

# Prediksi Muka Air Laut dari Sistem PUMMA Menggunakan SARIMA

Irfan Asfy Fakhry Anto<sup>1</sup>, Oka Mahendra<sup>2</sup>, Purnomo Husnul Khotimah<sup>3</sup>, Semeidi Husrin<sup>4</sup>

<sup>1,3</sup> Pusat Riset Sains Data dan Informasi Badan Riset dan Inovasi Nasional, KST Samaun Samadikun - Jl. Sangkuriang, Kota Bandung, 40135 INDONESIA (tel.: 022-2530050; email: [irfa013@brin.go.id](mailto:irfa013@brin.go.id), [purn005@brin.go.id](mailto:purn005@brin.go.id))

<sup>2</sup> Pusat Riset Mekatronika Cerdas Badan Riset dan Inovasi Nasional, KST Samaun Samadikun - Jl. Sangkuriang, Kota Bandung, 40135 INDONESIA (tel.: 022-2530050; email: [okam001@brin.go.id](mailto:okam001@brin.go.id))

<sup>4</sup> Pusat Riset Kebencanaan Geologi Badan Riset dan Inovasi Nasional, KST Samaun Samadikun - Jl. Sangkuriang, Kota Bandung, 40135 INDONESIA (tel.: 022-2530050; email: [seme001@brin.go.id](mailto:seme001@brin.go.id))

[Diterima: 30 Maret 2023, Revisi: 3 Juli 2023]

Corresponding Author: Irfan Asfy Fakhry Anto

**INTISARI** — Naiknya permukaan air laut berpotensi mengancam kehidupan di sepanjang pantai atau dataran rendah. Prediksi ketinggian permukaan air laut dapat mengurangi risiko ini. Prediksi dapat dilakukan dengan mengumpulkan informasi tentang kemungkinan kenaikan permukaan laut di masa mendatang. Kementerian Kelautan dan Perikanan Indonesia mengembangkan sebuah sistem, yaitu Perangkat Ukur Murah untuk Muka Air Laut (PUMMA), untuk mengukur ketinggian permukaan air laut. PUMMA terletak di stasiun pemantauan jarak jauh di sekitar wilayah laut Indonesia. Saat ini, sistem PUMMA masih belum memiliki fitur prediksi. Tujuan penelitian ini adalah untuk memodelkan prediksi muka air laut dengan menggunakan *dataset* selama satu tahun yang diambil dari bulan Juli 2021 hingga bulan Juli 2022. Metode *seasonal autoregressive integrated moving average* (SARIMA) digunakan karena metode ini terbukti merupakan metode yang fleksibel dan serbaguna untuk *dataset* yang tidak memiliki sifat kompleks yang memiliki pola musiman. Makalah ini mengembangkan beberapa model SARIMA. Kinerja model dievaluasi menggunakan metrik *mean absolute percentage error* (MAPE), *R-squared*, *mean square error* (MSE), dan *root mean square error* (RMSE). Model SARIMA(1, 1, 0)(1, 1, 1) 12 mencapai *error* prediksi terendah dengan *R-squared* 0,508, MSE 0,0479, dan RMSE 0,069. Berdasarkan kinerjanya, model SARIMA(1, 1, 0)(1, 1, 1) 12 dianggap layak untuk memprediksi muka air laut menggunakan *dataset* PUMMA.

**KATA KUNCI** — Permukaan Laut, Prediksi, PUMMA, Prediksi Permukaan Laut, Rata-Rata Pergerakan Musiman.

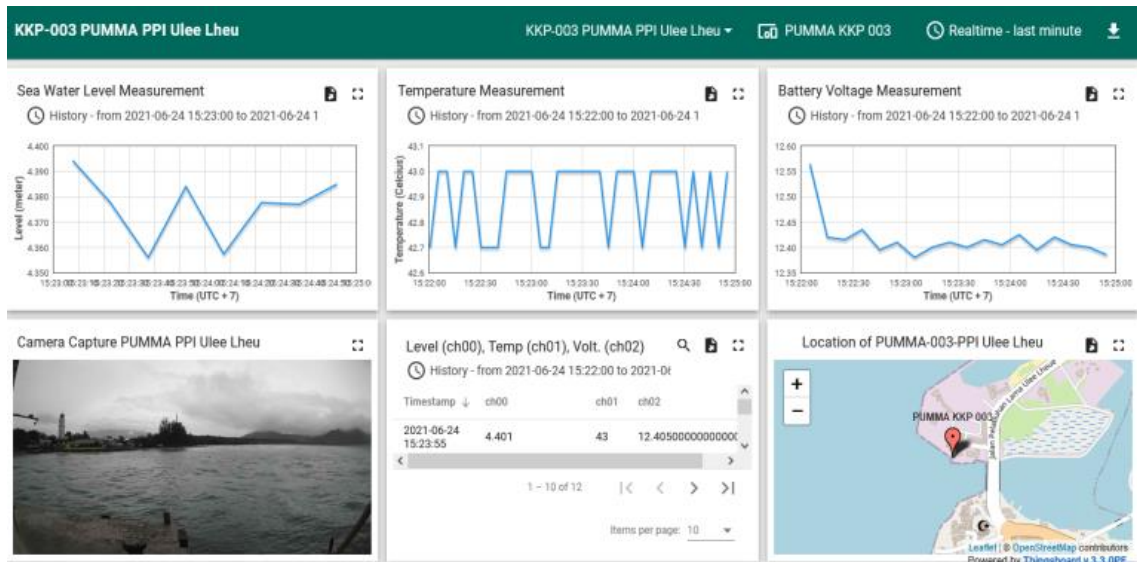
## I. PENDAHULUAN

Tsunami Aceh yang terjadi pada tahun 2004 telah menyadarkan pemerintah Republik Indonesia untuk membangun infrastruktur peringatan dini tsunami yang disebut Sistem Peringatan Dini Tsunami Indonesia atau dikenal dengan *Indonesian Tsunami Early Warning System* (InaTEWS) [1]. Sistem InaTEWS merupakan sistem peringatan dini tsunami yang masih mengandalkan sensor seismik untuk mendeteksi tsunami [2]. Salah satu data yang dibutuhkan oleh InaTEWS adalah data permukaan air laut [3]. Pemerintah Indonesia terus mengembangkan sistem prototipe untuk mengukur tinggi muka air laut, yaitu Perangkat Ukur Murah untuk Muka Air Laut (PUMMA), yang dikelola oleh Kementerian Kelautan dan Perikanan RI. PUMMA merupakan sistem pengukur tinggi muka air laut berbiaya rendah yang menyediakan beberapa sensor untuk merekam data tinggi muka air laut dan anomali yang terjadi. Sistem PUMMA disiapkan untuk mendukung dan terintegrasi dengan InaTEWS [4]. Masyarakat pesisir dan dataran rendah di dekat laut dapat terpengaruh oleh kenaikan permukaan laut, yang merupakan parameter geofisika fundamental yang relevan dengan banyak subbidang ilmu bumi, termasuk oseanografi, geodesi, dan klimatologi [5], [6]. Prediksi muka air laut bertujuan untuk mengumpulkan informasi tentang kemungkinan kenaikan muka air laut di masa mendatang untuk mengurangi risiko gangguan terhadap masyarakat pesisir dan infrastruktur, seperti banjir dan genangan [7].

Banyak penelitian telah dilakukan dari berbagai aspek prediksi atau peramalan muka air laut. Referensi [8] menggunakan dua metode prediksi, yaitu *autoregressive integrated moving average* (ARIMA) dan teknik pembelajaran mesin (*support vector regression*). Kedua metode tersebut digunakan untuk memperkirakan anomali permukaan laut

berdasarkan pengukuran ketinggian dari satelit di Samudera Pasifik, untuk mengembangkan dan memvalidasi pendekatan *sea level anomalies* (SLAs). Hasil yang diperoleh dibandingkan dengan nilai yang diamati, kemudian perkiraan permukaan laut dihitung menggunakan rutinitas model ARIMA. Penelitian lain menggunakan data tinggi permukaan air laut di Laut Cina Selatan dan perairan sekitarnya pada tahun 2017 untuk membangun model *seasonal ARIMA* (SARIMA) berdasarkan periodisitas rata-rata bulanan dan memprediksi tinggi muka air laut di wilayah tersebut [9].

Penelitian menggunakan metode berbasis pembelajaran mesin lebih banyak dilakukan. Referensi [10] menggunakan data muka air laut dari kegiatan pemantauan pantai di Turki. Kemudian kinerja prediksi muka laut dibandingkan dengan mempertimbangkan skenario yang berbeda menggunakan metode prediksi yang berbeda, seperti *multiple linear regression* dan *adaptive neuro-fuzzy inference system*, yang disajikan secara linear dan nonlinear yang dapat digunakan untuk operasi pelabuhan dalam memprediksi permukaan laut oleh para praktisi. Penelitian lain menggunakan jaringan saraf untuk meningkatkan produksi model hidrodinamika, yang memungkinkan integrasi semua informasi yang tersedia untuk membuat prediksi yang andal berdasarkan data yang diperoleh dari pantai tenggara Brazil [11]. Kemudian, data permukaan laut regional di Brittany, Perancis barat, digunakan untuk memprediksi kenaikan permukaan laut maksimum dengan metode *machine learning* [12]. Meskipun metode berbasis pembelajaran mesin menunjukkan beberapa prospek, metode berbasis ARIMA masih merupakan metode yang lazim untuk memprediksi permukaan laut, seperti yang ditunjukkan dalam penelitian terbaru [13]. Studi ini menganalisis anomali permukaan laut menggunakan dua metode, yaitu SARIMA dan *long short-term memory* (LSTM). Dalam hal ini, tren



Gambar 1. Dasbor untuk PUMMA PPI Ulee stasiun Lheu.

perubahan muka air laut diprediksi menggunakan model SARIMA untuk kondisi musiman, sedangkan LSTM digunakan untuk kondisi stokastik.

SARIMA adalah perpanjangan dari ARIMA. Metode ini adalah salah satu metode statistik paling terkenal untuk peramalan model deret waktu karena fleksibilitas, keserbagunaan, dan kualitasnya [14]. Model SARIMA memperhitungkan faktor musiman dalam deret waktu dan dapat memprediksi tren serta periode musiman dengan lebih baik dalam deret waktu. Model ini dapat memprediksi nilai masa depan dari nilai masa lalunya sendiri. SARIMA dapat dengan memuaskan menjelaskan deret waktu yang menunjukkan perilaku tidak konsisten di dalam dan di antara musim [15], [16]. Metode SARIMA cocok untuk penelitian ini karena sederhana, fleksibel, serbaguna, dan sesuai untuk *dataset* dari PUMMA yang memiliki pola musiman.

Tujuan makalah ini adalah melakukan analisis data dan menemukan model yang cocok untuk prediksi permukaan laut. Kontribusi makalah ini adalah: a) memanfaatkan data yang dihasilkan oleh PUMMA, yang merupakan sistem pemantauan permukaan laut yang baru untuk kebutuhan prediksi; b) menjelaskan secara mendalam pengembangan model prediksi menggunakan SARIMA, yang belum ada pada penelitian sebelumnya. Selanjutnya, bagian-bagian dari makalah ini disusun sebagai berikut. Pada Bagian II dan Bagian III, diberikan gambaran umum tentang PUMMA dan dijelaskan secara singkat aspek teoretis tentang metode ARIMA dan SARIMA. Bagian IV menjelaskan metodologi, sedangkan Bagian V berisi hasil dari setiap langkah, termasuk pengimporan dan prapemrosesan data, data stasioner, model estimasi, serta penemuan, prediksi, dan evaluasi. Bagian terakhir merupakan kesimpulan.

## II. GAMBARAN UMUM SISTEM PUMMA

PUMMA adalah sistem prototipe yang dikembangkan oleh Kementerian Kelautan dan Perikanan Indonesia untuk mengukur permukaan laut. PUMMA menyediakan data tinggi permukaan laut dari stasiun pemantauan jarak jauh, yang dipasang di wilayah laut Indonesia. Ini adalah sistem prototipe alat pengukur tinggi muka air laut murah untuk membantu deteksi dini bencana laut. Hingga tahun 2022, sistem PUMMA memiliki empat stasiun ukur, yaitu pelabuhan pendaratan ikan

(PPI) Ulee Lheu, Aceh; Mangrove Desa Gebang, Lampung; Muara Gembong, Bekasi; dan Pangandaran, Jawa Barat.

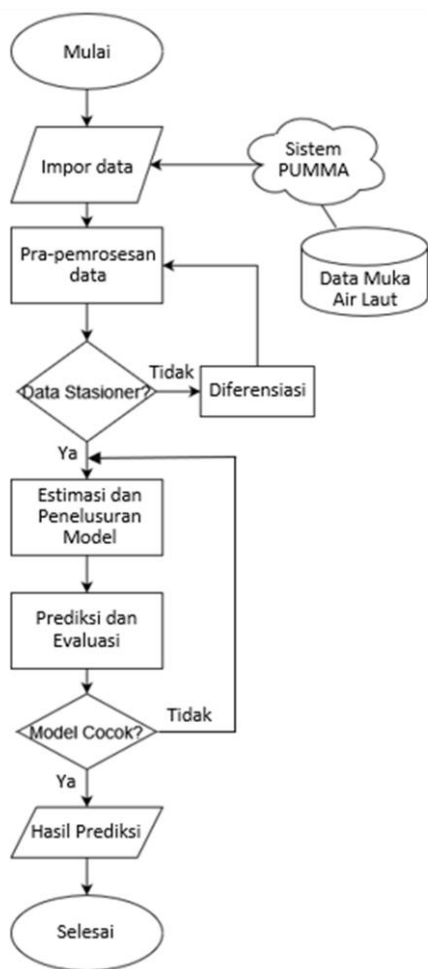
Setiap stasiun memiliki *remote monitoring system* (RMS). Komponen RMS terdiri atas *remote terminal unit* (RTU). Komponen RTU terdiri atas modem *wireless* 4G GSM, *wireless router*, CCTV, dan sensor ultrasonik. Untuk mengukur tinggi muka air laut, digunakan sensor *ultrasonic water level* dengan keluaran RS-232. Dipilih *dataset* KKP-003 PUMMA PPI Ulee Stasiun Lheu karena stasiun ini terletak di Aceh, yang pernah mengalami tsunami. Selain itu, di lokasi ini sistem PUMMA telah diinstal dan mulai mengukur sejak tahun 2021, sehingga menyimpan data paling banyak dibandingkan dengan stasiun lainnya. Dasbor stasiun ditampilkan pada Gambar 1.

Semua perangkat keras peralatan yang digunakan untuk membuat PUMMA merupakan komponen berbiaya rendah. Misalnya, kisaran harga pasaran modem 4G adalah 1 juta rupiah dan 5 juta rupiah, lalu dipilih modem yang harganya sekitar 500 ribu rupiah. Selanjutnya, dipilih *router* yang harganya kurang dari 300 ribu rupiah, padahal di pasaran banyak pilihan yang harganya jutaan rupiah. Lalu, dipilih CCTV yang harganya kurang dari 1 juta rupiah, padahal model lain harganya jutaan rupiah. Untuk sensor ultrasonik, dipilih sensor yang harganya 2 juta rupiah, sedangkan barang lain di pasaran harganya puluhan juta rupiah bahkan lebih.

## III. PREDIKSI DERET WAKTU

Analisis prediksi deret waktu adalah metode menganalisis serangkaian titik data historis yang dikumpulkan selama periode berurutan yang mencakup berbagai periode, seperti menit, jam, hari, bulan, dan kuartal. ARIMA, yang juga dikenal sebagai Box-Jenkins, menganalisis autokorelasi dalam deret waktu dengan memodelkannya secara langsung [17]. Metode ini terdiri atas rangkaian proses *autoregressive* (AR), *moving average* (MA), dan *stationary* (I). Model ARIMA menggabungkan dua metode, yaitu model AR dan MA [18].

ARIMA terdiri atas tiga parameter, yaitu  $p$ ,  $d$ , dan  $q$ . Parameter  $p$  adalah interpretasi AR, parameter  $d$  adalah interpretasi MA, sedangkan parameter  $q$  adalah urutan komponen MA. Selanjutnya, ARIMA membutuhkan data yang stasioner [19]. Jika data tersebut merupakan data nonstasioner, diperlukan proses diferensiasi untuk menghasilkan data yang stasioner. Persamaan ARIMA ditunjukkan dalam (1) [17], [19].



Gambar 2. Metodologi.

$$ARIMA(p, d, q)(P, D, Q)^S \quad (1)$$

dengan  $[p, d, q]$  adalah bagian nonmusiman dan  $[P, D, Q]$  adalah bagian musiman. Parameter  $S$  adalah jumlah periode per musim.

#### IV. METODOLOGI

Dalam makalah ini, teknik analisis data yang digunakan adalah metode SARIMA. Langkah-langkah penerapan metode SARIMA untuk memprediksi muka air laut di PUMMA ditunjukkan pada Gambar 2.

##### A. IMPOR DATA DAN PRAPEMROSESAN DATA

Untuk mengimpor data permukaan laut dari sistem PUMMA dan pembersihan data, dikembangkan *web service* HTTP-GET. Sistem ini menyimpan hasil RESPONSLayanan web dalam format JSON dan menyimpan data dalam server yang dapat diakses publik [20]. Berikut adalah cuplikan contoh format respons untuk GET *web service*.

```
{
  "ch00": [
    {
      "ts": 1645720408000,
      "value": "0.5894999999999999"
    },
    {
      "ts": 1645720418000,
      "value": "0.535"
    }
  ],
}
```

```
{
  "ts": 1645720418000,
  "value": "0.535"
}
]
```

*ID ch00* adalah *ID* perangkat untuk mengukur permukaan laut, *ts* adalah stempel waktu untuk tanggal-waktu pengumpulan, dan *value* adalah data pengukuran permukaan laut yang sebenarnya dalam satuan meter (m). Skrip berikut digunakan untuk tugas pengimporan data.

```
import json
import requests
import warnings
import datetime as dt
import pandas as pd
import matplotlib.pyplot as plt
import statsmodels.api as sm

URL =
'http://pumma.kkp.go.id/latestdata/pummadata-01july21-31july22.json'
json_object =
json.loads(requests.get(URL).text)
# original json string
TempData = pd.json_normalize(json_object)
```

```
data_muka_airlaut=TempData
data_muka_airlaut.drop(columns=['tanggal',
'ts'])
data_muka_airlaut.drop(data_muka_airlaut.columns[1], axis=1, inplace=True)
data_muka_airlaut.index =
pd.to_datetime(data_muka_airlaut.tanggal)
data_muka_airlaut.drop(columns=['tanggal'])
```

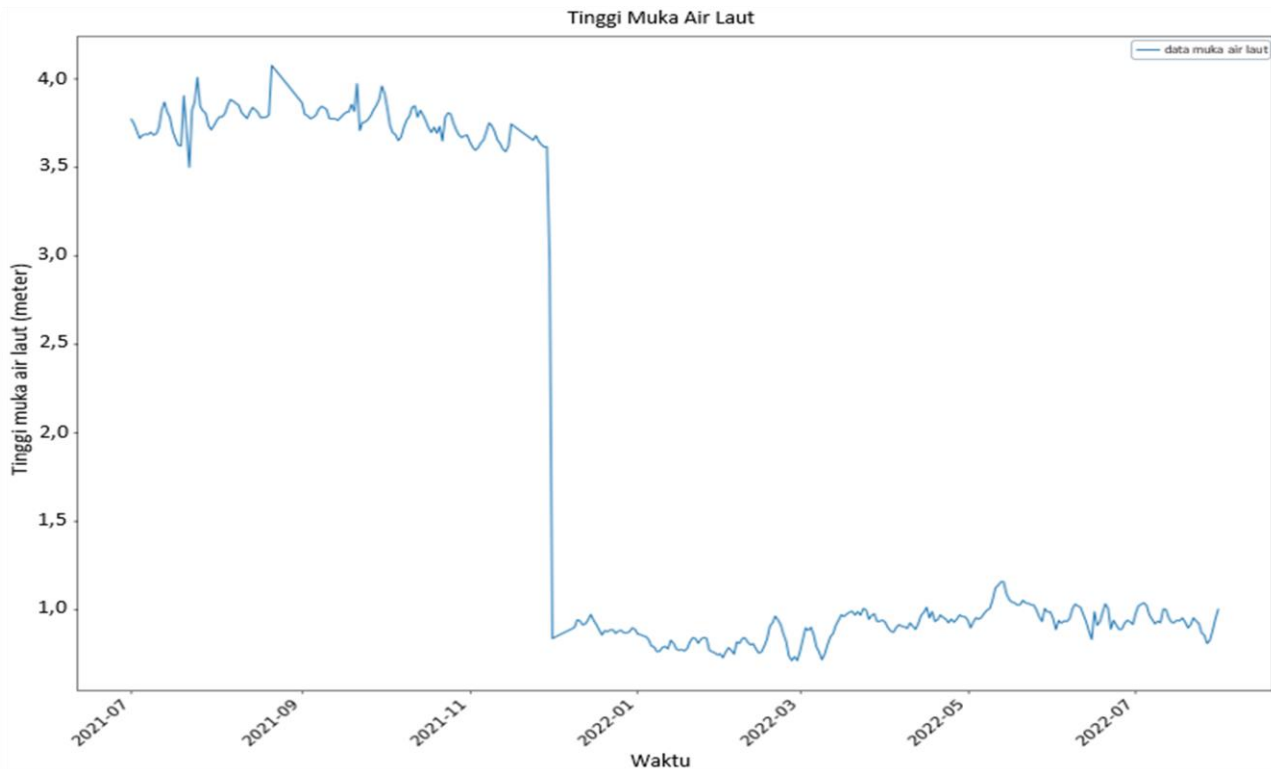
Skrip ini mendapatkan kumpulan data dari URL. Pertama, layanan web mengumpulkan kumpulan data per menit, kemudian data rata-rata per hari. Jumlah keseluruhan data per hari adalah sekitar 1.440 *instances*. Tahap selanjutnya adalah *resampling*, untuk mendapatkan sampel data rata-rata per hari, sehingga diperoleh sekitar 371 contoh data rata-rata harian.

##### B. UJI STASIONERITAS

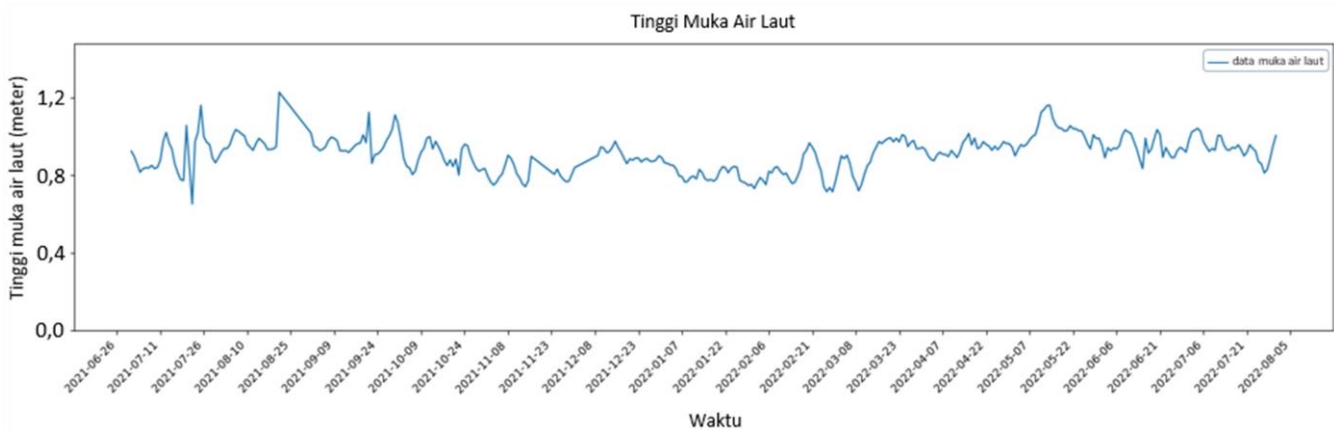
Pengujian stasioneritas adalah langkah dalam pemodelan AR. Bagian ini menggunakan *augmented Dickey Fuller* (ADF) untuk uji stasioner data karena ini adalah uji statistik yang paling sering digunakan untuk mengetahui suatu deret stasioner atau tidak. Prosedur untuk menentukan stasioneritas suatu data adalah dengan membandingkan nilai statistik ADF dengan nilai kritis dan *p-value*-nya [21]. Pada makalah ini, dipilih ADF yang merupakan salah satu uji statistik yang paling umum digunakan. Dalam ADF, proses perbedaan adalah metode untuk mengonversi deret waktu dari nonstasioner menjadi stasioner dengan menghitung selisih antara pengambilan.

##### C. ESTIMASI DAN PENELUSURAN MODEL

Pada tahap ini, model sementara diidentifikasi menggunakan *autocorrelation function* (ACF) dan *partial*



Gambar 3. Data permukaan laut per hari dari 1 Juli 2021 hingga 31 Juli 2022.



Gambar 4. Data permukaan laut per hari dari 1 Juli 2021 hingga 31 Juli 2022 setelah normalisasi.

*autocorrelation function* (PACF). Plot fungsi ini kemudian digunakan untuk menentukan kombinasi parameter  $p$ ,  $d$ , dan  $q$  pada SARIMA. Model yang sesuai dapat ditemukan dengan mengidentifikasi parameter  $p$ ,  $d$ , dan  $q$ . Dari identifikasi tersebut, dapat diperoleh nilai koefisien terbaik. Pada penelitian ini, dilakukan proses pemilahan model dengan menggunakan *Akaike information criterion* (AIC) dan uji signifikansi.

#### D. PREDIKSI DAN EVALUASI

Untuk prediksi, model terdefinisi terbaik dipilih dan dievaluasi menggunakan metrik *mean absolute percentage error* (MAPE), *R-squared*, *mean square error* (MSE), dan *root mean square error* (RMSE). Rumus metrik ini diberikan dalam persamaan berikut [22]–[24].

$$MAPE = \frac{1}{m} \sum_{t=1}^m \left[ \frac{Y_t - X_t}{Y_t} \right] \quad (2)$$

$$R - squared = \frac{\sum_{t=1}^m (X_t - Y_t)^2}{\sum_{t=1}^m (Y_t - Y_t)^2} \quad (3)$$

$$MSE = \frac{1}{m} \sum_{t=1}^m (X_t - Y_t)^2 \quad (4)$$

$$RMSE = \sqrt{\frac{1}{m} \sum_{t=1}^m (X_t - Y_t)^2} \quad (5)$$

dengan  $X_t$  adalah prediksi nilai  $t^{th}$ ,  $m$  adalah jumlah titik data, dan elemen  $Y_t$  adalah nilai  $t^{th}$  yang sebenarnya. Metode regresi memprediksi elemen  $X_t$  untuk elemen  $Y_t$  yang sesuai. Pemodelan SARIMA dikembangkan menggunakan skrip pemrograman berbasis Python3, termasuk *library* “Panda” sebagai *library* utama untuk proses analisis data.

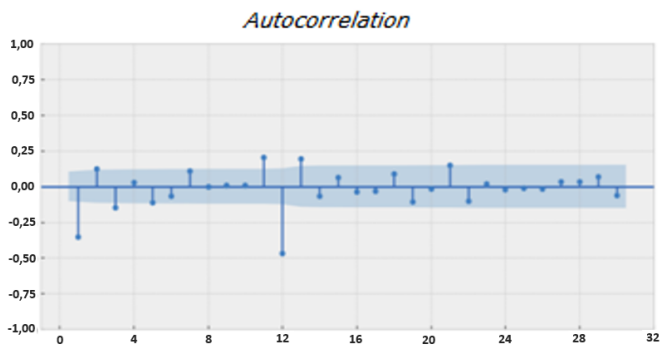
#### V. HASIL

Bagian ini menyajikan hasil dan langkah-langkah dalam mendapatkan model SARIMA untuk memprediksi tinggi muka air laut menggunakan *dataset* dari sistem PUMMA.

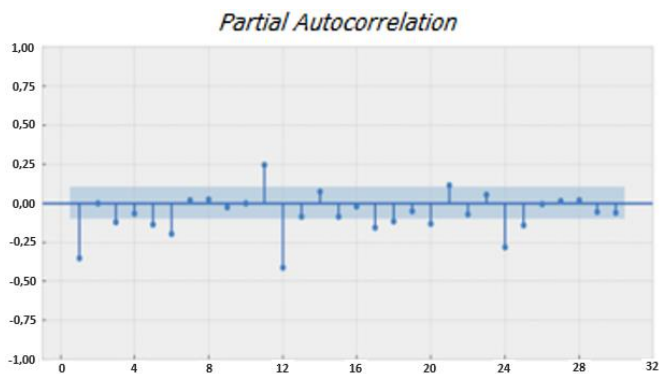
##### A. IMPOR DATA DAN PRAPEMROSESAN DATA

Impor data dan prapemrosesan data menghasilkan data pengamatan pengukuran tinggi muka air laut selama satu tahun, yang diambil dari tanggal 1 Juli 2021 pukul 00:00:00 sampai dengan 31 Juli 2022 pukul 23:59:59. Jumlah sampel data asli yang digunakan dalam penelitian ini adalah sekitar 534.200





Gambar 5. Plot ACF.



Gambar 6. Plot PACF.

data selama periode yang dipilih. Kumpulan data asli yang disediakan divisualisasikan seperti pada Gambar 3.

Berdasarkan Gambar 3, terjadi perubahan signifikan pada tinggi muka air laut rata-rata dari tanggal 29 November 2021 sampai dengan 1 Desember 2021, yaitu dari 3,61 m turun menjadi 0,83 m. Perubahan ini disebabkan oleh reposisi sensor pada tanggal 30 November 2021. Pemindehan sensor ke lokasi yang lebih dangkal dilakukan untuk mencegah kerusakan sistem di masa mendatang dan untuk mempermudah perawatan. Untuk menyediakan data pengukuran yang benar, modifikasi ini memerlukan penyesuaian di area pengaturan parameter dasbor *backend platform internet of things (IoT)*.

Oleh karena itu, *dataset* tersebut kemudian dinormalisasi dengan menghitung  $\delta$  sebagai selisih rata-rata muka air laut sebelum dan sesudah dilakukan reposisi sensor, yaitu  $\mu$  dan  $\mu'$  secara berturut-turut. Rata-rata muka air laut sebelum dilakukan reposisi,  $\mu$ , adalah 3,748 m dan nilai rata-rata muka air laut setelah dilakukan reposisi,  $\mu'$ , adalah 0,908 m. Jadi,  $\delta$  adalah 2,839 m. Selanjutnya, semua data sebelum dilakukan reposisi dikurangi dengan nilai  $\delta$ . Data yang dinormalisasi ditunjukkan pada Gambar 4.

**B. UJI STASIONERITAS**

Skrip berikut adalah keluaran dari hasil uji stasioneritas ADF.

The test statistic: -2.531014  
 p-value: 0.108109  
 Critical Values:  
 1%: -3.449  
 5%: -2.870  
 10%: -2.571

dengan *p-value* yang diperoleh adalah 0,108109. Nilai *p-value* ini lebih besar dari 0,05, sehingga dapat dikatakan data tidak stasioner. Oleh karena itu, data tersebut memerlukan proses *differencing* dengan nilai  $d = 1$ . Setelah dilakukan *differencing*,

TABEL I  
 NILAI AIC DARI ESTIMASI MODEL

Model	Nilai AIC
SARIMA( $p=1, d=1, q=1$ )( $P=0, D=1, Q=1$ ) <sup>12</sup>	-576,20
SARIMA( $p=1, d=1, q=1$ )( $P=1, D=1, Q=1$ ) <sup>12</sup>	-567,20
SARIMA( $p=0, d=1, q=1$ )( $P=0, D=1, Q=1$ ) <sup>12</sup>	-562,37
SARIMA( $p=0, d=1, q=1$ )( $P=1, D=1, Q=1$ ) <sup>12</sup>	-560,36
SARIMA( $p=1, d=1, q=0$ )( $P=0, D=1, Q=1$ ) <sup>12</sup>	-558,94
SARIMA( $p=1, d=1, q=0$ )( $P=1, D=1, Q=1$ ) <sup>12</sup>	-152,64

TABEL II  
 ESTIMASI PARAMETER MODEL

SARIMA ( $p,d,q$ ) ( $P,D,Q$ ) <sup>s</sup>	Parameter	Koef	z	P> z	Keputusan
SARIMA (1, 1, 1) (0, 1, 1) <sup>12</sup>	ar.L1	0,49	9,700	0,000	Signifikan
	ma.L1	0,91	-23,770	0,000	Signifikan
	ma.S.L12	-0,99	-5,230	0,000	Signifikan
SARIMA (1, 1, 1) (1, 1, 1) <sup>12</sup>	ar.L1	0,48	9,909	0,000	Signifikan
	ma.L1	-0,91	-25,110	0,000	Signifikan
	ar.S.L12	0,13	2,464	0,014	Signifikan
SARIMA (0, 1, 1) (0, 1, 1) <sup>12</sup>	ma.L1	-0,48	-15,610	0,000	Signifikan
	ma.S.L12	-0,70	-13,420	0,000	Signifikan
SARIMA (0, 1, 1) (1, 1, 1) <sup>12</sup>	ma.L1	-0,45	-24,910	0,000	Signifikan
	ar.S.L12	0,11	1,755	0,079	Tidak Signifikan
SARIMA (1, 1, 0) (0, 1, 1) <sup>12</sup>	ma.S.L12	-0,90	-15,340	0,000	Signifikan
	ar.L1	-0,35	-28,080	0,000	Signifikan
SARIMA (1, 1, 0) (1, 1, 1) <sup>12</sup>	ma.S.L12	-0,86	-17,860	0,000	Signifikan
	ar.L1	-0,36	-28,110	0,000	Signifikan
	ar.S.L12	0,09	1,447	0,148	Tidak Signifikan
ma.S.L12	-0,91	-14,800	0,000	Signifikan	

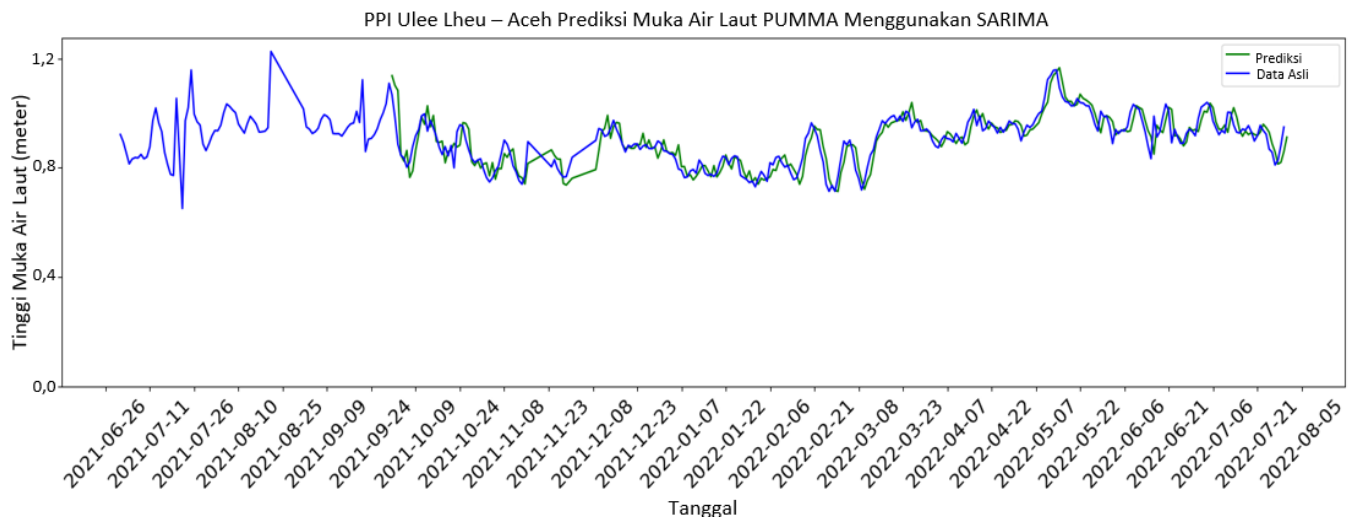
*p-value* yang diperoleh berhasil direduksi menjadi 0,000, yang lebih kecil dari 0,05. Skrip berikut adalah keluaran dari hasil uji stasioneritas setelah proses *differencing* dengan  $d = 1$  dari uji ADF.

The test statistic: -10.090788  
 p-value: 0.0000  
 Critical Values:  
 1%: -3.449  
 5%: -2.870  
 10%: -2.571

Dengan demikian, data setelah *differencing* menjadi stasioner. Setelah prosedur ini, data dibagi menjadi dua himpunan bagian. Subset pertama digunakan sebagai sampel untuk proses pelatihan, yang diambil dari tanggal 1 Juli 2021 hingga 31 Desember 2021, dan subset kedua digunakan untuk proses pengujian, yang diambil dari tanggal 1 Oktober 2021 hingga 31 Juli 2022.

**C. ESTIMASI DAN PENELUSURAN MODEL**

Setelah data menjadi stasioner, langkah selanjutnya adalah mengidentifikasi beberapa model sementara dengan melihat *plot ACF* dan *PACF*. *Plot ACF* seperti pada Gambar 5 menunjukkan *break off* pada *lag* ke-1 dan *lag* ke-12 untuk *lag* musiman. Adapun *plot PACF*, ditampilkan pada Gambar 6, menunjukkan *break off* pada *lag* ke-1 dan *lag* ke-12, ke-24 (kelipatan 12), kemudian  $S = 12$ . Berdasarkan hasil tersebut,



Gambar 7. Prediksi permukaan laut untuk SARIMA(1,1,0)(1,1,1)<sup>12</sup>.

beberapa model diperkirakan, seperti yang ditunjukkan pada Tabel I.

Selanjutnya, dilakukan uji signifikansi dan uji AIC untuk mendapatkan model yang sesuai dengan nilai AIC terkecil. Hasil pengujian AIC memiliki enam model dengan nilai AIC terendah, seperti ditunjukkan pada Tabel I. Di antara model-model ini, uji signifikansi dilakukan, seperti yang ditunjukkan pada Tabel II. Kolom Koef pada Tabel II adalah bobot dari setiap fitur dan menunjukkan dampaknya pada deret waktu. Jika nilai  $P > |z|$  kurang dari 0,05, parameter dapat dikatakan signifikan secara statistik.

Berdasarkan Tabel II, diperoleh model signifikan untuk semua parameter yang merupakan model estimasi. Estimasi model diuji menggunakan nilai MAPE sebagai berikut.

- SARIMA(1, 1, 1)(0, 1, 1)<sup>12</sup>, MAPE adalah 0,0634.
- SARIMA(1, 1, 1)(1, 1, 1)<sup>12</sup>, MAPE adalah 0,0622.
- SARIMA(0, 1, 1)(0, 1, 1)<sup>12</sup>, MAPE adalah 0,0647.
- SARIMA(1, 1, 0)(1, 1, 1)<sup>12</sup>, MAPE adalah 0,0615.

Model SARIMA terpilih dengan MAPE terkecil adalah SARIMA(1, 1, 0)(1, 1, 1)<sup>12</sup>, yang akan digunakan sebagai model prediksi.

SARIMA(1, 1, 0)(1, 1, 1)<sup>12</sup> harus memenuhi uji residual *white-noise* yang terdiri atas distribusi normal, nonautokorelasi, dan homogenitas seperti yang dipersyaratkan dalam pemodelan SARIMA sebagai berikut [17].

- Kolmogorov-Smirnov (*p-value*) adalah 5,15, yang lebih besar dari 0,05, sehingga dapat disimpulkan bahwa residual mengikuti distribusi normal. Maka, asumsi normalitas terpenuhi.
- Ljung-Box (*p-value*) adalah 0,35, yang juga lebih besar dari 0,05. Maka, dapat disimpulkan bahwa residual tidak autokorelasi. Dengan demikian, asumsi nonautokorelasi terpenuhi.
- *Heteroscedasticity* (*p-value*) adalah 0,34, yang lebih besar dari 0,05, sehingga dapat disimpulkan bahwa residual bersifat homogen.

Bentuk residualnya adalah *white-noise*, sehingga model SARIMA(1, 1, 0)(1, 1, 1)<sup>12</sup> dapat digunakan untuk prediksi berdasarkan *dataset* PUMMA.

#### D. PREDIKSI DAN EVALUASI

Hasil pencarian parameter optimal untuk model SARIMA disajikan pada Tabel II. Berdasarkan hasil estimasi, ditemukan

bahwa SARIMA(1, 1, 0)(1, 1, 1)<sup>12</sup> yang dianggap layak untuk memprediksi muka air laut untuk *dataset* PUMMA. Hasil prediksi diperlihatkan pada Gambar 7. Warna biru merupakan data asli dan warna hijau merupakan hasil prediksi. Misalnya, 29 Juli 2022 memiliki data asli 0,883170 dan nilai prediksi 0,820528, sehingga delta = 0,06. *Error* keseluruhan (MAPE) antara data prediksi dan data aktual adalah 6,13%. Langkah terakhir adalah mengevaluasi hasil prediksi untuk mengetahui kesalahan. Evaluasi dilakukan menggunakan metrik seperti *R-squared*, MSE, MAPE, dan RMSE.

- *R-squared* adalah 0,508 (*R-squared* bernilai antara 0 dan 1).
- Estimasi nilai *error* dalam prediksi (MSE) sebesar 0,00479. MSE yang rendah atau mendekati nol menandakan bahwa model dapat digunakan untuk perhitungan prediksi pada periode yang akan datang.
- Nilai akurasi ramalan pada metode prediksi (MAPE) sebesar 6,13%, yang berarti kurang dari 10%.
- Besarnya tingkat kesalahan prediksi (RMSE) adalah 0,069, yang berarti makin kecil atau mendekati 0.

Berdasarkan hasil evaluasi, model SARIMA(1, 1, 0)(1, 1, 1)<sup>12</sup> cocok untuk data prediksi muka air laut untuk *dataset* PUMMA.

## VI. KESIMPULAN

Penelitian ini bertujuan untuk membuat model prediksi muka air laut. Kumpulan data yang digunakan disediakan oleh sistem PUMMA. Dalam makalah ini, dirinci prosedur untuk menemukan model prediksi terbaik. Sebagai hasil dari penelitian yang dilakukan, SARIMA(1, 1, 0)(1, 1, 1)<sup>12</sup> adalah model prediksi yang cocok untuk muka air laut dari *dataset* PUMMA. Model dievaluasi berdasarkan nilai *R-squared*, MSE, MAPE, dan RMSE. Model mendapatkan *R-squared* sebesar 0,98, MSE sebesar 0,01731, dan RMSE sebesar 0,13. Dari hasil tersebut, model SARIMA(1, 1, 0)(1, 1, 1)<sup>12</sup> dianggap layak untuk prediksi dalam penelitian ini. Dengan semakin suksesnya pembelajaran mesin di banyak bidang, pendekatan berbasis jaringan saraf akan menjadi arah yang memungkinkan untuk pekerjaan di masa depan.

## KONFLIK KEPENTINGAN

Penulis makalah dengan judul “Prediksi Muka Air Laut dari Sistem PUMMA menggunakan SARIMA” ini menyatakan bahwa makalah ini bebas dari konflik kepentingan.

## KONTRIBUSI PENULIS

Konseptualisasi dan metodologi, Irfan Asfy Fakhry Anto; perangkat lunak, Oka Mahendra; supervisor sistem PUMMA, Semeidi Husrin; akuisisi data, Oka Mahendra dan Semeidi Husrin; analisis data, Irfan Asfy Fakhry Anto; analisis formal, Irfan Asfy Fakhry Anto; supervisi, Purnomo Husnul Khotimah; penelaahan, Purnomo Husnul Khotimah;

## UCAPAN TERIMA KASIH

Penelitian ini bekerja sama dengan Pusat Penelitian Kelautan (PUSRISKEK - Kementerian Kelautan dan Perikanan).

## REFERENSI

- [1] Syamsidik dkk., "Post-Tsunami Survey of the 28 September 2018 Tsunami Near Palu Bay in Central Sulawesi, Indonesia: Impacts and Challenges to Coastal Communities," *Int. J. Disaster Risk Reduct.*, Vol. 38, hal. 1-13, Agu. 2019, doi: 10.1016/j.ijdrr.2019.101229.
- [2] P. Harjadi, "Indonesia Tsunami Early Warning System (InaTEWS): Concept and Implementation," dipresentasikan pada *Int. Workshop Post Tsunami Soil Manage.*, Bogor, Indonesia. 2008.
- [3] H. Anwar dkk., "Guideline for Tsunami Risk Assessment in Indonesia Scientific Proposal for Practitioner and End Users Indonesian - German Working Group on Tsunami Risk Assessment," 2019, doi:10.13140/RG.2.2.20247.11685.
- [4] S. Husrin. (2020). PUMMA untuk Penguatan TEWS Indonesia [PowerPoint Slides]. [Online], [https://kkp.go.id/an-component/media/upload-gambar-pendukung/DitJaskel/publikasi-materi-2/pumma/semi\\_idsl\\_bungkus.pdf](https://kkp.go.id/an-component/media/upload-gambar-pendukung/DitJaskel/publikasi-materi-2/pumma/semi_idsl_bungkus.pdf)
- [5] S. Rahmstorf, "Sea-Level Rise: Towards Understanding Local Vulnerability," *Environ. Res. Lett.*, Vol. 7, No. 2, hal. 1-3, Jun. 2012, doi: 10.1088/1748-9326/7/2/021001.
- [6] Y. Yokoyama dan A. Purcell, "On the Geophysical Processes Impacting Palaeo-Sea-Level Observations," *Geosci. Lett.*, Vol. 8, hal. 1-19, Apr. 2021, doi: 10.1186/s40562-021-00184-w.
- [7] G. Griggs, "Rising Seas in California — An Update on Sea-Level Rise Science," in *World Scientific Encyclopedia of Climate Change: Case Studies of Climate Risk, Action, and Opportunity Volume 3*, J.W. Dash, Ed., Singapura: World Scientific Publishing Co. Pte. Ltd., 2021, hal. 105–111, doi: 10.1142/9789811213960\_0016.
- [8] M. Imani dkk., "Spatiotemporal Prediction of Satellite Altimetry Sea Level Anomalies in the Tropical Pacific Ocean," *IEEE Geosci., Remote Sens. Lett.*, Vol. 14, No. 7, hal. 1126–1130, Jul. 2017, doi: 10.1109/lgrs.2017.2699668.
- [9] S. Yang dkk., "Long-Term Prediction of Significant Wave Height Based on SARIMA Model in the South China Sea and Adjacent Waters," *IEEE Access*, Vol. 7, hal. 88082–88092, Jun. 2019, doi: 10.1109/access.2019.2925107.
- [10] R. Tur, E. Tas, A.T. Haghghi, dan A.D. Mehr, "Sea Level Prediction Using Machine Learning," *Water*, Vol. 13, No. 24, hal. 1-15, Des. 2021, doi: 10.3390/w13243566.
- [11] B.V.P. de Siqueira dan A. de M. Paiva, "Using Neural Network to Improve Sea Level Prediction Along the Southeastern Brazilian Coast," *Ocean Model.*, Vol. 168, hal. 1-14, Des. 2021, doi: 10.1016/j.ocemod.2021.101898.
- [12] N. Guillou dan G. Chapalain, "Machine Learning Methods Applied to Sea Level Predictions in the Upper Part of a Tidal Estuary," *Oceanologia*, Vol. 63, No. 4, hal. 531–544, Okt.-Des. 2021, doi: 10.1016/j.oceano.2021.07.003.
- [13] Q. Sun, J. Wan, dan S. Liu, "Estimation of Sea Level Variability in the China Sea and Its Vicinity Using the SARIMA and LSTM Models," *IEEE J. Sel. Top. Appl. Earth Observ., Remote Sens.*, Vol. 13, hal. 3317–3326, Jun. 2020, doi: 10.1109/JSTARS.2020.2997817.
- [14] O. Fathi, "Time Series Forecasting Using a Hybrid ARIMA and LSTM Model," *Velv. Consult.*, pp. 1–7, 2019. Tanggal akses: 9-Des-2022.
- [15] G.E.P. Box, G.M. Jenkins, and G.C. Reinsel, *Time Series Analysis: Forecasting and Control*, 4th ed. Hoboken, AS: John Wiley & Sons, Inc., 2008.
- [16] S. Wang, J. Feng, dan G. Liu, "Application of Seasonal Time Series Model in the Precipitation Forecast," *Math., Comput. Model.*, Vol. 58, No. 3–4, hal. 677–683, Agu. 2013, doi: 10.1016/j.mcm.2011.10.034.
- [17] A. Pankratz, *Forecasting with Univariate Box-Jenkins Models: Concepts and Cases*. Hoboken, USA: John Wiley & Sons, Inc, 1983.
- [18] U.A. Bhatti dkk., "Time Series Analysis and Forecasting of Air Pollution Particulate Matter (PM2.5): An SARIMA and Factor Analysis Approach," *IEEE Access*, Vol. 9, hal. 41019–41031, Feb. 2021, doi: 10.1109/ACCESS.2021.3060744.
- [19] G.E.P. Box, G.M. Jenkins, G.C. Reinsel, and G.M. Ljung, *Time Series Analysis: Forecasting and Control*, 5th ed. Hoboken, USA: John Wiley & Sons, Inc, 2016.
- [20] *Dataset Sea Level KKP-003 PUMMA PPI Ulee Lheu*, Des. 2022, [Online], <http://pumma.kkp.go.id/latestdata/pummadata-01july21-%2031july22.json>.
- [21] D.A. Dickey dan W.A. Fuller, "Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root," *Econometrica*, Vol. 49, No. 4, hal. 1057–1072, Jul. 1981, doi: 10.2307/1912517.
- [22] S. Kim dan H. Kim, "A New Metric of Absolute Percentage Error for Intermittent Demand Forecasts," *Int. J. Forecast.*, Vol. 32, No. 3, hal. 669–679, Jul.-Sep. 2016, doi: 10.1016/j.ijforecast.2015.12.003.
- [23] D.T. Allen, "Mean Square Error of Prediction as a Criterion for Selecting Variables," *Technometrics*, Vol. 13, No. 3, hal. 469–475, Agu. 1971, doi: 10.2307/1267161.
- [24] J.S. Armstrong dan F. Collopy, "Error Measures for Generalizing About Forecasting Methods: Empirical Comparisons," *Int. J. Forecast.*, Vol. 8, No. 1, hal. 69–80, Jun. 1992, doi: 10.1016/0169-2070(92)90008-w.