ISSN: 2581-1274

Ocean Tides Model in Eastern Indonesian Sea Using Data Assimilation from Altimetry, Tide Gauge, and Hydrodynamic Model

Misfallah Nurhayati [a,1,*], Dudy Darmawan Wijaya [b,2]

[a] *Department of Geomatics Engineering, Institute of Technology Sumatera, Jl. Terusan Ryacudu, Lampung Selatan 35365, Indonesia.*
[b] *Department of Geodesy and Geomatics Engineering, Institut Teknologi Bandung, Jl. Ganesha 10, Bandung,40116, Indonesia*
[1]*misfallah.nurhayati@gt.itera.ac.id*,* [2]*dudy.wijaya@itb.ac.id*
** corresponding author*

ARTICLE INFO

Article history:
Published

Keywords:
Altimetry
Assimilation
Shallow water
Eastern Indonesia
Hydrodynamic model

ABSTRACT

Tides play a crucial role in various coastal and marine activities. Despite numerous tidal models developed in Indonesia, modeling tides in shallow waters with complex topographies remains challenging, particularly in the Eastern Indonesian Sea. This study aims to develop a high-accuracy tidal model for Eastern Indonesian Sea using data assimilation techniques integrating observational data and ocean dynamics model. The study tested various scenarios, including different numbers of tide gauge observations, representers, and decorrelation length values in dynamic equations. By assimilating data from altimetry, tide gauges, and hydrodynamic models, significant improvements in model accuracy were achieved. Results show that of the 11 validator tide stations, while the rest stations have the predicted RMS values below 16 cm, and seven stations have the predicted RMS values below 9 cm. These findings highlight the potential use of assimilation technique for accurate tidal predictions in shallow waters with complex topographies, enhancing the safety and efficiency of coastal and marine activities in Eastern Indonesia.

I. Introduction

Accurate tidal modeling is crucial for shaping coastal environments and influencing a wide range of marine activities, making precise tidal predictions essential for navigation, fishing, coastal management, and various scientific applications [1][2][3]. In Indonesia, particularly in the eastern region, the complexities of shorelines and varied depths [4] pose significant challenges to traditional tidal modeling methods. Previous studies have highlighted the urgent need for improved tidal models that can account for the unique geographical and hydrodynamic characteristics of Eastern Indonesia. The intricate shoreline and significant depth variations necessitate a modeling approach that integrates high-resolution observational data and dynamic ocean processes [5].

The advent of satellite altimetry has revolutionized our ability to observe and measure sea surface heights over large areas with high temporal resolution [6]. However, altimetry data alone may not sufficient for accurate tidal predictions in regions with complex topographies [7][8]. Tide gauges, which provide precise local sea level measurements, offer critical ground truth data but are limited in spatial coverage [7]. Hydrodynamic models, which simulate ocean currents and tidal flows based on physical principles, can fill these gaps but require accurate initial and boundary conditions to perform effectively [9]. Data assimilation techniques, which integrate observational data with model simulations, have emerged as a powerful tool to enhance tidal predictions [10]. By combining altimetry, tide gauge data, and hydrodynamic models, data assimilation can provide more accurate and comprehensive tidal predictions [11].

This study aims to develop a high-accuracy tidal model for the Eastern Indonesia Sea through the integration of satellite altimetry, tide gauge observations, and hydrodynamic modeling using data assimilation techniques. The primary objective is to improve tidal predictions in this region [12], which is characterized by shallow waters and complex topographies. By addressing these challenges, the model seeks to provide reliable tidal information crucial for various coastal and marine activities.

The remainder of this paper is organized as follows: Section II describes the data sources and methodology used for the tidal model development. Section III presents the results of the model validation and analysis. Finally, Section IV concludes the study and suggests directions for future research.

II. Method

To develop a high-accuracy tidal model for Eastern Indonesian Sea, we integrated data from satellite altimetry, tide gauges, and a hydrodynamic model. This approach aimed to address the complexities associated with the region's intricate shorelines and variable depths [9]. Our primary data sources included TOPEX altimetry data from 1992 to 2002, providing high-resolution sea surface height measurements. Additionally, we utilized tide gauge data from 58 stations within the study area, with observations from the years 2017-2019, supplied by the Indonesian Geospatial Information Agency (Badan Informasi Geospasial, BIG). Despite the temporal gap, the TOPEX/Poseidon data is chosen for its precision in measuring sea surface heights. The tide gauges offered precise local sea level measurements essential for validating our model. This dual approach allows us to leverage the high accuracy of TOPEX/Poseidon data while using the recent tide gauge data to ensure the reliability and validity of our results. By combining these datasets, we aim to provide comprehensive insights into sea level dynamics in the study area.

In addition to serving as validation data, this data is also used as input to determine the amplitude and phase of tidal components for assimilation processes. In this study, eight primary tidal components are considered: M2, S2, O1, K1, N2, K2, P1, and Q1. The M2 component is the main semi-diurnal tide of the Moon, with a period of approximately 12.42 hours, while the S2 component is the main semi-diurnal tide of the Sun, with a period of approximately 12.00 hours. Diurnal components include O1 dan K1, where O1 is generated by the Moon's gravitational force with a period of about 25.82 hours, and K1 is generated by the combined gravitational forces of the Moon and Sun with a period of about 12.00 hours. N2 is a larger lunar elliptic semi-diurnal tide caused by the eccentricity of the Moon's orbit, with a period of about 12.66 hours, while K2 is a lunar-solar semi-diurnal tide with a period of about 11.97 hours. P1 is the principal solar diurnal tide with a period of about 24.07 hours, and Q1 is a larger lunar elliptic diurnal tide with a period of about 26.87 hours. Each of these components has different frequencies and amplitudes, collectively contributing to the variations in tidal patterns at different locations.

The amplitude and phase values of these components can be derived from tidal analysis results. To simulate ocean currents and tidal flows, we employed a hydrodynamic model that incorporated dynamic equations adjusted for tidal components. Fig.1 shows the distribution of satellite altimetry data and tide gauge station data used in this study, with the blue dots indicating the TOPEX altimetry data and red dots indicating tide gauge stations. The research area boundaries are defined as 16°S to 12°N latitude and 116°E to 142°E longitude.

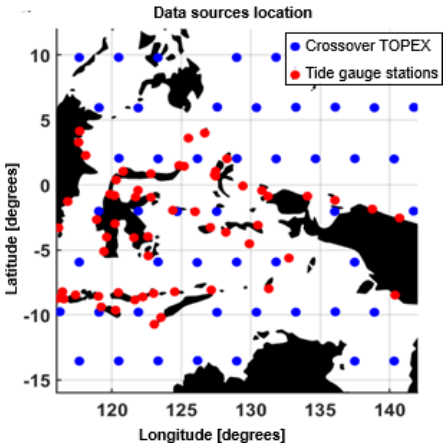

Fig. 1. TOPEX and tide gauge station locations as data sources

Before integrating these data sources, we undertook a rigorous preprocessing phase to ensure data quality and accuracy. The altimetry data were corrected for atmospheric effects, orbital errors, and

tidal aliasing using standard correction algorithms. Simultaneously, the tide gauge data were processed to remove noise and outliers, depicted by Fig. 2, through wavelet analysis and filtering techniques, maintaining the integrity of the observational data. The data reformatting process included resampling and the removal of data jumps and outliers. To identify jumps or outliers in the data, it is necessary to first detect sudden changes in the time series and reduce these changes relative to zero point. Subsequently, any sea level height values greater than three times the standard deviation are considered outliers.

After the observation data is free from outliers and jumps, shown by Fig. 3, the next step is wavelet decomposition analysis. The purpose of wavelet analysis is to divide the wave spectrum into several period ranges in the spatial domain. This division helps to focus the filtering process on specific wavelets, making the subsequent filtering steps easier. The noise identified from the wavelet decomposition analysis is subsequently filtered out. Fig. 4 shows the wavelet analysis result prior to filtering.

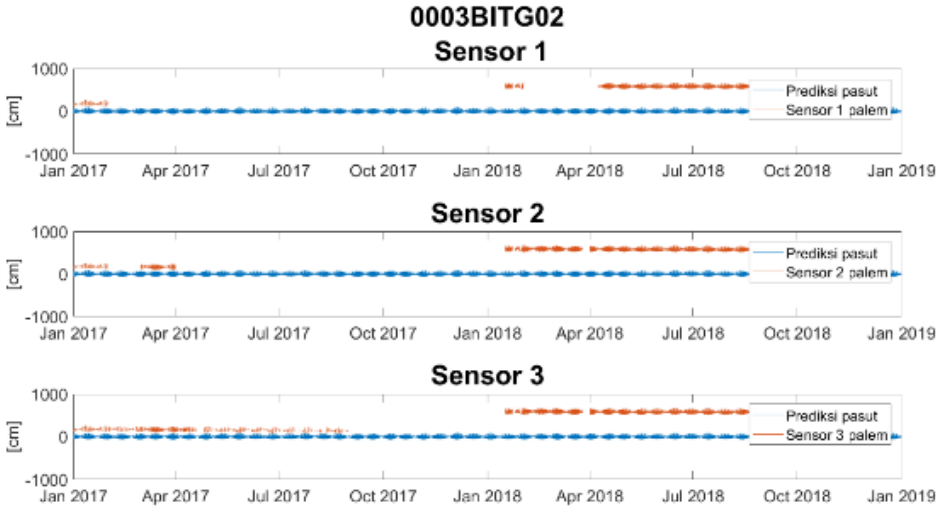

Fig. 2 Sample of tide gauges with jumps and outliers

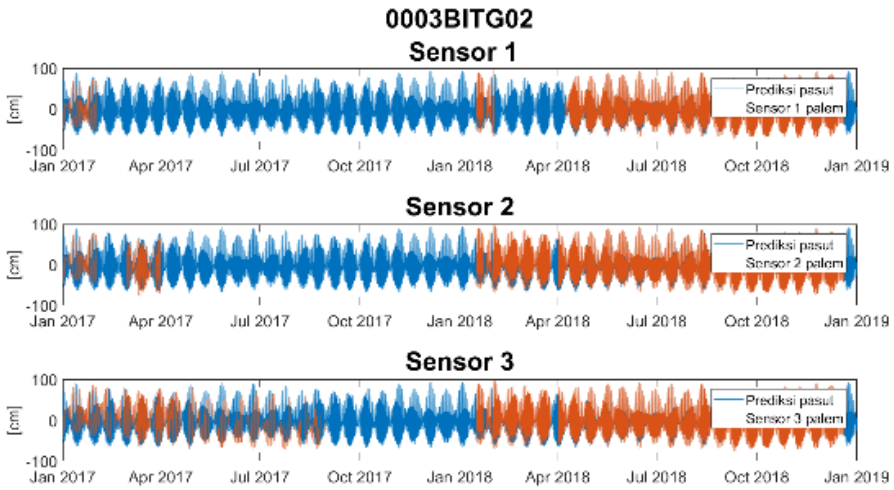

Fig. 3. Sample of tide gauges data correction without jumps and outliers

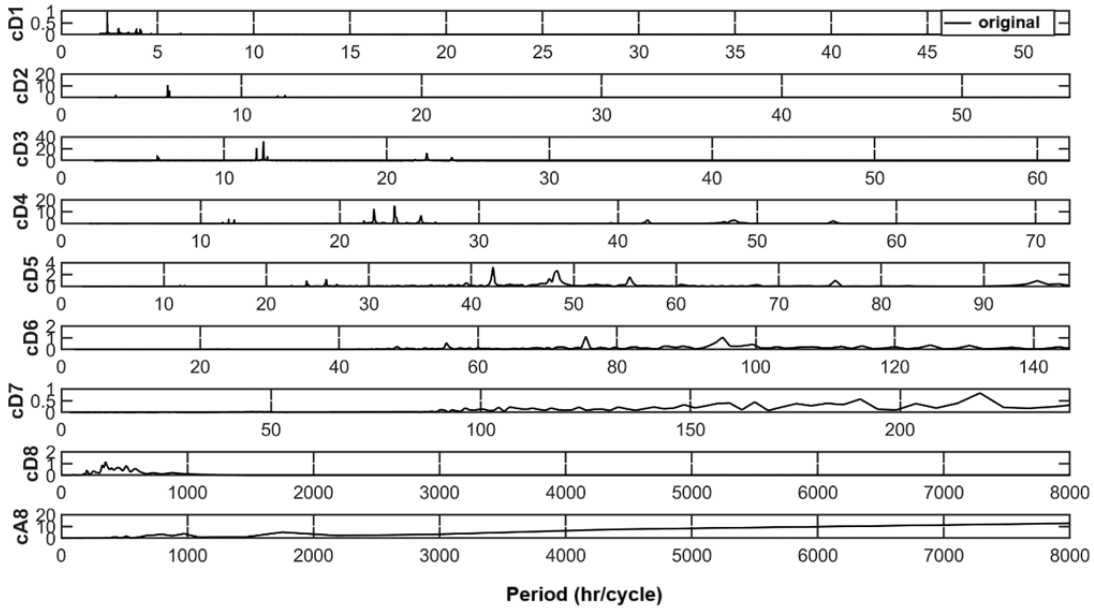

Fig. 4. Wavelet analysis result in frequency domain

Filtering is a step to concentrate the signal by removing noise components from the data. Noise interference in the data can cause significant errors in calculations. From Fig. 4, we only use cD2 and cD6 wavelets to remove the noise as they contain the main tidal component. Fig. 5 shows the filtering results in the temporal domain.

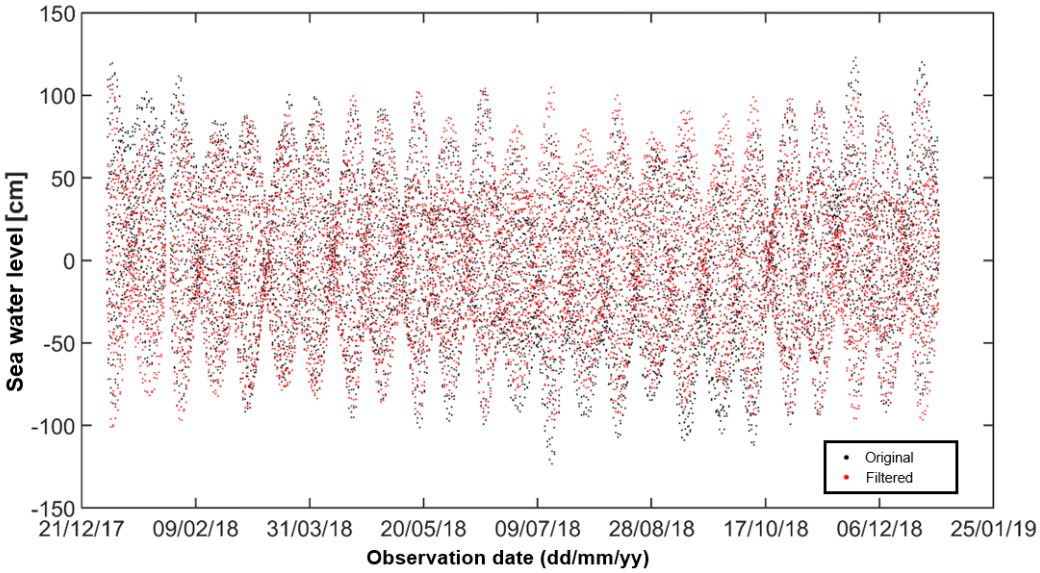

Fig. 5. Wavelet decomposition analysis result

The hydrodynamic model was initialized with accurate bathymetric data and boundary conditions reflective of the study area's unique geographical features. Bathymetric data contains information about sea depth, which is necessary for creating a grid model. In this study, the bathymetric data used is the TOPO 13.1 model, the latest bathymetric model from Smith and Sandwell [13]. TOPO 13.1 has a resolution of 1–12 kilometres, combining sounding data with marine gravity data from the Geosat and ERS-1 satellites. Bathymetric resolution is crucial in tidal modelling, as high-resolution bathymetry is essential for creating accurate data assimilation tidal models. The grid resolution used in this study is 2 minutes. An open boundary condition is required in the creation of a data assimilation tidal model to define the model domain boundaries and to provide input data. Model domain boundary information is typically extracted from a global tidal model. In this study, the model

domain used is TPXO7.2.

Data assimilation techniques played a crucial role in integrating these diverse data sources, thereby enhancing the model's accuracy. In this study, the assimilation method used is a generalization of the inverse method, as detailed by Egbert et al. [9]. One example of a tidal model created by Egbert et al. [14] is in Indonesian waters. The assimilation process was carried out using the Oregon Tidal Inverse Software (OTIS). We began by initializing the hydrodynamic model with the corrected altimetry data and the filtered tide gauge measurements, providing a robust foundation for the simulation. An advanced assimilation algorithm was then employed to combine the observational data with the model's output. This algorithm iteratively adjusted the model parameters to minimize discrepancies between observed and simulated sea surface heights. Continuous validation against independent tide gauge data ensured the model's reliability, with necessary adjustments made to maintain high accuracy.

The creation of the tidal model was conducted by applying several scenarios that considered factors influencing tidal modeling (Table 1). These factors included variations in the number of representers used ('rep set'), variations in the decorrelation length values as a covariance function affecting the inversion process (abbreviated as 'dl'), and variations in the amount of observational data from the tide gauge stations (indicated by the number after 'TG'). The different sets of representers aimed to examine the influence of the number and distribution of representers on tidal modeling.

The representer points were divided into two data sets, as shown in Fig 6. The first set consisted of 139 points (rep set 1), and the second set consisted of 206 points (rep set 2). These representer points included crossover data from the TOPEX satellite altimetry and randomly selected points based on the complexity and topography of the modeling area.

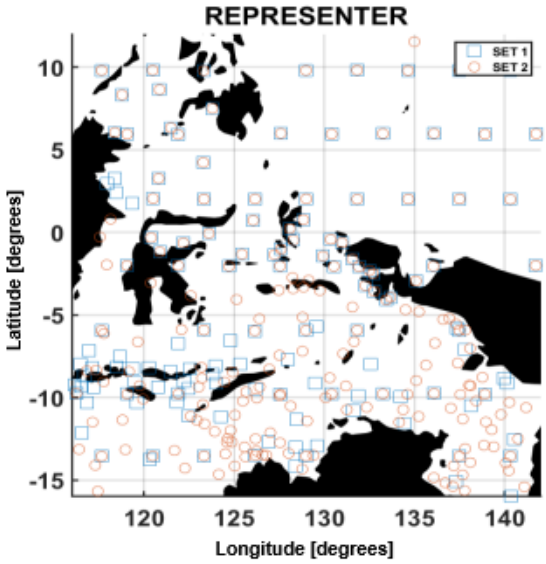

Fig. 6. Representer data set distribution

Decorrelation length is the maximum distance at which a component will still show correlation in the modeling process. This value can affect the dynamical error covariance, which in turn influences the inversion calculation process [15]. The final scenarios, which are the main focus of this study, involve the addition of observational data from tide gauge stations. This addition aims to evaluate the development of the tidal modeling.

Table 1. Scenarios in tidal model creation (v = used, bold = main focus)

No.	Model	pathfinder	**Tide gauge data set 33**	**Tide gauge data set 48**	**rep set 1**	**rep set 2**	**dl 100km**	**dl 200km**	**dl 400km**
a	Path-only	v				v		v	

No.	Model	pathfinder	Tide gauge data set 33	Tide gauge data set 48	rep set 1	rep set 2	dl 100km	dl 200km	dl 400km
b	TG-33.2-200km	v	v			v		v	
c	TG-48.1-200km	v		v	v			v	
d	**TG-48.2-200km**	**v**		**v**		**v**		**v**	
e	TG-48.2-400km	v		v		v			v
f	TG-48.2-100km	v		v		v	v		

The model's performance was validated using observational data from 11 tide stations across the study area. We calculated the Root Mean Square (RMS) error between the predicted and observed sea levels at each station to quantify the model's accuracy. Additionally, we conducted a spatial and temporal analysis of the model's predictions to ensure consistent performance across different regions and periods.

III. Result and Discussion

Before performing data assimilation, the amplitude and phase values of each tidal component need to be determined. These values can be obtained from tidal analysis. In this study, tidal analysis was conducted on the observation data from BIG stations. For the satellite altimetry data, tidal analysis was conducted separately using the Oregon State University Tidal Inversion Software (OTIS) [14]. After obtaining the amplitude and phase values, each data set was converted into a binary format. This conversion was necessary to enable the assimilation of these data sets with the altimetry pathfinder data in the OTIS. Fig. 7 shows the modeling results for scenario (d) for the K1 and M2 tidal component.

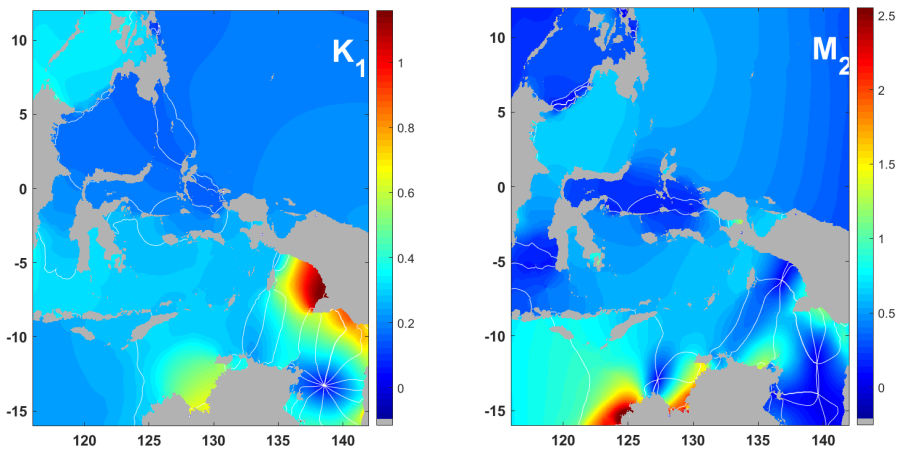

Fig. 7. Amplitude and phase of the K1 and M2 tidal components from assimilation

The highest amplitude values of the K1 component occur around the Arafura Sea, ranging from approximately 0.6 to 1 meter. In the Banda Sea, Flores Sea, Boni Bay, and Savu Sea, the K1 values predominantly range from about 0.25 to 0.3 meters. The K1 values decrease as the wave propagates towards the Seram Sea, Molucca Sea, Halmahera Sea, Tomini Bay, Celebes Sea, and further north into the open ocean, ranging from 0.05 to 0.2 meters. The K1 values rise again to around 0.3 meters when reaching the Sulu Sea. Unlike K1, which experience a decrease in amplitude towards the north, M2 constituents show a decrease towards the east. The amplitude values decrease in the regions of Tomini Bay *(Teluk Tomini)*, the Molucca Sea *(Laut Maluku)*, and the Sulu Sea *(Laut Sulu)*. This behavior is quite different from what is observed in diurnal tides. The highest amplitude values of the M2 component are found around the Arafura Sea, reaching more than 1 meter.

The tidal model developed for the Eastern Indonesia Sea was validated using observational data from 11 tide stations, demonstrating a significant improvement in model accuracy. The tide gauge

stations used as validators in this study are Balikpapan, Tanjung Batu, Parigi, Palopo, Namlea, Gebe, Tual, Jayapura, Tidore, Kupang, and Biak. The first RMS prediction calculation was performed by comparing the model with different sets of representer data. The second calculation compared the model with varying decorrelation length values, and the third compared the model with the addition of tide gauge station data. Table 2 and Fig. 8 shows the RMS results between the models. The models were created using additional data from 48 tide gauge stations with a decorrelation length value of 200 km, as referenced by [14]. Based on the table, it can be concluded that using the second set of representer data (48.2) yields smaller RMS prediction values compared to the first set of representers (48.1). However, at certain stations such as Kupang, Balikpapan, and Namlea, the first set of representers provides better RMS values than the second set.

Table 2. RMS model with different sets of representations

Station index	Station name	TG-48.1-200km (cm)			TG-48.2-200km (cm)		
		s1	*s2*	*s3*	*s1*	*s2*	*s3*
0011PLPO02	Palopo	6,4	70,77	12,55	6,37	70,62	***12,9***
0012KPNG02	Kupang	16,18	9,16	14,54	***16,63***	***9,36***	14,83
0015BIAK03	Biak	9,35	10,21	45,42	9,28	10,16	44,88
0017TUAL03	Tual	11,9	12,25	12,74	10,66	11,01	11,51
0020BLPP02	Balikpapan	6,95	NaN	87,22	***7,29***	NaN	87,63
0025JYPR03	Jayapura	5,12	5,09	4,43	5,12	5,08	4,43
0107NAML03	Namlea	14,79	17,96	17,93	***15,12***	***18,33***	18,31
0123GEBE03	Gebe	6,02	6,38	6,29	5,81	6,18	6,08
0133TJBT02	Tanjung Batu	16,56	17,76	4,91	15,67	17,12	5,73
0153PARI02	Parigi	7,42	3,75	8,57	7,32	***3,79***	8,42
0160TDRE03	Tidore	6,1	5,91	5,97	5,97	5,84	***6,04***

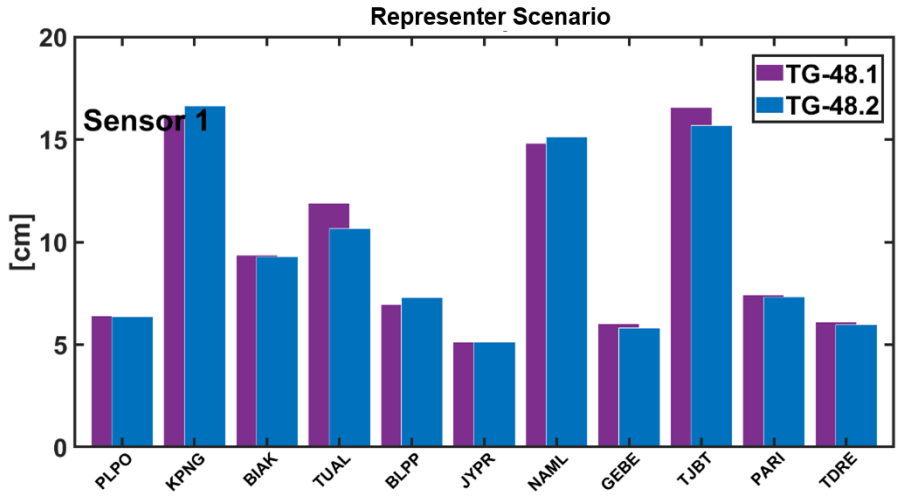

Fig. 8 RMS model with different sets of representations, all using 200 km decorrelation length

The second RMS prediction comparison is based on scenarios with different decorrelation length values. Table 3 and Fig. 9 show the RMS values for different decorrelation length variations. These results highlight the impact of decorrelation length on the accuracy of tidal modeling. The data demonstrates that a decorrelation length of 200 km provides the most accurate predictions, with lower RMS values compared to the 100 km and 400 km scenarios. This suggests that the 200 km decorrelation length is optimal for modeling tides in Indonesia. The increased or decreased decorrelation lengths result in higher RMS values, indicating less accurate predictions.

Table 3. RMS model with different decorrelation lengths

Station Index	Validator Name	TG-48.2-200km (cm)			TG-48.2-400km (cm)			TG-48.2-100km (cm)		
		s1	*s2*	*s3*	*s1*	*s2*	*s3*	*s1*	*s2*	*s3*
0011PLPO02	Palopo	6,37	70,62	12,9	NaN	70,72	12,81	NaN	70,72	12,8

Station Index	Validator Name	TG-48.2-200km (cm)			TG-48.2-400km (cm)			TG-48.2-100km (cm)		
		s1	*s2*	*s3*	*s1*	*s2*	*s3*	*s1*	*s2*	*s3*
0012KPNG02	Kupang	16,63	9,36	14,83	16,56	NaN	16,99	16,51	NaN	16,94
0015BIAK03	Biak	9,28	10,16	44,88	10,73	11,46	44,69	10,73	11,46	44,68
0017TUAL03	Tual	10,66	11,01	11,51	11,66	12,33	12,49	11,68	12,35	12,51
0020BLPP02	Balikpapan	7,29	NaN	87,63	NaN	NaN	89,14	NaN	NaN	89,15
0025JYPR03	Jayapura	5,12	5,08	4,43	5,73	6	NaN	5,7	5,97	NaN
0107NAML03	Namlea	15,12	18,33	18,31	17,81	18,21	18,19	17,81	18,22	18,19
0123GEBE03	Gebe	5,81	6,18	6,08	NaN			NaN		
0133TJBT02	Tanjung Batu	15,67	17,12	5,73	20,07	21,58	NaN	20,06	21,58	NaN
0153PARI02	Parigi	7,32	3,79	8,42	8,49	NaN	8,43	8,49	NaN	8,43
0160TDRE03	Tidore	5,97	5,84	6,04	NaN			NaN		

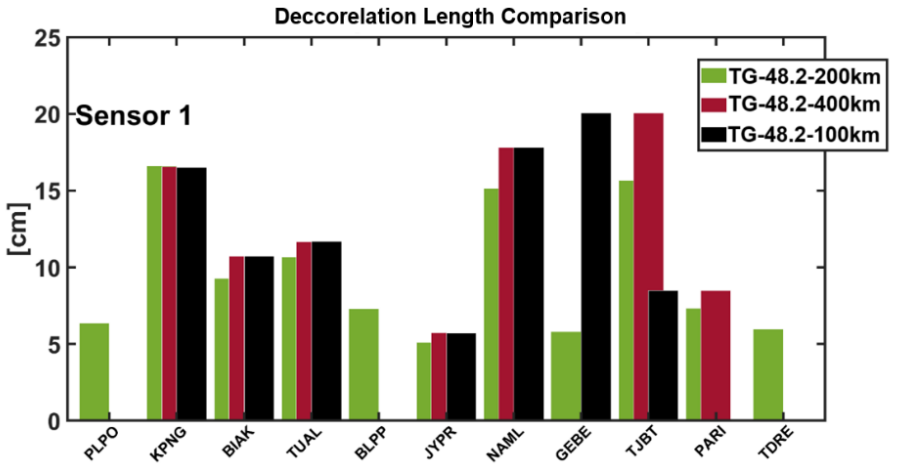

Fig. 9 RMS model with different decorrelation lengths

The next comparison involves evaluating the RMS prediction of the model with varying amounts of observational data from tide gauge stations. This scenario was conducted three times: using only pathfinder data (a), using pathfinder data plus 33 tide gauge stations (b), and using pathfinder data plus 48 tide gauge stations (c). Table 4 shows the RMS values for scenarios (a), (b), and (c). Fig. 10 illustrates the impact of adding input data to the modeling process.

Table 4. RMS model with varying amounts of observational data

Validator Index	Validator name	path_only (cm)			TG-33 (cm)			TG-48.200km (cm)		
		s1	*s2*	*s3*	*s1*	*s2*	*s3*	*s1*	*s2*	*s3*
0011PLPO02	Palopo	6,38	70,64	12,9	6,37	70,63	12,9	6,37	70,62	12,9
0012KPNG02	Kupang	16,52	9,37	14,75	16,52	9,37	14,75	16,63	9,36	14,83
0015BIAK03	Biak	9,26	10,14	45,04	9,26	10,14	45,03	9,28	10,16	44,88
0017TUAL03	Tual	10,67	11,02	11,5	10,71	11,07	11,54	10,66	11,01	11,51
0020BLPP02	Balikpapan	7,33	NaN	87,55	7,3	NaN	87,48	7,29	NaN	87,63
0025JYPR03	Jayapura	5,11	5,08	4,43	5,11	5,08	4,43	5,12	5,08	4,43
0107NAML03	Namlea	15,15	18,36	18,34	15,15	18,36	18,34	15,12	18,33	18,31
0123GEBE03	Gebe	5,81	6,18	6,08	5,81	6,18	6,08	5,81	6,18	6,08
0133TJBT02	Tanjung Batu	15,16	16,99	5,48	15,75	17,15	5,86	15,67	17,12	5,73
0153PARI02	Parigi	7,32	3,79	8,42	7,31	3,79	8,41	7,32	3,79	8,42
0160TDRE03	Tidore	5,98	5,85	6,05	5,97	5,85	6,04	5,97	5,84	6,04

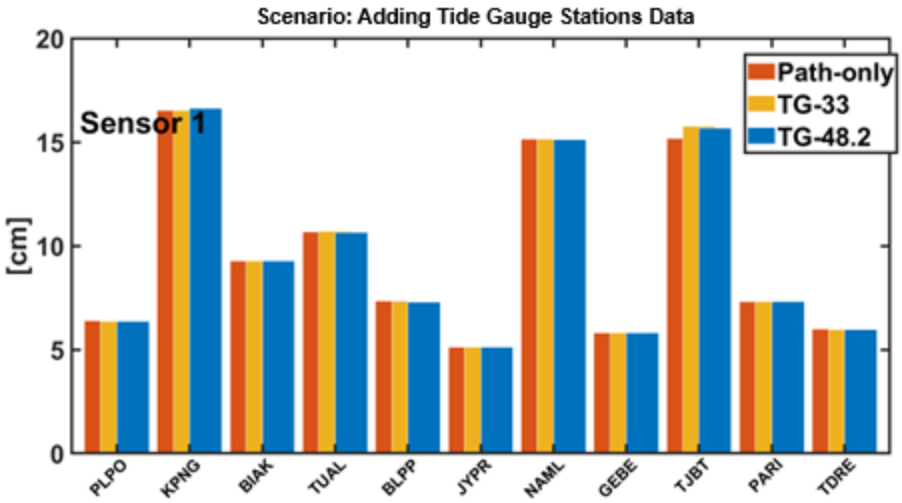

Fig. 10. RMS model with varying amounts of observational data

Based on the Fig. 10 it appears that the inclusion or exclusion of tide gauge station data does not significantly affect the RMS prediction results. This suggests that, so far, adding tide gauge station data has not improved the accuracy of the modelling in this research. The only noticeable improvement was observed in Tanjung Batu.

After testing the effects of different representers, decorrelation lengths, and the addition of tide gauge data, it was found that the second set of representers and a decorrelation length of 200 km provided the best RMS prediction values. The addition of tide gauge stations did not significantly improve the RMS predictions. Consequently, the final model developed uses the second set of representers, a 200 km decorrelation length, and includes 48 tide gauge stations.

This model was then compared to existing models, specifically the INA-BIG prediction model and the global tidal model TPXO9. As shown in Table 5 and Fig. 11 the research model, TG-48.2-200km outperformed the INA-BIG model in RMS prediction accuracy, demonstrating the effectiveness of the chosen parameters.

Table 5. RMS between the pathfinder model predictions, TG-48.2-200km model, TPXO9 model, and INA-BIG model

Validator Index	Validator Name	path_only (cm)			TG-48.2-200km (cm)			TPXO9 (cm)			INA-BIG (cm)		
		s1	*s2*	*s3*	*s1*	*s2*	*s3*	*s1*	*s2*	*s3*	*s1*	*s2*	*s3*
0011PLPO02	Palopo	6,38	70,64	12,9	6,37	70,62	12,9	6,37	70,62	12,9	NaN	69,55	11,98
0012KPNG02	Kupang	16,52	9,37	14,75	16,63	9,36	14,83	16,63	9,36	14,83	22,43	NaN	23,66
0015BIAK03	Biak	9,26	10,14	45,04	9,28	10,16	44,88	9,28	10,16	44,88	10,98	11,64	45,26
0017TUAL03	Tual	10,67	11,02	11,5	10,66	11,01	11,51	10,66	11,01	11,51	12,7	13,36	13,34
0020BLPP02	Balikpapan	7,33	NaN	87,55	7,29	NaN	87,63	7,29	NaN	87,63	NaN	NaN	88,69
0025JYPR03	Jayapura	5,11	5,08	4,43	5,12	5,08	4,43	5,12	5,08	4,43	5,74	6	NaN
0107NAML03	Namlea	15,15	18,36	18,34	15,12	18,33	18,31	15,12	18,33	18,31	16,63	17	16,92
0123GEBE03	Gebe	5,81	6,18	6,08	5,81	6,18	6,08	5,81	6,18	6,08	NaN	NaN	NaN
0133TJBT02	Tanjung Batu	15,16	16,99	5,48	15,67	17,12	5,73	15,67	17,12	5,73	14,46	19,34	NaN
0153PARI02	Parigi	7,32	3,79	8,42	7,32	3,79	8,42	7,32	3,79	8,42	9,65	NaN	9,7
0160TDRE03	Tidore	5,98	5,85	6,05	5,97	5,84	6,04	5,97	5,84	6,04	NaN	NaN	NaN

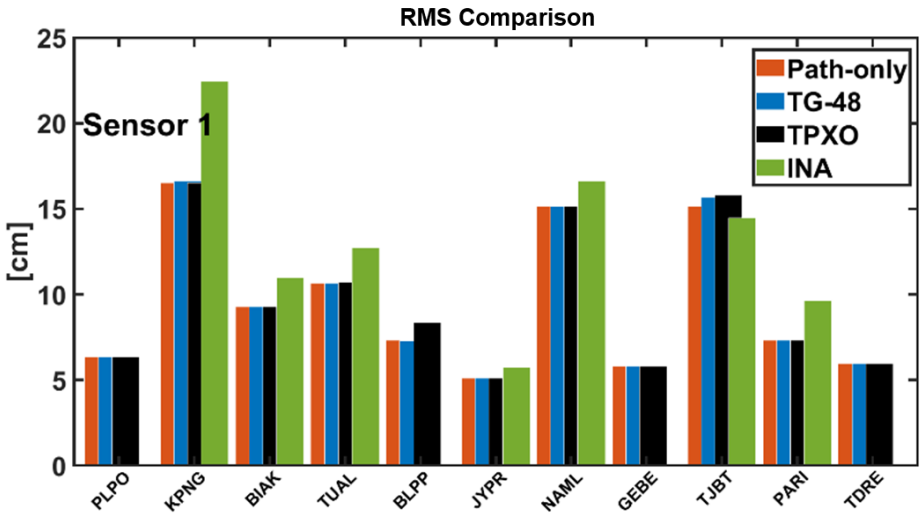

Fig. 11. RMS between the pathfinder model predictions, TG-48.2-200km model, TPXO9 model, and INA-BIG model

Fig. 11 illustrates the RMS comparison between the assimilation tidal model (with and without tide gauge data), the global tidal model, and the tidal model from BIG, known as INA-BIG. The tidal model from BIG serves as a validation reference. However, the INA-BIG model does not comprehensively cover all tide gauge stations in the Eastern Indonesian Sea. As shown in Fig. 11 the RMS values of the assimilation tidal model developed in this study are generally comparable to those of the global tidal model but differ from the INA-BIG model. The RMS values in the INA-BIG model are typically larger, except at the Tanjung Batu (TJBT) station. This discrepancy occurs because the TJBT station is located in shallow water, which impedes the performance of satellite altimetry in measuring tides, as satellite altimetry is most effective at distances greater than 40 km from the shoreline. Satellite data within a 40 km range still experiences interference due to coastal contamination affecting the signals returning to the satellite, and there is inherent bias during data acquisition.

The smallest RMS value was observed at the Jaya Pura (JYPR) tide gauge station. This could be due to the station's exposure to the open sea, which allows for effective data assimilation from altimetry, tide gauge, and hydrodynamic models. The largest RMS value was also observed at the Kupang (KPNG) station, particularly for the INA-BIG model. The high RMS values in the global model and the assimilation model at the KPNG station are due to the tide gauge station's location in a bay area. This setting causes the altimetry data to be less optimal.

The model demonstrated remarkable performance, particularly in regions characterized by complex topography and shallow waters. The causes of large RMS values can be influenced by various factors. These include water conditions or the propagation of errors at each stage of tidal model development. Errors from the instruments used, the resolution of the bathymetry employed, or other factors affecting water movement such as sedimentation type, air friction, water, and seabed friction, as also demonstrated by [12]. These findings indicate the high effectiveness of integrating altimetry, tide gauge data, and hydrodynamic modeling using data assimilation techniques.

IV. Conclusion

In conclusion, the comprehensive analysis and testing of various representers, decorrelation lengths, and tide gauge data incorporation have led to the development of a highly accurate tidal model for the Eastern Indonesia Sea. The optimal configuration, which includes the second set of representers and a 200 km decorrelation length, significantly improves RMS prediction values. Thus, research demonstrates improved accuracy compared to the BIG prediction model. The RMS prediction values increased by a range of 0.5 to 5.8 cm. Out of the 11 tide gauge validator stations, four exhibited RMS prediction values of less than 16 cm, and seven stations had RMS prediction values of less than 9 cm. Although the addition of tide gauge stations did not markedly enhance the model's accuracy, the final model, which integrates 48 tide gauge stations, still outperforms existing

models such as the INA-BIG prediction model and the global tidal model TPXO9. These findings underscore the effectiveness of the chosen parameters and highlight the potential for further refinement and application of the model in tidal prediction and coastal management.

Future research should emphasize the importance of high-resolution bathymetric data and rigorous preprocessing of observational data to maintain tidal model accuracy. Enhancing the integration of diverse data sources through advanced data assimilation algorithms is crucial for capturing intricate tidal variations. Refining the current model with more extensive observational data and exploring its application in other regions with similar complexities will improve the accuracy and reliability of tidal predictions. This approach will benefit a broader range of coastal and marine operations, providing practical solutions for navigation, fishing, and coastal management in areas with challenging geographical features.

References

[1] L. Qiang, Y. Bing-Dong, and H. Bi-Guang, "Calculation and Measurement of Tide Height for the Navigation of Ship at High Tide Using Artificial Neural Network," *Polish Maritime Research*, vol. 25, no. s3, 2018, doi: 10.2478/pomr-2018-0118.

[2] D. M. M. Bezerra, D. M. Nascimento, E. N. Ferreira, P. D. Rocha, and J. S. Mourão, "Influence of tides and winds on fishing techniques and strategies in the mamanguape River Estuary, Paraíba State, NE Brazil," *An Acad Bras Cienc*, vol. 84, no. 3, 2012, doi: 10.1590/S0001-37652012005000046.

[3] J. A. Reis-Filho, F. Barros, J. D. A. C. D. C. Nunes, C. L. S. Sampaio, and G. B. G. De Souza, "Moon and tide effects on fish capture in a tropical tidal flat," *Journal of the Marine Biological Association of the United Kingdom*, vol. 91, no. 3, 2011, doi: 10.1017/S0025315410001955.

[4] D. Nugroho, A. Koch-Larrouy, P. Gaspar, F. Lyard, G. Reffray, and B. Tranchant, "Modelling explicit tides in the Indonesian seas: An important process for surface sea water properties," *Mar Pollut Bull*, vol. 131, 2018, doi: 10.1016/j.marpolbul.2017.06.033.

[5] T. Hatayama, T. Awaji, and K. Akitomo, "Tidal currents in the Indonesian Seas and their effect on transport and mixing," *J Geophys Res Oceans*, vol. 101, no. C5, 1996, doi: 10.1029/96JC00036.

[6] C. K. Shum, "Accuracy assessment of recent ocean tide models," *J Geophys Res Oceans*, vol. 102, no. C11, 1997, doi: 10.1029/97JC00445.

[7] D. B. Chelton, J. C. Ries, B. J. Haines, L.-L. Fu, and P. S. Callahan, "Chapter 1 Satellite Altimetry," in *International Geophysics*, vol. 69, L.-L. Fu and A. Cazenave, Eds., Academic Press, 2001, pp. 1–ii. doi: https://doi.org/10.1016/S0074-6142(01)80146-7.

[8] J. F. Turner, J. C. Iliffe, M. K. Ziebart, and C. Jones, "Global Ocean Tide Models: Assessment and Use within a Surface Model of Lowest Astronomical Tide," *Marine Geodesy*, vol. 36, no. 2, pp. 123–137, Jan. 2013, doi: 10.1080/01490419.2013.771717.

[9] G. D. Egbert, A. F. Bennett, and M. G. G. Foreman, "TOPEX/POSEIDON tides estimated using a global inverse model," *J Geophys Res*, vol. 99, no. C12, 1994, doi: 10.1029/94jc01894.

[10] P. Mazzega and M. Bergé, "Ocean tides in the Asian semienclosed seas from TOPEX/POSEIDON," *J Geophys Res Oceans*, vol. 99, no. C12, pp. 24867–24881, Dec. 1994, doi: 10.1029/94JC01756.

[11] G. A. Kivman, "Assimilating data into open ocean tidal models," *Surv Geophys*, vol. 18, no. 6, 1997, doi: 10.1023/a:1006535821489.

[12] R. D. Ray, G. D. Egbert, and S. Y. Erofeeva, "A brief overview of tides in the Indonesian seas," *Oceanography*, vol. 18, no. SPL.ISS. 4, 2005, doi: 10.5670/oceanog.2005.07.

[13] W. H. F. Smith and D. T. Sandwell, "Global sea floor topography from satellite altimetry and ship depth soundings," *Science (1979)*, vol. 277, no. 5334, 1997, doi: 10.1126/science.277.5334.1956.

[14] G. D. Egbert and S. Y. Erofeeva, "Efficient inverse modeling of barotropic ocean tides," *J Atmos Ocean Technol*, vol. 19, no. 2, 2002, doi: 10.1175/1520-0426(2002)019<0183:EIMOBO>2.0.CO;2.

[15] D. N. Arabelos, D. Z. Papazachariou, M. E. Contadakis, and S. D. Spatalas, "A new tide model for the Mediterranean Sea based on altimetry and tide gauge assimilation," *Ocean Science*, vol. 7, no. 3, 2011, doi: 10.5194/os-7-429-2011.