

A Statistical Method to Quantify the Tide-surge Interaction Effects with Application in Probabilistic Prediction of Extreme Storm Tides along the Northern Coasts of the South China Sea

Abstract

Nonlinear interactions between tides and storm surges, i.e. the tide-surge interactions (TSI), play a critical role in modulating extreme coastal water levels. However, the TSI effects are often omitted in probabilistic modeling of extreme water levels, due to the extensive computational efforts caused by simulating numerous combinations of storm timings and random tidal phases. In this study, we proposed a statistical approach for fast estimation of TSI effects with applications in predicting extreme water levels. Storm tides in the northern expanse of the South China Sea (SCS) during typhoon events from 1979 to 2020 are simulated, based on which contributions of TSI to the total water level are extracted and analyzed. The results reveal a predominant negative influence of TSI on the total storm tide levels. It was found that for a given coastal location, statistically significant correlations exist between TSI divided by the peak storm tide and surge divided by the peak storm tide. Based on the correlations, location-dependent, multiple regression models for predicting TSI effects on storm tide levels were established using predicted surge and tidal levels as input, which avoids modeling numerous scenarios of different tidal phases and storm timings. By comparing with long term history data from a tide gauge, the proposed approach was shown to be able to reproduce the TSI contributions to extreme storm tides accurately and efficiently. By superposing the estimated TSI effects with predicted surge and random tides, a method for calculating the extreme water level with certain return periods is developed. Finally, an application of the method was demonstrated by calculating the 50-year return period water levels along the northern coast of the South China Sea.

Keywords: Storm tides; tide-surge interaction; extreme water level; South China Sea;

risk assessments

1. Introduction

Storm tide is the water level rise due to the combination of the storm surge and the astronomical tide during tropical/extratropical cyclones (NOAA, 2019). Storm tides are among the most disastrous natural events in the coastal zone, which can flood the low-lying areas, damage coastal and offshore infrastructures and erode the shorelines (Wu et al., 2018; Zeng et al., 2021). Annually, these events lead to property damage and casualties reaching billions to hundreds of billions of dollars in countries and regions worldwide that experience frequent storm surge occurrences (Mas et al., 2015; Muis et al., 2019; Needham et al., 2015; Yang et al., 2019).

As one of the most important environmental design parameters for coastal and offshore engineering, return-period sea levels are usually estimated by extrapolation of observed storm tide extremes (e.g., annual maxima) using extreme value theories, therefore heavily relies on history data. Nevertheless, the sparse distribution of tide gauges, along with limited duration of observations, poses limitations on accurately estimating extreme water levels (Marsooli and Lin, 2018; Muis et al., 2016). Some researchers use synthetic TC tracks in addition to or instead of historical TC tracks to overcome the limitation of insufficient storm samples (Dullaart et al., 2021; Khan et al., 2022; Leijnse et al., 2022; Li et al., 2018; Martín et al., 2023). However, storm tides usually deviate from the linear superpositions of storm surges and tidal levels, due to the intricate nonlinear interactions between them. Furthermore, owing to the independence between the astronomical tides and the storm events, peak surges can occur at any tidal phases, thus it is necessary to consider all possibilities of combining surge timings and tidal phases in a probabilistic prediction of extreme sea levels. The above-mentioned method leads to a substantial increase by several orders of magnitude in the overall count of numerical experiments, resulting in a noteworthy rise in computational costs. Some scholars omitted the nonlinear TSI and combined storm surge driven by stochastic TCs with random astronomical tides in a linear manner to

obtain peak water levels, thus enabling the estimation of storm tide return periods (Dullaart et al., 2021; Orton et al., 2016). However, omitting the TSI effects may lead to potential underestimation or overestimation of extreme sea levels (Lin and Chavas, 2012; Rego and Li, 2010; Wuxi et al., 2018). Recent studies demonstrated that disregarding the TSI effects can result in up to 30% differences in estimated extreme sea levels (Arns et al., 2020).

The fundamental physics behind nonlinear TSI lie in the mutual phase alterations between astronomical tides and storm surges (Flather, 2001; Zhang et al., 2010). In other words, high and low tides induce changes in storm surges via alterations in water depth. Conversely, storm surges impact the propagation velocity of tides by modifying water depth, thereby causing a phase shift in astronomical tides (Bernier and Thompson, 2007; Rossiter, 1961). Due to the presence of TSI, the maximum storm surge consistently mismatches the moments of high and low tides (Horsburgh and Wilson, 2007), and shows a higher occurrence frequency during rising tides (Idier et al., 2012; Rossiter, 1961; Zhang et al., 2017). Moreover, this interaction significantly amplifies negative surge amplitudes while reducing positive surge amplitudes (Prandle and Wolf, 1978; Wankang et al., 2019). Moreover, TSI exhibits pronounced spatial variations, particularly in areas characterized by shallow water depths, large tidal ranges, gentle slopes, and concave shorelines (Costa et al., 2023; Idier et al., 2012; Rego and Li, 2010; Rossiter, 1961; Zhang et al., 2017). TSI is primarily governed by three factors: the nonlinear bottom friction effect (Davies and Lawrence, 1994; Tang et al., 1996; Yang et al., 2023; Zhang et al., 2017; Zhang et al., 2010; Zheng et al., 2020), the shallow water effect (Bernier and Thompson, 2007; Wolf, 1981; Zheng et al., 2020), and the nonlinear advection effect (Hu et al., 2023; Rego and Li, 2010; Wankang et al., 2019; Yang et al., 2023; Zheng et al., 2020). The contributions of these factors to TSI exhibit variability in response to changes in water depths, astronomical tides, and storm surge dynamics (Zhang et al., 2010).

Previous studies have demonstrated the indispensability of considering TSI when assessing extreme storm tides (Arns et al., 2020; Costa et al., 2023; Marsooli and Lin,

2018; Ragno et al., 2023). However, investigations into TSI have predominantly concentrated on feature analysis, contribution evaluations, and physical mechanism, while rarely providing a practical method that can be readily applied to estimating extreme water levels (EWLs). In the present study, firstly, we simulated the typhoon-induced storm surges, storm tides and the astronomical tides in the northern South China Sea (SCS) spanning the years 1979 to 2020. Based on the model results, we analyzed the contributions of TSI to the storm tides. The primary objective is to develop a statistical approach to quantify the effects of TSI on storm tides, and apply the quantification method to estimate extreme water levels with certain return periods (RPs), which are essential environmental parameters for coastal and offshore structure design.

The structure of this paper is as follows. Section 2 provides a description of numerical models used, including details on parameter configuration, model validation, and the method employed to isolate TSI. Section 3 elaborates on the statistical quantification approach for TSI, presents an overview of extreme value statistical methods, and outlines the methodology for estimating storm tides RPs with a narrower 95% confidence interval. In Section 4, a comprehensive summary is presented regarding the spatiotemporal distribution characteristics of TSI along the northern coast of the South China Sea (SCS). The estimated TSI effects on the extreme storm tides is validated, and the method proposed in this article is utilized to calculate the exceedance probability curves for several grid points. Section 5 presents the estimation of 50-year return period (RP50) storm tides in the northern area of the SCS, and engages in a thorough discussion on the application and limitations of proposed methodology in this study. Finally, the conclusions of this study are summarized in Section 6.

2. Data and Model

2.1 Model setup

This study focuses on the typhoon-induced storm tides in the coastal regions of northern SCS. Based on the best track data of Japan Meteorological Agency, TC tracks from 1979~2020 that have been at least within 200 km from the study area are

considered and modeled (Figure 1a).

The FVCOM (Chen et al., 2003) model was used to model the storm tides caused by the 416 TCs affecting the target area (Figure 1a). The model encompasses the entire SCS with an unstructured mesh, stretching from the Vietnam coasts to the eastern coasts of the Philippines. The mesh extends from Zhejiang Province in the north and reaches Malaysia and northern Indonesia in the south. The longitudinal range of the model spans from 99°E to 128°E, while the latitudinal range covers from 1°N to 29°N (Figure 1b). Model mesh with grid resolution, and bathymetry are depicted in Figure 1b-c.

The mesh consists of triangular grids with 97,814 nodes and 183,610 elements. Grid resolution is around 500 m along the Chinese mainland coast, and gradually decreases to 5 km and 15 km at water depths of 20 m and 50 m, respectively. Resolution coarsens to 30~50 km in the ocean areas. The bathymetric data are derived from the combinations of digitized nautical charts and GEBCO (Macnab, 2003), with the depth reference plane adjusted to the local mean sea level. The model utilizes the global tidal model FES2014 (Carrere et al., 2015) for the tidal boundary conditions. Eight major tidal constituents (S2, M2, K1, K2, O1, N2, P1, and Q1) are used. The average tidal range (representing the mean values of simulated astronomical tidal ranges over a span of 18.6 years) are also shown in Figure 1d.

2.2 Meteorological forcing

The ERA5 reanalysis wind product (Hersbach et al., 2018) by the European Centre for Medium-Range Weather Forecasts (ECMWF), which has a temporal resolution of 1 hour and a spatial resolution of $0.25^{\circ} \times 0.25^{\circ}$, was used to provide the air pressure and peripheral wind field beyond 7 times the radius to the maximum wind. The Holland parametric TC wind model (Holland, 1980) was used to derive the wind field within twice the radius to the maximum wind speed. In between, the two wind fields are blended proportionally (Willoughby and Rahn, 2004). The TC track information is sourced from the Japan Meteorological Agency's Regional Specialized Meteorological Centers (RMSCs) in Tokyo (<https://www.data.jma.go.jp/yoho/typhoon/index.html>). This dataset includes the typhoon's center position, central pressure, and 10-minute

maximum wind speed, which can be used as input data for the Holland wind model.

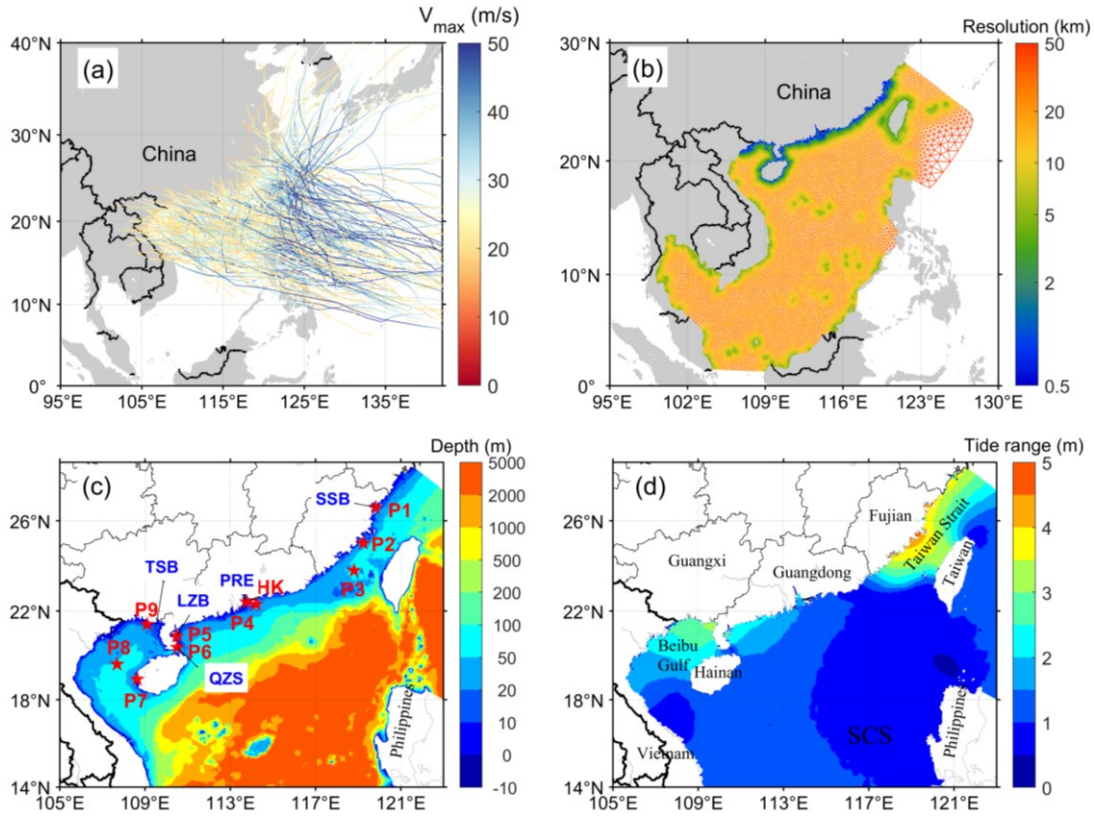


Figure 1. (a) Typhoons that affected the northern SCS from 1979 to 2020. (b) Computational mesh and resolution. (c) Bathymetry map of the northern SCS. SSB, Sansha Bay; PRE, Pearl River Estuary; LZB, Leizhou Bay; TSB, Tieshan Bay; QZS, Qiongzhou Strait. (d) Average tidal range map of the northern SCS.

2.3 Quantitative error metrics

We employ several metrics, including root mean square error (RMSE), average absolute error (AAE), mean relative error (MRE), and Pearson correlation coefficient (CC), to assess the model performance. RMSE quantifies the standard deviation between the true and the estimated values. AAE is the average of the absolute deviation between the true and the estimated value. MRE measures the absolute difference between the true and the estimated value, normalized by the absolute true value. The CC value, ranging from 0 to 1, measures the linear correlation between observed and simulated values. The formulas for these metrics are as follows:

$$RMSE = \left[\sum_1^n (M_i - O_i)^2 / n \right]^{0.5} \quad (1)$$

$$AAE = \sum_1^n |M_i - O_i| / n \quad (2)$$

$$MRE = \frac{1}{n} \sum_1^n |M_i - O_i / O_i| \quad (3)$$

$$CC = \sum_1^n (M_i - \bar{M}_i)(O_i - \bar{O}_i) / \sqrt{\sum_1^n (M_i - \bar{M}_i)^2} \sqrt{\sum_1^n (O_i - \bar{O}_i)^2} \quad (4)$$

where M_i represents the i th estimated value, O_i represents the i th true value. The overbar signifies the averaged value across all data points. A perfect correspondence between model outcomes and measurements yields RMSE, AAE and MRE values of zero, accompanied by a CC value of one. 2.4 Model validation

Validations of wind, air pressure, tides and storm surges are conducted by comparing to multiple-sourced data. Data sources include the University of Hawaii Sea Level Center (UHSLC, <https://uhslc.soest.hawaii.edu/>), the National Science & Technology Infrastructure - National Marine Data Center (<http://mds.nmdis.org.cn/>), and the "Collection of Storm Surge Disasters Historical Data in China 1949-2009". Validation results in Appendix A demonstrate the accuracy in simulating the tidal dynamics and the storm surges under the TC meteorological forcing, making it suitable for subsequent research on the influence of TSI.

3. Method

3.1 Numerical experiments design

Three numerical experiments were conducted using the method similar to Idier et al. (2012). These experiments include EXP-PS (pure surges, considering only wind and air pressure as model forcing), EXP-PT (pure tides, considering only astronomical tidal forcing), and EXP-ST (storm surge and tides fully coupled, considering both meteorological and tidal forcing). The nonlinear TSI effects on the water levels (TSIWL) can be quantified by:

$$\eta_l = \eta_{st} - \eta_s - \eta_t \quad (5)$$

where pure storm surge η_s , astronomical tide η_t , and total storm tide η_{st} .

3.2 Statistical dependence and multiple regression model

Aiming to seek an effective indicator of the magnitude of the TSI effects, we first analyze the dependencies between several variables including η_t , $\eta_{lp} = \eta_t / (\eta_s + \eta_t)$ and $\eta_{sp} = \eta_s / (\eta_s + \eta_t)$, η_s , η_t , $\eta_s + \eta_t$ corresponding to the extreme storm tides. The Spearman's rank correlation coefficient ρ was employed to assess the strength of the correlation between two variables based on the monotonic relationship,

$$\rho = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)} \quad (6)$$

where d_i represents the disparity between the ranks of each value, n denotes the total number of values. The Spearman's rank correlation assumes values ranging from +1 to -1, where a score of +1 indicates a flawless correspondence of ranks, a score of 0 implies no association between ranks, and a score of -1 signifies a complete negative relationship among ranks.

The thermal diagram depicting the Spearman correlation coefficient matrix among the parameters is presented in Figure 2a. It is apparent that the absolute values of the rank correlation coefficients for variables η_t , η_{lp} and η_t , $\eta_s + \eta_t$ are all below 0.4, indicating a lack of significant correlations. However, noteworthy correlations are observed between variables η_t , η_{lp} and η_s , η_{sp} , with coefficients reaching approximately -0.7, signifying a substantial relationship. It is important to note that the rank correlation coefficients, while effective in capturing monotonic relationships between variables, presents limitations when it comes to discerning non-monotonic relationships. To identify variables more suited for establishing statistical relationships with TSI, we take point P6 (Figure 1c) as an example and generate scatter plots comparing the variables η_t with η_s and η_{lp} with η_{sp} (Figure 2b and 2c). Analysis of these plots reveals a distinct pattern: during the initial half of the curve, both variables exhibit a decreasing trend in TSI as the x-axis increases. However, in the latter half, no clear correlation is evident between η_t and η_s , while an increasing trend is observed between η_{lp} and η_{sp} . This observation underscores the effectiveness of variable η_{sp} .

as an indicator for quantifying TSI. The correlation between η_{lp} and η_{sp} in Figure 2a indicates that when astronomical tides dominates in extreme storm tides, TSIWL increases with the increase of storm surges. However, when storm surges dominate, the influence of the astronomical tides diminishes, resulting in a weakening of the corresponding nonlinear effects.

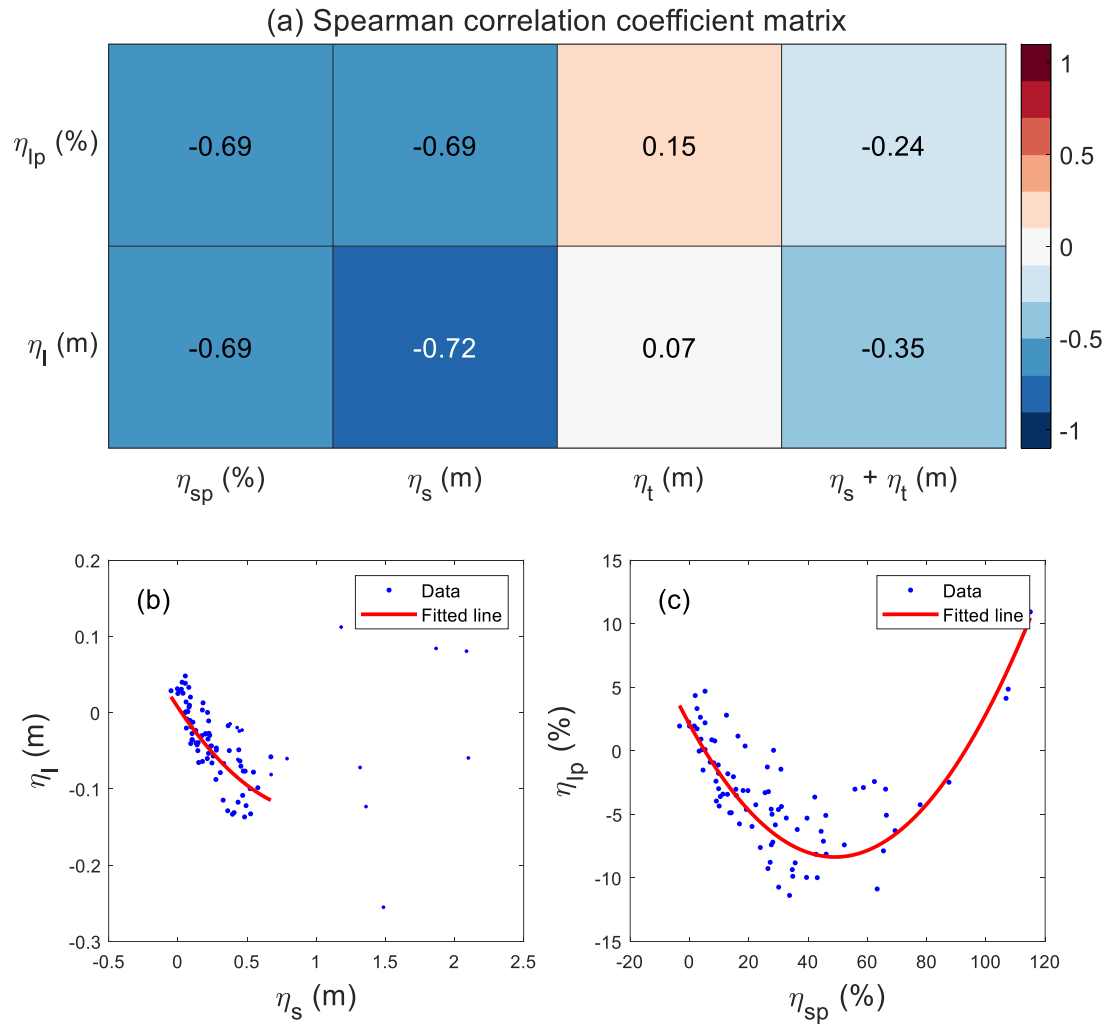


Figure 2. (a) Spearman correlation coefficient matrix. (b) Scatter plots between η_l and η_s at the peak storm tide at P6. (c) Scatter plots between η_{lp} and η_{sp} at the peak storm tide at P6.

Owing to the regulatory influence of TSI on extreme water levels, the temporal occurrences of $(\eta_s + \eta_t)_{\max}$ and $\eta_{st,\max}$ often do not align. After analyzing the model results, it is found that the time difference between the two peaks mostly falls into a two-hour window as shown in Figure B.1 in Appendix B. We aim to estimate the maximum storm tide $\eta_{st,\max}$ based on the known value $(\eta_s + \eta_t)_{\max}$. Consequently,

overlooking the errors arising from a minor temporal discrepancy between $(\eta_s + \eta_t)_{\max}$ and $\eta_{st,\max}$, we denote $\eta_t = \eta_{st,\max} - (\eta_s + \eta_t)_{\max}$, $\eta_{lp} = \eta_t / (\eta_s + \eta_t)_{\max}$, $\eta_{sp} = \eta_s / (\eta_s + \eta_t)_{\max}$ in this context. To construct a multiple regression model, we conduct a sample fitting for each grid point within the nearshore area, establishing a relationship between η_{lp} and η_{sp} . The model is represented by the following equation:

$$\eta_{lp} = a \times \eta_{sp}^3 + b \times \eta_{sp}^2 + c \times \eta_{sp} + d \quad (7)$$

Note that the regression model differs for different location (grid points), i.e., the model is location-dependent. For each grid point, the sample data used to fit the equation comprises of only storm tides generated by typhoons travelling within 200 km from the grid point.

3.3 TSIWL estimation and validation

After establishing the regression statistical relationships for each point, the estimation of TSIWL can be derived by utilizing the individually calculated η_s and η_t . The following steps outline the procedure:

- a) superimpose the modeled pure surge η_s onto the modeled synchronous astronomical tide η_t ;
- b) determine the peak water level $(\eta_s + \eta_t)_{\max}$ and its corresponding time; calculate η_{sp} based on η_s and η_t corresponding to the time of $(\eta_s + \eta_t)_{\max}$;
- c) substitute η_{sp} into the statistical relationship equation (7) to obtain an estimate of η_{lp} ;
- d) employ equation (8) to estimate TSIWL $\eta_{I,est}$ corresponding to the peak storm tide.

$$\eta_{I,est} = \eta_{lp} \times (\eta_s + \eta_t)_{\max} \quad (8)$$

The method was implemented across all grid points within the nearshore region to obtain the TSIWL. The estimated TSI are then compared with the modeled results to assess the performance of the proposed method. Performance metrics such as RMSE, AAE, and Pearson CC are computed to evaluate the accuracy and reliability. The computational grid points in the nearshore area are segregated into nine groups (P1-P9): Sansha Bay, Fujian Coast, Pearl River Estuary, Guangdong Coast, Leizhou Bay, Hainan Coast, Beibu Gulf, and Guangxi Coast. The estimation of TSIWL in these nine regions

is validated individually. Additionally, the calculated $\eta_{I,est}$ is added to $(\eta_I + \eta_s)_{\max}$ to obtain the estimated maximum storm tide. The estimated maximum storm tide is then compared against observed values to validate its accuracy.

Data from the Hong Kong (HK) tidal station (Figure 1a), which is one of the few stations that have long-term water level observations covering the 1979~2020, are used to validate the proposed method. We compared the exceedance probability curves calculated from observed data and simulated samples $(\eta_{I,est} + (\eta_I + \eta_s)_{\max})$ for validation. Both simulated and observed storm tides samples are fitted to the Generalized Pareto Distribution (Hosking and Wallis, 1987) to derive the exceedance probability curve for typhoon-induced storm tides. It is important for the tidal station records to cover the same time span as the simulations, encompassing the typhoons between 1979 and 2020.

3.4 Estimating storm tide return periods

The peaks-over-threshold (POT) method is used to sample extreme storm tides exceeding a certain value (Davison and Smith, 1990; Liang et al., 2019; Walton Jr, 2000). The extreme value samples are then fitted to the Generalized Pareto Distribution (GPD) to compute the exceedance probability curves. The GPD function is defined as follows:

$$F_u(x) = 1 - \left(1 + \xi \frac{x - \mu}{\sigma} \right)^{-\frac{1}{\xi}}, \quad x \geq \mu, \quad 1 + \xi(x - \mu) / \sigma > 0 \quad (9)$$

In this case, $\mu \in R$ denotes the location parameter, $\xi \in R$ represents the shape parameter, and $\sigma > 0$ signifies the scale parameter. It is generally accepted that in a specific region, the frequency of typhoon-induced storm tides (λ_μ) surpassing a threshold μ per year follows a Poisson process (Vickery et al., 2000). The 1/p-year return value, denoted as z_p , for such events analyzed using the POT/GPD method, is determined by the following equation:

$$z_p = \mu + \frac{\sigma}{\xi} \left[(\lambda_\mu p)^\xi - 1 \right] \quad (10)$$

When an ample number of samples are available, the calculation of extreme storm tides can be accomplished without resorting to distribution fitting. Instead, the empirical

cumulative frequency method can be employed. This method involves arranging the final extreme storm tide samples in descending order, calculating the cumulative frequency for each sample, plotting them on a probability coordinate paper, and ultimately directly determining the storm tide level for the corresponding RP. The formula for calculating the empirical cumulative frequency is as follows:

$$P_{exc}(\eta) = \frac{i}{n+1} \times \frac{n}{m} \quad (11)$$

$$T(\eta) = 1/P_{exc}(\eta) \quad (12)$$

In the given formulas, $P_{exc}(\eta)$ represents the probability of exceeding a specific extreme tidal level η at a given rank i . Here, n denotes the number of samples, m represents the length of the samples in years, and the return period $T(\eta)$ is the reciprocal of the exceedance probability $P_{exc}(\eta)$.

The Kolmogorov-Smirnov test is used to evaluate the fit goodness for the GPD model, calculating the probability value (p-value) associated with rejecting the null hypothesis. When the p-value exceeds 0.05, we consider the goodness of fit to be significant at a 0.05 confidence level, indicating the reliability of the estimates obtained through the extreme value distribution.

Procedures for estimating extreme storm tide RPs using simulated storm surges during typhoons can be outlined in the following five steps.

- a. Simulate typhoon-induced pure storm surges. Since this study only considered historical typhoons from 1979 to 2020, we regard the historical typhoons that impacted the northern region of the SCS as a short-term stochastic typhoon dataset.
- b. Simulate astronomical tides lasting for 18.6 years to encompass the entire tidal cycle. The simulated astronomical tide in this study encompasses the time period from April 1, 2002, to December 31, 2020.
- c. The storm surges induced by typhoons are combined with randomly generated astronomical tides through linear superposition. This approach aims to encompass a wide range of storm surge and astronomical tide combinations that have not been observed in historical records. To achieve this, each set of storm surge sequences is paired with 500 sets of randomly generated astronomical tides ("500" here refers to

the number of astronomical tides samples. The stability analysis of the astronomical tides sampling quantity will be discussed in Section 5.1). Similar methodologies have been employed in previous studies, such as Orton et al. (2016) who combined simulated historical storm surges with 50 random astronomical tides and Dullaart et al. (2021) who combined simulated storm surges with 1000 random astronomical tides. During this process, the timing of astronomical tides combined with storm surge is extracted using the Monte Carlo method, taking into account the seasonal characteristics of historical typhoon formation. Each selected time series of astronomical tides matches the length of the storm surge data. Assuming there are n typhoons, conducting this procedure yields $500 \times n$ segments of water level sequences, corresponding to a time span of 42×500 years.

- d. Estimate the TSIWL (refer to Equation [7]) and incorporate it in the total water level. From the water level time series obtained in the previous step, we can extract $500 \times n$ distinct peaks. For each peak moment, the TSIWL $\eta_{l,est}$ was estimated based on the surge η_s and tide η_t (specific methods shown in section 3.2). Subsequently, $\eta_{l,est}$ is combined with $(\eta_s + \eta_t)_{\max}$, resulting in $500 \times n$ samples of extreme storm tides.
- e. Determine the threshold and employ GPD to fit the samples surpassing the threshold, thereby obtaining the exceedance probability curve of storm tide. Subsequently, calculate the corresponding confidence interval for the exceedance probability. In scenarios where an ample number of samples are available, the empirical distribution method can also be utilized to directly determine the storm tides RPs.

We used the aforementioned method to calculate the return periods of extreme storm tides at the HK station and compare them with the extreme water levels calculated based on measured data. We selected hourly water level records from the HK station from 1979 to 2020. To ensure independence, the samples were extracted within a 72-hour window centered around each peak in the observed tide levels. The POT/GPD method is then utilized to estimate the RPs of extreme sea levels at the HK station.

The method proposed in this article incorporates TSIWL into the evaluation of

extreme storm tides, thereby improving the accuracy of the calculations. Additionally, a larger sample size ensures a narrower confidence interval for the computed results, ensuring their robustness. The framework employed in this paper to quantify TSIWL and estimate the RPs of storm tides can be divided into three main sections: numerical modeling, TSIWL estimation, and extreme value analysis. The flow chart is shown in Figure 3.

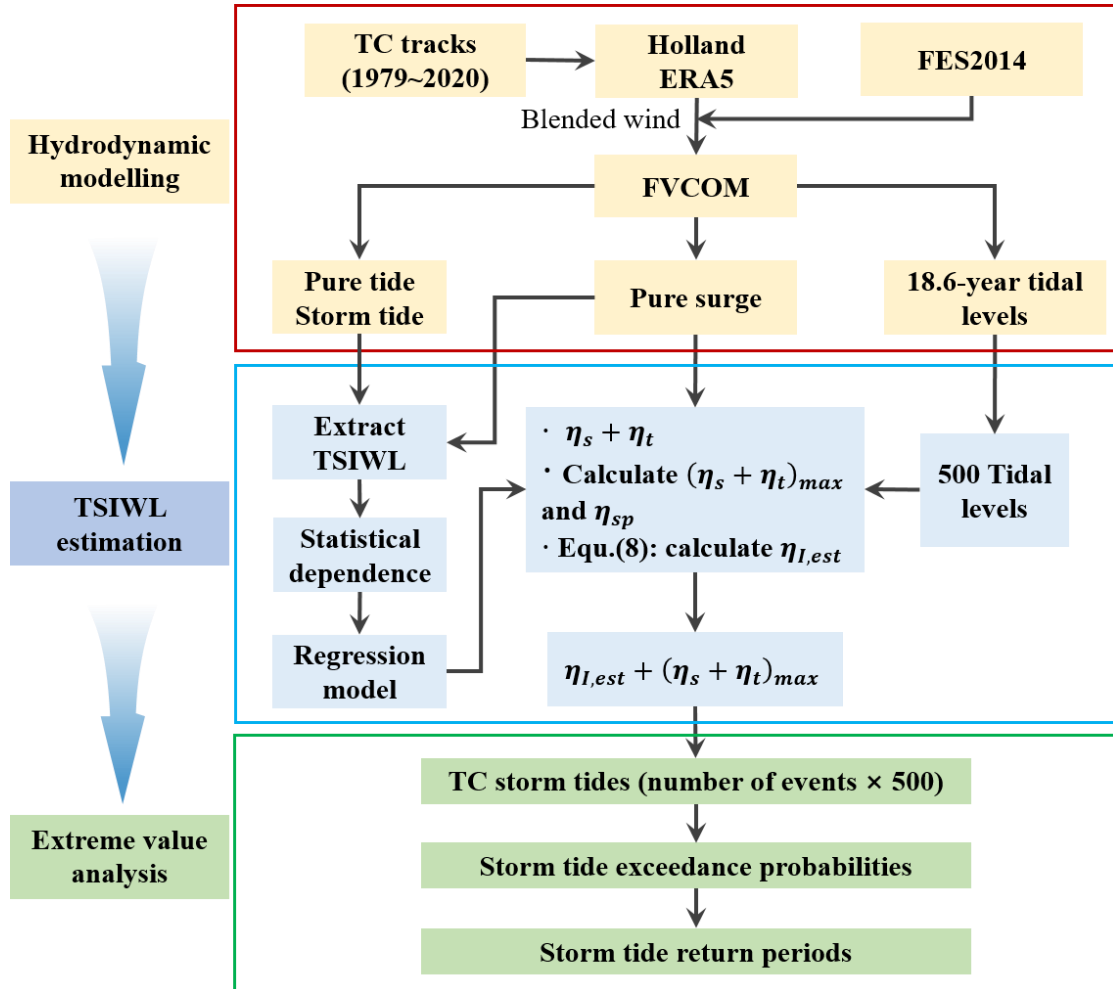


Figure 3. Method framework. Flowchart showing the hydrodynamic modeling (red box), TSIWL estimation (blue box) and extreme value analysis (green box).

4. Results

4.1 General characteristics of historical extreme storm tides

For a storm tide event, if the proportion of astronomical tides surpasses 50%, it is classified a tide-dominated event; otherwise, it is categorized as a surge-dominated

event. Figure 4a displays the percentage of astronomical tides within the 98th percentile of extreme storm tides, representing the top 2% of storm tides arranged in descending order. Notably, for the majority of areas encompassing the Fujian Coast, Guangxi Coast, Beibu Gulf, and the Taiwan Strait, the contribution of astronomical tides in the 98th percentile of EWLs exceeds 50%, signifying the dominance of astronomical tides in storm tides. Conversely, across the majority of grid points adjacent to the Guangdong Coast, the eastern coast of Hainan, and the Qiongzhou Strait area, the influence of astronomical tides in the 98th percentile of extreme storm tides is found to be below 50%, suggesting a prevalence of surge-dominated events. Figure 4b illustrates the ratio of events where the EWLs are primarily influenced by astronomical tides, compared to the total number of events, at each grid point. In regions such as the Beibu Gulf and Taiwan Strait, astronomical tides play a dominant role in all extreme water levels, while in the coastal areas of Guangdong, 10%~30% of extreme water levels are dominated by surges.

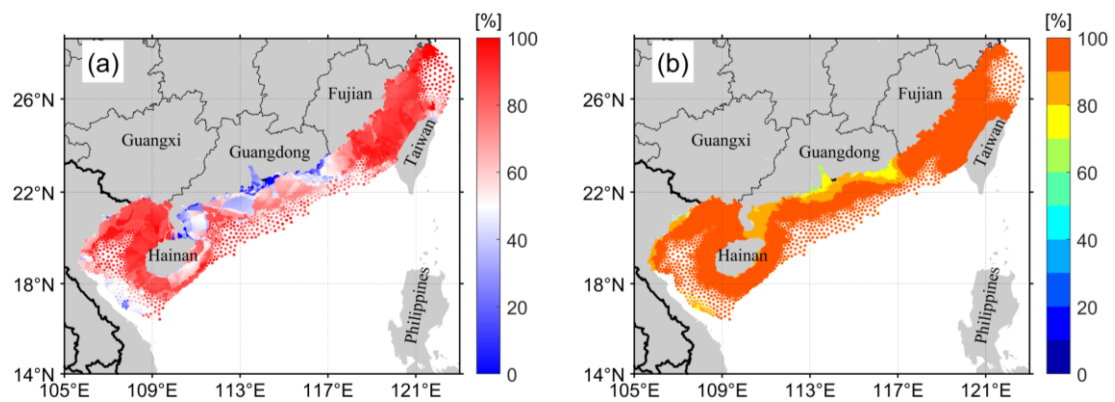


Figure 4. (a) The percentage of astronomical tides in the 98th percentile of extreme storm tides. (b)

The proportion of events dominated by tides at each grid point, relative to the total number of events. Only grid points near the northern SCS with a water depth below 200 m are displayed.

4.2 Characteristics of TSIWL

Figure 5a-d depict the maximum negative and positive water level due to TSI, as well as the maximum negative and positive water level induced by TSI during the peak storm tide $\eta_{st,max}$. It can be observed that TSI exerts significant influence on storm tides in bays and estuarine areas. For instance, in the Sansha Bay of Fujian, the maximum

negative water level induced by TSI exceeds -0.9 m, while the maximum positive water level induced by TSI exceeds 0.6 m. Near Xiamen Island of Fujian, the maximum negative water level induced by TSI exceeds -0.8 m, while the maximum positive water level surpasses 0.6 m. In the Pearl River Estuary, Leizhou Bay, and Tieshan Bay, the influence of TSI is relatively weaker. However, even in these areas, the maximum negative water level induced by TSI still reaches -0.7 m, while the maximum positive water level reaches 0.5 m. Moreover, substantial TSI is observed in most areas of the Taiwan Strait and Beibu Gulf, where the maximum negative water level induced by TSI ranges from -0.7 m to -0.3 m, and the maximum positive water level ranges from 0.2 m to 0.5 m. The spatial distribution characteristics of the nonlinear water levels induced by TSI bear resemblance to the patterns observed in the distribution of tidal range and bathymetry (Figure 1a-b). As the water depth decreases and the tidal range increases, the influence of TSI becomes more pronounced. The amplification of tides within bays and the enhancing effect of surges contribute to the heightened influence of TSI. During the extreme storm tide, TSI predominantly exerts a negative influence (Figure 5c-d). Among all events, approximately 74% exhibit negative contributions from TSI to the maximum storm tide. However, the magnitude of such negative contributions is noticeably smaller compared to the maximum positive and negative water levels induced by TSI throughout the storm tide process (Figure 5a-b). This indicates that during the peak storm tide $\eta_{st,max}$, nonlinear water levels induced by TSI may not necessarily reach the highest levels, but they still constitute a significant proportion of the EWLs.

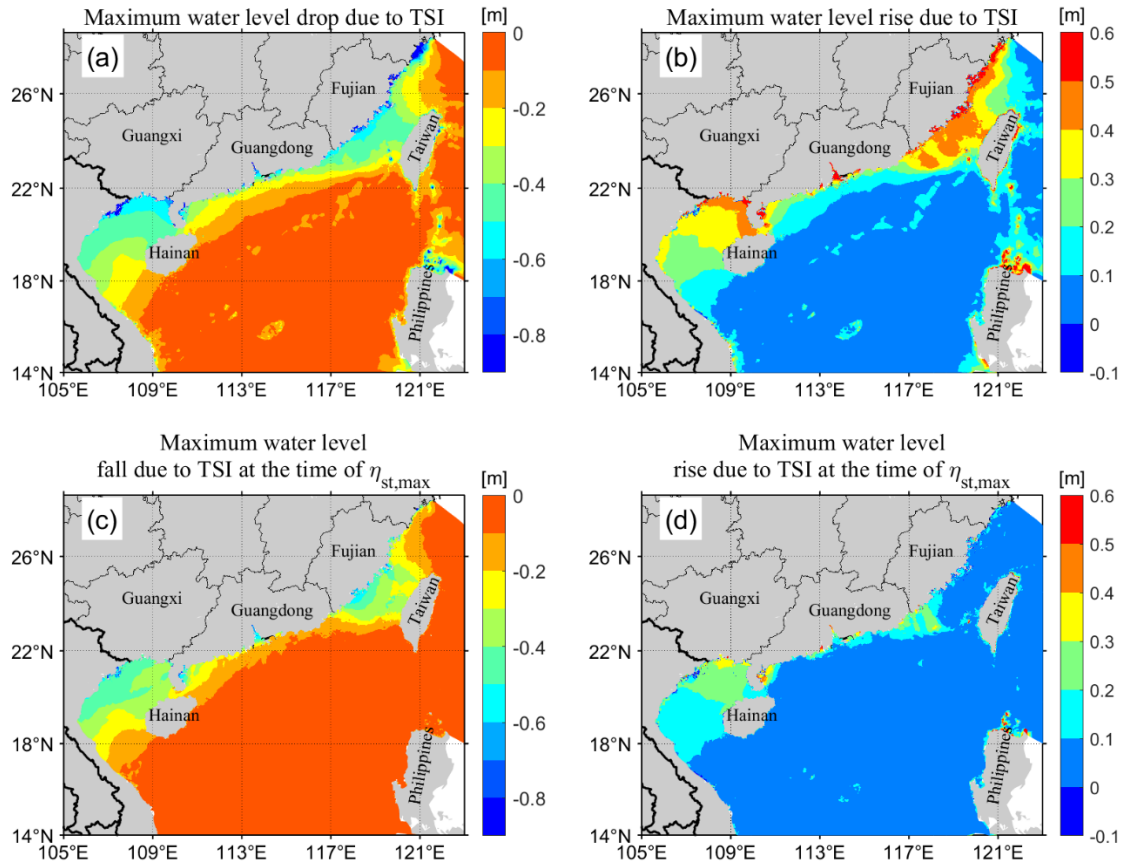


Figure 5. (a) Maximum negative water level induced by TSI. (b) Maximum positive water level induced by TSI. (c) Maximum negative water level induced by TSI at the time of maximum storm tide $\eta_{st,max}$. (d) Maximum positive water level induced by TSI at the time of maximum storm tide

$$\eta_{st,max}$$

4.3 Estimation and validation of TSIWL

We establish statistical relationships between η_{lp} and η_{sp} for each of the nearshore grid points and determine the regression models. Along the significant coastal bays and estuaries of Northern SCS (including Sansha Bay, Taiwan Strait, Pearl River Estuary, Leizhou Bay, Qiongzhou Strait, Beibu Gulf, Guangxi Coast, and the western side of Hainan Island), nine sample points (P1~P9) have been selected (Figure 1). Figure 6 illustrates the scatter plots of variables η_{lp} and η_{sp} (blue dots) and the fitting effect of the multiple regression equation (red line) for these nine sample points. Furthermore, the determination coefficient (R^2) is calculated to quantify the goodness of fit. Among these selected points, P7 exhibits the shallowest water depth at 5 m, while P8 has the greatest water depth at 57 m. The determination coefficient ranges from 0 to

1, with higher values indicating a favorable model fit. Analyzing the graph reveals that for P1, P2, P3, and P7, the η_{sp} values of all storm tide events are below 0.5, indicating dominance of astronomical tides, and η_{lp} decreasing as η_{sp} increases, exhibiting negative monotonic correlation between the two. On the other hand, for the remaining points, where η_{sp} of certain storm tide events exceeds 0.5, indicating dominance of storm surge in EWL, and η_{lp} initially decreases and then increases with the increasing value of η_{sp} . The determination coefficients of P2-P9 exhibit greater than 0.6, indicating a good fit, whereas P1 has a determination coefficient of only 0.41, suggesting a slightly weaker fit.

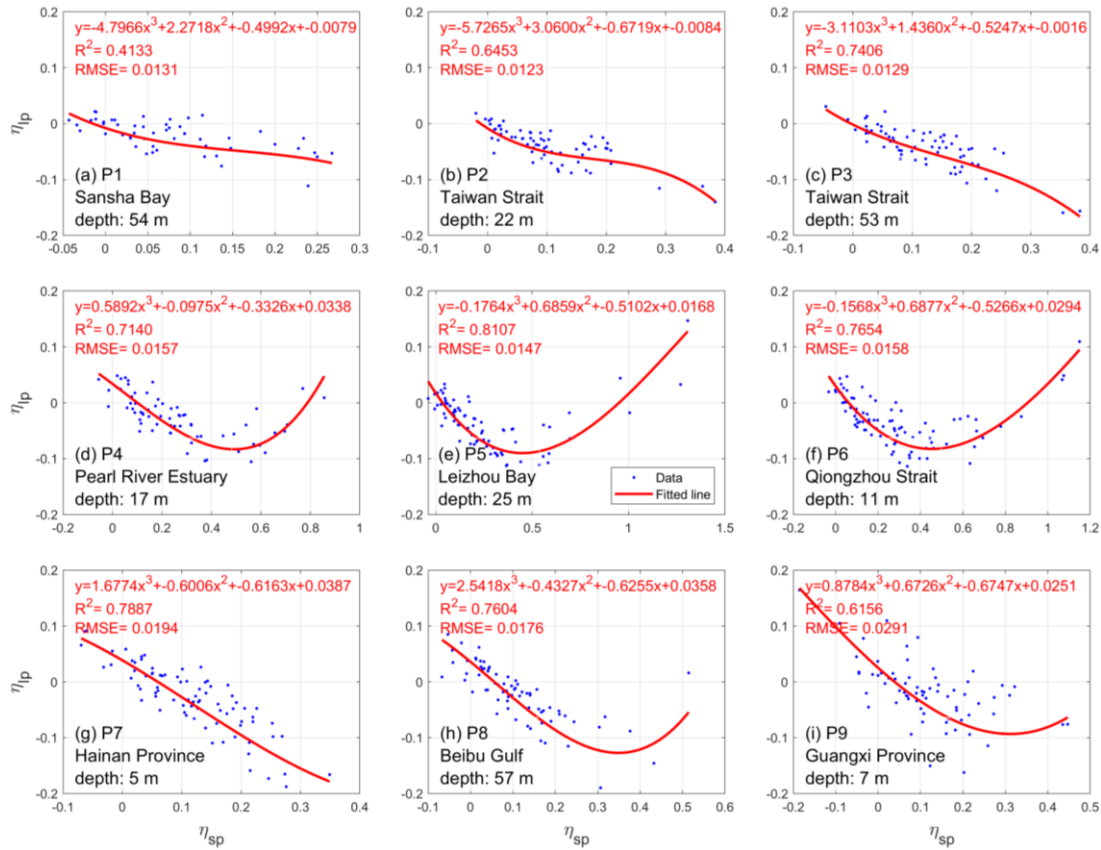


Figure 6. The dependency relationship between η_{lp} and η_{sp} at P1-P9, along with the fitting effect of multiple regression equation.

Upon establishing individual multiple regression models for grid points with water depths below 200 m along the northern coastal region of the SCS, TSIWL can be estimated by separately simulating pure surge and tides. By applying the proposed method to all nearshore grid points, the TSIWL $\eta_{I,est}$ at the time of maximum storm

tides $\eta_{st,max}$ is estimated and compared with the original results derived from numerical modeling. Various error evaluation metrics, including RMSE, AAE, and Pearson CC, are computed. The nearshore grid points are categorized into nine distinct regions based on important bays, estuaries, and administrative divisions: P1~P9, Sansha Bay, Fujian Coast, Pearl River Estuary, Guangdong Coast, Leizhou Bay, Hainan Coast, Beibu Gulf, and Guangxi Coast. Validation of the estimated TSIWL is conducted individually for the nine regions mentioned (Figure 7).

The scatter density plot depicted in Figure 7 reveals that, apart from a few scattered points exhibiting noticeable overestimation or underestimation, the majority of data points representing the estimated TSI demonstrate a concentrated distribution along the 1:1 diagonal line. Dense clusters of data points are observed within the range of -0.3 to 0 m. The magnitude of errors is influenced by the strength of TSI, whereby areas with more prominent TSI, such as Sansha Bay, exhibit larger errors, whereas regions with weaker TSI, such as the coastal areas of Guangdong, exhibit smaller overall errors. The RMSE of the overall estimation ranges from 4.06 to 5.49 cm, the AAE ranges from 2.88 to 3.95 cm, and the Pearson correlation coefficient fluctuates between 0.81 and 0.84.

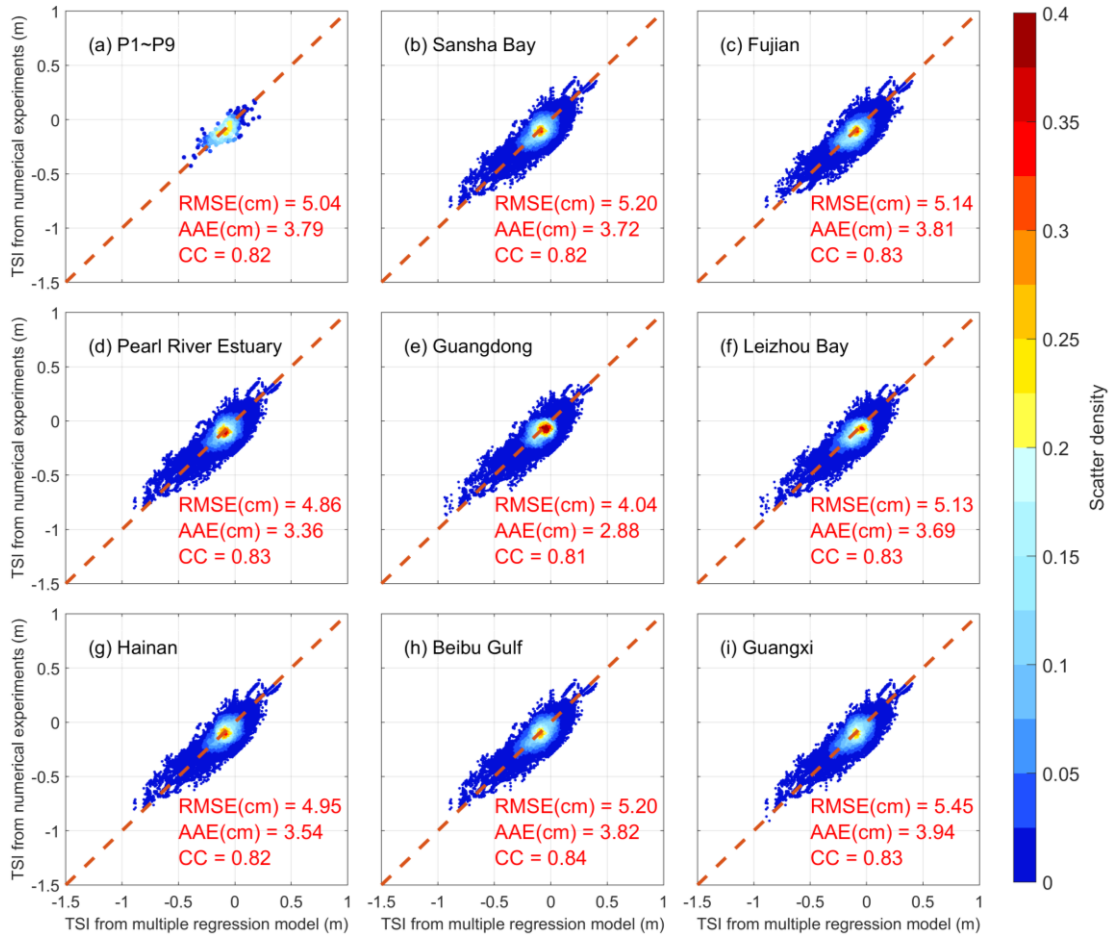


Figure 7. Scatter density map of TSIWL computed by numerical model simulation and multiple regression model.

4.4 Validation with data at the HK station

To validate the method, a linear superposition of the estimated $\eta_{I,est}$ and $(\eta_s + \eta_t)_{max}$ was performed, and the superposition results were compared with the observed data. We selected the observed tidal levels recorded during the typhoon periods from 1979 to 2020 at the HK station. Using the POT/GPD method, the exceedance probability curve for typhoon-induced storm tides at the HK station was calculated. The flood exceedance probability curves based on simulated and measured samples are compared to verify the estimated TSIWL.

In Figure 8, it is evident that the blue sample (ignoring TSI) exhibits a notably higher elevation of approximately 5~20 cm compared to the red, orange, and green samples (include TSI). The resemblance between the red dots (observed data) and the orange dots (coupled simulated storm tide) indicates an accurate simulation of storm

tides, demonstrating a satisfactory level of precision. With the exception of a few samples, the majority of the green and orange dots exhibit close proximity, signifying a relatively accurate estimation of TSIWL. Naturally, given the intricate physical mechanisms underlying TSI, some degree of error is inevitable. Upon examining the exceedance probability curves derived from various samples, we observe that the red, orange, and green lines are relatively close to each other, exhibiting a noticeable distinction from the blue line. This suggests that the statistical estimation of TSIWL at the HK station is relatively accurate, with errors within an acceptable range.

Using the method outlined in section 3.5, we compute the storm tide RPs utilizing the recorded and simulated tidal levels at the HK station, and compare the two. Figure 9 serves as a visual representation of the outcomes achieved when estimating the storm tide RPs at the HK station. Notably, the green dots and line depict the observed instances of EWL, alongside the corresponding exceedance probability curve derived through GPD fitting. The green shaded area represents the 95% confidence interval and the 95% confidence interval length for RP50 is 0.44 m. The maximum recorded extreme samples were 2.12 m (during typhoon “Mangkhut” in 2018), 1.8 m (during typhoon “Hagibis” in 2008), and 1.7 m (during typhoon “Hato” in 2017), correlating with respective return periods of 180 years, 37 years, and 20 years, respectively. The blue dots, denoted as $(\eta_s + \eta_t)_{\max}$, represents the peak value derived from the linear superposition of the modeled surge and 500 randomly generated astronomical tides, without accounting for TSI. The blue line represents the exceedance probability curve obtained by fitting the blue dots based on GPD. Similarly, the blue shaded area visually represents the 95% confidence interval, with a confidence interval length of 0.07 m for RP50 storm tide. The red dot represents $(\eta_s + \eta_t)_{\max} + \eta_{I,est}$, and the red line represents the exceedance probability curve obtained through GPD fitting. The red shaded area represents the 95% confidence interval, which, for RP50, has a confidence interval length of 0.07 m.

A comparison between the red and green lines reveals that employing the new method yields a larger sample size, resulting in narrower confidence intervals and more stable computed outcomes. Conversely, the original observed data, with its limited

sample size, leads to wide confidence intervals and significant uncertainty in the calculated results. Within the range of 20-year to 50-year return periods, we observe a closer overlap between the red and green lines, indicating the consistency between the results obtained from the new method and those based on observational data. However, for storm tides with a return period of less than 20 years, the green line shows a gradual decline from an exceedance probability of 0.05 to 1, while the red line exhibits a more pronounced decrease. This discrepancy arises from the fact that the measured data samples rely on a continuous long-time series of water levels, and the samples coverage time range is wider. In contrast, the green line considers only samples from typhoon events, disregarding a portion of the extreme storm tides caused by non-TCs. Consequently, for extreme storm tides falling within the range of lower return periods (below 20 years), the estimations derived from observed samples (green dots) surpass the results obtained from numerical simulation estimates (red dots). Furthermore, when considering the exceedance probability curve (green line) calculated based on observation data, it is important to note that the absence of sufficient samples to fit a distribution leads to an accelerated rise rate for storm surges with a return period exceeding 200 years. Consequently, we regard this specific segment of the calculated results as unreliable. Based on this analysis, it can be concluded that when estimating extreme sea levels with a return period exceeding 20 years at the HK station, it is adequate to rely exclusively on storm tide data during typhoons. The difference between the blue and red lines primarily arises from the consideration or exclusion of TSI. When TSI is not taken into account, the RP50 storm tide amounts to 2 m. However, considering TSI reduces it to 1.87 m, resulting in a difference of 0.13 m. As the return period increases, this difference progressively amplifies due to the growing influence of surge on peak storm tides.

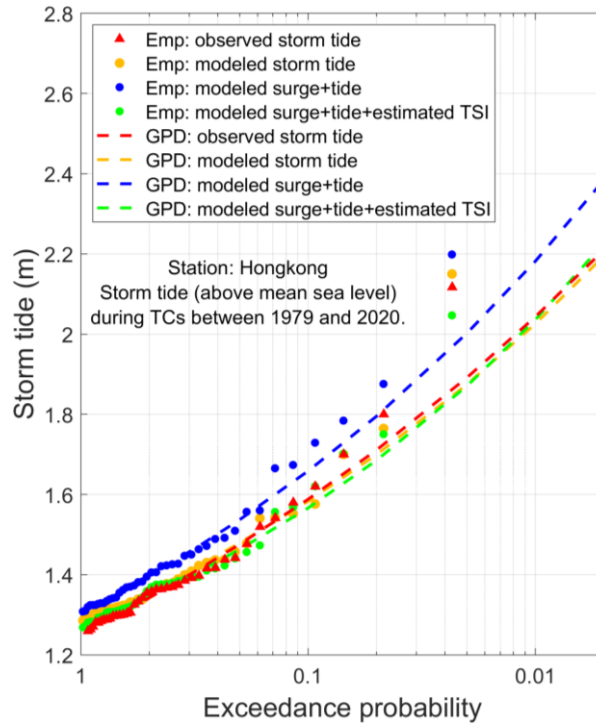


Figure 8. The flood exceedance probability curve for the HK station is calculated using the POT/GPD method based on typhoon-induced storm tides from 1979 to 2020. The red dots represent the observed tidal levels, the orange dots denote the coupled simulation of storm tide $\eta_{st,max}$, the blue dots represent $(\eta_s + \eta_t)_{max}$ (the linearly superposition of modeled pure surge and astronomical tide), and the green dots indicate $(\eta_s + \eta_t)_{max} + \eta_{l,est}$ (the linearly superposition of surge, astronomical tide, and estimated TSIWL). The exceedance probability curves for storm tide are estimated using the POT/GPD method with the aforementioned sample data, visually represented by the red, orange, blue, and green lines. "Emp" denotes the empirical cumulative frequency method, while "GPD" represents the fitting method based on the generalized Pareto distribution.

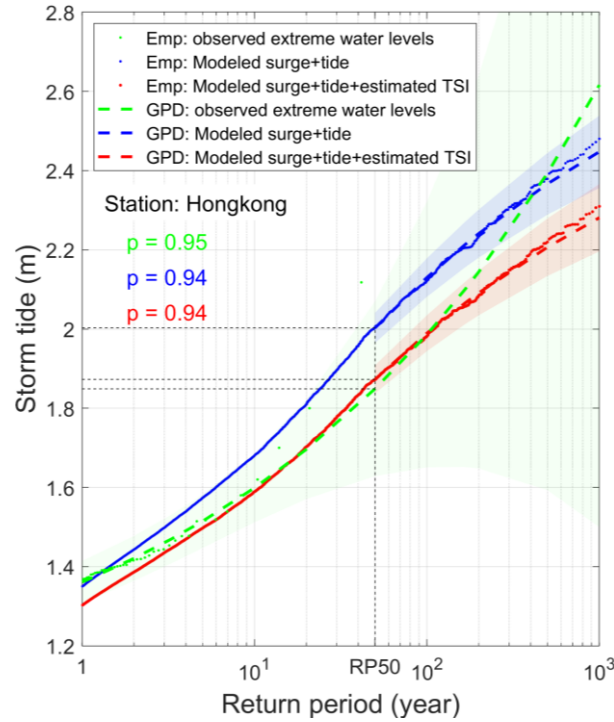


Figure 9. The return period for storm tide level. The green dots signify the observed extreme tidal levels recorded at the HK station between 1979 and 2020. The blue and red dots represent the simulated samples of typhoon-induced storm tides, with the blue dots representing the simulated $(\eta_s + \eta_t)_{\max}$ during the typhoon periods from 1979 to 2020 at the HK station. The red dots, which build upon the blue dots, incorporate the estimation of TSIWL. The green, blue, and red lines correspond to the exceedance probability curves obtained by fitting GPD to these three sets of samples. The shaded area represents the 95% confidence interval.

4.5 Estimation of extreme storm tide RPs

Using the methodology outlined in Section 3.5, the exceedance probability curves for storm tides at P1, P4, P5, P6, P7, and P9 were calculated (Figure 10), similar to the HK station. Examining the graph reveals that neglecting TSI results in overestimation (P1, P6, P7, P9) or underestimation (P4, P5) extreme storm tides. Moreover, as the return period increases, the magnitude of overestimation or underestimation becomes more pronounced. This phenomenon arises from the influence of TSI. Higher extreme storm tides indicate larger storm surges and astronomical tides, consequently leading to greater values for the nonlinear water level due to TSI. Notably, an intriguing observation emerges: only at P4 and P5 does the red line surpass the blue line,

signifying a positive contribution of TSI to extreme storm tides. This occurrence can be attributed to the intense typhoons making landfall in the Pearl River Estuary and Leizhou Bay, which result in significant storm surges. When combined with astronomical tides to calculate $\eta_{I,est}$, the larger proportion of the storm surge leads to a positive value for η_{Ip} (Figure 6d-e). Upon sorting the final samples in descending order and plotting them on a logarithmic scale, the red dots are positioned above the blue dots. The RP50 storm tide at these river six stations are 3.60, 2.91, 4.04, 2.27, 1.75, and 3.01 m, respectively. The corresponding contribution of TSI for the RP50 storm tides are -0.24, 0.1, 0.08, -0.13, -0.1, and -0.08 m, accounting for -6.8%, 3.1%, 1.8%, -5.3%, -5.7%, and -2.1% of the total water level, respectively.

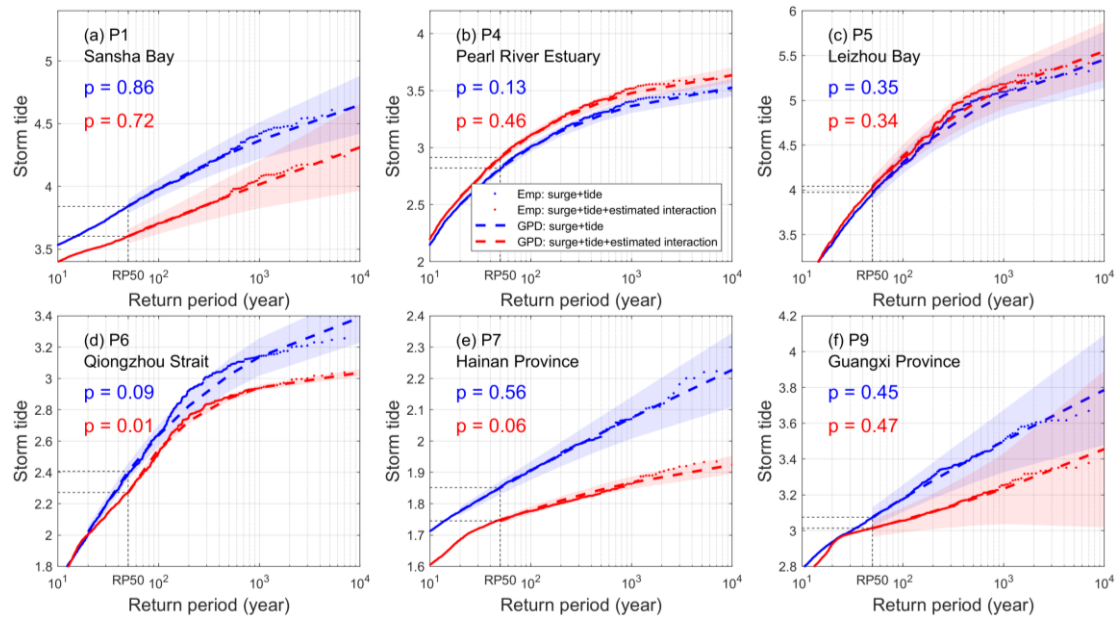


Figure 10. The exceedance probability curve for storm tides at P1, P4, P5, P6, P7, P9. The meanings of blue dots, red dots, blue lines, red lines, blue shaded area and red shaded area are the same as those in Figure 9.

5. Discussion

5.1 Stability analysis of astronomical tides sampling quantity

In the previously proposed method for estimating storm tide with a 2% exceedance probability, the typhoon-induced storm surge is linearly superimposed with 500

randomly generated astronomical tides. This is then combined with the estimated TSIWL to obtain samples for calculating extreme storm tides. The superposition with random astronomical tides is necessary to account for combinations of storm events and tides that have not occurred within the limited historical records (Dullaart et al., 2021; Orton et al., 2016). Along the northern coast of SCS, astronomical tides manifest as both diurnal and semidiurnal occurrences. Despite the presence of at least one high tide and one low tide daily, there are variations in the height between successive high tides. A complete cycle of astronomical tides occurs every 18.6 years. Therefore, a larger astronomical tides sampling quantity ensures a more comprehensive capture of potential storm surge and tide combinations.

However, exceeding a certain threshold of superimpositions introduces a notable degree of redundancy in the samples, leading to the wastage of samples and an excessive computational burden. Consequently, a stability analysis was conducted on the superposition number (Figure 11), confirming that employing 500 astronomical tides for superimposition ensures stable estimation results. Figure 11 displays the RP50 storm tide at P1, P4, and P5 calculated by superimposing storm surges with 10, 50, 100, 200, 500, 1000, 2000, and 5000 randomly generated astronomical tides (blue solid line). The red dashed line represents the 95% confidence interval length corresponding to each scenario. For P1, P4, and P5, a superposition of 100, 100, and 200 random astronomical tides, respectively, can be considered stable for calculating the RP50 storm tide. The 95% confidence interval length at P1 and P4 begins to be less than 10 cm when the superposition number is 50, while for P5, it becomes less than 10 cm when the superposition number reaches 200. Therefore, for these three points, a superposition number of 200 with random astronomical tides already provides sufficient accuracy for estimating RP50. To account for possible variations at specific points, we have relaxed this critical threshold to 500 to ensure a fully stable computation of the results.

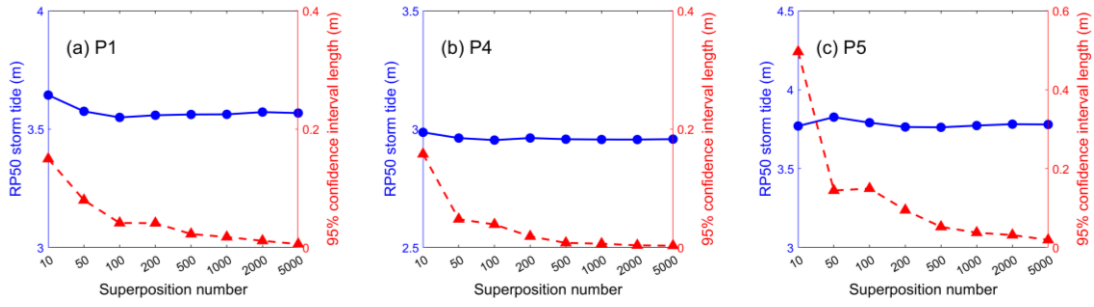


Figure 11. Stability analysis of the number of superposition with random astronomical tides at P1, P4, and P5. The blue solid line represents the RP50 storm tide, calculated using varying numbers of superposition with astronomical tides, while the red dashed line represents the 95% confidence interval length.

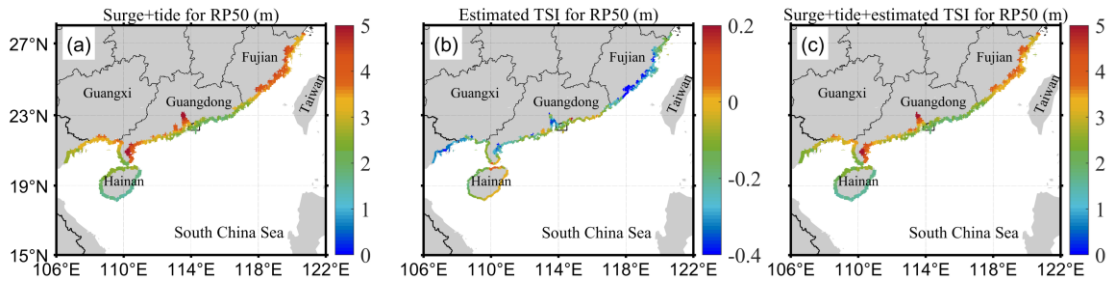


Figure 12. (a) The RP50 storm tide calculated based on $(\eta_s + \eta_t)_{\max}$. (b) The contribution of estimated TSIWL to the RP50 storm tide. (c) The RP50 storm tide calculated based on

$$(\eta_s + \eta_t)_{\max} + \eta_{t,est}$$

5.2 TSI contributions to the RP50 storm tide

When calculating storm tide RPs, the samples are usually selected from continuous water level sequences. The extreme value of the samples often comes from a combination of powerful TC-induced significant storm surges and astronomical high tides. Solely relying on extreme samples during typhoons can only provide estimate for storm tides with longer RPs. When estimating storm tides with shorter return periods, employing samples induced by typhoons would underestimate the results. For instance, in Figure 9, the tidal level corresponding to RPs of less than 20 years at the HK station exhibits a relatively gradual decline (from RP20 to RP1) in the fitted distribution curve based on measured data, while the extreme storm tide calculated using samples during typhoons shows a rapid decline, with the former being higher than the latter. This indicates that using tidal data during typhoons is only appropriate for estimating

extreme sea levels with higher RPs, and estimating extreme sea levels with lower RPs necessitates a comprehensive consideration of both non-TC-induced and TC-induced storm surges.

We introduce the notion of a critical RP in this context, which refers to the return period corresponding to the intersection point between the exceedance probability curves calculated from long-term continuous water level sequences and extreme tidal levels during TCs. The point of intersection between the green and red lines depicted in Figure 9 (corresponding to a RP of 20 years) signifies the critical RP at the HK station. The critical RP of storm tide, calculated based on two distinct datasets, exhibits variation across different coastal locations. Stations that are less affected by TCs or have larger tidal ranges tend to have higher critical RP. Conversely, stations that are more influenced by TC-induced storm surges or have smaller tidal ranges tend to have smaller critical RP. As an illustration, the critical RP calculated for the HK station in this study is approximately 20 years, while Dullaart et al. (2021) computed a critical RP of 30 years for the Quanzhou station along the Fujian Coast. The HK station is situated in the Pearl River Estuary area, with a relatively small tidal range, and is vulnerable to the direct impact of powerful typhoons. The Quanzhou station, with a larger tidal range, experiences weaker typhoon impacts in terms of frequency and intensity compared to the HK station, so the critical RP of the station is relatively greater. Although each station along the northern SCS coast exhibits its own distinct characteristics, considering the relatively high frequency of typhoon impacts, we posit that using tidal level data during typhoons is adequate for calculating the RP50 storm tides.

Figure 12a and 12c depict the RP50 storm tides calculated using sample $(\eta_s + \eta_t)_{\max}$ and sample $(\eta_s + \eta_t)_{\max} + \eta_{I,est}$, respectively, and Figure 12b presents the contribution of TSI effects to the RP50 storm tides. The figures clearly indicate that the highest RP50 storm tides are observed along the coastline from the Pearl River Estuary to the Leizhou Bay in Guangdong, as well as the coastal areas of Fujian (Figure 12c). At the top of the Pearl River Estuary and Leizhou Bay, the maximum values exceed 5 m, while the largest extreme storm tide for RP50 along the Fujian Coast occurs within

the Sansha Bay, with a maximum value of 4.5 m. In Guangdong Province's coastal areas, the larger extreme sea levels are primarily attributed to the direct impacts of strong typhoons, whereas in the coastal areas of Fujian, the influence is mainly due to the larger tidal range. Excluding the Qiongzhou Strait, in most of the northern coastline of the SCS, neglecting TSI would result in an overestimation of the RP50 extreme sea levels. The most significant overestimation is observed along the southern coast of Fujian, with an overestimation magnitude of 40 cm. In the northern region of the Fujian coastal area, TSI contributes to a decrease of approximately 30 cm in the RP50 extreme storm tides. Along the coastal areas of Guangdong Province, TSI contributes between -30 cm to 0 cm to the RP50 extreme storm tides, while in the coastal areas of Guangxi, the contribution ranges from -30 cm to -10 cm. The western region of Hainan Island has a TSI contribution of approximately -10 cm, while the eastern region of Hainan Island exhibits a smaller TSI contribution, within -10 cm.

5.3 Application and prospect

Given the pronounced regional attributes inherent to TSI-induced nonlinear water levels, their manifestation is heavily contingent not only on surge magnitudes but also on a host of influential factors encompassing local bathymetry, tidal ranges, coastal geometry, and bottom friction. The proposed statistical method for quantifying the TSI effects only needs to calculate the storm surge and one cycle of astronomical tides in advance, avoiding the multiple-scenario tide-surge coupled simulations. While providing a simple and efficient method for evaluating extreme water levels in a probabilistic framework, the specific formulations in the study are site-dependent, thus requires modifications for applications in other coastal areas.

6. Summary and conclusions

This article establishes multiple regression models for nearshore areas η_{lp} and η_{sp} by simulating storm surges and tide-surge interaction (TSI) during typhoons in the northern SCS from 1979 to 2020, and estimates the TSIWL during extreme storm events. We proposed a novel method that incorporates the TSI effects on storm tides and

estimates the RPs of storm tides. This method represents a significant improvement over previous methodologies, offering enhanced accuracy and efficiency in computing extreme storm tides.

Moreover, the study estimates the RP50 storm tides along the northern coast of the SCS, while also assessing the contribution of TSI. The results reveal that the maximum values exceeding 5 m are observed at the Pearl River Estuary and Leizhou Bay, whereas the highest RP50 storm tide along the Fujian Coast occurs within Sansha Bay, reaching 4.5 m. Excluding the Qiongzhou Strait, neglecting TSI leads to an overestimation of the RP50 storm tides in most of the northern coastline of the SCS.

Acknowledgement

The authors would like to acknowledge the support of the National Key R&D Program of China (2023YFC3007900) and the National Natural Science Foundation of China (Grant No. 52071305, No. 52088102 and No. 42006154).

Appendix A

The appendix materials A provide the validation results of the simulated wind speed and air pressure in the blended wind field, as well as the verification results of the numerical model for astronomical tides and storm surges. The anemometers and tide gauges are represented by squares, circles, and asterisks in the Figure A.1a and A.1b, respectively. Specifically, the astronomical tide constituents were derived from the tide gauges labeled with a red asterisk in Figure A.1a, while extreme wind speeds and minimum atmospheric pressures were extracted from the anemometers marked with blue and red squares in A.1a. The recorded data for extreme storm tides and storm surges were obtained from the tide gauges identified by blue and red circles in Figure A.1b.

Figure A.1c and A.1d illustrate the validation of the amplitudes and phases of the four primary astronomical tidal constituents (M2, K1, S2, O1) at multiple tide gauge stations. The validation data for each of the four tidal components is represented by

blue squares, red triangles, magenta circles, and green diamonds, respectively. It is observed that the model tends to overestimate the amplitude of the K1 constituent. However, the scatter plots for the remaining three constituents (M2, S2, O1) cluster closely along the 1:1 diagonal line, demonstrating a total amplitude RMSE of merely 7 cm and a CC of 0.98. The scatter plots for the phase of the four constituents also exhibit a concentration near the 1:1 diagonal line, with only a few points slightly deviating and mostly falling within the $\pm 30^\circ$ error range. The overall RMSE for the phase is merely 18.56° , accomplished by a Pearson CC of 0.98. Upon analyzing the overall RMSE, AAE, MRE, and Pearson CC, it is evident that the simulated values align well with the observed values. This indicates that the hydrodynamic model based on FVCOM accurately captures the behavior of astronomical tides.

The validation of blended wind fields, air pressure, storm surge, and storm tide primarily focuses on extreme values. Figure A.2a and A.2b illustrate the comparison between observed and simulated data for maximum wind speeds and minimum central atmospheric pressures. The comparison is visualized using red circles, with varying grayscale representing different typhoons. The dataset for wind speed comprises 50 typhoons, while the observed atmospheric pressure covers 31 typhoons. Due to space limitations, specific typhoon IDs are not explicitly provided in the text. By integrating the comparison of simulated extreme wind with observed data, along with statistical metrics such as RMSE, AAE, MRE, and Pearson CC, it can be concluded that within an acceptable margin of error, the blended wind fields demonstrate the capability to accurately simulate wind speeds and air pressures.

Figure A.2c and A.2d displays scatter plots comparing the observed and simulated data for extreme storm surges and storm tides. In Figure A.2c, the observed storm surge is represented by the residual water level obtained by subtracting the modeled astronomical tide and the nonlinear water level resulting from TSI from the total water level. The simulated storm surge, denoted by red squares, represents the pure storm surge (EXP-PS) without considering the influence of TSI. The shades of the circles and squares indicate different typhoons, with data points from 44 typhoon events included

in Figure A.2c. From the graph, it is evident that the RMSE, AAE, MRE, and CC for the simulated pure surge compared to the observed surge are 0.22 m, 0.19 m, 0.18, and 0.98, respectively. The majority of scattered data points cluster along the 1:1 diagonal, exhibiting minimal errors and high correlation coefficients, with no points exceeding the error range of ± 0.5 m. In Figure A.2d, the vertical axis represents the coupled simulation results of storm tides, incorporating the effects of TSI. The grayscale shading within the circles and squares represents different typhoons, encompassing a dataset of 50 typhoons. The RMSE, AAE, MRE, and CC of the two sets of data were 0.25 m, 0.21 m, 0.12, and 0.93, respectively. The majority of scattered data points fall within an error range of ± 0.5 m and are evenly distributed along a 1:1 diagonal. In conclusion, the model effectively reproduces extreme storm surges and storm tides, with errors falling within an acceptable range.

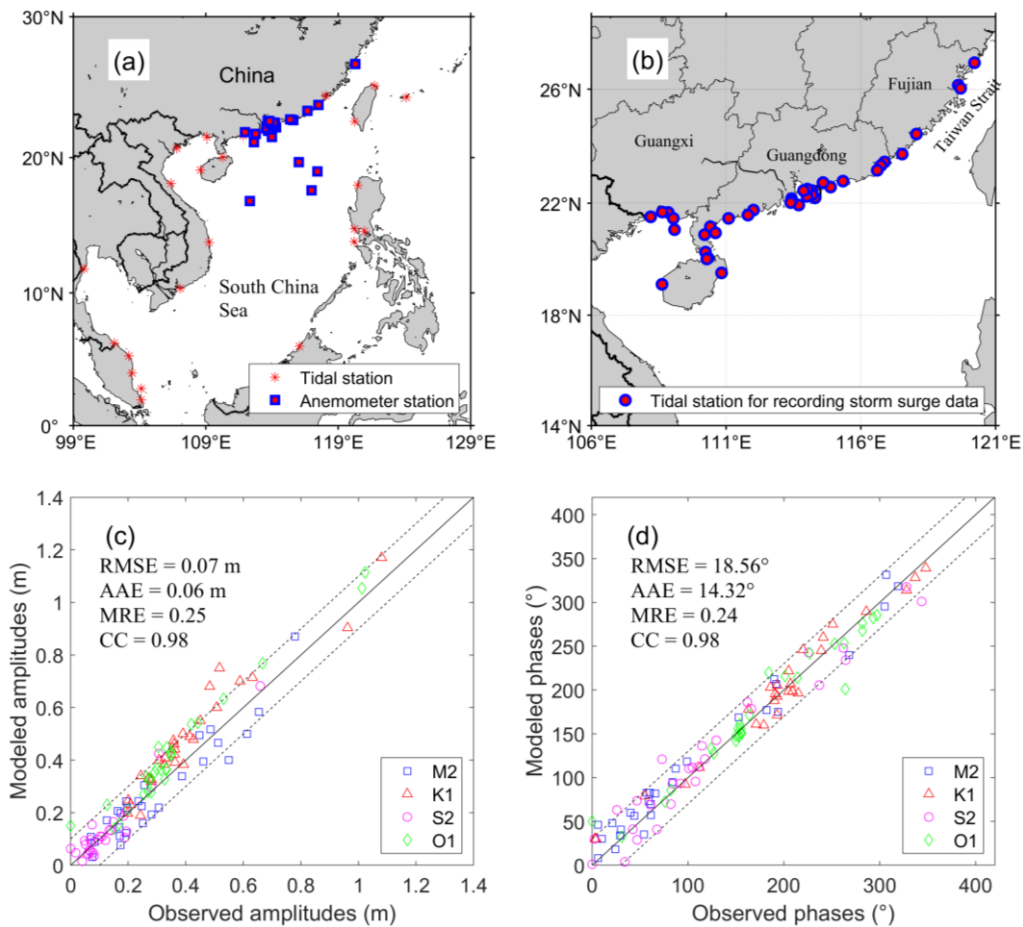


Figure A.1. (a) The red asterisks represent tide gauge stations used to validate astronomical tide constituents, while the blue and red squares denote anemometers employed to validate wind speed

and air pressure. (b) The blue and red circles denote tide gauge stations used to validate storm surge. (c) and (d) show scatter plots comparing simulated and observed values for the amplitude and phase of astronomical tide constituents. In these plots, the blue squares represent the M2 constituent, the red triangles represent K1, the magenta circles represent S2, and the green diamonds represent O1.

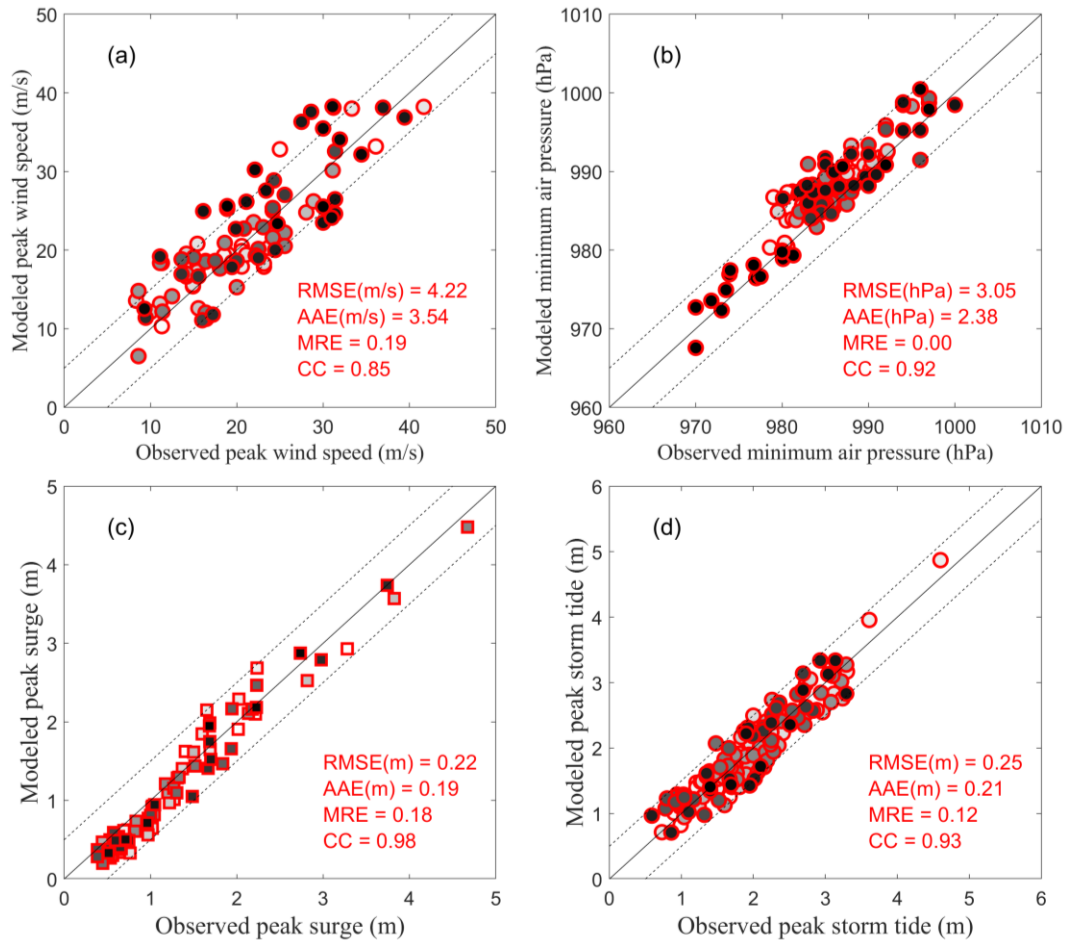


Figure A.2. (a-b) Observed and modeled values of peak wind speed and minimum central air pressure. The observed data is sourced from the anemometers in Figure A.1a. (c) Observed and modeled values of peak surge. (d) Observed and modeled values of peak storm tides. The observed data of tidal levels is from the tidal stations in Figure A.1b. The different grayscales inside the circles and squares symbolize different typhoons.

Appendix B

Appendix materials B consist of a bar chart illustrating the distribution of event proportions based on different time differences between $(\eta_s + \eta_t)_{\max}$ and $\eta_{st,\max}$. The

statistical data includes simulated storm tide events occurring at grid points near the shoreline. It is evident that over 80% of the events exhibit no significant impact of TSI on the timing of $(\eta_s + \eta_t)_{\max}$ and $\eta_{st,\max}$ (falling within the 0-hour interval). Only 16.5% and 1.3% of the events demonstrate a deviation of 1 to 2 hours in the occurrence of the peak value due to TSI (falling within the ± 1 and ± 2 -hour range, respectively). The slight time difference between $(\eta_s + \eta_t)_{\max}$ and $\eta_{st,\max}$ will not exert a substantial influence on TSIWL. Therefore, we denote $\eta_l = \eta_{st,\max} - (\eta_s + \eta_t)_{\max}$.

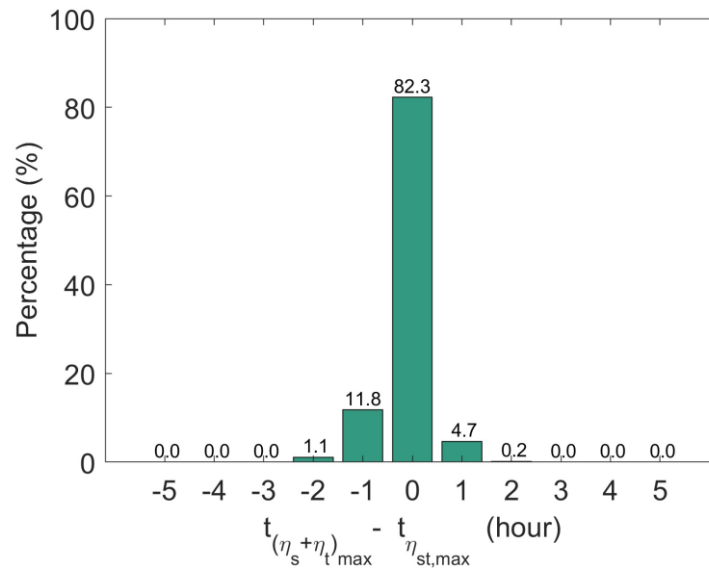


Figure B.1. Bar chart representing the proportion of events corresponding to different time differences between $(\eta_s + \eta_t)_{\max}$ and $\eta_{st,\max}$.

References

- Arns, A. et al., 2020. Non-linear interaction modulates global extreme sea levels, coastal flood exposure, and impacts. *Nature communications*, 11(1): 1918.
- Bernier, N. and Thompson, K., 2007. Tide-surge interaction off the east coast of Canada and northeastern United States. *Journal of Geophysical Research: Oceans*, 112(C6).
- Carrere, L., Lyard, F., Cancet, M. and Guillot, A., 2015. Fes 2014, a new tidal model on the global ocean with enhanced accuracy in shallow seas and in the arctic region. *EGU general assembly conference abstracts*, Vienna, Austria, p. 5481.
- Chen, C., Liu, H. and Beardsley, R.C., 2003. An unstructured grid, finite-volume, three-dimensional, primitive equations ocean model: Application to coastal ocean and estuaries. *Journal of atmospheric and oceanic technology*, 20(1): 159-186.
- Costa, W., Bryan, K.R., Stephens, S.A. and Coco, G., 2023. A regional analysis of tide-

- surge interactions during extreme water levels in complex coastal systems of aotearoa new zealand. *Frontiers in Marine Science*, 10: 1170756.
- Davies, A. and Lawrence, J., 1994. Examining the influence of wind and wind wave turbulence on tidal currents, using a three-dimensional hydrodynamic model including wave–current interaction. *Journal of Physical Oceanography*, 24(12): 2441-2460.
- Davison, A.C. and Smith, R.L., 1990. Models for exceedances over high thresholds. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 52(3): 393-425.
- Dullaart, J.C. et al., 2021. Accounting for tropical cyclones more than doubles the global population exposed to low-probability coastal flooding. *Communications Earth & Environment*, 2(1): 135.
- Flather, R., 2001. Storm surges. *Encyclopedia of ocean sciences*: 2882-2892.
- Hersbach, H. et al., 2018. Era5 hourly data on pressure levels from 1940 to present. Copernicus climate change service (c3s) climate data store (cds).
- Holland, G.J., 1980. An analytic model of the wind and pressure profiles in hurricanes.
- Horsburgh, K. and Wilson, C., 2007. Tide-surge interaction and its role in the distribution of surge residuals in the north sea. *Journal of Geophysical Research: Oceans*, 112(C8).
- Hosking, J.R. and Wallis, J.R., 1987. Parameter and quantile estimation for the generalized pareto distribution. *Technometrics*, 29(3): 339-349.
- Hu, S. et al., 2023. Quantification of the nonlinear interaction among the tide, surge and river in pearl river estuary. *Estuarine, Coastal and Shelf Science*, 290: 108415.
- Idier, D., Dumas, F. and Muller, H., 2012. Tide-surge interaction in the english channel. *Natural Hazards and Earth System Sciences*, 12(12): 3709-3718.
- Khan, M.J.U. et al., 2022. Storm surge hazard over bengal delta: A probabilistic–deterministic modelling approach. *Natural Hazards and Earth System Sciences*, 22(7): 2359-2379.
- Leijnse, T.W.B., Giardino, A., Nederhoff, K. and Caires, S., 2022. Generating reliable estimates of tropical-cyclone-induced coastal hazards along the bay of bengal for current and future climates using synthetic tracks. *Natural Hazards and Earth System Sciences*, 22(6): 1863-1891.
- Li, N., Yamazaki, Y., Roeber, V., Cheung, K.F. and Chock, G., 2018. Probabilistic mapping of storm-induced coastal inundation for climate change adaptation. *Coastal Engineering*, 133: 126-141.
- Liang, B., Shao, Z., Li, H., Shao, M. and Lee, D., 2019. An automated threshold selection method based on the characteristic of extrapolated significant wave heights. *Coastal Engineering*, 144: 22-32.
- Lin, N. and Chavas, D., 2012. On hurricane parametric wind and applications in storm surge modeling. *Journal of Geophysical Research: Atmospheres*, 117(D9).
- Macnab, R., 2003. The history of gebco, 1903-2003. Desmond scott. 2003. Published by gitc bv.

- Marsooli, R. and Lin, N., 2018. Numerical modeling of historical storm tides and waves and their interactions along the us east and gulf coasts. *Journal of Geophysical Research: Oceans*, 123(5): 3844-3874.
- Martín, A., Amores, A., Orfila, A., Toomey, T. and Marcos, M., 2023. Coastal extreme sea levels in the caribbean sea induced by tropical cyclones. *Natural Hazards and Earth System Sciences*, 23(2): 587-600.
- Mas, E. et al., 2015. Field survey report and satellite image interpretation of the 2013 super typhoon haiyan in the philippines. *Natural Hazards and Earth System Sciences*, 15(4): 805-816.
- Muis, S. et al., 2019. Spatiotemporal patterns of extreme sea levels along the western north-atlantic coasts. *Scientific Reports*, 9(1): 3391.
- Muis, S., Verlaan, M., Winsemius, H.C., Aerts, J.C. and Ward, P.J., 2016. A global reanalysis of storm surges and extreme sea levels. *Nature communications*, 7(1): 11969.
- Needham, H.F., Keim, B.D. and Sathiaraj, D., 2015. A review of tropical cyclone-generated storm surges: Global data sources, observations, and impacts. *Reviews of Geophysics*, 53(2): 545-591.
- Orton, P.M. et al., 2016. A validated tropical-extratropical flood hazard assessment for new york harbor. *Journal of Geophysical Research: Oceans*, 121(12): 8904-8929.
- Prandle, D. and Wolf, J., 1978. The interaction of surge and tide in the north sea and river thames. *Geophysical Journal International*, 55(1): 203-216.
- Ragno, E., Antonini, A. and Pasquali, D., 2023. Investigating extreme sea level components and their interactions in the adriatic and tyrrhenian seas. *Weather and Climate Extremes*, 41: 100590.
- Rego, J.L. and Li, C., 2010. Nonlinear terms in storm surge predictions: Effect of tide and shelf geometry with case study from hurricane rita. *Journal of Geophysical Research: Oceans*, 115(C6).
- Rossiter, J.R., 1961. Interaction between tide and surge in the thames. *Geophysical Journal International*, 6(1): 29-53.
- Tang, Y.M., Sanderson, B., Holland, G. and Grimshaw, R., 1996. A numerical study of storm surges and tides, with application to the north queensland coast. *Journal of Physical Oceanography*, 26(12): 2700-2711.
- Vickery, P., Skerlj, P. and Twisdale, L., 2000. Simulation of hurricane risk in the us using empirical track model. *Journal of structural engineering*, 126(10): 1222-1237.
- Walton Jr, T.L., 2000. Distributions for storm surge extremes. *Ocean Engineering*, 27(12): 1279-1293.
- Wankang, Y. et al., 2019. The effect of nonlinear factors on tide-surge interaction: A case study of typhoon rammasun in tieshan bay, china. *Estuarine, Coastal and Shelf Science*, 219: 420-428.
- Willoughby, H. and Rahn, M., 2004. Parametric representation of the primary hurricane

- vortex. Part i: Observations and evaluation of the holland (1980) model. *Monthly Weather Review*, 132(12): 3033-3048.
- Wolf, J., 1981. Surge-tide interaction in the north sea and river thames. Floods due to high winds and tides: 75-94.
- Wu, G., Shi, F., Kirby, J.T., Liang, B. and Shi, J., 2018. Modeling wave effects on storm surge and coastal inundation. *Coastal Engineering*, 140: 371-382.
- Wuxi, Q., Li, J. and Nie, B., 2018. Effects of tide-surge interaction and wave set-up/set-down on surge: Case studies of tropical cyclones landing china's zhe-min coast. *Theoretical and Applied Mechanics Letters*, 8(3): 153-159.
- Yang, J. et al., 2019. A comparative study of typhoon hato (2017) and typhoon mangkhut (2018)—their impacts on coastal inundation in macau. *Journal of Geophysical Research: Oceans*, 124(12): 9590-9619.
- Yang, S. et al., 2023. Non-linear interactions between tides and storm surges during extreme weather events over the eastern canadian shelf. *Ocean Dynamics*: 1-23.
- Zeng, X., Shi, W., Michailides, C., Zhang, S. and Li, X., 2021. Numerical and experimental investigation of breaking wave forces on a monopile-type offshore wind turbine. *Renewable Energy*, 175: 501-519.
- Zhang, H., Cheng, W., Qiu, X., Feng, X. and Gong, W., 2017. Tide-surge interaction along the east coast of the leizhou peninsula, south china sea. *Continental Shelf Research*, 142: 32-49.
- Zhang, W.Z., Shi, F., Hong, H.S., Shang, S.P. and Kirby, J.T., 2010. Tide-surge interaction intensified by the taiwan strait. *Journal of Geophysical Research: Oceans*, 115(C6).
- Zheng, P. et al., 2020. Tide-surge interaction in the pearl river estuary: A case study of typhoon hato. *Frontiers in Marine Science*, 7: 236.