An Attribute-based Model to Retrieve Storm Surge Disaster Cases

Ke Wang
Institute of Public Safety Research, Department of Engineering Physics, Tsinghua University, Beijing, China
Beijing Key Laboratory of City Integrated Emergency Response Science, Tsinghua University, Beijing, China
wangke16@mails.tsinghua.edu.cn

Yongsheng Yang
Institute of Public Safety Research, Department of Engineering Physics, Tsinghua University, Beijing, China
Beijing Key Laboratory of City Integrated Emergency Response Science, Tsinghua University, Beijing, China
yang-ys18@mails.tsinghua.edu.cn

Genserik Reniers
CEDON, KU Leuven, Brussels, Belgium
Faculty of Technology, Policy and Management, Safety and Security Science Group (S3G), TU Delft, the Netherlands
Faculty of Applied Economics, Antwerp Research Group on Safety and Security (ARGoSS), University of Antwerp, Antwerp Belgium
genserik.reniers@kuleuven.be

Jian Li
Institute of Public Safety Research, Department of Engineering Physics, Tsinghua University, Beijing, China
jianli@tsinghua.edu.cn

Quanyi Huang
Institute of Public Safety Research, Department of Engineering Physics, Tsinghua University, Beijing, China
qyhuang@tsinghua.edu.cn

ABSTRACT

In China, storm surge disasters cause severe damages in coastal regions. One of the most critical tasks is to predict affected regions and their relative damage levels to support decision-making. This study develops a two-stage retrieval model to search the most similar past disaster case to complete prediction. Based on spatial attributes of cases, the top-ranking past cases with a similar location to the target case are selected. Among these past cases, the most similar past case is selected by disaster attribute similarities. Three typical storm surge case studies have been used and implemented into this proposed model, and the results show that all the most affected regions can be predicted. The proposed model simplifies the prediction process and updates results quickly. This study provides valuable information for the government to make real-time response plans.

CoRe Paper – Planning, Foresight and Risk Analysis
Proceedings of the 18th ISCRAM Conference – Blacksburg, VA, USA May 2021
Anouck Adrot, Rob Grace, Kathleen Moore and Christopher Zobel, eds.

Keywords

Storm surge disaster, multiple attributes, retrieval model, affected region prediction.

INTRODUCTION

A storm surge is an abnormal increase in the sea surface caused by cyclones (Resio and Westerink, 2008), divided into the typhoon storm surge (caused by the tropical cyclone) and extratropical storm surge (caused by the extratropical cyclone) (State Oceanic Administration of China, 2017). The force of the winds and the low pressure associated with intense cyclones push water toward the shore (Feng, 1982). The rise of water level along the coast, which is caused by a combination of the storm surge, the astronomical tide, and the wave, can lead to a storm surge disaster (Liu et al., 2019). If the seawater invades inland, it can lead to coastal flooding. In China, storm surge disasters cause huge damages and have devastating impacts on coastal areas every year, especially typhoon storm surge disasters (Yang et al., 2016). For example, the amount of economic losses of typhoon storm surge disasters exceeds 10 billion RMB (1.5 billion dollars) in 2019, accounting for 99.4% of the total amount of marine disasters, according to Bulletin of China Marine Disaster (2019). In the past few decades, the frequency of storm surge disasters has been rising in China (Fang et al., 2017). Meanwhile, climate change and population growth in coastal regions are likely expected to increase the risk of this disaster in the future (Helderop and Grubesic, 2019; Lin et al., 2012; Ling et al., 2011). To respond this rapid-onset disaster and reduce the related losses, it is necessary to predict affected regions and their relative damage levels.

Researchers used hydrodynamic simulation models, such as Advanced Circulation and MIKE21 (Liu et al., 2018; Reffitt et al., 2020; Shi et al., 2020; Tanim and Goharian, 2020; Zheng and Sun, 2020), to predict affected regions and their damages of storm surge disasters. One of these simulation aims is to predict inundated areas to make evacuation plans, which requires various data and complex computations to obtain the flood water depth of storm surge disasters in coastal areas. The damages of affected areas can be evaluated by damage curves (Hsu et al., 2018). Before the hydrodynamic simulation, the data, such as the wind field of cyclones, topography data, and astronomical tide for simulated regions, are required and obtained from the public platforms. However, it is challenging to obtain the corresponding damage curves because they change with different affected targets and different inundated areas. One cyclone triggers a storm surge which can impact several province-level coastal regions. Under the circumstance of multi-region getting affected, the hydrodynamic simulation, which is usually performed on a county-level or city-level, fails to provide national-level disaster managers suggestions despite its high prediction accuracy. Furthermore, the hydrodynamic simulation may take several hours or even days, and it will also take so much time to update the simulated results if getting new information. Before a storm surge disaster occurs, the national-level disaster managers should know the affected scope and the relative damage levels in each affected region to manage this disaster. Therefore, a new model needs to be developed to predict the related information in a short time.

Storm surge disasters are influenced by various factors, and possible reasons for rich disaster events vary over time and regions. It is difficult to quantify and integrate all related factors to predict affected regions. If the upcoming disaster happens under similar conditions to past disasters, it is possible to complete prediction through similar past events. Case-based reasoning (CBR) is a knowledge method to solve a new problem by reusing solutions to old problems (Fei and Feng, 2020). Compared with simulations, the CBR method could simplify the calculation process to obtain prediction results quickly. Many studies introduced this method to predict disasters. For example, Pedro et al. (2005) integrated CBR and Fuzzy multicriteria decision making to forecast future positions of tropical cyclone tracks. Deng and Li (2020) adopted CBR for assessing the risk of geological disasters by revising spatial past cases. Chen et al. (2011) applied CBR to predict the economic loss of typhoon disasters. According to statistics and analyses (Shi et al., 2015), storm surge disasters have specific spatial and temporal patterns, including prone areas and times, which also provides a base to predict affected regions and their relative damage levels by similar disaster cases.

One essential step of the CBR cycle is to create a retrieval model to search similar cases. However, few studies emphasized the retrieval model of storm surge disasters. A cyclone triggers a storm surge disaster, and its spatial locations and intensities directly link to the affected areas and damages of storm surge disasters. Therefore, the retrieval model needs to consider corresponding cyclone information. Previous studies focused on retrieving similar cyclones to predict cyclone tracks by considering multiple attributes (e.g., the shape and location information) (Fraedrich et al., 2003; Li et al., 2018; Roy and Kovordányi, 2012; Sievers et al., 2000). These retrieval methods are so complex that they reduce the retrieval applicability of storm surge disasters. It is necessary to develop a simple retrieval strategy of cyclone tracks only with its locations to obtain similar storm surge disasters with close locations. To search the most similar case for prediction, this retrieval model also requires

disaster information, such as hazard and vulnerability data. Although many studies provide disaster indexes to represent storm surge disasters (Guo and Li, 2020), some data are not available during the prediction stage. For example, the data related to storm surges are confidential to the public before a storm surge disaster occurs. The available data of disaster indexes should be selected to predict storm surge disasters.

This study creates a two-stage retrieval model, consisting of spatial attribute and disaster attribute search, to predict affected regions and their relative damage levels quickly. The rest of this paper is organized as follows. Section 2 describes the data sources. In Section 3, the retrieval model is explained, followed by the case study in Section 4. Then in Section 5, the conclusion is formulated.

DATA COLLECTION

Typhoon storm surge disasters with the high frequency and severe damages are selected as study objects. Data of 177 typhoon storm surge disasters between 1983 and 2019 are collected from different sources. In this study, spatial attributes are 6-hourly locations (latitudes and longitudes) of tropical cyclone track points, and they are from the Best Track Dataset (Ying et al., 2014), which are available on the platform "China Meteorological Administration Tropical Cyclone Data Center" (http://tcdata.typhoon.org.cn/en/index.html).

This study describes attributes of typhoon storm surge disasters from 3 dimensions: the typhoon storm surge hazard, the region vulnerability, and the disaster prevention and mitigation capacity. In the prediction stage, attribute indexes of typhoon storm surge disasters should be representative and available. Through analyzing the characteristics and investigating literature, attribute indexes are selected as follows.

The hazard of a typhoon storm surge disaster is represented by a surge height or water level. However, data about the hazard cannot be obtained through public platforms in time during the prediction. Other available data are selected to replace it according to the characteristics of typhoon storm surge disasters. The most representative indexes are the tropical cyclone intensity and astronomical tide water level. The astronomical tide is one part of calculating the water level. When the storm surge is over the astronomical tide, they are nonlinear coupling to generate the rise water level, varying over time and locations. Considering the complex computation in the prediction, another attribute related to the astronomical tide is selected to replace it. Actually, a storm surge combining with the astronomical high tide will lead to a more severe disaster than only a storm surge, and whether a storm surge coincides with an astronomical high tide can be forecasted before a disaster happens. Therefore, three indexes of tropical cyclones and the astronomical tide, i.e., the minimum pressure (MP) near the tropical cyclone center of the track point, two-minute mean maximum sustained wind (MSW) near the tropical cyclone center of the track point, and whether a storm surge combines with an astronomical high tide, are selected as the representation of the storm surge hazard (e.g., Feng and Liu, 2017). In this study, if a typhoon storm surge coincides with an astronomical high tide, the value of this attribute is assigned as "1"; otherwise, the value is represented by "0". The indexes of region vulnerability are retrieved from damage records of Bulletin of China Marine Disaster (1983-2019), including five kinds of factors. Given a lack of data, four categories (6 indexes) (e.g., Jin et al., 2018; Zhao and Hao, 2013) of region vulnerability are gathered (seen Table 1). Furthermore, two indexes, regional Gross Domestic Product (GDP) and accumulated length of the dike up to the standard (ALDS), are collected to reflect regional disaster prevention and mitigation capacity (e.g., Guo and Li, 2020).

Table 1. Six indexes of region vulnerability

Name of the index group	Name of each index	
Farmland indexes	Area of grain sown;	Yield of grain per unit area
Aquaculture indexes	Area of mariculture;	Aquaculture production of marine products
Seawall index	Length of the dike	
Vessel index	Number of motorized fishing boats at year-end	

Data of tropical cyclone intensity is from the same source of spatial attributes, and data about the astronomical high tide are recorded in the Bulletin of China Marine Disaster (1989-2019) and Collection of Storm Surge Disaster Historical Data in China 1949-2009 (Yu et al., 2015). Data of other indexes are all obtained in an online database "China economic and social big data research platform" (http://data.cnki.net/YearData/Analysis).

Though all 14 provincial-level coastal regions have been affected by storm surge disasters, data of only 11 regions (study regions) are gathered because there are no disaster records of Hong Kong, Macao, and Taiwan in the data sources. Therefore, the study area is 11 coastal regions: Liaoning, Hebei, Tianjin, Shandong, Jiangsu, Shanghai, Zhejiang, Fujian, Guangdong, Guangxi, and Hainan (Figure 2).

METHOD

This study applies a two-stage retrieval method to predict affected regions and their relative damage levels (shown in Figure1). To easily search tropical cyclones which induce typhoon storm surge disasters, the same "Case Number Identity (CNI)" is set for one typhoon storm surge disaster and the corresponding tropical cyclone. The first stage is to spatially search tropical cyclones by calculating the similarity of key track point locations within spatial retrieval scope. According to CNIs of selected tropical cyclones, the same CNIs of typhoon storm surge disasters are chosen. Then, by comparing values of disaster attribute similarities between past cases and the target case, the most similar past disaster case is selected to predict affected regions and their relative damage levels.

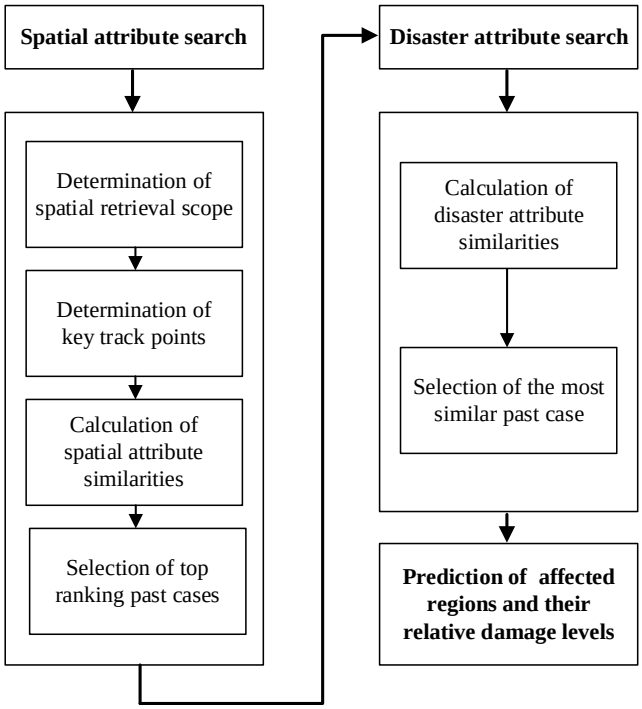

Figure 1. A retrieval model of storm surge disasters

First-Retrieval Based on Spatial Attributes

Most tropical cyclones affecting China are generated in the northwestern Pacific (Yin et al., 2013), and segments of their tracks close to coastal areas are critical to trigger typhoon storm surge disasters. Before performing the retrieval model, the initial steps are to determine the spatial search scope and select the key tropical cyclone track points for measuring the spatial similarity. The boundary line of the spatial search scope is drawn by a buffer of the regional coastline in the Geographic Information System (GIS), and this scope consists of the land scope and sea scope (Figure 2). Considering that most typhoon storm surges occur before tropical cyclones make landfall, the sea search scope is created based on the closest track points in the sea to the regional coastline.

Taking Guangdong province as an example, its sea search scope is created as follows. Firstly, all track points of cyclones inducing typhoon storm surge disasters in Guangdong are inputted to GIS. Next, points that are located at sea and closest to the regional coastline of tracks are highlighted. Then, a Guangdong coastline-centered buffer is generated, including all highlighted points. The final sea search scope is produced under expert subjective judgment to ensure the buffer is reasonable (not too big or small). Meanwhile, the Guangdong land search scope can be produced by the buffer symmetry. During this process, the land search scope is limited within the area of

Guangdong province. It is worth noting that the spatial search scope may change within a small range with different scholar's analyses, but it influences little for selecting key track points due to their sparse distribution. The method for generating the spatial search scope is applied to the rest of the regions, and then, the buffers of different provinces are transformed into one union buffer as the whole spatial search scope. In this study, the buffer radius of each coastal region is seen Table 2.

Table 2. Buffer radius of each region

	Liaoning	**Hebei**	**Tianjin**	**Shandong**	**Jiangsu**	**Shanghai**
Land buffer radius (km)	100	100	100	100	150	150
Sea buffer radius (km)	100	100	100	300	300	300
	Zhejiang	**Fujian**	**Guangdong**	**Guangxi**	**Hainan**	
Land buffer radius (km)	150	150	150	150	_	
Sea buffer radius (km)	300	100	100	250	100	

Two key points of tropical cyclone tracks are selected to retrieve the spatial similar tropical cyclones. The first key point is located at sea and closest to the regional coastline due to its huge impact on triggering typhoon storm surges, and it also reflects which region is most likely to be affected. The rest of the points of one tropical cyclone track in the spatial search scope are also important because they may occur in the buffer of other provinces causing damages in a larger scope of several regions. Therefore, the last point of one tropical cyclone track within the spatial search scope is chosen as another key point. Within the spatial search scope, the point located at sea closest to the regional coastline (Starting Point) and the last occurring point (Ending Point) of one track in buffers are key points to show the location of this track.

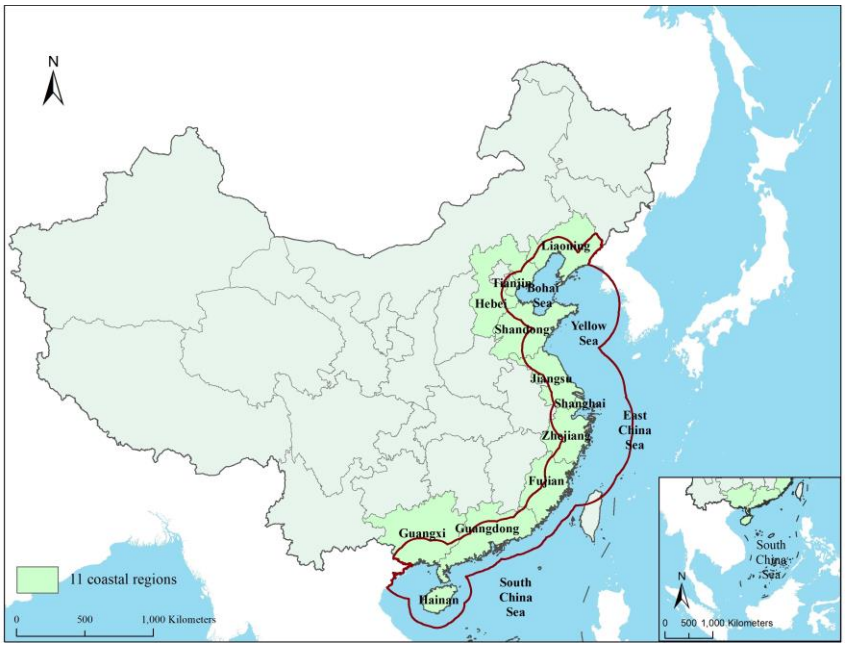

Figure 2. The area formed by the red line is the spatial retrieval scope. The scope is created based on the boundary line of 11 coastal regions.

The spatial similarity is measured by the Euclidean distance of key points between the target case and the past case. The distance of two Starting Points of the target case and one past case is calculated (Equation (1)), as well as the distance of two Ending Points (Equation (2)). These two distances are standardized, shown as Equation (3) and Equation (4), respectively. Then, these standardized distances are averaged as spatial attribute distances between the target case and one past case (Equation (5)). Finally, the similarity is a value of 1 minus the average standardized distance (Equation (6)). After calculating similarities between the target case and past cases, all values of similarities are arranged in descending order. The mathematical formulas of spatial attribute similarity are as follows:

$$D_{iT}^{S} = \left[\left(x_i^S - x_T^S \right)^2 + \left(y_i^S - y_T^S \right)^2 \right]^{\frac{1}{2}} \tag{1}$$

$$D_{iT}^{E} = \left[\left(x_i^E - x_T^E \right)^2 + \left(y_i^E - y_T^E \right)^2 \right]^{\frac{1}{2}} \tag{2}$$

$$D_{iT}^{S*} = \frac{D_{iT}^{S} - \min_i D_{iT}^{S}}{\max_i D_{iT}^{S} - \min_i D_{iT}^{S}} \tag{3}$$

$$D_{iT}^{E*} = \frac{D_{iT}^{E} - \min_i D_{iT}^{E}}{\max_i D_{iT}^{E} - \min_i D_{iT}^{E}} \tag{4}$$

$$D_{iT} = \frac{1}{2}(D_{iT}^{S*} + D_{iT}^{E*}) \tag{5}$$

$$S_{iT} = 1 - D_{iT} \tag{6}$$

Where x_i^S and y_i^S are the longitude and latitude of the Starting Point of the ith past tropical cyclone track, x_T^S and y_T^S are the longitude and latitude of the Starting Point of the target track, x_i^E and y_i^E are the longitude and latitude of the Ending Point of ith past tropical cyclone track, x_T^E and y_T^E are the longitude and latitude of the Ending Point of the target track, D_{iT}^S is the distance of Starting Points between the target case and the ith past case, D_{iT}^E is the distance of Ending Points between the target case and the ith past case, $\min_i D_{iT}^S$ and $\max_i D_{iT}^S$ are minimum and maximum of Starting Point distances among all cases, $\min_i D_{iT}^E$ and $\max_i D_{iT}^E$ are minimum and maximum of Ending Point distances among all cases, D_{iT}^{S*} and D_{iT}^{E*} are standardized distances of D_{iT}^S and D_{iT}^E, D_{iT} is the standardized distance of spatial attributes between the target case and the ith past case, S_{iT} is the similarity of spatial attributes between the target case and the ith past case.

Second-Retrieval Based On Disaster Attributes

Although the tropical cyclone locations contain spatial information about typhoon storm surge, predicting affected regions only by spatial data of tropical cyclones is incomprehensive. The spatially similar tropical cyclones can cause typhoon storm surge disasters with similar locations. Thus, the next step is to retrieve the most similar case with respect to disaster attributes from the selected cases of the first retrieval stage. Five top-ranking past cases are selected from the first retrieval into the next stage to assess disaster attribute similarities of them to the target one. The formulas of disaster attribute similarity are introduced depending on different types of attribute values. This study includes two types of disaster attributes: the symbol attribute and numerical attribute. For the symbol attribute, its similarity is given by

$$SIM(S_i^s, T^s) = \begin{cases} 1 & if\ S_i^s = T^s \\ 0 & if\ S_i^s \neq T^s \end{cases} \tag{7}$$

Where S_i^s is the value of the sth symbol attribute index of the ith past case, T^s is the value of the sth symbol attribute index of the target case, $SIM(S_i^s, T^s)$ is the similarity of the sth symbol attribute index of the ith past case to the target case.

The similarity of the numerical attribute is given by

$$SIM(S_i^j, T^j) = 1 - \frac{\left|S_i^j - T^j\right|}{\max\limits_{i+1} A^j - \min\limits_{i+1} A^j} \tag{8}$$

Where S_i^j is the value of the jth numerical attribute index of the ith past case, T^j is the value of the jth numerical attribute index of the target case, $\min\limits_{i+1} A^j$ and $\max\limits_{i+1} A^j$ are the minimum and maximum values of the jth numerical attribute index of all past cases and the target case, $SIM(S_i^j, T^j)$ is the similarity of the jth numerical attribute index of the ith past case to the target case.

The similarity of disaster attributes between the ith past case and the target case is given by

$$SIM(S_i, T) = \sum_{s=1}^{n} w_s \times SIM(S_i^s, T^s) + \sum_{j=1}^{m} w_j \times SIM(S_i^j, T^j) \tag{9}$$

Where w_s is the weight of the sth symbol attribute index, w_j is the weight of the jth numerical attribute index. These weights can be obtained by combining the expert judgment with literature investigation (e.g., Su, 2015; Yin, 2011; Zhang, 2013). The criterion for deciding weights is the importance of indexes to disasters. According to the importance of different types of indexes and the average distribution of the same type of indexes, the exact weights are shown in Table 3. After calculating the overall similarities between the target case and five past cases, the values of these similarities are arranged in descending order. The first ranked case is the most similar, and its recorded affected regions and their relative damage levels are the predicted result of the target case.

This two-stage retrieve simplifies the prediction for affected regions and their relative damage levels, taking minutes to retrieve similar past cases. If the predicted tropical cyclone information updates, the new retrieval can also be derived within minutes to update the predicted results.

Table 3. Weights of disaster attribute indexes

Name of the index group	Name of each index	Weight
Hazard indexes	MP of the Starting Point (H_1)	H_1: 0.075
	MSW of the Starting Point (H_2)	H_2: 0.075
	MP of the Ending Point (H_3)	H_3: 0.075
	MSW of the Ending Point (H_4)	H_4: 0.075
	Whether a typhoon storm surge combines with an astronomical high tide (H_5)	H_5: 0.3
Farmland indexes	Area of grain sown (V_1)	V_1: 0.05
	Yield of grain per unit area (V_2)	V_2: 0.05
Aquaculture indexes	Area of mariculture (V_3)	V_3: 0.05
	Aquaculture production of marine products (V_4)	V_4: 0.05
Seawall index	Length of the dike (V_5)	V_5: 0.1
Vessel index	Number of motorized fishing boats at year-end (V_6)	V_6: 0.1
Mitigation capacity indexes	GDP (M_1)	M_1: 0.4
	ALDS (M_2)	M_2: 0.6

CASE STUDY

In this section, three target cases of typhoon storm surge disasters are selected to verify the retrieval model. The CNI of the first target case is “1918”, and its tropical cyclone track was located in the East China Sea and passed through the coastal area of Zhejiang. The second one is the “1904” case, and the corresponding tropical cyclone passed through Hainan and entered the South China Sea. The tropical cyclone of the last one, the “1822” case, made landfall in Guangdong and entered Guangxi, and then disappeared. These tropical cyclone tracks are shown

in Figure 3, and their key points are bigger than other track points. The location information of key points is shown in Table 4.

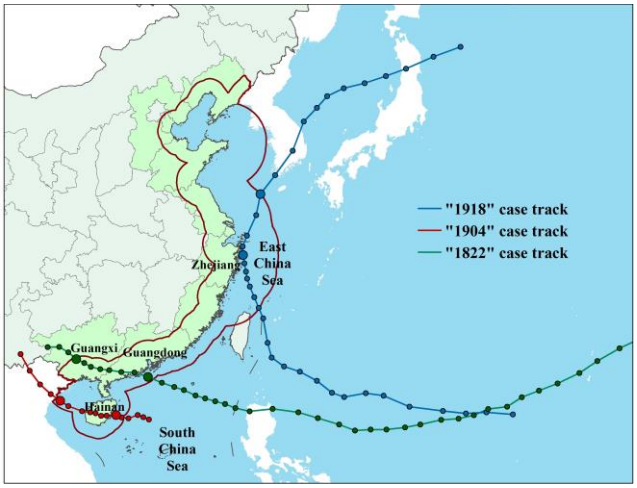

Figure 3. Tropical cyclone tracks of three target cases

Table 4. Spatial attributes of three target cases

Spatial attributes	"1918" Starting Point	"1918" Ending Point	"1904" Starting Point	"1904" Ending Point	"1822" Starting Point	"1822" Ending Point
Latitude (°N)	29.3	33.3	18.9	20.1	21.5	23.1
Longitude (°E)	122.2	124.6	110.9	106.8	113.5	108.1

The top five similar tracks and corresponding similarities are illustrated in Figure 4 and Table 5. It is found that the selected tracks in the spatial search scope are close to their target track, which can contribute to the same affected regions as target ones.

Table 5. Top five similar cases of each target case based on spatial attributes

Similar tracks of the "1918"		Similar tracks of the"1904"		Similar tracks of the "1822"	
CNI of past case	Similarity	CNI of past case	Similarity	CNI of past case	Similarity
1416	0.950	0508	0.995	9316	0.986
0407	0.932	0917	0.991	0915	0.983
1509	0.916	1508	0.983	1208	0.978
0205	0.907	9106	0.981	1713	0.971
0515	0.901	0518	0.978	0606	0.963

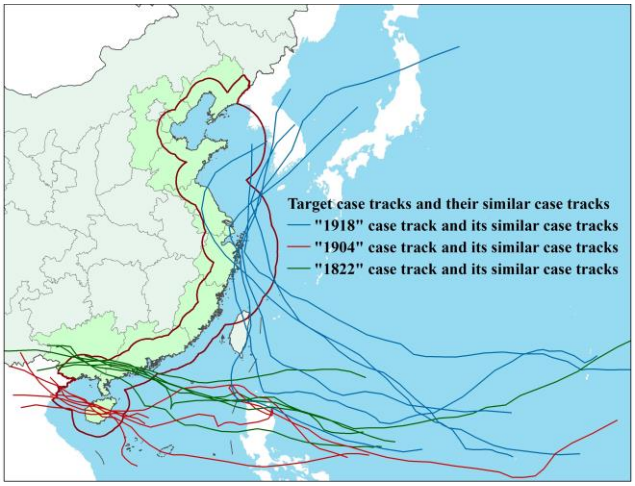

Figure 4. Three target tracks and their top-ranking five tracks

According to Equations (7)-(9), the similarities based on disaster attributes are listed in Table 6.

Table 6. Ranking five similar cases of each target case based on disaster attributes

Similar cases of the“1918”		Similar cases of the “1904”		Similar cases of the “1822”	
CNI of past case	Similarity	CNI of past case	Similarity	CNI of past case	Similarity
1509	1.599	1508	1.576	1208	1.560
1416	1.467	0917	1.026	1713	1.498
0407	0.641	0508	0.864	0915	0.917
0515	0.555	0518	0.650	0606	0.738
0205	0.500	9106	0.368	9316	0.218

After two steps to rank the most similar past case of each target case, the recorded affected regions of each selected past case are treated as the predicted affected regions of each target case. The proportions of the direct economic loss in different regions of the past case are regarded as the relative damage levels among predicted affected regions of the target case. For the “1918” typhoon storm surge disaster, the recorded affected region is Zhejiang (Figure 5(b)), whereas the predicted scope is larger, including the other three regions (Fujian, Jiangsu, and Shanghai) (Figure 5(c)). Although more regions are predicted to be affected, the predicted result can also provide valuable information to make a decision considering the relative damage levels among affected regions. The most affected region of the predicted result is Zhejiang, which is also the only affected region of the actual record. Before this typhoon storm surge disaster occurs, the four province-level governments need to pay more attention to the development of the disaster, especially Zhejiang officials, and make mitigation plans. The affected regions by typhoon storm surge disaster triggered by this type of cyclone track are difficult to predict accurately due to its long track in the sea search scope and few past cases. Furthermore, this predicted result can be reasonable because of statistic error and reducing loss by effective measures in lightly affected regions.

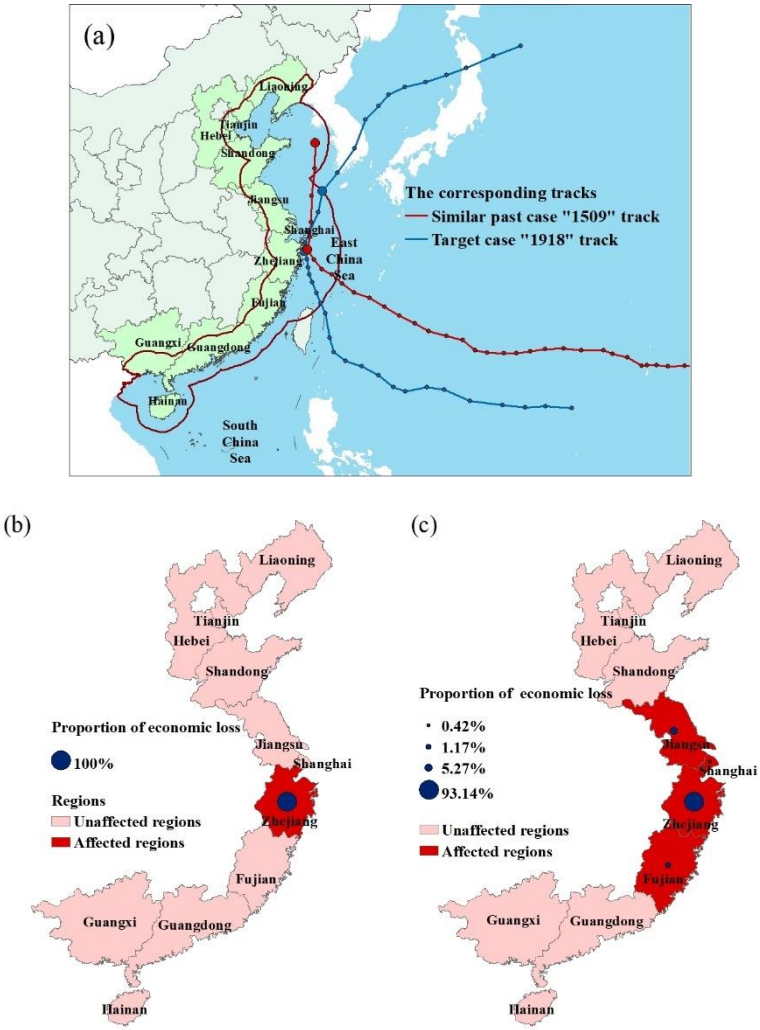

Figure 5. (a) Tracks of the "1918" case and its similar case "1509". (b) The recorded affected region of the "1918" case. (c) The predicted result of the "1918" case

The "1904" typhoon storm surge disaster is induced by another kind of tropical cyclone tracks, making landfall in one region and then moving to the sea. From Figure 6(a), Guangdong, Guangxi, and Hainan are possibly affected by the "1904" typhoon storm surge disaster. The recorded and predicted results both consist of the same province: Guangxi (Figure 6(b) and (c)), which is mainly resulted from high similarities of spatial and disaster attributes between the past "1508" case and target "1904" case. It is worth noting that this typhoon storm surge disaster records no economic loss in Hainan, though it passed through Hainan. This phenomenon can be related to multiple factors, such as the climate, geomorphology, and prevention capacity. The affected region can be predicted fast and accurately by simplifying related influencing factors based on similar disaster conditions. Despite the low intensity of this typhoon storm surge disaster, the Guangxi government also needs to respond to it to reduce damages.

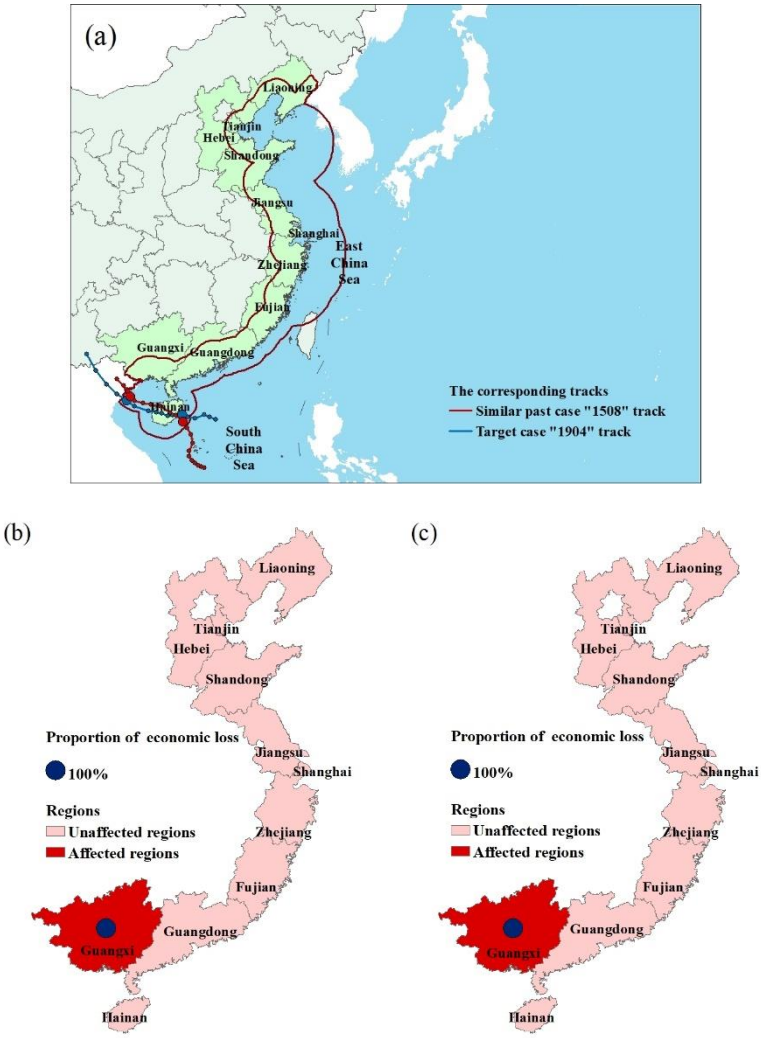

Figure 6. (a) Tracks of the "1904" case and its similar case "1508". (b) The recorded affected region of the "1904" case. (c) The predicted result of the "1904" case

The predicted result of the "1822" case includes two affected regions: Guangdong and Guangxi (Figure 7(c)), smaller than the recorded affected scope (Figure 7(b)). However, the direct economic loss in Fujian, the extra actual affected region, only accounts for 0.08% of the total loss, which reflects the relatively low damage level. Given that statistic error of the direct economic loss and attribute differences of the similar past case to this target one, the result of an unpredicted region with slight loss can be acceptable. It is clear that whether for the recorded or predicted result of the "1822" case, the damage degree in Guangdong reaches a high level. The result of predicted regions provides an essential base for decision makers to undertake response at different levels among these regions.

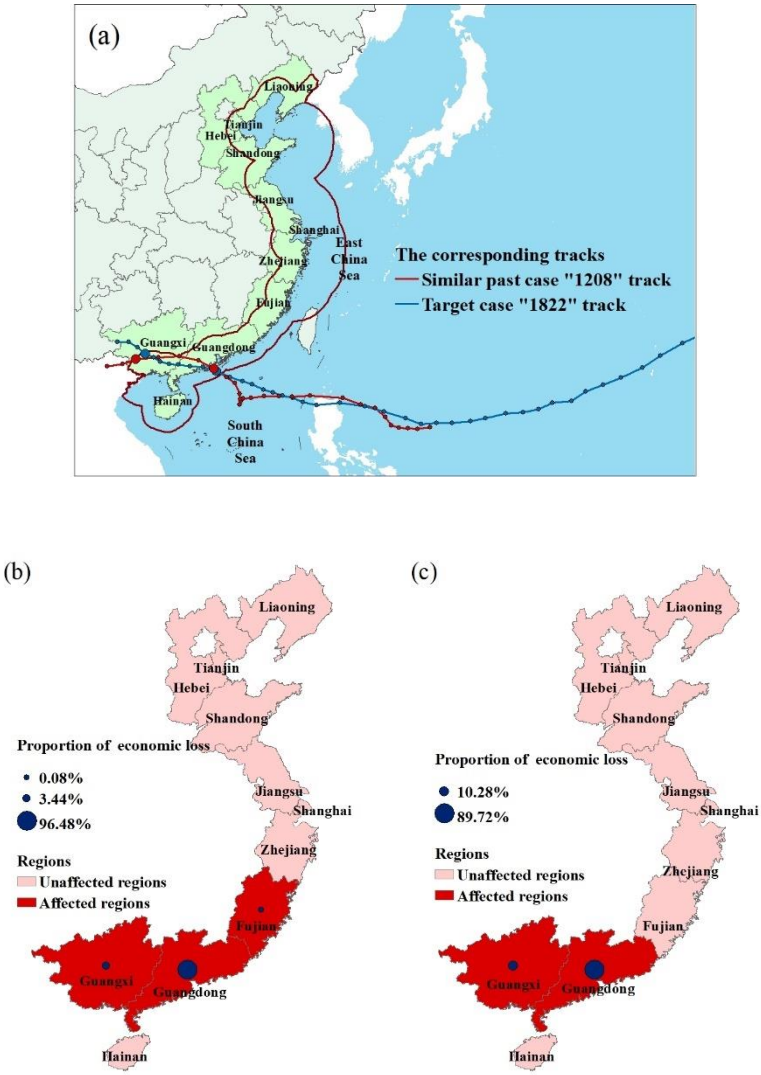

Figure 7. (a) Tracks of the "1822" case and its similar case "1208". (b) The recorded affected regions of the "1822" case. (c) The predicted result of the "1822" case

CONCLUSION

This study establishes a two-stage model to retrieve the most similar past case to predict affected regions and their relative damage levels of an upcoming typhoon storm surge disaster. The first stage is to search past cases by assessing spatial similarities of tropical cyclones which trigger typhoon storm surge disasters. Secondly, the best past case is chosen by ranking disaster similarities among selected top five past cases from the first stage, and finally, the affected regions and their relative damage levels of the target case are estimated as outputs of our retrieval model. Given lacking data of attributes in the prediction, some attributes are replaced by analyzing disaster characteristics to complete the prediction. In addition, this predicted result only takes minutes to be updated as the input data change. This model helps make real-time decisions to mitigate disasters.

From the results of three kinds of target cases, this model performs well. The most affected regions in all three target cases can be forecasted to warn the relevant government to implement further preparedness. The proportions of the direct economic loss of the most similar case can reflect the relative damage levels among predicted regions, which can be related to different levels of emergency response plans for the storm surge disaster in different provinces. There are also some biases for predicting the affected regions owing to manual statistic error of the direct economic loss and lack of similar past cases. However, the predicted requirement for these extra or neglected regions is not strict regarding their very light damages.

This retrieval model provides a possibility to predict affected regions and their relative damage levels of storm

surge disaster by the most similar past case. Different combinations of case attributes and weights may retrieve different similar past cases. We need to collect more data and test different attribute weights to better predict affected regions at the city-level or county-level. Furthermore, it is necessary to propose the adjustment algorithm to predict the spatial distribution of the storm surge disaster damages. For the prediction result, we can also develop a probabilistic methodology to describe affected regions and their relative damage levels by many similar cases in future work.

ACKNOWLEDGMENTS

This work was supported by the National Key Research and Development Plan of China [Grant Number 2017YFC1405301], National Natural Science Foundation of China [Grant Number 42007420], and China Postdoctoral Science Foundation [Grant Number 2020M670284].

REFERENCES

Chen, W. K., Sui, G. J. and Tang, D. L. (2011) Predicting the economic loss of typhoon by case base reasoning and fuzzy theory, *2011 International Conference on Machine Learning and Cybernetics*, 1, 254–257.

Deng, S. G. and Li, W. S. (2020) Spatial case revision in case-based reasoning for risk assessment of geological disasters, *Geomatics, Natural Hazards and Risk*, 11, 1, 1052–1074.

Fang, J., Hinkel, J., Shi, P., Nicholls, R. J., Saini, Y., Brown, S., Liu, W., & Shi, X. (2017). Spatial-temporal changes of coastal and marine disasters risks and impacts in Mainland China. Ocean & Coastal Management, 139, 125–140.

Fei, L. and Feng, Y. (2020) A novel retrieval strategy for case-based reasoning based on attitudinal Choquet integral, *Engineering Applications of Artificial Intelligence*, 94, 103791.

Feng, Q., & Liu, Q. (2017) Pre-assessment for the loss caused by storm surge based on the SVM-BP neural network (in Chinese), *Marine Environmental Science*, 36, 4, 615–621.

Feng, S. Z. (1982) Introduction to the Storm Surge, Science Press, Beijing.

Fraedrich, K., Raible, C. C. and Sielmann, F. (2003) Analog ensemble forecasts of tropical cyclone tracks in the Australian region, *Weather and Forecasting*, 18, 1, 3–11.

Guo, T. and Li, G. (2020) Study on methods to identify the impact factors of economic losses due to typhoon storm surge based on confirmatory factor analysis, *Natural Hazards*, 100, 2, 515–534.

Helderop, E. and Grubesic, T. H. (2019) Social, geomorphic, and climatic factors driving U.S. coastal city vulnerability to storm surge flooding, *Ocean and Coastal Management*, 181, 104902.

Hsu, C.-H., Olivera, F., and Irish, J. L. (2018) A hurricane surge risk assessment framework using the joint probability method and surge response functions, *Natural Hazards*, 91, 1, 7–28.

Jin, X., Shi, X. X., Gao, J. T., Xu, T. Bin, and Yin, K. D. (2018) Evaluation of loss due to storm surge disasters in China based on econometric model groups, *International Journal of Environmental Research and Public Health*, 15, 4, 604.

Li, Y., Xu, L., Hu, B. and Li, Y. (2018) A Similarity Retrieval Approach for Weighted Track and Ambient Field of Tropical Cyclones, In F. Wu and P. Zhou (Eds.), *Advances in Energy Science and Environment Engineering Ii*, 1944, 020024.

Lin, N., Emanuel, K., Oppenheimer, M. and Vanmarcke, E. (2012) Physically based assessment of hurricane surge threat under climate change, *Nature Climate Change*, 2, 6, 462–467.

Ling, Z., Wang, G. H., Wang, C. Z. and Fan, Z. S. (2011) Different effects of tropical cyclones generated in the South China Sea and the northwest Pacific on the summer South China Sea circulation, *Journal of Oceanography*, 67, 3, 347–355.

Liu, Q. Z., Shi, X. W., Guo, Z. X., Wang, X. N., Gong, M. X., Lu, M., Wang, Y. X., and Chen, F. Y. (2019) Technical directives for risk assessment and zoning of marine disaster-Part 1: Storm surge (HY/T 0273-2019).

Liu, X., Jiang, W., Yang, B. and Baugh, J. (2018) Numerical study on factors influencing typhoon-induced storm surge distribution in Zhanjiang Harbor, *Estuarine Coastal and Shelf Science*, 215, 39–51.

Pedro, J. S., Burstein, F. and Sharp, A. (2005) A case-based fuzzy multicriteria decision support model for tropical cyclone forecasting, *European Journal of Operational Research*, 160, 2, 308–324.

Reffitt, M., Orescanin, M. M., Massey, C., Raubenheimer, B., Jensen, R. E. and Elgar, S. (2020) Modeling Storm Surge in a Small Tidal Two-Inlet System, *Journal of Waterway, Port, Coastal, and Ocean Engineering*, 146, 6, 04020043.

Resio, D. T. and Westerink, J. J. (2008) Modeling the physics of storm surges, *Physics Today*, 61, 9, 33–38.

Roy, C. and Kovordányi, R. (2012) Tropical cyclone track forecasting techniques - A review, In *Atmospheric Research*, 104, 40-69.

Shi, X. W., Liu, S., Yang, S. N., Liu, Q. Z., Tan, J. and Guo, Z. X. (2015) Spatial-temporal distribution of storm surge damage in the coastal areas of China, *Natural Hazards*, 79, 1, 237–247.

Shi, X. W., Yu, P. B., Guo, Z. X., Sun, Z. L., Chen, F. Y., Wu, X. G., Cheng, W. L. and Zeng, J. (2020) Simulation of storm surge inundation under different typhoon intensity scenarios: case study of Pingyang County, China, *Natural Hazards and Earth System Sciences*, 20, 10, 2777–2790.

Sievers, O., Fraedrich, K. and Raible, C. C. (2000) Self-adapting analog ensemble predictions of tropical cyclone tracks, *Weather and Forecasting*, 15, 5, 623–629.

State Oceanic Administration of China. (2017) Bulletin of China Marine Disaster.

Su, C. N. (2015) Amendment of PPDC model and its application research in the risk zoning of storm surge disaster (in Chinese), Ocean University of China.

Tanim, A. H. and Goharian, E. (2020) Developing a Hybrid Modeling and Multivariate Analysis Framework for Storm Surge and Runoff Interactions in Urban Coastal Flooding, *Journal of Hydrology*, 125670.

Yang, S., Liu, X. and Liu, Q. (2016) A storm surge projection and disaster risk assessment model for China coastal areas, *Natural Hazards*, 84, 1, 649–667.

Yin, J. (2011) Study on the risk of typhoon storm tide in China coastal area (in Chinese), East China Normal University.

Yin, J., Yin, Z. N. and Xu, S. Y. (2013) Composite risk assessment of typhoon-induced disaster for China's coastal area, *Natural Hazards*, 69, 3, 1423–1434.

Ying, M., Zhang, W., Yu, H., Lu, X. Q., Feng, J. X., Fan, Y. X., Zhu, Y. T. and Chen, D. Q. (2014) An Overview of the China Meteorological Administration Tropical Cyclone Database, *Journal of Atmospheric and Oceanic Technology*, 31, 2, 287–301.

Yu, F. J., Dong, J. X. and Ye, L. (2015) Collection of Storm Surge Disasters Historical Data in China 1949–2009 (in Chinese), China Ocean Press, Beijing.

Zhang, Y. H. (2013) Risk assessments on Storm surge and coastal zoning management (in Chinese), Ocean University of China.

Zhao, L. Di, and Hao, Q. (2013) Classification and Prediction of Economic Losses - Storm Surge Disasters in Guangdong Province of China, *Advanced Materials Research.*, 709, 928–935.

Zheng, Y. and Sun, H. (2020) An Integrated Approach for the Simulation Modeling and Risk Assessment of Coastal Flooding, *Water*, 12, 8, 2076.