

Effective Storm Surge Risk Assessment and Deep Reinforcement Learning Based Evacuation Planning: A Case Study of Daya Bay Petrochemical Industrial Zone

Chuanfeng Liu², Yan Li¹, Wenjuan Li³, Hao Qin², Lin Mu^{1,4}, Si Wang^{1,4}, Darong Liu², and Kai Zhou³

¹College of Life Science and Oceanography, Shenzhen University

²College of Oceanic and Atmospheric Sciences, Ocean University of China

³Shenzhen Marine Development and Promotion Center

⁴College of Computer Science and Software Engineering, Shenzhen University

Correspondence: Yan Li (liyan_ocean@szu.edu.cn), Wenjuan Li (lwenjuan@hotmail.com)

Abstract. Storm surge is one of the most destructive marine disasters, characterized by abnormal and temporary rises in water levels during intense storms, leading to extreme inland flooding in the coastal area. storm surge risk assessment and evacuation planning, play a crucial role in saving lives and mitigating disasters. Conventional risk assessment struggles to meet the demands of refined risk evaluation research for small-scale elements, such as roads, and current evacuation plans are generally based on broader regional scales, failing to provide effective road-level evacuation planning for evacuees. This study developed ~~four~~ five typical typhoon scenarios for the coupled ADCIRC-SWAN model to simulate storm surge inundation. Combining these simulations with road network, storm surge risk assessment was conducted in the Daya Bay Petrochemical Industrial Zone, a vulnerable low-lying coastal region of Huizhou City, China, ~~which is frequently affected by storm surge-driven flooding~~. Based on the risk assessment, a combination of the Deep Q-Network (DQN) model and raster environment was employed to develop real-time evacuation plans during storm surge events. To address the DQN model's convergence challenges, compressed search space and ~~multi-reward~~ navigational reward methods were proposed, ~~which were specifically designed to enhance the DQN model's capacity to path planning problems within large-scale raster environments~~. 1000 starting points were randomly selected for path-route planning, and the results indicate that the proposed method is highly effective in devising optimal evacuation routes with minimal deviation, offering valuable guidance for evacuees during real-world storm surges.

Copyright statement.

1 Introduction

A storm surge is an abnormal and temporary rise of water that occurs during intense storms. This sudden rise in sea level can lead to extreme inland flooding in coastal communities ~~especially when an advancing surge coincides with astronomical high tide~~ (Wang et al., 2021b). Storm surge is one of the most dangerous and destructive natural hazards to life and property

along the coastline and kilometers inland in the world (UNISDR et al., 2015). In the US, coastal flooding from storm surge was responsible for 49 % of ~~hurricane-or-tropical-storm-related~~ storm-related fatalities during the period from 1994 to 2003 (Rappaport, 2014). When Hurricane Katrina struck the southeastern United States in 2005, an estimated 1,577 people died, causing ~~USD-108 billion in-property-damages~~ USD in property losses (Rhome and Brown, 2006). The storm surge from Super Typhoon ~~Haiyan'~~ Haiyan', hitting the Philippines in 2013, ~~was-estimated-to-be-7.5-high-leading-to~~ led to more than 7,000 persons losing their lives ~~along-the-coastline~~ (Mas et al., 2015). In China, from 1998 to ~~2019~~ 2020, the average annual economic losses resulting from storm surge flooding each year is approximately 10.17 billion RMB, which is equivalent to 96 % of the total direct economic losses from all types of marine disasters (China Marine disaster bulletin, 2021). Recent studies ~~suggested-that~~ indicated a potential escalation in the number of ~~people-facing-the-risk-of~~ individuals at risk from storm surge flooding ~~and-the-losses-in-property-and-human-lives-during-typhoon-events-could-continue-to-increase-in-the-future-due-to~~ accelerating sea-level rise and more intense hurricanes, along with associated property damage and loss of life (Merkens et al., 2016; Oppenheimer et al., 2019; Snaiki et al., 2020). Physical barriers, ~~which-run-while~~ parallel to the shoreline, ~~alone-cannot prevent-all-possible-damages-that-are-insufficient-to-prevent-all-potential-damages-to~~ urban settlements and infrastructure ~~can suffer~~ during storm surge ~~floodingevents~~. With increasing potential victims and economic losses, it is of paramount importance to perform risk assessments and develop evacuation plans to mitigate the ~~risks-risk~~ associated with storm ~~surgesurge~~.

Storm surge risk assessment ~~entails-the-identification-and-evaluation-of-potential-hazards-and-associated-risks-within-a~~ specified region, as well as the severity of their potential consequences (Wang et al., 2021a). Conducting effective and practical storm surge risk assessments lays the groundwork and establishes the premise for devising well-founded evacuation plans. Generally, based on simulated inundation scenarios, can provide predictive and analytical information on the danger of storm surge disasters in a region (Wang et al., 2021a). The risk assessment approach aims to quantify the damage and risk into value, where the value is the interaction between the geographic coverage of the hazard, the exposure of elements to the inundation depths and extents are regarded as the most common factors to measure the tangible risk hazard, and the computation of potential storm surge heights and the maximum inundated area under different hypothetical typhoon events is one of the major components of vulnerability of exposed elements (Granger, 2003; Kron, 2005; Lavell et al., 2012; Koks et al., 2015). In recent years, various storm surge risk assessment models have been established in various countries, and some case studies have been conducted (e.g. Zerger, 2002; Benavente et al., 2006; Lin et al., 2010). In the study of storm surge risk assessment (Thieken et al., 2007; Merz et al., 2010). The computation of maximum potential impact due to storm surge is usually performed using the numerical model by taking into account the atmospheric pressure, landfall location, varying forward speed, the radius of maximum wind, and tracking. So far, many hydrological models such as the, ~~a-comprehensive-assessment-of-storm surge risk was obtained by using numerical models to simulate the inundation scenarios and considering the damage of different types of vulnerable elements. The~~ Advanced Circulation (ADCIRC, Luettich et al.) model, ~~Finite-Volume-Community-Ocean Model (FVCOM, Chen et al.), and designed to address two- and three-dimensional hydrodynamic free surface circulation challenges, has been widely applied to simulate tide- and wind-driven circulations. The~~ Simulating Waves Nearshore (SWAN, Delft University of Technology) model have been applied to simulate tide, wave and storm surge in different regions in the world, obtaining good prediction accuracy in previous studies. In this study, the coupled ~~Booij et al.) model is a numerical~~

wave model, which computes the wave action density spectrum by solving the wave action balance equation. By employing a bidirectional coupling approach, the ADCIRC+SWAN model was employed to simulate storm surge flooding, and quantitative risk assessments were conducted by combining the extents and depths of the flooding is a widely used model to simulate coastal storm surges and tidal floods induced by tropical cyclones, and has shown good performance in previous studies (e.g. Dietrich et al., 2011; Suh et al., 2015; Wang et al., 2018).

The research on emergency evacuation originated in the early 20th century, and its core task is the development of evacuation plans, including the identification of disaster shelters and the planning of evacuation routes (Alsnih and Stopher, 2004). In some coastal areas vulnerable to storm surges, there might be a lack of designated disaster shelters. Consequently, it is essential to identify suitable facilities to serve as shelters, and a framework was proposed based on several criteria, including structural stability, waterproofing, and capacity. Conventional shortest-path algorithms, such as Dijkstra's and A* algorithms, have been enhanced and employed in emergency evacuation planning (Samah et al., 2015; Astri et al., 2020, e.g.,). Additionally, heuristic approaches, including the PSO algorithm, genetic algorithm, and ant colony algorithm have been introduced to identify optimal routes within intricate environments (Li et al., 2020; Goerigk et al., 2014; Forcael et al., 2014, e.g.,). Nevertheless, existing emergency evacuation plans predominantly adopts the perspective of administrators, furnishing evacuees with complete environmental information. In realistic circumstances, evacuees often encounter obstacles in obtaining the entire environmental information, instead possessing merely a limited awareness of their immediate surroundings. And the techniques employed for issuing evacuation mandates and cautionary advisories lack the timeliness in guiding individuals to (e.g., Li et al., 2020; Goerigk et al., 2014; Forcael et al., 2014). In recent years, reinforcement learning (RL), has been a hot topic in the field of machine learning, and the most appropriate escape routes where actual evacuation distances may fluctuate as the evolving environment of the catastrophe. Deep reinforcement learning (DRL) is a new paradigm of deep learning, and since its inception, it has rapidly evolved with the proposal of various basic idea is to learn the optimal strategy by maximizing the cumulative reward obtained from the interaction between the intelligent agent and the environment (Sutton and Barto, 2018). In the increasingly complex real-world tasks, deep learning (DL) can be utilized to better learn the abstract features of large-scale input data, which can then be used to optimize the RL's strategy learning process. In 2013, Mnih et al. proposed the Deep Q-Network (DQN), which was applied to solve visual-based control decision problems. Since then, the field of deep reinforcement learning (DRL) has experienced rapid development, arising numerous efficient algorithms, such as Deep Q-Network (DQN, Mnih et al.), Deterministic Policy Gradient (DDPG, Lillicrap et al.), Asynchronous Advantage Actor-Critic (A3C, Mnih et al.), and Proximal Policy Optimization (PPO, Schulman et al.). DRL has found widespread application in diverse domains, including game-playing (Mnih et al., 2015), autonomous navigation (Sallab et al., 2017), and industrial regulation (Levine et al., 2018), demonstrating strong learning and generalization capabilities, based on Markov decision processes and deep neural networks, offers an effective solution for the optimization of evacuation route planning. Yu et al. utilized DRL model to develop a navigation system for an agent in a maze. Zhang et al. created a scenario with obstacles in a room and addressed the problem of emergency evacuation within the room using DRL. Ni et al. simulated buildings in a fire scenario and used an improved double deep Q network to solve evacuation route planning. The merit of DRL algorithms in emergency

90 evacuation stems from their capacity to function without a priori comprehension of the entire environment. ~~In this study, an enhanced-~~

Storm surge risk assessment and evacuation route planning can provide decision-making support for local governments in disaster prevention and reduction, and provide timely and effective emergency evacuation guidance for affected area residents. To date, there are currently no cases of comprehensive storm surge disaster mitigation research integrating these two aspects.

95 Based on the current studies, the main problems faced include:(1) Storm surge endanger coastal regions, with the existing strategies falling short due to their neglect of small-scale elements like road risk assessment, hindering fine-tuned disaster management. (2) Conventional evacuation route planning is limited by its oversimplification of complex environments and inadequate consideration of victims' needs. The unpredictability of storm surges and incomplete regional information further complicate the maintenance of effective plans. (3) While deep reinforcement learning ~~method was proposed to assist individuals~~

100 ~~in evacuating to the nearest shelter based on their surrounding environment. Additionally, to tackle convergence challenges in path planning using DRL in~~ shows promise in route planning, it faces convergence challenges with large-scale areas, two compression methods were proposed to significantly reduce the problem size.

The case study area of this research focuses applications. This study focused on the low-elevation coastal regions of the Daya Bay district, ~~which is periodically exposed to tropical cyclones and frequently affected by storm surge driven coastal flooding.~~

105 ~~The work in this study involves: 1. employing~~ utilizing the coupled ADCIRC+SWAN numerical model to simulate storm surge flooding for risk assessments; 2. identifying disaster shelters; 3. constructing a simulation environment; 4. utilizing enhanced DQN for evacuation path planning. ~~And the main contribution is to develop a real-time effective emergency evacuation plan for individuals with limited awareness of their surrounding environment in a large-scale region. The rest of the paper is organized~~

scenarios, and subsequently conducting a comprehensive risk assessment of the road network. Based on the risk assessment,

110 a DRL model was utilized to provide intelligent evacuation route planning for evacuees. The main contributions of this work are summarized as follows: ~~Section 3 delineates the coupled numerical model utilized to simulate storm surge flooding during typhoon events and the deep reinforcement learning algorithm used to explore optimal routes to evacuate from the floods on the road network. Thereafter, in the Section 4, the implementation of the proposed method to simulate depths and extents of storm surge flooding and recommended evacuation routes is demonstrated with the example of the coastal area of Huizhou.~~

115 ~~Subsequently, the findings and analyses pertaining to the numerical simulations of inundation depths and extents of storm surge flooding, as well as the emergency flood evacuation simulations based on DRL in coastal regions, are presented. Finally, the conclusions and suggestions for future work were discussed in Section 5~~

- By analyzing historical typhoon data, five typical typhoon intensities affecting the study area were identified. Using the coupled ADCIRC+SWAN model, the storm surge inundation process in the study area under the five typhoon scenarios was simulated.
 - Developing a refined road risk assessment model for storm surges, taking into account the extent of inundation, inundation depth, and type of road, to analyze the traffic conditions and conduct a fine-grained risk assessment of the road network in the study area.
- 120

- Reframing the route planning problem in a raster environment as a continuous decision-making problem, and using a DRL algorithm to plan evacuation routes on road networks, which can provide real-time and effective evacuation guidance based on the limited environment around evacuees.
- To address the challenges faced by the DQN model in a large raster environment for route planning, this study proposed compressed search space and navigational rewards methods to optimize the traditional DQN model, enabling it to better suit evacuation route planning.

2 Study Area

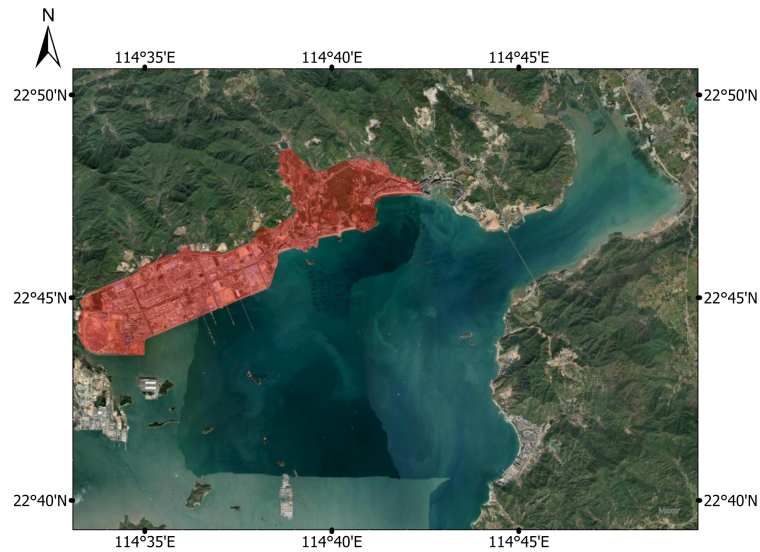


Figure 1. The Daya Bay. The study area is in red.

Daya Bay district is located in the southern region of Huizhou City, Guangdong Province. It has a total land area of 293 km² and a population of 0.45 million, which is concentrated most highly in coastal areas, in 2021. ~~In addition, the~~ The Daya Bay Petrochemical Industrial Zone, situated in the north-eastern part of Daya Bay, was listed as a national petrochemical industrial base. It has formed an annual production capacity of 22 million tons of oil refining and 2.2 million tons of ethylene in 2021, which ranks first in China in terms of the scale of petrochemical-refining integration. Industrial facilities and critical infrastructure in this area are vulnerable to storm surge-driven coastal flooding during typhoon events, leading to devastating losses of life and property. ~~The peak water level induced by severe storm surge during super typhoon Mangkhut on 16 September 2018 at Huizhou gauging station, closest to the Daya Bay district, rose to 349, causing estimated economic losses of 577.39 million RMB.~~ In the context of substantial sea-level rise and urban extent along low-lying coastal areas, most communities across the

140 Daya Bay district will likely face higher storm surge flooding risk in the future. It is crucial and essential to create the storm surge risk maps for raising awareness about areas at risk and making evacuation plans to minimize the loss and damage. The study area is shown in Fig. 2-1.

3 Methodology

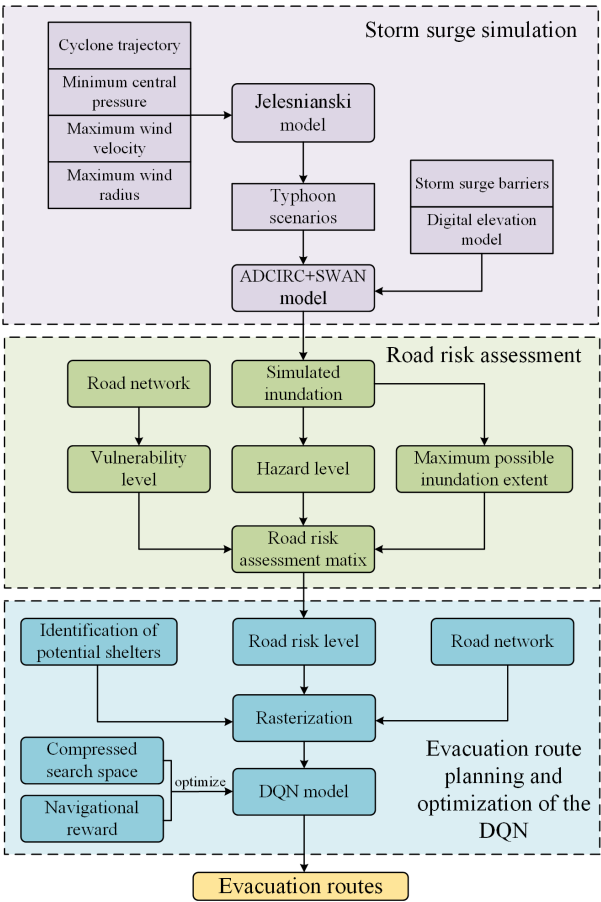


Figure 2. The Daya Bay overview of the work. The study area is in red.

4 Methodology

145 As storm surges assail, the goal of storm surge evacuation planning is to provide guiding paths to shelters for the people in the affected area, despite their limited environmental knowledge. The risk assessment and evacuation route planning for storm

surge play a crucial role in saving lives and mitigating disasters. The significance of this work lies in the ability to simulate storm surge flooding for risk assessments and to find routes to the nearest disaster ~~based on the road network.~~ In this study, ~~an effective method for planning the evacuation path during a storm surge is proposed.~~ The specific research method includes two segments: environmental modeling and deep reinforcement learning. ~~In Section ??, the process of simulating storm surges using numerical techniques and establishing an authentic raster environment model for the study area is delineated.~~ In Section ??, the principles and procedures of the Markov Decision Process (MDP) are described. Based on the MDP, we redefined the path planning problem with the raster environment and proposed two optimization methods to reduce the problem scale. the method used to perform risk assessment and evacuation route plan consists of following four sections: (1) storm surge simulation; (2) road risk assessment; (3) evacuation route planning; (4) optimization of the DQN. Please refer to Fig. 1-2 for an overview and the sections for more information. ~~The overview of the work.~~

3.1 ~~Environmental modeling~~Storm surge simulation

~~The purpose of environmental modeling lies in the accurate simulation of the environmental information of the study area. This work involves storm surge simulation and risk assessments, regional road network modeling, and identification of potential disaster shelter facilities. To facilitate computational engagement, all environmental data have been rasterized. The rasterized environment, with a three-dimensional construct, consists of regional risk levels, road networks, and disaster shelters, and these constituent elements are stored within three separate matrices. The whole flow chart of environmental modeling is depicted in Fig. 3. The flow chart of environmental modeling.~~

3.1.1 ~~Simulation of storm surge and risk assessments~~

~~Storm surge is a phenomenon wherein seawater levels experience a marked elevation due to turbulent atmospheric perturbations, such as typhoons and cyclones. The ADCIRC model, designed to address two- and three-dimensional hydrodynamic free surface circulation challenges, has been widely applied to simulate tide- and wind-driven circulations. The SWAN model is-~~ Given the difficulty in obtaining measured typhoon data, a numerical wave model, which computes the wave action density spectrum by solving the wave action balance equation widely adopted method is to construct a theoretical wind field model using the wind gradient formula (Jelesnianski, 1965; Willoughby and Rahn, 2004). In this study, the Jelesnianski hurricane model (Jelesnianski, 1965) was utilized to generate wind field, and the coupled ADCIRC+SWAN model, which integrates storm surge and wave interactions, was employed to simulate typhoon-induced storm surge flooding during typhoon events. The computational domain in this study covered the coastal region of Huizhou, as shown in Fig. 4. The model contained grids consisting of 74328 units and 38407 nodes. The coverage extended from 106.0° E to 128.0° E in longitude and from 13.0° N to 28.0° N in latitude. The 11 major tidal components of semidiurnal and diurnal frequencies (M2, N2, S2, K2, K1, O1, P1, Q1, MS4, M4, M6) were included. The Jelesnianski hurricane model calculates the air pressure and the wind. The ADCIRC model transfer wind field, water levels, and currents to the SWAN model every 600, while the SWAN model passes the wave radiation back to the ADCIRC model on the same unstructured finite element mesh, which can simulate storm surge and produce coastal flooding for the study area. The computational domain.

180 The coupled model is driven by the wind field, rendering the accuracy of storm surge model outcomes intrinsically reliant on the quality of the wind field model. Given the challenges in obtaining measured data, a prevalent approach involves constructing a theoretical wind field model based on the wind gradient formula (Jelesnianski, 1965; Willoughby and Rahn, 2004). Utilizing the Jelesnianski model necessitates four used to provide wind forcing. This approach requires the specification of four critical parameters: the cyclone track-trajectory T , the minimum central pressure P_c , the maximum wind velocity V_{max} , and the maximum wind radius R_{max} . The R_{max} may be approximated through several empirical equations (Vickery et al., 2000; Cheung et al., 2007), as delineated below:

$$R_{max_1} = \exp(2.635 - 0.00005086\Delta P^2 + 0.0394899\theta) \quad (1)$$

$$R_{max_2} = 1119 \times \Delta P^{-0.806} \quad (2)$$

$$R_{max_3} = R_k - 0.4 \times (P_c - 900) + 0.01 \times (P_c - 900)^2 \quad R_{max_4} = 260.93 \times \Delta P^{-0.512} \quad (3)$$

190 where ΔP indicates the pressure difference between the minimum central pressure P_c and ambient pressure, and in this study, the ambient pressure is 1010 hPa; θ represents the latitude of the storm's center; R_k is an empirical constant usually taking the value range of [30,60], and we take the $R_k = 50$. The $\{R_{max_1}, R_{max_2}, R_{max_3}, R_{max_4}\}$ are all estimates of R_{max} and the final R_{max} we adopted was their average, i.e.

$$R_{max} = \frac{R_{max_1} + R_{max_2} + R_{max_3} + R_{max_4}}{4} \quad (4)$$

195 Given a pressure difference ΔP , the maximum wind velocity V_{max} can be estimated by a empirical equation (Atkinson and Holliday, 1977):

$$V_{max} = 3.7237 \times \Delta P^{0.6065} \quad (5)$$

200 To simulate the inundation extents and depths under a specific typhoon According to the Tropical Cyclone Dataset of the China Meteorological Administration (CMA, Ying et al. (2014); Lu et al. (2021)), from 1949 to 2022, Huizhou was impacted by tropical cyclones of typhoon intensity or greater, account for 58.9% of all events, with central pressures ranging from 900 hPa to 960 hPa. Among these cyclones, 57.7% moved a westerly or northwesterly path. In this study, the trajectory of Super Typhoon 'Mangkhut' (NO. 1822) was selected to construct a simulated wind field for storm surge inundation simulations for its northwestward movement and its status as the most strong cyclone in the Asia-Pacific region in recent years. Employing the trajectory of 'Mangkhut' is significant, as it potentially introduces the maximum storm surge inundation to the study area. As shown in Fig. 3, the trajectory of 'Mangkhut' spanned a vast geographical area, originating in the tropical Pacific region, traversing the Philippine Sea, and culminating in landfall along the southern coast of China. Based on Jelesnianski model, five typhoon scenarios were constructed as demonstrated in Table 1.

205 Using the Jelesnianski hurricane model (Jelesnianski, 1965) to provide wind forcing, and combining the coupled ADCIRC+SWAN model with topographic data, bathymetric data, and barrier data to simulate the storm surge inundation in the study area. The computational domain is discretized using an unstructured triangular mesh, with the mesh shown in as shown in Fig. 4. The

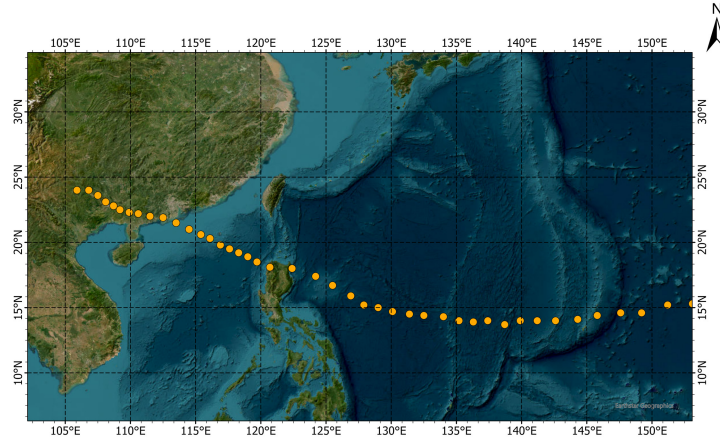


Figure 3. The trajectory of 'Mangkhut'.

Table 1. Typhoon scenarios parameter

<u>Typhoon scenarios</u>	<u>Minimum central pressure P_c (hPa)</u>	<u>Maximum wind velocity V_{max} (m/s)</u>	<u>Maximum wind radius R_{max} (km)</u>	<u>Trajectory</u>
<u>Scenario1</u>	<u>910</u>	<u>57</u>	<u>31</u>	<u>'Mangkhut'</u>
<u>Scenario2</u>	<u>920</u>	<u>53</u>	<u>33</u>	<u>'Mangkhut'</u>
<u>Scenario3</u>	<u>930</u>	<u>49</u>	<u>35</u>	<u>'Mangkhut'</u>
<u>Scenario4</u>	<u>940</u>	<u>45</u>	<u>38</u>	<u>'Mangkhut'</u>
<u>Scenario5</u>	<u>950</u>	<u>41</u>	<u>45</u>	<u>'Mangkhut'</u>

predominant tidal type in the northern China' South Sea is the 8 major tidal components of semidiurnal and diurnal frequencies (M2, S2, both the Digital Elevation Model (DEM) dataset and the Storm Surge Barriers (SSB) dataset were employed. The DEM dataset, available from Huizhou Land and Resource Bureau, is a raster dataset containing the land elevation in Huizhou. N2, K2, K1, O1, P1, Q1). The elevation data of the land grid was obtained from the digital elevation model (DEM) data of Huizhou acquired in 2015. The SSB dataset, available from Huizhou Oceanic Administration, contains information on storm surge barriers, such as location, height, and slope. Storm surge risk assessment was conducted by evaluating the difficulties of traversing areas with data in study area were obtained from the actual measured barrier engineering data and elevation measurement data. The ADCIRC+SWAN model was evaluated using real historical disaster events. (Detailed validation results between observed and simulated water levels is provided in Appendix A).

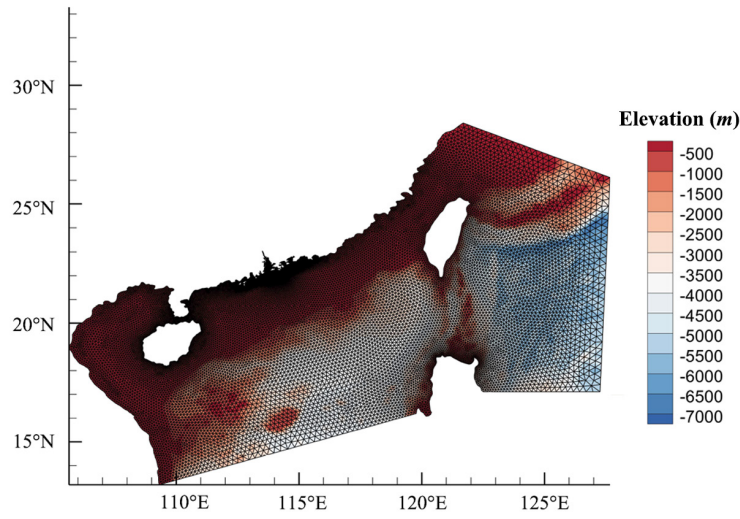


Figure 4. The computational domain. The model's computational domain covers the central and northern China' South Sea, as well as part of the eastern Philippine Sea and some areas of the northwest Pacific, extended from 106.0° E to 128.0° E and from 13.0° N to 28.0° N. The grid resolution on the open boundaries ranges from approximately 5 to 110 km, while the grid resolution along the coast of Huizhou is approximately 150 m. The mesh is optimized using local truncation error analysis, with a minimum grid resolution of approximately 100 m. The model contained grids consisting of 74328 units and 38407 nodes.

220 3.2 Road risk assessment

In this study, a fine-grained road risk assessment was conducted by comprehensively combining exposure, vulnerability, and hazard. The maximum possible extent of inundation under the storm surge scenario was determined to identify the roads that would be affected. To conduct hazard assessment, the roads were categorized into different vulnerability levels based on their traffic conditions. The evaluation of storm surge-related hazards is conducted by quantifying the extent of damage sustained by road infrastructures at varying inundation depths(Guidelines for Risk Assessment and Zoning of Marine Hazards Part 4: Storm Surge, 2019), and areas are classified into five risk levels, as displayed in Table 1,. The work of Huizinga et al. in 2017 provided a reference for the correlation between inundation depth and facility damage rate (Huizinga et al., 2017). Among all examined vulnerable elements, transportation facilities exhibited the highest damage rate due to their sensitivity to flooding and their importance in urban functions, as shown in Fig. 5. Based on the correlation, this study conducts a storm surge hazard assessment, dividing the inundation depth into five hazard levels, as illustrated in Table 2.

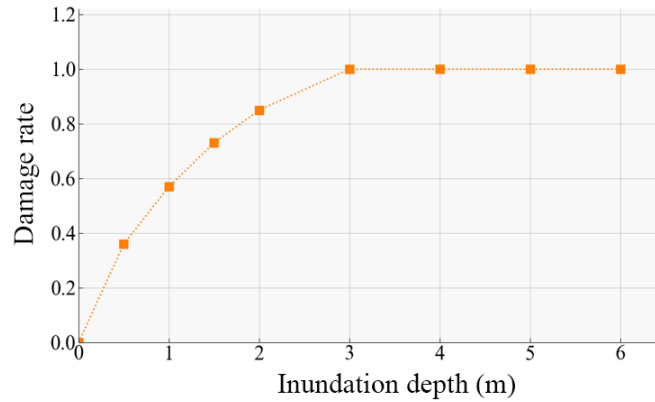


Figure 5. Relationship of inundation depth and transportation facilities damage rate. As the inundation depth increases from 0 m to 3 m, the damage rate of the transportation facilities also increases. However, this curve also delineates a saturation threshold in the damage rate when the inundation depth reaches 3 m. This suggests that once road transport facilities reach a state of severe damage, further increases in inundation depth cannot cause a higher damage.

In the construction of transportation facilities, roads are categorized based on criteria such as significance, capacity, and function, as follows: (1) Primary roads, also known as arterial roads, include highways, national and provincial roads. These roads typically have the highest design standards and firmer constructions. Consequently, their vulnerability is lowest, and they can maintain service even under extreme weather conditions. (2) Secondary roads typically serve to link pivotal urban areas, facilitating intra-city transportation. Although designed to lower standards than primary roads, they are still required to accommodate substantial traffic volumes and meet stringent safety demands. (3) Tertiary roads, encompassing rural pathways and minor unnamed thoroughfares, typically represent the extremities of the transportation network. These roads are designed and constructed to the lowest standards, primarily serving residents in rural areas. The vulnerability of the three categories of roads in ascending order is as: primary roads < secundar roads < tertiary roads. When considering the roads affected by the maximum inundation extent, combining the vulnerability of roads with the hazard level, a refined road risk assessment matrix was proposed, as shown in Table 3. where in Risk

3.2.1 Rasterization of road network

In contrast to the orderly thoroughfares of urban areas, suburban roads tend to exhibit meandering and disjointed configurations. To enhance the road network within the study area, we collected additional unnamed trails by using Geographic Information System (GIS) techniques. These trials are branches and extensions of existing roads and highways. The road network

3.3 Evacuation route planning

Table 2. Risk levels and inundation depths

Correlation between inundation depth and hazard level

	Inundation depth (m)	Hazard level
Risk level	0~15	I 0.0~0.15
	15~50	II 0.15~0.5
	50~120	III 0.5~1.0
	120~300	IV 1.0~2.0
	300~∞	V 2.0~∞

Table 3. Road risk assessment matrix

Risk level	Road categories	Primary roads
I	areas, passage is relatively unimpeded; Risk-	very low
II	areas present moderate obstacles to traversal; Risk-	low
III	areas pose a certain level of danger, making them impassable for vulnerable individuals such as children and the elderly; Risk-	moderate
IV	exhibit elevated risks, with traversal incurring considerable costs; finally, Risk-	moderate
V	areas present extreme danger, rendering passage virtually impossible.	high

Although the road network in the study area is depicted in Fig. 5, the road network in study area. The red line represents the country roads, the blue line represents the highways, and the green line represents the unnamed trails. Although the road network exhibits full connectivity, its complexity hampers the construction of a the topological structure. To facilitate spatial analysis and deep reinforcement learning on the road network, vector road data need to be rasterized. Raster data, akin to an image stored in a matrix, boasts a simplistic structure and readily engages in computational processes. In this study, an efficient method was proposed to convert road vector data into raster data. Road rasterization entails rasterizing each constituent vector segment. For a vector v that starts at (x_s, y_s) and ends at (x_e, y_e) , the rasterization algorithm is described in Algorithm ??, Vector Rasterization

where $\{c_{i,j}\} = \{c_{0,0}, c_{1,0}, c_{1,1}, \dots, c_{M,N}\}$ are the shorthand notations for the set of cells; (x_0, y_0) is the center coordinate of the starting cell; $\max(|x_p - x_c|, |y_p - y_c|)$ is the Chebyshev distance from (x_p, y_p) to (x_c, y_c) . In this study, the Chebyshev distance is used to measure the distance between two cells in a raster setting. The procedure delineated in Algorithm ?? permits parallelization, enabling simultaneous rasterization of multiple vectors and thereby significantly accelerating the rasterization process for the entire road network. To distinctly differentiate individual road branches, the cell dimensions of w apply DRL algorithm for evacuation route planning, this study utilized a raster environmental modeling; raster data can be perceived as images stored in matrices, with a concise structure that facilitates subsequent computational processing. The study area, a rectangular region of 9 km $\times$ h were established as 15 km, is divided into a cell of 16 m $\times$ 16 m. For a given area of size $H \times W$, the area can be rasterized into $M \times N$ cells and $M = \lceil W/w \rceil$, $N = \lceil H/h \rceil$, where $h \times w$ is the cell size. Each cell is independent, with the center coordinates of every cell representing the cell's location coordinates. A vector road is composed of

265 a series of interconnected vector segments, arrayed sequentially from beginning to end, as illustrated in Fig. 6. The rasterization of a road. A red straight line is a vector segment and it can be described using Δx and Δy ; The gray cells represent the result of the rasterization; The green lines are the boundaries of the adjacent cells that the vector traverse and the yellow points are the intersections.

3.3.1 Identification of disaster shelters

270 Identifying potential safe shelters is crucial to enable swift evacuation of affected individuals to structurally sound and watertight facilities. Contrary to urban areas, Identifying potential shelters for swift evacuation is essential, especially in suburban areas where dedicated disaster shelter facilities are scarce in suburbs, necessitating alternative options limited. This study employs three criteria—employed three criteria: structural stability, waterproofing, and capacity —to identify two primary facility types as to assess potential shelters. Hospitals are inherently designed to withstand various natural disasters. As per Chinese
275 codes GB50011-2021 Hospitals, inherently resilient to natural disasters, are prioritized due to their Class II waterproofing and GB50345-2021, hospital buildings are classified as Class II waterproof and Class I seismic-resistant structures, featuring multiple waterproof amenities and robust earthquake resistance. Consequently, hospitals are prioritized as disaster shelters during storm surges. In addition to hospitals, schools are designated as potential shelters. According to Chinese codes seismic
resistance as per GB50011-2021 and GB50223-2021, school buildings are classified as Class III waterproof, and their seismic-resistant
280 construction must surpass local residential buildings by one degree. Furthermore, hospitals and schools are typically situated in spacious, open areas, minimizing the likelihood of water accumulation. As public facilities, they can accommodate numerous evacuees. Fig. 7 depicts the spatial distribution of the chosen disaster avoidance facilities. The spatial distribution of potential shelters in study area.

In this study, path planning relies on the road network, necessitating both the starting point and destination to be situated
285 on roads. Given that shelter facilities (hospitals and schools) are generally not directly GB50345-2021. Schools, with Class
III waterproofing and enhanced seismic construction (GB50011-2021 and GB50223-2021), are also considered suitable. By
integrating local population density, traffic, and disaster risks, this research has identified several potential shelters. Their spatial
distribution is depicted in Fig. 6, which clearly marks the positions of each recommended shelter within the road network. For
route planning, since shelters are not typically located on roads, proximity to a shelter is defined by reaching road cells within a
290 specified range from a shelter is deemed equivalent to reaching the shelter itself. This range is referred to as the ‘shelter range’.
With the road network cell size of 16×16 , the ‘shelter range’ is defined as the Chebyshev distance of 8 cells (while the
actual distance is ‘shelter range’ of 128 m) from the shelter. All road network cells within the “shelter range” are considered as
destinations, equating to access to the shelter itself.

3.4 Deep Reinforcement Learning

295 Deep reinforcement learning is a new deep learning paradigm that focuses on formulating suitable policies and taking action to achieve a specific goal. A DRL agent learns autonomously through continuous interactions with a complex environment by performing actions and receiving rewards without explicit supervision. The interaction model between the agent and environ-

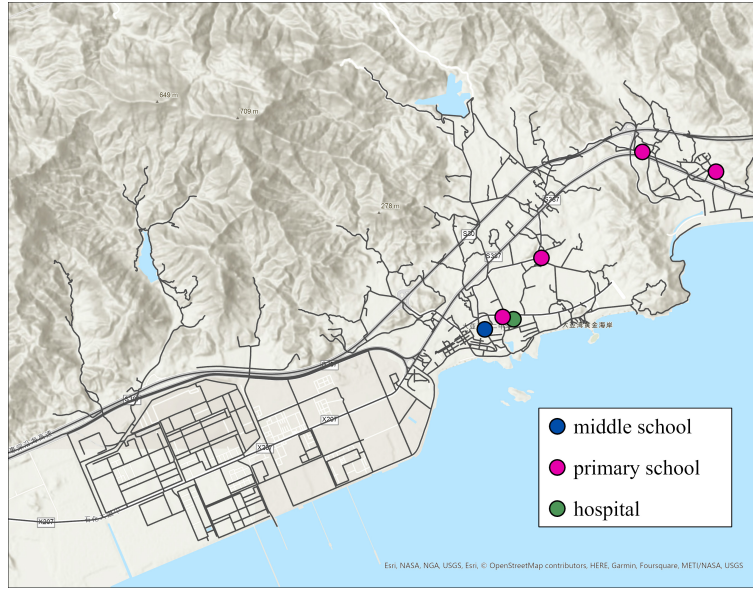


Figure 6. The spatial distribution of potential shelters in study area.

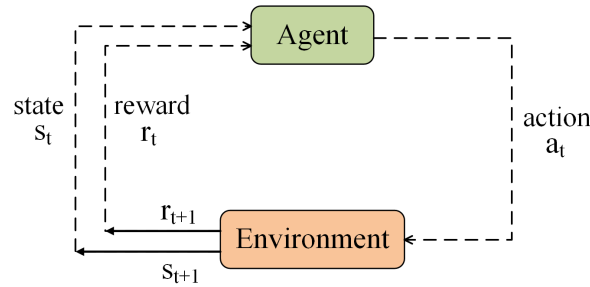


Figure 7. The DRL model. The environment is objective and refers to everything outside the agent that interacts with him, and the knowledge of the agent is reinforced through the interaction agent-environment. Consider State (s_t) represents the term-current state as-of the agent, which contains the current environmental feature, i.e.-all-and the information-that state space (S) represents the agent-perceives-from-set of all possible environmental states. Action (a_t) represents the environment-at-action taken by the current-moment. In-order-agent in state s_t according to assess-a-policy specifically π , and the action space (A) comprises all possible actions that an agent can take in a specific state. The feedback mechanism called reward r_{t+1} is introduced to define-quantify the value of a-the state s_t and to-quantify the effect-of-an-agent's action a_t .

ment is illustrated in Fig. 87. The primary objective of the agent is to maximize cumulative rewards, which can be advantageous for evacuation route planning. In this study, the Deep Q-Network was employed to maximize the cumulative rewards.

300 3.3.1 Markov Decision Process and Deep-Q Network

In the one-order Markov chain, the probability of a state s_t is only related to the preceding state s_{t-1} , and such a property is called the Markov property. Based on the one-order Markov chain, given a set of states The route planning can be reformulated as a continuous position state transition process based on Markov chains. This approach discretizes the route into states and actions, employing the Markov Decision Process (MDP) to address the uncertainties inherent in navigation. Given a state space

305 S , ~~a set of actions~~ an action space A , and a set of rewards R , the Markov decision process (MDP, Bellman) and the Markov reward process (MRP, Bertsekas) are defined as Eq. ~~7.6~~ and Eq. ~~7.7~~:

$$P_{s,s'}^a = P(s_{t+1} = s' | s_t = s, a_t = a) \quad (s, s' \in S, a \in A) \quad (6)$$

$$R_{s,s'}^a = R(r_{t+1} | s_t = s, a_t = a, s_{t+1} = s') \quad (s, s' \in S, a \in A, r \in R) \quad (7)$$

where $P_{s,s'}^a$ represents the transition probability from state s to state s' after performing action a , and $R_{s,s'}^a$ is the reward
310 obtained after transition (s, a, s') . MDP can be regarded as a continuous decision-making process, and the next action to be performed is only dependent on the current state.

~~Value-based DRL algorithms try to estimate the value of states and actions. State value is the expected reward that the agent can obtain from a state to give an estimate of how good a state is, and action value is the expected reward that the agent can obtain from a state after performing a specific action, providing an estimate of the action's utility. under a policy π , the value of a state s is denoted as $V_\pi(s)$ and the value of a action a in the state s is denoted as $Q_\pi(s, a)$:~~

315 ~~of a state s is denoted as $V_\pi(s)$ and the value of a action a in the state s is denoted as $Q_\pi(s, a)$:~~

$$V_\pi(s) = \mathbb{E}_\pi(G_t | s_t = s) = \sum_a P_\pi(a|s) Q_\pi(s, a)$$

$$Q_\pi(s, a) = \mathbb{E}_\pi(G_t | s_t = s, a_t = a) = \sum_{s'} P_{s,s'}^a [R_{s,s'}^a + \gamma V_\pi(s')]$$

~~where G_t is the total discounted reward from state s_t and $\gamma \in [0, 1]$ is the discount factor.~~

$$G_t = R_{t+1} + \gamma R_{t+2} + \dots = \sum_{k=0}^{\infty} \gamma^k R_{t+1+k}$$

320 ~~The goal of DRL is to find an optimal policy to maximize the state value and action value. Under the optimal policy π^* , the optimal state value function $V_*(s)$ and the optimal action value function $Q_*(s, a)$ can be obtained by:~~

$$V_*(s) = \max_a \sum_{s'} P_{s,s'}^a [R_{s,s'}^a + \gamma V_*(s')]$$

$$Q_*(s, a) = \sum_{s'} P_{s,s'}^a [R_{s,s'}^a + \gamma \max_{a'} Q_*(s', a')]$$

~~which are called the Bellman optimality equations (Bellman, 2010). Consequently, the optimal state value is the highest
325 attainable discounted reward from the state.~~

The path planning problem can be transformed into a finite and continuous decision-making process, wherein the In the setting of raster environment, the route planning can be described as: the agent chooses the subsequent action based on the current state until it reaches the destination. The location of the agent serves as the state, and the state transition is memoryless, satisfying the Markov property. Under the fixed-size raster environment and the same-size cells setting, the basic state space is $S: \{c_1, c_2, \dots, c_{M \times N} | c_i = (x_i, y_i), 0 \leq x_i \leq M, 0 \leq y_i \leq N\}$, $S: \{s_1, s_2, \dots, s_{M \times N} | s_i = (x_i, y_i), 0 \leq x_i \leq M, 0 \leq y_i \leq N\}$, and the basic action space is potential moves to the 8 adjacent cells $A: \{a_1, a_2, \dots, a_8 | a_i = (x, y), (x, y) \in \{-1, 0, 1\}^2 - (0, 0)\}$, while the following state transition equation is available:

$$\underline{cs}' = \underline{cs} + a \quad (\underline{cs}, \underline{cs}' \in S; a \in A) \quad (8)$$

The path-route planning problem in raster environment based on the MDP is defined as follows:

$$P_{\underline{c}, \underline{c}', \underline{s}, \underline{s}'}^a = P(\underline{cs}_{t+1} | \underline{cs}_t = \underline{cs}, a_t = a) = P(\underline{cs}_t = \underline{cs}, a_t = a) = 1 \quad (9)$$

$$\sum_{a \in A} P_{\underline{c}, \underline{c}', \underline{s}, \underline{s}'}^a = 8 \quad (10)$$

$$R_{\underline{c}, \underline{c}', \underline{s}, \underline{s}'}^a = R(r_{t+1} | \underline{cs}_t = \underline{cs}, a_t = a, \underline{cs}_{t+1} = \underline{cs}') = R(r_{t+1} | \underline{cs}_t = \underline{cs}, a_t = a) = f(d(\underline{cs}, \underline{cs}_e) - d(\underline{cs}', \underline{cs}_e)) \quad (11)$$

where $d(\underline{c}, \underline{c}_e) - d(\underline{s}, \underline{s}_e)$ is the Chebyshev distance between cell $\underline{c}$ and the destination cell $\underline{c}_e$. In the study area, the destination is represented by a cluster of shelter facilities, collectively constituting the destination set D . Let $d(\underline{c}) = \min_{\underline{c}_e \in D} d(\underline{c}, \underline{c}_e)$ signifies the distance to the nearest shelter facility. We defined the reward $R_{\underline{c}, \underline{c}'}^a - R_{\underline{s}, \underline{s}'}^a$ as a function of the difference $d(\underline{c}, \underline{c}_e) - d(\underline{c}', \underline{c}_e) - d(\underline{s}, \underline{s}_e) - d(\underline{s}', \underline{s}_e)$, implying that the reward is related to the agent's proximity to the destination. If $\underline{c}'$ is closer to the destination, the reward is positive, and vice versa. The corresponding state value function and action value function are:-

$$V_\pi(c) = \sum_{a \in A} P_\pi(a|c) Q_\pi(c, a)$$

$$Q_\pi(c, a) = f(d(c, c_e) - d(c + a, c_e)) + \gamma V_\pi(c + a) \quad (12)$$

A DQN is a multi-layered neural network, capable of approximating the optimal action value function $Q_*(c, a)$. Essentially, this function maps the n-dimensional state space to the action space. DQN is a value-based DRL algorithm where the output for a given state $\underline{c}_t - \underline{s}_t$ is a vector of action values denoted as $Q(\underline{c}_t, \cdot; \theta) Q(\underline{s}_t, \cdot; \theta)$, with θ representing the parameters of the online network. The agent's policy is to perform the action associated with the highest value. Moreover, DQN employs A DQN is a multi-layered neural network that estimates the value of states and actions. State value is the expected reward that the agent can obtain from a state to give an estimate of how good a state is, and action value is the expected reward that the agent can obtain from a state after performing a specific action. DQN incorporates an experience replay mechanism (Mnih et al., 2015), where past experiences are stored in a memory buffer and randomly sampled for training, to break the temporal correlation between samples and enable the which mitigates temporal correlations by randomly drawing from a stored memory

355 ~~of past experiences. This approach allows the~~ agent to learn ~~from infrequent events. Another important feature of DQN is the~~
~~use of~~ effectively from rare events. Additionally, DQN utilizes a separate target network (Mnih et al., 2015) ~~for estimating to~~
~~estimate~~ the $Q_*(c, a)$, thereby enhancing the stability of the learning process. ~~The DQN~~ Under a policy π , the value of a state
 ~~s is denoted as $V_\pi(s)$ and the value of an action a in the state s is denoted as $Q_\pi(s, a)$:~~

$$\underbrace{V_\pi(s) = \mathbb{E}_\pi(G_t | s_t = s)}_{\text{state value}} = \sum_a P_\pi(a|s) Q_\pi(s, a) \quad (12)$$

360 $\underbrace{Q_\pi(s, a) = \mathbb{E}_\pi(G_t | s_t = s, a_t = a)}_{\text{action value}} = f(d(s, s_e) - d(s + a, s_e)) + \gamma V_\pi(s + a) \quad (13)$

where G_t is the total discounted reward from state s_t and $\gamma \in [0, 1]$ is the discount factor.

$$\underbrace{G_t = R_{t+1} + \gamma R_{t+2} + \dots}_{\text{total discounted reward}} = \sum_{k=0}^{\infty} \gamma^k R_{t+1+k} \quad (14)$$

The goal of DQN is to find an optimal policy π^* to maximize the state value and action value. Under the optimal policy π^* , the optimal state value function $V_*(s)$ and the optimal action value function $Q_*(s, a)$ can be obtained by:

365 $\underbrace{V_*(s) = \max_a \sum_{s'} P_{s,s'}^a [R_{s,s'}^a + \gamma V_*(s')]}_{\text{optimal state value function}} \quad (15)$

$$\underbrace{Q_*(s, a) = \sum_{s'} P_{s,s'}^a [R_{s,s'}^a + \gamma \max_a Q_*(s', a')]}_{\text{optimal action value function}} \quad (16)$$

The DQN searches for the optimal policy to maximize the largest long-term cumulative reward that the target is:

$$Y_t^{\text{DQN}} \equiv r_{t+1} + \gamma \max_a Q(c_{t+1}, a_t; \theta_t^-) \quad (17)$$

~~and~~ And using the Mean squared error loss (RMSE) as loss function, the DQN can be trained by optimizing the following loss:

370 $\text{Loss}(\theta_t) = \mathbb{E}[(r_t + \gamma \max_{a_{t+1}} Q(c_{t+1}, a_{t+1}; \theta_t^-) - Q(c_t, a_t; \theta_t)]^2] \quad (18)$

The update process is based on the Monte Carlo method. By continuously interacting with the environment, the agent observes immediate rewards and accumulates them to count value information, which can then be transformed into a regression problem.

3.3.1 ~~Addressing Convergence Challenges in DQN~~

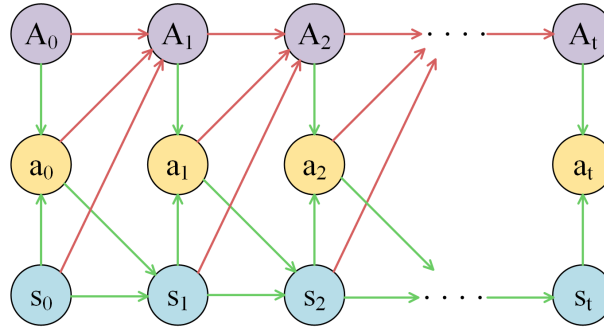
~~The Deep Q-Network (DQN)~~

375 **3.4 Optimization of the DQN**

The DQN model, when applied in a raster environment ~~with each cell possessing eight neighboring cells~~, faces significant convergence challenges from two aspects: 1. (1) the extensive search space, ~~and~~ 2. (2) the issue of sparse rewards. The

fundamental state space, search space is composed of the state space and the action space. In this study, the rasterized environment consisting of over 19,000 road cells, results cells with each cell possessing eight actions, resulting in a vast search space of around 8^{19000} . Coupled with the issue of sparse rewards, where the agent receives feedback only occasionally, the model's training process becomes even more complex and the convergence becomes notably difficult.

To counteract the vastness of the search space, two innovative methods were proposed: the masked action space and the masked state space. These methods effectively reduce the search space, thereby aiding in the model's convergence. In the basic action space, eight moves are available from the current cell in various directions. However, many of these actions might be meaningless. Typically, a cell in the action space allows movement in eight directions, but not all are relevant. By leveraging a mask, the compressed masked action space method efficiently narrows down the available actions from 8 to an approximate average of 3 per cell. This method focuses particularly on the transition of the action space, leading to dynamic action spaces where the action space transition is intricately linked with the state transition. For each cell (state) c , there is an associated action space A_c , where $|A_c| \leq 8$, and the action space emphasizing the interaction between action and state transitions. The action space transition is dependent on the state transition, as illustrated in Fig. 9, 8, where A_s denotes the action space associated with state s . According to Eq. 15, the action space of the current state relies on the preceding state and the last performed



The action space transition: where the red line represents the state transition and the green line represents the action space transition.

Figure 8. The action space transition. The red line represents the state transition and the green line represents the action space transition.

action, rendering the transition of the action space in compliance with the Markov property. The transition of action space action space is defined as:

$$P_{a,a'}^{c,s} = P(a_{t+1} = a' | c s_t = c s, a_t = a) = \pi(c s + a); \quad (19)$$

where $a \in A_c$, $a' \in A_{c+a}$, $a \in A_s$, $a' \in A_{s+a}$, $P_{a,a'}^c$ represents the probability of taking action a' after taking action a in cell c . There are two special cases in path planning in a raster environment. In raster-based route planning, two exceptional scenarios arise: (1) For a state transition $c' = c + a$, $s' = s + a$, an action a' exists such that $c = c' + a'$, in this case $s = s' + a'$, then the action a' is deemed meaningless-redundant for state s' . (2) Moving one cell in the diagonal direction is essentially equivalent to

moving one cell both horizontally and vertically. In this case Diagonal movement, which equates to a combination of horizontal and vertical movements, presents a unique case. Specifically, given the following transitions: $e + a = e'$, $e' + a' = e''$, and $e + a'' = e''$, the action $s + a = c'$, $s' + a' = s''$, and $s + a'' = s''$, then a' is meaningless after $c + a$ redundant following the initial transition $s + a$. Fig. 10-9 depicts two of eight compressed action spaces, namely action patterns. There are a total of 8 action patterns, which can be saved in 8 binary matrices. As path-

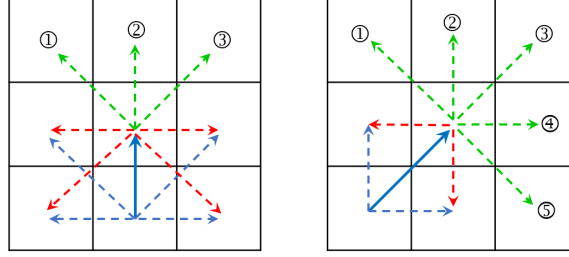


Figure 9. Action patterns for up and up-right. The blue solid line denotes the last action, while the red dashed lines represent meaningless redundant next actions. The green dashed lines signify the compressed-masked action space.

Additionally, as route planning operates on the extant-known-road network, exploring areas devoid of roads is considered futile. By focusing exclusively on state transitions within the road network, the action space can be further compressed. The compressed action space $A(e')A(s')$ for each state transition $e + a \rightarrow s + a$ can be calculated using the raster road network and the eight action patterns: action patterns:

$$A(\underline{cs'}) = I \circ A_P(a) \circ R_N(\underline{cs} + a) \quad (20)$$

where I is the basic action space of size 3×3 with all elements equal to 1, R_N is the matrix of the road network and $R_N(e)$ is the mask $R_N(s)$ is the road information of size 3×3 centered at cell es , the $A_P(a)$ is the action pattern for a , and the operator $\circ$ is the Hadamard product. For each pair $(e, a)(s, a)$, the $A_P(a) \circ R_N(c + a)$ is the mask of the action space. All the masks can be were computed in parallel and saved in a table prior to initiating training. Given a transition $e' = e + a, s' = s + a$, consulting the table using the tuple $(e, a)(s, a)$, and the action space of e' can be subsequently obtained. A single state corresponds to eight compressed action spaces of varying size, and the distribution of state-averaged action space size is shown in Fig. 11. The distribution of mean action space size per state. On average, the action space size for each state is less than 3. s' can be obtained.

Given start and end points, In large-scale path planning often necessitates information pertaining solely to a limited region, as the majority of roads remain untraversed. The challenge lies in discerning which roads are route planning, focusing on a 'premium region'—where the optimal route is most likely to be traversed and which are not. In this study, we introduce the concept of the 'premium region', which encompasses the area the optimal path may traverse, and we proposed a found—is

essential, rather than considering the entire network of roads. A masked state space ~~compression-method-method~~ was proposed to determine the ‘premium-region’.

The base path serves as an auxiliary path, reflecting the optimal path’s trend and providing ‘premium region’. We employed a base path, derived from a low-resolution raster image of the original map, to provide guidance in determining the ‘premium region’. We employed an additional low-resolution raster image, obtained through resampling the original raster image. The shortest path on the low-resolution image can be readily computed ‘premium region’. This base path is quickly identified using a breadth-first search algorithm, with this shortest path functioning as the base path. Expansion of the areas traversed by the base path yields the ‘premium region’. The shortest path from any start point on the road network to the nearest shelter can be precomputed, and during training, only the state value within the ‘premium’ region determined by the base path requires updates on the downsampled image. The low-resolution image represents an equally scaled-down projection of the original high-resolution image, as demonstrated in Fig. 12. The low-resolution image and the high-resolution image. For each cell $c(i, j)$ in the low-resolution image, uniquely corresponds to a rectangular area $rec(i, j): i \times M \leq x \leq i \times M + M, j \times N \leq y \leq j \times N + N$ in the high-resolution image. The cell size of the low-resolution image is and the cell size was set to $128 \text{ m} \times 128 \text{ m}$, which is 64 times larger than the cell size of the high-resolution image ($16 \text{ m} \times 16 \text{ m}$), and the ‘shelter-range’ is 1 cell (128). Given the base path $p_t = \{e_0, e_1, \dots, e_n\}$, the masked state space is described in Algorithm ??.

Computation of the masked state space

where m is a 0-1 matrix of size 937×546 serving as a mask, and $\delta b = \{s_1(x_1, y_1), s_2(x_2, y_2), \dots, s_n(x_n, y_n)\}$ in the low-resolution image, the ‘premium region’ $P = p_1 + p_2 + \dots + p_n$, where p_i is a rectangular region of $x = x_i \times 8 - \delta$ to $(x_i + 1) \times 8 + \delta$ and $y = y_i \times 8 - \delta$ to $(y_i + 1) \times 8 + \delta$. $\delta = 96 \text{ m}$ is the tolerance range. In this study, a suitable tolerance range is $\delta = 6$ (corresponding to an actual distance of 96 and an actual area of 320×320). In the example depicted in Fig. 13, based on a red base path derived from the low-resolution image, the blue region in the high-resolution image represents the ‘premium-region’ premium region. By considering only the road information situated in the ‘premium-region’ during the path ‘premium region’ during the route planning, the state space can be substantially compressed, with a compression ratio below 0.4.

Furthermore, to address Additionally, to tackle the issue of sparse rewards, we proposed the tri-aspect navigational reward mechanism. This mechanism offers a structured approach to reward-distribution, categorizing-categorize rewards into three distinct aspects: basic rewards, distance rewards, and risk rewards. Basic rewards encourage the agent to reach the goal (shelters) in the fewest steps as possible, with goal cells assigned a substantial positive reward (+2000), while other cells receive a negative reward of -1. Distance rewards guide the agent towards the goal, providing a +2 reward for moving closer and a -1 punishment otherwise. Risk rewards are negative incentives, designed to deter the agent from high-risk cells-roads whenever feasible. There are five risk levels (as depicted in Table 1), and corresponding rewards are The settings of navigational reward mechanism as illustrated in Table 2.

Risk level	Reward
0	-4
1	-8
2	-16
3	-32

4. Such a multi-layered reward structure provides the agent with more frequent and meaningful feedback, ensuring a consistent learning trajectory and fostering faster convergence.

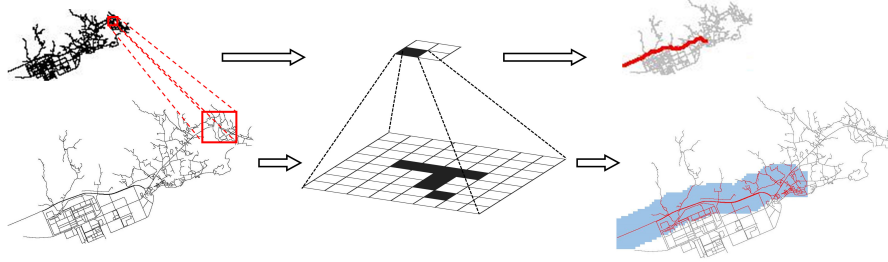


Figure 10. An example for masked state space compression. The low-resolution image and the high-resolution image. For each cell $c(i, j)$ in the low-resolution image, uniquely corresponds to a rectangular area $rec(i, j): i \times M \leq x \leq i \times M + M, j \times N \leq y \leq j \times N + N$ in the high-resolution image. The blue region in the high-resolution image represents the 'premium region'.

Table 4. Settings of navigational reward

Reward categories	Reward content	Value
basic rewards	destination reward	+100
	step reward	-1
distance rewards	distance reduction reward	+2
	distance increase reward	-1
risk rewards	very low risk reward	0
	low risk reward	-4
	moderate risk reward	-8
	high risk reward	-16
	very high risk reward	-32

¹Noted that, the rewards in the table are relative values, and need to be normalized for applications.

455 3.4.1 Training process

Prior to training, the state-action space for each cell within the high-resolution image was computed and stored in a table. Then, Each training episode begins with the agent at a randomly chosen road cell s_0 . The 'premium region' is defined by the base path from the breadth-first search algorithm was applied to the low-resolution image to ascertain the shortest path for each cell to the shelter. During each training episode, the agent commences from a randomly-selected road cell c_0 . The
 460 'premium region' is determined according to the corresponding shortest path in the low-resolution image, and an episode concludes upon the agent reaching the episode ends when the agent reaches a shelter. At each time-step, the input consists of the current cell-centered environmental observation. This observation agent receives an environmental observation centered on the current cell, with dimensions $(2r_{ob} + 1, 2r_{ob} + 1, 4)$, incorporates which includes information about roads, shelters, risk levels, and minimum Chebyshev distances to shelters within a $(2r_{ob} + 1) \times (2r_{ob} + 1)$ rectangular area, where r_{ob} represents
 465 the observation range. In this study the observation area. Here, the r_{ob} was set to 10 (equivalent to an actual distance of 160 m), reflecting the human field of vision in real-world scenarios. The output comprises is a sequence of length 8, corresponding to the values of 8 actions, and the subsequent 8-action values, with action a_i executed by the agent is selected based on the probability chosen probabilistically as follows:

$$P(a_i | \underline{cs}) = \epsilon \cdot \frac{1}{|A(\bar{a}, c)|} \frac{1}{|A(\bar{a}, s)|} + (1 - \epsilon) \cdot \frac{\mathcal{F}[Q_\pi(c, a_i), \max(Q_\pi(c, a_j))]}{\mathcal{F}[Q_\pi(s, a_i), \max(Q_\pi(s, a_j))]} \quad (a_i, a_j \in A(\bar{a}, \underline{cs})) \quad (21)$$

470 where $\mathcal{F}()$ is defined in Algorithm ??; $A(\bar{a}, c) \cdot \mathcal{F}(a, b) = 1$, if $a = b$; $\mathcal{F}(a, b) = 0$, otherwise. $A(\bar{a}, s)$ denotes the action space and it can be obtained by consulting the state-action space table using the current state c and last performed action $\bar{a}$. $0 \leq \epsilon \leq 1$ dictates the degree to which selection favors random exploration over the highest-value action. During the early training stages, a larger ϵ encourages agents to explore the unknown environment more extensively. As the model converges, ϵ should decrease to facilitate agent focus on high-value states and actions. The agent performs an action to the next cell, subsequently receiving
 475 a reward, which serves as an evaluation metric for the selected action.

Under the DQN framework, the training process is demonstrated in Fig. 14.11.

4 Simulation and Results and discussions

4.1 Results of road risk assessment

In this study, five distinct wind fields were employed, each characterized by a minimum central pressure of 910, 920, 930, 940, and 950 hPa were constructed using Jelesnianski model. These wind fields were subsequently integrated into the coupled
 480 ADCIRC-SWAN model to simulate the storm surge for risk assessments, and DQN was utilized to search for the optimal evacuation path under these scenarios. Three of these scenarios were used as training data, while the remaining two were used as test data. The study area is a part of coastal suburb of Huizhou with a size of 15×9 . Historical storm data was utilized in conjunction with the Jelesnianski empirical model to generate a wind field, which was then provided to the coupled

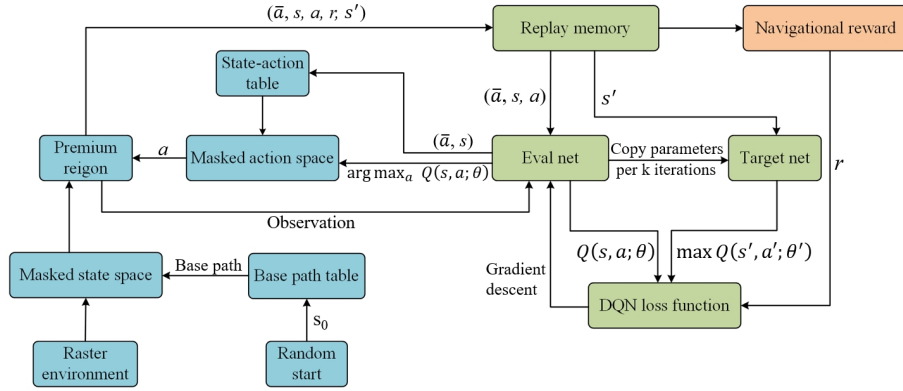


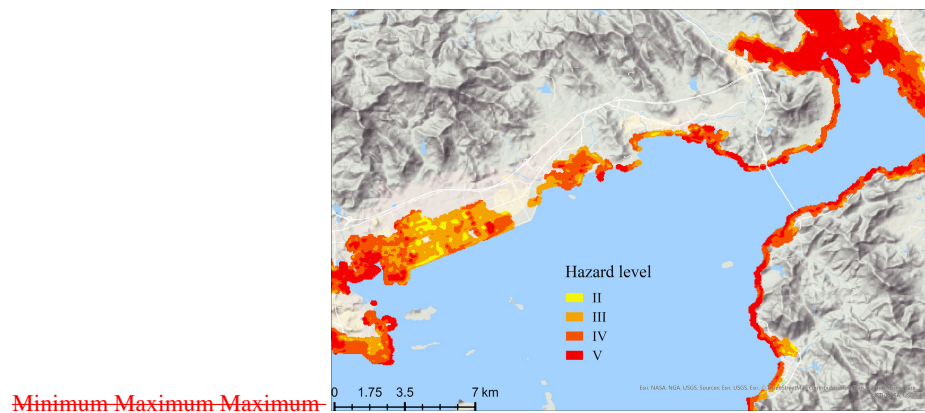
Figure 11. The DQN training process with the masked action and state spaces compression, features three components: green sections represent the classical DQN, blue sections correspond to masked state compressed search space compression method, and orange sections indicate masked action space compression navigational reward.

ADCIRC-SWAN model for calculating water levels and simulating inundation extents and depths within the study area. The DQN was employed to develop evacuation plans, focusing on generating real-time evacuation routes to predetermined disaster shelters from any given starting position. Extensive experiments were performed, demonstrating the effectiveness of the proposed method.

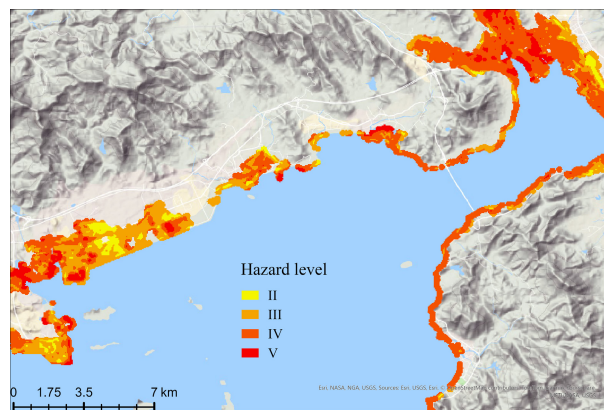
4.2 Storm surge simulation and risk assessments

To make the hypothetic wind field reasonable, the historical tropical cyclone dataset released by extent and depth of storm surge inundation within the study area. Considering the maximum inundation extent to conduct exposure assessment, and according to the China Meteorological Administration (CMA) was used. The dataset contains the location and intensity of tropical cyclones in the northwest Pacific Ocean, recorded every six hours from 1949 to 2018. According to the CMA dataset, the minimum central pressure of tropical cyclones affecting China ranges from 880 to 1000. Mild storm surges barely penetrate inland, while severe storm surges can inundate the entire area. In this study, five representative typhoon scenarios were defined, with their parameters displayed in Table 3. Typhoon scenarios Table 2, the storm surge hazard assessment across the five wind scenarios is graphically represented in Fig. 12. The results indicate that the intensification of typhoons correlates with more extensive and severe inundation, elevating the associated hazard levels. For instance, Fig. 12(a) illustrates a scenario where, at a central pressure of 910 31-61-2-920-33-57-3 hPa, nearly half of the study area is submerged, with a significant portion experiencing high-hazard conditions. Fig. 12(c) delineates a pivotal moment: at 930 35-53-4-940-38-49-5-950-41-45-

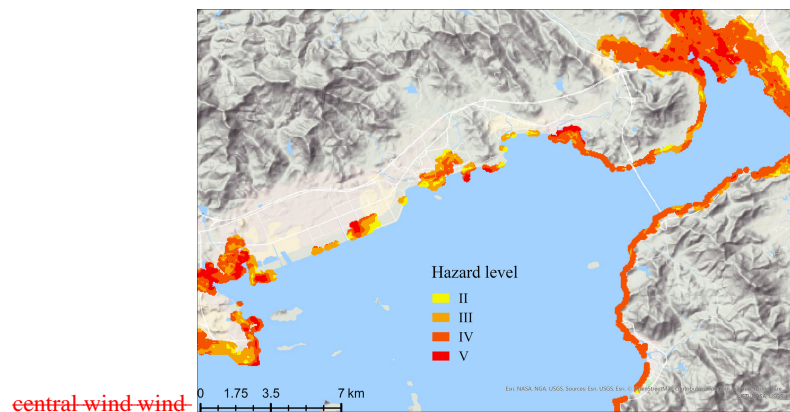
Additionally, Super Typhoon Mangkhut, which caused the most significant storm surge disaster to hPa, the study area on record, had its track chosen as the simulated cyclone's track. The track of Mangkhut is shown in Fig. 15. The track of Mangkhut.



(a) Scenario 1.



(b) Scenario 2.



(c) Scenario 3.

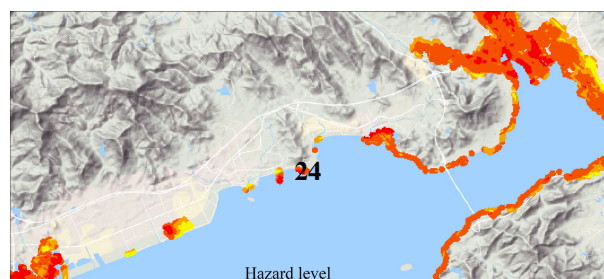


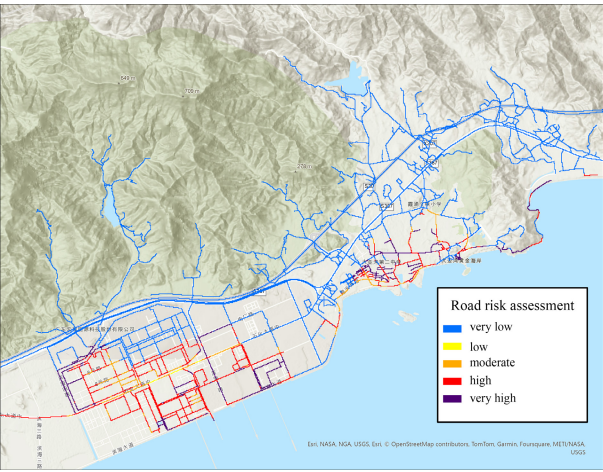


Figure 13. The road network and road categories in the study area.

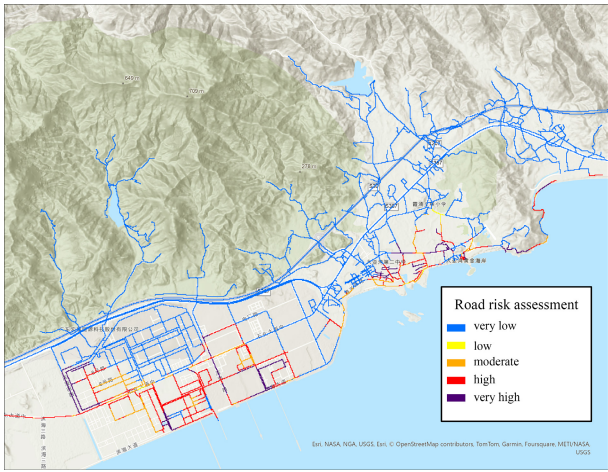
Under the wind field generated by the Jelesnianski model, typhoon's central pressure marks a significant decrease in both the inundation's extent and severity. In the last scenario, as depicted in Fig. 12(e), the inundation is confined to the peripheral parts of the study area, exhibiting limited capacity to affect the inland road network. Consequently, the coupled ADCIRC-SWAN model was run on the datasets to simulate surge water. The inundation depth was calculated by subtracting the DEM from the height of simulated surge water. According to Table 1, the risk assessments were conducted based on the inundation depths. When the inland flooding attains its maximum extent, the risk assessments for five storm surge scenarios are depicted in last scenario was excluded from consideration in the subsequent road risk assessment and evacuation route planning processes.

This study focuses on the road risk assessment of storm surge in the study area, which is located in a coastal suburb characterized by winding and discontinuous roads. To enhance the road network within the study area, additional road data was collected using geographic information system (GIS) technology, including minor and unnamed roads that often represent fine branches and extensions of existing roads and highways, which became an important basis for classifying roads and evaluating their vulnerability. The road network and road categories in the study area is shown in Fig. 13.

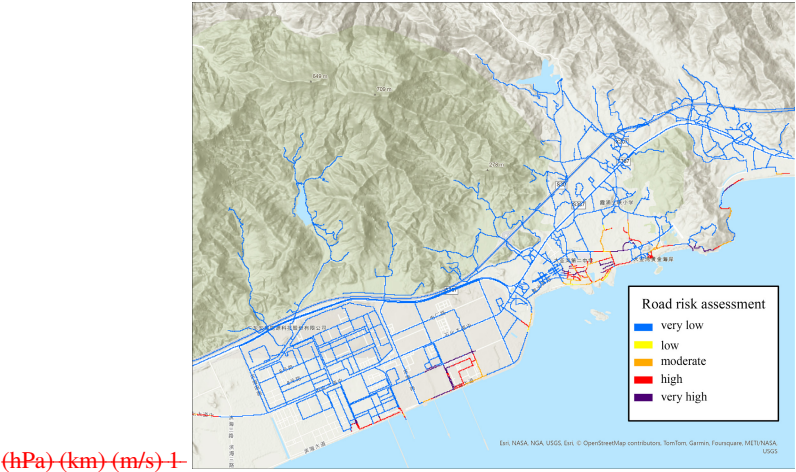
As presented the Table 3, the road risk assessment for the study area under the four scenarios is shown in Fig. 14, where storm surge mainly threatened the secondary and tertiary roads in the coastal road network, mainly concentrated in the southern part of the Daya Bay petrochemical industrial zone and the Daya Bay Golden Coast vicinity. In the scenarios depicted in Fig. 16. Risk assessments for five typhoon scenarios. The results were organized and displayed in ArcGis Pro 3.0 software. Risk areas are considered un-flooded safe areas and are therefore not depicted on the figure. Scenario 1 and 5 were used as test data, while Scenario 2-4 were employed as training data. To rasterize the inundation area, take the average inundation depth in a cell as the inundation depth of the cell. 14(a) and 14(b), a storm surge causes widespread flooding in the petrochemical industrial



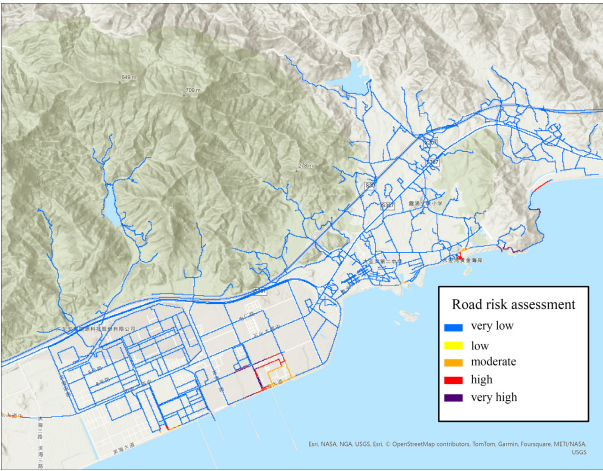
(a) Scenario 1.



(b) Scenario 2.



(c) Scenario 3.



(d) Scenario 4.

Figure 14. The road risk assessment.

zone, rendering most roads at high risk and impassable. In such situations, evacuation routes should guide individuals north into primary roads and then eastward to shelters. Samely, the area surrounding the Daya Bay Golden Coast faces a critical situation, necessitating swift evacuation along secondary and tertiary roads to the northwest or northeast towards shelters. Fig. 14(c) presents a pivotal scenario, demonstrating a significant reduction in the storm surge's impact on the roads of petrochemical industrial zone. However, the Golden Coast vicinity remains significantly affected, primarily due to the prevalence of vulnerable tertiary roads. As depicted in Fig. 14(d), when the typhoon's central pressure drops to 940 hPa, the storm surge's impact on the Golden Coast's roads diminishes substantially. Consequently, the typhoon's minimum central pressure at 930 hPa and 940 hPa are critical thresholds for the storm surge's impact. Below 940 hPa, the roads in the Daya Bay Golden Coast area are at high risk, and when the pressure is below 930 hPa, the Daya Bay Petrochemical Industrial Zone's roads experience severe disruption.

4.2 Model Performance of DQN model for evacuation route planning

The goal of this work is to enable real-time ~~planning-of-the-optimal-path~~ route planning to the shelter from any given cell-start based on the surrounding environment. The ~~proposed-model~~ achieves high-accuracy evacuation path-planning according to the surrounding environment. To evaluate the overall performance of the model, experiment utilized three simulated scenarios (Scenario 2, 3, and 4) designated for training DQN model, while Scenario 1 and 5 were utilized, and served as the test case. 1000 starting cells were randomly ~~chosen-for-each-test-scenario-to-conduct-path-planning~~, respectively. The selected in Scenario 1 to conduct route planning and the enumeration method is used to find the true optimal paths under the reward-setting for these 2000 locations, which are called the target paths, and the 'optimal' paths output by the model are called the eval paths routes. In Scenario 1, the ~~target-paths-cover-about-64-generated-routes~~ covered about 71 % of the road network, with an average length of 4286-4776 m. In Scenario 5, the target paths cover about 61 % of the road network, with an average length of 3758. Refer to Appendix B for exemplifications of the route planning.

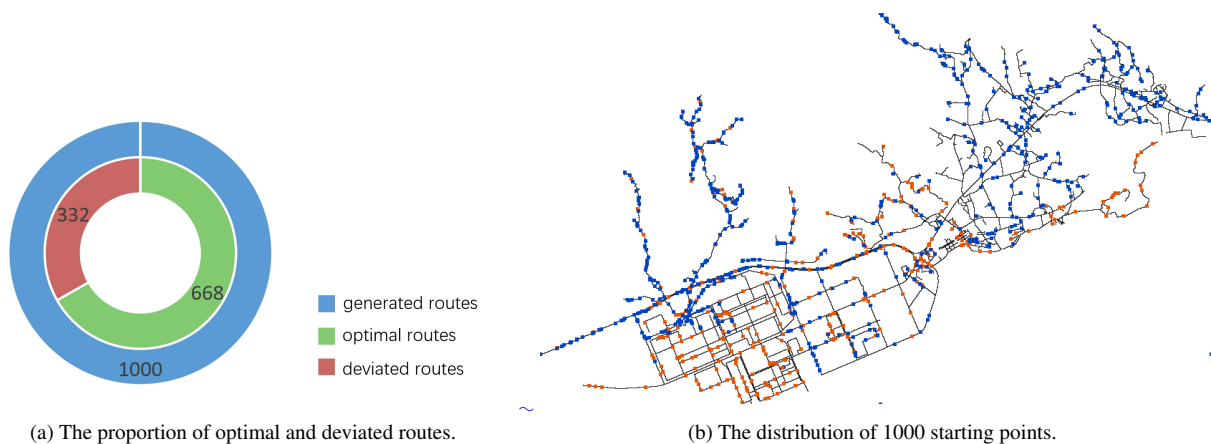


Figure 15. The proportion of optimal and deviated routes and the distribution of 1000 starting points.

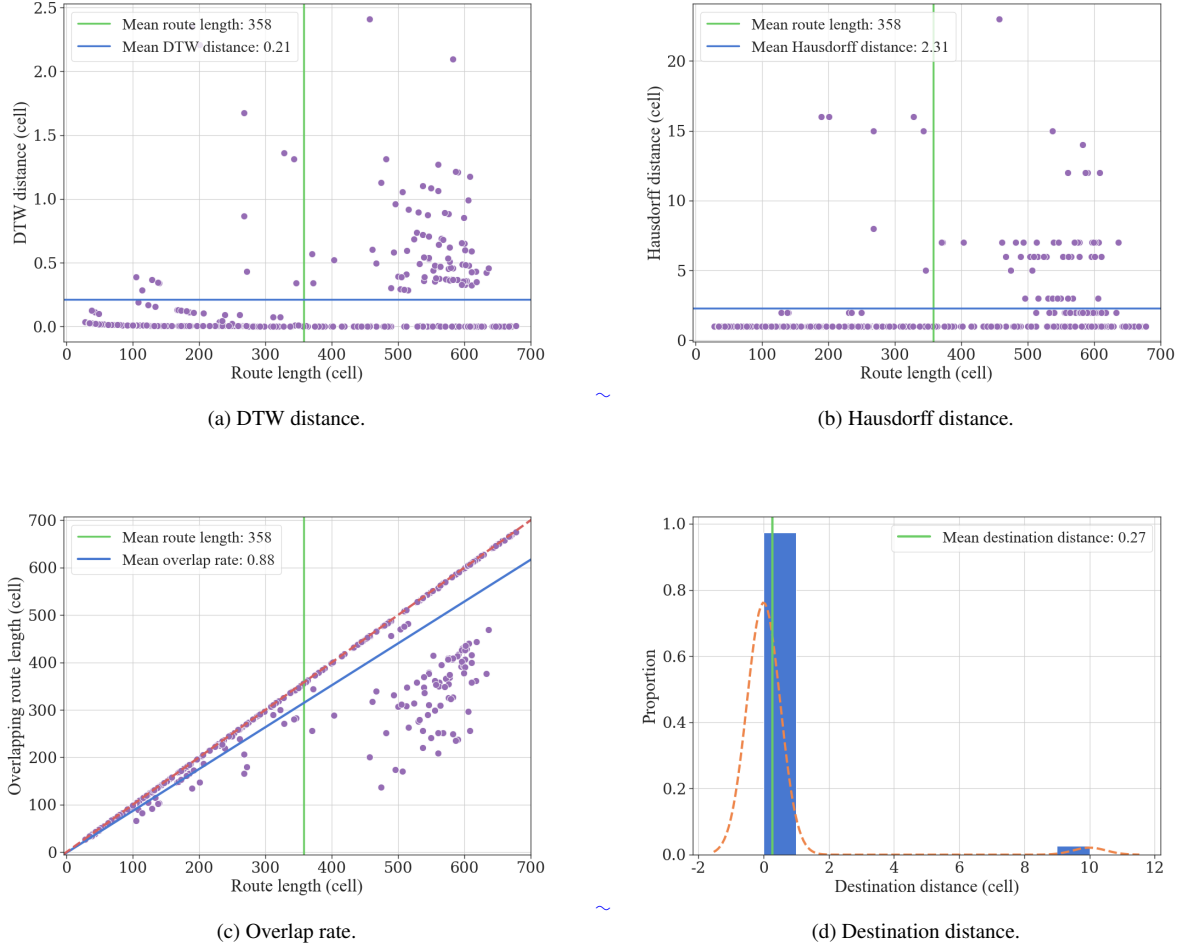


Figure 16. Model evaluation metrics for DTW, Hausdorff, destination distance and overlap rate with one cell representing 16 m.

To evaluate the overall model performance, ~~the differences between the target paths and the eval-paths were measured from two aspects: path similarity and distance to the true destination. Three evaluation~~ five metrics were introduced ~~to measure the path similarity including the dynamic~~, including: (1) proportion of optimal routes, (2) Dynamic time warping (DTW, Müller) rate, the Hausdorff (Huttenlocher et al., 1993) distance, (3) Hausdorff distance (Huttenlocher et al., 1993), (4) overlap rate, and the overlap rate. Given a target path p_{tar} of length l_{tar} and an eval-path p_{eval} of length l_{eval} (5) destination distance. The proportion of optimal routes measures the degree of correspondence between the generated routes and the optimal routes by calculating the percentage of generated routes that completely coincide with the optimal routes. The DTW evaluates the similarity between two time series, and the DTW distance is used to measure the average deviated distance of the generated route from the optimal route. Given a optimal route p_{opt} of length l_{opt} and a generated path p_{gen} of length l_{gen} , the DTW can

be discribed as a dynamic programming (DP):

$$\text{Minimize}(D \cdot W) \quad (22)$$

where $D = \{d(p, q)\}$ and $d(p, q)$ is the Chebyshev distance between the cell $p \in p_{tar}$ and cell $q \in p_{eval}$ and cell $q \in p_{gen}$. $W = \{w(i, j)\}$ and $w(i, j)$ is the binary DP variable $i, j \in \mathbb{Z}^+, i \leq l_{tar}, j \leq l_{eval}$. The above DP is subject to:-

$$w(i, j) = \{0, 1\}$$

$$w(1, 1) = 0$$

$$w(l_{tar}, l_{eval}) = 1$$

$$w(i, j) = [w(i-1, j) + w(i-1, j-1) + w(i, j-1)] \times w(i, j)$$

The DTW distance is used to measure the average deviated distance of the eval path from the target path that can be obtained by:-

$$d_{DTW}(p_{tar}, p_{eval}) = \frac{D \cdot W}{n}, \quad n = \sum w(i, j) \quad (w(i, j) \in W)$$

while DTW rate, defined as: d_{DTW}/l_{tar} , quantifies the DTW distance per unit length. Consider a path-

$$d_{DTW}(p_{opt}, p_{gen}) = \frac{D \cdot W}{n}, \quad n = \sum w(i, j) \quad (w(i, j) \in W) \quad (23)$$

Consider a route as a set of locations, and utilize the Hausdorff distance $d_{Haus}(p_{tar}, p_{eval}), d_{Haus}(p_{opt}, p_{gen})$ to measure the distance between the two set of paths, can be computed by: two sets, which can be calculated by:

$$d_{Haus}(p_{tar}, p_{eval}) = \max\{\sup \inf d(p, q), \sup \inf d(q, p)\} \quad (p \in p_{tar}, q \in p_{eval}) \quad (24)$$

where sup is the supremum and inf is the infimum. The Hausdorff distance measures the maximum deviated distance of the eval path, and the Hausdorff rate is: d_{Haus}/l_{tar} , which quantifies the Hausdorff distance per unit length. Let l_{ot} generated route.

Let l_{ov} denote the length of the longest overlapping sub-path between the target path and the eval path common sub-route between the generated and optimal routes, the overlap rate is defined as the ratio $r_{ov} = l_{ov}/l_{opt}$. The destination distance is a metric for assessing the efficacy of evacuation route planning, reflecting the ratio $r_{ot} = l_{ot}/l_{tar}$. In addition to the similarity, the distance between the destinations of the target path and the eval path can be measured by Chebyshev distance endpoint of the generated route and the designated shelter.

By comparing with the target paths, the eval paths can be categorized into optimal paths and deviated paths, where the deviated path differs from the corresponding target path by one or more cells. In Scenario 1, there are 668 Among generated routes, a route that deviates from the optimal paths and 332 deviated paths, while in Scenario 5, there are 620 optimal paths and 380 deviated paths, as shown in Fig. ?? The number of optimal paths and deviated paths for Scenario 1 and 5. route is termed a

deviated route. Fig. 15 depicts the proportion of optimal and deviated routes, as well as the distribution of their starting points, where blue and orange markers represent the optimal and deviated routes, respectively.

The 1000 test cases evaluated the model's performance amidst severe storm surges and complex inundation environments. Among the 1000 generated routes, 668 routes are optimal with 332 routes deviated. Focusing on the deviated paths, the DTW distance, the Hausdorff distance, the overlap rate, and the destination distance under Scenario 1 and 5 are illustrated as Fig. ?? and Fig. ??, where one cell is equal to an actual distance of 16 meters. Scenario 1 evaluates the model's performance amidst severe storm surges and complex inundation environments, while Scenario 5 assesses the model's performance in environments with minimal flooding. Among the 2000 eval paths tested in Scenario 1 and 5, 668 and 620 paths are optimal, respectively. For the remaining deviated paths, on average, the DTW distance is 76.8 16. The mean length of 332 deviated routes is 5728 mper path and 0.04, indicating that longer routes tend to exhibit greater deviations. In Fig. 16(a), the average DTW distance for deviated routes is a mere 3.36 mper meter in Scenario 1, and 65.6, and the majority of these routes exhibit DTW distances below this average, although a minority exceed 16 mper path and 0.03 per meter in Scenario 5. The average Hausdorff distance is 168 36.96 mper path and 0.07 per meter in Scenario 1, and 252.8 per path and 0.08 per meter in Scenario 5, signifying that the eval paths exhibit a remarkable morphological similarity to the target paths. The mean. An analysis combining Fig. 16(a) and Fig. 16(b) indicates that the routes exhibit minor deviations from the optimal routes in the majority of cases, with significant deviations occurring infrequently. In Fig. 16(c), the average overlap rate is 0.74 and 0.71, respectively, signifying that a substantial portion of the eval paths coincides with the target paths. Furthermore, the mean destination distance is 78.4 0.88, with most cases approaching a perfect overlap of 1, demonstrating that the deviated routes mostly remain consistent with the optimal routes. Finally, in Fig. 16(d), the destination distances were calculated with an average value of 4.32 m and 107.2. The majority of destination distances fall within the 0~16 m, ensuring that evacuees reach a location very close to the most suitable shelter. Notably, the mean lengths of the deviated paths (5584 range, although a few outlier routes exceed 144 m and 4832), are longer than the mean lengths of eval paths (4826 and 3758), revealing the model's inadequacies in effectively planning long path. Overall, these deviations are negligible for the study area of 135, suggesting that deviated paths, despite process discrepancies from the optimal routes, can still reach the optimal shelter with high accuracy. The experimental results demonstrate that the proposed method exhibits strong performance in generality, providing emergency evacuation path route planning for the entire study area. Model performance in Scenario 1.

Model performance in Scenario 5.

5 Conclusions

This study presents a comprehensive approach to emergency evacuation conduct the storm risk assessment and evacuation route planning in the Daya Bay Petrochemical Industrial Zone. By coupling a risk assessment of storm surges with a road network, a raster environment that reflects real-world scenarios has been constructed. The DQN model was employed to develop a real-time evacuation plan, providing efficient and effective guidance for individuals during storm surge events. To enhance the adaptability of the DQN model for rasterized road network, masked state space, masked action space, and tri-aspect

reward mechanism were proposed, significantly enhancing the model's convergence. The coupled ADCIRC-SWAN model and Jelesnianski method were used to create the simulation environment of storm surges under different typhoon scenarios. Additionally, potential safe shelters were identified for It facilitates a thorough understanding for local government regarding the spatial distribution of road risks and aids residents in swiftly devising optimal evacuation routes to shelters, contingent upon their immediate surroundings. This approach significantly bolsters efforts in storm surge disaster prevention, mitigation, and contributes to the sustainable development of the region.

This study utilized the ADCIRC+SWAN model to simulate five storm surge scenarios, identifying maximum inundation levels within the study area to provide more evacuation options.

Two distinct storm surge scenarios with Integrating inundation data with road network information facilitated a fine-grained risk assessment, revealing high vulnerability in the petrochemical industrial zone and Golden Coast region's road networks to storm surges. The minimum central pressure of typhoons, with key thresholds at 930 and 940 hPa, serves as a critical factor in determining the impact severity: roads in the minimum central pressures of 910 and 950 Golden Coast are at risk when the central pressure is below 940 hPa, while central pressure below 930 hPa were used as the test environments, and path plan for significantly disrupt the petrochemical zone's roads. Focusing on evacuation route planning, the study developed a high-resolution raster environmental model to explore deep reinforcement learning methods for large-scale raster environments. To address DQN model's convergence challenges, a compressed search space and a navigational reward mechanism were introduced, enhancing the DQN model's capacity in route planning. In Scenario 1, 1000 randomly selected starting cells were conducted in each scenario. By comparing the eval paths with the target paths, for both scenarios, over 60 % of the eval paths were optimal, while the remaining 40 % exhibited only minor deviations from the optimal paths, with an average difference of merely 4 centimeters per meter and an average overlap rate of exceeding starting points were randomly sampled to generate evacuation routes, with nearly 70 %. Moreover, the average destination distance between the deviated paths and the optimal paths was approximately 100 . Overall, the proposed method proved highly effective in planning optimal evacuation routes with a minimal deviation that can be of great assistance for the evacuee during a real-world storm surge, proving optimal and the rest showing minimal deviation, averaging a DTW distance of 3.36 m and an overlap rate of approximately 0.9.

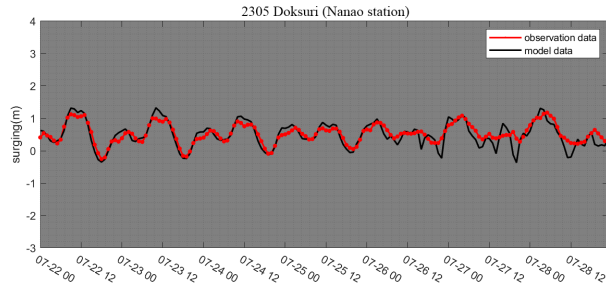
Based on the results and findings presented in this study, the proposed method showed effectiveness in enhancing real-time This study demonstrates the efficacy of the proposed method in assessing road risks and enhancing emergency evacuation plans and demonstrated the potential of employing . It underscores the worth of leveraging advanced modeling techniques to improve emergency response and preparedness in vulnerable areas. However, there is still room for improvement and future work can be done to further optimize the evacuation plan. One potential direction involves replacing the raster environment with a topological environment and utilizing graph neural network techniques. Due to the limitations of the raster environment, the proposed method trains slowly and is difficult to apply on a larger scale area (such as a city area). The reason for adopting the raster environment in this paper is that it is simple to construct and can directly correspond with the flooding extents. Another possible avenue for future research is to incorporate more advanced machine learning algorithms or data-driven models, as there is potential for further improvement with currently 60 % of eval paths being optimal. Additionally, more environmental information could be utilized to meet more complex demands, such as incorporating population restrictions

for shelters, identifying different types of roads' passing conditions and costs, and so on. This study is a promising start in developing. Yet, there are opportunities for refinement. Current road risk assessments consider flooding and road types but could be improved to include road width, population density and other factors to enhance the effectiveness of road risk assessment.

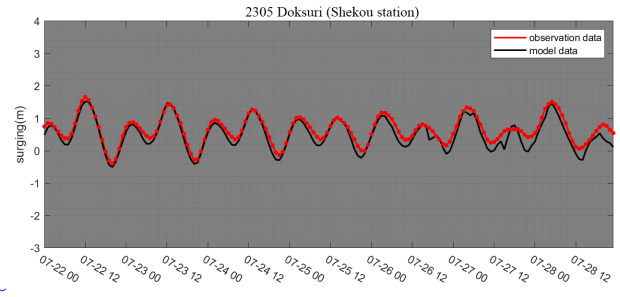
650 Looking ahead, transitioning from raster to topological environments, potentially through graph neural networks, presents a promising path to overcome the limitations in training speed and large data handling. Additionally, the current model's exclusion of infiltration and drainage dynamics may affect inundation depth accuracy. Future research should integrate these factors for more precise storm surge simulations. In summary, this study represents a pivotal step in developing storm surge risk assessment and real-time emergency evacuation plans, and we eagerly anticipate future advancements in this field. evacuation

655 planning. The scope for advancement is considerable, inviting ongoing research and innovation in this critical domain.

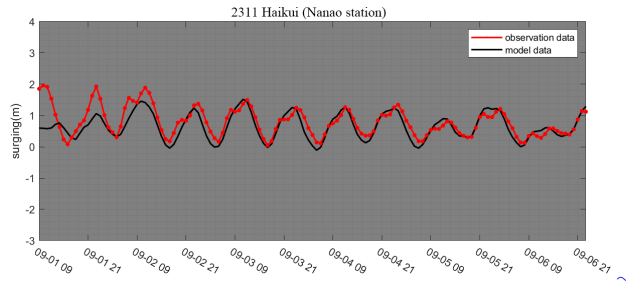
Appendix A: The validation of ADCIRC+SWAN model



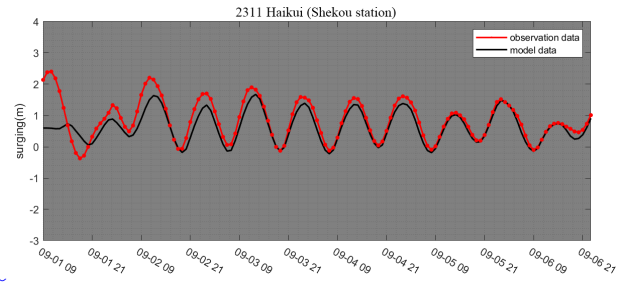
(a) 2305 Doksuri (Nanao station).



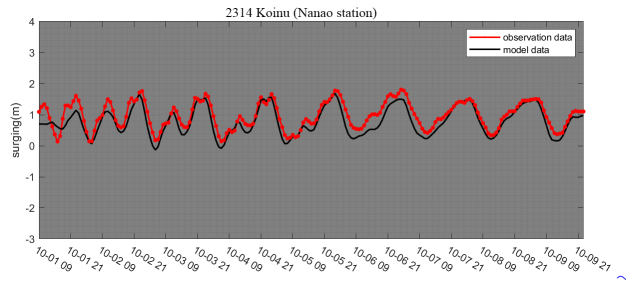
(b) 2305 Doksuri (Shekou station).



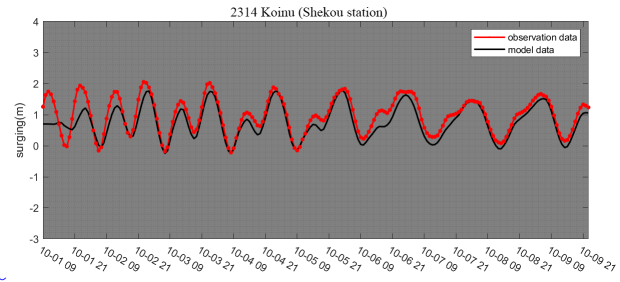
(c) 2311 Haikui (Nanao station).



(d) 2311 Haikui (Shekou station).



(e) 2314 Koinu (Nanao station).



(f) 2314 Koinu (Shekou station).

Figure A1. The validation of ADCIRC+SWAN model. The real historical typhoon events (2305, 2311, 2314) were used to validate water levels. The black line represents the simulated water levels and the redline represents the observed water levels record by Nanao station and Shekou station.

Appendix B: Examples of evacuation route planning

Author contributions. Yan Li and Wenjuan Li designed the research and optimized the overall structure of this paper. Chuanfeng Liu and Si Wang completed most of the main work, including the programming, debugging of parameters, and final drafting of the article. Lin Mu and
660 Darong Liu contributed some important algorithm ideas and completed the work of the comparison part. Hao Qin and Kai Zhou provided the original algorithm ideas and framework for this study and provides valuable suggestions for program optimization and parameter adjustment.

Competing interests. The authors declare that they have no conflict of interest.

Financial support. This work was supported by Shenzhen Science and Technology Program (Grant No.KCXFZ20211020164015024).

References

- China Marine disaster bulletin, <https://www.mnr.gov.cn/sj/sjfw/hy/gbagg/zghyzhgb/>, 2021.
- Alsnihi, R. and Stopher, P. R.: Review of procedures associated with devising emergency evacuation plans, *Transportation Research Record*, 1865, 89–97, 2004.
- Astri, R. et al.: Implementation of A-Star Algorithm for Searching Routes Near the Tsunami Evacuation Shelter Point, *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 4, 254–259, 2020.
- Atkinson, G. D. and Holliday, C. R.: Tropical cyclone minimum sea level pressure/maximum sustained wind relationship for the western North Pacific, *Monthly Weather Review*, 105, 421–427, 1977.
- Bellman, R.: A Markovian decision process, *Journal of mathematics and mechanics*, pp. 679–684, 1957.
- Bellman, R. E.: *Dynamic programming*, Princeton university press, 2010.
- Benavente, J., Del Río, L., Gracia, F. J., and Martínez-del Pozo, J. A.: Coastal flooding hazard related to storms and coastal evolution in Valdelagrana spit (Cadiz Bay Natural Park, SW Spain), *Continental Shelf Research*, 26, 1061–1076, 2006.
- Bertsekas, D.: *Dynamic programming and optimal control: Volume I*, vol. 1, Athena scientific, 2012.
- Booij, N., Ris, R. C., and Holthuijsen, L. H.: A third-generation wave model for coastal regions: 1. Model description and validation, *Journal of geophysical research: Oceans*, 104, 7649–7666, 1999.
- Chen, C., Liu, H., and Beardsley, R. C.: An unstructured grid, finite-volume, three-dimensional, primitive equations ocean model: application to coastal ocean and estuaries, *Journal of atmospheric and oceanic technology*, 20, 159–186, 2003.
- Cheung, K. F., Tang, L., Donnelly, J. P., Scileppi, E. M., Liu, K.-B., Mao, X.-Z., Houston, S. H., and Murnane, R. J.: Numerical modeling and field evidence of coastal overwash in southern New England from Hurricane Bob and implications for paleotempestology, *Journal of Geophysical Research: Earth Surface*, 112, 2007.
- Dietrich, J., Zijlema, M., Westerink, J., Holthuijsen, L., Dawson, C., Luettich Jr, R., Jensen, R., Smith, J., Stelling, G., and Stone, G.: Modeling hurricane waves and storm surge using integrally-coupled, scalable computations, *Coastal Engineering*, 58, 45–65, 2011.
- Forcael, E., González, V., Orozco, F., Vargas, S., Pantoja, A., and Moscoso, P.: Ant colony optimization model for tsunamis evacuation routes, *Computer-Aided Civil and Infrastructure Engineering*, 29, 723–737, 2014.
- Goerigk, M., Deghdak, K., and Heßler, P.: A comprehensive evacuation planning model and genetic solution algorithm, *Transportation research part E: logistics and transportation review*, 71, 82–97, 2014.
- Granger, K.: Quantifying storm tide risk in Cairns, *Natural Hazards*, 30, 165–185, 2003.
- Huizinga, J., De Moel, H., and Szewczyk, W.: *Global flood depth-damage functions: Methodology and the database with guidelines*, Tech. rep., Joint Research Centre, 2017.
- Huttenlocher, D. P., Klanderman, G. A., and Rucklidge, W. J.: Comparing images using the Hausdorff distance, *IEEE Transactions on pattern analysis and machine intelligence*, 15, 850–863, 1993.
- Jelesnianski, C. P.: A numerical calculation of storm tides induced by a tropical storm impinging on a continental shelf, *Monthly Weather Review*, 93, 343–358, 1965.
- Koks, E. E., Jongman, B., Husby, T. G., and Botzen, W. J.: Combining hazard, exposure and social vulnerability to provide lessons for flood risk management, *Environmental science & policy*, 47, 42–52, 2015.
- Kron, W.: Flood risk= hazard• values• vulnerability, *Water international*, 30, 58–68, 2005.

- 700 Lavell, A., Oppenheimer, M., Diop, C., Hess, J., Lempert, R., Li, J., Muir-Wood, R., Myeong, S., Moser, S., Takeuchi, K., et al.: Climate change: new dimensions in disaster risk, exposure, vulnerability, and resilience, in: *Managing the risks of extreme events and disasters to advance climate change adaptation: Special report of the intergovernmental panel on climate change*, pp. 25–64, Cambridge University Press, 2012.
- Levine, S., Pastor, P., Krizhevsky, A., Ibarz, J., and Quillen, D.: Learning hand-eye coordination for robotic grasping with deep learning and
705 large-scale data collection, *The International journal of robotics research*, 37, 421–436, 2018.
- Li, C., Bi, C., and Li, Z.: Crowd evacuation model based on improved PSO algorithm, *Journal of System Simulation*, 32, 1000, 2020.
- Lillicrap, T. P., Hunt, J. J., Pritzel, A., Heess, N., Erez, T., Tassa, Y., Silver, D., and Wierstra, D.: Continuous control with deep reinforcement learning, *arXiv preprint arXiv:1509.02971*, 2015.
- Lin, N., Emanuel, K. A., Smith, J. A., and Vanmarcke, E.: Risk assessment of hurricane storm surge for New York City, *Journal of Geophysical Research: Atmospheres*, 115, 2010.
710
- Lu, X., Yu, H., Ying, M., Zhao, B., Zhang, S., Lin, L., Bai, L., and Wan, R.: Western North Pacific tropical cyclone database created by the China Meteorological Administration, *Advances in Atmospheric Sciences*, 38, 690–699, 2021.
- Luettich, R. A., Westerink, J. J., Scheffner, N. W., et al.: ADCIRC: an advanced three-dimensional circulation model for shelves, coasts, and estuaries. Report 1, Theory and methodology of ADCIRC-2DD1 and ADCIRC-3DL, 1992.
- 715 Mas, E., Bricker, J., Kure, S., Adriano, B., Yi, C., Suppasri, A., and Koshimura, S.: Field survey report and satellite image interpretation of the 2013 Super Typhoon Haiyan in the Philippines, *Natural Hazards and Earth System Sciences*, 15, 805–816, 2015.
- Merkens, J.-L., Reimann, L., Hinkel, J., and Vafeidis, A. T.: Gridded population projections for the coastal zone under the Shared Socioeconomic Pathways, *Global and Planetary Change*, 145, 57–66, 2016.
- Merz, B., Kreibich, H., Schwarze, R., and Thieken, A.: Review article" Assessment of economic flood damage", *Natural Hazards and Earth
720 System Sciences*, 10, 1697–1724, 2010.
- Mnih, V., Kavukcuoglu, K., Silver, D., Graves, A., Antonoglou, I., Wierstra, D., and Riedmiller, M.: Playing atari with deep reinforcement learning, *arXiv preprint arXiv:1312.5602*, 2013.
- Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., Graves, A., Riedmiller, M., Fidjeland, A. K., Ostrovski, G., et al.: Human-level control through deep reinforcement learning, *nature*, 518, 529–533, 2015.
- 725 Mnih, V., Badia, A. P., Mirza, M., Graves, A., Lillicrap, T., Harley, T., Silver, D., and Kavukcuoglu, K.: Asynchronous methods for deep reinforcement learning, in: *International conference on machine learning*, pp. 1928–1937, PMLR, 2016.
- Müller, M.: Dynamic time warping, *Information retrieval for music and motion*, pp. 69–84, 2007.
- Ni, L., Huang, X., Li, H., and Zhang, Z.: Research on Fire Emergency Evacuation Simulation Based on Cooperative Deep Reinforcement Learning, *Journal of System Simulation*, 34, 1353–1366, 2022.
- 730 of Natural Resources, M.: Guidelines for Risk Assessment and Zoning of Marine Hazards Part 1: Storm Surge, (HY/T 0273-2019), 2019.
- Oppenheimer, M., Glavovic, B., Hinkel, J., van de Wal, R., Magnan, A. K., Abd-Elgawad, A., Cai, R., Cifuentes-Jara, M., Deconto, R. M., Ghosh, T., et al.: Sea level rise and implications for low lying islands, coasts and communities, 2019.
- Rappaport, E. N.: Fatalities in the United States from Atlantic tropical cyclones: New data and interpretation, *Bulletin of the American Meteorological Society*, 95, 341–346, 2014.
- 735 Rhome, J. and Brown, D. P.: Tropical Cyclone Report: Hurricane Katrina, 2006.
- Sallab, A. E., Abdou, M., Perot, E., and Yogamani, S.: Deep reinforcement learning framework for autonomous driving, *arXiv preprint arXiv:1704.02532*, 2017.

- Samah, K., Hussin, B., and Basari, A.: Modification of Dijkstra's algorithm for safest and shortest path during emergency evacuation, *Applied Mathematical Sciences*, 9, 1531–1541, 2015.
- 740 Schulman, J., Wolski, F., Dhariwal, P., Radford, A., and Klimov, O.: Proximal policy optimization algorithms, arXiv preprint arXiv:1707.06347, 2017.
- Snaiki, R., Wu, T., Whittaker, A. S., and Atkinson, J. F.: Hurricane wind and storm surge effects on coastal bridges under a changing climate, *Transportation research record*, 2674, 23–32, 2020.
- Suh, S. W., Lee, H. Y., Kim, H. J., and Fleming, J. G.: An efficient early warning system for typhoon storm surge based on time-varying
745 advisories by coupled ADCIRC and SWAN, *Ocean Dynamics*, 65, 617–646, 2015.
- Sutton, R. S. and Barto, A. G.: *Reinforcement learning: An introduction*, MIT press, 2018.
- Thieken, A. H., Kreibich, H., Müller, M., and Merz, B.: Coping with floods: preparedness, response and recovery of flood-affected residents in Germany in 2002, *Hydrological Sciences Journal*, 52, 1016–1037, 2007.
- UNISDR, C. et al.: *The human cost of natural disasters: A global perspective*, 2015.
- 750 Vickery, P., Skerlj, P., and Twisdale, L.: Simulation of hurricane risk in the US using empirical track model, *Journal of structural engineering*, 126, 1222–1237, 2000.
- Wang, S., Mu, L., Qi, M., Yu, Z., Yao, Z., and Zhao, E.: Quantitative risk assessment of storm surge using GIS techniques and open data: A case study of Daya Bay Zone, China, *Journal of environmental management*, 289, 112 514, 2021a.
- Wang, S., Mu, L., Yao, Z., Gao, J., Zhao, E., and Wang, L.: Assessing and zoning of typhoon storm surge risk with a geographic information
755 system (GIS) technique: a case study of the coastal area of Huizhou, *Natural Hazards and Earth System Sciences*, 21, 439–462, 2021b.
- Wang, Y., Mao, X., and Jiang, W.: Long-term hazard analysis of destructive storm surges using the ADCIRC-SWAN model: A case study of Bohai Sea, China, *International journal of applied earth observation and geoinformation*, 73, 52–62, 2018.
- Willoughby, H. and Rahn, M.: Parametric representation of the primary hurricane vortex. Part I: Observations and evaluation of the Holland (1980) model, *Monthly Weather Review*, 132, 3033–3048, 2004.
- 760 Ying, M., Zhang, W., Yu, H., Lu, X., Feng, J., Fan, Y., Zhu, Y., and Chen, D.: An overview of the China Meteorological Administration tropical cyclone database, *Journal of Atmospheric and Oceanic Technology*, 31, 287–301, 2014.
- Yu, H., Zhang, H., and Xu, W.: A deep compositional framework for human-like language acquisition in virtual environment, arXiv preprint arXiv:1703.09831, 2017.
- Zerger, A.: Examining GIS decision utility for natural hazard risk modelling, *Environmental modelling & software*, 17, 287–294, 2002.
- 765 Zhang, Y., Chai, Z., and Lykotrafitis, G.: Deep reinforcement learning with a particle dynamics environment applied to emergency evacuation of a room with obstacles, *Physica A: Statistical Mechanics and its Applications*, 571, 125 845, 2021.

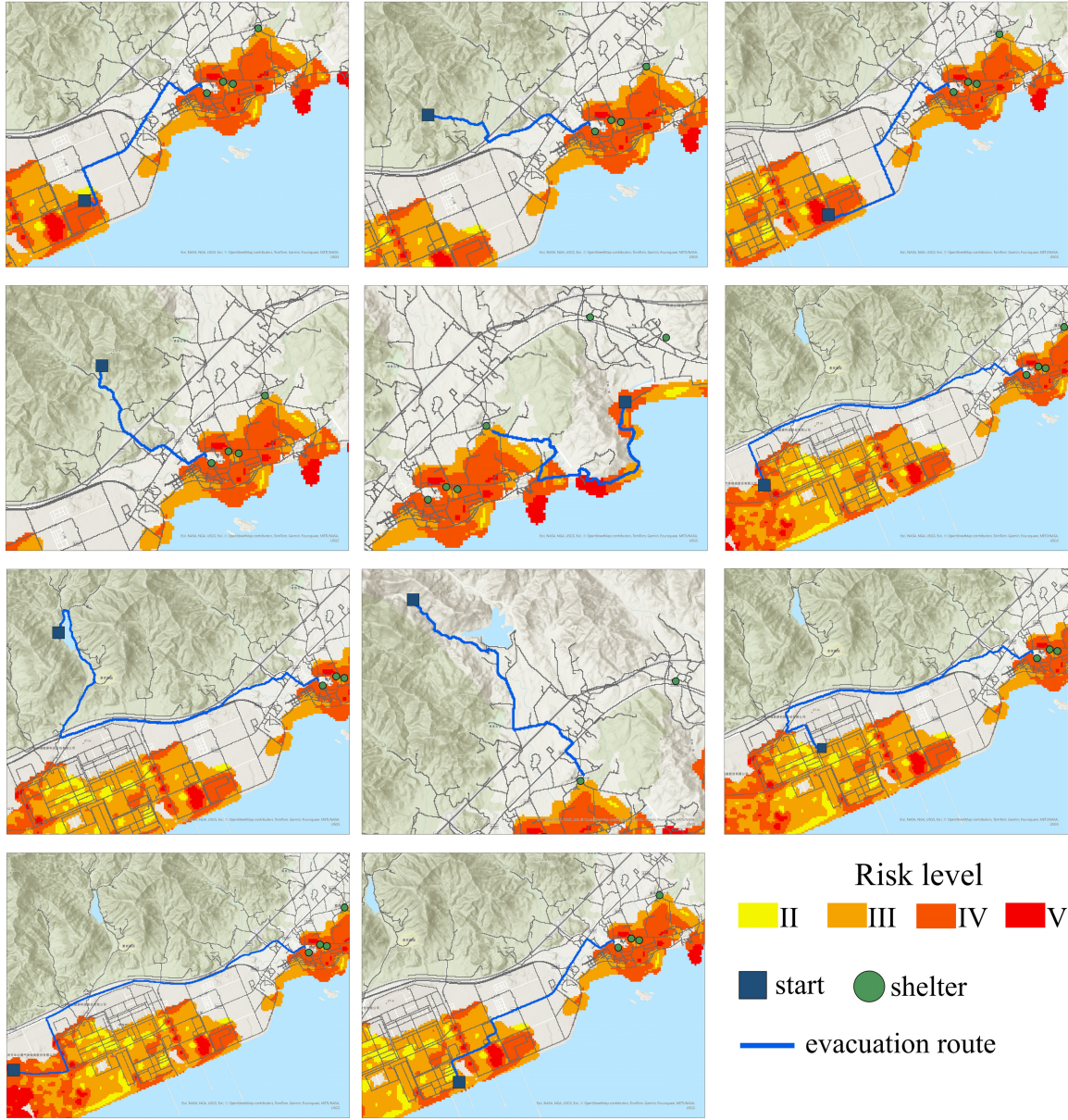


Figure B1. Examples of evacuation route planning. In this study, the goal of evacuation route planning is to find the optimal route from arbitrary starting points to designated shelters. While the inclusion of planning routes from non-flooded areas to shelters may be not reasonable in real-world applications, this approach effectively evaluate the model's route planning capabilities.

$$\text{Minimize}(D \cdot W)$$

where $D = \{d(p, q)\}$ and $d(p, q)$ is the Chebyshev distance between the cell $p \in p_{opt}$ and cell $q \in p_{gen}$. $W = \{w(i, j)\}$ and $w(i, j)$ is the binary DP-variable $i, j \in \mathbb{Z}^+, i \leq l_{opt}, j \leq l_{gen}$. The above DP is subject to:-

$$w(i, j) = \{0, 1\}$$

$$w(1, 1) = 0$$