Prediction of the Sea Level from the PUMMA System Using SARIMA

Irfan Asfy Fakhry Anto[1], Oka Mahendra[2], Purnomo Husnul Khotimah[3], Semeidi Husrin[4]

[1,3] Research Center for Data and Information Science National Research and Innovation Agency, KST Samaun Samadikun - Sangkuriang Street, Bandung, 40135 INDONESIA (tel.: 022-2530050; email: [1]irfa013@brin.go.id, [3]purn005@brin.go.id)
[2] Research Centre for Smart Mechatronics National Research and Innovation Agency, KST Samaun Samadikun - Sangkuriang Street, Bandung, 40135 INDONESIA (tel.: 022-2530050; email: okam001@brin.go.id)
[4] Research Centre for Geological Disasters National Research and Innovation Agency, KST Samaun Samadikun - Sangkuriang Street, Bandung, 40135 INDONESIA (tel.: 022-2530050; email: seme001@brin.go.id)

[Received: 30 March 2023, Revised: 3 July 2023]
Corresponding Author: Irfan Asfy Fakhry Anto

ABSTRACT — The rising sea levels can threaten millions of people residing along the coast or lowlands. The risk can be mitigated by the sea-level prediction done by collecting information on the likelihood of rising sea levels. The Ministry of Marine Affairs and Fisheries of Indonesia has developed *Perangkat Ukur Murah untuk Muka Air Laut* (Inexpensive Device for Sea Level Measurement, PUMMA) to measure sea levels. PUMMA is located in remote monitoring stations based on Indonesian maritime area. The PUMMA system currently lacks a prediction feature. This objective of this study is to model the sea-level prediction using the dataset for one year, from July 2021 until July 2022. The seasonal autoregressive integrated moving average (SARIMA) method was used because SARIMA proved to be a flexible and versatile method for a dataset having noncomplex nature and seasonal patterns. This study has developed several models of the SARIMA. The model performance was evaluated using the mean absolute percentage error (MAPE), R-squared, mean square error (MSE), and root mean square error (RMSE) metrics. The SARIMA$(1, 1, 0)(1, 1, 1)^{12}$ model achieved the lowest prediction error with an R-squared of 0.508, MSE of 0.0479, and RMSE of 0.069. Based on the performance, SARIMA$(1, 1, 0)(1, 1, 1)^{12}$ model is feasible for predicting sea levels using the PUMMA dataset.

KEYWORDS — Sea Level, Prediction, PUMMA, Sea Level Prediction, Seasonal Moving Average.

I. INTRODUCTION

The Aceh tsunami in 2004 has made the Indonesian government more aware of the significance of the tsunami early warning infrastructure called the Indonesian Tsunami Early Warning System (InaTEWS) [1]. The InaTEWS system is a tsunami early warning system that still relies on seismic sensors to detect tsunamis [2]. One of the data that is necessitated by the InaTEWS is the sea-level data [3]. The Indonesian government continues to develop a prototype system for measuring the sea level. This system is known as *Perangkat Ukur Murah untuk Muka Air Laut* (Inexpensive Measuring Device for Sea Level, PUMMA) and is maintained by the Ministry of Marine Affairs and Fisheries of Indonesia. It is a low-cost sea-level measuring system that provides several sensors to record the sea-level data and occurrences of anomalies. The PUMMA system is prepared for supporting and being integrated with the InaTEWS [4]. The sea-level rise, which can risk people residing along coastal and lowland, is a fundamental geophysical parameter pertinent to many Earth science subfields, including oceanography, geodesy, and climatology [5], [6]. The sea-level prediction aims to collect information on the possible future sea-level rise to mitigate the risk of flooding and inundation, which may affect coastal communities and infrastructures [7].

Many studies have been conducted from various aspects of the sea level prediction or forecasting. Two prediction methods: autoregressive integrated moving average (ARIMA) and machine learning techniques (support vector regression), were used to estimate sea level anomalies (SLAs) based on altitude measurements from satellites in the Pacific Ocean [8]. The obtained results were compared to the observed values and the estimation of the sea level was calculated using the ARIMA routine model for the purpose of developing and validating the SLAs approach. Sea-level data were also utilized in 2017 to build a SARIMA model based on monthly average periodicity and predict the wave's height in the South China Sea and its surrounding waters [9].

More studies have been done using machine learning-based methods. The sea-level data from Turkish coastal monitoring activities were used [10]. Then, the sea-level prediction performance was compared by considering different scenarios using different prediction methods, such as the multiple linear regression and adaptive neuro-fuzzy inference system. In addition, linear and nonlinear models, which practitioners can use for port operations for sea-level prediction, were presented. Neural networks were employed to enhance the production of hydrodynamic models, allowing the integration of all available information to create reliable predictions based on data obtained from the southeast coast of Brazil [11]. After that, the regional sea-level data in Brittany-western France was used to predict the maximum sea-level rise using the machine learning method [12]. Recent studies demonstrate that ARIMA-based method remains a prevalent method for predicting the sea level, despite machine learning-based method shows some prospects. A study analyzed sea-level anomalies using the seasonal autoregressive integrated moving average (SARIMA) and long short-term memory (LSTM) [13]. In this study, the SARIMA was employed to predict the trend of sea level change by considering seasonal conditions, whereas the LSTM considers stochastic conditions.

SARIMA is an extension of the ARIMA. It is one of the best-known statistical methods for forecasting time series models due to its flexibility, versatility, and quality [14]. SARIMA model takes into account seasonal factors in the time series and can better predict seasonal trends and periods in the time series. It can predict future values from its own past values. It can satisfactorily explain time series that shows

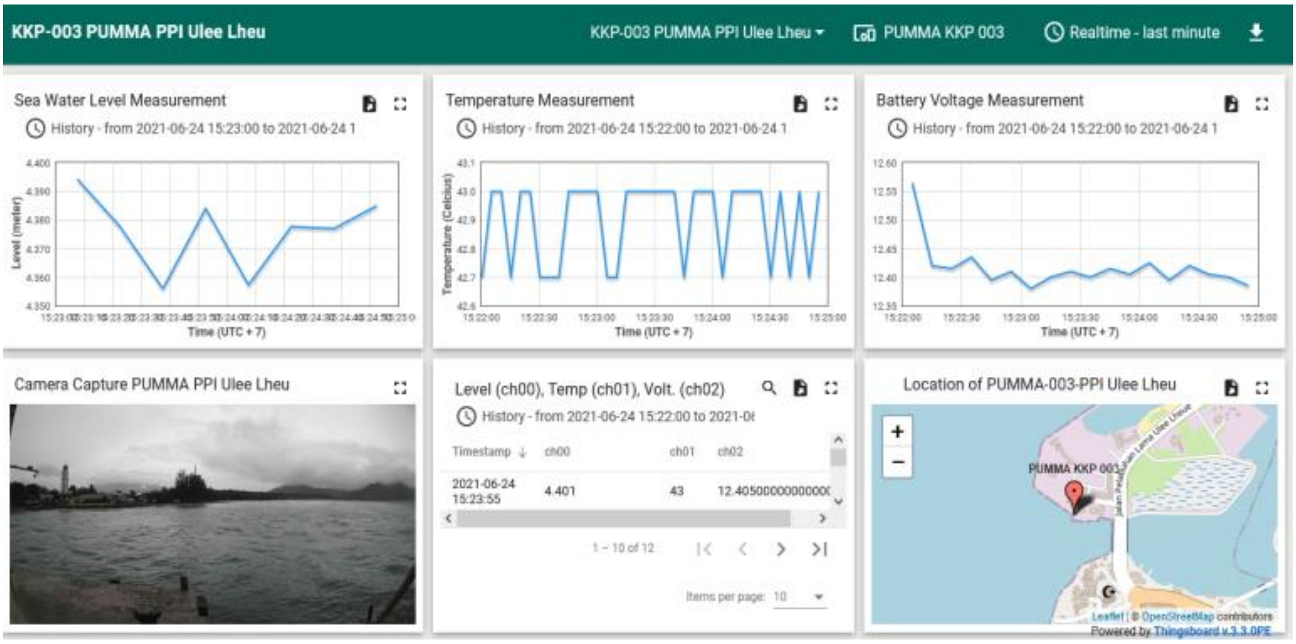

Figure 1. PUMMA dashboard of the PPI Ulee Lheu station.

inconsistent behavior within and between seasons [15], [16]. SARIMA method is suitable for this study because it is simple, flexible, versatile, and appropriate for the PUMMA dataset that has a seasonal pattern.

The study objective is to conduct data analysis and find a suitable model for sea-level prediction. The study contributions are a) the utilization of the data generated by PUMMA, which is a new sea-level monitoring system, and b) the exhaustive explanation of the prediction model's development using SARIMA, which must be added to previous studies. The remaining sections are organized as follows. Section II and Section III provide an overview of PUMMA and briefly explains the theoretical aspect of the ARIMA and SARIMA methods. Section IV explains the research method. Section V discusses the results of each step, including importing and data preprocessing, stationary data, estimation and finding models, prediction, and evaluation. The final section is a conclusion.

II. PUMMA SYSTEM OVERVIEW

PUMMA is a prototype system developed by the Ministry of Marine Affairs and Fisheries of Indonesia for measuring sea levels. It provides sea-level data from remote monitoring stations installed in Indonesia's maritime area. PUMMA is a prototype system for low-cost sea-level measuring instruments to assist early detection of ocean disasters. Until 2022, the PUMMA system has four measuring stations: fish landing port (*pelabuhan pendaratan ikan*, PPI) Ulee Lheu, Aceh; mangrove area in Gebang, Lampung; Muara Gembong, Bekasi; and Pangandaran, West Java.

Each station is equipped with a remote monitoring system (RMS). The RMS component consists of a remote terminal unit (RTU) comprising a 4G GSM wireless modem, wireless router, CCTV, and ultrasonic sensor. The sensor used to measure sea levels is the ultrasonic water level with an RS-232 output. The dataset used in this study was obtained from the KKP-003 PUMMA PPI Ulee Lheu station because it is located in Aceh. In addition, the PUMMA system in this location was installed in 2021 and has been operating since. Therefore, it stores the most data than any other station. The station's dashboard is displayed in Figure 1.

The equipment hardware to create PUMMA used all low-cost components. For example, pick a modem that costs roughly 500 thousand rupiahs, despite the fact that the price range for 4G modems on the market is between 1 million and 5 million rupiahs. Furthermore, for routers, choose those that cost less than 300 thousand rupiahs, even though numerous options in the market cost millions of rupiahs. Then, choose CCTV that costs less than one million rupiahs, despite other models costing millions of rupiahs. For ultrasonic sensors, pick a sensor that costs 2 million rupiahs, whereas other goods on the market cost tens of millions of rupiahs or more.

III. TIME SERIES PREDICTION

Time series prediction analysis is a method of analyzing a series of historical data points collected over consecutive periods spanning various periods such as minutes, hours, days, months, and quarters. ARIMA, also known as Box and Jenkins, analyzes autocorrelation in time series by modeling it directly [17]. This method consists of autoregressive (AR), moving average (MA), and stationary (I) series processes. ARIMA models combine two methods: autoregressive (AR) and moving average (MA) models [18].

ARIMA is comprised of three parameters, including p, d, and q. The parameter p represents the AR, d represents the MA, and q represents the order of the moving average components. Next, ARIMA requires stationary data [19]. If the data are nonstationary, they must undergo the differencing process to produce stationary data. The ARIMA equation is shown as [17], [19]:

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

where [p, d, q] is the nonseasonal part and [P, D, Q] is the seasonal part. The parameter S is the number of periods per season.

IV. METHODODOLOGY

In this study, the SARIMA method was used for data analysis techniques. The steps for applying the SARIMA method to predict sea level in the PUMMA as illustrated in Figure 2.

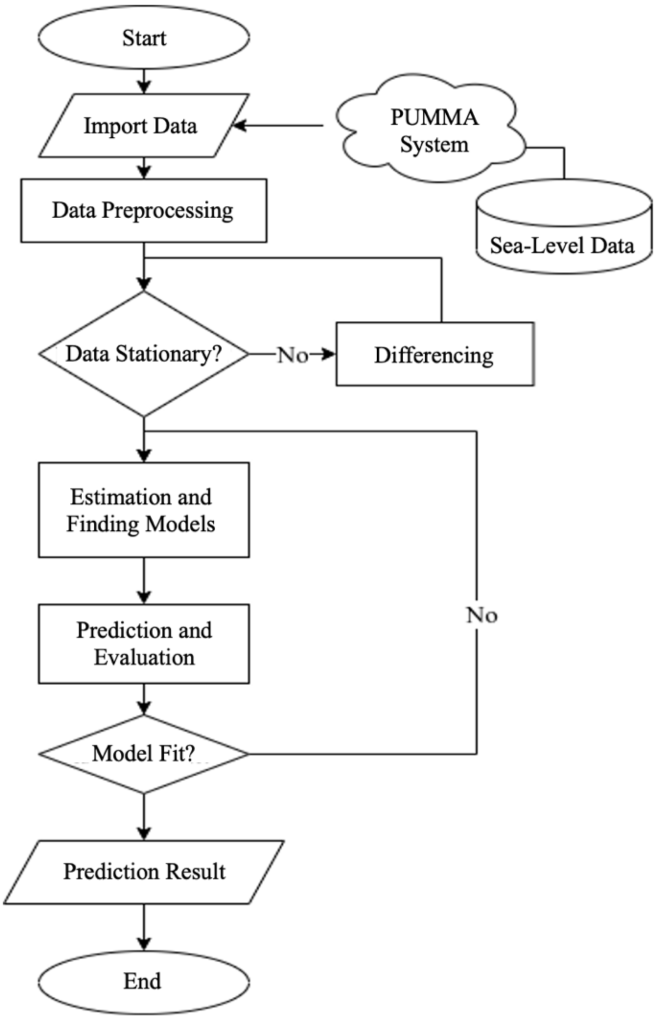

Figure 2. Methodology.

A. IMPORT DATA AND DATA PREPROCESSING

A web service HTTP-GET was developed to acquire the sea-level measurement data for importing sea-level data from the PUMMA system and data cleaning. It stored the web service RESPONSE results in JSON format and stored the data in a dedicated web server that is accessible to the public [20]. The following is a snippet of an example of a response format for the GET web service.


```
{
  "ch00": [
    {
      "ts": 1645720408000,
      "value": "0.5894999999999"
    },
    {
      "ts": 1645720418000,
      "value": "0.535"
    },
    {
      "ts": 1645720418000,
      "value": "0.535"
    }
  ]
}
```


The *ch00* is a device ID for measuring the sea level, *ts* is the timestamp for the collection of date-time, and *value* is the real sea-level measurement data in meters unit (m). The following script is used for the data importing task.

```
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/pumm
adata-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=['tangga
l', 'ts'])
data_muka_airlaut.drop(data_muka_airlau
t.columns[1], axis=1, inplace=True)
data_muka_airlaut.index =
pd.to_datetime(data_muka_airlaut.tangga
l)
data_muka_airlaut.drop(columns=['tangga
l'])
```

This script retrieved the datasets from the URL. First, a web service collected the dataset per minute. All data were averaged by day. Daily data totaled approximately 1,440 instances. The next stage involved resampling to obtain daily sample average data. Hence, approximately 371 instances of daily data were collected.

B. STATIONARITY TEST

Testing for stationarity is a common task in autoregressive modeling. This task used augmented Dickey-Fuller (ADF) for a data stationary test because it is the most commonly used statistical test to determine if a series is stationary. The procedure for determining the stationarity of data involved comparing the statistical value of the ADF with its critical value and p-value [21]. This study utilized ADF, which is one of the most commonly used statistical tests. In ADF, a differencing process is a method to convert time series from nonstationary to stationary by computing the difference between outtakes.

C. ESTIMATION AND FINDING MODELS

In this stage, the autocorrelation function (ACF) and partial autocorrelation function (PCF) were used to identify temporary models. These function plots were then used to determine the combination of parameters p, d, and q in SARIMA. The appropriate models can be found by identifying the parameters p, d, and q. This identification allows for the best coefficient values to be obtained. In this study, the model was sorted using the Akaike information criterion (AIC) and the significance test.

D. PREDICTION AND EVALUATION

For the prediction, the best-defined model was selected and evaluated using metrics: mean absolute percentage error

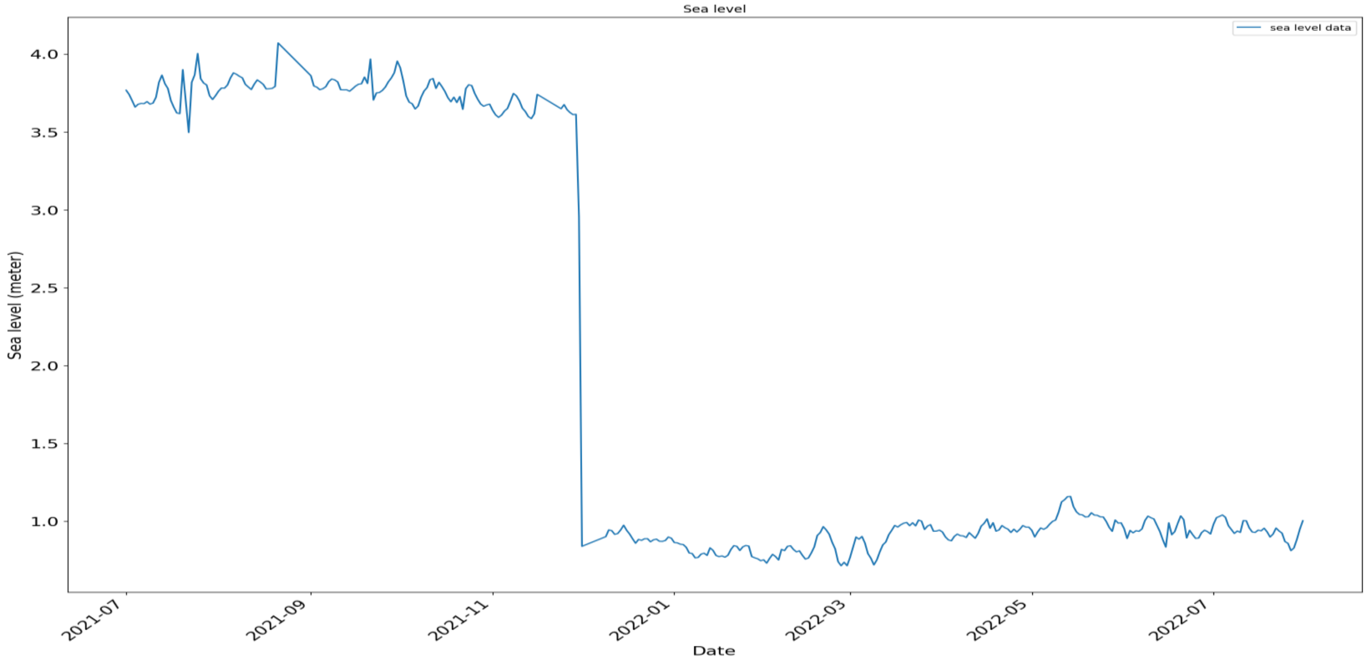

Figure 3. Sea-level data per day from 1 July 2021 to 31 July 2022.

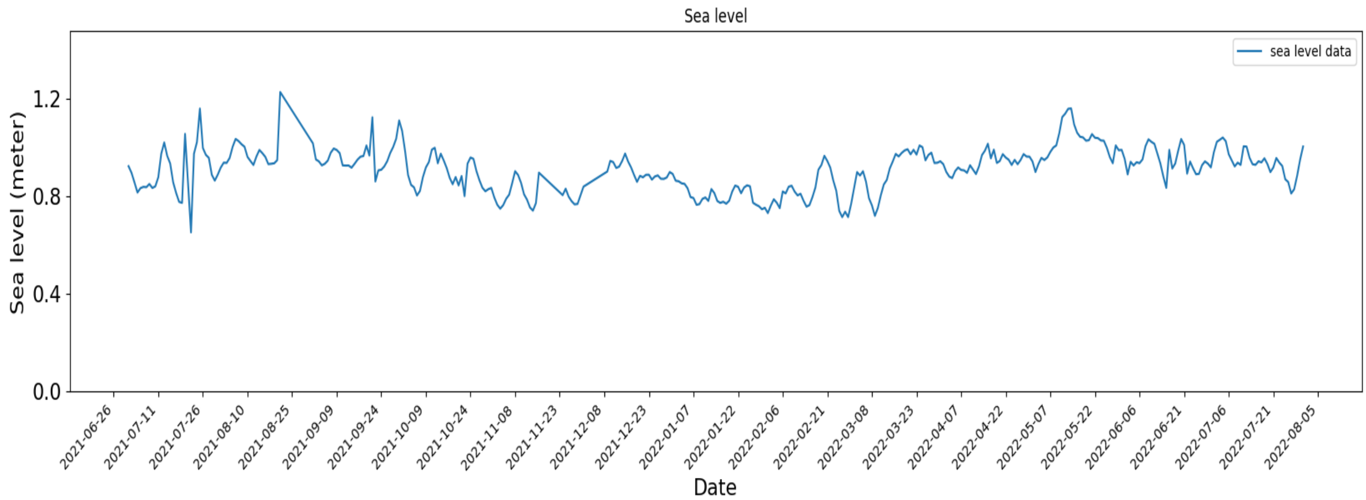

Figure 4. Sea-level data per day from 1 July 2021 to 31 July 2022 after being normalized.

(MAPE), R-squared, mean square error (MSE), and root mean square error (RMSE). These metric formulas are given in the (2)–(5) [22]–[24].

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

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

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

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

where X_t is the predicted t^{th} value, m is the total amount of data points, and the Y_t element is the actual t^{th} value. The regression method predicts the X_t element for the corresponding Y_t element of the ground truth dataset. The SARIMA modeling experiment was developed using a programming script based on Python 3 including the "Panda" library as the main library for the data analysis process.

V. RESULT AND DISCUSSION

This section provides results and steps to obtain the SARIMA model for the sea-level prediction using a dataset from the PUMMA system.

A. *IMPORTING AND DATA PREPROCESSING*

The data import and preprocessing stages produce observational data on sea level measurements for one year taken from 1 July 2021 at 00:00:00 to 31 July 2022 at 23:59:59. Approximately 534,200 original data samples were used in this study during the designated period. Figure 3 depicts a visualization of the original data set provided.

Figure 3 shows a significant drop in the average sea level from 29 November 2021 to 1 December 2021, from 3.61 m down to 0.83 m. This change was the result of the sensor repositioning on 30 November 2021. The sensor was relocated to a shallower location to prevent future system damage and make maintenance easier. This modification necessitates an adjustment in the parameter settings area of the Internet of things (IoT) platform's backend dashboard to provide correct measurement data.

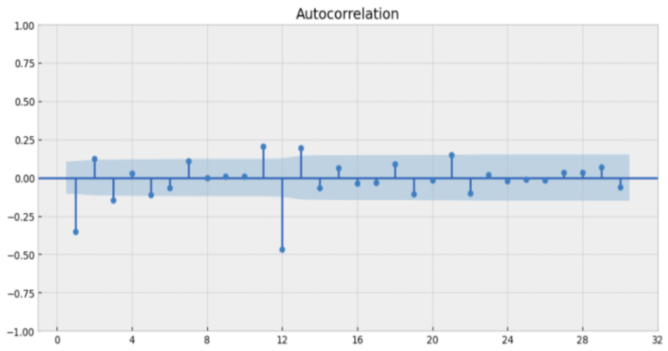

Figure 5. ACF plots.

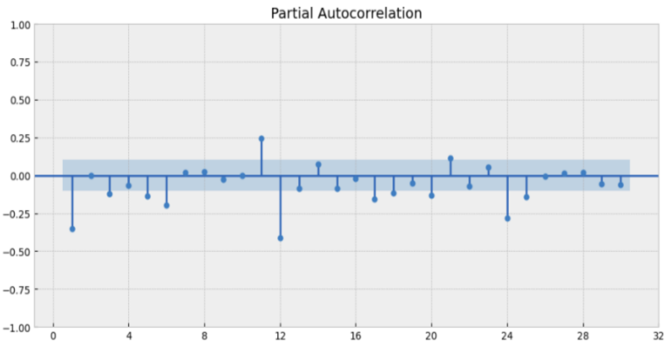

Figure 6. PACF plots.

Therefore, the dataset was then normalized by calculating δ as the difference of the sea-level average from before and after the sensor repositioning, namely μ and μ', respectively. For the sea-level average before repositioning, μ was 3,748 m; for the sea-level average value after repositioning, μ' was 0,908 m. Hence, δ was 2,839 m. Afterwards, all data before the repositioning were subtracted by the δ value. The normalized data are shown in Figure 4.

B. STATIONARITY TEST

The following script is the output of stationarity results of the ADF test.

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

the p-value obtained was 0.108109. This p-value is greater than 0.05, so the data are not stationary. Therefore, the data need a differencing process with a value of d = 1. After differencing, the p-value was successfully reduced to 0.000, which is less than 0.05. The following script is the output of stationarity results after differencing d = 1 of the ADF test.

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

Hence, the data after differencing became stationary. After this procedure, the data were then divided into two subsets. The first subset was used as a sample for the training process and was collected from 1 July 2021 to 31 December 2021. The second subset was used for testing process and was collected from 1 October 2021 until 31 July 2022.

TABLE I
AIC VALUE OF ESTIMATED MODELS

Model	AIC Value
SARIMA(p=1, d=1, q=1)(P=0, D=1, Q=1)^{12}	-576.20
SARIMA(p=1, d=1, q=1)(P=1, D=1, Q=1)^{12}	-567.20
SARIMA(p=0, d=1, q=1)(P=0, D=1, Q=1)^{12}	-562.37
SARIMA(p=0, d=1, q=1)(P=1, D=1, Q=1)^{12}	-560.36
SARIMA(p=1, d=1, q=0)(P=0, D=1, Q=1)^{12}	-558.94
SARIMA(p=1, d=1, q=0)(P=1, D=1, Q=1)^{12}	-152.64

TABLE II
PARAMETER ESTIMATION OF MODELS

SARIMA $(p,d,q)(P,D,Q)^S$	Parameters	Coef	z	P>|z|	Decision
SARIMA (1, 1, 1)(0,1, 1)^{12}	ar.L1	0.49	9.700	0.000	Significant
	ma.L1	-0.91	-23.770	0.000	Significant
	ma.S.L12	-0.99	-5.230	0.000	Significant
SARIMA (1, 1, 1)(1, 1, 1)^{12}	ar.L1	0.48	9.909	0.000	Significant
	ma.L1	-0.91	-25.110	0.000	Significant
	ar.S.L12	0.13	2.464	0.014	Significant
	ma.S.L12	-0.98	-5.169	0.000	Significant
SARIMA (0, 1, 1)(0, 1, 1)^{12}	ma.L1	-0.48	-15.610	0.000	Significant
	ma.S.L12	-0.70	-13.420	0.000	Significant
SARIMA (0, 1, 1)(1, 1, 1)^{12}	ma.L1	-0.45	-24.910	0.000	Significant
	ar.S.L12	0.11	1.755	0.079	Not significant
	ma.S.L12	-0.90	-15.340	0.000	Significant
SARIMA (1, 1, 0)(0, 1, 1)^{12}	ar.L1	-0.35	-28.080	0.000	Significant
	ma.S.L12	-0.86	-17.860	0.000	Significant
SARIMA (1, 1, 0)(1, 1, 1)^{12}	ar.L1	-0.36	-28.110	0.000	Significant
	ar.S.L12	0.09	1.447	0.148	Not Significant
	ma.S.L12	-0.91	-14.800	0.000	Significant

C. ESTIMATION AND FINDING MODEL

After the data were stationary, the next step was to identify a few temporary models by looking at the autocorrelation function (ACF) and partial autocorrelation function (PACF) plots. The ACF plot in Figure 5 shows the break-off in the first lag and the twelfth for the seasonal lag. At the same time, the PACF plot in Figure 6 depicts the break-off in the first, twelfth, and 24th lags (multiples of 12), then S = 12. Based on these results, several models are estimated, as shown in Table I.

Next, the significance and AIC tests were carried out to obtain the suitable model by the smallest AIC value. The AIC test results had six models with the lowest AIC values, as shown in Table I. A significance test was performed between these models, as shown in Table II. Column Coef in Table II is the weight of each feature and shows the impact on the time series. If P value > |z| is less than 0.05, the parameter can be said to be statistically significant.

Based on Table II, the significant models for all parameters that an estimated model was obtained. The estimated model was tested using the following MAPE value.

- SARIMA(1, 1, 1)(0, 1, 1)^{12} , MAPE is 0.0634

p-ISSN 2301–4156 | e-ISSN 2460–5719

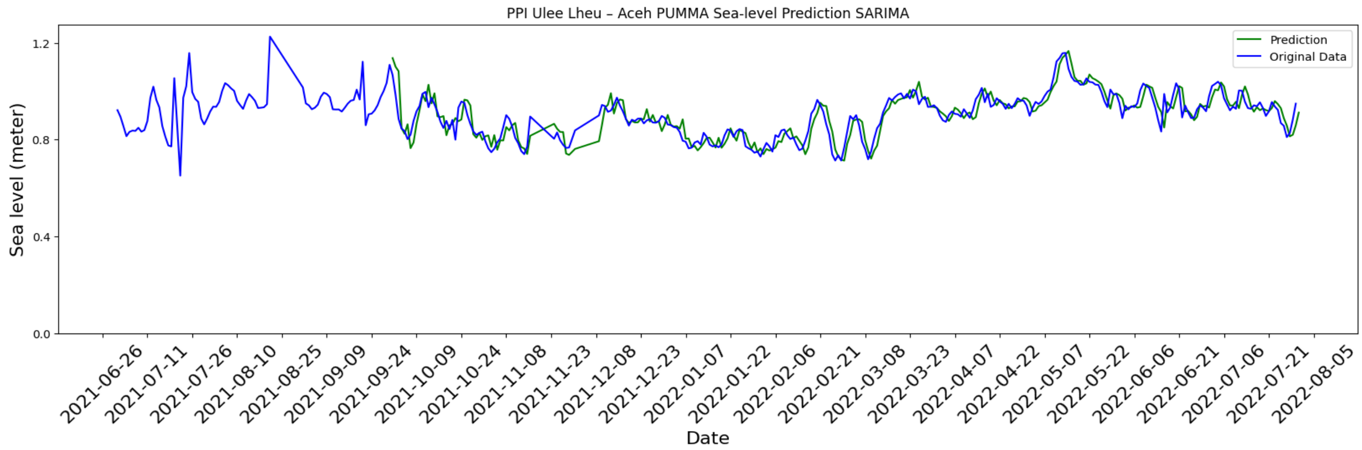

Figure 7. Sea-level prediction for SARIMA(1, 1, 0)(1, 1, 1)^{12}.

- SARIMA(1, 1, 1)(1, 1, 1)^{12} , MAPE is 0.0622
- SARIMA(0, 1, 1)(0, 1, 1)^{12}, MAPE is 0.0647
- SARIMA(1, 1, 0)(1, 1, 1)^{12}, MAPE is 0.0615.

The selected SARIMA model with the smallest MAPE was SARIMA(1, 1, 0)(1, 1, 1)^{12} , which was as the prediction model. SARIMA(1, 1, 0)(1, 1, 1)^{12} must fulfill the residual white-noise test which consists of: normal distribution, no autocorrelation, and homogeneity, as required in SARIMA modelling [17].

- The Kolmogorov-Smirnov (p-value) is 5.15, which is greater than 0.05. It can be concluded that the residuals follow a normal distribution, so the assumption of normality is met
- The Ljung-Box (p-value) is 0.35, which is greater than 0.05. It can be concluded that the residuals are not autocorrelated, thus, the no autocorrelation assumption is fulfilled
- The heteroskedasticity (p-value) is 0.34, which is greater than 0.05. It can be concluded that the residuals are homogeneous.

The residuals form is white noise so the SARIMA(1, 1, 0)(1, 1, 1)^{12} model can be used for prediction based on PUMMA dataset.

D. PREDICTION AND EVALUATION

Table II displays the results of determining the optimal parameters for the SARIMA model. On the basis of estimation results, the SARIMA(1, 1, 0)(1, 1, 1)^{12} was found, which is considered feasible to predict sea level for the PUMMA dataset. Prediction results can be seen in Figure 7. The blue color represents original data, while the green color represents the result of the prediction. For instance, the original data for 29 July 2022 were 0.883170, the predicted value was 0.820528, and the delta was 0.06. The overall error (MAPE) between predictive and actual data was 6.13%. The final step involved the evaluation of the prediction results to determine the errors. The evaluation utilized metrics such as R-squared, MSE, MAPE, and RMSE.

- R-squared was 0.508, R-squared is between 0 and 1.
- The estimated error value in prediction (MSE) was 0.00479. A low MSE or close to zero indicates that it can be used for prediction calculations in the future period.
- The value of forecast accuracy in the prediction method (MAPE) was 6.13%, which is less than 10%.
- The magnitude of the prediction error rate (RMSE) was 0.069, which is smaller or closer to 0.

Based on the evaluation results, SARIMA(1, 1, 0)(1, 1, 1)^{12} model is suitable for the sea-level prediction data for the PUMMA dataset.

VI. CONCLUSION

This study seeks to create a model for predicting sea-level rise. The PUMMA system provided the dataset used. This study outlined the procedure for determining the best prediction model. The research results suggest that SARIMA(1, 1, 0)(1, 1, 1)^{12} is the suitable model for predicting sea level from the PUMMA dataset. The model was evaluated based on the R-squared, MSE, MAPE, and RMSE values. The model yielded an R-squared of 0.98, an MSE of 0.01731, and an RMSE of 0.13. From the results, SARIMA(1, 1, 0)(1, 1, 1)^{12} model is considered feasible for the prediction for this study. With the growing success of machine learning in numerous fields, a neural network-based approach would be a possible direction for future work.

CONFLICT OF INTEREST

The authors of paper "Prediction of the Sea Level from the PUMMA System Using SARIMA" declare that this paper is free from conflict of interest.

AUTHOR CONTRIBUTION

Conceptualization and methodology, Irfan Asfy Fakhry Anto; software, Oka Mahendra; PUMMA system supervisor, Semeidi Husrin; data acquisition, Oka Mahendra, Semeidi Husrin; data analytic tool, Irfan Asfy Fakhry Anto; formal analysis, Irfan Asfy Fakhry Anto; supervision, Purnomo Husnul Khotimah; review, Purnomo Husnul Khotimah.

ACKNOWLEDGMENT

This research is in collaboration with the Research Center for Marine (PUSRISKEL - The Ministry of Marine Affairs and Fisheries).

REFERENCES

[1] Syamsidik *et al.*, "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, pp. 1–13, Aug. 2019, doi: 10.1016/j.ijdrr.2019.101229.

[2] P. Harjadi, "Indonesia Tsunami Early Warning System (InaTEWS): Concept and Implementation", presented at the *Int. Workshop Post Tsunami Soil Manage.*, Bogor, Indonesia. 2008.

[3] H. Anwar *et al*., *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/semeidi_idsl_bungus.pdf

[5] S. Rahmstorf, "Sea-Level Rise: Towards Understanding Local Vulnerability," *Environ. Res. Lett.*, Vol. 7, No. 2, pp. 1–3, Jun. 2012, doi: 10.1088/1748-9326/7/2/021001.

[6] Y. Yokoyama and A. Purcell, "On the Geophysical Processes Impacting Palaeo-Sea-Level Observations," *Geosci. Lett.*, Vol. 8, pp. 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., Singapore: World Scientific Publishing Co. Pte. Ltd., 2021, pp. 105–111, doi: 10.1142/9789811213960_0016.

[8] M. Imani *et al*., "Spatiotemporal Prediction of Satellite Altimetry Sea Level Anomalies in the Tropical Pacific Ocean," *IEEE Geosci., Remote Sens. Lett.*, Vol. 14, No. 7, pp. 1126–1130, Jul. 2017, doi: 10.1109/lgrs.2017.2699668.

[9] S. Yang *et al*., "Long-Term Prediction of Significant Wave Height Based on SARIMA Model in the South China Sea and Adjacent Waters," *IEEE Access*, Vol. 7, pp. 88082–88092, Jun. 2019, doi: https://doi.org/10.1109/access.2019.2925107.

[10] R. Tur, E. Tas, A.T. Haghighi, and A.D. Mehr, "Sea Level Prediction Using Machine Learning," *Water*, Vol. 13, No. 24, pp. 1–15, Dec. 2021, doi: 10.3390/w13243566.

[11] B.V.P. de Siqueira and A. de M. Paiva, "Using Neural Network to Improve Sea Level Prediction Along the Southeastern Brazilian Coast," *Ocean Model.*, Vol. 168, pp. 1–14, Dec. 2021, doi: 10.1016/j.ocemod.2021.101898.

[12] N. Guillou and G. Chapalain, "Machine Learning Methods Applied to Sea Level Predictions in the Upper Part of a Tidal Estuary," *Oceanologia*, Vol. 63, No. 4, pp. 531–544, Oct.–Dec. 2021, doi: 10.1016/j.oceano.2021.07.003.

[13] Q. Sun, J. Wan, and S. Liu, "Estimation of Sea Level Variability in the China Sea and Its Vicinity Using the Sarima and LTSM Models," *IEEE J. Sel. Top. Appl. Earth Observ., Remote Sens.*, Vol. 13, pp. 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. Access date: 9-Dec-2022.

[15] G.E.P. Box, G.M. Jenkins, and G.C. Reinsel, *Time Series Analysis: Forecasting and Control*, 4th ed. Hoboken, USA: John Wiley & Sons, Inc., 2008.

[16] S. Wang, J. Feng, and G. Liu, "Application of Seasonal Time Series Model in the Precipitation Forecast," *Math., Comput. Model.*, Vol. 58, No. 3–4, pp. 677–683, Aug. 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 *et al.*, "Time Series Analysis and Forecasting of Air Pollution Particulate Matter (PM2.5): An SARIMA and Factor Analysis Approach," *IEEE Access*, Vol. 9, pp. 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*, Dec. 2022, [Online], http://pumma.kkp.go.id/latestdata/pummadata-01july21-%2031july22.json

[21] D.A. Dickey and W.A. Fuller, "Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root," *Econometrica*, Vol. 49, No. 4, pp. 1057–1072, Jul. 1981, doi: 10.2307/1912517

[22] S. Kim and H. Kim, "A New Metric of Absolute Percentage Error for Intermittent Demand Forecasts," *Int. J. Forecast.*, Vol. 32, No. 3, pp. 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, pp. 469–475, Aug. 1971, doi: 10.2307/1267161.

[24] J. S. Armstrong and F. Collopy, "Error Measures for Generalizing About Forecasting Methods: Empirical Comparisons," *Int. J. Forecast.*, Vol. 8, No. 1, pp. 69–80, Jun. 1992, doi: 10.1016/0169-2070(92)90008-w.