Maritime Technology and Society
Vol. 5(1): 91-98, February 2026
https://doi.org/10.62012/mp.vi.48211

Regular Research Article

Implementation of the Least Square Method in Tidal Harmonic Analysis Using Python Programming

Suci Fitra Rahmadania, Ilham Cahya*, Ashari Wicaksono, and Inayatul Lailiyah

Marine Science Study Program, Department of Marine Science and Fisheries, Faculty of Agriculture, Trunodjoyo Madura University, Indonesia

*ilham.cahya@trunojoyo.ac.id ; Tel.: +62 821-3247-6178

Abstract: This study aims to analyze the tidal characteristics in the southern coastal area of Madura using the least squares-based harmonic analysis method with Python programming. This method is used to obtain tidal harmonic constants and determine the tidal type in the study area. The analysis results show that the M_2 component is the dominant component with an amplitude of 0.498 m, followed by the K_1, O_1, and S_2. The Formzahl value of 1.056 indicates that the tidal type in the southern coastal area of Madura is a mixed tide tending to be semidiurnal. Model validation using the Root Mean Square Error (RMSE) produced a value of 0.201 m, indicating that the modeling results have a fairly good level of accuracy. The results of this study are expected to be a reference in the planning and management of coastal areas in Madura waters.

Keywords: Madura strait; Google Colab; Formzahl; Coding; Tidal type

1. Introduction

Coastal areas are transitional ecosystems between land and sea, with complex environmental dynamics. Tidal activity is a major factor influencing coastal morphology, sediment transport, and human activities in coastal areas. Tides are the periodic rise and fall of sea levels due to the gravitational pull of celestial bodies, especially the moon and sun. Information on tides is an important basis for various coastal development and management activities such as port planning, navigation, fisheries activities, reclamation, and mitigating the impacts of climate change. The availability of accurate tidal data and analysis is essential for Indonesia's vast and diverse coastal areas [1].

Madura Island has a geographical location with coordinates of 112° 40’ 32”E to 114° 37’ 17”E and 6° 52’ 42”S, where this position is strategic because it is located in the north of East Java Province and is surrounded by waters with different characteristics [2]. The northern coast of Madura borders the Java Sea which is relatively open so that it is more affected by the dynamics of the open sea. The southern coast of Madura borders the Madura Strait which is a semi-enclosed water and receives tidal influences from the Java Sea and the Indian Ocean through the Bali Strait. These geographical and hydrodynamic conditions give rise to the suspicion of significant variations in tidal components, both in terms of amplitude and phase of the main harmonic components.

Harmonic tidal analysis can be performed using various methods. The least squares method is considered more accurate and flexible because it minimizes the sum of the squares of the differences between observed data and tidal models. The resulting harmonic constant estimates are superior to conventional methods such as the Admiralty method. This method is capable of processing data with limited series length and is more robust to data noise. The application of the least squares method to compare tidal components in two

Received: 2025-11-10; Accepted: 2026-01-13

coastal areas with different hydrodynamic characteristics in Madura is still rare [3]

Tides are not only influenced by celestial bodies but can also be influenced by the basic bathymetry of a body of water and are influenced by the coastline, especially for closed sea areas such as bays. Tidal variations not only determine the differences in harmonic constants between regions, but also affect sediment transport through the currents formed, thereby regulating the pattern of deposition and erosion in shipping lanes [4]. Detailed information about the variations in tidal components on the south coast of Madura is also important for data-based coastal area development and management planning. Tidal harmonic constants such as M_2, S_2, K_1, and O_1 can be used to determine the type of tides in a body of water and can also help estimate variations in sea level elevation, tidal currents, and their impacts on port activities, reclamation, and coastal abrasion. Tide types are divided into semidiurnal, diurnal, and mixed [5].

Accurate knowledge of tidal components in coastal areas plays a crucial role in supporting sustainable coastal planning and management [6]. Madura Island, as an area with distinct water characteristics between the north and south, requires studies capable of presenting tidal parameters clearly and representatively. Previous research has been limited in uncovering tidal harmonic constants, particularly in the southern coastal areas of Madura, and few studies have used programming in tidal analysis in the region. This study was conducted by applying a least squares-based harmonic analysis method using the Python programming language to obtain more precise and systematic tidal parameters. The results of this study are expected to provide a scientific contribution to the development of tidal analysis and serve as a reference for coastal management in Madura and similar studies in other archipelagos.

2. Materials and Methods

2.1 Materials

This research was conducted in Bangkalan Regency, East Java Province, at the TADDAN (South) station in Sampang Regency. The tidal data used in this study is primary. The primary data, sea level data, was obtained from the Seamless Reference Geodetic Information (SRGI) database managed by the Geospatial Information Agency (BIG) with an observation period of one year, from August 1, 2024 to August 31, 2025, accessible through the website (https://srgi.big.go.id/tides). The location of the Taddan tidal station in Sampang Regency was selected based on the availability of tidal data in the study area, as shown in Figure 1.

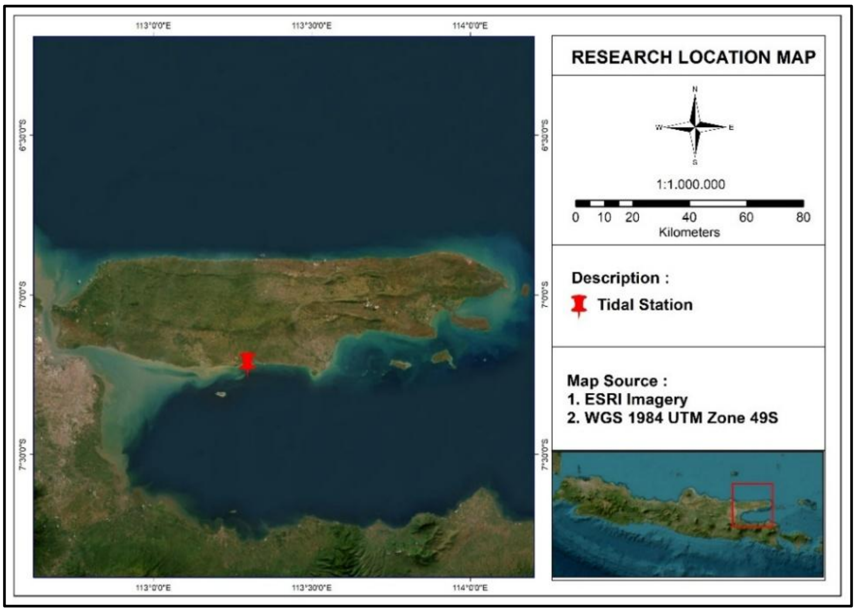

Figure 1. Sampling Location Map

2.2 Methods

Python programming is used for the entire tidal data processing process. Data is read and cleaned with the pandas and numpy libraries to ensure correct time formats and numerical values. Sea level elevation datum adjustments are first performed through a python script so that all data are at the same height reference. Analysis of the principal harmonic constants is then performed using the least squares method using (numpy.linalg.lstsq) to obtain the amplitude and phase of each tidal component. The harmonic constant values are then used in the python script to calculate the Formzahl number so that the tidal type in the study area can be determined. The Root Mean Square Error (RMSE) calculation is carried out at the final stage to evaluate the accuracy of the tidal model results against the observation data. The Root Mean Square Error (RMSE) method is used because it can describe the overall model accuracy and is sensitive to large errors, making it suitable for evaluating the accuracy of tidal predictions.

Tidal elevation is assumed to follow simple harmonic motion and can be represented as a superposition of sinusoidal functions.

$$\eta(t) = S_0 + \sum_{i=1}^{N}[a_i \cos(\omega_i t) + b_i \sin(\omega_i t)] \quad (1)$$

The tidal elevation, denoted as η(t)**,** is expressed as a function of time and represents the variation of sea level relative to the mean sea level (S_0)**.** The tidal signal is decomposed into a finite number of harmonic components, where a_i and b_i are the cosine and sine coefficients of the *i*-th tidal harmonic component, respectively, describing the amplitude and phase characteristics of each constituent. The parameter $\boldsymbol{\omega_i}$ represents the angular frequency of the *i*-th tidal component, while t denotes time measured from the beginning of the observation period. The total number of harmonic components included in the analysis is given by N**,** which determines the level of detail and accuracy of the tidal representation.

The coefficients a_i and b_i were estimated using the least squares method by minimizing the sum of squared differences between observed and modeled tidal elevations. The least squares formula according to [7] is :

$$\eta(t) = S_0 + \sum_{i=1}^{N} A_i \cos(\omega_i - P_i) \quad (2)$$

The tidal elevation, denoted as *η(t)***,** represents the variation of sea level as a function of time relative to the mean sea level (S_0)**.** The tidal signal is modeled as the superposition of N harmonic components, where A_i and P_i correspond to the amplitude and phase of the *i*-th tidal component, respectively. Time is represented by ***T***, while the angular frequency of each component, ω_i, is defined as $\omega_i = 2\pi / T_i$**,** where $\mathbf{T_i}$ denotes the period of the *i*-th tidal constituent. This formulation enables an accurate representation of tidal behavior by capturing the periodic characteristics of individual tidal components. According to a journal written by [8] the RMSE equation is as follows:

$$\text{RMSE} = \sqrt{\frac{\sum(x_{1i} - x_{2i})^2}{n}} \quad (2)$$

The accuracy of the model predictions is evaluated using the Root Mean Square Error (RMSE)**,** which quantifies the average magnitude of the differences between observed and predicted values. In this formulation, the summation symbol **∑** represents the aggregation of errors from the first to the *n*-th data point, where x_{1i} denotes the *i*-th observed or actual data value and x_{2i} represents the corresponding predicted value generated by the model. The parameter **n** indicates the total number of observations–prediction pairs used in the analysis. RMSE provides a single measure of predictive performance, with lower values indicating a closer agreement between the model results and the observed data.According to [8] RMSE values are classified as 0–0.5 (very good), 0.5–0.6 (good), 0.6–0.7 (sufficient), and >0.7 (poor).

According to [9], the Formzahl number formula is as follows:

$$F = \frac{H_{K_1} + H_{O_1}}{H_{M_2} + H_{S_2}} \quad (4)$$

The Formzahl number (F) is a dimensionless parameter used to classify tidal types based on the relative dominance of diurnal and semidiurnal tidal constituents. In this formulation, HK_1 represents the amplitude of the principal diurnal tidal component generated by the combined gravitational influences of the moon and the sun, while HO_1 denotes the amplitude of the principal diurnal tidal component caused solely by the moon's gravitational force. Meanwhile, HM_2 corresponds to the amplitude of the principal semidiurnal tidal component induced by the moon, and HS_2 represents the amplitude of the principal semidiurnal tidal component generated by the sun. These tidal constituents are collectively used to determine the prevailing tidal regime at the study location.

The determination of the tidal type can be done after obtaining the Formzahl value. The classification of tidal types is divided into four, where a value of $0.00 < F \leq 0.25$ indicates a semidiurnal tidal type. A value of $0.25 < F \leq 1.50$ indicates a mixed tidal type tending to be semidiurnal, $1.50 < F \leq 3.00$ indicates a mixed tidal type tending to be diurnal. A value of $F > 3.00$ indicates a diurnal tidal type.

2.3 Python Programming Procedures

Data processing and harmonic tidal analysis in this study were conducted using the Python programming language. To clarify the workflow of data processing and harmonic tidal analysis conducted in this study, the overall procedure is summarized in the form of a flowchart, The research stages are shown in Figure 2.

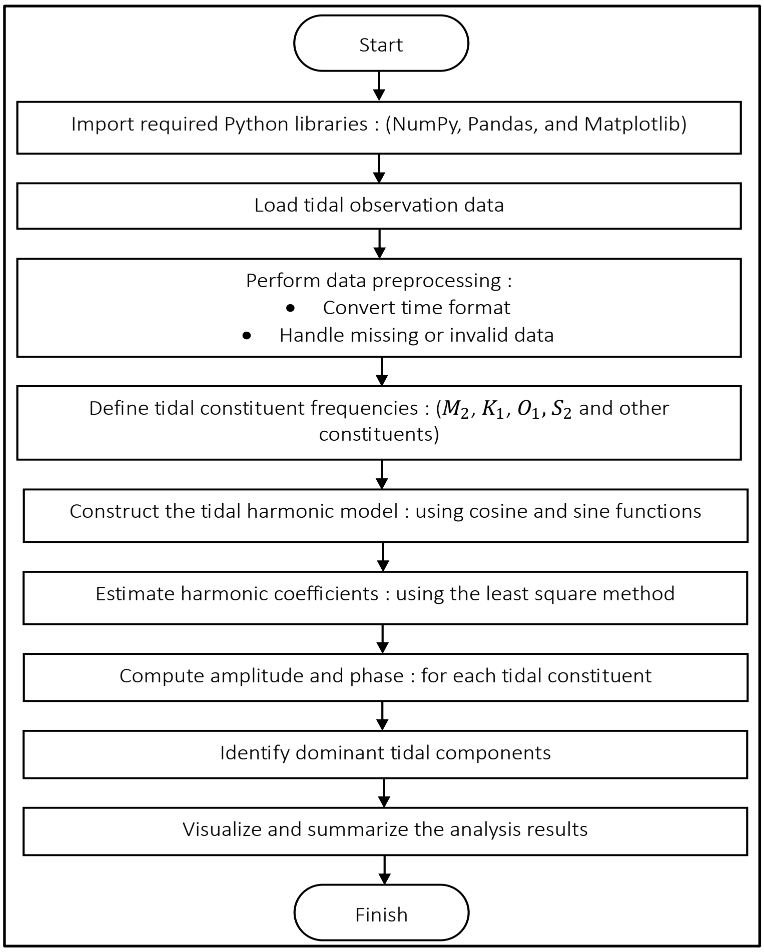

Figure 2. Research Stages

3. Results

3.1 Tidal Components Least Square Method

The results of processing tidal components from stations in the coastal area of Madura Island, namely the TADDAN (south) station in Sampang Regency during the period of August 1, 2024 to August 31, 2025 using the least square method are shown in Table 1 below:

Table 1. Tidal Components of the Least Square Method

Constant	TADDAN Station (south)	
	Amplitude (meter)	Phase (-180° to 180°)
S_0	–0.005	0
M_2	0.498	15.021
S_2	0.255	115.565
N_2	0.099	75.700
K_2	0.089	-147.743
K_1	0.485	-30.207
O_1	0.310	-72.127
Q_1	0.064	-4.390
P_1	0.132	48.561
M_4	0.004	-140.445
MS_4	0.005	3.075
M_f	0.024	-136.955
M_m	0.014	-51.686

The results of the tidal harmonic analysis using the least square method indicate that the TADDAN (southern) station has a mean sea level (MSL) component of approximately –0.005 m. The largest tidal amplitude is associated with the M_2 component, with a value of 0.498 m, while the smallest amplitude is observed in the M_4 component, with a value of 0.004 m. The highest phase value is recorded for the S_2 component at approximately 115.565°, whereas the lowest phase value reaches –147.743°. The dominance of the M_2 tidal component at the TADDAN (southern) station indicates that lunar semidiurnal tidal forcing is the primary factor controlling tidal variability in the study area. The M_2 constituent is widely recognized as the strongest tidal component in many coastal and shallow-water regions. Its relatively high amplitude compared to other tidal constituents suggests that semidiurnal tidal waves can propagate effectively in the waters surrounding Madura Island. This condition is influenced by regional hydrodynamic characteristics, such as bathymetry and coastal geometry, which can enhance the amplification of semidiurnal tidal signals.

3.2 Root Mean Square Error (RMSE) Test

The Root Mean Square Error (RMSE) test results obtained from data processing are 0.201. The classification of RMSE calculations according to the journal written by [8], that the RMSE value of 0.201 is included in the range of 0-0.5, so that the results of the Root Mean Square Error (RMSE) test of observation data and prediction models are classified as very good, this shows that, the average difference between model predictions and observation data is in this range. This RMSE value is relatively small compared to the observed tidal amplitude variations, indicating that the applied harmonic model has a good level of accuracy and is acceptable for tidal prediction purposes at the research location. Therefore, the developed tidal model is able to represent tidal characteristics with a practically acceptable level of error. The tidal graph is shown in Figure 3 below. The graph shows tidal data at TADDAN Station (south) with sea level elevations ranging from approximately –1.5 meters to over 1.8 meters. Observational data show the highest sea level values reaching approximately 1.8–1.85 meters, while the lowest values are in the range of –1.4 to –1.5 meters. The prediction model displays elevations with a slightly narrower range, with peak tides around 1.4–1.5 meters and the lowest ebb tides over –1.6 meters. This comparison indicates that the prediction model is able to follow the main tidal patterns, including high and low tide fluctuations, although there are still differences in the highest and lowest extreme values. This indicates that the model is good enough to represent the general tidal trend at the location.

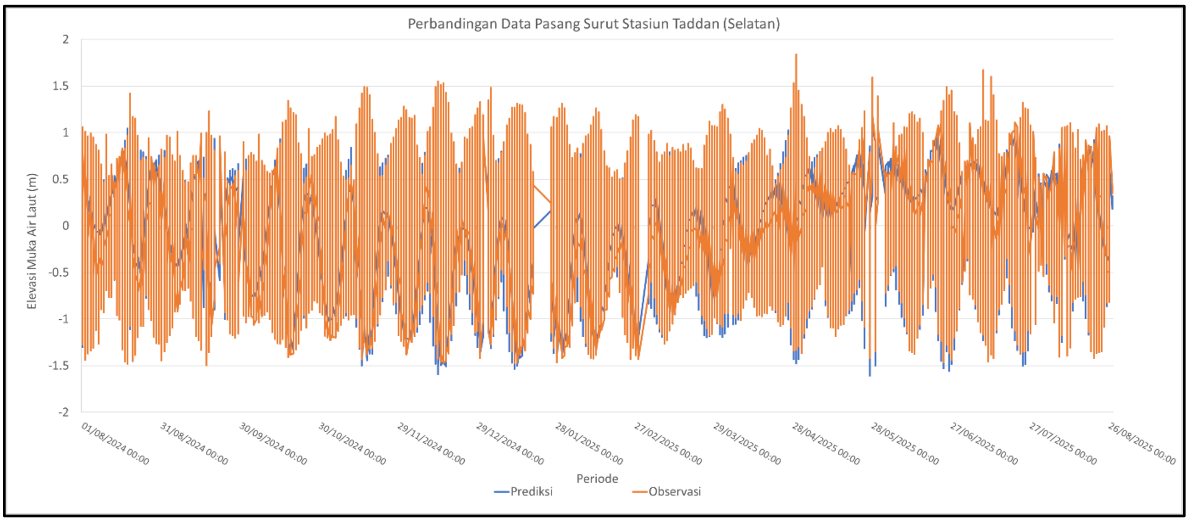

Figure 3. comparison graph of observational data with the prediction model

3.3 Tidal Type

The determination of tidal types in this study is based on the results of the amplitude and phase analysis of tidal harmonic components. The formzhal values of the amplitudes of the principal harmonic constants, such as M_2, S_2, K_1, and $O_1$are then calculated to serve as a reference in determining the tidal type classification. The calculation of the Formzahl value at TADDAN station (south) was carried out using the equation proposed by [9], and the results obtained were as below. The results of the calculation of the formzhal values for the two stations are presented in Table 3.

Table 3. Tidal Types in the Coastal Area of TADDAN Station (South Madura)

Location	Period	*Formzhal* (*least square*)	Tidal Type
TADDAN station (south)	August 1, 2024 to August 31, 2025	1.056	Mixed, semidiurnal leaning

The results of tidal type analysis at the Taddan (southern) station during the observation period from 1 August 2024 to 31 August 2025 show a Formzahl value of 1.056. Based on the classification proposed by [10], this value falls within the range of 0.25 < F ≤ 1.50, indicating that the tidal type at this location is classified as mixed tide with a tendency toward semidiurnal characteristics. This classification suggests that semidiurnal tidal components remain dominant; however, diurnal components also exert a noticeable influence, resulting in differences between the first and second high tides within a single day. These results are consistent with the findings of [7], who classified the tidal type in the Mandangin Island waters, Sampang Regency, as mixed with semidiurnal dominance. With a Formzahl value of 1.056, the present study falls within the same category, supporting the conclusion that tidal dynamics in the Sampang region are still predominantly controlled by semidiurnal components. This similarity in tidal classification indicates a relatively stable tidal pattern in the Sampang area despite the presence of local variations among observation points.

3.4 Sea Level Elevation

The results of sea level elevation produce values from several datums, namely the average sea level (Mean Sea Level) or MSL, the average high sea level (Mean High Water Level) or MHWL, the average high sea level at full tide (Mean High Water Springs) or MHWS, the highest high sea level (Highest High Water Level) or HHWL, the average low water level (Mean Low Water Level) or MLWL, the average lowest sea level at full tide (Mean Low Water Springs) or MLWS, the lowest low water level (Lowest Low Water Level) or LLWL, the astronomically highest receding sea level (Highest Astronomical Tide) or HAT and the astronomically lowest receding sea level (Lowest Astronomical Tide) or LAT. The calculated values of sea level elevation at TADDAN station (south) during the observation period of August 1, 2024 to August 31, 2025 are shown in Table 4.

Table 4. Sea Level Elevation

Taddan Station (South) August 1, 2024 to August 31, 2025	
Sea level elevation datum	Value (meter)
MSL	-0.001
MHWL	0.469
MHWS	1.065
HHWL	1.840
MLWL	-0.412
MLWS	-1.167
LLWL	-1.490
HAT	1.295
LAT	-1.551

The observation results indicate a Mean Sea Level (MSL) of –0.001 m, with a Highest High Water Level (HHWL) of 1.840 m and a Lowest Low Water Level (LLWL) of –1.490 m. The estimated Highest Astronomical Tide (HAT) reaches 1.295 m, while the Lowest Astronomical Tide (LAT) is –1.551 m.

4. Discussion

The results of the study indicate that the tidal characteristics on the southern coast of Madura, particularly at Taddan Station, are dominated by the semidiurnal component, with the highest amplitude in the M_2 component. This condition indicates that the influence of the Moon's gravitational pull-on sea level in the region is greater than the diurnal component. This finding supports the hypothesis that the southern coast of Madura has complex tidal characteristics due to the interaction between astronomical forces and the morphology of semi-enclosed waters.

The mixed tidal type of classification, leaning towards semidiurnal, indicates that two high tides and two low tides occur in a single day, with different tide heights. This indicates the relatively strong influence of the diurnal component on the semidiurnal tidal system. Dynamically, this can be explained by the location of the southern coast of Madura directly facing the Madura Strait and connected to the Java Sea and the Indian Ocean via the Bali Strait, allowing for interference of tidal waves from two different ocean systems.

The calculated Root Mean Square Error (RMSE) of 0.201 indicates that the least squares method is capable of accurately representing tidal patterns. This value falls into a very good category, indicating that the model has a low error rate in representing the observed data. The effectiveness of the least squares method in this study aligns with the findings of [3] and [5], who stated that this method provides more stable estimates of harmonic constants than the Admiralty method, especially for data with a long observation period.

Despite the good performance of the harmonic least squares model, this study has several limitations. The analysis focuses primarily on astronomical tidal components and does not explicitly incorporate meteorological effects such as wind forcing, atmospheric pressure variations, and storm surges, which may contribute to shorter sea level fluctuations. In addition, the model assumes linear superposition of tidal constituents and spatial homogeneity at the observation station, which may oversimplify complex hydrodynamic processes in semi-enclosed coastal waters. Therefore, the results are most representative of tidal behavior under normal conditions and should be interpreted with these limitations in mind.

Scientifically, the dominance of the M_2 component amplitude in the southern Madura region indicates that the influence of semidiurnal tides is still very significant, even though the region has semi-enclosed geographic characteristics. This result is in line with the research conducted in study [1], where the tidal type obtained was also classified as mixed leaning semidiurnal. The similarity of these results indicates that regional oceanographic conditions, such as coastline configuration and bathymetric characteristics, play an important role in controlling the variations in the amplitude and phase of tidal components in the coastal area of Madura.

The practical implications of this study are quite broad. The mixed, semi-diurnal tidal pattern in the southern Madura coastal region can be used as a basis for determining shipping schedules, designing coastal structures, and managing ports and coastal reclamation. The accuracy of the least squares model implemented using Python programming also opens opportunities for developing a tidal prediction system based on efficient digital data that can be updated in real time.

Further research is recommended to integrate harmonic analysis with two- or three-dimensional numerical hydrodynamic modeling to obtain a more comprehensive spatial representation of tidal dynamics in the Madura region. Furthermore, further studies are needed to examine the influence of meteorological factors such as wind, air pressure, and ocean surface currents on sea level fluctuations. This approach is expected to strengthen scientific understanding of temporal tidal variations and improve the accuracy of tidal prediction systems in Indonesian waters.

5. Conclusions

The results of processing and analyzing tidal data in this study can be concluded that the tidal components obtained are S_0, M_2, S_2, N_2, K_2, K_1, O_1, Q_1, P_1, M_4, MS_4, M_f, and M_m. The highest tidal amplitude component at TADDAN station (south) is M_2 of 0.498 meters, the lowest amplitude is M_4 of 0.004 meters. The results of the tidal type analysis at Taddan station (south) show a Formzahl value of 1.056, so the tide at this location is classified as a mixed tide tending to be semidiurnal. This classification indicates that the twice-daily (semidiurnal) tidal component is still dominant, but there is a significant influence of the daily (diurnal) component so that the water level of the first and second tides in one day can be different.

Acknowledgments

We would like to express our gratitude to Trunodjoyo Madura University, through the Institute for Research and Community Service (LPPM), which has provided funding, support, direction, and facilities so that this research can be carried out well.

References

[1] D. Fitriana, M. P. Patria, And E. Kusratmoko, "Karakteristik Pasang Surut Surabaya Diamati Selama 5 Tahun (2015-2020)," *Jgise J. Geospatial Inf. Sci. Eng.*, Vol. 5, No. 1, Pp. 1–7, 2022.

[2] N. S. Agustin And A. F. Syah, "Analisis Perubahan Garis Pantai Di Pulau Madura Menggunakan Citra Satelit Landsat 8," *Juv. J. Ilm. Kelaut. Dan Perikan.*, Vol. 1, No. 3, Pp. 427–436, 2020.

[3] V. Setia, R. Fadly, And A. Zakaria, "Analisis Perbandingan Hasil Pengolahan Data Pasang Surut Menggunakan Metode Least Square Dan Admiralty Di Pelabuhan Bakauheni Lampung," *Datum J. Geod. Geomatics*, Vol. 4, No. 2, Pp. 32–41, 2024.

[4] I. Cahya And I. Lailiyah, "Analisis Tingkat Sedimentasi Saat Musim Barat Di Kolam Dermaga Koarmatim Surabaya Menggunakan Perbandingan Pengukuran Batimetri," *Juv. J. Ilm. Kelaut. Dan Perikan.*, Vol. 5, No. 3, Pp. 307–312, 2024.

[5] L. M. Prayogo, "Metode Kuadrat Terkecil Untuk Analisis Konstanta Harmonik Pasang Surut Air Laut Di Pulau Gili Raja, Kabupaten Sumenep, Madura," *Pena Akuatika J. Ilm. Perikan. Dan Kelaut.*, Vol. 20, No. 1, 2021.

[6] O. F. Rumihin, "Pengaruh Hidrodinamika Terhadap Letak Alternatif Hard Structure Pantai Bemu, Pulau Seram," *Teras J. J. Tek. Sipil*, Vol. 14, No. 2, Pp. 561–573, 2024.

[7] L. M. Prayogo, "Perbandingan Metode Admiralty Dan Least Square Untuk Analisis Pasang Surut Di Pulau Mandangin Kabupaten Sampang, Jawa Timur (Comparison Of Admiralty And Least Square Methods For Tidal Analysis In Mandangin Island, Sampang Regency, East Java)," *J. Perikan. Dan Kelutan P–Issn*, Vol.

2089, P. 3469, 2021.

[8] A. Milasari, D. H. Ismunarti, E. Indrayanti, F. Muldiyatno, A. Ismanto, And A. Rifai, "Model Arus Permukaan Teluk Lampung Pada Musim Peralihan Ii Dengan Pendekatan Hidrodinamika," *Bul. Oseanografi Mar.*, Vol. 10, No. 3, Pp. 259–268, 2021.

[9] D. Pugh, *Tides, Surges, And Mean Sea-Level*. Chichester; New York: J. Wiley, 1987.

[10] M. G. I. Mardika And M. I. Pratama, "Analisis Pasang Surut Menggunakan Metode Least Square Di Wilayah Perairan Muara Sungai Poso," *J. Infrastruct. Plan. Des.*, Vol. 1, No. 2, Pp. 1–8, 2021.