Applying Parametric Model Correction on Sea State Bias to Estimate Sea Surface Height over the Savu Sea

Alifia Rusydah Fitria Dewi[1], *Aries* Dwi Siswanto[1,2*], *Ashari* Wicaksono[1,2], and *Ary* Giri Dwi Kartika[1,2]

[1]Department of Marine Science, University of Trunojoyo Madura, Jl. Raya Telang PO BOX 2 Kamal-Bangkalan, East Java, Indonesia

[2] Laboratory of Oceanography, Department of Marine Science, University of Trunojoyo Madura, Jl. Raya Telang PO BOX 2 Kamal-Bangkalan, East Java, Indonesia

* Email: ariesdwisiswanto@trunojoyo.ac.id

Abstract. Range, the distance between the satellite and the ocean surface covered by the radar trace, is measured by tracking the received waveform. The range measurement errors may be caused by sea state bias (SSB) which is the main source of measurement errors in altimeter applications. Particularly, the error due to SSB is related to atmospheric dynamics and sea surface roughness. Sea surface height (SSH) is determined by advanced satellite remote sensing in the fields of earth sciences, such as geophysics, geodesy, and oceanography. The research aims to apply the parametric model on the SSB to estimate the SSH over the Savu Sea. We analyzed data along the cycle along the track across the line consisting of the GDR Ku Band. We concluded that the parametric model can be used to effectively improve the SSH correction on the Sentinel 3A altimeter.

Keywords: Altimeter, sea state bias (SSB), sea surface height (SSH)

1. Introduction

Sea surface height (SSH) measurements are based on ocean currents and eddies [1], ocean geodesy and geophysics (e.g. rift zones, ocean floor relief, tectonic plate movement and gravity anomalies), and global climate change. It was important for understanding and tracking sea level rise and El Niño phenomena. In recent decades, tide gauges have been the primary method used to acquire SSH. Satellite altimetry is the most widely used technology in the past three decades [2, 3, 4]. Exploring ocean circulation and observing the topograhy of the seabed, is particularly advantageous when using satellite altimetry [5].

In February 2016, Sentinel-3A was launched, which is the first satellite covers the global ocean and has a short repetition cycle (27 days compared to 369 days), making Sentinel-3A suitable for tracking ocean dynamics [6]. In addition, Sentinel-3A has a few more sensors, i.e. a microwave radiometer, DORIS system, laser retroreflector, and color instrument for the ocean and land. The first dual frequency near-nadir radar altimeter operating exclusively in SAR mode is called SRAL. Sentinel-3A runs at 13.575 GHz and 5.41 GHz (Ku/C bands). In fact, the initial purpose of satellite altimetry was to estimate sea level precisely in open ocean environments based on the waveform or the shape of the pulse-limited radar altimeter echo after reflection on the sea surface [7]. Land and ocean changes can be detected on a regular basis through satellite altimetry. Theoretically, sea surface height (SSH) can be calculated by deducting the height of the satellite's ellipsoid from the recorded instantaneous distance. However, altimeters are complex measurement systems that are subject to many errors [2]. The largest error is known as sea state bias (SSB).

By providing the mean sea level observed by the altimeter is lower than the actual sea level, the SSB error edge spans from millimeters to decimeters [5, 8, 9]. Sea state bias (SSB) is used to characterize the overall mistake related to sea state. Electromagnetic bias, skewness bias, and tracker bias are the three well-known components that make up SSB [3, 4, 8, 10, 11]. Typically, an empirical link between wind speed (U_{10}) and significant wave height (SWH) determines electromagnetic bias. The discrepancy between the mean and median sea surface specular heights is referred to the skewness bias [11]. The third contribution is actually the total of the errors related to the altimeter's tracking of the returning echoes, which has been called tracker bias in the past. Based on empirical estimates, this contribution is part of the overall SSB correction [8].

This is one of the largest sources of altimeter error and must be corrected [2, 4, 12, 13]. Therefore, estimating the SSB becomes an important project to correct the SSH measurement [11]. Using the SSH difference instead of the SSH measurement itself is a simple method to directly remove the unknown geoid signal from the estimation process [12]. To estimate SSB, the empirical models are classified into parametric [14] and non-parametric [9]. In the previous study, parametric models have the advantage of being scalable, intuitive, and easy to use [4]. Therefore, in the current reasearch, we apply parametric model corrections to the sea state errors to estimate the SSH of sub-oceans. The combination of parametric model and SSB was estimated by the BM4 empirical model [15], which expresses the relationship between SSB, SWH, and wind speed at 10 m (U_{10}) height or radar cross section σ_0 [16]. The BM4 model was chosen because from several studies conducted by Gaspar *et al.*,[15], Guo *et al.*, [2], and Guoshou *et al.*, [4], it showed good results using the crossover differences method which resulted in an increase to the correlation coefficient between SSH without SSB correction and SSH using correction.

2. Materials and Method

We used satellite data from Sentinel-3A, i.e SWH and SSH cross-over and then observe the global sea level (cycles 67–68). The research study focused on many dynamic interactions of atmosphere and ocean. We collected data from ascending and descending (**Figure 1)** which provided by European Organisation for the Exploitation of Meteorological Satellites EUMETSAT (https://www.eumetsat.int/). The wind speed data were used the ERA5 reanalysis and obtained from ECMWF (https://www.ecmwf.int/). The ERA5 reanalysis data provides the hourly estimation of waves, atmosphere, and sea surface in the horizontal resolution of 0.25 × 0.25. The tidal datasets provided by the University of Hawaii Sea Level Center (UHSLC) (https://uhslc.soest.hawaii.edu/) and were levels on Research Quality Data (RQD) for altimeter data quality assessment.

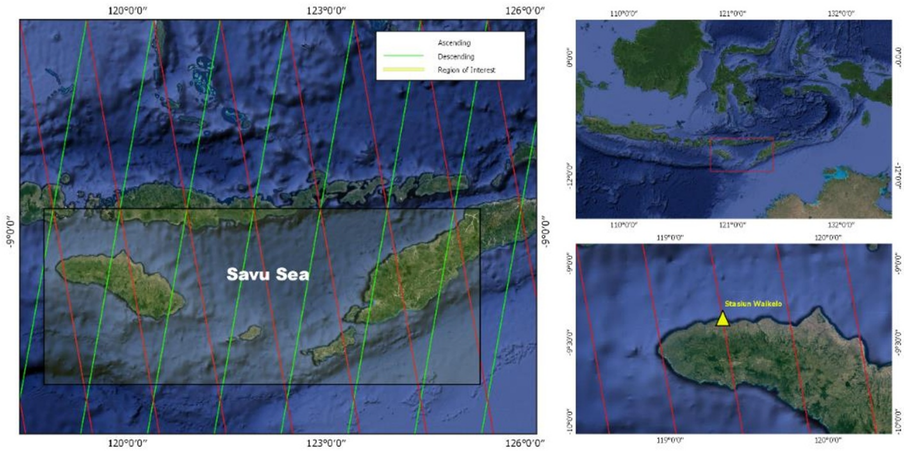

Figure 1. Research Location

Data were eliminated with stringent applied data after we collected is all measurement without unreliable value. Also discarded are the measurements, SWH less than 0 m, SWH more than 11m, U less than 0 m/s or U more than 21m/s [4]. Significant wave height, sea surface height, and wind speed in cross over point at the ascending and descending along track generate 1820 group after we controlled it.

2.1 Theory of Modelling

Several models used to estimate the SSB value using altimeter, such as a parametric model. According to Gaspar *et al.*, [15] the SSB value calculate using taylor expansion and linear regression.

2.1.1 SSH Noise Processing

The altimeter was used to measure the SSH. When the SSB correction is not applied, a raw SSH' includes the height of geoid, N; topographical of dynamic ocean, η; and additional error of altimeter, ε'. With the exception of SSB, all instrumental and geophysical error corrections are included in the ε'. The following is an expression for the SSH:

$$SSH = SSB + \eta + \varepsilon' \quad (1)$$

The dynamic ocean topography and the portion of the geoid height [1] can be removed using the difference of SSH crossover. The following is an expression for the SSH' at the intersection:

$$\Delta SSH' = \Delta SSB + \Delta\eta + \Delta\varepsilon' \quad (2)$$

The dynamic ocean topography (Δη) is defined that the varies time. All measurement error correction. the residual error terms are included in $\Delta\varepsilon'$, except for SSB. The majority of error in altimetry are caused by instrumentation malfunction, tropospheric wet and dry delay error, dynamic atmosphere errors, solid earth error, ionospheric errors, polar tide errors loading tide errors, and ocean tide errors.

2.1.2 Methodology

The parametric model of sea state bias can be stated as follows, assuming that the relationship between SSB and SWH is one of linear dependence speed [15].

$$SSB = f(x,\theta)SWH \tag{3}$$

The sea conditions is defined with x. θ is the coefficient vector, and f is a nondimensional negative, also known as the coefficient of SSB. These variables, which can be any combination of SWH, U, backscatter coefficient, or other measured all the parameters who sea-state related by the altimeters, were typically selected from prior empirical studies. U and the backscatter coefficient have a strong correlation (U is typically retrieved from : those are the two alternatives). U was thus used as a variable. Typically, a second-order Taylor expansion based on wind speed (U_{10}) and significant wave height (SWH) is used to expand Equation (1) [9]. The SSB parameter model can be written such as:

$$SSB = SWH \cdot (a1 + a2 \cdot SWH + a3 \cdot U + a4 \cdot SWH^2 + a5 \cdot U^2 + a6 \cdot SWH \cdot U) \tag{4}$$

SWH is denoted the significant wave height, wind speed is defined with U and the six parameters are $a1, a2, a3, a4, a5, a6$ are to be estimated. However, there many conceivable order combinations, and each one is known as a parameter. The number of parameters was increased from one parameter models to six parameter models, with the first parameter always being present.

2.2 Basis for Parametric SSB Estimation

2.2.1 Linear Regression Estimation

By substituting Equation (4) into Equation (2), we can generate the following :

$$\Delta SSH' = \sum_{i=0}^{6} a_i \Delta X_i + \Delta \eta + \Delta \varepsilon' \tag{5}$$

where $\Delta SSH'$ is the crossover SSH discrepancy except correction of SSB, and ΔX_i (i = 1, . . . , 6) is $\Delta SWH, \Delta SWH^2, \Delta(SWH \cdot U), \Delta SWH^3, \Delta(SWH \cdot U^2), (SWH^2 \cdot U)$, respectively. a_0 is the offset, the noise have equal zero of mean value is defined with e. Therefore, the sum of a_0 and ε a can be expressed for all errors. The equation can be expressed as follows:

$$\Delta SSH' = \sum_{i=0}^{6} a_i \Delta X_i + \varepsilon \tag{6}$$

where ΔX_0 is a dummy variable equal to unity [9]. Problem in multiple linear regression can be seen in Equation (6). The least square method can resolve the parameter, when obtained (ΔSSH, ΔX). The equation can be expressed as follows:

$$a = (\Delta X^T \Delta X^{-1}) \Delta X^T \Delta SSH' \tag{7}$$

2.2.2 Optimization of Model

Analyzing the sample data concentration around the regression line and evaluating the quality of the sample data representation using equatio of the regression serve as indicators for the goodness-of-fit test. The method in regression analysis can two commonly employed are test of regression coefficient adn t test [2]. The equation of these method are follow:

$$t = \sqrt{R^2} \times \sqrt{\frac{n-2}{1-R^2}} \tag{8}$$

$$R^2 = \frac{ESS}{TSS} \tag{9}$$

The data of observation denotes n , TSS stands for total deviation squares sum, ESS for regression squares sum, and R2 for determination coefficient. The goodness fit of the regression equation and the determination coefficient increase as the percentage of the squares sum of the total deviation from the regression squares sum. Equations (8) and (9) report that, under the assumption that the significance level was set at 0.05, the t-test and determination coefficient test were performed for 32 distinct types of SSB parametric models in this investigation. The larger the determination coefficient, the better the model fit the data. The determination coefficient was used to choose the optimal parametric model.

3. Result and Discussion

3.1 Parametric Model Correction

The number of other parameters was increased by preserving the parameter a1, afterwards the calculation of the coefficient value based on Equation (7) and compliance with the modeling fundamental for the parameter model based on Equation (4). Subsequently, thirty-two different types of parametric models were obtained and each model's parameters were estimated sequentially. Equations (8) and (9) are used to solve the t-values and coefficients of determination of the 32 parametric models at a significance level of 0.05. Table 1 provides the parameter estimates, t-values, and coefficients of determination for the 32 parameter models.

The t-values of the 32 models are all greater than the critical values indicated in Table 1 at a significance level of 0.05. Since the six-parameter parametric model fit the data the best and had the highest coefficient of determination, it was selected as the best possible SSB parameter model.

$$SSB = SWH \cdot (-0.67438 - 0.1602 \cdot SWH - 0.3760 \cdot U + 0.03406 \cdot SWH^2 + 0.03011 \cdot U^2 + 0.06867 \cdot SWH \cdot U) \tag{10}$$

This parametric model used with six parameters and showed in Equation (10). These parametric data determined the SWH GDR from Sentinel 3A and the wind speed from ECMWF. By interpolating the grid table based on U and SWH using the GDR model as a guide, SSB is calculated. The correction of SSB value is calculated by the estimation of parametric model regression constructed in the current study. We then compared the values of this model with the SSB from the GDR to assess the parametric model effectiveness in regression estimation. The fitting scatterplot of the parametric model is shown in **Figure 2**.

Table 1. Value of parametric model

a1	*a2*	*a3*	*a4*	*a5*	*a6*	*t*	R^2
			One parameter				
-0.04344						76.098	0.9594
			Two parameter				

-0.04573	0.00076					104.09	0.9778
-0.03791		-0.00109				90.174	0.9707
-0.04573			0.00076			105.35	0.9784
-0.01361				0.00006		91.758	0.9717
-0.01110					0.00010	77.867	0.9611
Three parameter							
-0.01216	0.00186	-0.00744				115.47	0.9819
-0.02876	0.00494		-0.00031			105.42	0.9784
-0.01693	0.00177			-0.00231		118.11	0.9827
-0.02207	0.00322				-0.00098	110.77	0.9804
-0.01693		0.00178	-0.00231			118.41	0.9828
-0.02876		0.00494		-0.00031		91.76	0.9717
-0.02207		0.00323			-0.00098	101.90	0.9769
-0.01120			-0.00224	0.00012		121.02	0.9835
-0.01078			-0.00435		0.00102	111.97	0.9808
-0.01211				0.00022	-0.00089	101.07	0.9766
Four parameter							
-0.0512	0.01393	-0.00133	-0.00286			118.58	0.9828
-0.04555	0.00447	-0.00066		-0.00006		118.56	0.9828
-0.04244	0.00338		-0.00165		0.00011	116.23	0.9822
-0.05343	0.01277		-0.00243	-0.00012		121.13	0.9836
-0.05424	0.01319		-0.00181		-0.00065	112.13	0.9809
-0.04688	0.00419			-0.00014	0.00009	119.10	0.9830
-0.05342		0.01277	-0.00243	-0.00013			
-0.05424		0.01318	-0.00181		-0.00065	119.79	0.9832
-0.04687		0.00419		-0.00013	0.00009	121.22	0.9836
-0.82898			-0.27878	0.00040	-0.00592	122.39	0.9839
Five parameter							
-0.02516	-0.0010	0.00453	-0.00108	-0.00021		121.29	0.9836
-0.01640	-0.0012	0.00161	-0.00126		0.00010	120.88	0.9835
-0.01477	-0.0163	0.00497		-0.00042	0.00150	119.11	0.9830
-0.00171	-0.0126		-0.00203	-0.00010	0.00188	122.91	0.9840
-0.02348		0.00413	-0.00401	-0.00029	0.000906	122.427	0.98392
Six parameter							
-0.67438	-0.1602	-0.3761	0.034061	0.030112	0.068667	123.153	0.98410

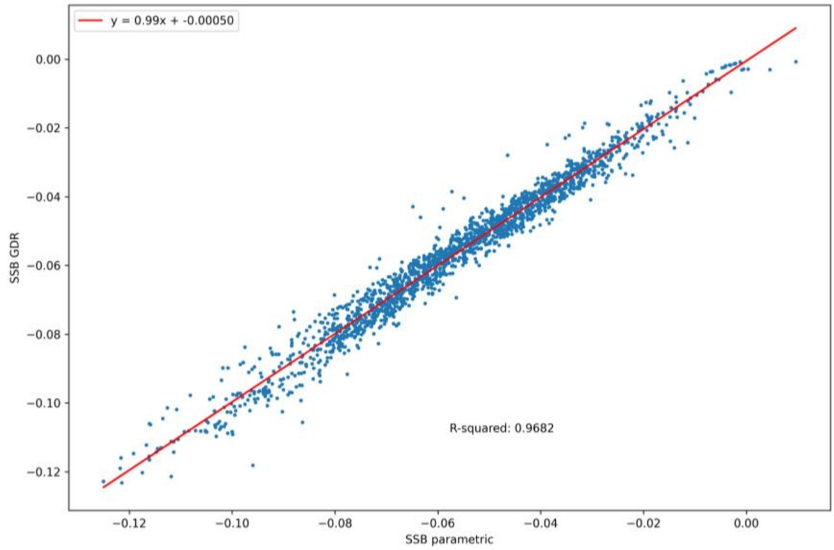

Figure 2. The parametric model (SSB and GDR) scatterplot

SWH and wind data were obtained from the ERA5 reanalysis data. The altitude data were used from cycle 67 to cycle 68. Also, using the width difference method [19], we performed data quality controls by finding cross points without calculating SSB correction values. After determining the SSH value without performing SSB correction, we were applied linear regression to determine the coefficient value a (include uncorrected SWH and U crossover differences and delta SSH values).

The coefficient value a is assigned to equation (4) with SWH and U. The two variables have many combinations and form one of the parameters for calculating the SSB value. The first parameter is preserved by increasing the parameter value from 1 to 6. Depending on the permutations and combinations of parameters, 32 models are obtained. This corresponds to [15]. The 32 models consist of 1-parameter models, 5 2-parameter models, 10 3-parameter models, 10 4 parameter models, 5 5-parameter models, and 1 6-parameter model.

To evaluate the performance of model, we were used t-test (equation 8) and the coefficient determination (equation 9) [19]. The best parametric model is selected based on the coefficient determination values. As a result of the current research, the best model was found at six parameter model (t-test value = 123.153; coefficient determination = 0.9841).

According to **Figure 2**, it showed there is a positive correlation between the SSB and SSB-GDR, and all data are well-fitted. This model shows both statistical differences. Statistically, the graph shows that the standard deviation is 0.0217 and the relative error is ~7.72%. Once we have parametrically corrected SSH values, we check the accuracy of the SSH measurements by comparing to the tide gauge data.

3.2 Accuracy SSH from Tidal Gauge Record

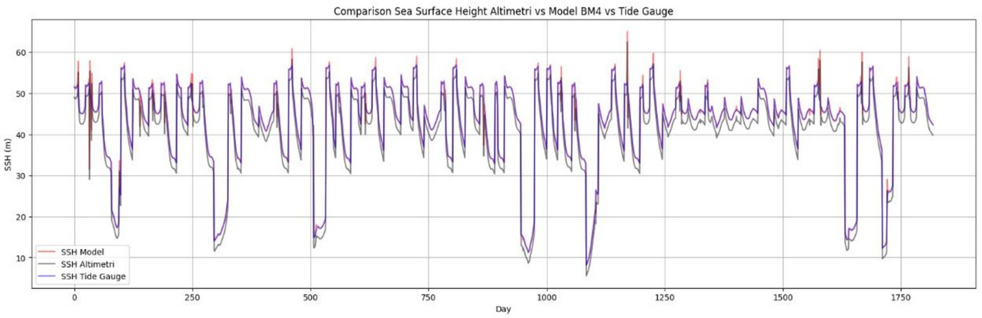

Figure 3. Time series of parametric model and tide gauge records by altimetry

Figure 3 was shown the SSH time series data corrected by parametric model and tidal gauge records based on substracting respective mean values. The horizontal axis define to the cumulative days calculated since 2018, and the vertical axis refers to the SSH after removing the mean value. **Table 2** shows the correlation coefficients and standard deviations between SSH and tide gauge records based on parametric models and GDR model.

Table 2. Statistic calculate from parametric and GDR Model with Tidal Gauge Records

Model	STD	Correlation Coefficient
SSB parametric	0.021532	5.68%
SSB GDR	0.021713	5.34%

We selected hourly data from the 2018 tide gauge at Waikelo Station, which is closest to Path 462. Figure 1 illustrates the precise location of the tidal observatory as well as the satellite transport path. Since the local zero water level is the basis for tide gauge records, the radar altimeter is based on a reference ellipsoid [18]. Because a reference ellipsoid is different, the discrepancy between the altimeter data and the tidal record will remain constant, assuming a constant geoid difference between the tide station geoid and nadir altitude. As a result, we created each elevation time series independently using a parametric model, a GDR model, and a regression estimation model.

Compare the sea surface elevation value without sea state correction or parametric model with the sea surface elevation value obtained from the tide gauge with the average value removed. **Figure 3** shows the SSH value in the height measurement (black line) is lower than the parametric model (red line). These results show that the polynomial model provides good accuracy. Compared to the GDR model, the SSH determined using parametric regression estimation has a higher correlation and a lower standard deviation, indicating that the estimation of parametric regression model was constructed with improved SSH accuracy.

4. Conclusion

SSH is one of the parameters used to understand and determine the movement of sea level rise and natural phenomena, such as ENSO. SSH can be calculated by utilizing altimetry satellites. However, the calculations made by altimetry satellites contain many errors due to bias during the radar reflection process. These errors usually occur due to sea state conditions, with the largest error commonly called sea state bias (SSB). SSB can be eliminated by using a parametric model.

In the current research, the intersection data set are based on the Seninel-3A altimeter GDR and generated using ECMWF. The performance of SSB parametric model on the SSH estimation was constructed according to the taylor expansion of wind speed and SWH. The coefficients for each parametric model were calculated using linear regression methods. After evaluation and comparison, the best model was selected as a 6-parameter model. The model efficiency was evaluated by the t-test and the determination coefficient. When compared to the original SSB estimation model based on Sentinel-3A GDR, the determination coefficient of the model was found 0.984. The good accuracy are also showed by SSH measurement and model to the tide gauge measurements (standard value 0.022, correlation coefficient 5.68%).The parametric model can be used to effectively improve the SSH correction on the Sentinel 3A altimeter.

Actknowledgment: We would like thanks to the reviewers who have provided constructive suggestions that helped improve this paper. We also thank the provider of sentinel 3A data taken from the European Organization for the Exploitation of Meteorological Satellites EUMETSAT (https://www.eumetsat.int/), wind data from European Centre for Medium-Range Weather Forecasts ECMWF (https://www.ecmwf.int/), and hourly tide gauge data from the University of Hawaii Sea Level Center (UHSLC) (https://uhslc.soest.hawaii.edu/).

References

1. Mathers E L. An Intercomparison of Parametric Models of Sea State Bias for the TOPEX, Poseidon and Jason-1 Altimeters **126** (2003).
2. Guo J Zhang H Li Z Zhu C and Liu X. On Modelling Sea State Bias of Jason-2 Altimeter Data Based on Significant Wave Heights and Wind Speeds *Remote Sens.* **15** (2023).
3. Badulin S I Grigorieva V G Shabanov P A Sharmar V D and Karpov I O. Sea state bias in altimetry measurements within the theory of similarity for wind-driven seas *Adv. Sp. Res.* **68** (2021).
4. Zhang G Miao H Wang G Xin W and Jie Z. Study on parametric models of estimating the sea state bias based on the HY-2 altimeter *Proc. - 2015 8th Int. Congr. Image Signal Process. CISP 2015* Cisp (2016).
5. Miao H Wang X Wang G Zhang G and Zhang J. Study on satellite altimeter sea state bias estimation comprehensive model *Proc. 2015 Int. Conf. Electr. Comput. Eng. Electron.* **24** (2015).
6. Peng F and Deng X. Validation of Sentinel-3A SAR mode sea level anomalies around the Australian coastal region *Remote Sens. Environ.* **237** (2020).
7. Mostafavi M Delpeche-Ellmann N and Ellmann A. Accurate Sea Surface heights from Sentinel-3A and Jason-3 retrackers by incorporating High-Resolution Marine Geoid and Hydrodynamic Models *J. Geod. Sci.* **11** (2021).
8. Passaro M Nadzir Z A and Quartly G D. Improving the precision of sea level data from satellite altimetry with high-frequency and regional sea state bias corrections *Remote Sens. Environ.* **218** (2018).
9. Gaspar P Labroue S Ogor F Lafitte G Marchal L and Rafanel M. Improving nonparametric estimates of the sea state bias in radar altimeter measurements of sea level *J. Atmos. Ocean. Technol.* **19** (2002).
10. Peng F and Deng X. Improving precision of high-rate altimeter sea level anomalies by removing the sea state bias and intra-1-Hz covariant error *Remote Sens. Environ.* **251** (2020).
11. Zhong G Liu B Guo Y and Miao H. Sea State Bias Estimation with Least Absolute Shrinkage and Selection Operator (LASSO) *J. Ocean Univ. China* **17** (2018).
12. Laboue S *et al.*, . Nonparametric estimates of the sea state bias for the Jason-1 radar

altimeter *Mar. Geod.* **27** (2004).
13. Cheng Y Xu Q Gao L Li X Zou B and Liu T. Sea State Bias Variability in Satellite Altimetry Data *Remote Sens.* **11**, (2019).
14. Tran N *et al.*, . Sea state bias in altimeter sea level estimates determined by combining wave model and satellite data *J. Geophys. Res. Ocean.* **115** (2010).
15. Gaspar P Ogor F Le Traon P Y and Zanife O Z. Estimating the sea state bias of the TOPEX and POSEIDON altimeters from crossover differences *J. Geophys. Res.* **99** (1994).
16. Elfouhaily T and Thompson D R. Improved electromagnetic bias theory electromagnetic of a single-pass not satisfied , then Srokosz ' s theory will not be apl • lichble Glazman **105** (2000).
17. Tran, N.; Labroue, S.; Philipps, S.; Bronner, E.; Picot, N. Overview and update of the sea state bias corrections for the Jason-2, Jason-1 and TOPEX missions. Mar. Geod. **33** (2010).
18. Hou, K.; Zhang, S.; Kong, X. Quality assessment of HY-2A altimeter data through tide gauge comparisons. Haiyang Xuebao, **41** (2019).
19. Yuan, J.; Guo, J.; Niu, Y.; Zhu, C.; Li, Z. Mean Sea Surface Model over the Sea of Japan Determined from Multi-Satellite Altimeter Data and Tide Gauge Records. Remote Sens. **12** (2020).