Analysis and Prediction of Tidal Measurement Data from Temporary Stations using the Least Squares Method

Andi Rusdin 1*, Hideo Oshikawa 2, Andi M. A. Divanesia 3, Muksan P. Hatta 3

1 Department of Civil Engineering, Faculty of Engineering, Tadulako University, Palu 94119, Indonesia.
2 Department of Civil Engineering and Architecture, Saga University, Saga 840-8502, Japan.
3 Department of Civil Engineering, Faculty of Engineering, Hasanuddin University, Gowa 9217, Indonesia.

Received 19 September 2023; Revised 03 January 2024; Accepted 08 January 2024; Published 01 February 2024

Abstract

This research was conducted by equipping three temporary tidal stations located in three places inside Palu Bay with pressure-type tidal gauges. The stations recorded tidal series fluctuations for 4 months with a 5-minute sampling interval (Dt). Moreover, the simple and widely used least squares method (LSM) was applied to separate the harmonic constants of constituents, including amplitudes (Hi) and phases (gi), from the observed tidal series. A total of 11 dominant constituents were selected based on the largest magnitudes of tidal generating potential (CE), and these include M2, K1, S2, O1, P1, N2, Mf, K2, Mm, Q1, and Ms, which are diurnal, semidiurnal, and long-period constituents. The results showed that the semidiurnal constituents generated higher amplitudes than the diurnal constituents, while the long-period constituents produced quite small amplitudes. Furthermore, the ratios of amplitudes recorded showed that tidal in Palu Bay was mainly mixed with semidiurnal constituents. The difference between the observed and predicted values was quite small, and this showed the validity of the measurement conducted at the temporary tidal stations. The performance indicators applied also showed that LSM had acceptable accuracy compared to other methods. Moreover, tidal datums were calculated using the peak approach, and the average tidal range (RA) of Palu Bay was found to be 2.39 m.

Keywords: Palu Bay; Tidal; Least Squares Method; Coastal Engineering.

1. Introduction

Palu Bay is located on the west coast of Sulawesi Island, stretching from south to north between 0°53′9.32″S and 0°38′41.21″S. The eastern boundary of the bay is at 119°52′58.27″E, while the western boundary is at 119°44′19.50″E. Palu City, which is the capital of Central Sulawesi Province, is located at the southern end of the bay, as presented in Figure 1. Palu Bay is long and narrow, with an estimated length of 30.5 km and a width of 7 km [1, 2], with a width of 9.5 km at its broadest cross-section. The east and west coasts have steep bathymetry slopes of 1:2.5 and 1:1.6, respectively. A deep channel appears near the centerline of the bay and deepens towards the mouth, reaching a maximum depth of 850 m [3].

The region is significantly benefiting economically from the bay because it serves as a source of income for the people working as fishermen in coastal communities [4]. The area also has several important permanent ports for general trade as well as semi-permanent ports for loading construction materials from nearby mines. Additionally, there are several tourist spots on the beaches of Palu Bay [5]. However, on Friday, September 28, 2018, at 10:02 UTC or 18:02

*Corresponding author: andi.rusdin@untad.ac.id

http://dx.doi.org/10.28991/CEJ-2024-010-02-03

© 2024 by the authors. Licensee C.E.J, Tehran, Iran. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license [http://creativecommons.org/licenses/by/4.0/].
local time, a shallow strike-slip faulting earthquake of 7.5 Mw occurred near Palu City. This was followed by tsunamis in low-lying areas along Palu Bay, resulting in several losses and over 3,300 casualties [6]. After the disaster, the government and the community strived to revive social and economic activities by providing infrastructure such as coastal protection structures, port reconstructions, repairs to recreational and tourist facilities, and the construction of fishing boat facilities for fishermen. However, the construction of coastal structures and infrastructure requires reference levels of height from the seawater surface, which are known as tidal datums [7, 8]. These tidal datums are also often described as the average height of a specific tidal level.

Figure 1. Location of Palu Bay in Sulawesi Island; (a) Sulawesi Island; (b) Palu Bay and Palu City

Tides are the vertical, alternating rise and fall of the seawater surface caused by the astronomical tidal-generating forces associated with the change in the gravitational effect of the moon and the sun. This often occurs during the process of these elements altering their positions and becoming unbalanced by the centripetal acceleration of the earth in its orbital movements. The tidal oscillation has the same frequency as the astronomical tidal generating potential [9]. However, the intricate nature of both the orbital motion of the moon around the Earth and the Earth around the sun creates several tidal generating potentials with different frequencies and strengths [10]. Tidal generating potential at each frequency is generally represented by tidal harmonic constituents. Meanwhile, the pairs of amplitudes ($H_i$) and phase lags ($g_i$) of the constituents are defined as harmonic constants. This shows that tides can be explained as the combination of the harmonic constants of the constituents.

Tidal analysis is normally used to determine amplitudes and phase lags of dominant constituents from tidal records. Moreover, tidal prediction is usually achieved by recombining harmonic constants of the constituents for a certain period [11]. Tidal datums are also commonly determined by setting the length of the time series data for tidal prediction at 19 years. This specific value is selected because it is the average of the variations in tides caused by the 18.6-year lunar nodal cycle. Furthermore, the two primary harmonic methods generally applied for tidal analysis and prediction are the Fourier Transform Method (FFM) [12] and the Least Squares Method (LSM) [13–16]. FFM can be used to calculate the constants of each tidal constituent separately while LSM determines tidal constants simultaneously [12].

Tidal analysis is normally used to determine the amplitudes and phase lags of dominant constituents from tidal records. Moreover, tidal prediction is usually achieved by recombining the harmonic constants of the constituents for a certain period [11]. Tidal datums are also commonly determined by setting the length of the time series data for tidal prediction at 19 years. This specific value is selected because it is the average of the variations in tides caused by the 18.6-year lunar nodal cycle. Furthermore, the two primary harmonic methods generally applied for tidal analysis and prediction are the Fourier Transform Method (FFM) [12] and the Least Squares Method (LSM) [13–16]. FFM can be used to calculate the constants of each tidal constituent separately, while LSM determines tidal constants simultaneously [12].

LSM has generally been used for tidal harmonic analysis programs since the availability of computers in the 1960s, compared to FFM [12]. This is because the algorithm is quite simple to program on a personal computer and requires a
small amount of memory. The method is exceptionally efficient in obtaining the harmonic constants of constituents from the time series data of tidal observation [19]. The need for quick provision of tidal predictions for coastal engineering constructions, port constructions, and navigation has led to a focus on improving the calculation methods of tidal analysis, which aim to achieve high accuracy by using only a short record of tidal observations [20–23]. However, these methods have some restrictions, such as less precision often experienced when used for long-term tidal predictions [24]. The long-term predictions are crucial for determining the tidal datums, where all harmonic aspects should be included in the calculation. The length of the tidal series to assign the tidal datum is usually not less than 18.6 years, which is one period of the lunar cycle [10]. LSM is confirmed as the most popular modeling approach for long-tidal prediction because of its accuracy advantages and computational efficiency [24, 25].

Previous research has used temporary tidal stations equipped with tidal staff to measure tidal fluctuations inside Palu Bay. For instance, Setiyawan et al. [26] measured tidal series for 15 days with a 1-hour sampling interval (Δt) and applied LSM to analyze the harmonic constants, including amplitudes and phases, to predict tidal fluctuations for 19 years. Tidal datums were obtained by averaging specific stages of tidal elevations from 19 years of extended predicted tidal series. Another study by Sabhan et al. [27] also examined tidal patterns for three days in March 2021, with a sampling interval of 1 hour. The observed tidal series were used to extract the harmonic constants, while tidal datums were subsequently calculated using specific combinations of constituent amplitudes. Both prediction processes were conducted using nine tidal constituents consisting of three diurnal, four semi-diurnal, and two shallow-water constituents, including K1, O1, P1, M2, S2, K2, N2, M4, and MS4. The result showed that no further explanation was required regarding the use of constituents in tidal analysis. These studies found that the type of tide in Palu Bay is mixed mainly semi-diurnal, which agrees with the type of tide in Makassar Strait [28], where Makassar Strait is connected to Palu Bay. Furthermore, the tidal series had considerably short records and was not sufficient to be used in separating the harmonic constants of constituents according to the Rayleigh Criterion [16, 29]. Another factor that needed to be considered was the sampling interval (Δt). Hall & Davies [30] found that the accuracy of harmonic analysis improved significantly when the sample rate was reduced from 60 minutes to 10 minutes. However, obtaining tidal series with a high frequency of sampling interval is impossible to achieve from manual observation using a tidal staff. In the present time, the sampling interval has been decreased to six minutes due to the opportunity to apply an automatic tide gauge instrument [10, 31]. An alternative gauge type that can be used for tidal observation and can capture a high frequency of the sampling interval (Δt) is the pressure transducer-type tidal gauge [32]. The pressure-type tidal gauge is widely used at remote sites because of its simplicity of installation and operation [33, 34].

This background information led to the conduct of this research with the initial focus on the description of tidal time series data observed and recorded in Palu Bay using pressure-type tidal gauges. This was followed by tidal analysis which focused on using LSM to disentangle harmonic constants of constituents. Similar research on tidal analysis and prediction using observed tidal series has been conducted by Madah [35] using observed tidal series from two tidal stations in the Gulf of Aden. The harmonic constants found using LSM were applied to generate the predicted 19 years of tidal series data which was slightly longer than the 18.6 year lunar nodal cycle. Finally, the predicted data were used to compute tidal datums.

2. Research Methodology

Tidal analysis and prediction were conducted through four main stages of activities. At the initial stage, temporary tidal stations were set up in selected locations and equipped with pressure-type gauges to automatically record tidal fluctuations. The next stage was to conduct tidal analysis by disentangling the tidal series observations obtained from the stations using LSM to obtain amplitudes and phases of specific constituents. This was followed by the application of amplitudes and phases to the tidal harmonic equation to generate 19 years of tidal series prediction data. Finally, the predicted data were used to estimate tidal datums and tidal range (RA) as shown in Figure 2.

![Figure 2. The stages of tidal analysis and predictions](image)
2.1. Locations of Tidal Observations

Secondary [36] or temporary tidal stations [37], defined as stations operated for less than 1 year, were established in three locations inside Palu Bay to acquire local tidal information by recording the time series data of sea level fluctuation. The selection of these locations was mainly based on the absence of disturbances to avoid poor accuracy of tidal records. Moreover, the common criteria usually used to set up tidal gauges are ease of installation, operation, and maintenance [11], as well as some other detailed constraints [31]. Some of these include the need for the site to protect against the impact of waves and currents as well as the capacity to withstand the worst climate conditions. It is also important to have sufficient water depth during low tides to ensure the gauge does not dry out. Impounding areas with the ability to isolate the gauge from the open sea during low tides should be avoided. The same operation is also required to be applied to sandbars just below the surface between the location and the open sea, which can cause the observation of unusual values. Sharp headlands, which can lead to abnormal tidal currents, need to be avoided, and the location should be kept away from the possibility of marine traffic disturbance or damage. Furthermore, adequate access should be made for the installation and maintenance, as well as safety against vandalism or theft. The existence of structures, such as wharves, jetties, fish stakes, or bridge pilings, can also provide stable and firm support for gauge installation. A sturdy and stable structure can produce relatively precise tidal data series for a long period, which is suitable to be used for defining the tidal datum [37].

This research was conducted using three tidal gauges to collect tidal time series data at three different locations. The gauges were placed and arranged to represent the inlet, middle, and end sections of the bay, as presented in Figure 3. One was installed near the bay inlet at the jetty of the general cargo port of Donggala Port (DP) in Donggala City. The second was set up on the jetty of Naval Headquarters, situated in Watusampu (WS), located relatively in the middle of the bay. These two locations have relatively deep port basins considered suitable for setting up the gauges and have good access roads. Meanwhile, the third tidal gauge was placed in the sea area near the end of Palu Bay and projected to the waterfront area of Palu City. The sea area is a gently sloping beach, and this means the gauge was put offshore northward to keep sufficient water depth while remaining submerged during low tide. The existing bamboo fishing platform located offshore of Taman Ria (TR) Beach and owned by a fisherman was used as the installation point due to the relatively deep water and quite far distance from the estuary of the Palu River to avoid significant fluctuations in the seawater density. All three gauges were set up on the west side of Palu Bay, with calmer sea surfaces compared to the east side, characterized by the consistent impact of waves generated by winds blowing from the northwest [38].

![Figure 3. The Map of Palu Bay and temporary stations; (a) locations of the temporary stations; (b) Structures used as tidal stations; (c) Three pressure-type tidal gauges used simultaneously](image-url)

2.2. Tidal Record Devices

Tidal can be measured using manual tabulation of the data recorded by tidal staff [39] or through the application of automatic recording devices such as stilling well and float, pressure, acoustic, and radar systems [29]. The pressure type
of water deep loggers was used to record the sea level fluctuation at Palu Bay. The gauge was observed to be pocket-sized, and all instrument circuits, batteries, and memories were integrated into a titanium body as presented in Figure 3. Moreover, electrical power was obtained from a small battery, and a flash memory with the capacity to store data for one year at a one-minute sampling interval was employed to record the data. The accuracies of the gauges used in this research were less than 1% of the full scale of measurements (1% F.S.). A standard of less than 1.0 cm of accuracy or resolution is required for tidal poles and automatic tidal gauges by the Indonesian National Standard (SNI). The Intergovernmental Oceanographic Commission [31] also requires a maximum reading accuracy of 1.0 cm of tidal gauge. However, the small resolution of the reading accuracy of a device to log tidal data is not a strict requirement. A tidal reading accuracy of 1.0 cm is difficult to achieve at temporary stations unless the measurement is made under very sheltered sea surface conditions. The reading accuracies of tidal gauges installed at DP, WS, and TR were 0.5 cm, 1.0 cm, and 5.0 cm, respectively. The complete specification of tidal gauges is presented in the following Table 1.

Table 1. Descriptions and specifications of tidal gauges

<table>
<thead>
<tr>
<th>Times</th>
<th>Locations</th>
<th>Donggala Port (DP)</th>
<th>Watusampu (WS)</th>
<th>Taman Ria (TR)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Location (UTN)</td>
<td>E 805544.00 m</td>
<td>E 812975.00 m</td>
<td>E 816563.00 m</td>
</tr>
<tr>
<td></td>
<td>Type of tidal gauge</td>
<td>Pressure logger</td>
<td>Pressure logger</td>
<td>Pressure logger</td>
</tr>
<tr>
<td>Range (m)</td>
<td>0 – 50</td>
<td>0 – 100</td>
<td>0 – 200</td>
<td></td>
</tr>
<tr>
<td>Resolution (m)</td>
<td>0.005</td>
<td>0.01</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>Accuracy</td>
<td>±0.3% F.S.</td>
<td>±1.0% F.S.</td>
<td>±1.0% F.S.</td>
<td></td>
</tr>
<tr>
<td>Start time (UTC)</td>
<td>17/05/2014 04:00</td>
<td>17/05/2014 04:00</td>
<td>17/05/2014 04:00</td>
<td></td>
</tr>
<tr>
<td>Finish time (UTC)</td>
<td>12/09/2014 04:00</td>
<td>12/09/2014 04:00</td>
<td>12/09/2014 04:00</td>
<td></td>
</tr>
<tr>
<td>Record length (days)</td>
<td>118</td>
<td>118</td>
<td>118</td>
<td></td>
</tr>
<tr>
<td>Record length (hours)</td>
<td>2833</td>
<td>2833</td>
<td>2833</td>
<td></td>
</tr>
<tr>
<td>Sampling interval (minutes)</td>
<td>5.0</td>
<td>5.0</td>
<td>5.0</td>
<td></td>
</tr>
<tr>
<td>Number of records</td>
<td>33985</td>
<td>33985</td>
<td>33985</td>
<td></td>
</tr>
</tbody>
</table>

The gauges were tied and hung using ropes submerged beneath the seawater under the jetty and fish platform. Concrete cube ballasts with dimensions of 15×15×15 cm³ and a weight of 7.425 kg were attached below the gauge at the end of the ropes to ensure stability. Moreover, the points where the gauges were hung were also considered to prevent disturbances from waves, as presented in Figure 4. At DP, the tidal gauge was hung just in the middle and below the piles-type jetty. Meanwhile, the gauge was placed on the lee side of the Naval Jetty in WS and the southern part of the bamboo fishing platform in TR because the point was on the lee side of the platform and sheltered from waves that mainly came from the north.

![Initial](image1.png)
![Installation](image2.png)
![After](image3.png)

**Figure 4. Installations of tidal gauges; (a) Donggala Port, DP; (b) Watusampu, WS (c.); Taman Ria, TR**
2.3. Record Length and Sampling Interval

The minimum length of tidal record \( T \) required to split the harmonic constants in the main tidal constituents is 29 days [11]. The Indonesian National Standard [39] also demands a minimum of 29 days of tidal record observation and a maximum of 1 hour of sampling interval to analyze tidal constituents. This generally means that 1 hour of sampling interval is appropriate because it provides the opportunity to study all tidal constituents. A shorter interval can be used when shorter-term variations in sea level are needed for tidal analysis, but the maximum possible is 1 hour [31]. The acceptable interval needs to be selected in advance because it is impossible to resample data with a more frequent interval.

Observations were made inside Palu Bay from May 17, 2014, at 04:00:00 to September 12, 2014, at 04:00:00 UTC. This showed that the record length \( T \) for tidal data was 118 days, and this was four times longer than the minimum of 29 days required for a temporal station [39]. Moreover, the pressure-type digital loggers could be set up with at least 1 second of sampling interval for 9.6 days of \( T \) and 1 minute of sampling interval for 355 days of \( T \). In this research, all the gauges were set up at a 5-minute sampling interval \((\Delta t = 5.0 \, \text{minutes}, \text{or} 1/12 \, \text{hours})\), which was lower than the value suggested by the Intergovernmental Oceanographic Commission (IOC) [29].

2.4. Least Squares Method

Harmonic analysis is defined as a mathematical operation to separate observed tidal data series from any location into harmonic constants of tidal constituents. These constants are normally used to reconstruct the predicted tidal time series. Suppose the data series of tidal height \( h(t) \) was assembled by \( K \) tidal harmonic constants of tidal constituents. The equation of tidal height fluctuation [13, 14, 29, 40–44] could be written as follows:

\[
h(t) = h_o + \sum_{i=1}^{K} H_i \cos(\omega_i t - g_i)
\]

(1)

where, \( h(t) \) is sea water elevation (m), \( h_o \) is mean sea level (m) or the mean value of tidal height series, \( H_i \) is constituent amplitude (m), \( \omega_i \) is constituent angular velocity (°/hour), \( g_i \) is phase (°), and \( t \) is time (hour). The reduction of the tidal height series to obtain the mean value equal to zero shows that \( h_o \) needs to be excluded from Equation 1 to produce the following harmonic equation [19, 42]:

\[
h(t) = \sum_{i=1}^{K} H_i \cos(\omega_i t - g_i)
\]

(2)

The harmonic formulation for tidal fluctuation in Equation 2 can be re-expressed using a trigonometric identity. The purpose is to ensure it becomes a comparable equation with the sum of harmonic cosines and sinus terms as follows:

\[
h(t) = \sum_{i=1}^{K} A_i \cos(\omega_i t) + \sum_{i=1}^{K} B_i \sin(\omega_i t)
\]

(3)

The linear fitting equation has two unknown parameters which are \( A_i = H_i \cos(g_i) \) and \( B_i = H_i \sin(g_i) \), where:

\[
H_i = \sqrt{A_i^2 + B_i^2}
\]

(4)

\[
g_i = \tan^{-1}\left(\frac{B_i}{A_i}\right)
\]

(5)

The purpose of LSM is to minimize the total quadratic aspect of the residual time series, \( E \), between the predicted values of seawater elevation, \( h(t_m) \), in Equation 3, and observed seawater elevation data, \( h_m \) as indicated in the following equation:

\[
E = \sum_{m=1}^{N}[h(t_m) - h_m]^2
\]

(6)

where, \( m \) is the index of the time record and \( N \) is the record length. The substitution of Equation 3 into Equation 6 changes the formulation of \( E \) to the following:

\[
E = \sum_{m=1}^{N}[\sum_{i=1}^{K} A_i \cos(\omega_i t_m) + \sum_{i=1}^{K} B_i \sin(\omega_i t_m) - h_m]^2
\]

(7)

The magnitude of \( E \) is minimum when the partial derivatives of the residual Equation 6 are applied with respect to the unknown parameters \( A_i \) and \( B_i \) and define the derivatives as equal to zero.

\[
\frac{\partial E}{\partial A_j} = \frac{\partial E}{\partial B_j} = 0 \quad j = 1, 2, \ldots, K
\]

(8)

where, \( j \) is an index notation indicating \( E \) is derived with respect to the order of constituents. The application of Equation 8 to Equation 7 can be used to obtain the following fitting equations:
\[
\frac{\partial E}{\partial A_i} = \left\{ \sum_{m=1}^{N} [A_i \cos(\omega_i t_m) + \sum_{i=1}^{K} B_i \sin(\omega_j t_m) - h_m] \cos(\omega_i t_m) \right\} = 0
\] (9)

\[
\frac{\partial E}{\partial B_i} = \left\{ \sum_{m=1}^{N} [A_i \cos(\omega_i t_m) + \sum_{i=1}^{K} B_i \sin(\omega_j t_m) - h_m] \sin(\omega_i t_m) \right\} = 0
\] (10)

Rearranging Equation 9 and Equation 10, the optimum tidal series predictions equations become:

\[
\left\{ \sum_{m=1}^{N} [A_i \cos(\omega_i t_m) \cos(\omega_j t) + \sum_{i=1}^{K} B_i \sin(\omega_j t_m) \cos(\omega_j t_m)] \right\} = \sum_{m=1}^{N} h_m \cos(\omega_j t_m)
\] (11)

\[
\left\{ \sum_{m=1}^{N} [A_i \cos(\omega_i t_m) \sin(\omega_j t_m) + \sum_{i=1}^{K} B_i \sin(\omega_i t_m) \sin(\omega_j t_m)] \right\} = \sum_{m=1}^{N} h_m \sin(\omega_j t_m)
\] (12)

Equations 11 and 12 can be written as \(2K\) linear equations and the number of parameters \(A_i\) and \(B_i\) is \(K\) respectively in each equation, where \(K\) is the number of constituents.

\[
A_1CC_{11} + A_2CC_{12} + \cdots + A_KCC_{1K} + B_1SC_{11} + B_2SC_{12} + \cdots + B_KSC_{1K} = hC_1
\]

\[
A_1CC_{21} + A_2CC_{22} + \cdots + A_KCC_{2K} + B_1SC_{21} + B_2SC_{22} + \cdots + B_KSC_{2K} = hC_2
\]

\[\vdots\]

\[
A_1CC_{K1} + A_2CC_{K2} + \cdots + A_KCC_{KK} + B_1SC_{K1} + B_2SC_{K2} + \cdots + B_KSC_{KK} = hC_K
\]

\[
A_1CS_{11} + A_2CS_{12} + \cdots + A_KCS_{1K} + B_1SS_{11} + B_2SS_{12} + \cdots + B_KSS_{1K} = hS_1
\]

\[
A_1CS_{21} + A_2CS_{22} + \cdots + A_KCS_{2K} + B_1SS_{21} + B_2SS_{22} + \cdots + B_KSS_{2K} = hS_2
\]

\[\vdots\]

\[
A_1CS_{K1} + A_2CS_{K2} + \cdots + A_KCS_{KK} + B_1SS_{K1} + B_2SS_{K2} + \cdots + B_KSS_{KK} = hS_K
\]

Equation 13 can also be reconstructed in the following matrix form:

\[
D Z = Y
\] (14)

where, \(D\) is a \((K \times K)\) matrix consisting of the summation of sine and cosine terms, \(Z\) is a \((K \times 1)\) matrix containing unknown parameters \(A_i\) and \(B_i\), and \(Y\) is a \((K \times 1)\) matrix composing the summation of the observed tidal height data series.

\[
D = \begin{bmatrix}
CC_{11} & CC_{12} & \cdots & CC_{1K} & SC_{11} & SC_{12} & \cdots & SC_{1K} \\
CC_{21} & CC_{22} & \cdots & CC_{2K} & SC_{21} & SC_{22} & \cdots & SC_{2K} \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
CC_{K1} & CC_{K2} & \cdots & CC_{KK} & SC_{K1} & SC_{K2} & \cdots & SC_{KK} \\
SC_{S1} & CS_{S1} & \cdots & CS_{SK} & SS_{S1} & SS_{S2} & \cdots & SS_{SK} \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
SC_{SK1} & CS_{SK2} & \cdots & CS_{KK} & SS_{SK1} & SS_{SK2} & \cdots & SS_{KK}
\end{bmatrix};
Z = \begin{bmatrix}
A_1 \\
A_2 \\
\vdots \\
A_K \\
B_1 \\
B_2 \\
\vdots \\
B_K
\end{bmatrix};
Y = \begin{bmatrix}
hC_1 \\
hC_2 \\
\vdots \\
hC_K \\
hS_1 \\
hS_2 \\
\vdots \\
hS_K
\end{bmatrix}
\] (15)

The elements of matrix \(D\) and \(Y\) are described as follows:

\[
CC_{ij} = \sum_{m=1}^{N} \cos(\omega_i t_m) \cos(\omega_j t_m)
\]

\[
CS_{ij} = \sum_{m=1}^{N} \cos(\omega_i t_m) \sin(\omega_j t_m)
\]

\[
SS_{ij} = \sum_{m=1}^{N} \sin(\omega_i t_m) \sin(\omega_j t_m)
\]

\[
SC_{ij} = \sum_{m=1}^{N} \sin(\omega_i t_m) \cos(\omega_j t_m)
\]

\[
hC_i = \sum_{m=1}^{N} h_m \cos(\omega_j t_m)
\]

\[
hS_i = \sum_{m=1}^{N} h_m \sin(\omega_j t_m)
\]

where, \(i\) and \(j\) are indexes of the constituents having a maximum value equal to the number of constituents, \(K\), applied in the calculation. Moreover, the unknown parameters \(A_i\) and \(B_i\) in matrix \(Z\) can be obtained by calculating the inverse of matrix \(D\), and the elements of matrix \(Z\) can be determined using:

\[
Z = D^{-1}Y
\] (22)

This matrix was computed using the FORTRAN computer program and the inverse of matrix \(D\) was solved through Gauss-Jordan Elimination.
2.5. Dominant Tidal Constituents

Doodson & Lamb [45] were reported to have identified 389 tidal harmonic constituents, as shown by Chelton and Endfield [46]. However, the National Ocean Service Center for Operational Oceanographic Products and Services (NOS CO-OPS), USA, listed 37 as the standard constituents to be used for tidal analysis [10]. Additional constituents can be necessary when the observation site is in a shallow waterway where nonlinear effects are usually larger. For example, Zetler and Cummings [47] used 114 constituents to analyze tides at Anchorage, Alaska, and the same number was used by Rossiter & Lennon [48] to calculate tides in the Thames Estuary, United Kingdom. Hanxing [49] applied 128 constituents for tidal prediction analysis in Shanghai Port, China, while Foreman [16] utilized 45 astronomical and 24 shallow water constituents for tidal harmonic analysis. Moreover, Boon III and Kiley [19] used 37 harmonic constituents in a computer program for tidal analysis and prediction through the application of LSM.

Some of the constituents included in tidal harmonic analysis have considerably small amplitudes. Those with very small amplitudes are considered non-tidal continuum or unreal harmonic constants, which can be rejected in the analysis [10]. Moreover, the computational time for analysis usually experiences a roughly proportional increase to the square of the number of constituents included in the calculation [16]. Godin & Taylor [50] applied a computer program for tidal prediction and discovered that the time required for computation increased when the number of tidal constituents included in the analysis was increased. Therefore, most of the shallow water constituents in several tidal stations, specifically those located in the open sea, are unimportant [16]. This is because the constituents with insignificant amplitudes can be neglected in the analysis [46].

The selection of harmonic constituents for tidal analysis and prediction is based on two factors. The first factor is the contribution to tidal fluctuation based on tidal generating potential (\(CE\)), which was calculated by Cartwright and Edden [9]. The second factor is the sufficient record length (\(T\)) of the tidal series needed to separate the harmonic constants of neighboring constituents. Additionally, the time needed to resolve the effects of two neighboring tidal constituents is known as the synodic period [12], and it is characterized by the time difference between two successive period connectives, \(1/f\), of the two constituents. The minimum length of tidal record needed to separate two constituents is known as the Rayleigh Criterion [16, 29], and it can be expressed using the following mathematical relationship:

\[
|f_m - f_i|T_R \geq RC
\]  

(23)

where, \(f_m\) is the frequency of a specific constituent to be included in the tidal analysis, \(f_i\) is the frequency of the comparison neighboring constituent in the similar frequency band, \(T_R\) is the minimum record length needed to separate harmonic constants, and \(RC\) is the Rayleigh Criterion Value with the magnitude usually specified to be equal to one.

Table 2 shows 11 dominant constituents selected based on the magnitudes of tidal generating potential coefficients (\(CE\)) [9]. The table also provides the periods and angular speeds of constituents. \(M_2\) generally has the largest contribution to the equilibrium tidal based on \(CE\), followed successively by \(S_2\), \(N_2\), and \(K_2\) among the semidiurnal tidal constituents, while \(K_1\) has the highest value of \(CE\) for diurnal constituents, followed by \(O_1\), \(P_1\), and \(Q_1\). The lunar-solar fortnightly \(M_4\) and \(M_6\) were also included as long-period constituents. The last constituent used for the tidal harmonic constants is the lunar monthly constituent, \(M_m\). The period (\(1/f\)) of the principal semidiurnal constituent, \(M_2\), is approximately 12.4206 hours, or half a lunar day, and its harmonic constant can be determined with a \(T_R\) of at least 13 hours. The principal diurnal constituent \(K_1\) requires at least 24 hours for tidal analysis because its period is 23.9345 hours. The principal long-period constituent of the lunar-solar fortnightly \(M_d\) has a period of 354.3671 hours and requires 355 hours for tidal analysis.

The Rayleigh Criterion showed that the separation of the harmonic constants of the next primary semidiurnal constituent \(S_2\) from those of \(M_2\) required a minimum \(T_R\) of 355 hours. The frequencies of \(M_2\) and \(S_2\) serve as the Rayleigh comparison frequencies (\(f_i\)) when evaluating the frequencies (\(f_m\)) of other constituents in the semidiurnal band. To distinguish \(N_2\) from \(M_2\) and \(K_2\) from \(S_2\), a minimum tidal record duration of 662 hours and 4383 hours is required, respectively. For the diurnal band, the frequency of \(K_1\) functions as \(f_i\) for \(f_m\) of \(O_1\), and the minimum \(T_R\) needed for tidal separation analysis between \(K_1\) and \(O_1\) is 328 hours. Moreover, the application of \(K_1\) and \(O_1\) frequencies as \(f_i\) for two other diurnal band frequencies (\(f_m\)) of \(P_1\) and \(Q_1\) requires a minimum \(T_R\) of 4383 hours and 662 hours, respectively, to successfully differentiate \(P_1\) from \(K_1\) and \(Q_1\) from \(O_1\). \(f_1\) of \(M_d\) and \(f_m\) of \(M_m\) yield a 764-hour tidal record to separate the harmonic constants of \(M_m\) from \(M_d\). The adherence to the Rayleigh Criterion also showed that a minimum tidal series record of 4383 hours is essential to separate the components of constituent \(M_2\) from \(M_d\). The record length (\(T\)) of the tidal data series acquired from the temporary tidal stations is 118 days, or 2832 hours, less than the 4383 hours of \(T_R\) required to analyze the harmonic constants of \(P_1\), \(K_2\), and \(M_2\). However, the \(T\) of the tidal series can be shorter than the time length suggested by the Rayleigh Criterion when the harmonic constants are analyzed using LSM [10, 29, 51].
Another criterion considered in selecting the constituents is $\Delta t$ [10, 29, 51]. This is because the maximum frequency to analyze harmonic constants depends on the sampling interval and is known as the Nyquist frequency or folding frequency ($f_n$), which is presented in the following equation:

$$f_n = \frac{1}{2\Delta t} \quad (24)$$

The Nyquist criterion states that the harmonic constants of a constituent cannot be resolved in the harmonic analysis when the period is smaller than twice the sampling interval. This means a sinusoidal-type oscillation can only be resolved by using a minimum of two sampling intervals or three data points. Since $\Delta t = 5$ minutes or 1/12 hours, the highest frequency that can be calculated is 6.0 hours$^{-1}$ ($f_n = 6.0$), and the minimum period to be included in tidal analysis is 1/6 or 0.167 hours (10 minutes). This shows that all 11 constituents fulfill this requirement because the period recorded for each is greater than 0.167 hours. For example, $K_2$ has 11.967 hours, which is the smallest, thereby indicating the sampling interval is small enough to resolve tidal analysis. However, the Nyquist criterion is not a strict requirement [10, 51] because tidal observations are made inside a deep bay and not in shallow water areas such as a river or a waterway, which sometimes have a high-frequency sampling rate.

2.6. Type of Tide

The type of tidal can be classified based on tidal form factors or form numbers ($F$) calculated using amplitudes of four dominant constituents, including lunisolar diurnal $K_1$, principal lunar diurnal $O_1$, principal lunar semidiurnal $M_2$, and principal solar semidiurnal $S_2$ [52, 53]. The term was first introduced by Van der Stok [54] with three tidal types, which were further increased to four by Courtier [55]. A similar method has been used in some specific regions by Parker [56], Amin [57], Daher et al. [58], Godin & Taylor [50], and Lee & Chang [59]. The classification was conducted based on the ratio between amplitudes of diurnal and semidiurnal, as shown in the following equation:

$$F = \frac{H_{K_1} + H_{O_1}}{H_{M_2} + H_{S_2}} \quad (25)$$

where, $H_{K_1}$ is the amplitude of $K_1$, $H_{O_1}$ is for $O_1$, $H_{M_2}$ is for $M_2$, and $H_{S_2}$ is for $S_2$. Tidal types are divided into four categories including $F < 0.25$ for semidiurnal, $0.25 \leq F < 1.50$ for mixed mainly semidiurnal, $1.50 \leq F < 3.00$ for mixed mainly diurnal, and $F \leq 3.00$ for diurnal.

3. Results

3.1. Time Series of Sea Level Fluctuations

The time series data of sea level fluctuation obtained from three observation stations are presented in Figure 5 after the mean values have been subtracted to produce zero median values. The variations identified in the data showed the continuous coverage of observations to present the spring and neap patterns of tidal. Moreover, the phase data [60] associated with the orbit revolution of the moon encircling the earth relative to the sun were obtained from Scientific Visualization Studio, Goddard Space Flight Center, and the National Aeronautics and Space Administration (SVS, GVFC, NASA), covering between May and September 2014, are also included in the figure. The new moon is represented by 0.0%, while the full moon is 100%. It is indicated that the spring tide occurred during the new and full moon, while the neap tide occurred during the first and third quarters of the moon. Moreover, a completed revolution of

<table>
<thead>
<tr>
<th>No.</th>
<th>Symbol</th>
<th>Type</th>
<th>Period (hour)</th>
<th>Angular speed (°/hour)</th>
<th>Tide generating potential, $CE_i$</th>
<th>Required minimum record length, $T_s$ (hours)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>$K_1$</td>
<td>Diurnal</td>
<td>23.93447213</td>
<td>15.0410686</td>
<td>0.53011</td>
<td>24</td>
</tr>
<tr>
<td>2.</td>
<td>$O_1$</td>
<td>Diurnal</td>
<td>25.81933871</td>
<td>13.9430356</td>
<td>0.37694</td>
<td>328</td>
</tr>
<tr>
<td>3.</td>
<td>$P_1$</td>
<td>Diurnal</td>
<td>24.06588766</td>
<td>14.9589314</td>
<td>0.17543</td>
<td>4383</td>
</tr>
<tr>
<td>4.</td>
<td>$Q_1$</td>
<td>Diurnal</td>
<td>26.36835000</td>
<td>13.3986609</td>
<td>0.07217</td>
<td>662</td>
</tr>
<tr>
<td>5.</td>
<td>$M_2$</td>
<td>Semidiurnal</td>
<td>12.42060120</td>
<td>28.9841042</td>
<td>0.90809</td>
<td>13</td>
</tr>
<tr>
<td>6.</td>
<td>$S_2$</td>
<td>Semidiurnal</td>
<td>12.00000000</td>
<td>30.00000000</td>
<td>0.42248</td>
<td>355</td>
</tr>
<tr>
<td>7.</td>
<td>$N_2$</td>
<td>Semidiurnal</td>
<td>12.65834751</td>
<td>28.4397295</td>
<td>0.17386</td>
<td>662</td>
</tr>
<tr>
<td>8.</td>
<td>$K_2$</td>
<td>Semidiurnal</td>
<td>11.96723606</td>
<td>30.0821373</td>
<td>0.11498</td>
<td>4383</td>
</tr>
<tr>
<td>9.</td>
<td>$M_4$</td>
<td>Fortnightly</td>
<td>327.8599387</td>
<td>1.0980331</td>
<td>0.15647</td>
<td>4383</td>
</tr>
<tr>
<td>10.</td>
<td>$M_6$</td>
<td>Fortnightly</td>
<td>354.3670666</td>
<td>1.0158958</td>
<td>0.01369</td>
<td>355</td>
</tr>
<tr>
<td>11.</td>
<td>$M_n$</td>
<td>Monthly</td>
<td>661.3111655</td>
<td>0.5443747</td>
<td>0.08254</td>
<td>764</td>
</tr>
</tbody>
</table>
the orbit period of the moon encircling the earth relative to the sun, also known as the synodic month or one lunation, was estimated at 29.5306 days [61]. Spring and neap tides occur every 14.7653 days, or a half of a synodic month, and neap tide occurs 7.3827 days after spring tide. The results showed that the three stations had fairly similar tidal patterns, and this could be because the distance between DP and WS is 18.69 km and between WS and TR is 7.55 km. The wavelength of the tide could be thousands of kilometers [61, 62], however, the stations were separated within 26.24 km, which is quite small compared to the wavelength of the tide, and the discrepancies in tide fluctuations among stations in Palu Bay are significantly small.

![Figure 5. Observed tidal series; (a) Donggala Port, DP; (b) Watusampu, WS; (c) Taman Ria, TR](image)

High Water Level (HWL) and Low Water Level (LWL) magnitudes of the observed seawater surface elevations are presented in Table 3. It was discovered that the highest tidal fluctuations for DP, WS, and TR were 1.29 m, 1.34 m, and 1.33 m, respectively, while the lowest were -1.12 m, -1.22 m, and -1.22 m. This showed that the highest and lowest water surface values of WS and TR were quite similar, but DP had slightly smaller values. The difference between the HWL of DP and the two other stations was found to be 5.0 cm, while the LWL was 10.0 cm. Moreover, tidal ranges (RA) based on the difference between HWL and LWL were recorded to be 2.41 m, 2.56 m, and 2.55 m for DP, WS, and TR, respectively. These values were discovered to be quite similar, and the average value was recorded to be 2.51 m.

<table>
<thead>
<tr>
<th>Description</th>
<th>Maximum (HWL)</th>
<th>Minimum (LWL)</th>
<th>Tidal Range, RA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Donggala Port (DP)</td>
<td>1.29</td>
<td>-1.12</td>
<td>2.41</td>
</tr>
<tr>
<td>Watusampu (WS)</td>
<td>1.34</td>
<td>-1.22</td>
<td>2.56</td>
</tr>
<tr>
<td>Taman Ria (TR)</td>
<td>1.33</td>
<td>-1.22</td>
<td>2.55</td>
</tr>
<tr>
<td>Average</td>
<td>1.32</td>
<td>-1.19</td>
<td>2.51</td>
</tr>
</tbody>
</table>

### 3.2. Amplitudes and Phases of Constituents

The harmonic constant magnitudes of tidal constituents consisting of amplitudes are presented in Figure 6 in the form of bar charts. The amplitudes were plotted sequentially according to the types of constituents, including diurnal and semidiurnal, as well as those for a long period, such as fortnightly and monthly constituents. The results showed that the magnitudes of amplitudes for each tidal constituent from the three observation locations tended to be similar, with only small discrepancies. The largest was recorded for $M_2$, followed by $S_2$, $K_1$, $O_1$, $N_2$, $K_2$, $P_1$, $Q_1$, $M_n$, $M_i$, and $M_{st}$. This showed that the semidiurnal constituents had a more significant contribution to tidal fluctuation than the diurnal constituents, while long-period constituents provided considerably less effect. The lunar monthly constituent, $M_m$, had the third smallest amplitude, while the second-smallest and the smallest were recorded for $M_i$ and $M_{st}$ which were both lunar-solar fortnightly constituents. The magnitude of amplitudes for the semidiurnal and diurnal constituents...
constituents ranged from $M_2$, $S_2$, $K_1$, $O_1$, $K_2$, $P_1$, to $N_2$ and this trend was similar to the results of Sabhan et al. [27] and Setiyawan et al. [26]. However, the researchers did not include $Q_1$, a diurnal constituent, in the calculation. Amplitudes calculated for $Q_1$ in DP, WS, and TR were 2.07 cm, 2.14 cm, and 2.21 cm, respectively, and these were observed to have some effects on tidal fluctuation, leading to the inclusion of the constituent in tidal analysis. Meanwhile, shallow water constituents, $M_4$ and $M_{S4}$, had extremely small amplitudes, which were less than 0.3 mm [26, 27], showing lesser influence on tidal level variations. This showed that $M_4$ and $M_{S4}$ were insignificant in Palu Bay and could be neglected in tidal analysis.

![Figure 6. Amplitudes of tidal constituents](image)

$M_2$ was the most significant constituent based on both tidal generating potential and amplitudes. However, there were variations in determining the sequence of significance of the other constituents based on these factors. This led to the introduction of Magnitude Ratio ($MR$), a nondimensional value, to compare the significance of the constituents based on tidal generating potential [29] and amplitudes. $MR$ was computed using two methods and the first was $MR_{Ci}$ which was achieved by normalizing tidal generating potential values of constituents with the values of tidal generating potential of $M_2$ ($CE_{M2}$) as follows:

$$MR_{Ci} = \frac{CE_i}{CE_{M2}}$$  \hspace{1cm} (26)

$MR_{Ci}$ values are calculated based on tidal generating potential. $MR_{Ci}$ values are found to be similar to those mentioned in the Canadian Tidal Manual [11]. The second method is $MR_{Hi}$ which is specified as the ratio between the amplitudes of constituents and the amplitudes of $M_2$ ($HM_2$) in the following equation:

$$MR_{Hi} = \frac{H_i}{HM_2}$$  \hspace{1cm} (27)

Based on $MR$ equations, the $MR_{Hi}$ and $MR_{Ci}$ values of $M_2$ are equal to 1. The $MR_{Hi}$ of three tidal stations and $MR_{Ci}$ of tidal generating potential are presented in Figure 7. Most $MR_{Hi}$ values acquired from amplitudes of semidiurnal constituents were found to be larger than the $MR_{Ci}$ calculated using tidal generating potential values. This shows that the tidal generating potential of semidiurnal constituents inside Palu Bay is larger than the potential of other constituents and contributes more to distance variations of tidal fluctuations. $M_2$ is identified as the most dominant constituent based on tidal generating potential values and amplitudes. Similarly, $S_2$ contributes more to tidal fluctuation because $MR_{Hi}$ is 37% larger than $MR_{Ci}$. $K_2$ also shows the same trend with a 17% difference and only $N_2$ had a slightly smaller value of 15%. The $MR$ magnitudes of diurnal constituents show distinct patterns, where generally $MR_{Hi}$ magnitudes of $K_1$, $O_1$, $P_1$, and $Q_1$ are significantly smaller than $MR_{Ci}$ magnitudes of the diurnal constituents about 29%, 58%, 27%, and 51%, respectively. The tidal analysis indicated that diurnal constituents contribute less to distance variations of tidal fluctuations than the contributions expected from the order of $CE_i$ values. For the long-period constituents, $MR_{Ci}$ values have extreme discrepancies compared to $MR_{Hi}$. $MR_{Hi}$ values of $M_f$, $M_{sf}$, and $M_m$ were smaller than their $MR_{Ci}$ about 95%, 74%, and 76%, respectively.
The tidal generating potential \(CE_i\) introduced by Cartwright & Edden [9] was merely computed from the tidal generating force, which is produced by the gravitational attractions of the moon, sun, and the Earth based on the ephemeris of the moon and sun. This force is combined with the centrifugal force of the rotational movement of the Earth and the moon system. The magnitudes of the tidal generating forces are different at every location on the Earth because of their position relative to the moon and sun. Furthermore, the differences in amplitudes of constituents at every location are tidal propagations influenced by the Coriolis force of the Earth, frictions with the ocean or sea floors as well as resonances caused by the depths and shapes of the ocean basins, seas, bays, straits, and other marginal seas. Additionally, amplitudes are also affected by hydrodynamic and meteorological phenomena such as river discharge and storm surge.

The phases of constituents \(g_i\) for the three observation locations are presented in Figure 8 while the mean values \(\bar{g}_i\) and standard deviations \(\sigma_i\) are listed in Table 4. The fortnightly constituents, \(M_f\) and \(M_{sf}\), were observed to have significantly different phase values from the stations based on the standard deviations. This is possible because the local response of Palu Bay can alter the phase of constituents as tidal water enters or exits the bay [63].
Table 4. Tidal constituents included in tidal analysis and prediction

<table>
<thead>
<tr>
<th>No.</th>
<th>Symbol</th>
<th>$\bar{g}_i$ (degrees)</th>
<th>$\sigma$ (degrees)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>K_1</td>
<td>299.93</td>
<td>±4.61</td>
</tr>
<tr>
<td>2</td>
<td>O_1</td>
<td>304.61</td>
<td>±1.38</td>
</tr>
<tr>
<td>3</td>
<td>P_1</td>
<td>216.56</td>
<td>±7.17</td>
</tr>
<tr>
<td>4</td>
<td>Q_1</td>
<td>273.86</td>
<td>±6.08</td>
</tr>
<tr>
<td>5</td>
<td>M_2</td>
<td>261.05</td>
<td>±0.46</td>
</tr>
<tr>
<td>6</td>
<td>S_2</td>
<td>237.08</td>
<td>±0.79</td>
</tr>
<tr>
<td>7</td>
<td>N_2</td>
<td>209.37</td>
<td>±2.47</td>
</tr>
<tr>
<td>8</td>
<td>K_2</td>
<td>109.92</td>
<td>±2.60</td>
</tr>
<tr>
<td>9</td>
<td>M_4</td>
<td>218.37</td>
<td>±44.54</td>
</tr>
<tr>
<td>10</td>
<td>M_d</td>
<td>264.04</td>
<td>±37.76</td>
</tr>
<tr>
<td>11</td>
<td>M_m</td>
<td>44.79</td>
<td>±7.63</td>
</tr>
</tbody>
</table>

3.3. Type of Tidal

The type of tidal was determined through the form factor or form number ($F$) which was the ratio between the major amplitudes of diurnal and semidiurnal constituents. $F$ was calculated using Equation 25 with amplitudes of the four most dominant constituents, $K_1$, $O_1$, $M_2$, and $S_2$, obtained from tidal analysis. These four constituents have been proven in Figure 6 to have the largest amplitudes. Tidal types of three tidal locations were identified based on $F$ values and were mixed mainly semidiurnal types as presented in Table 5. The $F$ values for the three locations, were 0.359, 0.360, and 0.357, respectively, showing that the tidal generating potentials of semidiurnal constituents were stronger than those of diurnal constituents. The types were identified quite similarly to those reported by Setiyawan et al. [26] and Sabhan et al. [27] with $F$ values of 0.440 and 0.357, respectively.

Table 5. Type of Tidal based on the magnitudes of $F$

<table>
<thead>
<tr>
<th>Location</th>
<th>$F$</th>
<th>Type of Tide</th>
</tr>
</thead>
<tbody>
<tr>
<td>Donggala Port (DP)</td>
<td>0.359</td>
<td>Mixed mainly semidiurnal</td>
</tr>
<tr>
<td>Watusampu (WS)</td>
<td>0.360</td>
<td>Mixed mainly semidiurnal</td>
</tr>
<tr>
<td>Taman Ria (TR)</td>
<td>0.357</td>
<td>Mixed mainly semidiurnal</td>
</tr>
</tbody>
</table>

3.4. Performance Evaluation

The quantitative performance of the LSM was assessed based on its level of accuracy using statistical parameters. Meanwhile, multiple statistical tests were required due to the possibility of single tests, such as the coefficient of correlation, having consistency errors. Previous research also applied quantitative methods and statistical parameters to evaluate the performance of models in forecasting wave and tidal patterns, such as the works of Ris et al. [64], Ardhuin et al. [65], Akpınar et al. [66], Mentaschi et al. [67], Bryant et al. [68], Ding et al. [69], Yang et al. [70], Bradbury and Conley [71], and Zhang et al. [72]. The statistical parameters selected to measure the quality of results from tidal prediction analysis were biased ($b$), Root Mean Square Error ($RMSE$), coefficient of correlation ($R$), and symmetric slope ($SR$). The parameters are expressed mathematically as follows:

\[
b = \frac{1}{N} \sum_{m=1}^{N} (h(t_m) - h_m) \tag{28}
\]

\[
RMSE = \sqrt{\frac{1}{N} \sum_{m=1}^{N} (h(t_m) - h_m)^2} \tag{29}
\]

\[
R = \frac{\sum_{m=1}^{N} (h(t_m) - h_m)(h(t_m) - \overline{h(t_m)})}{\sqrt{\sum_{m=1}^{N} (h(t_m) - h_m)^2} \sqrt{\sum_{m=1}^{N} (h(t_m) - \overline{h(t_m)})^2}} \tag{30}
\]

\[
SR = \sqrt{\frac{\sum_{m=1}^{N} h(t_m)^2}{\sum_{m=1}^{N} h_m^2}} \tag{31}
\]

where, $\overline{h(t_m)}$ is the mean value of the predicted seawater elevation and $\overline{h_m}$ is the mean value of the observed seawater elevation. The length of the tidal series used to evaluate the accuracies of tidal predictions compared to observations was 118 days. This showed that the first 118 days of tidal prediction data were compared with 118 days of observed tidal series to assess the accuracies.
Figure 9 shows a comparison between predicted and observed tidal fluctuations across 15 days of tidal time series data from three stations. The predictions were found to be significantly similar to the observed data. The crests of water fluctuation in the predictions tended to be slightly lower than those in the observed time series data, while the troughs were slightly higher. The differences between the predicted and observed tidal data series for the three tidal stations were generally found to be in the range of ± 25 cm, as presented in Figure 10. These differences were identified as equivalent to those reported by Ali et al. [17]. This research was conducted using a 5-minute sampling interval to record tidal data series, which had a higher frequency than the 1-hour interval used by Ali et al. [17]. A high-frequency sampling interval captures more hydrodynamic disturbance on the sea surface than a low-frequency sampling interval. This shows that the 5-minute sampling interval tends to record more undulation than the 1-hour sampling interval. Generally, the residuals are fairly small. This indicates that seawater surface fluctuations are mostly driven by harmonic astronomical forces [35]. Therefore, meteorological forces such as wind stress, atmospheric pressure variations, and short waves, as well as hydrodynamic disturbances including marine traffic and river discharges, had small effects on sea level fluctuation inside Palu Bay. The $b$ values of the tidal data series were recorded to be $5.00 \times 10^{-10}$, $1.24 \times 10^{-9}$, and $1.32 \times 10^{-9}$ for DP, WS, and TR, respectively. These values were found to be very close to zero, and the tidal predictions were discovered to be the appropriate estimations. This is based on the criteria that a negative value shows underestimation, and a positive one denotes overestimation. Furthermore, the variation of the residual was assessed using RMSE, and the values were recorded to be 0.050, 0.061, and 0.064 for DP, WS, and TR stations, respectively. The closeness of these values to zero shows an almost perfect fit between the predicted and observed data. This demonstrates that the tidal observations combined with LSM for tidal analysis produce a good quality of tidal prediction.

![Figure 9. Comparison of 15 days predicted and observed tide data series; (a.) Donggala Port, DP; (b.) Watusampu, WS; (c.) Taman Ria, TR](image)

![Figure 10. Discrepancies between predicted and observed tide data series, b, and RMSE](image)
Figure 11 shows the scatter plots with the predicted tidal data on the vertical axis and the observed tidal data on the horizontal axis as well as the best fitted linear regression lines drawn with black dashed lines. The accuracy of the predicted data was determined based on the closeness of the dots plotted in the graph to the black dashed line. The distances between the dots above or below the black line show how much the predicted tidal data are overestimated or underestimated, respectively. The dashed green lines represent the ±25 cm limit of prediction error and this means the predictions have acceptable accuracy when the data are within the range of ±25 cm error lines. The red line represents the linear regression fitted line with the slope shown by SR and is ideally expected to be equal to 1.0. Values greater than 1.0 imply a persistent overestimation in the prediction while values below 1.0 denote a consistent underestimation. SR values recorded for DP, WS, and TR were 0.990, 0.987, and 0.985, respectively, thereby showing good accuracies for the prediction results. Another statistical parameter applied to evaluate the precision of the prediction was R and the results are presented in Figure 11. A value of 1.0 also denotes a perfect linearly positive relationship while 0.0 represents the absence of a linear relationship. R values for DP, WS, and TR were found to be 0.995, 0.993, and 0.992, respectively signifying that LSM produced an almost precise prediction for the tidal series.

The results of LSM applied for tidal analysis and prediction in this research were compared to the other five similar studies conducted using different methods, such as artificial neural networks (ANN) [23], inaction methods based on normal time-frequency transform theory (IM-NTFT) [3, 25], long-short-term memory recurrent neural networks (LSTM-RNN) [70], and the Fourier transform method (FFM) [73]. RMSE and R of all these methods, as well as the average values of RMSE and R of the three stations studied in this research, are presented in Table 6. It was discovered that the RMSE of LSM was slightly higher than other methods, by 10% to 17%. However, the RMSE of IM-NTFT-1 applied by Cai et al. [21] with a shorter T of 6.5 months was quite large, at approximately 50%, compared to the LSM used in this research. In contrast, IM-NTFT-2 used by Li et al. [25] with a T of 72 months also had a better RMSE of about 10%. The coefficient of correlation (R) magnitude produced recorded for LSM was found to be comparable to other methods, as shown by the quite small differences of -0.5%, 0.7%, 3.0%, and -0.1% to ANN, IM-NTFT-1, IM-NTFT-2, and LSTM-RNN, respectively. The positive values of differences refer to the prominence of LSM. Generally, the quality of tidal prediction associated with performance indicators indicates that the longest record length (T) and the smallest sampling interval (Δt) produced the highest level of accuracy.

<table>
<thead>
<tr>
<th>Researchers</th>
<th>Method</th>
<th>Record length, T (months)</th>
<th>Δt (minutes)</th>
<th>RMSE (cm)</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>Present study</td>
<td>LSM</td>
<td>4</td>
<td>5</td>
<td>5.8</td>
<td>0.993</td>
</tr>
<tr>
<td>Meena &amp; Agrawal (2015) [22]</td>
<td>ANN</td>
<td>7.5</td>
<td>15</td>
<td>4.8</td>
<td>0.998</td>
</tr>
<tr>
<td>Cai et al. (2018) [20]</td>
<td>IM-NTFT-1</td>
<td>6.5</td>
<td>60</td>
<td>8.9</td>
<td>0.986</td>
</tr>
<tr>
<td>Li et al. (2019) [24]</td>
<td>IM-NTFT-2</td>
<td>72</td>
<td>60</td>
<td>5.2</td>
<td>0.961</td>
</tr>
<tr>
<td>Yang et al. (2020) [70]</td>
<td>LSTM-RNN</td>
<td>252</td>
<td>60</td>
<td>4.9</td>
<td>0.994</td>
</tr>
<tr>
<td>Kusuma et al. (2021) [73]</td>
<td>FFM</td>
<td>1.5</td>
<td>1</td>
<td>5.2</td>
<td></td>
</tr>
</tbody>
</table>
in tidal data series, which create noises such as meteorological and hydrodynamic disturbances. In this research, the residuals and parameters of performance indicators showed an acceptable level of accuracy for LSM despite the inclusion of non-tidal harmonic fluctuations in the calculation.

### 3.5. Tidal Datums

A tidal datum at a specific location is generally referred to as the average elevation of a particular stage of a certain interval record of the tide. The tidal datum is the vertical distance of the sea surface and is used as the reference elevation from which height and depth measurements are taken. The average value of the tide variation is influenced by an 18.6-year lunar nodal cycle that is called lunar nodal regression. This refers to the fact that over 18.613 years, the angle between the plane through the equator and the plane through the moon's orbit changes from 18.3° to 28.5°. In this research, a 19-year tide data series is chosen to average out the tidal datum [10]. It is important to establish an observed time series long enough to average out the oceanic and hydraulic variability, as well as all significant tidal periods, to create a stable datum. It is significant to use longer tidal data series to avoid the possible bias in tidal datum values that could exist from a shorter period of data series [61].

The data series of 118 days, or roughly four months, of tide observations was extended to 19 years using Equation 1 and the harmonic constants in Figures 6 and 8. Subsequently, the 19-year tidal series was analyzed to determine the Highest Astronomical Tide (HAT) and Lowest Astronomical Tide (LAT). The tidal types in Palu Bay, as represented in Table 5, are mainly mixed semi-diurnal. The important tidal datums based on these tidal types are Mean High Water Springs (MHWS) and Mean Low Water Springs (MLWS) [7]. Other tidal datums calculated are Mean High Water (MHW) and Mean Low Water (MLW). The tidal datums were calculated based on the peak approach proposed by Palmer et al. [74]. HAT and LAT are obtained from the highest and lowest sea surface elevations of the tidal predictions over a period of 19 years. MHWS and MLWS are computed by averaging all peaks and troughs of the spring tide, respectively, over 19 years. MHW and MLW are the average values of all daily peaks and troughs estimated over 19 years of tidal prediction. For the semi-diurnal tide, peaks and troughs occur within a day.

The magnitudes of tidal datums based on 19 years of tidal prediction are presented in Table 7, and the values were nearly similar for the three tidal stations. However, DP showed slightly smaller HAT, MHWS, and MHW values and slightly larger LAT, MLWS, and MLW values than WS and TR. The results further showed that WS had the largest tidal datum and tidal range (RA) values, followed by TR and DP. This difference was associated with the changes in inward geometry, which led to a reduction in the depth and narrower width of the bay.

The RA of predictions and those of observations were compared. The RA of tidal prediction is the difference between HAT and LAT, and those of tidal observation is the difference between HWL and LWL, as shown in Table 3. It was discovered that RA for the observations was slightly larger by approximately 3.9%, 4.5%, and 5.8% for DP, WP, and TR, respectively. Moreover, the average RA magnitudes of the prediction and observation were 2.39 m and 2.51 m, with a 5.0% difference. This means the tidal range in Palu Bay can be classified as meso-tidal due to the range of values between 2.0 and 4.0 m [75].

The LAT of Palu Bay was previously calculated by Djunarsjah et al. [76] to be -1.2 m using long-term tidal observation data. Sabhan et al. [27] calculated LAT by using constituent amplitudes and obtained -1.34 m of LAT. While the LAT of observation is defined by the average LWL of the three tidal stations listed in Table 3, it is -1.19 m. The average LAT of three tidal stations for the tidal prediction in this research was -1.13 m, which was 5.3%, 5.8%, and 15.7% smaller than the absolute values recorded during tidal observations as well as by Djunarsjah et al. [76] and Sabhan et al. [27], respectively. Moreover, the LAT of tidal observations seemed nearly similar to the value reported by Djunarsjah et al. [76]. The difference in tidal datum values between observation and prediction can be associated with non-astronomical factors such as atmospheric pressure variation, local wind, short waves, discharge inflow, and ship disturbances [77] that are not removed from tidal observation data. Furthermore, the RA of tidal prediction and Sabhan et al. [27] are 2.39 m and 2.68 m. This showed that the RA value for Sabhan et al. [27] was 12.1% higher than the RA of tidal prediction.
value for tidal prediction. This overestimation was possibly due to the usage of a shorter tidal record length (T) of three days by Sabhan et al. [27] to separate the harmonic constants based on the $T_R$ required in Table 2, leading to the generation of some errors in amplitudes ($H_i$) and phase lags ($g_i$).

4. Conclusion

In conclusion, the acquisition of in situ tidal records was considered important for the construction of coastal infrastructure and navigation. However, it was discovered that precise tidal series records by authorized tidal stations were unavailable in some specific locations. This highlighted the need for temporary stations with proper tidal gauges to provide accurate, real-time tidal records. In Palu Bay, attempts were made to obtain good-quality tidal records by setting up three temporary tidal stations equipped with pressure-type tidal gauges. The time series data recorded for four months showed realistic harmonic patterns, indicating the reliability of the gauge.

Tidal datums used as reference levels for coastal infrastructure and navigation purposes must be calculated using a minimum of 18.6 years of tidal record, which represents a lunar nodal cycle. This led to the application of LSM to separate the harmonic constants, including amplitudes and phases of tidal constituents, using a short period of tidal record. These constants were further used to predict a 19-year tidal series in advance using 11 selected constituents based on the magnitude of the tidal-generating potential. The results showed that the most dominant constituent based on amplitude was $M_2$, followed by $S_2$, $K_1$, $O_1$, $N_2$, $K_2$, $P_1$, $Q_1$, $M_4$, $M_6$, and $M_{2d}$.

The performance of the tidal prediction method was assessed using statistical parameters, and the results of LSM showed relatively small differences compared to the values obtained from tidal observations. The level of accuracy of the method was also compared to other advanced methods, and it was found to be beneficial for tidal forecasting due to its simple algorithm and low specification requirements for personal computers.

5. Declarations

5.1. Author Contributions


5.2. Data Availability Statement

The data presented in this research are available on request from the corresponding author.

5.3. Funding

This research was funded by the University Development Research Scheme of DIPA BLU Universitas Tadulako, with grant number 769.k/UN28.2/PL/2022.

5.4. Acknowledgements

The authors are grateful to the Environmental Fluid Dynamics Laboratory at Kyushu University for the assistance on field observations as well as the Faculty of Engineering Tadulako University for providing computing facilities for this research.

5.5. Conflicts of Interest

The authors declare no conflict of interest.

6. References


Courtier, A. (1939). Classification of tides in four types. The International Hydrographic Review, 50-58


ICSM. (2005). The Factors Contributing to the level of Confidence in the Tidal Predictions Accuracy of Tidal Predictions. Intergovernmental Committee on Surveying and Mapping (ICSM), Sydney, Australia.