup>, Diana Nurlaily<sup>(2)</sup>, Elly Pusporani<sup>(3)</sup>

Institut Teknologi Kalimantan

Jl. Soekarno Hatta Km. 15 Karang Joang, Telp./Fax. 0542-8530800, Balikpapan Kalimantan Timur, 76127

e-mail: farida.nur@lecturer.itk.ac.id, diana.nurlaily@lecturer.itk.ac.id, elly.pusporani@gmail.com

---

### ABSTRAK

Kegiatan ekspor sangat penting untuk mengetahui jenis perekonomian negara dikatakan perekonomian terbuka atau tertutup. Semakin tinggi nilai ekspor suatu negara maka akan semakin terbuka perekonomiannya. Kegiatan ekspor juga menjadi penggerak perekonomian suatu negara. Mengingat pentingnya hal tersebut maka, diperlukan kebijaksanaan meningkatkan nilai ekspor khususnya pada komoditi non migas. Komoditi non migas dipilih pada penelitian ini karena adanya kemerosotan harga minyak bumi di pasaran dunia. Sehingga untuk mengatasi ketergantungan penerimaan negara yang berasal dari minyak dan gas bumi maka pemerintah berusaha meningkatkan ekspor komoditi dan jasa-jasa non migas. Kondisi ekspor luar negeri Kalimantan Timur didominasi oleh ekspor non migas. Oleh karena itu, prediksi atau peramalan nilai dari komoditas ekspor non migas di Kalimantan Timur dapat digunakan untuk mengetahui devisa/ sejumlah uang pada lingkup internasional yang akan diperoleh, agar dapat digunakan sebagai bahan pertimbangan dalam menentukan kebijakan oleh pemerintah. Pada penelitian ini, untuk meramalkan nilai ekspor non migas digunakan metode *Autoregresif Integrated Moving Average* (ARIMA). Langkah pertama dalam melakukan analisis metode tersebut adalah identifikasi model dengan cara melihat ACF dan PACF, selanjutnya melakukan penaksiran serta uji signifikansi parameter, setelah itu dilakukan diagnostic checking, dan yang terakhir adalah peramalan/prediksi. Hasil penelitian peramalan data ekspor menggunakan ARIMA menunjukkan bahwa model ARIMA yang paling baik dalam meramalkan data ekspor non migas di Kalimantan Timur berdasarkan nilai sMAPE, dan RMSE adalah ARIMA (0,1,[1,12]).

**Kata kunci :** *Ekspor, Ekspor non migas, ARIMA, RMSE, sMAPE*

### ABSTRACT

*Exports are very important to know the type of economy a country can be said to be open or closed economy. The higher the export value of a country, the more open the economy will be. Export activities are also a driver of a country's economy. Given the importance of this, discretion is needed in increasing exports, especially non-oil and gas commodities. Non-oil and gas commodities were chosen in this study because of the slump in petroleum prices in the world market. So as to overcome the dependence on state revenues derived from oil and gas, the government seeks to increase exports of commodities and non-oil and gas services. East Kalimantan's overseas export conditions are dominated by non-oil and gas exports. Therefore, prediction or forecasting of the value of non-oil and gas export commodities in East Kalimantan can be used to determine foreign exchange / amount of money in the international sphere to be obtained, so that it can be used as a consideration by the government to determine policy. In this study the method used to predict the value of non-oil and gas exports is Autoregressive Integrated Moving Average (ARIMA). The first step in analyzing ARIMA method is to identify the model by looking at ACF and PACF, then perform the assessment and parameter significance test, after that diagnostic checking parameter is obtained, and the last is forecasting/ prediction. The results of export data forecasting using ARIMA show that ARIMA model is best to forecast non-oil and gas export data in East Kalimantan based on sMAPE value, and RMSE is ARIMA (0.1,[1,12]).*

**Keywords :** *Non-oil and gas Export, ARIMA, RMSE, sMAPE*

## 1. PENDAHULUAN

Kegiatan ekspor sangat penting untuk mengetahui jenis perekonomian negara bisa dikatakan perekonomian terbuka atau tertutup. Semakin tinggi nilai ekspor suatu negara maka akan semakin terbuka perekonomiannya. Kegiatan ekspor juga menjadi penggerak perekonomian suatu negara. Hal ini membuat antar negara berusaha melakukan peningkatan kuantitas dan kualitas ekspornya agar laku dan lebih efisien di pasar Internasional (Badan Pusat Statistika, 2020). Perekonomian berbasis ekspor mempunyai kelebihan diantaranya adalah: kegiatan ekspor, dapat menyerap tenaga kerja, selain itu, dapat menghasilkan cash inflow berupa valuta asing sebagai bentuk pembayaran atas produk yang dijual ke luar negeri, terutama untuk ekspor non migas, kegiatan ekspor juga dapat memudahkan suatu negara untuk mencapai kemandiriannya dalam perekonomian (Wardhana, 2011). Karenanya diperlukan analisis untuk membuat suatu kebijaksanaan dalam meningkatkan nilai ekspor suatu barang khususnya komoditi non migas. Komoditi non migas tersebut dipilih karena adanya penurunan harga minyak bumi di pasaran dunia. Sehingga untuk mengatasi ketergantungan penerimaan negara yang berasal dari minyak dan gas bumi maka pemerintah berusaha meningkatkan ekspor komoditi dan jasa-jasa non migas (Huda, 2006). Menurut Bank Indonesia (2009), sektor non migas adalah sektor yang memberikan sumbangan devisa terbesar untuk Indonesia. Selama periode 2013-2020, peranan ekspor nonmigas di Indonesia mencapai kisaran angka 82% sampai 94%, sedangkan peranan ekspor migas berkisar antara 5% sampai 17% (Badan Pusat Statistika, 2020).

Komoditas sektor non migas dibagi menjadi kelompok dagangan umum dan kelompok barang lainnya seperti emas nonmoneter. Kelompok barang dagangan umum meliputi pertanian, sektor industry, sektor pertambangan dan sektor barang lainnya (Lestari, Widiharih, & Rahmawati, 2018). Di Kalimantan Timur, Komoditi ekspornya didominasi oleh komoditi ekspor batu bara dan turuannya, diikuti oleh produksi cpo dan sebagainya. Hal ini menunjukkan bahwa di Kalimantan Timur, ekspor non migas merupakan penyumbang terbesar nilai ekspor. Berdasarkan perkembangan ekspor non migas di Kalimantan Timur tahun 2011 sampai tahun 2017, pada tahun 2014 terjadi penurunan jumlah ekspor non migas yang tinggi jika dibandingkan jumlah ekspor pada tahun 2013. Penurunan jumlah ekspor non migas terus terjadi sampai tahun 2016, dimana tahun 2017 jumlah

ekspor non migas di Kalimantan Timur sudah mulai meningkat. Penurunan jumlah ekspor non migas ini disebabkan adanya pembatasan ekspor batu bara oleh pemerintah pusat (Nurmawaddah, 2019).

Prediksi atau peramalan nilai dari komoditas ekspor non migas di Kalimantan Timur dapat digunakan untuk mengetahui devisa /sejumlah uang pada lingkup internasional yang akan diperoleh sebagai bahan pertimbangan oleh pemerintah untuk menentukan kebijakan. Berdasarkan hal tersebut, diperlukan metode analisis yang tepat untuk mendapatkan prediksi data yang paling sesuai. Banyak penelitian berkaitan peramalan ekspor pernah dikaji sebelumnya diantaranya oleh Ruslan, Harahap, dan Sembiring pada tahun 2013 untuk meramalkan nilai ekspor di Provinsi Sumatera utara dengan menggunakan metode ARIMA Box Jenkins, dan menghasilkan model ARIMA (1,0,1). Amalina, (2019), juga pernah melakukan penelitian dengan meramalkan nilai ekspor migas dan non migas menggunakan Artificial Neural yang menghasilkan parameter yang paling optimal untuk meramalkan data ekspor dengan MAPE bernilai baik. Lestari, dkk (2018), juga meramalkan ekspor non migas menggunakan kalender variansi islam, dimana berdasarkan analisis didapatkan data ekspor non migas terdapat trend, pola musiman sehingga menggunakan model ARIMA. Pada penelitian ini, metode yang digunakan untuk melakukan *forecast*/peramalan data ekspor non migas adalah Autoregressif Integrated Moving Average (ARIMA). ARIMA merupakan salah satu metode yang digunakan untuk meramal suatu variabel dengan cepat, akurat dan sederhana. Langkah pertama dalam melakukan analisis metode ARIMA adalah mengidentifikasi model dengan cara melihat ACF dan PACF, selanjutnya melakukan penaksiran parameter serta uji signifikansi, setelah itu dilakukan diagnostic checking, dan yang terakhir adalah peramalan/prediksi. Banyak penelitian yang berkaitan dengan metode ARIMA salah satunya yaitu Hartati (2017) yang meneliti tentang peralaman pergerakan laju inflasi dengan metode ARIMA. Hasil penelitian menunjukkan bahwa hasil ramalan menggunakan metode ARIMA mampu mengikuti pergerakan data actual laju inflasi. Berdasarkan hasil uraian tersebut, pada penelitian ini akan dilakukan peramalan nilai ekspor non migas yang ada di Kalimantan Timur menggunakan Metode ARIMA Box Jenkins. Diharapkan penelitian ini dapat memberikan informasi terkait prediksi nilai ekspor non migas di Kalimantan Timur sehingga dapat digunakan pemerintah dalam

mengambil kebijakan untuk meningkatkan ekspor di Kalimantan Timur.

2. METODE PENELITIAN

2.1 ARIMA

Autoregressive Integrated Moving Average adalah salah satu metode peramalan/prediksi untuk data time series. Metode ini biasa disebut dengan ARIMA. Berdasarkan Wei (2006) model ARIMA(p,d,q) dapat dituliskan sebagai berikut:

$$\phi_p(B)(1 - B)^d Z_t = \theta_q(B)a_t$$

dimana

(p, d, q) : p menyatakan orde AR (Autoregressive), q menyatakan orde MA (Moving Average) dan d menyatakan differencing untuk data yang bukan musiman

$\phi_p(B)$  : koefisien AR dengan orde p dimana  $\phi_p(B) = 1 - \phi_1(B) - \phi_2(B^2) - \dots - \phi_p(B^p)$

$\theta_q(B)$  : koefisien MA dengan orde q dimana  $\theta_q(B) = 1 - \theta_1(B) - \theta_2(B^2) - \dots - \theta_q(B^q)$

$(1 - B)^d$  : differencing orde d

$a_t$  : nilai residual di t

Dalam menggunakan metode ARIMA terdapat syarat yang harus terpenuhi yaitu data penelitian harus Stasioner terhadap varians dan means. Setelah data telah memenuhi syarat tersebut, baru analisis menggunakan metode ARIMA dapat digunakan. Berdasarkan Box & Jenkins diperlukan empat langkah dalam menganalisis data dengan metode ARIMA yaitu:

1. Mengidentifikasi Model

Dalam tahapan ini dilakukan identifikasi Model Awal ARIMA. Langkah Awal yang dilakukan yaitu pengecekan stasioneritas data kemudian dilakukan pengecekan orde AR atau MA dari data dengan menggunakan ACF yang merupakan kepanjangan dari "Autocorrelation Function" dan PACF "Partial Autocorrelation Function". ACF adalah ukuran korelasi antara  $Z_t$  dan  $Z_{t+k}$  pada data sedangkan PACF adalah ukuran korelasi antara  $Z_t$  dan  $Z_{t+k}$  setelah mendapat pengaruh dependensi linier pada variabel yang mengintervensi variabel  $Z_{t+1}, Z_{t+2}, \dots, Z_{t+k-1}$  dihilangkan (Wei, 2006):

2. Mengestimasi Parameter

Pada tahap ini nilai parameter akan ditentukan dengan melakukan estimasi model awal. OLS (Ordinary Least Square), MLE (Maximum Likelihood Estimation) dan lain-lain adalah beberapa metode estimasi yang bisa digunakan. Selanjutnya dilakukan uji untuk memeriksa signifikansi, untuk mengetahui apakah parameter yang ditaksir sudah signifikan atau belum. Hipotesis dari uji signifikansi ini adalah sebagai berikut (Wei, 2006):

$H_0 : \phi_p = 0$  atau  $\theta_q = 0$

$H_1 : \phi_p \neq 0$  atau  $\theta_q \neq 0$

Statistik Uji :

$$t = \frac{\hat{\phi}_p}{SE(\hat{\phi}_p)} \text{ atau } t = \frac{\hat{\theta}_q}{SE(\hat{\theta}_q)}$$

Daerah Penolakan  $H_0$  adalah jika  $|t| > t_{\frac{\alpha}{2}, n-p}$  dimana nilai n menunjukkan banyak observasi dari data time series dan p menunjukkan jumlah parameter suatu model.

3. Diagnostic Checking

Diagnostic Checking dalam ARIMA adalah pengecekan apakah aumsi residual yaitu White Noise dan Distribusi Normal terpenuhi. Pengujian asumsi white noise dapat dilakukan menggunakan pengujian Ljung-Box (Wei, 2006) dengan hipotesis sebagai berikut

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

$H_1 : \text{minimal ada satu } \rho_i \neq 0, \text{ untuk } i = 1, 2, \dots, k$

Statistik Uji :

$$Q = n(n + 2) \sum_{i=1}^k \frac{\hat{\rho}_i^2}{(n-i)}$$

$H_0$  akan ditolak apabila  $Q > \chi_{\alpha, df}^2$  atau nilai p-value  $< \alpha$ , dengan  $df=k-p-q$

Sedangkan untuk mengecek asumsi distribusi normal digunakan uji menggunakan Kolmogorov Smirnov. Hipotesis dalam pengujian ini dapat diketahui sebagai berikut (Daniel, 1989) :

$H_0 : F(x) = F_0(x)$  (residual dari data analisis berdistribusi normal)

$H_1 : F(x) \neq F_0(x)$  (residual dari data analisis tidak berdistribusi normal)

Statistik uji :

$$D = \text{Sup}|S(x) - F_0(x)|$$

$H_0$  akan ditolak apabila nilai  $D > D_{(1-\alpha, n)}$  atau p-value  $< \alpha$

4. Forecasting

Apabila parameter modelnya sudah signifikan dan memenuhi uji asumsi residual, maka tahapan forecast pada data dapat dilakukan.

**2.2 Kriteria Penentuan Model Terbaik**

Ukuran untuk menentukan model terbaik dalam melakukan peramalan/prediksi dapat menggunakan RMSE atau yang biasa disebut *Root Mean Square Error* dan sMAPE atau *Symmetric Mean Absolute Percentage Error*. Persamaan dari kedua metode tersebut adalah sebagai berikut (Gooijer & Hyndman, 2006).

$$RMSE = \sqrt{\frac{1}{M} \sum_{t=1}^M (Z_t - \widehat{Z}_t)^2}$$

$$sMAPE = \sqrt{\frac{1}{M} \sum_{t=1}^M \frac{|Z_t - \widehat{Z}_t|}{1/2(Z_t + \widehat{Z}_t)}} \times 100\%$$

Dimana:

- M= banyak data *out sample*
- Z<sub>t</sub>= Data penelitian pada waktu ke t
- Ẑ<sub>t</sub>= Data hasil ramalan dari penelitian pada waktu ke t

**2.3 Langkah Analisis**

Pada penelitian ini, digunakan data sekunder tentang Ekspor Non Migas di Kalimantan Timur pada bulan Januari 2011-Desember 2020 yang didapat dari website BPS Kalimantan Timur. Data kemudian dilakukan analisis menggunakan metode ARIMA Box Jenkins untuk tujuan peramalan/prediksi nilai ekspor non migas. Langkah analisis dari metode ARIMA dapat dijelaskan sebagai berikut:

1. Melakukan pembagian data menjadi data in sampel/ training pada Januari 2011 sampai Desember 2019 dan out sampel/testing pada Januari 2020 sampai Desember 2020.
2. Melakukan pemeriksaan stasioner data dalam varians. Jika tidak terpenuhi perlu dilakukan transformasi data.
3. Melakukan pemeriksaan stasioner data dalam mean. Jika tidak terpenuhi perlu dilakukan differencing.
4. Mengidentifikasi model dengan menggunakan plot ACF dan PACF
5. Mengestimasi dan menguji parameter model ARIMA

6. Memeriksa asumsi white noise dan distribusi normal
7. Pemilihan model yang paling baik untuk melakukan peramalan/prediksi dengan menggunakan kriteria sMAPE dan RMSE
8. Melakukan forecast pada data Ekspor Non Migas di Kalimantan Timur

**3. HASIL DAN PEMBAHASAN**

Pada bab ini akan dibahas mengenai karakteristik dan peramalan/prediksi data Ekspor Non Migas di Kalimantan Timur.

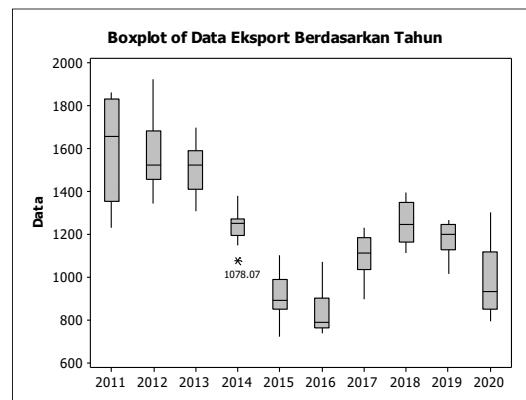
**3.1 Karakteristik Data**

Berikut merupakan gambaran karakteristik dari data Ekspor Non Migas di Kalimantan Timur pada bulan Januari 2011-Desember 2020.

**Tabel 1** Karakteristik Data

Variabel	Mean	StDev	Minimum	Maximum
Data Ekspor Non Migas	1219.2	290.9	722.0	1927.3

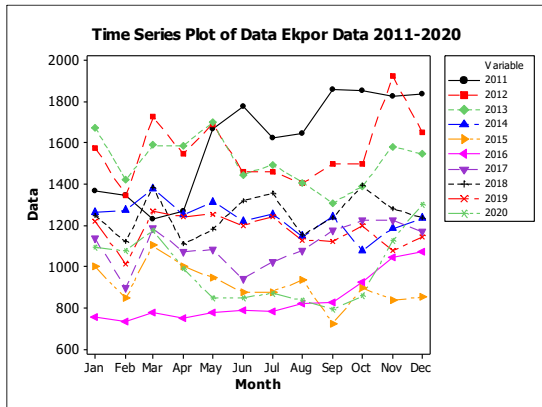
Tabel 1 menunjukkan rata-rata nilai Ekspor Non Migas di Kalimantan Timur adalah sebesar \$ 1219,2 juta, dengan nilai maksimum sebesar \$ 1927,3 juta terjadi pada November 2012, dan minimum sebesar \$ 722.0 juta pada bulan September 2015. Selanjutnya akan disajikan boxplot untuk mengetahui fluktuasi nilai Ekspor Non Migas per tahun.



**Gambar 1** Boxplot Data Per Tahun

Berdasarkan Gambar 1 diatas dapat diketahui bahwa nilai ekspor non migas di Kalimantan Timur pada tahun 2011-2016 semakin menurun, Data tersebut kemudian meningkat pada tahun-tahun

berikutnya sampai 2018 dan menurun lagi pada tahun 2019, 2020. Menurut BPS Penurunan nilai ekspor luar negeri pada tahun 2019 adalah akibat adanya penurunan harga komoditas non migas seperti Harga Batubara Acuan (HBA). Sedangkan penurunan nilai ekspor non migas Provinsi Kalimantan Timur pada tahun 2020 dipengaruhi oleh turunnya sebagian besar nilai ekspor ke beberapa negara. Selanjutnya untuk mengetahui pola data tiap bulan, dilakukan visualisasi data yang digambarkan pada Gambar 2 berikut:

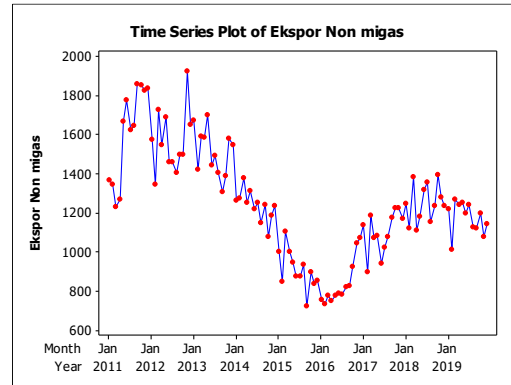


Gambar 2 Pola Data Ekpor Non Migas Per Bulan

Berdasarkan Gambar 2 dapat diketahui pola data ekspor non migas di Kalimantan Timur hampir sama di bulan Januari, Februari, Maret, dan April. Namun, Pada bulan April-Desember data ekspor non migas menunjukkan fluktuasi yang tidak menentu.

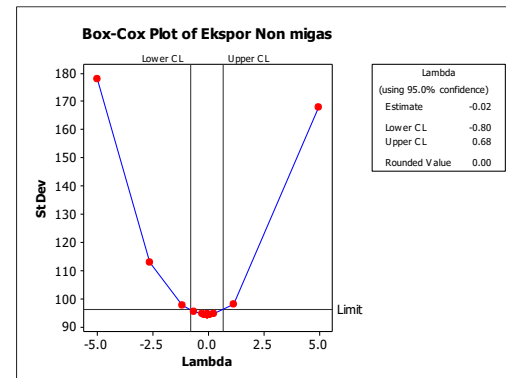
### 3.2 Analisis Data Metode ARIMA

Data Ekspor Non Migas di Kalimantan Timur dibagi menjadi dua, dengan pembagian data in sampel/training pada bulan Januari 2011 sampai Desember 2019 dan out sampel/testing pada Januari 2020 sampai Desember 2020. Selanjutnya data tersebut akan digambarkan dengan time series plot untuk mengetahui bentuk/pola datanya.



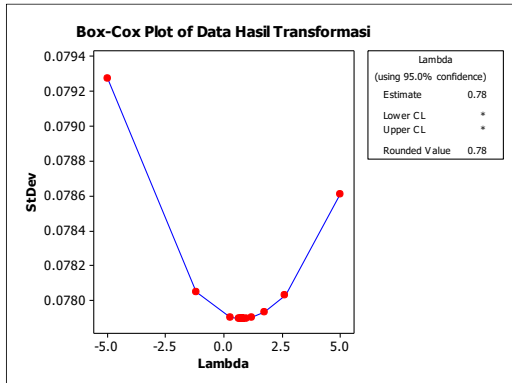
Gambar 3 Time Series Plot Data Ekpor Non Migas

Gambar 3 memperlihatkan bentuk/pola data Ekspor non migas di Kalimantan Timur yang menurun pada Januari 2011-Januari 2016. Data tersebut kemudian mengingkat pada Januari 2016-Januari 2018 dan menurun kembali pada Januari 2019 dan 2020. Setelah melakukan visualisasi data untuk mengetahui pola data yang akan dianalisis, Langkah selanjutnya yaitu pemeriksaan kestasioneran data terhadap varians. Salah satu caranya adalah melakukan Transformasi *Box-Cox* yang disajikan pada Gambar 4 berikut:



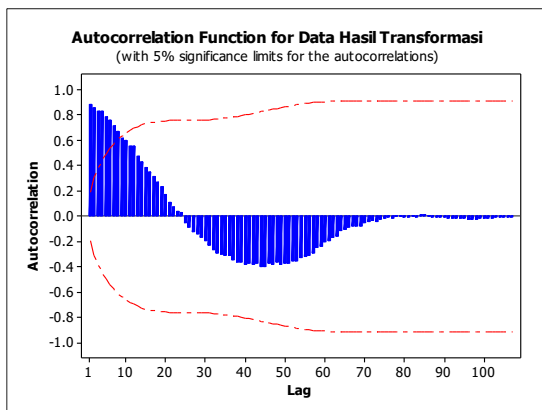
Gambar 4 Plot Box Cox Transformation Nilai Ekspor

Gambar 4 menunjukkan bahwa data ekspor non migas di Kalimantan Timur ternyata belum stasioner terhadap varians karena nilai *lower* dan *upper*nya berada pada range -0.80 sampai 0,68 yang masih belum melewati 1. Karena nilai tersebut belum melewati 1, maka transformasi logaritma natural perlu dilakukan pada data asli. Setelah dilakukan transformasi “ln” hasil yang didapat adalah:



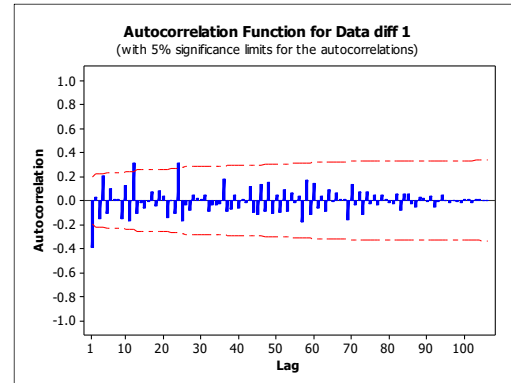
Gambar 5 Box Cox Data Transformasi

Gambar 5 menunjukkan bahwa data hasil transformasi yang telah dilakukan menggunakan fungsi logaritma natural, telah stasioner dalam varian. Hal tersebut terlihat dari nilai *lower* dan *upper*nya yang berada pada kisaran range  $-\infty$  sampai  $\infty$  (melewati 1). Karena data transformasi telah memenuhi asumsi stasioner terhadap varian, maka selanjutnya akan dilakukan pemeriksaan kestasioneran data terhadap *means* dengan menggunakan plot ACF. Hasil pemeriksaan dapat disajikan pada Gambar 6 berikut:



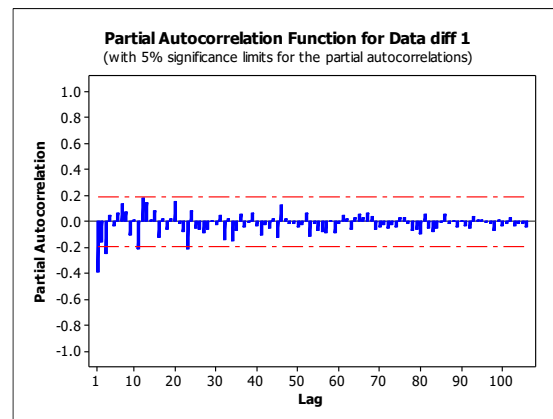
Gambar 6 ACF Data Transformasi

Berdasarkan hasil Plot ACF pada Gambar 6 maka plot ACF dari data yang dianalisis turun lambat/terdapat proses penurunan pada lag 1,2,3, 4 yang tidak terlalu signifikan. Hal ini menunjukkan bahwa data yang dianalisis belum stasioner terhadap mean. Hasil tersebut didukung oleh pemeriksaan dengan pengujian Dickey Fuller yang menunjukkan P-Value (0.8138) yang  $>$  dari  $\alpha$  (0,05). Sehingga perlu dilakukan *differencing* 1 yang dapat disajikan pada Gambar berikut:



Gambar 7 ACF Data Differencing

Plot ACF pada Gambar 7 menunjukkan bahwa data telah stasioner terhadap *mean*. Hasil tersebut juga didukung oleh pemeriksaan menggunakan pengujian Dickey Fuller yang memiliki nilai P-Value (0.000)  $<$   $\alpha$  (0,05). Sehingga data dapat dianalisis lebih lanjut untuk menentukan orde model ARIMA.



Gambar 8 PACF Data Differencing

Gambar 7 dan 8 menunjukkan bahwa Plot ACFnya *cut off* di Lag 1,12, dan 24 sedangkan PACF *cut off* di Lag 1,3,11,23 sehingga keluar adalah ARIMA dugaan berdasarkan lag yang keluar adalah ARIMA (0,1,[1,12]) dan ARIMA (1,1,[24]). Selanjutnya dilakukan uji signifikansi untuk menentukan apakah parameter model ARIMA yang diduga signifikan atau tidak. Hasil uji signifikansi dapat disajikan pada Tabel berikut:

Table 2 Signifikansi Parameter

Model	Type	Coef	SE Coef	T	P	Kesimpulan
ARIMA						
ARIMA (0,1,[1,12])	$\theta_1$	0,48246	0,08112	5,95	0,000	Signifikan
	$\theta_{12}$	-0,24697	0,08370	-2,95	0,004	Signifikan

ARIMA (1,1,[24])	$\phi_1$	-0.38331	0.09044	-4.24	<0.000	Signifikan
	$\theta_{24}$	-0.31264	0.10518	-2.97	0.004	Signifikan

Dari estimasi parameter yang telah dilakukan diketahui bahwa parameter model ARIMA (0,1,[1,12]) dan ARIMA (1,1,[24]) berpengaruh signifikan terhadap parameternya karena memiliki P-value yang nilainya  $(\alpha=0,05)$ . Setelah dilakukan uji signifikansi, dilakukan pemeriksaan asumsi white noise dan distribusi normal yang akan ditunjukkan pada Tabel 5.3.

**Table 3** Pemeriksaan White Noise dan Residual Distribusi Normal

Model	White Noise (WN)			Normal itas (p-value)
	Lag	P-value	Kesimpulan	
ARIMA (0,1,[1,12])	6	0.2018	Memenuhi	>0.150
	12	0.4917	Memenuhi	
	18	0.7432	Memenuhi	
	24	0.3550	Memenuhi	

**Table 4** Pemeriksaan White Noise dan Residual Distribusi Normal (Lanjutan)

ARIMA (1,1,[24])	6	0.1603	Memenuhi	>0.1500
	12	0.1223	Memenuhi	
	18	0.3419	Memenuhi	
	24	0.3681	Memenuhi	

Dari hasil pemeriksaan asumsi white noise dan distribusi normal, maka model yang diduga (ARIMA (0,1,[1,12]) dan ARIMA (1,1,[24])) untuk meramalkan nilai ekspor non migas telah memenuhi asumsi residualnya. Hal ini karena pada pengujian asumsi, model tersebut memiliki nilai P-Value yang lebih besar dari  $\alpha$  (0,05). Dua model ini selanjutnya akan dibandingkan untuk meramalkan data Ekspor Non Migas dengan mencari nilai sMAPE dan RMSE pada data outsample. Berikut merupakan nilai kriteria kebaikan model yang bisa dilihat dari nilai sMAPE, dan RMSE.

**Table 5** Penentuan Model Terbaik Data Outsample

Model	sMAPE	RMSE
<b>ARIMA (0,1,[1,12])</b>	<b>17.38724</b>	<b>203.7754</b>
ARIMA (1,1,[24])	18.31621	225.7532

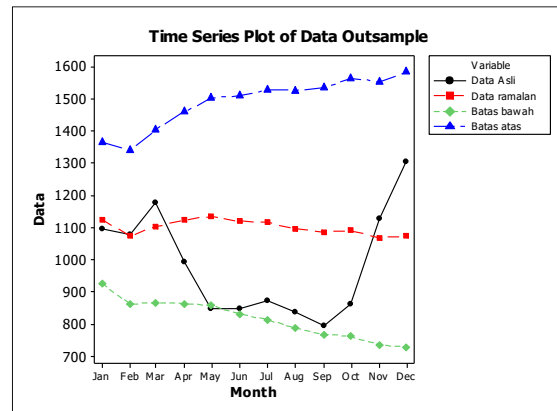
Berdasarkan kriteria untuk menentukan model terbaik menggunakan sMAPE dan RMSE dapat diketahui bahwa model ARIMA yang paling baik untuk meramalkan data Ekspor non migas di

Kalimantan Timur adalah ARIMA (0,1,[1,12]). Bentuk umum model dapat dituliskan :

$$(1 - B)^1 Z_t = (1 - \theta_1 B^1 - \theta_{12} B^{12}) a_t$$

$$Z_t = Z_{t-1} - 0,4824 a_{t-1} + 0,24697 a_{t-12} + a_t$$

Selanjutnya data out sample nilai non migas di Kalimantan Timur akan dibandingkan dengan hasil ramalan data outsample:



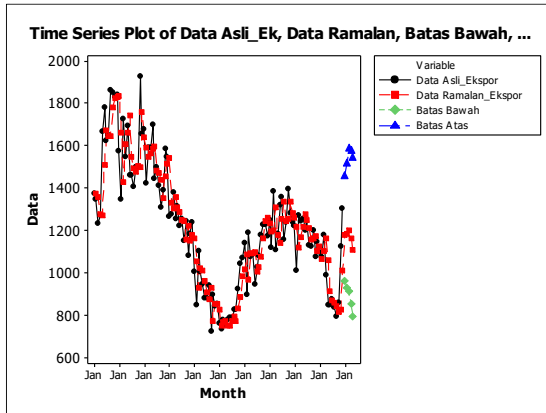
**Gambar 9** Perbandingan Data Outsample dan Hasil Ramalan

Berdasarkan perbandingan antara data outsample dan data ramalan dapat diketahui bahwa hasil ramalan menggunakan model ARIMA ternyata belum bisa menangkap pola data dengan baik apabila terdapat perubahan ekstrim dari data. Hal ini dapat dilihat dari plot yang berwarna hitam dan merah masih memiliki perbedaan, namun hasil analisis menunjukkan bahwa data outsample masih berada pada range batas atas dan bawah hasil ramalan sehingga model tersebut masih dapat digunakan untuk meramalkan data Ekspor non migas beberapa periode bulan ke depan. Hasil ramalan data ekspor untuk 5 bulan dapat diketahui sebagai berikut:

**Table 6** Peramalan Data Ekspor Bulan Januari-Mei 2021

	Batas Bawah	Hasil Ramalan	Batas Atas
Januari 2021	959.09	1180.21	1452.34
Februari 2021	926.71	1183.71	1511.99
Maret 2021	910.16	1200.76	1584.14
April 2021	854.39	1160.27	1575.66
Mei 2021	793.67	1106.66	1543.09

Hasil visualisasi peramalan data Ekspor Non Migas di Kalimantan Timur dapat disajikan pada Gambar berikut:



**Gambar 10** Perbandingan Data Ekspor Non Migas dan Hasil Ramalan

Berdasarkan plot pada Gambar 10 yang menunjukkan perbandingan antara data aktual dan data ramalan dapat diketahui bahwa hasil ramalan sudah bisa menangkap pola data dengan baik. Hal itu terlihat dari plot berwarna hitam dan merah yang saling berhimpit.

#### 4. KESIMPULAN DAN SARAN

Hasil penelitian peramalan data Ekspor Non Migas di Kalimantan Timur menggunakan metode ARIMA diperoleh kesimpulan yaitu:

1. Model ARIMA (0,1,[1,12]) merupakan model yang paling baik untuk meramalkan data Ekspor non migas di Kalimantan Timur karena memiliki nilai sMAPE dan RMSE yang minimum. Bentuk umum model dapat dituliskan :
 
$$Z_t = Z_{t-1} - 0,4824a_{t-1} + 0,24697a_{t-12} + a_t$$
2. Dari hasil perbandingan antara data aktual dan data ramalan dapat diketahui bahwa hasil ramalan sudah bisa menangkap pola data dengan baik karena plot nya yang saling berhimpit.

#### DAFTAR PUSTAKA

- Amalina, N. (2016). *Penerapan Metode Artificial Neural Network untuk Meramalkan Nilai Ekspor Migas dan Non Migas di Indonesia*. Surabaya: Institut Teknologi Sepuluh Nopember.
- Badan Pusat Statistika. (2020). *Analisis Komoditas Ekspor 2013-2020*. Jakarta: Badan Pusat Statistika.

- Daniel, W. W. (1989). *Statistik Nonparametrik Terapan*. Jakarta: PT Gramedia.
- Gooijer, J. D., & Hyndman, R. J. (2006). 25 Years of Time Series Forecasting. *Journal of Forecasting*, vol 22 no 443-473.
- Hartati. (2017). Penggunaan Metode ARIMA Dalam Meramal Pergerakan Inflasi. *Jurnal Matematika, Sains, dan Teknologi*, 1-10.
- Huda, S. (2006). Analisis Beberapa Faktor yang Mempengaruhi Ekspor Non Migas Indonesia ke Jepang. *Jurnal Ilmu-ilmu Ekonomi*, 117-224.
- Lestari, E., Widiyarih, T., & Rahmawati, R. (2018). Peramalan Ekspor Nonmigas dengan variasi Kalender Islam Menggunakan X-13 ARIMA SEATS. *Jurnal Gaussian*, 236-247.
- Nurmawaddah, S. (2019). Pengaruh Ekspor Non Migas Terhadap Pertumbuhan Ekonomi Provinsi Kalimantan Timur. *Akuntabel*, 1228-1333.
- Ruslan, R., Harahap, A. S., & Sembiring, P. (2013). Peramalan Nilai Ekspor di Provinsi Sumatera Utara dengan Metode ARIMA Box-Jenkins. *Saintia Matematika*, 579-589.
- Wardhana, A. (2011). Analisis Faktor-Faktor yang Mempengaruhi Ekspor Non Migas Indonesia ke Singapura Tahun 1990-2010. *Jurnal Manajemen Akutansi*, 99-102.
- Wei, W. (2006). *Time Series Analysis Univariate and Multivariate Methods (2nd ed.)*. New Jersey: Pearson Education Inc.

**Pemanfaatan Satelit Himawari-8 Untuk Estimasi Curah Hujan Dengan Metode *Convective Stratiform Technique* (CST) Dan *Modified Convective Stratiform Technique* (Mcst) Di Wilayah Ekuatorial Dan Monsunal (Studi Kasus Sulawesi 2020)**

Nur Habib Muzaki<sup>(1)</sup>, Eriska Febriati<sup>(2)</sup>, Yosafat Donni Haryanto<sup>(3)</sup>

Sekolah Tinggi Meteorologi Klimatologi dan Geofisika

Jl. Perhubungan 1 No. 5 Komplek Meteorologi, Pondok Betung, Bintaro 15221, Indonesia.

e-mail: [nur.habib.muzaki@stmkg.ac.id](mailto:nur.habib.muzaki@stmkg.ac.id)

**ABSTRAK**

Curah hujan merupakan salah satu parameter cuaca yang sangat penting bagi kehidupan. Informasi data curah hujan mampu menunjukkan pola tipe curah hujan di suatu wilayah. Kurangnya rapat persebaran alat pengamatan curah hujan menyebabkan cakupan wilayah menjadi sempit. Oleh karena itu, pemanfaatan metode estimasi curah hujan dengan menggunakan data satelit merupakan salah satu solusi untuk mendapatkan data curah hujan di wilayah yang tidak memiliki alat pengamatan curah hujan. Dalam penelitian ini, estimasi curah hujan menggunakan data satelit *Himawari-8* dengan menggunakan metode *Convective Stratiform Technique* (CST) dan *Modified Convective Stratiform Technique* (mCST). Metode CST merupakan metode yang memisahkan komponen awan konvektif dan *stratiform*, sedangkan metode mCST merupakan metode modifikasi intensitas curah hujan serta luasan area rata-rata yang dilingkupi piksel terhadap metode CST. Penelitian ini dilakukan di wilayah tipe hujan ekuatorial yang diwakili oleh Kabupaten Luwu Utara dan wilayah tipe hujan monsun yang diwakili oleh Kota Makassar. Penelitian ini dilakukan selama satu tahun dengan mengambil sampel bulan puncak curah hujan di kedua wilayah. Tujuan penelitian ini adalah untuk membandingkan kualitas hasil estimasi curah hujan dengan menggunakan metode CST dan metode mCST. Berdasarkan hasil estimasi curah hujan, metode CST menghasilkan nilai yang lebih baik dibandingkan dengan metode mCST di kedua wilayah yang ditandai dengan nilai korelasi yang lebih baik. Nilai eror RMSE berkisar 33.80 mm/jam hingga 42.66 mm/jam dan Nilai MAE berkisar 26.30 mm/jam hingga 34.55 mm/jam. Berdasarkan penelitian ini, kedua metode estimasi curah hujan ini, kurang mampu mempresentasikan data curah hujan di kedua wilayah.

**Kata kunci :** *Estimasi curah hujan, Himawari-8, CST, mCST.*

**ABSTRACT**

*Rainfall is one of the most important weather parameters for life. Rainfall data information is able to show the pattern of rainfall types in an area. The lack of a meeting of the distribution of rainfall observation instruments has resulted in a narrow coverage area. Therefore, the use of rainfall estimation methods using satellite data is one solution to obtain rainfall data in areas that do not have rainfall observation tools. In this study, the bulk estimation used Himawari-8 satellite data using the Convective Stratiform Technique (CST) method and the Modified Convective Stratiform Technique (mCST). The CST method is a method that separates the convective and stratiform cloud components, while the mCST method is a method of modifying rainfall intensity and the average area area covered by pixels to the CST method. This research was conducted in the equatorial rain type area represented by North Luwu Regency and the monsoon rain type area represented by Makassar City. This research was conducted for one year by taking samples of the peak months of rainfall in both regions. The purpose of this study was to compare the quality of the estimation results of rainfall using the CST method and the mCST method. Based on the results of rainfall estimation, the CST method produced better values compared to the mCST method in both regions which were characterized by better correlation values. RMSE error values range from 22.80 mm / hour to 42.66 mm / hour and MAE values range from 26.30 mm / hour to 34.55 mm / hour. Based on this research, these two rainfall estimation methods are not able to present rainfall data in the two regions.*

**Keywords :** *Rainfall estimation, Himawari-8, CST, mCST.*

## 1. PENDAHULUAN

Indonesia merupakan salah satu wilayah tropis dengan intensitas curah hujan yang tinggi. Hal ini disebabkan adanya banyaknya interaksi sirkulasi udara baik dalam skala lokal, regional, dan global (Hasanah dkk., 2017). Data curah hujan di manfaatkan untuk mengetahui pola curah hujan di suatu wilayah untuk mengetahui kapan terjadinya awal musim hujan, awal musim kemarau serta puncak musim hujan untuk mengantisipasi bencana seperti kekeringan dan banjir.

Data curah hujan umumnya didapatkan melalui pengamatan langsung secara konvensional dengan menakar curah hujan menggunakan alat pengukur hujan. Akan tetapi, alat pengukur hujan yang hanya tersebar di beberapa wilayah dengan sebaran yang kurang rapat sehingga hanya bisa mencakup wilayah yang sempit (Danamik dkk., 2018). Sehingga menyebabkan tidak tersedianya data curah hujan di beberapa wilayah yang tidak memiliki alat pengukur hujan. Oleh karena itu, dibutuhkan metode estimasi curah hujan untuk mengatasi masalah tersebut. Pemanfaatan alat penginderaan jauh seperti satelit diharapkan sebagai solusi terhadap tidak tersedianya data curah hujan di beberapa wilayah (Krisnayanti dkk., 2020). Salah satu satelit cuaca yang sering digunakan adalah satelit *Himawari-8*.

Salah satu metode estimasi curah hujan yang memanfaatkan data satelit cuaca berupa kanal *Infrared (IR)* dengan berdasarkan suhu puncak awan adalah *Convective Startiform Technique (CST)* dan *Modified Convective Startiform Technique (mCST)*. Metode *Convective Startiform Technique (CST)* merupakan metode estimasi curah hujan dengan memanfaatkan suhu kecerahan awan untuk memisahkan awan konvektif dan *stratiform* yang di temukan oleh Alder dan Negri (1998), sedangkan Metode *Modified Convective Startiform Technique (mCST)* merupakan metode modifikasi intensitas curah hujan serta luasan area rata-rata yang dilingkupi piksel (Endarwin, 2014). Dalam penelitian yang dilakukan di wilayah Pontianak bahwa kedua metode *Convective Startiform Technique (CST)* dan *Modified Convective Startiform Technique (mCST)* memiliki hasil estimasi curah hujan per jam dengan kualitas yang baik serta korelasi yang kuat (Andani dan Endarwin, 2016). Kemudian dalam penelitian di wilayah Pontianak sebagai wilayah dengan tipe curah hujan ekuatorial serta di wilayah Surabaya sebagai wilayah dengan tipe curah hujan monsunial bahwa metode *Convective Startiform Technique (CST)* dan *Modified Convective Startiform Technique (mCST)*

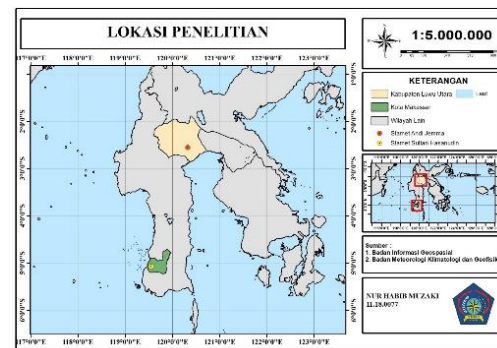
memiliki kemampuan yang cukup baik untuk estimasi curah hujan ringan dan sangat lebat (Andani dan Putra, 2016).

Dalam penelitian ini, metode *Convective Startiform Technique (CST)* dan *Modified Convective Startiform Technique (mCST)* digunakan dalam estimasi curah hujan pada tahun 2020 di pulau Sulawesi dengan wilayah yang memiliki tipe curah hujan ekuatorial dan monsunial. Untuk wilayah dengan tipe curah hujan ekuatorial yang memiliki dua puncak hujan dalam setahun diwakili oleh Kabupaten Luwu Utara dan tipe curah hujan monsunial diwakili oleh wilayah Kota Makassar. Tujuan penelitian ini adalah untuk menganalisis dan membandingkan kualitas hasil estimasi curah hujan dari metode *Convective Startiform Technique (CST)* dan *Modified Convective Startiform Technique (mCST)*.

## 2. METODE PENELITIAN

### 2.1. Lokasi penelitian

Penelitian ini dilakukan di Kabupaten Luwu Utara yang memiliki koordinat 20°30' sampai 2°37' Lintang Selatan dan 119°41' sampai 12°43' Bujur Timur dan Kota Makassar dengan koordinat 5° 30' hingga 5° 14' Lintang Selatan dan 19° 18' hingga 119° 32' Bujur Timur seperti yang digambarkan pada Gambar 1



Gambar 1. Peta wilayah penelitian.

### 2.2. Data Penelitian

Data yang digunakan dalam penelitian ini adalah data pengamatan curah hujan bulanan dan setiap satu jam selama satu tahun pada 2020 di Stasiun Meteorologi kelas III Andi Jemma dan Stasiun Meteorologi Kelas I Sultan Hasanuddin, serta data satelit *Himawari-8* kanal 13 atau kanal *Infrared Enhance (IR)* yang diperoleh dari <ftp://satelit.bmkg.go.id> dalam format netCDF (.nc) setiap satu jam pada bulan Juni dan Desember pada tahun 2020.

### 2.3. Metode Penelitian

Dalam penelitian ini, metode yang digunakan untuk estimasi curah hujan adalah metode *Convective Startiform Technique* (CST) dan *Modified Convective Startiform Technique* (mCST). Data yang sudah ada di olah menggunakan *Microsoft Excel* serta *Google Colab*. Adapun pengolahan data adalah sebagai berikut:

1. Mengumpulkan dan mengolah dalam *Microsoft Excel* data curah hujan bulanan sepanjang tahun 2020 di wilayah Kabupaten Luwu Utara dan Kota Makassar dalam bentuk garfik untuk melihat bulan puncak curah hujan.

2. Mengatur waktu dan koordinat lokasi penelitian

3. Konversi Data

Data satelit *Himawari-8* kanal *Infrared Enhance* (IR) setiap satu jam dalam bulan puncak hujan di konversi dalam ekstensi file *NetCDF* (.nc) sehingga diperoleh temperatur kecerahan awan ( $T_{BB}$ ) atau  $T$  minimum dalam satu atau multi piksel.

4. Penentuan Inti Konvektif

Setelah melakukan identifikasi terhadap suhu kecerahan awan ( $T_{BB}$ ), dilakukan pengukuran kekuatan awan konvektif dengan memperhitungkan *slope parameter* ( $S$ ) (Alder dan negri, 1988).

5. Pemisahan Awan Konvektif dan *Stratiform*

Pemisahan awan konvektif dan *stratiform* ini dilakukan dengan menggunakan data satelit *Himawari-8* kanal *Infrared* dengan mencari *slope parameter* ( $S$ ) dengan menggunakan rumus (Islam dkk., 2002; Endarwin, 2014):

$$S = k(T_{i-1,j-1} + T_{i-1,j} + T_{i+1,j+1} + T_{i+1,j} + T_{i,j-1} + T_{i,j} + 1 + T_{i+1,j-1} + T_{i-1,j+1} - 8T_{i,j}) \quad (1)$$

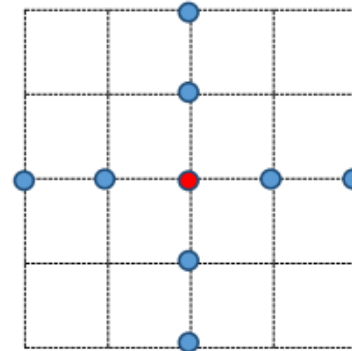
Keterangan:

$S$  = *slope parameter* ( $S$ ) (K)

$i$  dan  $j$  = posisi nilai piksel dimana  $S$  dihitung

$T$  = temperatur kecerahan awan ( $T_{BB}$ ) (K)

$K$  = konstanta (0.125)



Gambar 2. Lokasi 8 titik penentuan *slope parameter*.

Untuk pemisahan awan konvektif dan *stratiform* dilakukan dengan pembatasan yang menggunakan persamaan (Islam dkk, 2002; Endarwin, 2014):

$$S \geq \exp [0.0826 (T_{\min} - 207)] \quad (2)$$

Keterangan

$T_{\min}$  = temperature minimum relatif dari  $T_{BB}$  (K)

Dalam penelitian ini, dilakukan pemilihan satu piksel sebagai perwakilan koordinat stasiun, sehingga data  $T_{\min}$  merupakan temperatur pada piksel tersebut. Apabila nilai *slope parameter* ( $S$ ) memenuhi persamaan 2 maka merupakan inti awan konvektif dan sebaliknya (Islam dkk., 2002).

6. Penentuan luasan wilayah konvektif dan *stratiform*

Penentuan luasan wilayah konvektif dan *stratiform* ditentukan dengan menggunakan persamaan (Islam dkk., 2002; Endarwin, 2014). Untuk luasan awan konvektif ditentukan menggunakan rumus sebagai berikut:

$$\ln (A_c) = a T_{ci} + b \quad (3)$$

Keterangan:

$A_c$  = luasan area hujan konvektif ( $\text{km}^2$ )

$T_{ci}$  = temperatur puncak awan pada inti konvektif ke- $i$  (K)

$a = -0.0492$

$b = 15.27$

Sedangkan untuk luasan awan *stratiform* ditentukan dengan menggunakan rumus sebagai berikut:

$$\ln (A_s) = a T_s + b \quad (4)$$

Keterangan:

$A_s$  = luasan area hujan *stratiform* ( $\text{km}^2$ )

$T_s$  = temperatur puncak awan *stratiform* (K)

Dalam penelitian ini, hanya dipilih satu piksel yang mewakili koordinat stasiun, maka  $T_{\min}$  untuk  $T_{ci}$  dan  $T_s$  merupakan sebagai temperatur kecerahan awan pada piksel yang dibaca dalam penentuan  $T_s$ .

7. Tahapan estimasi curah hujan setiap jam. Estimasi curah hujan dengan menggunakan metode *Convective Startiform Technique* (CST) menggunakan persamaan sebagai berikut (Islam dkk., 2002; Enderwin, 2014):

$$\text{Curah Hujan Stratiform (mm)} = C(A_c/A)TR_c \tag{5}$$

Keterangan:

- C = baingan sel konvektif
- A<sub>c</sub> = luasan area hujan konvektif (km<sup>2</sup>)
- A = rata-rata area yang dilingkup piksel (km<sup>2</sup>)
- T = rata-rata periode estimasi (jam)
- R<sub>c</sub> = intensitas curah hujan konvektif (mm/jam)

$$\text{Curah Hujan Stratiform (mm)} = S_t(A_s/A)TR_s \tag{6}$$

Keterangan:

- S<sub>t</sub> = bilangan sel stratiform
- A<sub>s</sub> = luasan area hujan stratiform (km<sup>2</sup>)
- A = rata-rata area yang dilingkup piksel (km<sup>2</sup>)
- T = rata-rata periode estimasi (jam)
- R<sub>s</sub> = intensitas curah hujan stratiform (mm/jam)

Nilai T adalah 1, dikarenakan estimasi curah hujan dilakukan untuk curah hujan setiap jam serta nilai C dan S<sub>t</sub> juga adalah 1 karena pada penelitian ini hanya merujuk pada titik tertentu yang dipilih. Perbedaan spesifikasi metode *Convective Startiform Technique* (CST) dan *Modified Convective Startiform Technique* (mCST) yang digunakan dapat dilihat pada table table 1.

**Tabel 1.** Perbedaan spesifikasi metode CST dan mCST.

Metode	Intensitas Curah Hujan (mm/jam)		Luasan Rata-Rata Area Yang Dilingkupi
	Konvektif (R <sub>c</sub> )	Stratiform (R <sub>s</sub> )	
CST	20	3.5	121
Mcst	26	0.8	202.1243

8. Verifikasi hasil estimasi curah hujan. Verifikasi hasil estimasi curah hujan dilakukan dengan data pengamatan curah hujan dari stasiun yang telah dipilih. Verifikasi dilakukan untuk melihat bagus atau tidak nya kedua metode tersebut terhadap hasil data pengukuran curah hujan observasi. Menurut Pahlevi (2016) verifikasi hasil estimasi curah hujan terhadap curah hujan observasi dapat dilakukan dengan perhitungan sebagai berikut:

a. Koefisien Korelasi (C)

Korelasi merupakan salah satu metode statistik yang digunakan untuk mengetahui menguji ada atau tidaknya hubungan antar variabel, sedangkan koefisien korelasi (C) adalah bilangan yang

menyatakan besar kecilnya suatu hubungan (Pahlevi, 2016).

$$C = \frac{\sum_{i=1}^n (RO_i - \overline{RO})(RE_i - \overline{RE})}{\sqrt{\sum_{i=1}^n (RO_i - \overline{RO})^2} \sqrt{\sum_{i=1}^n (RE_i - \overline{RE})^2}} \tag{7}$$

**Tabel 2.** Kategori koefisien korelasi.

Interval Korelasi	Tingkat Hubungan
0.00 - 0.199	Sangat Lemah
0.20 - 0.399	Lemah
0.40 - 0.599	Sedang
0.60 - 0.799	Kuat
0.80 - 1.000	Sangat Kuat

b. *Root Mean Square Error* (RMSE)

RMSE digunakan untuk mengetahui besarnya penyimpangan yang terjadi antara data curah hujan hasil estimasi dengan data curah hujan hasil pengamatan atau observasi (Rizki, 2020).

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (RE_i - RO_i)^2}{n}} \tag{8}$$

c. *Mean Absolute Error* (MAE)

*Mean Absolute Error* (MAE) adalah rata-rata absolut dari kesalahan estimasi, tanpa menghiraukan tanda positif atau negatif.

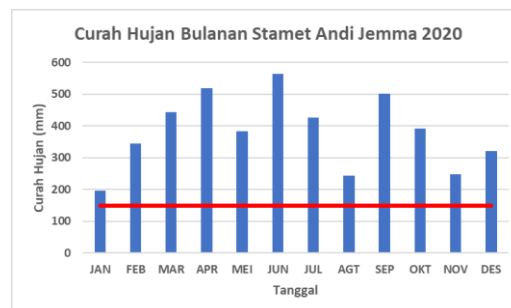
$$MAE = \frac{1}{n} \sum_{i=1}^n |(RE_i - RO_i)| \tag{9}$$

Keterangan:

- RE<sub>i</sub> dan RO<sub>i</sub> = nilai estimasi curah hujan dan observasi ke-i
- $\overline{RE}$  dan  $\overline{RO}$  = nilai rata-rata estimasi dan observasi
- n = jumlah data

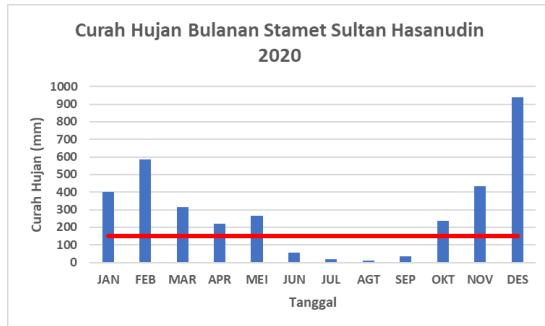
### 3. HASIL DAN PEMBAHASAN

#### 3.1. Penentuan Puncak Curah Hujan



**Gambar 3.** Curah hujan bulanan Stasiun Meteorologi Andi Jemma 2020.

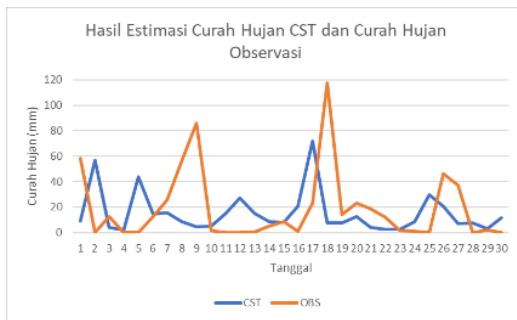
Berdasarkan Gambar 3, menunjukkan curah hujan bulanan stasiun meteorologi Andi Jemma pada tahun 2020. Terlihat bahwa puncak curah hujan terjadi pada bulan Juni sebesar 563.6 mm.



**Gambar 4.** Curah hujan bulanan Stasiun Meteorologi Sultan Hasanudin 2020.

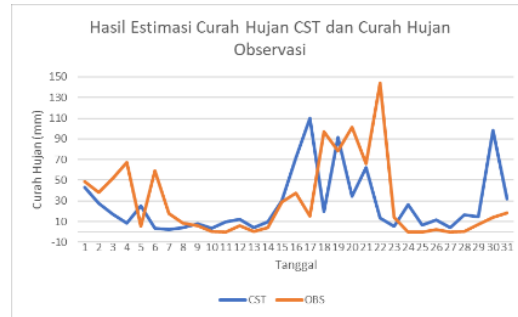
Berdasarkan Gambar 4, menunjukan curah hujan bulanan pada Stasiun Meteorologi Sultan Hasanudin pada tahun 2020. Terlihat bahwa puncak curah hujan terjadi pada bulan Desember sebesar 939.1 mm.

**3.2. Estimasi Curah dengan Metode Convective Startiform Technique (CST)**



**Gambar 5.** Hasil estimasi curah hujan CST dan curah hujan observasi di Stasiun Meterologi Andi Jemma.

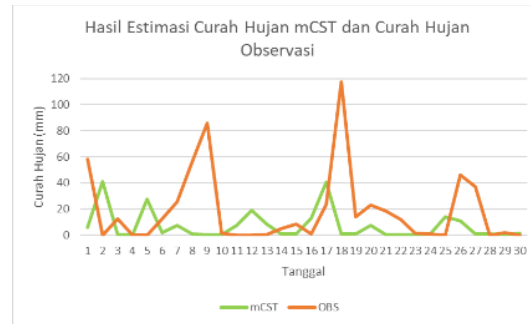
Berdasarkan Gambar 5, terlihat nilai estimasi curah hujan dengan metode Convective Startiform Technique (CST) di Kabupaten Luwu Utara, dengan menghasilkan estimasi curah hujan harian dengan nilai yang lebih rendah dibandingkan dengan data observasi, dengan selisih yang cukup tinggi. Data hasil estimasi curah hujan kurang baik dalam mempresentasikan curah hujan yang rendah di Kabupaten Luwu Utara.



**Gambar 6.** Hasil estimasi curah hujan CST dan curah hujan observasi di Stasiun Meteorologi Sultan Hasanudin.

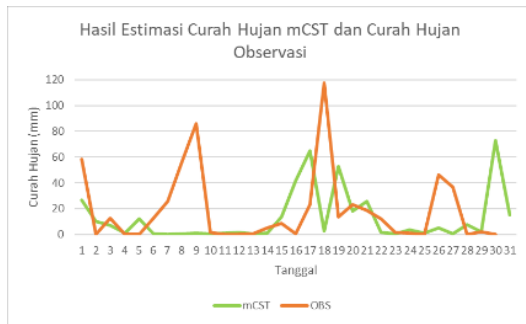
Berdasarkan Gambar 6, terlihat nilai estimasi curah hujan dengan metode Convective Startiform Technique (CST) di Kota Makassar, dengan menghasilkan estimasi curah hujan harian yang mampu mengikuti pola curah hujan dengan intensitas tinggi dari hasil data curah hujan observasi. Beberapa data hasil estimasi curah hujan juga dapat mendekati nilai curah hujan observasi.

**3.3. Estimasi Curah Hujan dengan Metode Modified Convective Startiform Technique (mCST)**



**Gambar 7.** Hasil estimasi curah hujan mCST dan curah hujan observasi di Stasiun Meterologi Andi Jemma.

Berdasarkan Gambar 7, terlihat nilai estimasi curah hujan dengan menggunakan metode Modified Convective Startiform Technique (mCST) di Kabupaten Luwu Utara. Pola estimasi curah hujan dengan metode mCST sangat berbeda dengan pola yang dihasilkan dari data curah hujan observasi. Nilai estimasi curah hujan yang berada di bawah nilai curah hujan observasi dengan selisih yang cukup besar.



**Gambar 8.** Hasil estimasi curah hujan mCST dan curah hujan observasi di Stasiun Meteorologi Sultan Hasanudin

Berdasarkan Gambar 8, terlihat nilai estimasi curah hujan dengan menggunakan metode *Modified Convective Startiform Technique* (mCST) di Kota Makassar. Pola estimasi curah hujan dengan metode mCST cukup baik dalam mempresentasikan curah hujan dalam intensitas yang rendah dengan selisih yang tidak terlalu besar. Akan tetapi tidak cukup baik dalam mempresentasikan data curah hujan observasi dalam intensitas yang tinggi, ditandai dengan besarnya selisih hasil data estimasi dan data observasi.

**3.4. Verifikasi Hasil Estimasi Curah Hujan dengan Data Observasi**

**Tabel 3.** Verifikasi data curah hujan hasil estimasi dengan data observasi.

	CST	mCST	Wilayah
<b>Korelasi</b>	-0.15	-0.17	<b>Luwu Utara</b>
<b>RMSE</b>	34.55	33.80	
<b>MAE</b>	23.27	21.46	
<b>Korelasi</b>	0.19	0.13	<b>Makassar</b>
<b>RMSE</b>	42.17	42.66	
<b>MAE</b>	26.93	26.30	

Berdasarkan Tabel 3, menunjukkan bahwa metode *Convective Startiform Technique* (CST) lebih baik dibandingkan dengan metode *Modified Convective Startiform Technique* (mCST) dengan nilai korelasi yang lebih besar yang menunjukkan bahwa curah hujan hasil metode *Convective Startiform Technique* (CST) lebih mendekati terhadap curah hujan hasil observasi. Nilai korelasi di Kota Makassar lebih besar dibandingkan dengan nilai korelasi di Kabupaten Luwu Utara yang menunjukkan bahwa metode *Convective Startiform Technique* (CST) di Kota Makassar lebih mendekati hasil observasi. Kemudian berdasarkan nilai RMSE, menunjukkan bahwa hasil kedua metode baik metode *Convective Startiform Technique* (CST) dan *Modified Convective Startiform Technique* (mCST) memiliki penyimpangan atau *error* yang besar terhadap data

observasi curah hujan, yaitu 33.8 nilai MAE atau nilai rata-rata sebesar 21.46 mm/jam hingga 26.30 mm/jam yang menunjukkan bahwa rata-rata penyimpangan estimasi curah hujan cukup besar dari kedua metode terhadap curah hujan observasi.

**4. KESIMPULAN DAN SARAN**

Berdasarkan hasil estimasi curah hujan setiap jam dengan metode *Convective Startiform Technique* (CST) dan *Modified Convective Startiform Technique* (mCST) di wilayah dengan tipe curah hujan ekuatorial yang diwakili oleh Kabupaten Luwu Utara serta wilayah dengan tipe curah hujan monsunial yang diwakili oleh Kota Makassar diperoleh kesimpulan sebagai berikut:

1. Metode estimasi curah hujan *Convective Startiform Technique* (CST) menghasilkan estimasi curah hujan yang lebih dibandingkan metode *Modified Convective Startiform Technique* (mCST) di kedua wilayah.
2. Hasil estimasi curah hujan dari kedua metode di Kota Makassar memiliki estimasi yang lebih mendekati curah hujan hasil observasi dibandingkan dengan di Kabupaten Luwu Utara yang dapat dilihat dengan nilai korelasi yang lebih besar di Kota Makassar sebesar -0.15 hingga 0.19.
3. Berdasarkan hasil verifikasi estimasi curah hujan dan data observasi, bahwa estimasi curah hujan dengan metode *Convective Startiform Technique* (CST) dan *Modified Convective Startiform Technique* (mCST) ini menghasilkan penyimpangan yang besar terhadap curah hujan hasil observasi, yaitu sebesar 33.80 mm/jam hingga 42.66 mm/jam. Sehingga, kedua metode ini tidak cocok digunakan untuk estimasi curah hujan di Kota Makassar dan Kabupaten Luwu Utara.

**DAFTAR PUSTAKA**

Andani, A.J.P. dan Enderwin., (2016). Kajian Penerapan Estimasi Curah Hujan Per Jam Memanfaatkan Metode *Convection Stratiform Technique* (CST) dan *Modified Convective Stratiform Technique* (mCST) di Pontianak. *Jurnal Meteorologi Klimatologi dan Geofisika*, 3(3), 9-20.

Andani, A.J.P. dan Putra, R.M. (2016). Penerapan Curah Hujan dengan Metode CST dan MCST di Wilayah Tipe Curah hujan Ekuatorial dan Monsunial (Studi kasus Pontianak, November 2014 dan Surabaya,

- Januari 2014). Prosiding Seminar Hari Meteorologi Dunia
- Danamik R.A., Dewi, N.K.T., Wulandari, A.V., Fadlan,, A., 2018. Penggunaan Metode *Modified Convective Stratiform Technique* (mCST) di Wilayah Tipe Curah Hujan Ekuatorial, Monsunal dan Lokal. Seminar Nasional Penginderaan Jauh.
- Enderwin, 2014, Modifikasi *Convective Stratiform Technique* dengan Kombinasi Data Satelit Gelombang Mikro Pasif dan Inframerah untuk Estimasi Curah Hujan di Indonesia, Disertasi, Program Studi Sains Kebumihan, Institut Teknologi Bandung, Bandung
- Hasanah, N., Ugiarto, M., dan Puspitasari, N., 2017, Sistem Pengelompokan Curah Hujan Menggunakan Metode K-Means Di wilayah Kalimantan Timur, Prosiding Seminar Nasional Ilmu Komputer dan Teknologi Informasi, Vol. 2. No. 2.
- Islam, M. N., Islam, A. K. M. S., Hayashi, T., Terao, T. dan Uyeda, H, 2002, Application of a Method to Estimate Rainfall in Bangladesh Using GMS- 5 Data, *Journal of Natural Disaster Science*, Vol. 24, no. 2 pp 83-89.
- Jiang, S., Ren, L., Yong, B., Yang, X. dan Shi, L, 2010, Evaluation of High Resolution Satellite Precipitation Products with Surface Rain Gauge Observations from Laohahe Basin in Northern China. *Water Science and Engineering*, Vol. 3, no. 4 pp 405- 417.
- Krisnayanti, D. S., Welkis, D. F. B., Hepy, F. M., dan Legono, D. 2020. Evaluasi Kesesuaian Data Tropical Rainfall Measuring Mission (TRMM) Dengan Data Pos Hujan Pada Das Temef Di Kabupaten Timor Tengah Selatan. *Jurnal Sumber Daya Air*, 16(1), 51-62.
- Pahlevi, A. R., 2016. Skema Parameterisasi Kumulus Untuk Prediksi Hujan di Wilayah Bandar Lampung. *Prosiding SNSA*.
- Rizki, M., Basuki, S., & Azhar, Y. 2020. Implementasi Deep Learning Menggunakan Arsitektur Long Short Term Memory Untuk Prediksi Curah Hujan Kota Malang. *Repositor*, 2(3), 331-338.

**Pemodelan Kasus Kumulatif Covid-19 di Pulau Jawa dan Bali  
Dengan Pendekatan Multiple Classification Analysis (MCA)**

Rezky Yayang Yakhamid<sup>(1)</sup>, Amelia Tri Wahyuni<sup>(2)</sup>, Nadidah Pengestika<sup>(3)</sup>, Hanifah<sup>(4)</sup>, Putu  
Adi Myarsithawan<sup>(5)</sup>, Risni Julaeli Yuhani<sup>(6)</sup>

Politeknik Statistika STIS  
Jalan Otto Iskandardinata No. 64C, Jakarta 13330  
e-mail: [211810559@stis.ac.id](mailto:211810559@stis.ac.id)

**ABSTRAK**

Pada Maret 2020, pandemi global Covid-19 mulai menyerang Indonesia. Tingginya jumlah kasus kumulatif Covid-19 mengakibatkan diberlakukannya kebijakan PSBB (Pembatasan Sosial Berskala Besar) di Indonesia. Meski kebijakan PSBB sempat dihapuskan karena penyebaran Covid-19 yang menurun, tetapi kemudian diberlakukan kebijakan baru berupa PPKM (Pemberlakuan Pembatasan Kegiatan Masyarakat) di Pulau Jawa dan Bali karena tingginya kasus kumulatif Covid-19 di wilayah tersebut. Tujuan dari penelitian ini mengetahui faktor-faktor yang mempengaruhi jumlah kasus kumulatif Covid-19 di Pulau Jawa dan Bali. Metode analisis yang digunakan dalam penelitian ini adalah MCA (Multiple Classification Analysis), yakni salah satu metode analisis yang bertujuan untuk mengestimasi pengaruh variabel penjelas yang bertipe kategoris terhadap variabel respons. Adapun variabel yang diduga memengaruhi jumlah kasus kumulatif Covid-19 yaitu klasifikasi daerah, kepadatan penduduk, persentase penduduk lansia, dan PDRB per kapita. Hasil penelitian menunjukkan bahwa pada tingkat signifikansi 5%, variabel klasifikasi daerah, kepadatan penduduk, persentase penduduk lansia, dan PDRB per kapita berpengaruh signifikan terhadap jumlah kasus kumulatif Covid-19 di Pulau Jawa dan Bali.

**Kata kunci :** Covid-19, MCA, Jawa, Bali, Spasial

**ABSTRACT**

*In March 2020, the global COVID-19 pandemic started to attack Indonesia. The high number of cumulative Covid-19 cases resulted in the implementation of the LSSR (Large Scale Social Restrictions) policy in Indonesia. Although the LSSR policy was removed because of the decreasing spread of Covid-19, a new policy named ELCA (Enforcement of Limitations on Community Activities) was implemented in Java and Bali islands due to the high number of cumulative Covid-19 cases in that region. The purpose of this study is to determine the factors that influence the number of cumulative active cases of COVID-19 in Java and Bali islands. The analytical method used in this study is MCA (Multiple Classification Analysis) that is one of the methods of analysis to estimate the effect of categorical explanatory variables on the response variables. The variables that are thought influence the number of cumulative active cases of COVID-19 are area classification, population density, percentage of the elderly population, and GRDP per capita. The results showed that at a significance level of 5%, area classification, population density, percentage of the elderly population, and GRDP per capita had a significant effect on the number of cumulative active cases of COVID-19 in Java and Bali islands.*

**Keywords :** Covid-19, MCA, Java, Bali, Spatial..

**1. PENDAHULUAN**

Pada Desember 2019 silam, dunia dikejutkan dengan penemuan penyakit Covid-19 di Wuhan, Hubei, Tiongkok. Penyebabnya adalah virus SARS-CoV-2 yang dapat menular antar manusia. Penyebaran virus ini dapat dikatakan cepat hingga merambat hampir ke seluruh belahan dunia. Sejak saat itu, pasien Covid-19 terus bertambah sehingga

WHO menetapkan Covid-19 merupakan pandemi global pada Maret 2020.

Di Indonesia, Covid-19 pertama kali ditemukan pada 2 Maret 2020 dan diumumkan langsung oleh Presiden Joko Widodo. Pasien tersebut merupakan Warga Negara Indonesia (WNI), ibu dan anaknya, yang berdomisili di Depok. Menurut Menteri Kesehatan yang menjabat saat itu, Terawan Agus Putranto, kedua WNI

Rezky Yayang Yakhamid<sup>1</sup>, Amelia Tri Wahyuni<sup>2</sup>, Nadidah Pengestika<sup>3</sup>, Hanifah<sup>4</sup>, Putu Adi Myarsithawan<sup>5</sup>, Risni Julaeli Yuhani<sup>6</sup>

tersebut terpapar virus Covid-19 dari WN Jepang yang sedang melakukan perjalanan ke Indonesia. Virus tersebut terus menyebar luas hingga ke pelosok-pelosok negeri.

Guna menekan penyebaran Covid-19, pemerintah menerapkan kebijakan-kebijakan yang dinilai efektif untuk mencegah penyebaran Covid-19. Awalnya, pemerintah mengeluarkan kebijakan Pembatasan Sosial Berskala Besar (PSBB). Kebijakan PSBB ini tertuang dalam Peraturan Menteri Kesehatan Nomor 9 Tahun 2020 tentang Pedoman Pembatasan Sosial Berskala Besar dalam Rangka Percepatan Penanganan Covid-19. Setelah penyebaran Covid-19 dinilai menurun, PSBB dihapuskan.

Seiring berjalannya waktu, kenaikan kasus kumulatif Covid-19 kembali naik. Tingginya kasus kumulatif ini membuat pemerintah kembali menerapkan kebijakan guna menekan penyebaran virus tersebut. Tak lagi menggunakan istilah PSBB, pemerintah memilih menggunakan istilah baru, yaitu Pemberlakuan Pembatasan Kegiatan Masyarakat (PPKM) Pulau Jawa dan Bali karena tingginya kasus kumulatif Covid-19 di wilayah tersebut.

Permasalahan yang terjadi adalah peningkatan Covid-19 yang belum dapat ditekan di Indonesia meskipun kebijakan telah diterapkan, khususnya Pulau Jawa dan Bali. Diketahui bahwa kasus Covid-19 di Indonesia mencapai 1,7 juta orang, pasien sembuh 1,6 juta orang, dan pasien meninggal mencapai 48 ribu orang per 20 Mei 2021. Provinsi Jawa Barat menyumbang jumlah kasus terkonfirmasi positif Covid-19 terbanyak, kemudian disusul oleh DKI Jakarta dan Jawa Tengah di urutan kedua dan ketiga. Selain penyumbang kasus terkonfirmasi Covid-19 terbanyak, menurut Juru Bicara Pemerintah Penanganan Covid-19, Wiku Adisasmito, Pulau Jawa dan Bali juga menjadi penyumbang kasus kematian tertinggi Covid-19 per 3 Januari 2021.

Peningkatan kasus terkonfirmasi Covid-19 tentunya harus ditangani dengan benar. Oleh karena itu, dilakukan penelitian mengenai faktor-faktor yang memengaruhi kasus Covid-19 di Pulau Jawa dan Bali. Selain itu, penelitian ini juga akan membahas pengaruh klasifikasi daerah, kepadatan penduduk, persentase lansia, dan PDRB per kapita terhadap kasus kumulatif Covid-19, khususnya Pulau Jawa dan Bali. Metode yang digunakan dalam penelitian ini adalah Multiple Classification Analysis (MCA), yakni salah satu metode analisis yang bertujuan untuk mengestimasi pengaruh

variabel penjelas yang bertipe kategoris terhadap variabel respons. Harapannya, hasil penelitian dapat digunakan sebagai acuan untuk menentukan kebijakan dalam menekan angka positif Covid-19 di Pulau Jawa dan Bali.

Peningkatan kasus Covid-19 di Indonesia dapat disebabkan oleh beberapa variabel, di antaranya adalah klasifikasi daerah dan persentase penduduk lansia. Ricardo Ramirez-Aldana (2020) dalam penelitiannya diperoleh kesimpulan bahwa perkotaan dengan penduduk tua yang mendominasi memiliki kasus Covid-19 yang tinggi. Menurut Ujii, et al (2020), rasio ketergantungan penduduk lansia memiliki hubungan dengan jumlah kumulatif kasus Covid-19 pada 16 Maret 2020 di Jepang. Hal ini disebabkan karena penduduk lansia cenderung mudah terpapar virus sebab imunitas yang menurun seiring bertambahnya umur.

Variabel lain yang dapat menyebabkan peningkatan kasus Covid-19 adalah kepadatan penduduk. Nelwan (2020) dalam penelitiannya yang bertujuan untuk mengetahui distribusi Covid-19 berdasarkan kepadatan penduduk dan ketinggian tempat dengan unit analisis kecamatan di Kota Manado. Hasil dari penelitian tersebut menunjukkan bahwa semakin tinggi kepadatan penduduk, maka semakin tinggi kasus Covid-19 di daerah tersebut.

Wong dan Li (2020) juga melakukan penelitian yang berjudul *Spreading of Covid-19 : Density matters*. Penelitian ini dilakukan karena menjaga jarak diperdebatkan sebagai salah satu cara yang efektif untuk mencegah penyebaran Covid-19 sebelum adanya vaksin. Peneliti juga menyebutkan bahwa sebagian besar model yang digunakan untuk memprediksi penyebaran Covid-19 di U.S. tidak memasukkan kepadatan penduduk secara terpisah sebagai salah satu variabel yang mempengaruhi sehingga peneliti menggunakan kepadatan penduduk sebagai variabel yang diteliti. Hasil penelitian ini menunjukkan bahwa kepadatan penduduk merupakan variabel yang efektif mempengaruhi kasus kumulatif Covid-19 di U.S. pada tingkat perkotaan.

Penelitian terkait pengaruh kepadatan penduduk terhadap kasus kumulatif Covid-19 juga dilakukan oleh Sy, et.al. (2021). Penelitian yang berjudul *Population Density and Basic Reproductive Number of Covid-19 across United States Counties* ini bertujuan untuk melihat hubungan antara kepadatan penduduk dengan kasus kumulatif Covid-19 di kabupaten di U.S. Penelitian ini menghasilkan kesimpulan bahwa

area yang padat dapat meningkatkan penularan virus Covid-19 karena tingkat kontak antarmanusia cenderung tinggi.

Selain variabel-variabel tersebut, Covid-19 juga dipengaruhi oleh PDRB per kapita. Pendapatan yang tinggi dapat memudahkan dalam memenuhi kebutuhan nutrisi dan kesehatan dibandingkan dengan pendapatan rendah (Dewi, 2020). Oleh karena itu, PDRB per kapita yang tinggi diindikasikan dapat menekan angka penyebaran Covid-19. Sebab, dengan tingginya PDRB per kapita, maka nutrisi dan kesehatan masyarakat akan terjaga.

## 2. METODE PENELITIAN

### 2.1. Ruang Lingkup dan Sumber Data

Unit analisis dalam penelitian ini adalah seluruh wilayah kabupaten dan kota (termasuk kota administratif Provinsi DKI Jakarta) di Pulau Jawa dan Bali yang terdiri atas 128 kabupaten/kota. Data jumlah kasus kumulatif Covid-19 didapatkan dari situs web Satuan Tugas Penanganan Covid-19 pada setiap provinsi di Pulau Jawa dan Bali. Dalam situs web tersebut terdapat data spasial jumlah kasus Covid-19 pada setiap kabupaten dan kota pada setiap provinsi di Pulau Jawa dan Bali.

Adapun sumber data variabel-variabel penjelas yang digunakan berasal dari data sekunder berbagai survei dan sensus oleh Badan Pusat Statistik pada tahun 2020. Sumber data yang digunakan antara lain hasil Sensus Penduduk 2020 untuk variabel kepadatan penduduk dan persentase penduduk lansia, serta variabel PDRB atas dasar harga konstan yang dikompilasi dari berbagai sumber oleh BPS. Adapun status daerah ditetapkan sesuai perundang-undangan yang berlaku yang mengklasifikasikan daerah otonomi tingkat II menjadi kabupaten atau kota madya.

### 2.2. Variabel Penelitian

Variabel respons (Y) dalam penelitian ini adalah jumlah kasus kumulatif Covid-19 di Indonesia per tanggal 21 Mei 2021. Kasus kumulatif Covid-19 dalam penelitian ini didefinisikan sebagai seseorang yang masih aktif terinfeksi Covid-19 sampai dengan tanggal 21 Mei 2021. Adapun pertimbangan dalam memilih titik potong tanggal 21 Mei 2021 antara lain adalah adanya kekhawatiran meningkatnya kasus Covid-19 pada momen setelah Idul Fitri 2021 akibat pelanggaran mudik dan pelanggaran protokol kesehatan saat Idul Fitri.

Sementara itu, variabel penjelas yang digunakan didasarkan atas beberapa penelitian

terkait yang relevan dan memberikan hasil signifikan. Variabel penjelas yang digunakan dalam penelitian ini adalah klasifikasi daerah (X1), kepadatan penduduk (X2), persentase penduduk lansia (X3), dan Pendapatan Domestik Regional Bruto (PDRB) per kapita (X4). Variabel penjelas tersebut kemudian dikategorisasi dengan menggunakan nilai median. Penggunaan ketegorisasi dengan menggunakan median dilakukan karena nilai statistik median *robust* terhadap *outlier*.

### 2.3. Metode Analisis

Metode analisis yang digunakan adalah analisis deskriptif dan analisis inferensial. Analisis deskriptif adalah metode analisis yang berkaitan dengan pengumpulan dan penyajian data sehingga memberikan informasi atau interpretasi yang bermanfaat (Walpole, 1988). Analisis deskriptif digunakan untuk mengetahui sebaran data dari variabel respons dalam bentuk diagram peta dan sebaran variabel respons pada masing-masing variabel penjelas dalam bentuk diagram batang. Adapun analisis inferensial adalah metode analisis yang berhubungan dengan analisis sampel untuk sampai pada penarikan kesimpulan populasi (Walpole, 1988). Analisis inferensial dalam penelitian ini menggunakan metode Multiple Classification Analysis (MCA) untuk mengetahui pengaruh dari variabel penjelas terhadap variabel respons. Pengolahan dan analisis data dalam penelitian ini menggunakan Software Google Sheet, Microsoft Excel, QGIS, dan SPSS versi 24.

### 2.4. Multiple Classification Analysis (MCA)

Multiple Classification Analysis (MCA) adalah salah satu metode analisis inferensial yang bertujuan untuk mengestimasi pengaruh variabel penjelas terhadap variabel respons. MCA merupakan sebuah teknik untuk menduga hubungan antara beberapa variabel penjelas dan satu variabel respons serta menentukan efek dari masing-masing variabel penjelas sebelum dan sesudah disesuaikan dengan variabel penjelas lainnya. MCA merupakan metode alternatif yang dapat digunakan apabila variabel penjelasnya berupa data kategori atau *dummy* sedangkan variabel bebasnya berupa interval/rasio (Sugiarto, 2018).

Adapun langkah-langkah dalam menerapkan analisis inferensial dengan metode MCA menurut Sugiarto (2018) adalah sebagai berikut :

1. Melakukan observasi/pengamatan dan identifikasi nilai ekstrem pada distribusi data

variabel respons. Salah satu syarat penggunaan model MCA adalah pada variabel responsnya tidak tidak terdapat outlier karena data outlier akan memengaruhi model yang terbentuk. Cara untuk mengatasinya adalah dengan melakukan transformasi data atau membuangnya. (Sugiarto, 2018).

2. Melakukan pengujian model secara simultan dan parsial. Uji simultan dan parsial digunakan untuk mengidentifikasi ada tidaknya pengaruh variabel penjelas terhadap variabel respons. Pengujian tersebut dilakukan dengan menguji nilai F main effect. Hipotesis yang digunakan sebagai berikut :

H0 = tidak terdapat pengaruh variabel penjelas terhadap variabel respons

H1 = minimal terdapat satu variabel penjelas yang berpengaruh terhadap variabel respons

Statistik uji yang digunakan untuk uji simultan adalah sebagai berikut :

$$F = \frac{E/(C-P)}{(T-E)/(N+P-C-1)} \quad (1)$$

dengan :

F = nilai F secara simultan

C = jumlah seluruh kategori

P = jumlah variabel independen

T = total sum of squares

E = explained sum squares kuadrat

Keputusan adalah tolak H0 jika p-value < α (tingkat signifikansi yang ditetapkan), atau  $F_{hitung} > F_{0,10}(v1,v2)$

Sedangkan, statistik uji yang digunakan untuk uji parsial adalah sebagai berikut :

$$F = \frac{U_a/(C_a-1)}{(T-U_a)/(N-C_a)} \quad (2)$$

dengan :

$F_a$  = nilai F secara parsial untuk variabel ke-a

$C_a$  = jumlah kategori variabel ke-a

$U_a$  = jumlah kuadrat antar kategori variabel ke-a

T = total sum of squares

Keputusan adalah tolak H0 jika p-value < α (tingkat signifikansi yang ditetapkan), atau  $F_{hitung} > F_{0,10}(v1,v2)$

3. Melakukan pengecekan ada tidaknya interaksi antar variabel bebasnya. Karena model bersifat aditif, MCA mensyaratkan bahwa nilai suatu variabel penjelas tidak dipengaruhi nilai variabel penjelas lainnya dalam hal memengaruhi variabel respons (Sugiarto, 2018). Untuk itu dilakukan uji signifikansi F pada pola interaksi antar variabel bebas yang terbentuk. Hipotesis sebagai berikut:

H0 = tidak terdapat interaksi antar variabel bebas

H1 = terdapat interaksi antar variabel bebas

Statistik ujinya adalah sebagai berikut :

$$F = \frac{\text{Mean square interaksi}}{\text{Mean square of residual}} \quad (3)$$

Keputusan adalah tolak H0 jika p-value < α (tingkat signifikansi yang ditetapkan) atau

$F_{hitung} > F_{0,10}(v1,v2)$

4. Pembentukan model aditif. Model yang digunakan pada MCA adalah model aditif (model penambahan) yang dibentuk sebagai berikut :

$$Y_{ij...n} = \bar{Y} + a_i + b_j + \dots + \epsilon_{ij...n} \quad (4)$$

di mana  $\bar{Y} = \frac{\sum_{k=1}^n Y_k}{n}$

dengan :

$Y_{ij...n}$  = skor pengamatan pada individu pada kategori ke-i dari variabel a, kategori ke-j dari variabel b, dan seterusnya

$\bar{Y}$  = grandmean (rata-rata keseluruhan) dari variabel respons

$Y_k$  = nilai individu ke -x pada variabel respons

n = banyak observasi

$a_i$  = pengaruh/efek dari grandmean pada kategori ke-i dan variabel a

$b_j$  = pengaruh/efek dari grandmean pada kategori ke-j dari variabel b

$\epsilon_{ij...n}$  = komponen error bagi individu yang bersesuaian dengan  $Y_{ij...n}$

Setelah mendapatkan model tersebut, MCA dapat dengan mudah diaplikasikan untuk mengestimasi nilai variabel respons observasi tertentu dengan memasukkan komponen-komponen variabel penjelas ke dalam model.

5. Evaluasi model dengan mengidentifikasi besar pengaruh yang diberikan variabel penjelas terhadap variabel respons. Dalam mengidentifikasi besarnya pengaruh dapat digunakan statistik nilai eta (η) dan beta (β). Nilai Eta diinterpretasikan sebagai pengaruh setiap variabel penjelas terhadap variabel respons sebelum dipertimbangkan pengaruh dari variabel penjelas lain (before adjustment). Sementara itu, nilai beta diinterpretasikan sebagai pengaruh setiap variabel penjelas terhadap variabel respons setelah dipertimbangkan variabel penjelas lain (after adjustment). Karakteristik eta dan beta juga disebut rasio korelasi (correlation rasio) (Susel, 2011). Formula perhitungan nilai eta (η) dan beta (β) adalah sebagai berikut.

a. Nilai eta (η) untuk variabel ke-a

$$\eta_a = \sqrt{\frac{U_a}{T}} \quad (5)$$

dengan  $\eta_a$  = nilai eta untuk variabel ke-a

Selain nilai eta yang mengukur pengaruh dari variabel penjelas, terdapat juga nilai eta kuadrat (*eta square*). Eta kuadrat ( $\eta^2$ ) menjelaskan proporsi varians dari variabel respons yang dapat dijelaskan oleh variabel penjelas sebelum *adjustment* (Susel, 2011). Berikut formula untuk menghitung eta kuadrat.

$$\eta_{adj}^2 = 1 - \left[ \frac{(T-E)/(N-C+P-1)}{T/(N-1)} \right] \quad (6)$$

b. Nilai beta ( $\beta$ ) untuk variabel ke-a

$$\beta_a = \sqrt{\frac{D_a}{T}} \quad (7)$$

di mana  $D_a = \sum_{(i=1)}^{C_a} N(\bar{Y}_{ai(adj)} - \bar{Y})^2$  dan  $T = \sum_k N(Y_k - \bar{Y})^2$

dengan :

$\beta_a$  = nilai beta untuk variabel ke-a

$D_a$  = jumlah kuadrat adjusted antar kategori pada variabel ke-a

$\bar{Y}_{ai(adj)}$  = rata-rata variabel respons untuk kategori ke-i pada variabel ke-a adjusted

Selain nilai eta yang mengukur pengaruh dari variabel penjelas, terdapat juga nilai eta kuadrat (*eta square*). Beta kuadrat ( $\beta^2$ ) menjelaskan proporsi varians dari variabel respons yang dapat dijelaskan oleh variabel penjelas setelah *adjustment* (Susel, 2011). Berikut formula untuk menghitung beta kuadrat

$$\beta_{adj}^2 = 1 - (1 - \beta^2) \left[ \frac{N-1}{N+P-C-1} \right] \quad (8)$$

6. Evaluasi model dengan cara mengidentifikasi hubungan atau besar pengaruh seluruh variabel penjelas terhadap variabel respons menggunakan nilai koefisien determinasi ( $R^2$ ). Perhitungan koefisien determinasi adalah sebagai berikut.

$$R^2 = E/T \quad (9)$$

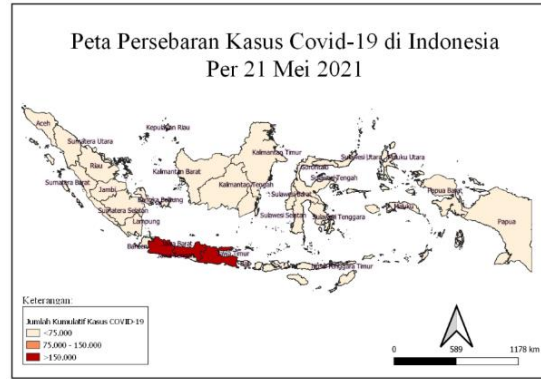
$$R_{adj}^2 = 1 - (1 - R^2)A \quad (10)$$

di mana  $A = (N - 1)/(N + P - C - 1)$  adalah derajat bebas adjusted.

### 3. HASIL DAN PEMBAHASAN

#### 3.1. Analisis Deskriptif

Pandemi Covid-19 yang terjadi di Indonesia cukup mengkhawatirkan, Hal ini disebabkan terus terjadinya penambahan kasus baru Covid -19 di tiap harinya. Peristiwa tersebut terjadi hampir di seluruh wilayah Indonesia, tak terkecuali di Pulau Jawa dan Bali.

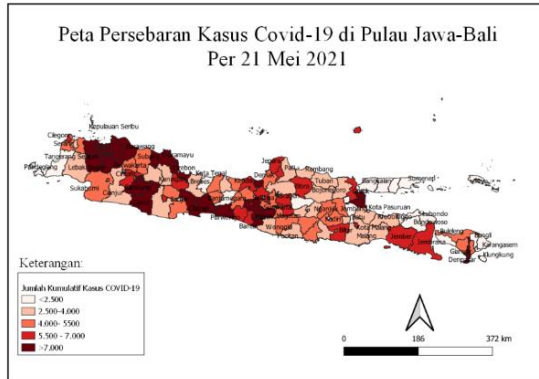


Sumber : Satgas Covid-19 (diolah)

**Gambar 1.** Sebaran Kasus Covid-19 di Indonesia Per 21 Mei 2021

Pada gambar 1 ditunjukkan sebaran kumulatif jumlah kasus Covid-19 di seluruh wilayah Indonesia per 21 Mei 2021. Berdasarkan gambar tersebut, terlihat bahwa hampir seluruh wilayah di Pulau Jawa memiliki jumlah kasus Covid-19 yang terbanyak di Indonesia dengan lebih dari 150 ribu kasus. Sedangkan untuk wilayah lainnya kebanyakan memiliki jumlah kasus covid-19 kurang dari 75 ribu kasus.

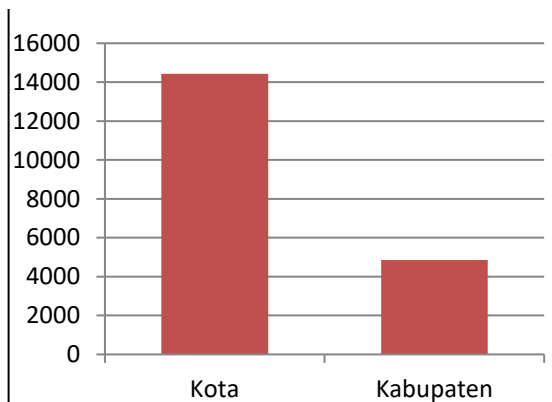
Hal tersebut menjadikan Pulau Jawa menyumbang paling banyak kasus covid-19 yang ada di Indonesia. Padahal, Pemerintah sudah berupaya dengan memberlakukan Pembatasan Sosial Berskala Besar (PSBB) untuk menekan peningkatan jumlah kasus covid-19 yang ada di Pulau Jawa. Akan tetapi upaya tersebut masih optimal sebab masih terdapat pelanggaran-pelanggaran akibat ketidakpatuhan masyarakat. Penerapan PSBB juga diterapkan pada wilayah di Pulau Bali demi menekan penyebaran covid-19 dengan nama Pembatasan Pemberlakuan Kegiatan Masyarakat atau PPKM. Hal itu mengingat Pulau Bali juga merupakan salah satu wilayah yang dapat dikatakan cukup tinggi dalam penambahan kasus covid-19. Menurut Juru Bicara Pemerintah Penanganan Covid-19, Wiku Adisasmito, Pulau Jawa dan Bali menyumbang kasus kematian Covid-19 tertinggi di Indonesia, yakni 66,7 persen atau 15.165 dari kumulatif kasus kematian akibat Covid-19 per 3 Januari 2021. Berikut ditunjukkan sebaran kumulatif jumlah kasus Covid-19 pada setiap kabupaten/kota di pulau Jawa dan Bali per 21 Mei 2021.



Sumber : Satgas Covid-19 (diolah)

**Gambar 2.** Sebaran Kasus Covid-19 di Pulau Jawa – Bali Per 21 Mei 2021

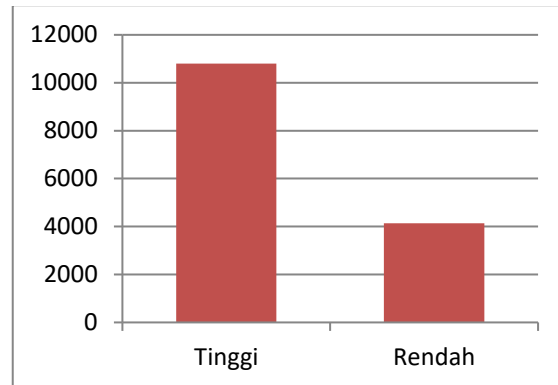
Apabila ditinjau lebih dalam menurut wilayah di Pulau Jawa-Bali, gambar 2 menunjukkan sebaran kasus kumulatif Covid-19 per satu juta penduduk di masing-masing wilayah kabupaten/kota di Pulau Jawa-Bali. Hal yang menarik dari gambar 2 yaitu seluruh wilayah Jabodetabek memiliki jumlah kasus Covid-19 yang tinggi dengan lebih dari 7.500 kasus. Hal tersebut juga terjadi di kota-kota besar di Provinsi Jawa Barat, Jawa Tengah, Jawa Timur dan Bali seperti Bandung, Cilacap, Surabaya, dan Denpasar. Jumlah kasus kumulatif Covid-19 terus bertambah setiap harinya. Hal tersebut tentu tidak terlepas dari faktor-faktor yang memengaruhi peningkatan kasus kumulatif Covid-19 di Pulau Jawa-Bali.



**Gambar 3.** Rata-rata Jumlah Kumulatif Kasus Covid-19 di Pulau Jawa – Bali Per 21 Mei 2021 Menurut Status Wilayah.

Pada Gambar 3 ditunjukkan rata-rata jumlah kumulatif kasus Covid-19 per satu juta penduduk yang terjadi di Pulau Jawa-Bali per 21 Mei 2021 menurut status wilayah. Dari gambar di atas diperoleh bahwa status wilayah kota memiliki rata-

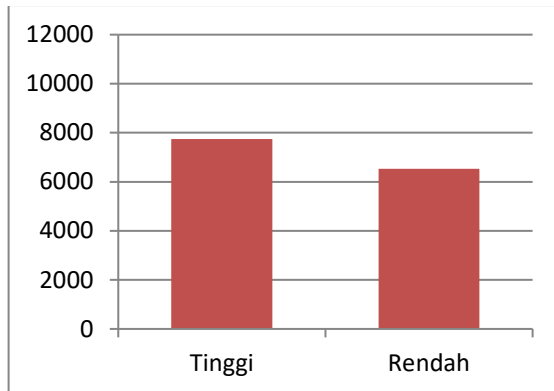
rata jumlah kumulatif kasus Covid-19 per satu juta penduduk lebih tinggi dibandingkan dengan status wilayah kabupaten dengan lebih dari 14 ribu kasus. Berbeda dengan status kota, wilayah dengan status kabupaten memiliki perbedaan rata-rata jumlah kumulatif kasus Covid-19 yang cukup jauh dibandingkan dengan status wilayah kota yaitu sekitar lima ribu kasus. Hal ini mengindikasikan bahwa terdapat pengaruh status wilayah kota dan kabupaten terhadap jumlah kasus kumulatif baru Covid-19 di Pulau Jawa-Bali.



**Gambar 4.** Rata-rata Jumlah Kumulatif Kasus Covid-19 di Pulau Jawa – Bali Per 21 Mei 2021 Menurut Kepadatan Penduduk

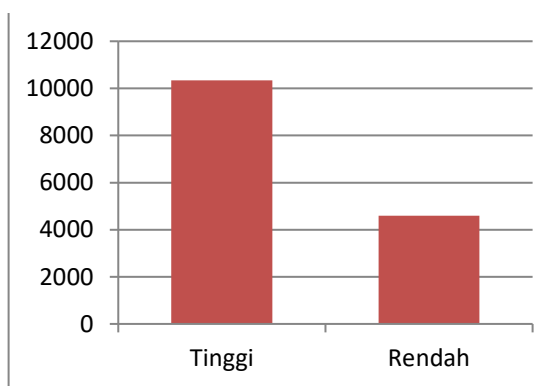
Faktor lain yang diindikasikan dapat memengaruhi penambahan jumlah kumulatif kasus Covid-19 adalah kepadatan penduduk. Menurut pakar kesehatan, virus Covid-19 cepat menyebar pada kelompok yang berkerumun. Kepadatan penduduk merupakan salah satu pemicu terbentuknya kerumunan akibat terbatasnya wilayah pergerakan penduduk. Oleh karena itu diindikasikan bahwa kepadatan penduduk memicu penambahan kasus baru Covid-19. Pada gambar 4 ditunjukkan rata-rata jumlah kumulatif kasus Covid-19 per satu juta penduduk di Pulau Jawa-Bali per 21 Mei 2021 menurut status kepadatan penduduk. Berdasarkan gambar di atas, diperoleh bahwa wilayah dengan status kepadatan penduduk yang tinggi memiliki rata-rata jumlah kumulatif kasus Covid-19 per satu juta penduduk lebih banyak dibandingkan dengan wilayah yang memiliki status kepadatan penduduk yang rendah. Pada wilayah dengan status kepadatan penduduk tinggi, diperoleh rata-rata jumlah kumulatif kasus Covid-19 per satu juta penduduk hampir 11 ribu kasus, sedangkan wilayah dengan status kepadatan penduduk rendah diperoleh sekitar empat ribu kasus. Hal ini mengindikasikan bahwa perbedaan status kepadatan penduduk memengaruhi

penambahan kasus kumulatif Covid -19 per satu juta penduduk di Pulau Jawa-Bali.



**Gambar 5.** Rata-rata Jumlah Kumulatif Kasus Covid-19 di Pulau Jawa – Bali Per 21 Mei 2021 Menurut Persentase Penduduk Lansia

Sementara itu, kelompok usia penduduk juga menjadi salah satu faktor terjadinya kasus baru Covid-19, terutama pada kelompok usia penduduk lansia yang memiliki risiko lebih tinggi untuk terinfeksi virus Covid-19. Berdasarkan gambar 5, terlihat bahwa wilayah di Pulau Jawa-Bali yang memiliki persentase penduduk lansia tinggi lebih banyak terjadi kasus Covid-19 dibandingkan dengan wilayah yang memiliki persentase penduduk lansia yang rendah. Rata-rata jumlah kasus Covid-19 per satu juta penduduk di wilayah yang memiliki persentase lansia yang tinggi hampir mencapai delapan ribu kasus. Sedangkan untuk wilayah dengan persentase lansia yang rendah memiliki sekitar enam ribu kasus. Hal ini mengindikasikan bahwa terdapat pengaruh persentase penduduk lansia dalam peningkatan jumlah kumulatif kasus Covid-19 di Pulau Jawa-Bali.



**Gambar 6.** Jumlah Kumulatif Kasus Covid-19 di Pulau Jawa – Bali Per 21 Mei 2021 PDRB per Kapita

Selanjutnya, pada gambar 6 ditunjukkan rata-rata jumlah kumulatif kasus Covid-19 per satu juta penduduk di Pulau Jawa-Bali per 21 Mei 2021 menurut PDRB per kapita. Berdasarkan gambar 3.6, terlihat bahwa wilayah yang memiliki PDRB per kapita tinggi lebih banyak terjadi kasus Covid-19 dibandingkan dengan wilayah yang memiliki PDRB per kapita yang rendah. Hal ini ditunjukkan dengan rata-rata jumlah kumulatif kasus Covid-19 per satu juta penduduk di wilayah yang memiliki PDRB per kapita tinggi sekitar sepuluh ribu kasus. Sementara itu, untuk wilayah dengan PDRB per kapita yang rendah hanya memiliki rata-rata jumlah kumulatif kasus Covid-19 per satu juta penduduk sekitar enam ribu kasus. Hal ini mengindikasikan bahwa adanya pengaruh PDRB per kapita dalam peningkatan jumlah kumulatif kasus Covid-19 di Pulau Jawa-Bali.

### 3.2. Analisis Inferensial

#### Multiple Classification Analysis pada Jumlah Kumulatif Kasus Covid-19 Per Satu juta Penduduk

Analisis MCA digunakan untuk menganalisis pengaruh variabel klasifikasi daerah, kepadatan penduduk, persentase penduduk lansia, dan Pendapatan Domestik Regional Bruto (PDRB) per kapita terhadap jumlah kumulatif kasus Covid-19 per satu juta penduduk. Pengaruh tersebut dianalisis melalui uji simultan dan uji parsial. Uji simultan digunakan untuk mengetahui pengaruh variabel penjelas secara bersama-sama terhadap jumlah kumulatif kasus Covid-19 per satu juta penduduk, sedangkan uji parsial digunakan untuk mengetahui pengaruh setiap variabel penjelas terhadap jumlah kumulatif kasus Covid-19 per satu juta penduduk.

#### Uji Simultan dan Uji Parsial

Sebelum dibentuk persamaan MCA, langkah pertama yang dilakukan adalah melihat pengaruh berbagai variabel penjelas terhadap jumlah kumulatif kasus Covid-19 per satu juta penduduk. Uji yang digunakan adalah melalui uji ANOVA. Hasil pengujian ANOVA ditunjukkan pada tabel (2) berikut :

Tabel 2. ANOVA

Variabel Penelitian	Hierarchical Method				
	Sum of Squares	df	Mean Square	F	Sig.
(Combined)	3153540853	4	788385213.3	32.307	0.000
Status wilayah	2331911879	1	2331911879	95.557	0.000
Main Effects					
Kepadatan penduduk	103524132.4	1	103524132.4	4.242	0.042
Persentase lansia	587849439.3	1	587849439.3	24.089	0.000
PDRB per kapita	130255402.3	1	130255402.3	5.338	0.023
Model	3153540853	4	788385213.3	32.307	0.000
Residual	3001605227	123	24403294.53		
Total	6155146080	127	48465717.17		

Pengaruh berbagai variabel penjelas terhadap jumlah kumulatif kasus Covid-19 per satu juta penduduk ditunjukkan melalui nilai signifikansi (*sig*) atau nilai *p-value* pada model. Dengan menggunakan tingkat signifikansi 5 persen, hasil uji simultan menghasilkan keputusan tolak H0 yang artinya signifikan. Sehingga dapat disimpulkan bahwa dengan tingkat signifikansi 5% dapat dikatakan bahwa terdapat minimal satu variabel penjelas yang memengaruhi jumlah kumulatif kasus Covid-19 per satu juta penduduk.

Untuk dapat mengetahui informasi yang lebih detail, dilakukan uji parsial dalam rangka mengetahui pengaruh setiap variabel penjelas terhadap jumlah kumulatif kasus Covid-19 per satu juta penduduk. Hasil pengujian parsial menunjukkan keputusan tolak H0 pada semua variabel penjelas, hal ini dikarenakan semua nilai signifikansi (*sig*) kurang dari 5 persen. Sehingga dapat disimpulkan bahwa dengan tingkat signifikansi 5% dari data yang tersedia terdapat cukup bukti untuk mengatakan bahwa secara parsial variabel klasifikasi daerah, kepadatan penduduk, persentase penduduk lansia dan Pendapatan Domestik Regional Bruto (PDRB) per kapita berpengaruh signifikan terhadap jumlah kumulatif kasus Covid-19 per satu juta penduduk.

**Pembentukan Model Aditif**

Tabel 3. Multiple Classification Analysis

Variabel	N	Predicted Mean	Deviation
Stat	kota	1298	5515
		1.7083	.40924
wilayah	kabupaten	5390	-
		.6074	2075.69165
Kepadatan penduduk	tinggi	9268	1801
		.1138	.81475
penduduk	rendah	5664	-
		.4843	1801.81475

entase lansia	gi	tinggi	7	9898	2432
		rendah	1	.6158	.31673
RB per kapita	PD	tinggi	4	8672	1206
		rendah	4	.588	.28891

Hubungan variabel penjelas terhadap jumlah kumulatif kasus Covid-19 per satu juta penduduk dapat dilihat dari *deviation*. Tanda negatif dari nilai tersebut menunjukkan bahwa daerah yang berada pada kategori tersebut memiliki jumlah kumulatif kasus Covid-19 per satu juta penduduk yang lebih sedikit dibandingkan dengan rata – rata jumlah kumulatif kasus Covid-19 per satu juta penduduk secara keseluruhan. Dari tabel 2 dapat disimpulkan bahwa wilayah dengan status kabupaten memiliki jumlah kumulatif kasus Covid-19 per satu juta penduduk lebih sedikit dibanding wilayah dengan status kota, wilayah dengan kepadatan penduduk yang rendah memiliki jumlah kumulatif kasus Covid-19 per satu juta penduduk lebih sedikit dibanding wilayah dengan kepadatan penduduk tinggi, wilayah dengan persentase penduduk lansia yang rendah memiliki jumlah kumulatif kasus Covid-19 per satu juta penduduk lebih sedikit dibanding wilayah dengan persentase penduduk lansia yang tinggi, wilayah dengan PDRB per kapita rendah memiliki jumlah kumulatif kasus Covid-19 per satu juta penduduk lebih sedikit dibanding wilayah dengan PDRB per kapita tinggi.

Persamaan MCA yang terbentuk merupakan persamaan linier aditif dari berbagai variabel penjelas. Nilai konstanta merupakan nilai *grand mean* atau rata – rata seluruh jumlah kumulatif kasus Covid-19 per satu juta penduduk dan nilai koefisien merupakan nilai *deviation* yang telah disesuaikan pada setiap variabel penjelas yang digunakan. Berdasarkan tabel (2) dan penghitungan nilai *grand mean*, persamaan MCA yang terbentuk adalah :

$$Y = 7866,3582 + X1_i + X2_j + X3_k + X4_l$$

Berdasarkan persamaan MCA di atas, secara umum wilayah yang memiliki jumlah kumulatif kasus Covid-19 per satu juta penduduk lebih banyak adalah wilayah dengan status kota dengan kepadatan penduduk, persentase penduduk lansia, dan PDRB per kapita yang tinggi. Sebaliknya wilayah dengan status kabupaten dengan kepadatan penduduk, persentase penduduk lansia, dan PDRB per kapita yang rendah memiliki jumlah kumulatif

kasus Covid-19 per satu juta penduduk lebih sedikit.

**Evaluasi Model**

Nilai eta dan beta menunjukkan besarnya pengaruh setiap variabel penjelas terhadap lama jumlah kumulatif kasus Covid-19 per satu juta penduduk. Nilai eta dan beta setiap variabel penjelas ditunjukkan oleh tabel berikut.

**Tabel 4.** Nilai Eta dan Beta

Variabel Penelitian	Eta	Beta
Status wilayah	0.616	0.488
Kepadatan penduduk	0.480	0.260
Persentase lansia	0.041	0.368
PDRB per kapita	0.415	0.174

Berdasarkan tabel 3 dapat dilihat secara berturut – turut variabel yang berpengaruh terhadap jumlah kumulatif kasus Covid-19 per satu juta penduduk adalah variabel klasifikasi daerah, kepadatan penduduk, persentase penduduk lansia, dan Pendapatan Domestik Regional Bruto (PDRB) per kapita. Sebelum mempertimbangkan pengaruh dari variabel penjelas lainnya (*before adjustment*), dengan melihat nilai Eta masing-masing variabel dapat dikatakan bahwa variabel yang paling kuat berpengaruh terhadap jumlah kumulatif kasus Covid-19 per satu juta penduduk adalah variabel klasifikasi wilayah sedangkan yang paling lemah berpengaruh adalah variabel persentase penduduk lansia.

Setelah mempertimbangkan pengaruh dari variabel penjelas lainnya (*after adjustment*), Variabel klasifikasi wilayah adalah variabel yang paling berpengaruh terhadap jumlah kumulatif kasus Covid-19 per satu juta penduduk yang dibuktikan dengan nilai Beta sebesar 0,488 paling tinggi jika dibandingkan dengan variabel penjelas yang lain. Sedangkan variabel dengan nilai Beta terendah adalah variabel yang paling lemah berpengaruh terhadap jumlah kumulatif kasus Covid-19 per satu juta penduduk, yaitu dimiliki oleh variabel PDRB per kapita dengan nilai Beta 0,174.

**Tabel 5.** Goodness of Fit

R	0.716
R Squared	0.512

Dari tabel 4 dapat dilihat bahwa Persamaan MCA yang terbentuk memiliki nilai R squared sebesar 0,512 yang berarti bahwa keragaman jumlah kumulatif kasus Covid-19 per satu juta penduduk sebesar 51,2%. Selain itu dapat dikatakan bahwa pengaruh variabel klasifikasi daerah, kepadatan penduduk, persentase penduduk lansia, dan Pendapatan Domestik Regional Bruto (PDRB) per kapita terhadap jumlah kumulatif kasus Covid-19 per satu juta penduduk adalah sebesar 51,2%.

**4. KESIMPULAN DAN SARAN**

Berdasarkan hasil analisis yang telah diuraikan, dapat disimpulkan bahwa pada tingkat signifikansi 5%, variabel yang berpengaruh signifikan terhadap jumlah kumulatif kasus Covid-19 per satu juta penduduk di Pulau Jawa dan Bali adalah klasifikasi daerah, kepadatan penduduk, persentase penduduk lansia, dan PDRB per kapita. Secara umum, wilayah yang memiliki jumlah kumulatif kasus Covid-19 per satu juta penduduk lebih banyak adalah wilayah daerah perkotaan dengan kepadatan penduduk, persentase penduduk lansia, dan PDRB per kapita yang tinggi. Sebaliknya, wilayah daerah perdesaan dengan kepadatan penduduk, persentase penduduk lansia, dan PDRB per kapita yang rendah memiliki jumlah kumulatif kasus Covid-19 per satu juta penduduk yang lebih sedikit. Kemudian, variabel penjelas dalam penelitian ini secara bersama-sama dapat menjelaskan variasi jumlah kumulatif kasus Covid-19 per satu juta penduduk sebesar 51,2%. Adapun sebesar 48,8% sisanya dijelaskan oleh variabel lain yang tidak tercakup dalam model. Oleh karena itu, diperlukan penelitian lebih lanjut untuk mengkaji variabel penjelas lainnya yang belum tercakup.

**DAFTAR PUSTAKA**

Ayo Bandung. (2021). *Alasan PPKM, Jawa dan Bali Sumbang Kasus Kematian Terbanyak Covid-19*. Ayo Bandung. Tersedia di <https://ayobandung.com/read/2021/01/08/172702/alasan-ppkm-jawa-dan-bali-sumbang-kasus-kematian-terbanyak-covid-19>

Detik Health. (2020). *Dugaan Kasus Pertama Virus Corona di China Terdeteksi pada November 2019*. Detik Health. Tersedia di <https://health.detik.com/berita-detikhealth/d-5004285/dugaan-kasus-pertama-virus-corona-di-china-terdeteksi-pada-november-2019>

- Indonesia. (2020). *Kasus Covid-19 Pertama, Masyarakat Jangan Panik. Portal Informasi Indonesia*. Tersedia di <https://indonesia.go.id/narasi/indonesia-dalam-angka/ekonomi/kasus-covid-19-pertama-masyarakat-jangan-panik>
- Kementerian Kesehatan RI. (2020). *Hindari Lansia dari Covid-19. Kemkes RI*. Tersedia di <http://padk.kemkes.go.id/article/read/2020/04/23/21/hindari-lansia-dari-covid-19.html>
- Kompas. (2021). *Update 20 Mei 2021 : Bertambah 5.797 Orang, Kasus Covid-19 di Indonesia Capai 1.758.898*. Kompas. Tersedia di <https://nasional.kompas.com/read/2021/05/20/16335371/update-20-mei-2021-bertambah-5797-orang-kasus-covid-19-di-indonesia-capai>
- Nelwa, J.E. (2020). Kejadian Corona Virus Disease 2019 berdasarkan Kepadatan Penduduk dan Ketinggian Tempat per Wilayah Kecamatan. *Indonesian Journal of Public Health and Community Medicine*, 1(2)
- Sugiarto. (2018). Multiple classification analysis (MCA) Sebagai Metode Alternatif Analisis Data Untuk Variabel Bebas yang Kategori. *Jurnal Statistika*, 6(2)
- Susel, Aleksander. (2011). Multiple Classification Analysis Theory and Application to Demography. *Folia Oeconomica*, 183-189.
- Sy,K.T.L., dkk. (2021). Population Density and Basic Reproductive Number of Covid-19 Across United States Counties. *PLoS ONE* 16(4) : e0249271.
- Walpole, Ronald E. (1988). *Pengantar Statistika* Edisi ke-3. Jakarta: Gramedia
- Wong, D.W.S., & Li, Yun. (2020). Spreading of Covid-19 : Density matters. *PLoS ONE* 15(12) : e0242398.

## ***Geographically Weighted Spline Nonparametric Regression dengan Fungsi Pembobot Bisquare dan Gaussian Pada Tingkat Pengangguran Terbuka Di Pulau Kalimantan***

Hillidatul Ilmi<sup>(1)</sup>, Sifriyani<sup>(2)</sup>, Surya Prangga<sup>(3)</sup>

Universitas Mulawarman

Jl. Barong Tongkok No. 4 Kampus Gunung Kelua, Samarinda, Kalimantan Timur 75123

e-mail: [hillidatul08ilmi@gmail.com](mailto:hillidatul08ilmi@gmail.com)<sup>(1)</sup>, [sifriyani@fmipa.unmul.ac.id](mailto:sifriyani@fmipa.unmul.ac.id)<sup>(2)</sup>,  
[suryaprangge@gmail.com](mailto:suryaprangge@gmail.com)<sup>(3)</sup>

### **ABSTRAK**

*Geographically weighted spline nonparametric regression* merupakan pengembangan regresi nonparametrik untuk data spasial dengan estimator parameter bersifat lokal setiap lokasi pengamatan yang diaplikasikan pada kasus tingkat pengangguran terbuka. Tingkat pengangguran terbuka menjadi alat ukur kualitas kesejahteraan di suatu wilayah yang mengindikasikan besarnya persentase penduduk usia kerja yang aktif secara ekonomi. Tujuan penelitian ini yaitu untuk mengidentifikasi faktor-faktor yang mempengaruhi tingkat pengangguran terbuka 56 Kabupaten/Kota di Kalimantan. Metode yang digunakan adalah *geographically weighted spline nonparametric regression* dengan pembobot fungsi kernel eksponensial. Model terbaik *geographically weighted spline nonparametric regression* dengan pembobot fungsi kernel eksponensial pada orde 1 titik knot 1 dengan nilai *R-Square* sebesar 86,410 persen, nilai AIC sebesar 12,152, nilai RMSE sebesar 0,584 serta nilai CV terkecil adalah fungsi kernel bisquare sebesar 77,175. Adapun faktor-faktor yang berpengaruh signifikan terhadap tingkat pengangguran terbuka yaitu tingkat partisipasi angkatan kerja, jumlah penduduk, indeks pembangunan manusia, harapan lama sekolah dan upah minimum.

**Kata Kunci :** Fungsi kernel bisquare, Fungsi kernel gaussian, *geographically weighted spline nonparametric regression*, tingkat pengangguran terbuka, titik knot.

### **ABSTRACT**

*Geographically weighted spline nonparametric regression* is a nonparametric regression development for spatial data with a local parameter estimator for each observation location which is applied to the case of the open unemployment rate. The open unemployment rate is a measure of the quality of welfare in an area which indicates the large percentage of the working age population who are economically active. The purpose of this study is to identify the factors that influence the open unemployment rate in 56 regencies/cities in Kalimantan. The method used is *geographically weighted spline nonparametric regression* with weighted exponential kernel function. The best model is *geographically weighted spline nonparametric regression* with an exponential kernel function weighting on the order of 1 point knot 1 with an *R-Square* value of 86,410 percent, an AIC value of 12,152, an RMSE value of 0.584 and the smallest CV value is a bisquare kernel function of 77.175. The factors that have a significant effect on the open unemployment rate are the level of labor force participation, population, human development index, school year expectations and minimum wages.

**Keywords :** bisquare kernel function, gaussian kernel function, *geographically weighted spline nonparametric regression*, open unemployment rate, knot points.

### **1. PENDAHULUAN**

Analisis regresi yang digunakan untuk memodelkan data spasial dikenal model *Geographically Weighted Regression* (GWR). GWR pertama kali diperkenalkan oleh Fotheringham pada tahun 1967. GWR adalah salah satu model spasial dengan vektor titik. GWR merupakan pengembangan dari model regresi linier menjadi model regresi terboboti dengan

memperhatikan efek spasial, sehingga menghasilkan penduga parameter yang hanya dapat digunakan untuk memprediksi setiap titik atau lokasi di mana data tersebut diamati dan disimpulkan. Melalui model GWR akan menghasilkan interpretasi setiap titik lokasi yang diteliti akan berbeda-beda (Fotheringham, dkk, 2002).

Model GWR yang dikembangkan masih dalam bentuk linear, yang hanya mampu mengatasi masalah analisis regresi spasial yang kurva regresinya diketahui dan linier. Dalam beberapa kenyataan yang dihadapi dalam analisis regresi tidak semua data berpola linier dan kurva regresinya diketahui. Model regresi yang baik, sebaiknya dipandang dari berbagai aspek dan menempatkan suatu persoalan pemodelan tepat pada porsinya. Perbedaan karakteristik lingkungan dan geografis antar lokasi pengamatan, mengakibatkan pengamatan memiliki variasi yang berbeda atau terdapat perbedaan pengaruh variabel prediktor terhadap variabel respon untuk setiap lokasi pengamatan. Bagaimana cara menyelesaikan apabila pengaruh variabel prediktor terhadap variabel respon tidak mengikuti pola tertentu dan terdapat pola yang berubah-ubah pada sub-sub interval tertentu. Dalam situasi ini model GWR belum mampu mengatasi masalah ini, sehingga sangat dipertimbangkan untuk mengembangkan penelitian dibidang regresi *nonparametric* dan statistika spasial.

Para ahli statistik mengembangkan penelitian tentang GWR diantaranya (Nakaya, Fotheringham, Brunson dan Charlton, 2005) mengembangkan *Geographically Weighted Poisson Regression Models*, selanjutnya (Mei, Wang dan Zhang, 2006) menghasilkan model yang menggabungkan model regresi global dan GWR yang dikenal *Mixed Geographically Weighted Regression Models*. Model GWR juga dikembangkan dalam bidang spasial yang berhubungan dengan timeseries yang dikenal dengan spatio-temporal, eksplorasi data spatio-temporal dianalisis dengan menggunakan GWR dan *analitic geovisual*, penelitian ini dilakukan oleh (Demsar, Fotheringham dan Charlton, 2008). Penelitian spatio-temporal juga digunakan oleh (Huang, Wu dan Barry, 2010) dalam *Geographically And Temporally Weighted Regression Model*. Selanjutnya (Yu, 2010) penelitian pada data panel spasial dengan menggunakan *Geographically Weighted Panel Regression*, dan penelitian oleh (Wrenn dan sam, 2014) digunakan model *Geographically and Temporally Weighted Likelihood Regression*. Penelitian sebelumnya mengembangkan model regresi *nonparametric* dalam GWR atau *geographically weighted spline nonparametric regression* (Sifriyani, 2018a).

Model *geographically weighted spline nonparametric regression* diaplikasikan untuk mendeteksi penyebaran kesejahteraan masyarakat melalui Tingkat Pengangguran Terbuka (TPT)

dalam konteks spasial. Kajian epidemiologi spasial yang membahas bagaimana pengaruh bidang pendidikan, kependudukan dan ekonomi mempengaruhi kenaikan atau penurunan TPT (BPS, 2020). Kajian epidemiologi spasial yang membahas bagaimana pengaruh bidang pendidikan, kependudukan dan ekonomi mempengaruhi kenaikan atau penurunan TPT. Penelitian ini tentang analisis TPT dalam Skala Daerah yang dapat membantu merumuskan kebijakan wilayah untuk 56 Kabupaten/Kota di Pulau Kalimantan dalam menurunkan penyebaran TPT yang berkelanjutan dan memperkuat pemahaman tentang epidemiologi spasial ekonomi tersebut. Artikel ini membahas identifikasi faktor-faktor yang mempengaruhi TPT 56 Kabupaten/kota di Pulau Kalimantan berdasarkan faktor di bidang pendidikan, bidang kependudukan dan ekonomi. Pada penelitian ini digunakan data penelitian TPT 56 Kabupaten/Kota di Pulau Kalimantan Tahun 2019.

Kebaharuan dalam penelitian ini, digunakan pengembangan model *geographically weighted spline nonparametric regression* dengan optimasi GCV yang menggunakan dua basis matriks fungsi polynomial dan fungsi *truncated*. Selanjutnya diperoleh uji signifikansi simultan dan parsial yang mendukung *model geographically weighted spline nonparametric regression*. Hasil model *geographically weighted spline nonparametric regression* setiap kelompok memiliki faktor pengaruh yang berbeda sehingga penanganan TPT setiap Kabupaten/Kota berbeda. Pengelompokan berdasarkan faktor-faktor yang mempengaruhi TPT yaitu Tingkat Partisipan Angkatan Kerja (TPAK), jumlah penduduk, Indeks Pembangunan Manusia (IPM), Harapan Lama Sekolah (HLS) dan upah minimum. Penelitian ini dapat memberikan kontribusi di bidang ekonomi dan pemerintah dapat menggunakan atau lebih memperhatikan faktor-faktor tersebut dalam membangun tingkat kesejahteraan di Pulau Kalimantan.

## 2. METODE PENELITIAN

### 2.1 Data dan Sumber Data

Data yang digunakan dalam penelitian ini adalah data cross section yang bersumber dari Badan Pusat Statistika Tahun 2019 dan Dinas Ketenagakerjaan Tahun 2019. Deskripsi dan Sumber data pengamatan dalam penelitian ditampilkan pada Tabel 1.

**Tabel 1.** Deskripsi Variabel dan Sumber Data Pengamatan

Variabel		Deskripsi Variabel	Satuan	Skala
Respon	y	TPT	Persen	Kontinu
Prediktor	x <sub>1</sub>	TPAK	Persen	Kontinu
	x <sub>2</sub>	Jumlah Penduduk	Jiwa	Diskrit
	x <sub>3</sub>	IPM	Persen	Kontinu
	x <sub>4</sub>	HLS	Tahun	Kontinu
	x <sub>5</sub>	Upah Minimum	Rupiah	Kontinu

**2.2 Analisis Data**

Analisis pada penelitian ini dilakukan dengan bantuan *microsoft excel*, *software R 3.6.3* dan *software octave*. Berdasarkan tujuan dalam penelitian ini, maka langkah-langkah sebagai berikut :

1. Membuat analisis statistika deskriptif dari masing-masing variabel yang terdiri dari mapping area, rata-rata, nilai minimum dan nilai maksimum.
2. Analisis pola linier menggunakan *scatterplot* antara variabel respon dengan masing-masing variabel prediktor.
3. Melakukan pendekteksian multikolineritas menggunakan nilai *variance inflation factor* (VIF) (Ghozali, 2016).

$$VIF = \frac{1}{1 - R_k^2} \tag{1}$$

Dengan  $R_k^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2}$ , dimana  $\hat{y}_i$  adalah

nilai prediksi model regresi nonparametrik variabel respon ke-*i* dan  $\bar{y}$  adalah nilai rata-rata variabel respon.

4. Heterogenitas spasial dengan metode *Breusch Pagan*, hipotesis : (Kosfeld, 2006).

$$H_0 : \sigma_1^2 = \sigma_2^2 = \dots = \sigma_n^2 = \sigma^2$$

$$H_1 : \text{Paling tidak ada satu } \sigma_1^2 \neq \sigma_n^2$$

dengan statistika uji :

$$BP = \frac{1}{2} \mathbf{f}^T \mathbf{Z} (\mathbf{Z}^T \mathbf{Z})^{-1} \mathbf{Z}^T \mathbf{f} \tag{2}$$

Dengan  $\mathbf{f} = \left( \frac{\varepsilon_i^2}{\sigma^2} - 1 \right)$  dimana vektor

$\mathbf{f} = (f_1, f_2, \dots, f_n)^T$  dan  $\varepsilon_i = y_i - \hat{y}_i$ . Matriks  $\mathbf{Z}$  merupakan matriks yang berisi vektor pengamatan

variabel respon. Tolak  $H_0$  jika  $BP > \chi_{(k)}^2$  atau  $p - \text{value} < \alpha$  dengan  $k$  adalah banyaknya prediktor.

5. Menghitung jarak *eucliden* antara lokasi ke-*i* yang terletak pada koordinat terhadap lokasi ke-*j* (Fotheringham, dkk, 2002).

$$d_{ij} = \sqrt{(u_i - u_j)^2 + (v_i - v_j)^2} \tag{3}$$

$d_{ij}$  adalah jarak *euclidian* antara lokasi ( $u_i, v_i$ ) dan lokasi ( $u_j, v_j$ ).

6. Menghitung matriks pembobot geografis, dengan menggunakan fungsi pembobot georafis (Sifriyani, dkk, 2019), (Sifriyani, dkk, 2019a), (Sifriyani, dkk, 2019b). Fungsi Kernel *Gaussian*

$$w_j(u_i, v_i) = \exp \left( -\frac{1}{2} \left( \frac{d_{ij}}{h} \right)^2 \right) \tag{4}$$

dan Fungsi Kernel *Bisquare*

$$w_j(u_i, v_i) = \begin{cases} \left( 1 - \left( \frac{d_{ij}}{h} \right)^2 \right)^2, & \text{untuk } d_{ij} \leq h \\ 0, & \text{untuk } d_{ij} > h \end{cases} \tag{5}$$

$h$  adalah parameter non-negatif yang diketahui dan biasanya disebut parameter penghalus (*bandwidth*).

7. Menentukan titik knot optimum dengan menggunakan metode pengembangan *Generalized Cross Validation* (GCV) (Eubank, 1998).

$$GCV = \frac{MSE(K)}{\left( n^{-1} \text{trace}[\mathbf{I} - A(K)] \right)^2} \tag{6}$$

Dengan  $MSE(K) = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{f}(x_i))^2$ ,  $\mathbf{I}$  adalah

matriks identitas,  $K$  adalah titik knot  $K_1, K_2, \dots, K_r$ ,  $A(K)$  adalah matriks

$\mathbf{X}(\mathbf{X}^T \mathbf{VX})^{-1} \mathbf{X}^T \mathbf{V}$  dan *trace* adalah jumlah dari elemen diagonal utama dari matriks kuadrat.

8. Melakukan estimator model *geographically weighted spline nonparametric regression* dengan pembobot geografis (Sifriyani, dkk, 2018a).

$$\begin{aligned} \hat{\boldsymbol{\beta}}(u_i, v_i) &= \mathbf{A}(\mathbf{K})\mathbf{y} \\ \hat{\boldsymbol{\delta}}(u_i, v_i) &= \mathbf{B}(\mathbf{K})\mathbf{y} \end{aligned} \tag{7}$$

Dimana  $\hat{\boldsymbol{\beta}}(u_i, v_i)$  merupakan estimator dari parameter  $\boldsymbol{\beta}(u_i, v_i)$  dengan vektor ukuran

$(1 + (d \times p)) \times 1$  dan  $\hat{\delta}(u_i, v_i)$  estimator dari parameter  $\delta(u_i, v_i)$  dengan vektor ukuran  $(d \times r) \times 1$ .

9. Melakukan pengujian kesesuaian model, hipotesis : (Sifriyani, dkk, 2018b).

$$H_0 : \beta_{kf}(u_i, v_i) = \beta_{kf} \text{ dan } \delta_{k,m+s}(u_i, v_i) = \delta_{k,m+s}$$

$$H_1 : \text{Paling tidak ada satu } \beta_{kf}(u_i, v_i) \neq \beta_{kf} \text{ atau } \delta_{k,m+s}(u_i, v_i) \neq \delta_{k,m+s}, k = 1, \dots, d$$

$f = 1, \dots, p; s = 1, \dots, r; i = 1, \dots, n$  dengan Statistika uji :

$$V = \frac{\mathbf{y}^T \mathbf{S} \mathbf{y}}{\frac{n - lm - 1}{\mathbf{y}^T \mathbf{D}(u_i, v_i) \mathbf{y}} \text{tr}((\mathbf{I} - \xi)^T (\mathbf{I} - \xi))} \quad (8)$$

dimana

$$\mathbf{D}(u_i, v_i) = (\mathbf{I} - \mathbf{W}(u_i, v_i) \mathbf{Q} (\mathbf{Q}^T \mathbf{W}(u_i, v_i) \mathbf{Q})^{-1} \mathbf{Q}^T) (\mathbf{I} - \mathbf{Q} (\mathbf{Q}^T \mathbf{W}(u_i, v_i) \mathbf{Q})^{-1} \mathbf{Q}^T \mathbf{W}(u_i, v_i))$$

$$\xi = \mathbf{Q} (\mathbf{Q}^T \mathbf{W}(u_i, v_i) \mathbf{Q})^{-1} \mathbf{W}(u_i, v_i)$$

Uji hipotesis untuk kesesuaian model, dengan  $\mathbf{y}$  adalah vektor variabel respon,  $n$  adalah banyaknya variabel respon,  $l$  adalah banyaknya variabel prediktor,  $m$  adalah orde fungsi *spline* dan  $\mathbf{I}$  adalah matriks identitas. Kriteria penolakan  $H_0$  adalah tolak  $H_0$  jika  $F_{hitung}(V) > F_{(\alpha; db_1; db_2)}$  atau  $p - value < \alpha$ .

10. Melakukan pengujian Signifikan Parameter Simultan, hipotesis : (Sifriyani, dkk, 2018b).

$$H_0 : \beta_{11}(u_i, v_i) = \beta_{12}(u_i, v_i) = \dots = \beta_{dp}(u_i, v_i) = \delta_{1,m+1}(u_i, v_i) = \delta_{1,m+2}(u_i, v_i) = \dots = \delta_{d,m+r}(u_i, v_i)$$

$$H_1 : \text{Paling tidak ada satu } \beta_{kf}(u_i, v_i) \neq 0 \text{ atau } \delta_{k,m+s}(u_i, v_i) \neq 0, k = 1, \dots, d; f = 1, \dots, p; s = 1, \dots, r; i = 1, \dots, n$$

dengan statistika uji :

$$V^* = \frac{\left( \frac{\mathbf{y}^T \mathbf{M}(u_i, v_i) \mathbf{y}}{\text{tr}((\mathbf{I} - \mathbf{B}_\omega)^T (\mathbf{I} - \mathbf{B}_\omega))} \right)}{\left( \frac{\mathbf{y}^T \mathbf{D}(u_i, v_i) \mathbf{y}}{\text{tr}((\mathbf{I} - \xi)^T (\mathbf{I} - \xi))} \right)} \quad (9)$$

Dimana  $\mathbf{M}(u_i, v_i) = (\mathbf{I} - \mathbf{B}_\omega)^T (\mathbf{I} - \mathbf{B}_\omega)$  dengan  $\mathbf{y}$  adalah vektor variabel respon dan  $\mathbf{I}$  adalah matriks identitas. Dan kriteria penolakan  $H_0$  adalah tolak  $H_0$  jika  $F_{hitung}(V^*) > F_{(\alpha; db_1; db_2)}$  atau  $p - value < \alpha$  (Sifriyani, Hayatmi, Budiantara, dkk., 2018b).

11. Melakukan pengujian hipotesis parsial atau pengujian pengaruh lokasi secara parsial pada setiap variabel prediktor, hipotesis : (Sifriyani, dkk, 2018b).

$$H_0 : \beta_{kf}(u_i, v_i) = 0 \text{ dan } \delta_{k,m+s}(u_i, v_i) = 0$$

$$H_1 : \text{Paling tidak ada satu } \beta_{kf}(u_i, v_i) \neq 0 \text{ atau}$$

$$\delta_{k,m+s}(u_i, v_i) \neq 0; k = 1, \dots, d; f = 1, \dots, p;$$

$$s = 1, \dots, r; i = 1, \dots, n$$

dengan statistika uji :

$$t = \frac{\hat{\eta}(u_i, v_i)}{SE(\hat{\eta}(u_i, v_i))} \quad (10)$$

$\hat{\eta}(u_i, v_i)$  adalah Estimator dari  $\eta(u_i, v_i)$  dan  $SE(\hat{\eta}(u_i, v_i))$  adalah Elemen diagonal ke  $k+1$  dari matrik  $(\mathbf{Q}^T \mathbf{W}(u_i, v_i) \mathbf{Q})^{-1} \hat{\sigma}^2(u_i, v_i)$ ,  $\mathbf{Q}$  adalah matriks ukuran  $n \times [1 + (d \times p) + (d \times r)]$  yang memuat variabel prediktor dan variabel prediktor yang mempunyai komponen *spline truncated*.

12. Mendapatkan faktor-faktor yang mempengaruhi TPT 56 Kabupaten/Kota di Pulau Kalimantan.

13. Kesimpulan.

### 3. HASIL DAN PEMBAHASAN

#### 3.1 Statistika Deskriptif

Deskripsi data penelitian meliputi analisis statistika deskriptif berupa nilai rata rata, nilai minimum serta nilai maksimum. Statistika deskriptif disajikan pada Tabel 1.

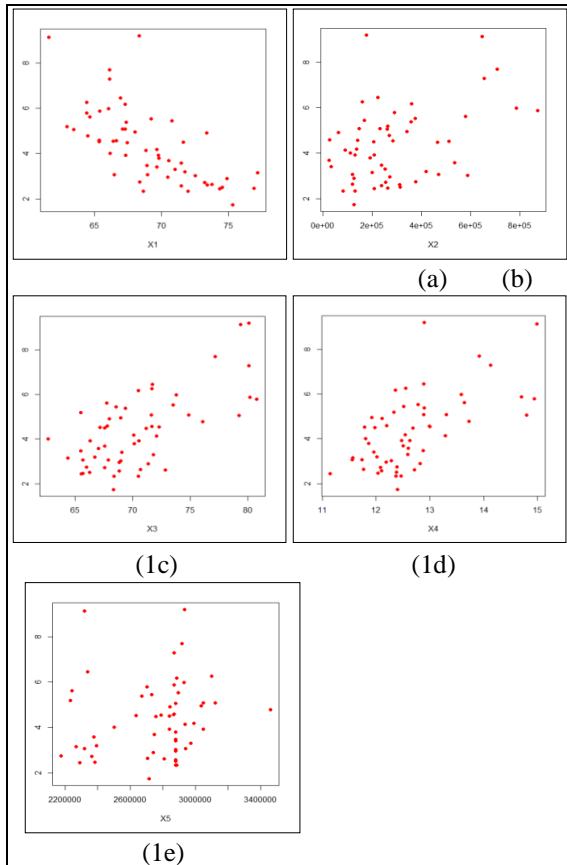
**Tabel 1.** Statistika Deskriptif Data Penelitian

	Rata-rata	Minimum	Maksimum
$y$	4,355	1,740	9,190
$x_1$	69,04	61,62	77,19
$x_2$	288.967	26.400	872.800
$x_3$	70,41	62,66	80,77
$x_4$	12,67	11,15	14,99
$x_5$	2.753.530	2.175.200	3.462.192

Berdasarkan Tabel 1 Statistika deskripsi data penelitian TPT dari 56 Kabupaten/Kota di Kalimantan dapat dilihat bahwa rata-rata yang mewakili sekumpulan data TPT adalah 4,355 persen dengan TPT terendah ada di Kabupaten Pulang Pisau sebesar 1,740 persen dan tertinggi ada di Kota Bontang.

**3.2 Pola Hubungan Antar Variabel**

Identifikasi pola hubungan menggunakan scatterplot, dimana data penelitian dapat berupa regresi parametrik, regresi semiparametrik atau regresi nonparametrik dengan melihat suatu pola tertentu (linier, kuadratik atau kubik). Berikut adalah hasil scatterplot pada TPT dengan beberapa variabel prediktor yang disajikan pada Gambar 1 :



**Gambar 1.** Pola Hubungan Variabel Respon dengan Variabel Prediktor

Pola hubungan yang terbentuk antara variabel respon dengan masing-masing variabel prediktor tidak mengikuti pola tertentu dan terdapat pola yang berubah-ubah pada sub-sub interval tertentu. Perubahan ini dikarenakan setiap daerah menggunakan karakteristik yang berbeda-beda. Oleh karena itu *metode geographically weighted spline nonparametric regression* layak digunakan.

**Tabel 2.** Pendekteksian Multikolinieritas

Nilai VIF				
$x_1$	$x_2$	$x_3$	$x_4$	$x_5$
1,551	1,456	4,132	4,370	1,488

Pendekteksian multikolinieritas bertujuan untuk menguji apakah adanya korelasi antara variabel prediktor. Pendekteksian multikolinieritas menunjukkan bahwa tidak terjadi kolerasi karena nilai VIF <10, sehingga seluruh variabel prediktor bisa digunakan. Sedangkan pengujian heterogenitas spasial bertujuan untuk mengetahui apakah ada hubungan antara variabel respon dengan lokasi geografis.

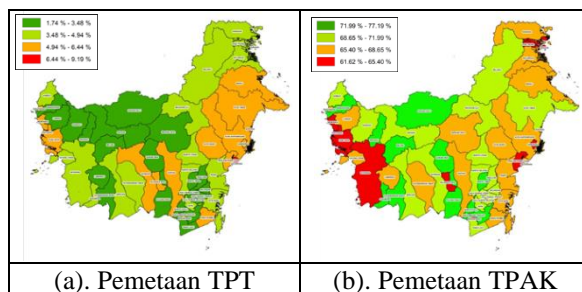
**Tabel 3.** Breusch-Pagan Test

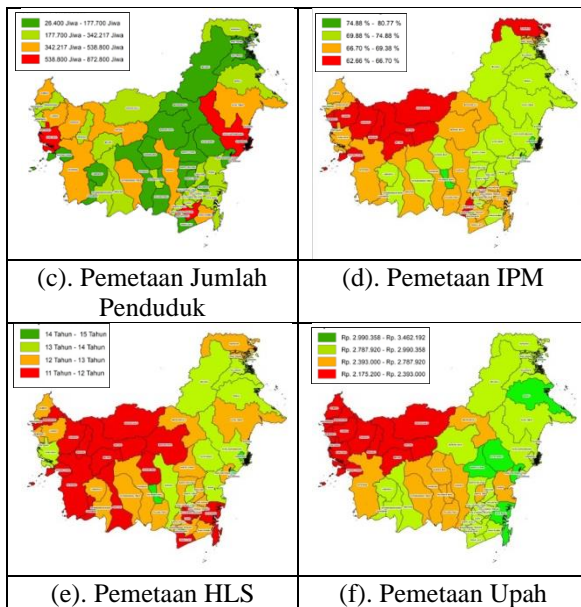
Breusch-Pagan	Df	P-Value
19,988	10	0,048

Hasil nilai maka keputusan ditolak. Pengujian aspek spasial terpenuhi yaitu terdapat efek heterogenitas spasial maka kasus tingkat pengangguran terbuka dapat diselesaikan dengan menggunakan pendekatan titik.

**3.3 Peta Penyebaran Spasial**

Peta penyebaran spasial bertujuan untuk melihat visualisasi data penelitian yang akan dikelompokkan secara pemetaan berdasarkan nilai masing-masing variabel setiap wilayah 56 Kabupaten/Kota di Kalimantan. Berikut masing-masing peta penyebaran data penelitian :





Gambar 1. Pemetaan Variabel Respon dan Variabel Prediktor 56 Kabupaten/Kota di Kalimantan

### 3.4 Pemilihan Model Terbaik dan Titik Knot Optimum

Pemilihan kriteria model terbaik dengan melihat nilai *R-squared* terbesar dari semua orde dan titik knot.

Tabel 4. Kriteria Model Terbaik

Orde	Titik Knot	<i>R-squared</i>	AIC	RMSE
1	1	<b>86,411</b>	12,152	0,584
	2	82,340	11,482	0,690
	3	81,319	11,303	0,722

Berdasarkan kriteria *R-squared*, AIC serta RMSE dapat dilihat bahwa model orde 1 dan titik knot 1 adalah model *geographically weighted spline nonparametric regression* yang terbaik, dengan nilai *R-squared* sebesar 86,411 persen. Berdasarkan kriteria model terbaik maka diperlukan satu titik knot yang optimum menggunakan metode GCV untuk setiap variabel prediktor.

Tabel 4. Pemilihan Titik Knot Optimum dengan Satu Titik Knot

No	$x_1$	$x_2$	$x_3$	$x_4$
1	71,585	568100	74,250	13,608
2	69,716	466530	72,077	13,147
3	69,872	474990	72,258	13,185
4	71,429	559630	74,069	13,569

5	69,561	458060	71,896	13,108
No	$x_5$	GCV		
1	2998900	<b>1,182</b>		
2	2844400	1,184		
3	2857300	1,185		
4	2986000	1,186		
5	2831600	1,187		

Nilai GCV terkecil pada pemilihan titik knot optimum dengan satu titik knot yaitu 1,182. Sehingga didapatkan titik-titik knot optimum yaitu:  $K_{11} = 71,585$ ,  $K_{21} = 56.100$ ,  $K_{31} = 74,250$ ,  $K_{41} = 13,608$  dan  $K_{51} = 2.998.900$ .

### 3.5 Pembobot Fungsi Kernel

Matriks pembobot Spasial dengan fungsi Kernel Gaussian dan fungsi Kernel Bisquare pada lokasi pengamatan ke-1, ke-2 sampai ke-56.

Tabel 4. Nilai CV untuk Fungsi Pembobot

Fungsi Pembobot	Nilai <i>Bandwidth</i>	GCV
Fungsi Kernel Gaussian	2,000	77,803
Fungsi Kernel Bisquare	4,451	<b>77,175</b>

Dari dua fungsi pembobot tersebut didapatkan fungsi pembobot terbaik adalah pembobot fungsi kernel bisquare dengan nilai GCV lebih kecil dibandingkan dengan nilai GCV dari fungsi kernel gaussian sebesar 77,175.

### 3.6 Uji Kesesuaian Model

Bentuk hipotesis uji kesesuaian model sebagai berikut :

$$H_0 : \beta_{kf}(u_i, v_i) = \beta_{kf} \text{ dan } \delta_{k,m+s}(u_i, v_i) = \delta_{k,m+s}$$

$$H_1 : \text{Paling tidak ada satu } \beta_{kf}(u_i, v_i) \neq \beta_{kf} \text{ atau}$$

$$\delta_{k,m+s}(u_i, v_i) \neq \delta_{k,m+s}; k = 1, \dots, 5; f = 1;$$

$$s = 1; i = 1, \dots, 56$$

Pengujian kesesuaian model antara menggunakan Persamaan (8), diperoleh  $V(1,756) > F_{(0,05;21;36)}(1,756)$  atau *p-value* (0,041) < 0,05 maka keputusan  $H_0$  ditolak. Oleh karena itu terdapat pengaruh faktor geografis pada model yang sesuai digunakan adalah *geographically weighted spline nonparametric regression*.

### 3.7 Uji Signifikansi Parameter Simultan

Bentuk hipotesis uji kesesuaian model sebagai berikut :

$$H_0 : \beta_{1,1}(u_i, v_i) = \beta_{1,2}(u_i, v_i) = \dots = \beta_{1,5}(u_i, v_i) = \delta_{1,6}(u_i, v_i) = \dots = \delta_{1,10}(u_i, v_i) = 0.$$

$H_1$  : Paling tidak ada satu  $\beta_{kf}(u_i, v_i) \neq 0$  atau

$$\delta_{k,m+s}(u_i, v_i) \neq 0; k = 1, \dots, 5; f = 1; s = 1; i = 1, \dots, 56$$

Pengujian signifikansi parameter simultan antara menggunakan Persamaan (9),  $V(4,035) > F_{(0,05;55;32)}(1,724)$  atau  $p\text{-value } (2,698 \times 10^{-5}) < 0,05$  maka keputusan  $H_0$  ditolak. Sehingga dapat diperoleh kesimpulan bahwa variabel prediktor pada *geographically weighted spline nonparametric regression* berpengaruh signifikan terhadap variabel TPT.

### 3.8 Uji Signifikansi Parameter Parsial

Hasil perhitungan dari pengujian signifikansi parameter secara parsial menunjukkan bahwa variabel prediktor yang berpengaruh berbeda-beda untuk setiap area. Hal ini mengakibatkan terdapat 14 kelompok yang terbentuk. Pembagian 14 kelompok yang terbentuk berdasarkan variabel prediktor yang signifikan dapat dijelaskan sebagai berikut :

1. Faktor-faktor yang berpengaruh (TPAK, jumlah penduduk, IPM, HLS dan upah minimum) terhadap TPT (Kelompok 1) dilokasi pengamatan Mempawah, Sanggau, Sintang, Kota Pontianak, Tanah Laut, Kotabaru, Tapin, Pulang Pisau, Gunung Mas, Barito Timur, Kota Palangka Raya, Malinau, Bulungan, Tana Tidung, Kota Nunukan dan Kota Tarakan.
2. Faktor-faktor yang berpengaruh (TPAK, jumlah penduduk, IPM dan HLS) terhadap TPT (Kelompok 2) dilokasi pengamatan Melawi, Kubu Raya, Paser, Penajam Paser Utara dan Kota Balikpapan..
3. Faktor-faktor yang berpengaruh (TPAK, jumlah penduduk, IPM dan upah minimum) terhadap TPT (Kelompok 3) dilokasi pengamatan Barito Kuala dan Kapuas.
4. Faktor-faktor yang berpengaruh (TPAK, jumlah penduduk, HLS dan upah minimum) terhadap TPT (Kelompok 4) dilokasi pengamatan Kotawaringin Timur.
5. Faktor-faktor yang berpengaruh (TPAK, IPM, HLS dan upah minimum) terhadap TPT (Kelompok 5) dilokasi pengamatan Kutai Timur dan Berau.
6. Faktor-faktor yang berpengaruh (TPAK, jumlah penduduk dan HLS) terhadap TPT (Kelompok 6) dilokasi pengamatan Kayong Utara dan Kotawaringin Barat.

7. Faktor-faktor yang berpengaruh (TPAK, IPM dan HLS) terhadap TPT (Kelompok 7) dilokasi pengamatan Kutai Barat, Kota Samarinda dan Kota Bontang.

8. Faktor-faktor yang berpengaruh (Jumlah penduduk, IPM dan upah minimum) terhadap TPT (Kelompok 8) dilokasi pengamatan Hulu Sungai Utara, Barito Selatan, Barito Utara dan murung Raya.

9. Faktor-faktor yang berpengaruh (TPAK dan IPM) terhadap TPT (Kelompok 9) dilokasi pengamatan Sambas, Bengkayang, Kota Singkawang, Kutai Kartanegara dan Mahakam Ulu.

10. Faktor-faktor yang berpengaruh (TPAK dan jumlah penduduk) terhadap TPT (Kelompok 10) dilokasi Seruyan.

11. Faktor-faktor yang berpengaruh (Jumlah penduduk dan HLS) terhadap TPT (Kelompok 11) dilokasi pengamatan Kapuas Hulu

12. Faktor-faktor yang berpengaruh (TPAK dan HLS) terhadap TPT (Kelompok 12) dilokasi Seruyan.

13. Faktor-faktor yang berpengaruh (IPM) terhadap TPT (Kelompok 13) dilokasi pengamatan Hulu Sungai Tengah, Tabalong dan Banjarbaru.

14. Faktor-faktor yang Tidak Signifikan berpengaruh terhadap TPT (Kelompok 14) dilokasi pengamatan Landak, Ketapang, Kota Banjar, Hulu Sungai Selatan, Tanah Bumbu, Kota Banjarmasin, Sukamara, Lamandau dan Katingan.

## 4. KESIMPULAN DAN SARAN

Model *geographically weighted spline nonparametric regression* dengan pendekatan *spline truncated* yang terbaik pada data tingkat pengangguran terbuka di Kalimantan Tahun 2019 adalah orde  $f = 1$  dan titik knot  $s = 1$ . Nilai *R-squared* sebesar 86,410 persen, hal ini mengindikasikan bahwa model mampu menjelaskan pengaruh variabel prediktor terhadap variabel respon TPT sebesar 86,410 persen sedangkan 13,590 persen sisanya dijelaskan oleh variabel lain yang tidak diteliti. Hasil pengujian parsial dari estimasi parameter mengelompokkan 56 Kabupaten/Kota menjadi 14 kelompok berdasarkan variabel prediktor yang berpengaruh terhadap TPT di Kalimantan Tahun 2019.

Masing-masing setiap titik lokasi pengamatan 56 Kabupaten/Kota di Kalimantan memiliki model yang berbeda-beda. Salah satu model *geographically weighted spline nonparametric regression* dari 56 Kabupaten/Kota di Kalimantan yaitu Kabupaten Pulang Pisau yang memiliki TPT terendah dari 56 area di Kalimantan :

$$\hat{y}_{37} = 0,009936 - 0,10157x_{37,1} - 0,30999(x_{37,1} - 71,585)_+ + 1,49 \times 10^{-6} x_{37,2} + 5,97 \times 10^{-6} (x_{37,2} - 568100)_+ + 0,16685x_{37,3} - 0,20796(x_{37,3} - 74,25)_+ + 0,25129x_{37,4} - 0,01111(x_{37,4} - 13,608)_+ - 1,42 \times 10^{-6} x_{37,5} + 9,82 \times 10^{-6} (x_{37,5} - 2998900)_+$$

Interpretasi : Untuk kenaikan TPAK sebesar 1 persen di Kabupaten Pulang maka TPT mengalami kenaikan sebesar 21,479 persen, untuk kenaikan jumlah penduduk sebesar 1.000 jiwa di Kabupaten Pulang maka TPT mengalami kenaikan sebesar 0,009 persen, untuk kenaikan IPM sebesar 1 persen di Kabupaten Pulang maka TPT mengalami kenaikan sebesar 0,031 persen, untuk kenaikan HLS sebesar 1 tahun di Kabupaten Pulang maka TPT mengalami kenaikan sebesar 0,250 persen dan untuk kenaikan upah minimum sebesar Rp1.000.000,00 di Kabupaten Pulang maka TPT mengalami kenaikan sebesar 0,009 persen.

Fungsi pembobot yang digunakan pada penelitian ini yaitu Fungsi kernel Gaussian dan fungsi kernel Bisquare, sehingga untuk penelitian selanjutnya dapat menggunakan pembobot fungsi kernel ekponensial dan fungsi kernel Tricube. Data yang digunakan pada penelitian ini adalah data tahun 2019. Untuk penelitian selanjutnya dapat menggunakan data keluaran terbaru.

#### DAFTAR PUSTAKA

Demsar, Fotheringham dan Charlton, 2008, Exploring the spatio-temporal dynamics of geographical processes with geographically weighted regression and geovisual analytics, *Inference Vis*, 7. 181–197.

BPS Kalimantan Barat. (2019). *Keadaan Angkatan Kerja*. Badan Pusat Statistik Provinsi Kalimantan Barat.

BPS Kalimantan Barat. (2020a). *Provinsi Kalimantan Barat dalam Angka*. Badan Pusat Statistik Provinsi Kalimantan Barat.

BPS Kalimantan Barat. (2020b). *Proyeksi Penduduk Kabupaten/Kota*. Badan Pusat Statistik Provinsi Kalimantan Barat.

BPS Kalimantan Selatan. (2019). *Keadaan Angkatan Kerja*. Badan Pusat Statistik Provinsi Kalimantan Selatan.

BPS Kalimantan Selatan. (2020). *Provinsi Kalimantan Selatan dalam Angka*. Badan Pusat Statistik Provinsi Kalimantan Selatan.

BPS Kalimantan Tengah. (2019a). *Keadaan Angkatan Kerja*. Badan Pusat Statistik Provinsi Kalimantan Tengah.

BPS Kalimantan Tengah. (2019b). *Statistika Ketenagakerjaan*. Badan Pusat Statistik Provinsi Kalimantan Tengah.

BPS Kalimantan Tengah. (2020a). *Provinsi Kalimantan Tengah dalam Angka*. Badan Pusat Statistik Provinsi Kalimantan Tengah.

BPS Kalimantan Tengah. (2020b). *Proyeksi Penduduk Kabupaten/Kota*. Badan Pusat Statistik Provinsi Kalimantan Tengah.

BPS Kalimantan Timur. (2019a). *Keadaan Angkatan Kerja*. Badan Pusat Statistik Provinsi Kalimantan Timur.

BPS Kalimantan Timur. (2019b). *Indeks Pembangunan Manusia*. Badan Pusat Statistik Provinsi Kalimantan Timur.

BPS Kalimantan Timur. (2020). *Provinsi Kalimantan Timur dalam Angka*. Badan Pusat Statistik Provinsi Kalimantan Timur.

BPS Kalimantan Utara. (2019). *Keadaan Angkatan Kerja*. Badan Pusat Statistik Provinsi Kalimantan Utara.

BPS Kalimantan Utara. (2020). *Provinsi Kalimantan Utara dalam Angka*. Badan Pusat Statistik Provinsi Kalimantan Utara.

Dinas Ketenagakerjaan. (2020). Penetapan Upah Minimum. Dinas Ketenagakerjaan Kalimantan Barat.

Dinas Ketenagakerjaan. (2020). Upah Minimum Regional. Dinas Ketenagakerjaan Kalimantan Tengah.

Draper dan Smith. (1992). *Applied Regression Analysis*, Second Edition. New York: John Wiley & Sons.

Eubank, R.L. (1998). *Spline Smoothing and Nonparametric Regression*. New York: Marcel Dekker.

Fotheringham, A.S., Brundson, C dan Charlton, M. (2002). *Geographically Weighted Regression*. New York: John Wiley and Sons.

Huang, Wu, dan Barry, 2010, Geographically and temporally weighted regression for modeling spatio-temporal variation in house prices. *International Journal Geographical Inference Science*. 24, 383–401.

Kosfeld, Reinhold. (2006). *Spatial Econometric*. Institute of Economics. University of Kassel.

Mei, Wang dan Zhang, 2006, Testing the importance of the explanatory variables in a mixed geographically weighted regression model, *Journal Environment and Planning*, 38(3), 587-598.

- Nakaya, Fotheringham, Brunsdon dan Charlton, 2005, Geographically Weighted Poisson Regression for disease association mapping, *Journal Statistics Medicine*, 24(17), 2695-2717
- Sifriyani., Haryatmi., Budiantara, I.N dan Gunardi.. (2018a). Development of Nonparametric Geographically Weighted Regression Using Truncated Spline Approach. *Songklanakar J. Sci. Technol*, Vol. 40(4), pp. 909-920.
- Sifriyani., Haryatmi., Budiantara, I.N dan Gunardi.. (2018a). (2018b). A New Method of Hypotesis Test for Truncated Spline Nonparametric Regression Influenced by Spatial Heterogenity and Application. *Abstract and Applied Analysis*, 2018. doi: 10.1155/2018/9769150.
- Sifriyani, Ruslan, F.H. Susanty. 2019(a). Evaluation of Forest Productivity and Governance on the Preservation of Tropical Rain Forests in Kalimantan Using The NGWR-TS Nonparametric Geospatial Method. *EurAsian Journal of BioSciences*, Vol 13 Issue 2, pp. 2373-2379.
- Sifriyani, Ruslan, F.H. Susanty. 2019(b). Mapping and Analysis Factors of Affecting Productivity Tropical Rain Forests in East Kalimantan, *Modern Applied Science*, Vol 13 No 10, DOI: 10.5539/mas.v13n10p112
- Sifriyani, Budiantara, I. N., Kartiko, S.H., and Gunardi. (2019). Evaluation of Factors Affecting Increased Unemployment in East Java Using NGWR-TS Method, *International Journal of Sciences: Basic and Applied Research*. 49(1), 123-142.
- Yu, D., 2010, Exploring spatiotemporally varying regressed relationships: the geographically weighted panel regression analysis. *Proceedings of the Joint International Conference on Theory, Data Handling and Modeling in GeoSpatial Information Science*, 134–139.

## Analisis Multilevel Kemiskinan Rumah Tangga Pertanian di Provinsi Sulawesi Selatan Tahun 2020

Volandio Ardhian Rastantra  
Badan Pusat Statistik Kabupaten Bulukumba  
Jl. Jendral Ahmad Yani No. 18 Bulukumba, Sulawesi Selatan 92512  
e-mail: [ardhian.rastantra@bps.go.id](mailto:ardhian.rastantra@bps.go.id)

### ABSTRAK

Sulawesi Selatan sebagai daerah agraris memiliki potensi sektor pertanian yang besar. Namun, masih terdapat beberapa permasalahan yang dihadapi oleh rumah tangga pertanian seperti rendahnya kesejahteraan petani dan tingginya tingkat kemiskinan penduduk yang bekerja di sektor pertanian. Terdapat perbedaan yang terlihat dalam persentase penduduk miskin yang bekerja di sektor pertanian pada masing-masing kabupaten/kota. Adanya struktur hierarki individu dan wilayah pada data yang digunakan, maka pemodelan harus mengakomodasi korelasi individu di dalam satu wilayah supaya estimasi parameter akan lebih efisien. Metode regresi multilevel sesuai dengan data berstruktur hierarki, selain itu dapat menjelaskan adanya keragaman dalam model. Karena variabel dependen berupa data kategorik dengan dua kategorik, maka penelitian ini menggunakan regresi multilevel logistik biner untuk menganalisis variabel rumah tangga dan wilayah yang memengaruhi kemiskinan rumah tangga pertanian serta mengetahui variasi wilayah terhadap kemiskinan. Dari penelitian diperoleh hasil bahwa variabel rumah tangga yang signifikan memengaruhi kemiskinan rumah tangga pertanian antara lain pendidikan KRT, jenis kelamin KRT, jumlah ART, tipe daerah, kepemilikan lahan, dan akses layanan keuangan. Jika dilihat dari variabel kabupaten/kota, IPM signifikan memengaruhi kemiskinan rumah tangga pertanian. Sebesar 11,56 persen keragaman status kemiskinan rumah tangga pertanian di Provinsi Sulawesi Selatan disebabkan oleh perbedaan karakteristik antar kabupaten/kota.

**Kata kunci :** *Kemiskinan, Rumah Tangga Pertanian, Regresi Multilevel, Provinsi Sulawesi Selatan.*

### ABSTRACT

*South Sulawesi as agricultural area has large potential for agricultural sector. However there are still some problems faced by agricultural households such as low welfare of farmers and high level of poverty of the population working in agricultural sector. There is difference in the percentage of poor working in agricultural sector in each district/city. There is hierarchical structure of individuals and region in data used, then modelling must accommodate individual correlations in area so parameter estimation be more efficient. Multilevel regression method fit data with hierarchical structure, besides that it can explain diversity in model. Because dependent variable is categorical data with two categories, this study uses binary multilevel logistic regression to analyze household and regional variables that affect agricultural households poverty and determine regional variations in poverty. From the research, it is found that household variables have a significant effect on agricultural household poverty include household head education, gender of household head, number of household members, type of area, land ownership, and access to financial services. From regional variable, HDI significantly affects agricultural household poverty. 11.56 percent of the diversity in the poverty status of agricultural households in South Sulawesi Province is caused by differences in characteristics between districts/cities.*

**Keywords :** *Poverty, Agricultural Household, Multilevel Regression, South Sulawesi Province.*

### 1. PENDAHULUAN

Sebagai daerah agraris, Sulawesi Selatan merupakan lumbung pangan beras nasional dan wilayah penghasil komoditas penting sektor pertanian lainnya. Sektor pertanian memiliki peranan penting dalam pembangunan di Sulawesi Selatan. Peranan tersebut digambarkan melalui

sumbangan lapangan usaha pertanian terhadap Produk Domestik Regional Bruto (PDRB) Sulawesi Selatan sebesar 21,70 persen pada tahun 2020 (BPS, 2020). Angka tersebut menjadikan pertanian memiliki share PDRB terbesar dibandingkan lapangan usaha lain seperti industri, konstruksi, dan perdagangan.

Berdasarkan hasil Sakernas Agustus 2020, sebanyak 1,93 juta penduduk bekerja di sektor pertanian atau sekitar 39,78 persen dari jumlah penduduk bekerja di Sulawesi Selatan. Hal itu menjadikan sektor pertanian unggul dibandingkan sektor lain dalam penyerapan tenaga kerja. Namun, dibalik potensi sektor pertanian yang begitu besar, terdapat beberapa masalah yang dihadapi oleh pekerja sektor pertanian di Sulawesi Selatan. Salah satunya adalah masih rendahnya Nilai Tukar Petani (NTP) Provinsi Sulawesi Selatan pada tahun 2020, yaitu sebesar 96,97. Hal tersebut menunjukkan bahwa petani mengalami defisit dan mengindikasikan rendahnya kesejahteraan petani. Masalah lain yang dihadapi adalah tingginya tingkat kemiskinan pada rumah tangga pertanian. Pada tahun 2020 terdapat 30,46 persen penduduk usia 15 tahun ke atas berstatus miskin dan bekerja di sektor pertanian (BPS, 2020). Lebih besar dibandingkan penduduk yang bekerja di sektor selain pertanian dan berstatus miskin sekitar 9 persen. Selain itu, persentase penduduk miskin di masing-masing kabupaten/kota pada sektor pertanian menunjukkan adanya perbedaan yang cukup terlihat. Hal ini membuat dugaan adanya pengaruh wilayah terhadap status kemiskinan rumah tangga pertanian.

Struktur data hierarki berasal dari populasi yang memiliki tingkatan dan berkelompok dalam satu unit. Data yang berstruktur secara hierarki memiliki efek dari kelompok yang membawahi unit penelitian (Hox, 2010). Adanya struktur hierarki individu dan wilayah pada data yang digunakan, maka pemodelan harus mengakomodasi korelasi individu di dalam satu wilayah supaya estimasi parameter akan lebih efisien. Goldstein (1995) memperkenalkan analisis regresi multilevel sebagai pengembangan untuk mengatasi masalah yang disebabkan oleh data yang memiliki struktur hierarki.

Model regresi multilevel adalah bagian dari *Generalized Linear Mixed Model* (GLMM) dimana dalam persamaan modelnya menggabungkan *fixed effect* dan *random effect*. Model ini pada dasarnya merupakan versi bertingkat model regresi berganda untuk menganalisis data dengan struktur hierarki (Hox, 2010). Metode ini sering digunakan dalam kasus-kasus sosial terutama untuk data berstruktur hierarki yang sering ditemukan pada penelitian survei dimana penarikan sampel dilakukan secara bertahap (*cluster sampling*). Selain itu, metode multilevel juga mampu menjelaskan keragaman dalam model.

Hox (2010) menyebutkan bahwa terdapat dua bentuk dasar model multilevel, yaitu model dengan *random intercept* dan model dengan *random slope*. Model *random intercept* memodelkan dimana setiap kelompok memiliki *intercept* yang berbeda, sedangkan model *random slope* memodelkan variabel independen pada level yang rendah sebagai *random effect* sehingga tiap kelompok memiliki garis kemiringan yang berbeda.

Dalam model regresi multilevel jika variabel respon memiliki distribusi binomial, maka dalam pendugaan parameter menggunakan fungsi hubungan logit ( $\ln\{\pi/(1-\pi)\}$ ) sehingga dinamakan model regresi multilevel logistik biner. Penelitian ini menggunakan model regresi multilevel logistik biner *random intercept* dengan asumsi pengaruh masing-masing variabel adalah sama pada level kabupaten/kota. Selain itu, estimasi model *random slope* kurang dapat dipercaya menurut Hox (2010). Penelitian yang membahas tentang pengaruh wilayah dalam kemiskinan rumah tangga pertanian di Indonesia masih belum banyak dilakukan. Mursalina dan Melly (2016) dalam kesimpulannya menyatakan variabel yang signifikan memengaruhi kemiskinan rumah tangga pertanian antara lain variabel umur KRT, status kepemilikan rumah, jumlah anggota rumah tangga, ukuran rumah, sumber air minum, sumber bahan utama memasak, dan kepemilikan sepeda motor. Selanjutnya, penelitian oleh Prayitno (2019) dihasilkan kesimpulan bahwa umur, pendidikan, rasio ketergantungan, kepemilikan lahan, akses kredit, dan kegagalan panen signifikan memengaruhi rumah tangga pertanian berstatus miskin di Indonesia.

Uraian latar belakang di atas mendasari pentingnya melakukan penelitian yang mengkaji hierarki tentang kemiskinan rumah tangga pertanian. Tujuan penelitian antara lain untuk mengetahui variabel pada level rumah tangga maupun level wilayah yang signifikan memengaruhi kemiskinan rumah tangga pertanian di Provinsi Sulawesi Selatan tahun 2020 serta mengetahui variasi wilayah terhadap kemiskinan.

## 2. METODE PENELITIAN

Penelitian ini menggunakan data sekunder berupa data mentah hasil Susenas KOR bulan Maret tahun 2020 serta data yang sudah diolah oleh BPS pada tahun 2020. Terdapat sampel sebanyak 7.088 rumah tangga pertanian dari Provinsi Sulawesi Selatan yang digunakan dalam penelitian. Kemiskinan rumah tangga pertanian merupakan variabel dependen dalam penelitian. Variabel

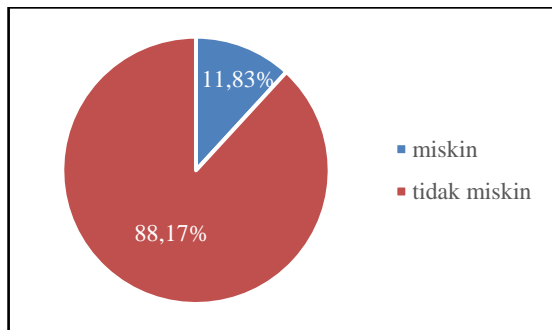
independen pada level rumah tangga adalah pendidikan KRT (X1), jenis kelamin KRT (X2), jumlah ART (X3), tipe daerah (X4), kepemilikan lahan (X5), dan akses layanan keuangan (X6), sedangkan pada level wilayah adalah rasio ketergantungan (X7) dan Indeks Pembangunan Manusia (X8).

Analisis inferensia yang digunakan adalah regresi multilevel logistik biner dengan *random intercept*. Tahapan analisis yang dilakukan dalam penelitian ini adalah sebagai berikut:

1. Melakukan analisis deskriptif kemiskinan rumah tangga pertanian
2. Melakukan pemeriksaan kasus multikolinearitas menggunakan kriteria VIF
3. Pengujian signifikansi *random effect*
4. Melakukan uji signifikan parameter secara simultan
5. Melakukan uji signifikan parameter secara parsial
6. Penghitungan *odds ratio*
7. Penghitungan *Intraclass Correlation (ICC)*

**3. HASIL DAN PEMBAHASAN**

Berdasarkan hasil pengolahan data Susenas Maret 2020, diketahui bahwa terdapat 11,83 persen rumah tangga pertanian di Sulawesi Selatan yang miskin dan 88,17 persen sisanya merupakan rumah tangga pertanian tidak miskin.



**Gambar 1.** Persentase kemiskinan rumah tangga pertanian di Sulawesi Selatan

Hasil karakteristik rumah tangga pertanian menurut variabel independen level rumah tangga didapatkan 10,98 persen rumah tangga pertanian miskin tinggal di daerah perdesaan dan 0,84 persen berstatus miskin di daerah perkotaan. Terdapat 10,47 persen rumah tangga pertanian dengan pendidikan tertinggi KRT  $\leq$  SMP yang berstatus miskin. Hal yang cukup menarik adalah 9,16 persen rumah tangga pertanian yang memiliki lahan di Sulawesi Selatan berstatus miskin.

Sedangkan, rumah tangga yang tidak memiliki lahan hanya 2,67 persen yang berstatus miskin. Hasil karakteristik rumah tangga pertanian di Provinsi Sulawesi Selatan disajikan secara menyeluruh pada Tabel 1. berikut:

**Tabel 1.** Karakteristik rumah tangga pertanian

Variabel	Kategori	Status Kemiskinan	
		Miskin	Tidak Miskin
Pendidikan KRT	$\leq$ SMP	10,47 %	89,53 %
	$>$ SMP	1,36 %	98,64 %
Jenis Kelamin KRT	Perempuan	2,16 %	97,84 %
	Laki-Laki	9,66 %	90,34 %
Jumlah ART	$\leq$ 4	4,19 %	95,81 %
	$>$ 4	7,64 %	92,36 %
Tipe Daerah	Perdesaan	10,98 %	89,02 %
	Perkotaan	0,84 %	99,16 %
Kepemilikan Lahan	Tidak	2,67 %	97,33 %
	Ya	9,16 %	90,84 %
Akses Layanan Keuangan	Tidak	10,30 %	89,70 %
	Ya	1,53 %	98,47 %

Sebelum analisis regresi, dilakukan pemeriksaan asumsi multikolinearitas. Hasil pengujian asumsi multikolinearitas disajikan dalam Tabel 2. berikut.

**Tabel 2.** Pengujian asumsi multikolinearitas

Variabel	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	X <sub>6</sub>
VIF	1,02	1,02	1,03	1,03	1,01	1,03

Berdasarkan Tabel 2. diketahui bahwa seluruh variabel independen memiliki nilai VIF dibawah 10, hal ini menunjukkan tidak adanya multikolinearitas dan dapat dilanjutkan ke analisis regresi.

Dalam analisis regresi, dilakukan pengujian *random effect* untuk mengetahui kecocokan metode multilevel untuk data yang digunakan dalam penelitian. Pengujian signifikansi *random effect* menghasilkan nilai *Likelihood Ratio Test* sebesar 204,47. Nilai tersebut lebih besar dari  $\chi^2_{0,05;1} = 3,84$  dan dapat diambil kesimpulan adanya *random effect* yang signifikan dengan tingkat kepercayaan 95%.

Selain itu, dibandingkan nilai AIC pada kondisi tanpa variabel penjelas (*null model*) antara model regresi logistik biner satu level dan regresi multilevel. Hasil yang didapat menunjukkan bahwa nilai AIC pada regresi multilevel lebih kecil daripada regresi satu level ketika kondisi tanpa variabel penjelas (*null model*). Sehingga, regresi multilevel logistik biner lebih baik untuk digunakan dalam menjelaskan data.

**Tabel 3.** Hasil nilai AIC pada *null model*

Model	AIC
Regresi satu level	4758,858
Regresi multilevel	4556,386

Untuk mengetahui adanya paling sedikit satu variabel independen yang berpengaruh secara signifikan terhadap status kemiskinan rumah tangga pertanian, dilakukan pengujian signifikansi parameter secara simultan. Statistik uji parameter secara simultan diperoleh dengan,

$$G = -2 \ln \left( \frac{L(\text{null model})}{L(\text{conditional model})} \right) = -2(-2276,1932 - (2082,4344)) = 387,5176$$

Berdasarkan pengujian parameter secara simultan, diperoleh nilai  $G$  yang lebih besar dari  $\chi^2_{0,05;8} = 15,51$ . Didapatkan kesimpulan bahwa terdapat setidaknya satu variabel independen yang signifikan dalam memengaruhi kemiskinan rumah tangga pertanian di Provinsi Sulawesi Selatan dengan tingkat kepercayaan 95%.

Selain menguji secara simultan, dilakukan uji parsial untuk memperoleh informasi signifikansi dan pengaruh setiap variabel independen terhadap kemiskinan rumah tangga pertanian. Uji parsial menggunakan hipotesis sebagai berikut:

$$H_0: \beta_k = 0$$

$$H_1: \beta_k \neq 0, k = 1, 2, \dots, 8$$

Tabel 4. menunjukkan hasil statistik uji estimasi parameter secara parsial dalam model regresi *random intercept* multilevel logistik biner. Terdapat enam variabel pada level rumah tangga dan satu variabel pada level wilayah yang memiliki pengaruh signifikan terhadap kemiskinan rumah tangga pertanian berdasarkan hasil pengujian parameter secara parsial.

Variabel level rumah tangga yang signifikan memengaruhi kemiskinan rumah tangga pertanian antara lain pendidikan KRT, jenis kelamin KRT,

jumlah ART, tipe daerah, kepemilikan lahan, dan akses layanan keuangan. Sedangkan pada level kabupaten/kota, variabel IPM signifikan memengaruhi kemiskinan rumah tangga pertanian.

**Tabel 4.** Hasil signifikansi parameter dan *odds ratio*

Variabel	Koefisien	S.E.	p-value	Odds Ratio
Konstanta	5,896	4,116	0,152	363,5813
Pendidikan KRT > SMP <= SMP	-0,637	0,126	0,000*	0,529
Jenis Kelamin Laki-laki Perempuan	-0,498	0,113	0,000*	0,608
Jumlah ART > 4 <= 4	1,345	0,087	0,000*	3,840
Tipe Daerah Perkotaan Perdesaan	-0,862	0,156	0,000*	0,422
Pemilikan Lahan Ya Tidak	-0,634	0,111	0,000*	0,531
Akses Layanan Keuangan Ya Tidak	-0,849	0,123	0,000*	0,428
Rasio Ketergantungan	0,085	0,056	0,125	1,089
IPM	-0,167	0,056	0,003*	0,846

Keterangan: \* adalah signifikan pada  $\alpha = 5\%$

Berdasarkan Tabel 4, dihasilkan persamaan regresi multilevel logistik biner sebagai berikut:

$$\ln \left( \frac{\hat{\pi}_{ij}}{1 - \hat{\pi}_{ij}} \right) = 5,896 - 0,637 \text{ pendidikan}_{ij} - 0,498 \text{ jenis kelamin}_{ij} + 1,345 \text{ jumlah ART}_{ij} - 0,862 \text{ tipe daerah}_{ij} - 0,634 \text{ lahan}_{ij} - 0,849 \text{ akses keuangan}_{ij} + 0,085 \text{ Rasio Ketergantungan}_{ij} - 0,167 \text{ IPM}_{ij}$$

Berdasarkan hasil pengujian signifikansi secara parsial, seluruh variabel pada level rumah

tangga signifikan memengaruhi kemiskinan rumah tangga pertanian di Provinsi Sulawesi Selatan. Pendidikan KRT berpengaruh terhadap status kemiskinan dengan nilai koefisien  $-0,637$ . Rasio kecenderungan yang diperoleh sebesar  $0,529$  artinya rumah tangga pertanian dengan KRT yang memiliki pendidikan tertinggi diatas SMP cenderung untuk miskin dibandingkan rumah tangga yang dikepalai oleh KRT dengan pendidikan tertinggi SMP kebawah sebesar  $0,529$  kali.

Jenis kelamin KRT signifikan dalam berpengaruh terhadap status kemiskinan rumah tangga pertanian. Rasio kecenderungan untuk miskin rumah tangga dengan laki-laki sebagai KRT sebesar  $0,608$  dibandingkan rumah tangga dengan KRT perempuan. Atau dapat dinyatakan rumah tangga dengan KRT perempuan memiliki kecenderungan lebih besar untuk miskin senilai  $1,645$  kali dibandingkan rumah tangga yang berjenis kelamin KRT laki-laki.

Selanjutnya, rumah tangga pertanian yang memiliki ART berjumlah lebih dari 4 mempunyai kecenderungan sebesar  $3,840$  kali untuk berstatus miskin daripada rumah tangga dengan jumlah ART tidak lebih dari 4. Rasio kecenderungan rumah tangga yang tinggal di daerah perkotaan untuk berstatus miskin sebesar  $0,422$  atau dapat dinyatakan rumah tangga pertanian yang tinggal di perdesaan mempunyai kecenderungan  $2,370$  kali untuk miskin dibandingkan rumah tangga pertanian yang tinggal di perkotaan.

Rumah tangga pertanian yang memiliki lahan mempunyai rasio kecenderungan  $0,531$ , maka kecenderungan rumah tangga yang tidak memiliki lahan untuk miskin sebesar  $1,883$  kali dibandingkan rumah tangga dengan status kepemilikan lahan. Sedangkan ditinjau dari akses layanan keuangan, rumah tangga pertanian yang tidak mendapatkan akses layanan keuangan berupa kredit memiliki kecenderungan untuk miskin sebesar  $2,336$  lebih besar dibandingkan yang mendapat akses layanan keuangan.

Dari variabel kabupaten/kota, ditemukan bahwa variabel Indeks Pembangunan Manusia signifikan memengaruhi status kemiskinan rumah tangga pertanian di Provinsi Sulawesi Selatan. Nilai negatif koefisien IPM menunjukkan bahwa kenaikan nilai IPM berdampak pada kemiskinan rumah tangga pertanian yang semakin berkurang dalam keadaan variabel penjelas lainnya konstan. Sedangkan rasio ketergantungan tidak signifikan memengaruhi kemiskinan rumah tangga pertanian di Provinsi Sulawesi Selatan.

Nilai ICC model regresi multilevel logistik biner yang diperoleh sebesar  $0,1156$ . Yang berarti

adanya perbedaan karakteristik antar kabupaten/kota menyebabkan keragaman status kemiskinan rumah tangga pertanian di Provinsi Sulawesi Selatan sebesar  $11,56$  persen.

#### 4. KESIMPULAN DAN SARAN

Kesimpulan yang diperoleh berdasarkan hasil dan pembahasan antara lain terdapat  $11,83$  persen rumah tangga pertanian berstatus miskin di Sulawesi Selatan tahun 2020. Perbedaan karakteristik antar kabupaten/kota menyebabkan adanya keragaman status kemiskinan rumah tangga pertanian di Provinsi Sulawesi Selatan sebesar  $11,56$  persen.

Saran bagi penelitian selanjutnya, dapat menggunakan metode spasial serta dapat menambahkan variabel lain yang diduga memengaruhi kemiskinan rumah tangga pertanian.

#### DAFTAR PUSTAKA

- Badan Pusat Statistik. (2014). *Analisis Sosial Ekonomi Petani di Indonesia*. Jakarta: BPS.
- Badan Pusat Statistik Provinsi Sulawesi Selatan. (2021). *Data dan Informasi Kemiskinan Provinsi Sulawesi Selatan 2020*. Makassar: BPS.
- Badan Pusat Statistik Provinsi Sulawesi Selatan. (2021). *Indeks Pembangunan Manusia Provinsi Sulawesi Selatan 2020*. Makassar: BPS
- Badan Pusat Statistik Provinsi Sulawesi Selatan. (2021). *Indikator Pasar Tenaga Kerja Provinsi Sulawesi Selatan 2020*. Makassar: BPS.
- Harlan, J. (2016). *Analisis Multilevel*. Depok: Gunadarma.
- Hox, J. J. (2010). *Multilevel Analysis Techniques and Applications (2nd ed)*. New York: Routledge.
- Jiang, Y., Huang, C., Yin, D., Liang, C., Wang, Y. (2020). Constructing HLM to examine multilevel poverty-contributing factors of farmer households: Why and how?. *Plos ONE*, 15(1): e0228032. <https://doi.org/10.1371/journal.pone.0228032>.
- Litbang Pertanian. (2019). *Strategi Penanggulangan Kemiskinan di Pertanian dan Perdesaan [Policy Brief]*. Jakarta: Badan Litbang Pertanian.
- Mursalina & Melly. (2016). Pengaruh Faktor Sosial-Ekonomi Terhadap Status Kemiskinan Rumah Tangga Pertanian di Aceh. *Jurnal Agriseip*, 17(2).

- Okunola, A. M. & Ojo, O. S. (2020). Household Poverty Measurement and Its Determinants Among Rural Farmers in Ondo State, Nigeria. *Journal Poverty & Public Policy*, 11(4), 277-290
- Prayitno, H. L. (2019). Kemiskinan Rumah Tangga Pertanian di Indonesia [Skripsi thesis, Universitas Airlangga. UNAIR Repository. <http://repository.unair.ac.id/id/eprint/86825>.

---

**Analisis Volatilitas Harga Saham Sekor Minyak dan Gas di Indonesia pada Masa Pandemi Covid-19 dengan Metode ARIMA-GARCH**Nanda Septiana<sup>(1)</sup>, Primadina Hasanah<sup>(2)</sup>, Annisa Rahmita Soemarsono<sup>(3)</sup>

Institut Teknologi Kalimantan

Jl. Soekarno-Hatta Km. 15, Karang Joang Telp. (0542) 8530800 Balikpapan 76127

e-mail: [nanda18septiana@gmail.com](mailto:nanda18septiana@gmail.com), [primadina@lecturer.itk.ac.id](mailto:primadina@lecturer.itk.ac.id), dan [annisarahmitas@lecturer.itk.ac.id](mailto:annisarahmitas@lecturer.itk.ac.id)

---

**ABSTRAK**

Pandemi Covid-19 memberi dampak yang signifikan terhadap berbagai sektor industri di Indonesia salah satunya saham sektor pertambangan Minyak Mentah dan Gas Bumi (MIGAS). Hal ini ditunjukkan pada penurunan harga minyak yang turun di bawah \$40 USD dan aktivitas eksplorasi di Indonesia menurun lebih dari 40% dibanding sebelum pandemi Covid-19. Selama pandemi, harga saham sektor pertambangan MIGAS mengalami volatilitas yang cukup tinggi sehingga cukup meresahkan sektor investasi di Indonesia. Volatilitas merupakan naik turunnya harga saham atau sekuritas dalam kurun waktu tertentu. Oleh karena itu, diperlukan suatu prediksi volatilitas harga saham sektor pertambangan MIGAS agar mampu memberikan informasi terhadap investor untuk melakukan manajemen portofolio. Pada penelitian ini, dianalisis volatilitas harga saham empat perusahaan pertambangan MIGAS, yaitu PT. Apexindo Pratma Duta (APEX), PT. Elnusa (ELSA), PT. Medco Energi Internasional (MEDC), dan PT. Radiant Utama Interinsco (RUIS) pada tanggal 01 Maret 2020 - 28 Februari 2021 dengan metode ARIMA-GARCH. ARIMA-GARCH merupakan metode analisis data dengan memodelkan suatu data stasioner dengan model *Autoregressive* (AR) dan *Moving Average* (MA) yang memiliki gejala heteroskedastisitas pada model ARIMA. Pada proses analisis, digunakan RStudio dengan pembentukan model ARIMA dilakukan terlebih dahulu kemudian dilanjutkan pembentukan model ARIMA-GARCH jika model ARIMA terdapat gejala heteroskedastisitas. Hasil dari penelitian ini, pada saham APEX, ELSA, dan RUIS terdapat gejala heteroskedastisitas pada model ARIMA dan didapatkan model ARIMA(0,1,1)GARCH(1,1) untuk perusahaan APEX, ELSA dan RUIS serta model ARIMA(4,1,4) untuk perusahaan MEDC. Berdasarkan hasil analisis, diperoleh bahwa terdapat asumsi autokorelasi, normalitas, dan heteroskedastisitas yang belum terpenuhi pada uji diagnostik. Nilai MAPE untuk APEX, ELSA, MEDC, dan RUIS, yaitu 67.32667%, 42.49374%, 5.269889%, dan 4.113383%. Dari hasil akurasi peramalan yang didapatkan, terdapat nilai MAPE di atas 10%, yaitu pada model APEX dan ELSA sehingga model tersebut belum dapat dikatakan baik untuk peramalan.

**Kata kunci :** ARIMA, GARCH, Volatilitas Harga Saham

**ABSTRACT**

*The Covid-19 pandemic has significant impact on various industrial sectors in Indonesia, one of them is the Crude Oil and Natural Gas (Oil and Gas) stocks. The impact was showed by the conditional of oil prices that have fallen below \$40 USD and exploration activities in Indonesia has decreased by more than 40% compared by condition before the pandemic. During the pandemic the stock price of Oil and Gas mining sector have experienced high volatility which is confusing the investement sectors in Indonesia. Volatility is the fluctuation of stocks over a certain period of time. Therefore, a prediction of stock price volatility in oil and gas mining sectors is needed in order to provide information for the investors to manage their portofolio. In this study, the stock prices volatility of four oil and gas companies was analyzed, namely PT. Apexindo Pratma Duta (APEX), PT. Elnusa (ELSA), PT. Medco Energi Internasional (MEDC), and PT. Radiant Utama Interinsco (RUIS) on 01 March 2020 - 28 February 2021 using the ARIMA-GARCH method. ARIMA-GARCH is a data analysis method by modeling stationary data with Autoregressive (AR) and Moving Average (MA) models which have heteroscedasticity in the ARIMA model. In the analysis process, Rstudio is used with the formation of the ARIMA was conducted first then it is*

continue by the ARIMA-GARCH model if the ARIMA model has the heteroscedasticity. The results of this study, have obtained the ARIMA(0,1,1)GARCH(1,1) model for APEX, ELSA and RUIS and the ARIMA(4,1,4) model for MEDC. Based on the results, it is found that there were autocorrelation, normality, and heteroscedasticity the residual that have not been met in the diagnostic test. MAPE values for APEX, ELSA, MEDC, and RUIS were 67.32667%, 42.49374%, 5.269889%, and 4.113383%. From the results of forecasting accuracy obtained that were MAPE value above 10%, namely in APEX and ELSA models so that the model is not suggested for the forecasting.

**Keywords :** ARIMA, GARCH, Stock Price Volatility.

## 1. PENDAHULUAN

Banyak sektor yang terpuuk pada saat Covid-19 mewabah di Indonesia. Sektor pertambangan khususnya pertambangan Minyak Mentah dan Gas Bumi (MIGAS) dilihat dari penurunan harga minyak yang turun di bawah \$40 USD dan aktivitas eksplorasi di Indonesia menurun lebih dari 40% dibanding sebelum pandemi Covid-19. Menurut Menteri Keuangan Republik Indonesia, sektor manufaktur dan pertambangan minyak mentah dan gas bumi terpuuk saat Covid-19, di mana permintaan menurun sangat signifikan. Hal ini berimbas pada penurunan saham pada pertambangan minyak mentah dan gas bumi (Widyastuti dan Hanan, 2020).

Menurut Organisasi untuk Kerja sama dan Pembangunan Ekonomi atau *Organization for Economic Co-operation and Development* (OECD), Covid-19 berpotensi menimbulkan krisis ekonomi dunia. Terbukti dengan adanya wabah virus ini telah memukul berbagai bidang ekonomi yang membawa dampak yang signifikan terhadap perdagangan di bursa. Penurunan dan peningkatan permintaan jasa berbanding lurus dengan rendah tingginya harga saham yang nantinya akan berimbas pada penurunan atau peningkatan harga saham (Darmayanti dkk., 2020). Hal ini yang disebut juga volatilitas harga saham. Volatilitas harga saham menggambarkan fluktuasi atau naik turun pada harga saham sehingga akan menunjukkan risiko pada investor (Raneo dkk., 2018).

Volatilitas tinggi menunjukkan bahwa harga saham naik dan turun dengan range yang besar. Sedangkan volatilitas rendah menunjukkan bahwa harga saham cenderung konstan atau jarang berubah. Hal ini ditunjukkan dengan turunnya harga saham perusahaan sektor pertambangan MIGAS pada awal maret yang merupakan awal mula Covid-19 menyebar di Indonesia yang mengalami volatilitas yang cukup tinggi, yaitu pada PT. Apexindo Pratama Duta Tbk (APEX) dan PT. Elnusa Tbk (ELSA) dari harga saham mencapai Rp.300 per lembar kemudian turun mejadi Rp.100 per lembar, PT. Medco Energi Internasional

(MEDC) dari Rp. 700 per lembar turun menjadi Rp. 500 per lembar, dan PT. Radiant Utama Interinsco (RUIS) dari Rp. 200 per lembar menjadi Rp.100 per lembar. Para investor menyukai saham yang memiliki volatilitas tinggi, karena dapat memungkinkan mereka memperoleh keuntungan yang besar dalam waktu yang singkat tetapi hal ini juga dapat mengakibatkan kerugian yang cukup besar. Cara untuk menghindari hal tersebut maka dibutuhkan model yang tepat untuk memprediksi volatilitas harga saham.

Beberapa penelitian sebelumnya, menganalisis volatilitas saham dengan metode dan studi kasus lain. Contoh penelitian sebelumnya mengenai model ARIMA-GARCH oleh Natsya Bella Yolanda dkk. pada tahun 2017, yaitu penerapan model ARIMA-GARCH untuk memprediksi harga saham Bank BRI. ARIMA-GARCH sendiri merupakan metode analisis data dengan memodelkan suatu data stasioner dengan model *Autoregressive* (AR) dan *Moving Average* (MA) yang memiliki gejala heteroskedastisitas pada model ARIMA. Digunakan metode ARIMA-GARCH karena pada hal volatilitas harga saham tidak terlepas dari gejala heteroskedastisitas, dimana data tersebut memiliki varian yang tidak konstan. Maka dari itu digunakan metode ARIMA-GARCH untuk mengatasi gejala heteroskedastisitas. Pada penelitiannya didapatkan kesimpulan bahwa model ARIMA-GARCH untuk prediksi harga saham BRI, yaitu ARIMA(2,1,1)GARCH(2,2) dengan nilai  $R^2$  sebesar 99.91%, di mana model dikatakan baik dalam memprediksi harga saham BRI.

Pada penelitian ini, dikaji mengenai analisis volatilitas harga saham sektor pertambangan minyak mentah dan gas bumi di Indonesia pada masa Covid-19. Studi kasus yang diambil, yaitu PT. Apexindo Pratama Duta Tbk (APEX), PT. ElnusaTbk (ELSA), PT. Medco Energi Internasional Tbk (MEDC), dan PT. Radiant Utama Interinsco Tbk (RUIS). Pada langkah analisis, terdapat unsur heteroskedastisitas pada pembentukan model ARIMA sehingga digunakan metode ARIMA-GARCH. Berdasarkan hasil

analisis, diperoleh ARIMA-GARCH terbaik untuk peramalan dengan nilai eror yang kecil. Maka dari itu, pada penelitian ini akan dianalisis tentang harga saham sektor pertambangan MIGAS, yaitu perusahaan APEX, ELSA, MEDC, dan RUIS pada masa Covid-19 dengan metode ARIMA-GARCH.

**2. METODE PENELITIAN**

Pada penelitian ini, akan dibentuk model ARIMA-GARCH dimana model tersebut secara umum dinyatakan sebagai berikut:

A. Model ARIMA

$$\phi_p(B)(1 - B)^d Y_t = \theta_q(B)e_t$$

dengan

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

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

di mana  $p$  merupakan orde AR,  $q$  merupakan orde MA,  $d$  merupakan orde differencing,  $\phi_p$  merupakan parameter AR,  $\theta_q$  parameter MA,  $e_t$  merupakan nilai residual, dan  $B$  merupakan operator *backshift*. Sebagai contoh dalam pembentukan persamaan model ARIMA, yaitu pada model ARIMA(1,1,1) sebagai berikut:

$$(1 - \phi_1 B)(1 - B)Y_t = (1 + \theta_1 B)e_t$$

$$(1 - B - \phi_1 B + \phi_1 B^2)Y_t = e_t + \theta_1 B e_t$$

$$Y_t - B Y_t - \phi_1 B Y_t + \phi_1 B^2 Y_t = e_t + \theta_1 B e_t$$

$$Y_t = B Y_t + \phi_1 B Y_t - \phi_1 B^2 Y_t + \theta_1 B e_t + e_t$$

$$Y_t = Y_{t-1} + \phi_1 Y_{t-1} - \phi_1 Y_{t-2} + \theta_1 B e_t + e_t$$

Berdasarkan Persamaan diatas, didapatkan persamaan matematis dari model ARIMA(1,1,1).

B. Model GARCH

$$\sigma_t^2 = \alpha_0 + \alpha_1 e_{t-1}^2 + \dots + \alpha_p e_{t-p}^2 + \beta_1 \sigma_{t-1}^2 + \dots + \beta_q \sigma_{t-q}^2$$

dengan  $\sigma_t^2$  merupakan variabel varian residual pada waktu  $t$ ,  $\alpha_0$  merupakan konstanta,  $\alpha_p$  koefisien ARCH pada orde ke- $p$ , dan  $e_{t-1}^2$  merupakan residual kuadrat pada waktu  $t - p$ ,  $\beta_q$  koefisien GARCH pada orde ke- $q$ , dan  $\sigma_{t-q}^2$  merupakan variabel varian residual pada waktu  $t - q$ .

Pada penelitian ini data yang digunakan dalam adalah data *daily closed price* atau harga saham penutupan harian dari 4 perusahaan sektor pertambangan MIGAS di Indonesia dan terdaftar di Bursa Efek Indonesia (BEI), yaitu yaitu PT. Apexindo Pratama Duta Tbk (APEX), PT. ElnusaTbk (ELSA), PT. Medco Energi Internasional Tbk (MEDC), dan PT. Radiant Utama Interinsco Tbk (RUIS). Data ini dapat diakses dari situs web [www.finance.yahoo.com](http://www.finance.yahoo.com) dimana data yang diambil yaitu pada periode

pandemi Covid-19 di Indonesia mulai tanggal 1 Maret 2020 sampai dengan 28 Februari 2021.

Langkah analisis dimulai dengan plot *time series* dan menghitung statistik deskriptif harga saham masing-masing perusahaan, langkah berikutnya dilakukan uji stasioneritas, yaitu stasioner terhadap mean dengan uji *Augmented Dickey-Fuller* (ADF) dan stasioner terhadap varian dengan melihat nilai lambda dari plot *Box-Cox* di mana jika lambda tidak bernilai 1 maka akan dilakukan transformasi *Box-Cox*. Data yang telah stasioner akan dilihat *Correlogram* ACF dan PACF yang digunakan untuk pendugaan orde ARIMA. Berdasarkan pendugaan orde ARIMA, dilakukan uji signifikansi parameter terhadap orde pendugaan dan orde yang telah *overfitting*. Model terbaik dipilih berdasarkan nilai *Mean Absolut Percentage Error* (MAPE) terkecil.

Pada model ARIMA terbaik diuji apakah terdapat gejala heteroskedastisitas dengan uji ARCH-LM. Jika terdapat gejala heteroskedastisitas maka akan dimodelkan dengan GARCH. Orde GARCH diperoleh dari *Correlogram* ACF dan PACF residual kuadrat dari model ARIMA. Selanjutnya dilakukan uji signifikansi dan uji diagnostik untuk model ARIMA-GARCH. Setelah didapatkan model terbaik maka dilakukan peramalan harga saham untuk 4 perusahaan dan dihitung akurasi peramalan dengan data aktual menggunakan MAPE.

**3. HASIL DAN PEMBAHASAN**

A. Statistika Deskriptif

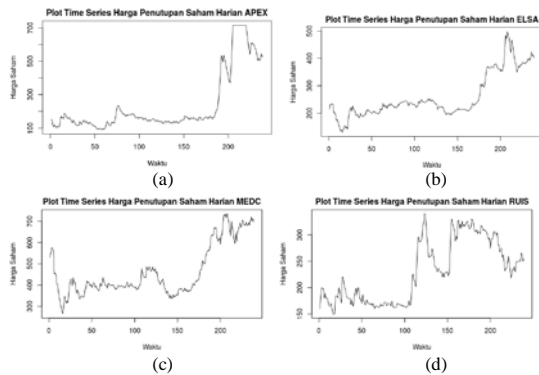
Pada penelitian ini, dilakukan statistika deskriptif untuk melihat karakteristik dari data harga penutupan saham (2020-2021). Hasil statistika deskriptif pada masing masing 4 perusahaan dapat dilihat dari Tabel 1.

**Tabel 1.** Statistika deskriptif

	APEX	ELSA	MEDC	RUIS
N	239	239	239	239
Mean	239	259.07	462.88	230.70
SD	185.62	81.418	119.85	58.415
Median	157	230	408	228
Min	89	129	264.12	150
Max	157	498	735	340

Ukuran statistika deskriptif dari 4 perusahaan ditampilkan pada Tabel 1. yang meliputi jumlah data, rata-rata, standar deviasi, median, nilai minimum, dan nilai maksimum. Pada statistika deskriptif, digunakan harga penutupan saham

harian dari tanggal 01 Maret 2020 sampai 28 Februari 2021. Pada statistika deskriptif, dapat dikatakan bahwa terjadi fluktuasi harga penutupan saham harian pada masing-masing 4 perusahaan. Untuk melihat lebih jelas, berikut plot *time series* harga penutupan saham untuk masing-masing 4 perusahaan.



**Gambar 1.** Plot *Time Series* harga penutupan saham harian (a) APEX, (b) ELSA, (c) MEDC, (d) RUIS

Berdasarkan plot *time series*, dapat dilihat bahwa terdapat pola fluktuatif pada data harga penutupan saham harian 4 perusahaan. Pada Gambar 1, dapat dilihat juga terdapat pola trend pada harga saham ELSA, MEDC, dan RUIS. Sedangkan untuk harga saham APEX cenderung konstan. Harga saham pada 4 perusahaan sektor pertambangan minyak mentah dan gas bumi ini mengalami volatilitas yang cenderung tinggi dan dapat menimbulkan resiko untuk investor yang akan berinvestasi maupun yang telah berinvestasi di salah satu 4 perusahaan, yaitu APEX, ELSA, MEDC, dan RUIS.

**B. Uji Stasioneritas**

Berdasarkan plot data, terlihat bahwa data belum stasioner, masih terlihat naik turunnya atau fluktuatif dari data harga penutupan saham 4 perusahaan. Oleh karena itu, data harus distasionerkan baik dalam rata-rata maupun varian karena dalam metode ARIMA-GARCH data yang dianalisis harus stasioner. Langkah awal dalam uji stasioner, yaitu uji stasioner terhadap rata-rata. Pada uji stasioner terhadap rata-rata data diuji dengan uji ADF dengan hipotesis  $H_0$  data belum stasioner dan taraf signifikansi  $\alpha = 0.05$ .

**Tabel 2.** Hasil uji ADF

Data	<i>p-value</i>	Kesimpulan
APEX	0.6826	Belum Stasioner

ELSA	0.4895	Belum Stasioner
MEDC	0.3498	Belum Stasioner
RUIS	0.5459	Belum Stasioner

Berdasarkan uji hipotesis untuk uji ADF tersebut didapatkan kesimpulan bahwa data harga penutupan saham harian 4 perusahaan belum stasioner terhadap rata-rata, sehingga dilakukan *differencing* atau proses diferensi agar data stasioner terhadap rata-rata. Setelah dilakukan *differencing*, maka dilakukan uji ADF kembali dari data yang sudah dilakukan *differencing*.

**Tabel 3.** Hasil uji ADF setelah *differencing*

Data	<i>p-value</i>	Kesimpulan
APEX	0.01	Stasioner
ELSA	0.01	Stasioner
MEDC	0.01	Stasioner
RUIS	0.01	Stasioner

Dapat dilihat uji hipotesis dari uji ADF didapatkan kesimpulan bahwa data stasioner. Selanjutnya dilakukan uji stasioner terhadap varian dengan melihat nilai lambda plot *Box-Cox*.

**Tabel 4.** Nilai lambda pada plot *Box-Cox*

Data	Nilai Lambda
APEX	0.552
ELSA	0.694
MEDC	1
RUIS	0.559

Dapat dilihat nilai lambda APEX, ELSA, dan RUIS tidak menunjukkan nilai 1 yang artinya belum stasioner terhadap varian, maka dari itu perlu adanya transformasi data di mana masing-masing data akan di transformasi karena nilai  $\lambda \neq 0$  dan untuk MEDC tidak perlu ada transformasi data karena nilai lambda menunjukkan nilai 1. Transformasi *Box-Cox* sebagai berikut:

$$Y_{trans} = \frac{Y^\lambda - 1}{\lambda}, \lambda \neq 0 \tag{1}$$

$$Y_{trans} = \ln Y, \lambda = 0 \tag{2}$$

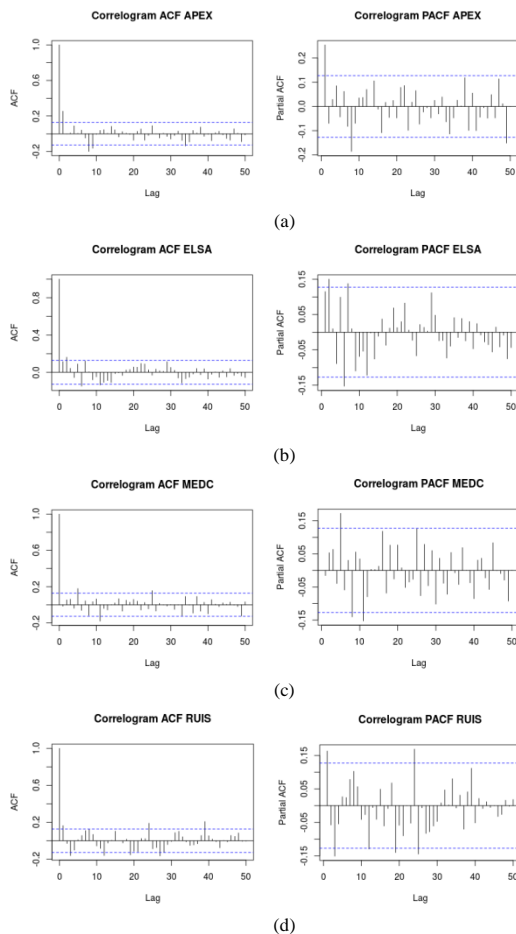
Dari persamaan (1) dan (2), untuk mengembalikan data transformasi ke data asli sebagai berikut

$$Y = ((\lambda \times Y_{trans}) + 1)^{1/\lambda}, \lambda \neq 0 \tag{3}$$

$$Y = e^{Y_{trans}} \quad , \lambda = 0 \quad (4)$$

C. Estimasi Model ARIMA

Pada estimasi model ARIMA(p,d,q) dilakukan dengan metode *maximum likelihood* dengan orde p,d, dan q dapat dilihat pada *correlogram* ACF dan PACF. Orde p merupakan orde AR yang dapat dilihat pada *correlogram* PACF, orde q merupakan orde MA dapat dilihat pada *correlogram* ACF, sedangkan orde d merupakan proses diferensi yang dilakukan pada pengolahan data. Berikut *correlogram* ACF dan PACF masing-masing perusahaan:



Gambar 2. Correlogram ACF dan PACF (a) APEX, (b) ELSA, (c) MEDC, (d) RUIS

Dalam pendugaan model dengan *correlogram* ACF dan PACF, masing-masing 4 perusahaan menunjukkan lag yang keluar batas atau *cut off*. Pada ACF APEX, terpotong pada lag 0 dan 1. Untuk PACF APEX terpotong pada lag 1. Sehingga pendugaan model ARIMA APEX, yaitu ARIMA(1,1,0) dan ARIMA(1,1,1). Pada ACF ELSA, terpotong pada lag 0 dan 2. Untuk PACF

ELSA terpotong pada lag 2. Sehingga pendugaan model ARIMA ELSA, yaitu ARIMA(2,1,0) dan ARIMA(2,1,2). Pada ACF MEDC, terpotong pada lag 0 dan 5. Untuk PACF MEDC terpotong pada lag 5. Sehingga pendugaan model ARIMA MEDC, yaitu ARIMA(5,1,0) dan ARIMA(5,1,5). Pada ACF RUIS, terpotong pada lag 0, 1, dan 3. Untuk PACF RUIS terpotong pada lag 1 dan 3. Sehingga pendugaan model ARIMA RUIS, yaitu ARIMA(1,1,0), ARIMA(1,1,1), ARIMA(1,1,3), ARIMA(3,1,0), ARIMA(3,1,1), dan ARIMA(3,1,3). Setelah didapatkan orde ARIMA dilakukan estimasi parameter untuk uji signifikan model pendugaan dan model *overfitting* dengan hipotesis  $H_0$  model tidak signifikan. Hasil estimasi parameter dari model ARIMA APEX disajikan pada Tabel 5.

Tabel 5. Hasil estimasi parameter model ARIMA APEX

Model	Parameter	Coefficient	p-value	Signifikan
ARIMA (1,1,0)	AR(1)	0.4325	$1.058 \times 10^{-13}$	Ya
	MA(1)	0.08872	0.529078	Tidak
ARIMA (0,1,1)	MA(1)	0.3902	$4.02 \times 10^{-14}$	Ya
ARIMA (0,1,2)	MA(1)	0.4484	$3.138 \times 10^{-12}$	Ya
	MA(2)	0.1508	0.01311	
ARIMA (2,1,1)	AR(1)	1.1940	$8.869 \times 10^{-8}$	Ya
	AR(2)	-0.3676	0.0002049	
	MA(1)	-0.7447	0.0008888	

Berdasarkan Tabel 5, didapatkan bahwa terdapat dua model ARIMA APEX yang signifikan, yaitu ARIMA(1,1,0), ARIMA(0,1,1), ARIMA(0,1,2), dan ARIMA(2,1,1). Hasil estimasi parameter dari model ARIMA ELSA disajikan pada Tabel 6.

Tabel 6. Hasil estimasi parameter model ARIMA ELSA

Model	Parameter	Coefficient	p-value	Signifikan
ARIMA (2,1,0)	AR(1)	0.090300	0.16075	Tidak
	AR(2)	0.157438	0.01434	
ARIMA (2,1,2)	AR(1)	0.194984	0.6864	Tidak
	AR(2)	-0.061237	0.8347	
	MA(1)	-0.105692	0.8233	
	MA(2)	0.222783	0.3887	
ARIMA (1,1,0)	AR(1)	0.107595	0.0971	Tidak
ARIMA (1,1,1)	AR(1)	0.63339	0.0008306	Ya
	MA(1)	-0.51001	0.0138721	
ARIMA (2,1,1)	AR(1)	0.150528	0.60067	Tidak
	AR(2)	0.151192	0.03672	
	MA(1)	-0.061706	0.82979	
ARIMA (3,1,1)	AR(1)	-0.771282	$2.862 \times 10^{-13}$	Ya
	AR(2)	0.225349	0.0049454	
	AR(3)	0.222413	0.0005866	
	MA(1)	0.873831	$2.2 \times 10^{-16}$	
ARIMA (3,1,2)	AR(1)	-1.143860	$2.2 \times 10^{-16}$	Ya
	AR(2)	-0.566170	$3.421 \times 10^{-7}$	
	AR(3)	0.231352	0.000687	

	MA(1)	1.281005	$2.2 \times 10^{-16}$	
	MA(2)	0.898268	$2.2 \times 10^{-16}$	

Berdasarkan Tabel 6, didapatkan model yang signifikan, yaitu ARIMA(1,1,1), ARIMA(3,1,2), dan ARIMA(3,1,1). Hasil estimasi parameter dari model ARIMA MEDC disajikan pada Tabel 7.

**Tabel 7.** Hasil estimasi parameter model ARIMA MEDC

Model	Parameter	Coefficient	p-value	Signifikan
ARIMA (5,1,0)	AR(1)	-0.00026448	0.996687	Tidak
	AR(2)	0.03843147	0.547416	
	AR(3)	0.05247552	0.409665	
	AR(4)	-0.06355707	0.318549	
	AR(5)	0.19845796	0.003566	
ARIMA (5,1,5)	AR(1)	-0.603619	0.996687	Tidak
	AR(2)	-0.108925	0.547416	
	AR(3)	-0.558597	0.409665	
	AR(4)	0.010860	0.318549	
	AR(5)	0.701730	0.003566	
	MA(1)	0.737452	0.996687	
	MA(2)	0.144162	0.547416	
	MA(3)	0.660708	0.409665	
	MA(4)	0.047215	0.318549	
	MA(5)	-0.671001	0.003566	
ARIMA (1,1,0)	AR(1)	-0.0093016	0.8861	Tidak
ARIMA (1,1,1)	AR(1)	-0.025690	0.02619	Ya
	MA(1)	-0.981238	0.03838	
ARIMA (0,1,1)	MA(1)	-0.0084149	0.8916	Tidak
ARIMA (2,1,2)	AR(1)	1.163614	$2.2 \times 10^{-16}$	Ya
	AR(2)	-0.970528	$2.2 \times 10^{-16}$	
	MA(1)	-1.173293	$2.2 \times 10^{-16}$	
	MA(2)	0.999935	$2.2 \times 10^{-16}$	
ARIMA (4,1,4)	AR(1)	-0.639064	$2.2 \times 10^{-16}$	Ya
	AR(2)	0.188087	0.002193	
	AR(3)	-0.641934	$2.2 \times 10^{-16}$	
	AR(4)	-0.936179	$2.2 \times 10^{-16}$	
	MA(1)	0.711261	$2.2 \times 10^{-16}$	
	MA(2)	-0.221940	$3.717 \times 10^{-8}$	
	MA(3)	0.712097	$2.2 \times 10^{-16}$	
	MA(4)	0.999032	$2.2 \times 10^{-16}$	

Berdasarkan Tabel 7, didapatkan model yang signifikan, yaitu ARIMA(1,1,1), ARIMA(2, 1, 2), dan ARIMA(4,1,4). Hasil estimasi parameter dari model ARIMA RUIS disajikan pada Tabel 8.

**Tabel 8.** Hasil estimasi parameter model ARIMA RUIS

Model	Parameter	Coefficient	p-value	Signifikan
ARIMA (1,1,0)	AR(1)	0.152241	0.01889	Ya
ARIMA (1,1,1)	AR(1)	-0.0032926	0.9901	Tidak
	MA(1)	0.1627043	0.5275	
ARIMA (1,1,3)	AR(1)	0.0260131	0.36411	Tidak
	MA(1)	-0.113621	0.68596	
	MA(2)	-0.050418	0.51139	
ARIMA (3,1,0)	MA(3)	-0.134409	0.03897	Tidak
	AR(1)	0.152400	0.01864	
	AR(2)	-0.048526	0.45959	
	AR(3)	-0.147198	0.02341	
ARIMA (3,1,3)	AR(1)	0.402169	$2.2 \times 10^{-16}$	Ya
	AR(2)	0.422054	$2.2 \times 10^{-16}$	
	AR(3)	-0.979148	$2.2 \times 10^{-16}$	
	MA(1)	-0.414517	$2.2 \times 10^{-16}$	
	MA(2)	-0.469010	$2.2 \times 10^{-16}$	

	MA(3)	0.941941	$2.2 \times 10^{-16}$	
ARIMA (2,1,1)	AR(1)	0.684559	0.001669	Ya
	AR(2)	-0.186383	0.004937	
	MA(1)	-0.525370	0.014103	
ARIMA (2,1,2)	AR(1)	1.401894	$2.2 \times 10^{-16}$	Ya
	AR(2)	-0.979374	$2.2 \times 10^{-16}$	
	MA(1)	-1.413665	$2.2 \times 10^{-16}$	
ARIMA (0,1,1)	MA(2)	0.943050	$2.2 \times 10^{-16}$	Ya
	MA(1)	0.159617	0.01183	
ARIMA (3,1,2)	AR(1)	1.544541	$2.2 \times 10^{-16}$	Ya
	AR(2)	-1.167195	$2.2 \times 10^{-16}$	
	AR(3)	0.136647	0.04689	
	MA(1)	-1.448667	$2.2 \times 10^{-16}$	
	MA(2)	0.956763	$2.2 \times 10^{-16}$	

Untuk mengetahui model dapat dilihat dari nilai MAPE. Nilai MAPE untuk masing-masing model ARIMA 4 perusahaan disajikan pada Tabel 9.

**Tabel 9.** Nilai MAPE model ARIMA

	Model	MAPE
APEX	ARIMA(1,1,0)	2.80301
	ARIMA(0,1,1)	2.764987
	ARIMA(0,1,2)	2.789863
	ARIMA(2,1,1)	2.805497
ELSA	ARIMA(1,1,1)	2.058361
	ARIMA(3,1,1)	1.98819
	ARIMA(3,1,2)	1.97583
MEDC	ARIMA(1,1,1)	3.193272
	ARIMA(2, 1, 2)	3.193878
	ARIMA(4,1,4)	2.957221
RUIS	ARIMA(3, 1, 3)	1.90784
	ARIMA(3, 1, 2)	1.89665
	ARIMA(2, 1, 1)	1.91154
	ARIMA(2, 1, 2)	1.906668
	ARIMA(1, 1, 0)	1.916048
	ARIMA(0, 1, 1)	1.921408

Berdasarkan Tabel 9, diperoleh model ARIMA terbaik untuk APEX, ELSA, MEDC, dan RUIS masing-masing, yaitu ARIMA(0,1,1), ARIMA(3,1,2), ARIMA(4,1,4), dan ARIMA(3, 1, 2).

D. Uji Heteroskedastisitas

Pada uji heteroskedastisitas ini digunakan uji Lagrange Multiplier atau sering disebut uji ARCH-LM dengan hipotesis  $H_0$  tidak terdapat heteroskedastisitas pada model. Hasil uji ARCH-LM pada model ARIMA terbaik disajikan pada Tabel 10.

**Tabel 10.** Hasil Uji ARCH-LM

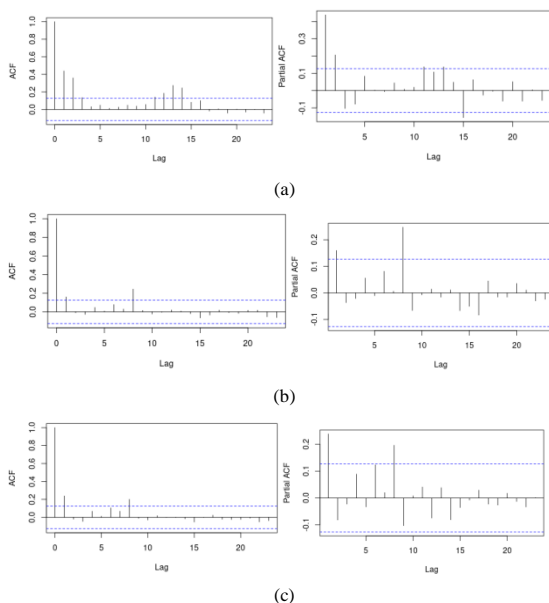
Data	Model	p - value	Kesimpulan
APEX	ARIMA (0,1,1)	$1.614 \times 10^{-11}$	Heteroskedastisitas
ELSA	ARIMA	$1.718 \times 10^{-5}$	Heteroskedastisitas

	(3,1,2)		
MEDC	ARIMA (4,1,4)	0.9466	Homoskedastisitas
RUIS	ARIMA (3,1,2)	0.002153	Heteroskedastisitas

Berdasarkan Tabel 10, diperoleh kesimpulan bahwa terdapat gejala heteroskedastisitas pada model ARIMA terbaik untuk data perusahaan APEX, ELSA, dan RUIS. Sedangkan untuk data perusahaan MEDC tidak terdapat efek heteroskedastisitas atau dapat dikatakan terdapat gejala homoskedastisitas pada model ARIMA yang artinya pada MEDC akan digunakan model ARIMA untuk *forecast*. Selanjutnya akan dilakukan pembentukan model GARCH pada model ARIMA APEX, ELSA, dan RUIS karena terdapat efek ARCH atau dengan kata lain model mengandung unsur heteroskedastisitas.

E. Estimasi Model ARIMA-GARCH

Pada estimasi model GARCH(*p, d, q*) dilakukan dengan metode *maximum likelihood* dengan orde *p* dan *q* dapat dilihat pada *correlogram* ACF dan PACF dari residual kuadrat model ARIMA terbaik. Orde *p* merupakan orde AR yang dapat dilihat pada *correlogram* PACF, orde *q* merupakan orde MA dapat dilihat pada *correlogram* ACF. Berikut *correlogram* ACF dan PACF dari residual kuadrat model ARIMA terbaik:



Gambar 3. Correlogram ACF dan PACF residual kuadrat model ARIMA (a) APEX, (b) ELSA, (c) MEDC, (d) RUIS

Pada ACF residual kuadrat APEX, terpotong pada lag 0, 1, dan 2. Untuk PACF residual kuadrat APEX terpotong pada lag 1 dan 2. Pada ACF residual kuadrat ELSA, terpotong pada lag 0 dan 1. Untuk PACF residual kuadrat ELSA terpotong pada lag 1. Pada ACF residual kuadrat RUIS, terpotong pada lag 0 dan 1. Untuk PACF residual kuadrat RUIS terpotong pada lag 1. Setelah didapatkan orde ARIMA-GARCH dilakukan estimasi parameter untuk uji signifikan model pendugaan dan model *overfitting* dengan hipotesis  $H_0$  model tidak signifikan. Hasil estimasi parameter dari model ARIMA-GARCH APEX disajikan pada Tabel 11.

Tabel 11. Hasil estimasi model ARIMA-GARCH APEX

Model	Parameter	Coefficient	p-value	Signifikan
ARIMA(0,1,1) GARCH(1,0)	$\mu$	24.32242	$2 \times 10^{-16}$	Ya
	MA(1)	0.69605	$2 \times 10^{-16}$	
	$\omega$	0.99651	$1.39 \times 10^{-10}$	
	$\alpha_1$	0.82816	$1.80 \times 10^{-13}$	
ARIMA(0,1,1) GARCH(1,1)	$\mu$	24.27698	$2 \times 10^{-16}$	Ya
	MA(1)	0.73146	$2 \times 10^{-16}$	
	$\omega$	0.22404	$2 \times 10^{-16}$	
	$\alpha_1$	0.49002	$1.60 \times 6$	
	$\beta_1$	0.48433	$1.39 \times 10^{-10}$	

Berdasarkan Tabel 11, didapatkan model yang signifikan, yaitu ARIMA(0,1,1)GARCH(1,0) dan ARIMA(0,1,1)GARCH(1,1). Hasil estimasi parameter dari model ARIMA-GARCH ELSA disajikan pada Tabel 12.

Tabel 12. Hasil estimasi model ARIMA-GARCH ELSA

Model	Parameter	Coefficient	p-value	Signifikan
ARIMA(3,1,2) GARCH(1,0)	$\mu$	-0.6369	0.371882	Tidak
	AR(1)	0.2655	0.124965	
	AR(2)	0.2855	0.082169	
	AR(3)	0.4605	0.000461	
	MA(1)	0.8335	$2.65 \times 10^{-9}$	
	MA(2)	0.5805	$2.07 \times 10^{-6}$	
	$\omega$	0.9966	$2.91 \times 10^{-9}$	
ARIMA(3,1,2) GARCH(1,1)	$\alpha_1$	0.9860	$1.94 \times 10^{-7}$	Ya
	$\mu$	0.05755	0.957140	
	AR(1)	0.01714	0.966657	
	AR(2)	0.51161	0.183560	
	AR(3)	0.47402	0.001823	
	MA(1)	1.00000	0.003236	
	MA(2)	0.52277	0.003327	
	$\omega$	0.46162	$2.1 \times 10^{-11}$	
	$\alpha_1$	0.71352	$2.81 \times 10^{-5}$	
	$\beta_1$	0.30903	0.000619	
ARIMA(0,1,1) GARCH(1,0)	$\mu$	59.5781	$2 \times 10^{-16}$	Ya
	MA(1)	0.7366	$2 \times 10^{-16}$	
	$\omega$	3.1623	$4.86 \times 10^{-9}$	
	$\alpha_1$	0.8675	$7.77 \times 10^{-15}$	
ARIMA(0,1,1) GARCH(1,1)	$\mu$	59.03431	$2 \times 10^{-16}$	Ya
	MA(1)	0.79956	$2 \times 10^{-16}$	
	$\omega$	0.38644	0.063	
	$\alpha_1$	0.51335	$2.83 \times 10^{-6}$	
	$\beta_1$	0.50757	$2.02 \times 10^{-10}$	

Berdasarkan Tabel 12, didapatkan model yang signifikan, yaitu ARIMA(0,1,1)GARCH(1,0), ARIMA(0,1,1)GARCH(1,1), dan ARIMA(1,1,0)GARCH(1,0). Hasil estimasi parameter dari model ARIMA-GARCH RUIS disajikan pada Tabel 13.

**Tabel 13.** Hasil estimasi model ARIMA-GARCH RUIS

Model	Parameter	Coefficient	p-value	Signifikan
ARIMA(3,1,2) GARCH(1,0)	$\mu$	0.4150	$2 \times 10^{-16}$	Tidak
	AR(1)	0.2990	$2 \times 10^{-16}$	
	AR(2)	0.9518	$2 \times 10^{-16}$	
	AR(3)	-0.2632	$2 \times 10^{-16}$	
	MA(1)	0.6626	$2 \times 10^{-16}$	
	MA(2)	-0.4235	$2 \times 10^{-16}$	
	$\omega$	0.5561	$9.11 \times 10^{-12}$	
ARIMA(0,1,1) GARCH(1,0)	$\alpha_1$	0.4372	0.1046	Ya
	$\mu$	30.60226	$2 \times 10^{-16}$	
	MA(1)	0.71526	$2 \times 10^{-16}$	
	$\omega$	0.88521	$1.44 \times 10^{-5}$	
ARIMA(0,1,1) GARCH(1,1)	$\alpha_1$	0.96839	$8.88 \times 10^{-16}$	Ya
	$\mu$	36.47111	$2 \times 10^{-16}$	
	MA(1)	0.85021	$2 \times 10^{-16}$	
	$\omega$	0.17574	0.080717	
	$\beta_1$	0.31886	0.000129	
		0.66445	$2 \times 10^{-16}$	

Berdasarkan Tabel 13, didapatkan model yang signifikan, yaitu ARIMA(0,1,1)GARCH(1,0) dan ARIMA(0,1,1)GARCH(1,1). Untuk mengetahui model terbaik dapat dilihat dari nilai AIC. Nilai AIC untuk masing-masing model ARIMA-GARCH 4 perusahaan disajikan pada Tabel 14.

**Tabel 14.** Nilai AIC model ARIMA-GARCH

	Model	AIC
APEX	ARIMA(1,1,0) GARCH(1,0)	4.612655
	ARIMA(0,1,1) GARCH(1,1)	2.543815
ELSA	ARIMA(1,1,0) GARCH(1,0)	5.778898
	ARIMA(0,1,1) GARCH(1,1)	5.65927
RUIS	ARIMA(1,1,0) GARCH(1,0)	4.873360
	ARIMA(0,1,1) GARCH(1,1)	4.849532

Berdasarkan Tabel 14, diperoleh model ARIMA-GARCH terbaik untuk APEX, ELSA, dan RUIS, yaitu ARIMA(0,1,1)GARCH(1,1).

**F. Uji Diagnostik**

Uji diagnostik dilakukan untuk menganalisis uji asumsi residual dari model yang didapatkan. Pada uji diagnostik ini, model akan di uji asumsi

residualnya yaitu uji autokorelasi, uji normalitas, dan uji homoskedastisitas. Pada uji diagnostik, digunakan *Jarque-Bera Test* untuk uji normalitas dengan hipotesis  $H_0$  residual data berdistribusi normal, *Ljung-Box Test* untuk uji autokorelasi dengan hipotesis  $H_0$  tidak terdapat autokorelasi pada residual data, dan *ARCH-LM Test* untuk uji Heteroskedastisitas dengan hipotesis  $H_0$  tidak terdapat gejala heteroskedastisitas pada residual data. Uji diagnostik model ARIMA-GARCH untuk perusahaan APEX, ELSA, dan RUIS serta model ARIMA untuk perusahaan MEDC disajikan pada Tabel 15.

**Tabel 15.** Uji diagnostik model ARIMA dan ARIMA-GARCH

Data	Model	p-value		
		Uji Normalitas	Uji Autokorelasi	Uji Heteroskedastisitas
APEX	ARIMA(0,1,1) GARCH(1,1)	0.3025	0	0.3735
ELSA	ARIMA(0,1,1) GARCH(1,1)	$3.898 \times 10^{-3}$	$8.182 \times 10^{12}$	$3.583 \times 10^{-8}$
MEDC	ARIMA(4,1,4)	$2.2 \times 10^{-16}$	0.9961	0.9466
RUIS	ARIMA(0,1,1) GARCH(1,0)	0.0029	0	0.00032

Berdasarkan Tabel 15, uji normalitas pada model 4 perusahaan dapat disimpulkan bahwa hanya model ARIMA-GARCH APEX yang memiliki residual data berdistribusi normal. Untuk uji autokorelasi pada model 4 perusahaan disimpulkan bahwa hanya pada model ARIMA MEDC yang tidak terdapat autokorelasi pada data. Untuk uji heteroskedastisitas pada model 4 perusahaan dapat disimpulkan bahwa hanya model ARIMA MEDC dan ARIMA-GARCH APEX tidak terdapat heteroskedastisitas pada residual data.

**G. Peramalan**

Setelah didapatkan model ARIMA dan ARIMA-GARCH pada 4 perusahaan, selanjutnya dilakukan *forecasting* atau peramalan dengan model yang telah didapatkan yang disajikan pada Tabel 16.

**Tabel 16.** Model ARIMA dan ARIMA-GARCH

Perusahaan	Model
APEX	ARIMA(0,1,1)GARCH(1,1)
ELSA	ARIMA(0,1,1)GARCH(1,1)

MEDC	ARIMA(4,1,4)
RUIS	ARIMA(0,1,1)GARCH(1,1)

Berdasarkan Tabel 16, didapatkan persamaan matematis untuk model ARIMA dan ARIMA-GARCH sebagai berikut:

1. Model ARIMA(0,1,1)GARCH(1,1) APEX  
 Pada model ARIMA-GARCH APEX didapatkan persamaan matematis sebagai berikut:

$$Y_t = 24.27968 + Y_{t-1} + 0.73146e_{t-1} + e_t$$

dan

$$\sigma_t^2 = 0.99651 + 0.49002e_{t-1}^2 + 0.48433\sigma_{t-1}^2 + e_t$$

2. Model ARIMA(0,1,1)GARCH(1,1) ELSA  
 Pada model ARIMA-GARCH ELSA didapatkan persamaan matematis sebagai berikut:

$$Y_t = 59.03431 + Y_{t-1} + 0.79956e_{t-1} + e_t$$

dan

$$\sigma_t^2 = 3.8644 + 0.51335e_{t-1}^2 + 0.50757\sigma_{t-1}^2 + e_t$$

3. Model ARIMA(4,1,4) MEDC  
 Pada model ARIMA-GARCH ELSA didapatkan persamaan matematis sebagai berikut:

$$Y_t = Y_{t-1} - 0.639064Y_{t-1} + 0.639064Y_{t-2} + 0.188087Y_{t-2} - 0.188087Y_{t-3} - 0.641934Y_{t-3} + 0.641934Y_{t-4} - 0.936179Y_{t-4} + 0.936179Y_{t-5} + 0.711261e_{t-1} - 0.221940e_{t-2} + 0.712097e_{t-3} + 0.999032e_{t-4} + e_t$$

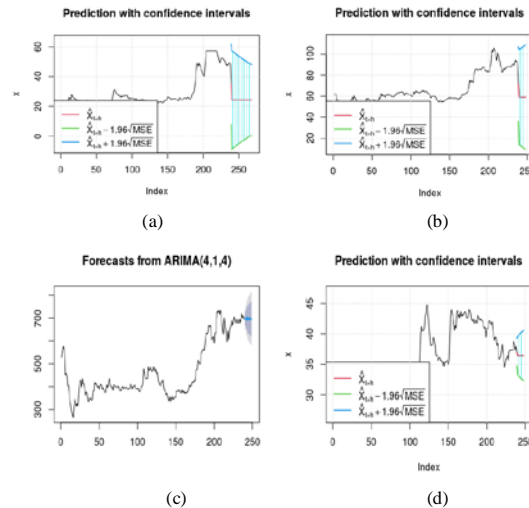
4. Model ARIMA(0,1,1)GARCH(1,0) RUIS  
 Pada model ARIMA-GARCH APEX didapatkan persamaan matematis sebagai berikut:

$$Y_t = 36.47111 + Y_{t-1} + 0.850216e_{t-1} + e_t$$

dan

$$\sigma_t^2 = 0.17574 + 0.31886e_{t-1}^2 + 0.66445\sigma_{t-1}^2 + e_t$$

Setah didapatkan model untuk 4 perusahaan, maka dilakuakn *forecasting*, yaitu meramalkan 10 data harga saham penutupan untuk masing-maisng 4 perusahaan untuk mengukur nilai akurasi dengan membandingkan data hasil peramalan dan data riil, di mana akan diramalkan harga saham penutupan dari tanggal 1 Maret 2021. Hasil peramalan 4 perusahaan dapat dilihat pada plot sebagai berikut:



Gambar 3. Plot hasil peramalan (a) APEX, (b) ELSA, (c) MEDC, (d) RUIS

Berdasarkan Gambar 3, hasil peramalan masih dalam bentuk transformasi. Oleh karena itu diperlukan pengembalian data transformasi ke data asli dengan menggunakan Persamaan (3). Hasil peramalan setelah pengembalian data transformasi ke data asli disajikan pada Tabel 17.

Tabel 17. Hasil peramalan ARIMA dan ARIMA-GARCH

No	Forecasting			
	APEX	ELSA	MEDC	RUIS
1	285.0054	285.4715	695.5060	245.5738
2	149.9544	218.0348	701.0539	239.6236
3	149.9544	218.0348	691.6119	239.6236
4	149.9544	218.0348	703.3553	239.6236
5	149.9544	218.0348	690.0395	239.6236
6	149.9544	218.0348	701.6252	239.6236
7	149.9544	218.0348	693.0176	239.6236
8	149.9544	218.0348	698.2515	239.6236
9	149.9544	218.0348	698.3164	239.6236
10	149.9544	218.0348	693.9386	239.6236

Berdasarkan Tabel 17, akan dibandingkan dengan data *out sample* dari harga saham penutupan masing-masing 4 perusahaan dari tanggal 1 Maret 2021. Perbandingan ini bertujuan untuk mengukur kesalahan dalam model ARIMA dan ARIMA-GARCH berdasarkan nilai *Mean Absolute Percentage Error* (MAPE) dengan rumus sebagai berikut:

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

Berdasarkan rumus MAPE didapatkan hasil akurasi peramalan untuk masing-masing 4 perusahaan disajikan pada Tabel 18.

**Tabel 18.** Hasil akurasi peramalan

No	Perusahaan	MAPE
1	APEX	67.32667%
2	ELSA	42.49374%
3	MEDC	5.269889%
4	RUIS	4.113383%

Berdasarkan Tabel 18, untuk model ARIMA-GARCH APEX dan ELSA didapatkan nilai MAPE di atas 10%. Maka untuk model ARIMA-GARCH belum dikatakan baik digunakan untuk peramalan harga saham perusahaan APEX dan ELSA. Untuk model ARIMA pada MEDC dan ARIMA-GARCH RUIS didapatkan nilai MAPE dibawah 10%. Maka untuk model dapat dikatakan baik digunakan untuk peramalan harga saham perusahaan MEDC dan RUIS.

#### 4. KESIMPULAN DAN SARAN

##### A. Kesimpulan

Berdasarkan hasil dan pembahasan diperoleh kesimpulan sebagai berikut:

1. Model ARIMA dan ARIMA-GARCH untuk peramalan harga penutupan saham harian sebagai berikut:
  - APEX didapatkan model mean atau ARIMA dengan orde  $AR(p)$  adalah 0 dan  $MA(q)$  adalah 1 dengan satu kali *differencing* sedangkan model varian atau GARCH didapatkan orde  $ARCH(p)$  adalah 1 dan orde  $GARCH(q)$  adalah 1.
  - ELSA didapatkan model mean atau ARIMA dengan orde  $AR(p)$  adalah 0 dan  $MA(q)$  adalah 1 dengan satu kali *differencing* sedangkan model varian atau GARCH didapatkan orde  $ARCH(p)$  adalah 1 dan orde  $GARCH(q)$  adalah 1.
  - MEDC didapatkan model mean atau ARIMA dengan orde  $AR(p)$  adalah 4 dan  $MA(q)$  adalah 4.
  - RUIS didapatkan model mean atau ARIMA dengan orde  $AR(p)$  adalah 0 dan  $MA(q)$  adalah 1 dengan satu kali *differencing* sedangkan model varian atau GARCH didapatkan orde  $ARCH(p)$  adalah 1 dan orde  $GARCH(q)$  adalah 1.
2. Hasil *forecasting* atau peramalan untuk 10 hari kedepan yang dimulai pada tanggal 01 Maret 2021. Dari hasil peramalan didapatkan nilai akurasi yang diperoleh dengan membandingkan data hasil peramalan dan data *out sample* berdasarkan nilai MAPE yaitu untuk APEX diperoleh nilai akurasi 67.32667%, untuk ELSA diperoleh nilai

akurasi 41.82999%, untuk MEDC diperoleh nilai akurasi 5.269889%, dan untuk RUIS diperoleh nilai akurasi 4.113383%. Dari nilai MAPE yang diperoleh, hasil akurasi untuk model peramalan ARIMA-GARCH untuk APEX dan ELSA masih di atas 10% yang di artikan model belum baik digunakan untuk peramalan. Sedangkan untuk nilai MAPE yang diperoleh untuk model peramalan ARIMA dan ARIMA-GARCH untuk MEDC dan RUIS sudah di bawah 10% yang di artiakan model sudah cukup baik digunakan untuk peramalan.

##### B. Saran

Saran yang dapat diberikan untuk penelitian selanjutnya yaitu membandingkan dengan metode lain seperti perbandingan dengan *Fuzzy Time Series Markov Chain*. Metode ARIMA memiliki keterbatasan yaitu hanya meramalkan untuk jangka pendek. Maka dari itu, untuk penelitian selanjutnya menggunakan metode peramalan lain yang dapat meramalkan untuk jangka waktu yang panjang seperti *Generalized Regression Neural Network*. Penelitian selanjutnya disarankan melanjutkan dengan menggunakan metode dari pengembangan GARCH yaitu seperti EGARCH, IGARCH, GARCH-M, dan TGARCH untuk mengatasi uji diagnostik yang belum terpenuhi dan menghasilkan nilai akurasi MAPE dibawah 10%.

#### DAFTAR PUSTAKA

- Angraeny, Nella. 2019. Penerapan Metode ARCH GARCH untuk Analisis Peramalan Nilai Ekspor Indonesia. Skripsi. FMIPA, Matematika, Universitas Negeri Semarang, Semarang.
- Ati, Enggar Niken Laras. 2015. Analisis Volatilitas Forecasting Sembilan Bahan Pokok menggunakan Metode GARCH dengan Program R. Skripsi. FMIPA, Matematika, Universitas Negeri Semarang, Semarang.
- Darmayati, Novi dkk. 2020. Dampak Vovid-19 Terhadap Perubahan Harga dan *Return* Saham. *Jurnal Ekonomi dan Keuangan*, Vol. 4 No. 4, Hal. 462-480
- Faustina, Riza Silvia dkk. 2017. Model *Hybrid* ARIMA-GARCH untuk Estimasi Volatilitas Harga Emas Menggunakan Software R. *UNNES Journal of Mathematics*. Vol. 6 No. 1, Hal. 11-24

- Farizah, Icak (2017), “Penerapan Model GARCH dalam Mengukur Risioko Berinvestasi”. Skripsi. Fakultas Sains dan Teknologi, Matematika, Universitas Islam Negeri Alaudin, Makassar.
- Hikmah, Nur Falilah Nurul Hikmah. 2018. Perbandingan Metode Arima-Garch Dan Fuzzy Time Series Markov Chain Dalam Peramalan Data Harga Minyak Mentah Dunia (Studi Kasus: Data Harga Minyak Mentah Dunia Tahun 2001-2017). Skripsi. FMIPA, Statistika, Universitas Islam Indonesia, Yogyakarta.
- Raneo, Agung Putra dan Fida Muthia 2018. Penerapan Model GARCH dalam Peramalan Volatilitas di Bursa Efek Indonesia. *Jurnal Manajemen dan Bisnis Sriwijaya*, Vol. 15 No. 3, Hal 194-202
- Wicaksono, Agung dan Rahandika Ivan Adyaksana. 2020. Analisis Reaksi Investor Sebagai Dampak Covid-19 pada Sektor Perbankan di Indonesia. *Jurnal Ilmiah Akuntansi Fakultas Ekonomi*. Vol. 6 No. 2, Hal. 129-138
- Widyastuti, Nur Laila dan Hanan Nugroho. 2020. Dampak Covid-19 terhadap Industri Minyak dan Gas Bumi: Rekomendasi Kebijakan untuk Indonesia. *The Indonesian Journal of Development Planning*. Vol. 4 No. 2, Hal. 166-176
- Yolanda, Natasya Bella dkk. 2017. Penerapan Model ARIMA-GARCH untuk Memprediksi Harga Saham Bank BRI. *Jurnal MIPA UNSRAT*. Vol. 6 No. 2, Hal. 92-96

## Analisis Kalman filter berbasis Google Trends untuk Prediksi Kedatangan Wisatawan Mancanegara di Indonesia Pasca Pandemi

Evita Purnaningrum<sup>(1)</sup>, Hanief Khooyir Nafah<sup>(2)</sup>

<sup>(1)</sup>Manajemen, Universitas PGRI Adi Buana Surabaya

<sup>(2)</sup>Akuntansi, Universitas PGRI Adi Buana Surabaya

Jl. Dukuh Menanggal XII No. 4 Surabaya 60234

e-mail: [purnaningrum@unipasby.ac.id](mailto:purnaningrum@unipasby.ac.id) dan [haniefknafah@gmail.com](mailto:haniefknafah@gmail.com)

### ABSTRAK

Pada tahun 2019 kunjungan wisatawan mancanegara (wisman) ke Indonesia mengalami peningkatan yang cukup signifikan. Sehingga, pariwisata diprediksi menjadi salah satu penopang terbesar dari penerimaan negara. Namun, saat wabah Coronavirus terjadi di akhir tahun 2019, sektor ini menjadi sektor industri yang paling terdampak dengan penurunan yang sangat tajam dan perkiraan akan membaik sekitar tahun 2035 hingga 2045. Kejadian tersebut mendorong penelitian untuk merumuskan model proyeksi terbaik bagi wisatawan asing pasca pandemi dengan menggunakan metode Kalman filter. Kalman filter merupakan model *state space* yang dapat diulang untuk menghasilkan nilai akurasi estimasi yang tinggi. Model ini didukung oleh analisis *google trends* yang mampu menangkap minat negara lain terhadap pariwisata Indonesia, terutama di masa pandemi. Hasil penelitian menunjukkan bahwa meskipun pandemi, beberapa negara masih memiliki minat terhadap objek wisata di Indonesia. Selain itu, Kalman filter memiliki akurasi yang tinggi dalam peramalan wisatawan asing.

**Kata kunci :** *peramalan, Kalman filter, Pariwisata, Pandemi, Google Analitik*

### ABSTRACT

*In 2019, foreign tourist visits to Indonesia experienced a significant increase. Thus, tourism is predicted to be one of the biggest pillars of state revenue. However, when the Coronavirus outbreak occurred at the end of 2019, this sector became the most affected industrial sector with a very sharp decline and is expected to improve around 2035 to 2045. This incident prompted research to formulate the best projection model for foreign tourists after the pandemic by using Kalman filter method. Kalman filter is a state space model that can be repeated to produce high estimation accuracy values. This model is supported by an analysis of Google Trends which is able to capture the interest of other countries in Indonesian tourism, especially during the pandemic. The results of the study show that despite the pandemic, several countries still have an interest in tourism objects in Indonesia. In addition, Kalman filter has high accuracy in forecasting foreign tourists.*

**Keywords :** *forecasting, Kalman filter, Tourism, Pandemic, Google Analytics.*

### 1. PENDAHULUAN

Pariwisata Indonesia salah satu sektor yang memiliki potensi untuk terus berkembang. Permintaan pariwisata pada 2019 mencapai 16,11 juta, naik 1,88% dari tahun sebelumnya (Badan Pusat Statistik, 2019). Meski demikian, pada 2020, pemerintah memutuskan untuk merevisi target kunjungan wisatawan yang sebelumnya mencapai 17,2 juta orang akibat pandemi Covid-19. Covid-19 adalah penyakit dari coronavirus jenis baru yang berasal dari paparan makanan laut di Wuhan, keberadaan virus ini pertama kali dilaporkan pada 31 Desember 2019 ke kantor negara WHO di China (World Health Organization (WHO), 2020).

Wabah Covid-19 telah mempengaruhi berbagai sektor baik formal maupun informal terutama sektor pariwisata. Hal tersebut menjadikan alasan pemerintah perlu mengambil keputusan dan langkah-langkah yang cukup besar untuk penanganan dan perbaikan dari kejadian yang ditimbulkan oleh pandemi. Salah satu dampak dari wabah tersebut adalah penutupan tempat-tempat wisata pada waktu yang tidak dapat diprediksi, termasuk penurunan permintaan pariwisata dan penghentian sementara industri pariwisata (Newsome, 2020), penurunan kinerja karyawan di sektor pariwisata, dan penurunan jumlah penumpang angkutan. Dengan kata lain, Covid-19

menyebabkan kedatangan wisman ke Indonesia turun tajam. Apalagi diperkirakan jumlah ini akan kembali normal sesuai target pada 2024 atau 2025. Namun, belum ada penelitian lebih lanjut mengenai perkiraan waktu aktivitas pariwisata bisa kembali normal. Oleh karena itu, penelitian ini bertujuan untuk memperkirakan kedatangan turis asing menggunakan analisis *Google Trends*. *Google Trends* merupakan salah satu big data yang dapat diterapkan untuk menyelesaikan masalah dengan cara yang sederhana (Höpken, Eberle, Fuchs, & Lexhagen, 2019; Kusumawardhani & Purnaningrum, 2021; Nafah & Purnaningrum, 2021; Evita Purnaningrum & Ariqoh, 2019; Evita Purnaningrum & Ariyanti, 2020; Sun, Wei, Tsui, & Wang, 2019).

Tingginya permintaan pariwisata terutama dari wisatawan asing dapat berdampak positif terhadap PDB. Berdasarkan literatur, nilai tukar (Cachia, Compañó, & Da Costa, 2007; De Vita, 2014), PDB per kapita dunia (Martins, Gan, & Ferreira-Lopes, 2017; Wakimin, Azlina, & Hazman, 2018) *Google Trends Index* (Yang, Pan, Evans, & Lv, 2015), dan Indeks Harga Konsumen (CPI) (Alamsyah & Friscintia, 2019) dapat mempengaruhi permintaan pariwisata. Oleh karena itu, faktor-faktor tersebut dapat membantu pemerintah dalam menetapkan kebijakan yang tepat untuk memikat minat wisatawan. Selain itu, Beberapa studi literatur sebelumnya telah membahas tentang pemesanan peramalan pariwisata khususnya di Indonesia seperti model Vector Error Correction Model (VECM) (Haqiq & Pharmasetiawan, 2019), Artificial Neural Network (ANN) (Alamsyah & Friscintia, 2019), ANN dan Holt winters (Supriatna, Hertini, Saputra, Subartini, & Robbani, 2019), variabel ekonomi (Mariyono, 2017), Model ARIMA (Kharista, Permanasari, & Hidayah, 2015). Namun, metode sebelumnya membutuhkan pengembangan untuk meningkatkan kinerja peramalan. Dalam kondisi yang tidak pasti seperti pandemi, diperlukan metode yang ampuh untuk menangkap kejadian tersebut. Penelitian ini menerapkan metode Kalman filter untuk memprediksi kondisi ketidakpastian yang telah digunakan peneliti sebelumnya dalam bidang ekonometrika (Bradford & Imstrand, 2018; Fauziyah & Purnaningrum, 2021; E Purnaningrum, 2018; E Purnaningrum, Cahyaningtias, & Kusumawardhani, 2021; Evita Purnaningrum, 2020; Wilcox & Hamano, 2017). Selain itu, Kalman Filter (KF) berbasis teknik model time series univariat jarang diterapkan dalam studi ekonomi empiris, khususnya pariwisata, dan belum pernah dilakukan studi untuk kasus Indonesia. Meskipun Rizal (2017)

telah menerapkan filter RNN-Extended Kalman untuk memprediksi kedatangan wisatawan di Lombok dengan akurasi 86%, namun dalam penelitian ini Kalman filter digunakan sebagai algoritma pelatihan untuk mendapatkan bobot pada RNN. Berbeda dengan hal tersebut, penelitian ini menerapkan metode Kalman filter sebagai model dan metode peramalan pariwisata di Indonesia. Oleh karena itu, penelitian ini ingin menambah literatur terbaru untuk pemahaman lebih lanjut tentang topik Kalman filter.

## 2. METODE PENELITIAN

### 2.1 Kalman filter untuk Kedatangan Wisatawan

Kalman filter diperkenalkan sebagai metode peramalan dengan model *state space* ekonometrik oleh Harvey pada tahun 1990 (Harvey, 1990). Selanjutnya model tersebut dimodifikasi dan diterapkan pada berbagai permasalahan ekonomi. Pada penelitian ini akan menggunakan model Kalman filter yang telah dibentuk oleh Durbin dan Koopman (2012) dengan penambahan variabel penjelas dari *google trends*. Persamaannya sebagai berikut:

$$\mathbf{y}_t = \mathbf{Z}_t \boldsymbol{\alpha}_t + \mathbf{X}_t \boldsymbol{\beta}_t + \boldsymbol{\varepsilon}_t, \quad \boldsymbol{\varepsilon}_t \sim N(\mathbf{0}, \mathbf{H}_t) \quad (1)$$

dimana  $\mathbf{y}$  adalah data kedatangan wisatawan, bulanan,  $\boldsymbol{\alpha}$  adalah vektor keadaan yang tidak teramati,  $\boldsymbol{\beta}$  adalah koefisien regresi.  $\mathbf{X}$  merupakan variabel penjelas atau merepresentasikan data *google trends*.  $\boldsymbol{\varepsilon}$  merupakan error dari model yang berdistribusi normal dengan  $\mu = 0$ .

Keuntungan dari model *state space* Kalman filter adalah operasi rekursif dari perhitungan untuk mendapatkan model berbasis waktu terbaik dan sebelumnya. Dalam hal peramalan secara real time menjadi sangat efektif dan efisien. Selanjutnya model filter Kalman tidak hanya memberikan prediksi yang akurat tetapi juga membentuk struktur model yang dapat dianalisis lebih lanjut untuk hubungan antar variabel.

### 2.2 Tahapan Kalman filter dalam software R

Model simulasi filter Kalman dilakukan di R open source dengan paket KFAS yang dimodifikasi untuk peramalan wisatawan (Helske, 2017). Langkah-langkah metodenya adalah sebagai berikut:

- Menemukan ketidakpastian terbaik  $Q$  (kesalahan kovarians dari keadaan tidak teramati) dengan syarat  $t < 10000, \text{abs}(RMSE_t - RMSE_{t-1}) > 10^{-7}, \text{dan } RMSE > 10^{-3}$ .

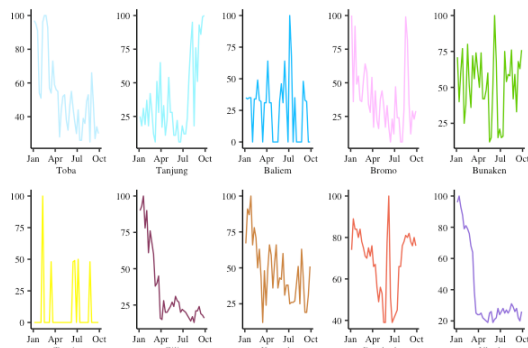
- Pembagian data dengan rasio 3:1 untuk data training dan testing.
- Model Terbaik untuk Kalman filter (Data training).
- Prediksi Masa Depan (Data testing).
- Pengukuran Akurasi (MAPE dan RMSE)

### 3. HASIL DAN PEMBAHASAN

Secara umum, hasil dan pembahasan memiliki dua topik. Topik pertama adalah hasil dari banyaknya negara yang masih tertarik dengan objek wisata di Indonesia selama masa pandemi. Hal ini penting untuk mempromosikan Indonesia kembali ke dunia internasional. Topik kedua membahas hasil metode Kalman filter dan mencari model terbaik berdasarkan MAPE dan RMSE.

#### 3.1. Google Trends Analitik

Pengumpulan data tren google menggunakan paket di R (gTrends) (Massicotte & Eddelbuettel, 2019) berdasarkan tiga kategori istilah pencarian, yaitu kata umum tentang coronavirus, kata lain terkait untuk coronavirus, dan sepuluh besar tempat wisata di Indonesia (<https://www.touropia.com/tourist-attractions-in-indonesia/>). Data google sebagai variabel penjelas dalam peramalan menggunakan metode Kalman filter. Data ini ditambang sejak awal deteksi virus corona, akhir Desember 2019 hingga September 2020, setiap minggu.

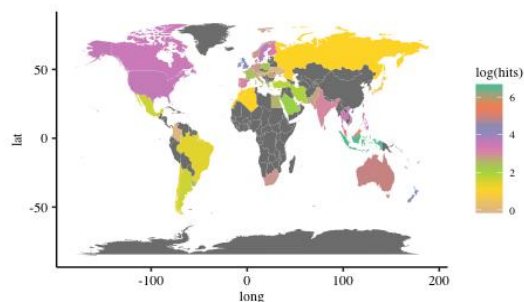


**Gambar 1.** Grafik Pencarian Obyek Wisata di Indonesia Saat Pandemi (*Google Trends*)

Berdasarkan google trend, kita bisa mengetahui negara mana saja yang tertarik dengan beberapa objek wisata di Indonesia. Kami telah membuat peta (ditunjukkan pada gambar 2) berdasarkan kata pencarian untuk objek wisata. Dengan demikian, hasil tersebut dapat bermanfaat bagi pemerintah atau pelaku usaha di bidang pariwisata untuk memetakan proyeksi target dan target promosi pariwisata ke depan. Lebih dari

separuh pencari informasi pariwisata Indonesia di kawasan ASEAN dan Australia masing-masing. Dengan kata lain, sekarang ini kita harus promosi yang luas tidak hanya menargetkan kawasan ASEAN tetapi ke negara-negara yang melihat pariwisata berdasarkan data google trend. Tak disangka, Inggris menjadi 5 negara teratas yang mencari objek wisata ini. Pada akhirnya, peta berdasarkan istilah pencarian berguna dalam pengambilan keputusan dalam hal perencanaan pariwisata.

Hasil dari model google trend analytic ini akan sangat membantu untuk mendukung para pekerja di sektor pariwisata dan pemerintah dalam menentukan kebijakan yang tepat dan perbaikan yang tepat sehingga dapat berdampak positif pada peningkatan jumlah permintaan pariwisata Indonesia. Selain itu, Penelitian ini dapat membantu industri pariwisata untuk mengetahui perkiraan negara wisatawan yang berlibur pada waktu tertentu sehingga dapat digunakan untuk memberikan diskon tiket pesawat dan mengadakan kegiatan budaya. Oleh karena itu, promosi tersebut membuat wisatawan asing datang kembali di masa depan. Selain untuk menambah pendapatan negara, kegiatan budaya yang dilakukan juga dapat dijadikan sebagai pengenalan budaya Indonesia kepada wisatawan.



**Gambar 2.** Peta pencarian objek pariwisata di Indonesia berdasarkan negara dengan *Google Trends*

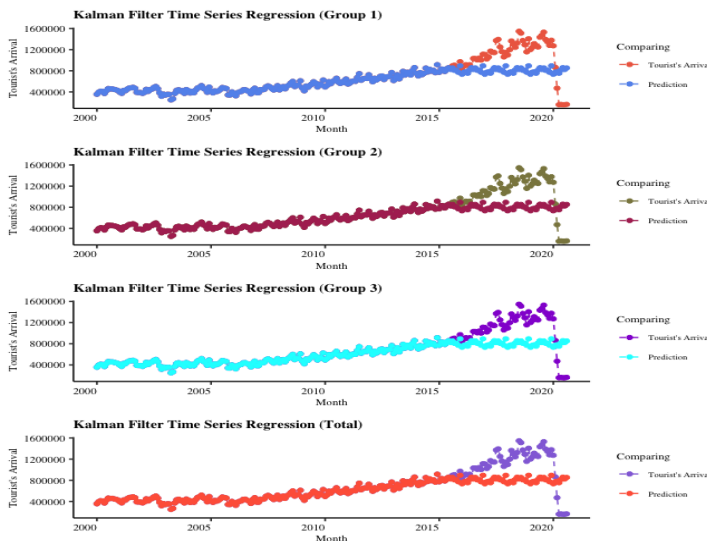
#### 3.2. Hasil Kalman filter

Setelah menyelesaikan analisis tren google, kami telah mensimulasikan filter Kalman untuk mendapatkan kesalahan kovarians yang memiliki RMSE terkecil. Studi ini membagi simulasi menjadi tiga kelompok dan membandingkan RMSE dan MAPE (ditunjukkan pada Tabel 1). Hasil tersebut menunjukkan bahwa Kalman filter mampu memprediksi kedatangan wisman ke Indonesia secara signifikan. Karena MAPE dari ketiga model yang disimulasikan di bawah 10%, maka model filter Kalman cocok untuk memprediksi data

ketidakpastian dan memproyeksikan masa depan. Tingkat akurasi setiap percobaan memiliki rata-rata yang sama, artinya model filter Kalman dengan variabel penjelas dari google trend stabil dalam memprediksi wisatawan mancanegara.

Tabel 1. Nilai RMSE dan MAPE

Error	Grup berdasarkan Kata kunci			Total
	Group 1	Group 2	Group 3	
RMSE	0.5965737	0.5965737	0.5965737	0.5965737
MAPE	0.07250522	0.07250522	0.07250522	0.07250522



Gambar 3. Hasil simulasi peramalan dengan menggunakan Kalman filter berbasis Google Trends

Perbandingan pergerakan kunjungan wisman dengan prediksi dapat dilihat pada gambar 3. Split ratio untuk data sebesar 75%. Seperti yang ditunjukkan pada Gambar 3, dibandingkan dengan kurva kunjungan wisatawan yang sebenarnya, dapat dilihat bahwa model yang diusulkan dalam makalah ini memiliki kinerja yang baik dalam memprediksi tren kunjungan wisatawan. Namun permasalahan outlier (seperti bencana alam, pandemi, wabah) perlu dikaji lebih lanjut dengan menyempurnakan model ruang keadaan Kalman filter, tidak hanya Time Series tetapi dengan penambahan model lainnya. Namun, model Kalman filter dapat mengikuti pergerakan fungsi yang sebenarnya dengan lebih sedikit kesalahan. Terakhir, penelitian ini menunjukkan bahwa model Kalman filter dapat menjadi metode peramalan alternatif untuk kasus ekonometrika, khususnya peramalan wisatawan asing. Akhirnya, berdasarkan MAPE (7%) kurang dari 10% yang berarti metode ini memiliki

peramalan yang sangat akurat (Montaño Moreno, Palmer Pol, Sesé Abad, & Cajal Blasco, 2013).

4. KESIMPULAN DAN SARAN

Pandemi adalah keadaan ketidakpastian yang tidak dapat diprediksi kapan akan berakhir. Di Indonesia, telah terjadi perlambatan ekonomi dan kemungkinan akan mengalami resesi. pascapandemi dan memasuki era new normal, perlu direncanakan strategi untuk mendukung peningkatan ekonomi. salah satu harapan terbesar adalah pariwisata. Dari penelitian ini didapatkan hasil bahwa kita dapat mengetahui potensi wisatawan berdasarkan sentiment masyarakat di google search dengan menggunakan google trends analitik. Di sisi lain, Kalman filter dapat memprediksi secara signifikan dan akurat kunjungan wisatawan asing ke Indonesia dengan perolehan nilai MAPE sebesar 7% dan RMSE 60%. Penambahan variabel baik yang merupakan faktor makroekonomi dan modifikasi dengan penggabungan model ekonometrik yang lain dapat digunakan untuk peningkatan performa prediksi kunjungan wisatawan mancanegara. Hasil ini dapat berguna dengan cara lebih lanjut untuk memprediksi kunjungan wisatawan asing.

DAFTAR PUSTAKA

Alamsyah, A., & Friscintia, P. B. A. (2019). Artificial neural network for Indonesian tourism demand forecasting. *2019 7th International Conference on Information and Communication Technology, ICoICT 2019*. <https://doi.org/10.1109/ICoICT.2019.8835382>

Badan Pusat Statistik. (2019). Jumlah kunjungan wisman ke Indonesia Desember 2019 mencapai 1,38 juta kunjungan. Retrieved from Perkembangan Pariwisata dan Transportasi Nasional Desember 2019 website: <https://www.bps.go.id/pressrelease/2020/02/03/1711/jumlah-kunjungan-wisman-ke-indonesia-desember-2019-mencapai-1-38-juta-kunjungan-.html>

Bradford, E., & Imsland, L. (2018). Economic Stochastic Model Predictive Control Using the Unscented Kalman Filter. *IFAC-PapersOnLine*, 51(18), 417-422. <https://doi.org/10.1016/j.ifacol.2018.09.336>

Cachia, R., Compañó, R., & Da Costa, O. (2007). Grasping the potential of online social networks for foresight. *Technological Forecasting and Social Change*. <https://doi.org/10.1016/j.techfore.2007.05.00>

6

- De Vita, G. (2014). The long-run impact of exchange rate regimes on international tourism flows. *Tourism Management*. <https://doi.org/10.1016/j.tourman.2014.05.001>
- Durbin, J., & Koopman, S. J. (2012). *Time Series Analysis by State Space Methods: Second Edition*. *Aging*. <https://doi.org/10.1017/CBO9781107415324.004>
- Fauziyah, & Purnaningrum, E. (2021). Optimization of Stock Portfolios Using Goal Programming Based on the Kalman-Filter Method. *Jurnal Matematika MANTIK*, 7(1), 20–30. <https://doi.org/https://doi.org/10.15642/mantik.2021.7.1.20-30>
- Haqiq, A., & Pharmasetiawan, B. (2019). Data Analytics for Forecasting Arrival of Tourism Visit in Indonesia. *Proceeding - 2019 International Conference on ICT for Smart Society: Innovation and Transformation Toward Smart Region, ICISS 2019*. <https://doi.org/10.1109/ICISS48059.2019.8969795>
- Harvey, A. C. (1990). Forecasting, Structural Time Series Models and the Kalman Filter. In *Forecasting, Structural Time Series Models and the Kalman Filter*. <https://doi.org/10.1017/cbo9781107049994>
- Helske, J. (2017). KFAS: Exponential family state space models in R. *Journal of Statistical Software*. <https://doi.org/10.18637/jss.v078.i10>
- Höpken, W., Eberle, T., Fuchs, M., & Lexhagen, M. (2019). Google Trends data for analysing tourists' online search behaviour and improving demand forecasting: the case of Åre, Sweden. *Information Technology and Tourism*. <https://doi.org/10.1007/s40558-018-0129-4>
- Kharista, A., Permanasari, A. E., & Hidayah, I. (2015). The performance of GM (1,1) and ARIMA for forecasting of foreign tourists visit to Indonesia. *2015 International Seminar on Intelligent Technology and Its Applications, ISITIA 2015 - Proceeding*. <https://doi.org/10.1109/ISITIA.2015.7219949>
- Kusumawardhani, D. A., & Purnaningrum, E. (2021). Penyebaran Pengguna Digital Wallet Di Indonesia Berdasarkan Google Trends Analytics. *INOVASI*, 17(2), 377–385. <https://doi.org/http://dx.doi.org/10.29264/jinv.v17i2.8069>
- Mariyono, J. (2017). Determinants of Demand for Foreign Tourism in Indonesia. *Jurnal Ekonomi Pembangunan: Kajian Masalah Ekonomi Dan Pembangunan*. <https://doi.org/10.23917/jep.v18i1.2042>
- Martins, L. F., Gan, Y., & Ferreira-Lopes, A. (2017). An empirical analysis of the influence of macroeconomic determinants on World tourism demand. *Tourism Management*. <https://doi.org/10.1016/j.tourman.2017.01.008>
- Massicotte, P., & Eddelbuettel, D. (2019). gtrendsR: Perform and Display Google Trends Queries. *R Package Version*.
- Montaño Moreno, J. J., Palmer Pol, A., Sesé Abad, A., & Cajal Blasco, B. (2013). Using the R-MAPE index as a resistant measure of forecast accuracy. *Psicothema*. <https://doi.org/10.7334/psicothema2013.23>
- Nafah, H. K., & Purnaningrum, E. (2021). Penggunaan Big Data Melalui Analisis Google Trends Untuk Mengetahui Perspektif Pariwisata Indonesia Di Mata Dunia. *Snhrp*.
- Newsome, D. (2020). The collapse of tourism and its impact on wildlife tourism destinations. *Journal of Tourism Futures*. <https://doi.org/10.1108/JTF-04-2020-0053>
- Purnaningrum, E. (2018). Renewable Stock Price Model Sebagai Pendukung Investasi Saham: Studi Kasus Saham Jii. *Kolegial*.
- Purnaningrum, E., Cahyaningtias, S., & Kusumawardhani, D. A. (2021). Augmentation time series model with Kalman filter to predict foreign tourist arrivals in East Java. *Journal of Physics: Conference Series*, 1869(1), 012116. <https://doi.org/10.1088/1742-6596/1869/1/012116>
- Purnaningrum, Evita. (2020). Pendekatan Metode Kalman Filter untuk Peramalan Pergerakan Indeks Harga Saham Terdampak Pandemi Coronavirus. *Majalah Ekonomi*.
- Purnaningrum, Evita, & Ariqoh, I. (2019). Google Trends Analytics dalam Bidang Pariwisata. *Majalah Ekonomi*.
- Purnaningrum, Evita, & Ariyanti, V. (2020). Pemanfaatan Google Trends Untuk Mengetahui Intervensi Pandemi Covid-19 Terhadap Pasar Saham Di Indonesia. *Jurnal.Unipasby.Ac.Id*.
- Rizal, A. A., & Hartati, S. (2017). Recurrent neural network with Extended Kalman Filter for

- prediction of the number of tourist arrival in Lombok. *2016 International Conference on Informatics and Computing, ICIC 2016*.  
<https://doi.org/10.1109/IAC.2016.7905712>
- Sun, S., Wei, Y., Tsui, K. L., & Wang, S. (2019). Forecasting tourist arrivals with machine learning and internet search index. *Tourism Management*.  
<https://doi.org/10.1016/j.tourman.2018.07.010>
- Supriatna, A., Hertini, E., Saputra, J., Subartini, B., & Robbani, A. A. (2019). The forecasting of foreign tourists arrival in indonesia based on the supply chain management: An application of artificial neural network and holt winters approaches. *International Journal of Supply Chain Management*.
- Wakimin, N. F., Azlina, A. A., & Hazman, S. (2018). Tourism demand in Asean-5 countries: Evidence from panel data analysis. *Management Science Letters*.  
<https://doi.org/10.5267/j.msl.2018.4.023>
- Wilcox, B. A., & Hamano, F. (2017). Kalman's Expanding Influence in the Econometrics Discipline. *IFAC-PapersOnLine*.  
<https://doi.org/10.1016/j.ifacol.2017.08.106>
- World Health Organization (WHO). (2020). Novel Coronavirus – China.
- Yang, X., Pan, B., Evans, J. A., & Lv, B. (2015). Forecasting Chinese tourist volume with search engine data. *Tourism Management*.  
<https://doi.org/10.1016/j.tourman.2014.07.019>

**JStatistika** : Jurnal Ilmiah Teori Dan Aplikasi Statistika ber ISSN online 2654-7511 dan ISSN cetak 2089-0028,

**JStatistika** merupakan media penerbitan karya ilmiah berupa literasi hasil penelitian ataupun kajian tentang teori, metodologi serta aplikasi dalam bidang ilmu statistika dan data sains. Jstatistika diterbitkan oleh Program Studi Statistika Fakultas Sains dan Teknologi Universitas PGRI Adi Buana Surabaya dan akan terbit dua kali dalam setahun setiap bulan juli dan desember.

**FAKULTAS  
SAINS DAN TEKNOLOGI**  
Universitas PGRI Adi Buana Surabaya

