MARITIME TRANSPORTATION RESEARCH AND EDUCATION CENTER
TIER 1 UNIVERSITY TRANSPORTATION CENTER
U.S. DEPARTMENT OF TRANSPORTATION

MarTREC
Maritime Transportation Research & Education Center

Prediction of Port Recovery Time after a Severe Storm Project

10/1/2022 – 8/31/2023

Zihao Li, Yunlong Zhang, and Bruce Wang,

3136 Zachry Department of Civil Engineering
Texas A&M University
College Station, TX 77843

Final Report Date: August, 2023

FINAL RESEARCH REPORT
Prepared for:
Maritime Transportation Research and Education Center

University of Arkansas
4190 Bell Engineering Center
Fayetteville, AR 72701
479-575-6021

ACKNOWLEDGEMENT

This material is based upon work supported by the U.S. Department of Transportation under Grant Award Number 69A3551747130. The work was conducted through the Maritime Transportation Research and Education Center at the University of Arkansas.

We would like to thank Yue Wu and Jiahao Chen for data collection and data visualization. Kelvin Wang also assisted with this project, particularly with data processing.

DISCLAIMER

ABSTRACT

Predicting the impact of incoming tropical cyclones on ports in terms of the number of days underperforming is crucial for the effective management of the ports. However, existing methods perform undesirably due to the limited data and the inherent uncertainty associated with cyclone trajectory forecasting. This study applies a recommendation algorithm to address these challenges by focusing on predicting port impact rankings instead of predicting the duration of port impacts, which is often inaccurate and unreliable. First, we have collected comprehensive features of ports and hurricanes in the Gulf of Mexican and employed a modular time-series regression model to determine the duration of port impacts due to tropical cyclones, leveraging vessel count data extracted from the Automatic Identification System (AIS). Inspired by the recommendation algorithm, we recast tropical cyclones and ports as "user" and "items," respectively, while the duration of port impacts represents their "interaction," offering an innovative approach to model and analyze cyclone effects on ports. The Factorization Machine (FM) is adopted to learn the relationship between features (i.e., ports and cyclones) and subsequently conduct the port impact ranking. Finally, utilizing the hurricanes Alex, Ian, and Nicole that happened in 2022 as testing cases, the FM-based model excels in prediction performance and robustness against uncertainties compared to the widely used distance-based method. This study aims to provide port authorities and other stakeholders with a trustworthy tool for informed disaster management decisions, thereby enhancing port resilience.

***Keywords*:** Tropical cyclones; Port resilience; Impacted ranking prediction; Recommendation algorithm

1. INTRODUCTION

Maritime transportation system constitutes a critical component of import and export activity in the United States. According to statistical data from Burau of Transportation statistics, approximately 69% of goods traded by U.S. are conveyed through waterways, primarily utilizing seagoing vessels (*1*). More than 95% of all U.S. trade relates to maritime transport, and vessels move around $11.4 trillion worth of products in and out of U.S. ports annually. With its significant role in the economy, maritime transport is an operation-efficient, cost-effective, and environment-friendly transportation mode, especially for long-distance transport (*2*, *3*). With the acceleration of global warming and climate change, the operation of the maritime transportation system confronts unprecedented challenges, such as sea level rise and frequent occurrence of inclement weather (*4*, *5*). Ports are vital to maritime transportation as they facilitate the transfer of goods between different modes and serve as critical nodes in the supply-chain network. However, ports are often located in areas prone to extreme weather events, such as coastal flooding, hurricanes, and tropical storms, which may lead to delays, operation disruption, or infrastructure damage (*6*–*8*). Port performance freight statistic programs (*9*) has revealed that the 2020 Atlantic hurricane season caused 41.1 billion dollars in total damage, and nearly every port along the Gulf and South Atlantic coast endured closures and operation shutdowns due to tropical cyclones (i.e., hurricane and tropical storm) in 2020.

In order to enhance port resilience and minimize the adverse impact of extreme weather conditions on port operation, the main mitigation strategy is proactive disaster preparedness for both port and vessel, including port shutdown and ship rerouting (*7*, *10*). The current hurricane preparation plan heavily relies on weather forecasting. For example, if the port falls within the forecast path of a hurricane, the port has to be closed to all vessel traffic. This approach only focuses on ensuring port operation safety but is ineffective in reducing economic loss due to the uncertainty of the hurricane prediction model (*2*). Moreover, the impact of tropical cyclones on port operations is a complicated and intercorrelated problem involving hurricane severity, landfall location, and port conditions. If the mitigation strategies for port disruption solely rely on weather forecasting, it is too simplistic and does not consider the complexity and multifaceted nature of port operation and disaster management (*10*, *11*). Therefore, a comprehensive prediction model for port's potential impact is imminent for efficient resource allocation and robust contingency planning, reinforcing the resilience of the maritime supply chain against tropical cyclones (*12*). Besides, gaining insights from the model can also assist vessels in making dynamic decisions about rerouting, reducing delays and economic loss from tropical cyclones (*13*).

Past studies mainly tried to quantify the port impact and recovery under inclement weather events based on operation data, including Automatic Identification System (AIS) data and port throughput data (*7*, *14*–*16*). Several studies collect empirical evidence on port disruption incidents to analyze the disruption duration pattern and resilience curve. The statistical analysis improves the understanding of the relationship between extreme weather events on potential consequences across different geographical scales. However, the understanding is inadequate to alleviate the adverse impact of tropical cyclones in the future. In order to better estimate the port performance, researchers have adopted qualitative methods, such as fuzzy analytical hierarch process (*17*), and quantitative methods, including simulation-based method (*18*, *19*) and Bayesian network (*20*, *21*) to assess the port performance under possible hurricane cyclones. While these studies have made contributions to port performance assessment in inclement weather, they are sensitive to model assumption and expert judgment, bringing uncertainty and subjectivity to the model. In recent decades, with the rich AIS data for vessels and the development of data-driven algorithms, artificial intelligence has gained increasing attention and depicts favorable results in port operation (*22*). Many studies have applied machine learning or deep learning algorithms to predict port or vessel operation performance, such as the estimated time of arrival (*23*, *24*) and port throughput (*25*). Nonetheless, these studies predominantly focus on predictions under regular operational conditions, often ignoring the influence of adverse weather conditions. Current research employing machine learning techniques for managing port operations under extreme weather conditions is considerably limited, primarily due to two key challenges. First, the infrequency of extreme weather events leads to insufficient datasets, which hampers the effective use of data-driven approaches. Second, the challenge of appropriately formulating the problem of port operation amidst extreme weather has not been sufficiently addressed.

Therefore, this study aims to overcome these challenges to provide a novel paradigm to predict port performance under tropical cyclones. From the data perspective, we select the Gulf of Mexico as the case and utilize five-year AIS data from 2017 to 2022 to evaluate the port performance. Note that AIS data for Year 2020 is neglected because the maritime transportation system is significantly impacted by the COVID-19 pandemic (*26*), making it difficult to discern whether the impacts were attributable to extreme weather events or the health crisis. Moreover, we collect comprehensive explanatory features about port infrastructure and hurricane records from the database of Department of Transportation (DOT) and National Hurricane Center (NHC), to enhance the prediction performance. In the methodology, we frame the port performance prediction problem as a typical recommendation problem (*27*). The recommendation algorithm is a machine learning algorithm that predicts a user's preferences for various products. These predictions will then be ranked and returned to the users. Contrary to prediction models, recommendation algorithms do not primarily hinge on absolute accuracy (*28*), but rather the correct preference ranking holds utmost significance. The characteristic of recommendation algorithms could potentially be beneficial for disaster management at port operations because the traffic prediction model under hurricanes always performs undesirably (*29*). A reliable port impact ranking considering multiple features can be a useful tool for port disaster management, to inform planning and preparedness for port and vessel (*30*, *31*). In this study, we define each tropical cyclone as a “user,” and each port in the Gulf of Mexico should be an “item.” The adopted recommendation algorithm utilizes the features from both users and items to determine the preference ranking (i.e., port impact ranking under an incoming tropical cyclone). In general, the contribution of this paper has twofold: (i) providing a unique and comprehensive dataset for port operation under tropical cyclones research; (ii) introducing the recommendation algorithm to port impact ranking issue.

This study is organized in sections outlined as follows. In section 2, data collection and the framework of data processing are provided. Section 3 introduces the designed framework of the recommendation algorithm. Then, Section 4 shows the model performance. Finally, the conclusion presents the primary result and the limitation of this study.

2. DATA COLLECTION AND PROCESSING

In this study, we aim to make a final recommendation for port operation based on impact day ranking. The term “impact day” refers to the days on which a port’s performance falls short of normal operations due to tropical cyclones. The impact day ranking is a ranking of ports based on the number of impact days in descending order. Higher ranks denote a more severe impact of a tropical cyclone on the port. In the context of predicting the impact of incoming hurricanes on ports operation, the similarity between the current hurricane and previous hurricanes, the similarity among ports, and historical port impact conditions are the critical indicators. Given its capability to consider similarities between user-item pairs and leverage their historical interactions for ranking user preferences towards items, the recommendation algorithm appears to be an apt solution in this context. Therefore, we characterize an incoming hurricane as the user, ports as items, and the number of impact days as the interaction. A desirable recommendation algorithm would accurately rank the effects of the incoming hurricane on each port. Specifically, the historical cyclones and ports information is combined with the forecast incoming hurricane path provided by National Oceanic and Atmospheric Administration (NOAA) to estimate the potential port impact ranking. In order to build up a more comprehensive and reliable model, the features of each port and cyclone should be collected. In this section, we first introduce the data collection and data processing for port infrastructure and tropical cyclones information. The processed data should be the input of the recommendation algorithm. Then, the approach to determine the impact day for each port due to cyclones is provided.

2.1 Port Infrastructure Data

The abundant oil and gas reserve and the convenience of shipping have made the Gulf Coast the center of the petrochemical industry and its related activities in the United States. However, extreme weather events significantly impact regional economies and port operations. For example, Hurricane Ike in 2008 caused more than 25 million dollars in damage along the coast and also led to hundreds of thousands of people being evacuated or losing their homes. Many ports on Gulf Coast closed part or all operations under Hurricane Laura in 2020. Therefore, this study selects 37 ports in the Gulf of Mexican, and the information on port infrastructure is collected from the port performance freight statistics program in U.S. DOT.

The description of the port features is listed in **Table 1**. The collected features comprehensively describe the port condition from various perspectives, including location, function, operation situation, infrastructure, and harbor condition. Note that the mean, standard deviation, minimal, and maximal values for numerical features (i.e., annual average tonnage) are provided.

Table 1. Descriptive statistics of port infrastructure data

Variable			Num.	Prop. (%)	Variable	Num.	Prop. (%)
State					***Dominate vessel type***		
Alabama (AL)			2	5.4	Tanker	15	40.6
Florida (FL)			8	21.6	Container	7	18.9
Louisiana (LA)			10	27.0	Dry bulk	13	35.1
Mississippi (MS)			6	16.2	Roll-on/roll-off	2	5.4
Texas (TX)			11	29.8	***Harbor type***		
Port governance					River natural	20	54.1
Other			3	8.1	Coastal natural	11	29.7
Municipal			3	8.1	Lake or Canal	4	10.8
Special district			13	35.1	Sea port	2	5.4
Country			13	35.2	***Harbor size***		
State			5	13.5	Very small	7	18.9
If it is container port					Small	16	43.3
Yes			12	32.4	Medium	10	27.0
No			25	67.6	Large	4	10.8
Annual average tonnage					***Shelter condition***		
(millions)					Average	7	33.3
Mean	Std.	Min	Max		Good	21	56.8
35.09	61.09	0.99	275.94		Excellent	9	24.3

Note: Num. stands for the Number and Prop. is for proposition.

2.2 Tropical Cyclones Data

In this study, we collect data on 31 named tropical cyclones that have traversed the Gulf of Mexico from 2017 to 2022 (except for 2020). NHC has recorded detailed hurricane information, including trajectory and corresponding central pressure and wind speed (*32*). Since the destructiveness of tropical cyclones is strongly related to wind speed (*33*), we collect more information specifically related to wind speed rather than central pressure. Given the hurricane trajectory, we extract the formation datetime, landfall datetime, dissipation datetime, and coordinates of the critical points (i.e., formation location, maximum wind speed location, and landfall location). The recorded coordinates are used for further creating more interaction features between port and hurricane, such as the distance between the port to landfall location of a tropical cyclone, which is beneficial for recommendation algorithm to predict more precise impact day ranking. Notes that if the tropical cyclone is not landed, we calculate the distance to the nearest hurricane point to the coastline to replace the distance to the landfall location. Finally, the features of tropical cyclones are summarized in **Table 2**.

Table 2. Descriptive statistics of tropical cyclone data

Variables	Num.	Prop. (%)		Variables	Num.	Prop. (%)	
Year				***Saffir-Simpson hurricane wind scale***			
2017	8	25.8		TS	12	38.7	
2018	5	16.1		H1	8	25.8	
2019	6	19.4		H2	2	6.4	
2021	7	22.6		H3	3	9.7	
2022	5	16.1		H4	3	9.7	
Lowest central pressure (millibars)				H5	3	9.7	
Mean	Std.	Min	Max	***Maximum wind speed (kts)***			
976.3	26.59	914	1003	Mean	Std.	Min	Max
Wind speed at landfall (kts)				77.6	34.2	35	155
Mean	Std.	Min	Max	***Duration from formation to landfall (day)***			
65.8	28.9	35	140	Mean	Std.	Min	Max
Duration from formation to dissipation (day)				4.4	3.3	0.71	15.5
Mean	Std.	Min	Max	***Distance between port and formed location (km)***			
6.7	4.3	2	19	Mean	Std.	Min	Max
Distance between port and maximum wind speed location (km)				1845.8	1570.7	42.2	7465.0
				Distance between port and landfall location (km)			
Mean	Std.	Min	Max	Mean	Std.	Min	Max
1131.0	823.6	8.7	3932.2	867.7	489.0	0.2	2272.3

Note: kts is knot per second and 1 kts = 1.852 km/h. The mean, standard deviation, minimal, and maximal values are calculated based on collected 2017, 2018, 2019, 2021, and 2022 data.

2.3 AIS Data

In the above sections, features of the port infrastructure and tropical cyclones data have been summarized. Then we need to process the impact day for recommendation algorithm. In this study, we adopt the commercial vessel count in the port area as the indicator to represent the port performance (*16*). Therefore, AIS data, also known as vessel traffic data, are used, which tracks the location and characteristics of vessels in real time. The data is valuable for a wide range of maritime applications, including vessel collision avoidance, vessel assignment, and port management (*34*). Because the AIS data is from recording of datetime, GPS location, vessel type, speed, course, heading, and so on every 30 seconds, the AIS data for a meaningful application is usually oversized. For example, the original AIS data for a year is around 300 gigabytes. Preprocessing of the data for use in the research is, therefore, necessary in order to reduce the data to a meaningfully small enough, tractable scale. In this research, we applied multiple criteria for preprocessing (*7*), as shown in the following:

- Removing the non-commercial vessel types (e.g., fishing boat, tug, and towing boat) and only keeping the commercial vessel types such as cargo, tanker, and container vessel.
- Defining the port boundary by specifying a range of coordinates, then counting the number of active vessels within the port boundary.
- Aggregating the hourly time-series vessel counts on a daily basis from the year 2019 to 2022 (except for 2020).

After the preprocessing, the daily vessel count for each port is obtained. The number of impact days could subsequently be determined based on the time-series data. Existing studies have used the threshold method (*7*) and Bayesian change point detection (*16*) to identify the impact days. However, they largely ignore the seasonality of time series data and are limited to certain distribution assumptions. In contrast, we have adopted a modular time-series regression model (*35*) to identify the impact days. The modular time-series regression model is a type of forecasting model that allows for different components of the time-series data, such as trend, seasonality, and other factors, to be modeled separately and then combined additively or multiplicatively. This modular approach makes it easier to interpret the model and accommodate changes in the data structure over time, thereby improving the model's flexibility and forecasting accuracy. Specifically, the additive regression model with components for trend and seasonality is trained by one-year traffic counts data to learn the general pattern of the vessel count for a specific port. For each data point, a prediction interval can be calculated based on the specified confidence interval. If the vessel counts in a day that falls below the lower limit of this prediction interval, it indicates underperformance with confidence and therefore is classified as an impact day. Furthermore, if this identified impact day coincides with a period during which a hurricane occurred, we attribute the impact to the presence of the hurricane. In **Figure 1**, the time-series pattern is successfully captured by additive regression, with most data points lying within the 90% confidence prediction interval. In the 2017 hurricane season, only Hurricane Harvey had a severe impact on the operations of Port Houston for 6 days. Note that a day with vessel counts exceeding the upper limit of the prediction interval is not considered as an impact day, and it is because of the production recapture, which means the port can make up for disruption by shifting more cargo when they become normally operational again (*7*). After analyzing all ports across different years, the impact day for ports under each hurricane is obtained. The mean, standard deviation, minimum and maximum values of the impact days for all selected ports from 2017 to 2022 (excluding 2020) due to tropical cyclones are 0.79, 1.53, 0, and 9 days, respectively.

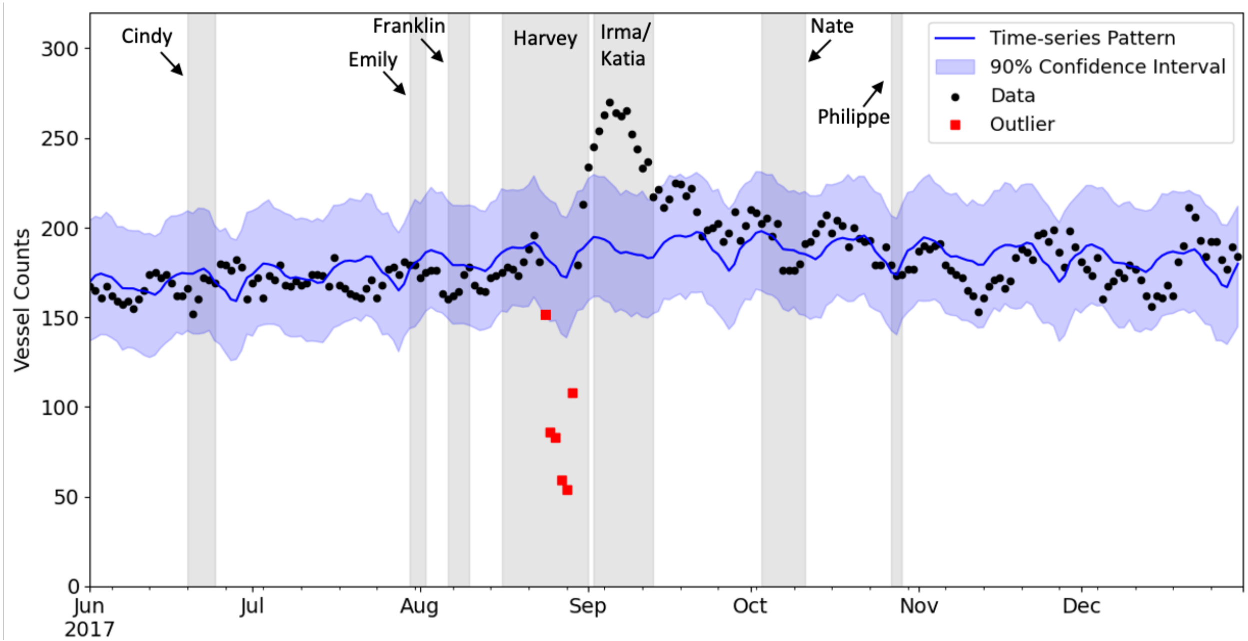

Figure 1. Performance of Port Houston during the 2017 hurricane season

3. METHODOLOGY

Our proposed recommendation system process has three main steps, mining, retrieving, and ranking, as shown in **Figure 2**. The mining step is to collect data about ports, tropical cyclones, and most importantly, their interaction. The retrieving is to select a subset of relevant ports potentially affected by tropical cyclones to efficiently identify the likely impacted port. Due to the limitation of data size with only 31 tropical cyclones available, the commonly used method, such as user-based or item-based collaborative filtering, is hard to understand the complex interaction between hurricanes and the ports (*36*). Therefore, we define a rule to select potentially impacted ports from the selected total of 37 ports in the Gulf of Mexico. Based on the forecast landfall location of each hurricane, we select the ten nearest ports to the landfall location for training and ranking. Note that the port impact ranking is used for future hurricanes, and the input of the recommendation algorithm should be the forecast hurricane trajectory because, in a current situation, the trajectory would not be observed but only forecast. However, the hurricane trajectory prediction has inherent uncertainly(*32*), especially for long-term forecast periods. According to the forecast error statistics from NHC (*37*), the forecast error for 48 hours is about 65 nautical miles (around 120 km). To incorporate the forecast error of hurricane trajectory, we introduce a random offset by generating a single point randomly from the coastline within a 120 km radius around the recorded landfall location. We repeat this process ten times to obtain ten probable "actual" landfall locations. Each "actual" landfall corresponds to a trajectory calculated from its location to the port. In this study, for each hurricane, we generate ten random landfall locations as input following a uniform distribution, assuming the actual landfall location is equally likely distributed on the coastline within the radius of 120 km to replace the actual landfall location. The actual location randomly generated may have an effect on the results to some extent. Given limited information, uniform distribution is currently an accepted assumption. However, the number of impact days is calculated based on the observed/recorded trajectory of the tropical cyclone. In other words, there are ten different landfall locations generated for a specific tropical cyclone, but their number of impact days is the same. Through this approach, we aim to enhance the robustness of the recommendation algorithm against the uncertainty associated with the forecasting of incoming cyclones' paths. Finally, there are a total of 11,470 (37 ports × 31 tropical cyclones × 10 random landfall locations) records about the instances, called interaction in this paper, between ports and tropical cyclones. Furthermore, we use the 2022 data as the test dataset and the 2017, 2018, 2019, and 2021 data as the training dataset. Details of the recommendation algorithm and performance measures are introduced in the following sections.

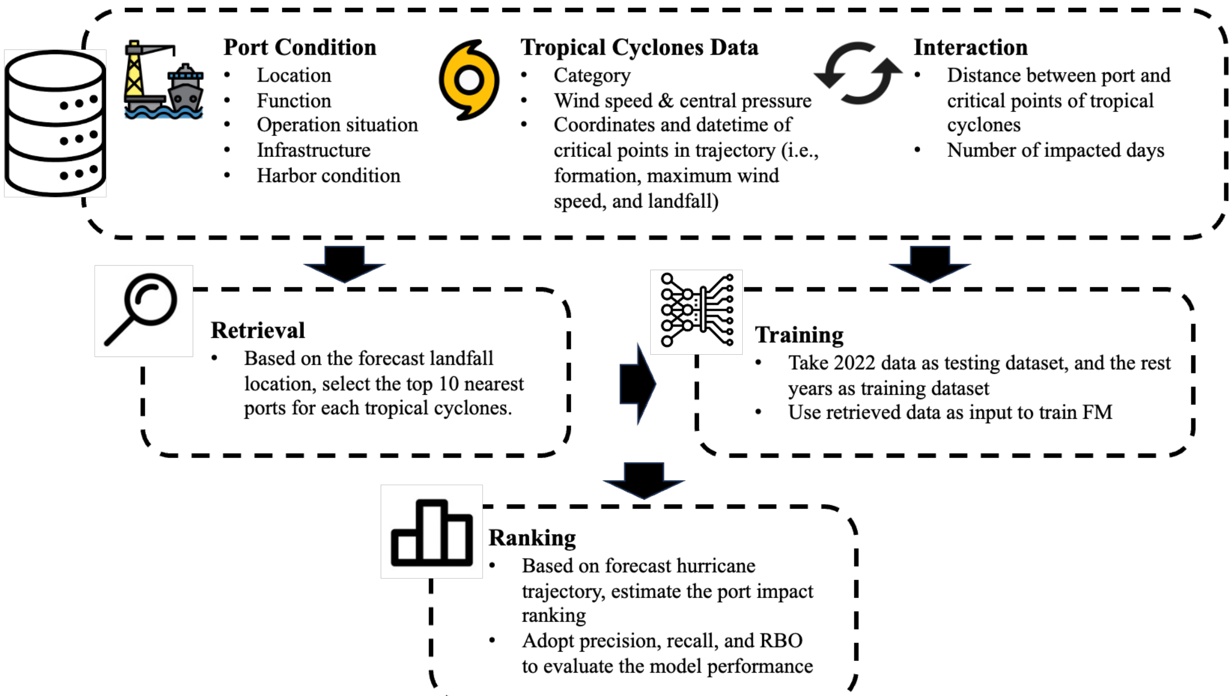

Figure 2. Framework of the proposed recommendation system for port impact ranking

3.1 Proposed Recommendation Algorithm – Factorization Machine

Factorization Machine (FM) is a machine learning model that efficiently captures higher-order feature interaction by factorizing the interaction matrix, making it well-suited for handling sparse and high-dimensional data in the recommendation system (*38*). Compared to linear models, FM can model second-order feature interaction through factorization of the interaction matrix into latent vectors. Therefore, FM is able to deduce interactions even with extremely sparse data. Note that the input of FM can be only categorical features. To generate the second-order interaction terms, we have discretized the continuous features in Section 2 into categorical measures using the 25th, 50th, and 75th percentiles in the same way as (*39*).

In FM, the model is combined with a linear regression component and feature interaction component, as in Eq. (1). However, it will simultaneously increase the computational complexity due to the interaction components. Therefore, FM model the feature interaction using latent factors v. Each feature x_i has a latent factor v_i of size k (i.e., hyperparameter), and two features' interaction are modeled as $\langle v_i, v_j \rangle = \sum_{f=1}^{k} v_{i,f} v_{j,f}$, where $\langle\ ,\ \rangle$ is the dot product of the two latent factors. Finally, the number of parameters for the interaction term reduces from n^2 to $n \times k$. Introducing matrix $\boldsymbol{V}$ is similar to factorizing the original weight matrix of the interaction term.

$$\hat{y}(\boldsymbol{x}) = \underbrace{w_0 + \sum_{i=1}^{n} w_i x_i}_{\text{linear regression}} + \underbrace{\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \langle v_i, v_j \rangle x_i x_j}_{\text{interaction}} \tag{1}$$

Where w_0 is the bias term and w_i is the weight corresponding to feature vector x_i. To further reduce the time computational complex from $O(kn^2)$ in Eq. (1) to $O(kn)$, Eq. (1) can be rewritten as Eq. (2) in light of (*38*).

$$\hat{y}(\boldsymbol{x}) = w_0 + \sum_{i=1}^{n} w_i x_i + \frac{1}{2} \sum_{f=1}^{k} \left(\left(\sum_{i}^{n} v_{i,f} x_i \right)^2 - \sum_{i=1}^{n} v_{i,f}^2 x_i^2 \right) \tag{2}$$

In the end, we use a stochastic gradient descent (SGD) and least squares error to train FM in this study. The gradients of the parameters w_0, $\boldsymbol{w}$, and $\boldsymbol{V}$ are as described below:

$$\frac{\partial}{\partial \theta} \hat{y}(\boldsymbol{x}) = \begin{cases} 1, & \text{if } \theta \text{ is } w_0 \\ x_i & \text{if } \theta \text{ is } w_i \\ x_i \sum_{j=1}^{n} v_{j,f} x_j - v_{i,f} x_i^2 & \text{if } \theta \text{ is } v_{i,f} \end{cases} \tag{3}$$

3.2 Performance Measures and Baseline Method

In this study, we evaluate the model performance in the testing dataset from two perspectives: first, the capability to accurately identify the actually impacted ports within the top L ranks, and second, the degree of concordance between the estimated port impact ranking and their corresponding ranking of ground truth impact. The measures, including precision and recall, are introduced to assess the model's ability to differentiate between impacted and non-impacted ports. Precision measures the proportion of the top L ports estimated that are actually impacted, and recall is the proportion of all the ports actually impacted that are correctly estimated, as shown in Eqs. (4) and (5). Both measures range from 0 to 1, and a higher value indicates a better performance.

$$\text{precision} = \frac{|U \cap L|}{|L|} \tag{4}$$

$$\text{recall} = \frac{|U \cap L|}{|U|} \tag{5}$$

Where U is the set of actually impacted ports, and L represents the set of top ports predicted to be impacted by incoming tropical cyclones. Here, we set the $|L| = 10$ for the FM-based method. $|\cdot|$ represents the size of a set.

Furthermore, Rank Biased Overlap (RBO) is a measure used to evaluate the similarity between two ranked lists while considering the importance of rank position. It quantifies the agreement between estimated and actual rankings of ports impacted, considering that ports ranked higher have higher importance. In other words, RBO provides a finer evaluation of the model's performance, reflecting its ability to rank the impacted ports more accurately.

$$RBO = (1 - p) \sum_{d=1}^{D} p^{d-1} \cdot \frac{|U_{1:d} \cap L_{1:d}|}{d} \quad (6)$$

Where D is the depth of the list L, which indicates the position or rank within the ranked list. p is the hyperparameter, and we set it as 0.6. The range of RBO is from 0 to 1, and the larger values indicate better performance.

Besides, a distance-based method is selected to be the baseline algorithm in this study, which calculates the impact based on the distance from the landfall location to the port, as detailed below. The distance-based method is intuitive and widely used in real-life scenarios. A port will be impacted by tropical cyclones if it falls within the radius of tropical cyclones, and the severity of the impact increase as the port's proximity to the landfall location decreases (*40*). Based on the definition of the hurricane strike cycle from NHC (*41*), the radius of the strike cycle is around 75 nautical miles (around 140 km). Considering the secondary disaster of tropical cyclones, including flooding and heavy precipitation, we choose to extend the radius to 300 km, centered on the hurricane's landfall location. Thus, if the distance of a port to the landfall location or nearest point of the cyclone is less than 300 km, the distance-based method assumes that the port is impacted by a specific tropical cyclone. The ranking of impacted ports aligns with the distance ranking, meaning that the closer the distance, the more severe the impact.

4. RESULT AND DISCUSSION

In this section, we use the hurricanes that happened in the Gulf of Mexico in 2022, including Hurricanes Alex, Ian, and Nicole, as the test dataset, as shown in **Figure 3**. The hurricane-related features for input are based on the forecast hurricane trajectory rather than the actual trajectory. As described earlier, in order to evaluate the sensitivity of the model to hurricane landfall uncertainty, we randomly generate ten landfall locations from the intersection of uncertainty cone of incoming hurricane and coastline, as input. An ideal method shall balance performance and robustness. In other words, the model shall perform expectedly well considering the uncertainty of the forecast hurricane trajectories.

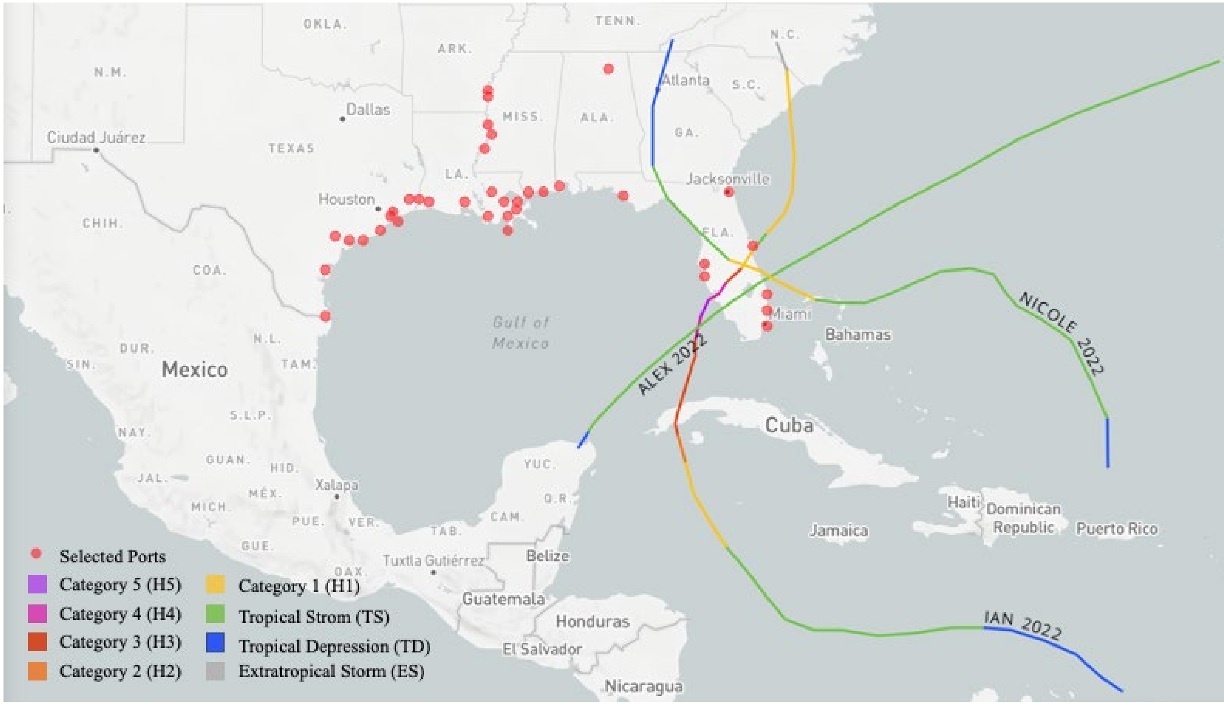

Figure 3. Selected ports and tropical cyclones in the Gulf of Mexico in 2022 (*42*)

In **Figure 4**, the precision of FM-based method is lower than distance-based method under three hurricanes. It is because the FM-based method is more conservative than distance-based, in which the method assumes the top 10 nearest ports to landfall location are impacted. Conversely, the distance-based method only considers the port within 300 km to the landfall location as an impact port. Nevertheless, the FM-based method remains desirable. Taking Hurricane Alex as an example, there are eight impacted ports for the top 10 ranked list, of which the highest precision for FM-based method is 0.8 (i.e., 8/10). Despite the FM-based method having a lower precision than the distance-based method, it has already reached the upper-performance threshold. Furthermore, the FM-based method exhibits relatively higher recall. A higher recall indicates that the model has a stronger ability to capture as many impacted ports as possible, demonstrating its effectiveness in capturing impacted ports. This property enables the port authorities to reduce the potential economic loss due to unpreparedness and inadequate planning.

Moreover, the accuracy of the ranking sequence is also crucial for performance evaluation. The higher RBO of the FM-based method demonstrates that the proposed recommendation algorithm performed well not only in identifying whether a port is impacted or not but also in the ranking of the impact. The error bars in **Figure 4** represent the standard deviation. Three measures of distance-based method have larger standard deviations due to the impact of uncertainty from hurricane trajectory forecasting. In contrast, the FM-based method generates results robust to input uncertainty. The upper threshold of some measures in distance-based method (e.g., RBO in Hurricane Nicole) is higher than in FM-based method. The result suggests that if a hurricane trajectory is accurate, the distance-based method is a more suitable choice. However, considering the existence of uncertainty and the importance of reliability in disaster management, the FM-based method is more desirable for predicting the impact of incoming cyclones on port operations.

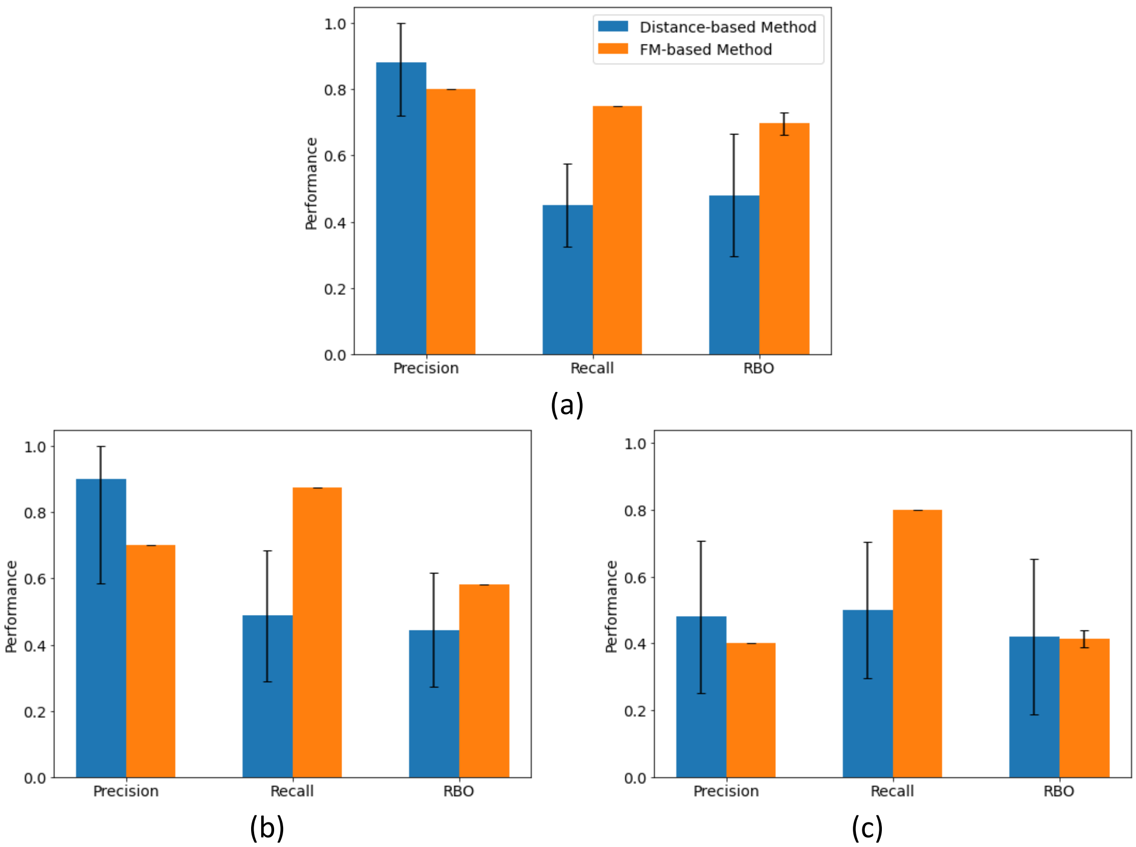

Figure 4. Performance metrics of distance-based and FM-based methods under hurricane Alex (a), Ian (b), and Nicole (c). (Note: the actual number of impacted ports for three hurricanes is 8, 8, 5, respectively)

To better demonstrate the performance of the FM-based method, we randomly select three forecasted hurricane trajectories of Hurricane Alex within the uncertainty cone, and we list their estimated port impact rankings correspondingly as shown in **Table 3**. The detailed port impact rankings also show that the FM-based method outperforms distance-based method in terms of accuracy and robustness. The unreliability of the distance-based method may be exacerbated by an increase in forecasting duration because long-term forecasting often leads to amplification in forecasting errors. The distance-based method has two main disadvantages. First, the estimated port impact is unreliable, especially with long-term forecasting (e.g., 48/96-hour). If the ports around the Gulf prepare based on the predicted impact with it, it may lead to some ports over-prepare while others under-prepared, both of which can increase safety risk and economic losses. When the distance-based method becomes reliable (i.e., under short-term forecasting), there may not be enough time to react appropriately, rendering the distance-based method undesirable overall. In contrast, the proposed FM-based method can alleviate these disadvantages and provide a more reliable and robust estimation of port impact ranking for incoming tropical cyclones.

Table 3. Comparison of estimated port impact ranking and actual ranking

Port Impact Ranking	Actual Impact	Distance-based 1	Distance-based 2	Distance-based 3	FM-based 1	FM-based 2	FM-based 3
1	Everglades	Manatee	PlamBeach	Jacksonville	Everglades	Everglades	Everglades
2	PlamBeach	Tampa	Everglades	Canaveral	Canaveral	Canaveral	PlamBeach
3	Canaveral	Canaveral	Miami	Manatee	Tampa	Tampa	Miami
4	Miami	Miami	\	Tampa	Miami	PlamBeach	Tampa
5	Mobile	Everglades	\	\	PlamBeach	Miami	Canaveral
6	Panama	PlamBeach	\	\	Manatee	Manatee	Panama
7	Pascagoula	\	\	\	Jacksonville	Jacksonville	Mobile
8	Tampa	\	\	\	Panama	Panama	Pascagoula
9	\	\	\	\	Mobile	Mobile	Manatee
10	\	\	\	\	Pascagoula	Pascagoula	Jacksonville

Note: The names of ports are all abbreviations, and the full names should have the prefix “Port of” added before the abbreviation. Distance-based 1 means the port impact ranking from distance-based method based on trajectory 1 of Hurricane Alex.

5. CONCLUSIONS

This study introduces a recommendation system designed to predict impacted ports in anticipation of incoming tropical cyclones. Traditional prediction models often fall short of expectations because they typically demand tremendous, detailed data. In contrast, the proposed recommendation algorithm circumvents the issue by focusing on predicting port impact ranking rather than predicting the specific duration of port impacts. In the first phase, we collect features pertinent to ports and cyclones and employed an additive regression model to identify the days impacted, as reflected in the Automatic Identification System (AIS) vessel count data. Subsequently, a recommendation system is proposed with two steps after data preparation: retrieval and ranking. In the retrieval step, ten potentially impacted ports are filtered from all ports based on the distance of each hurricane landfall. The ranking step incorporates multiple features about the ports, historical tropical cyclones, and their interaction to conduct the port impact ranking. An FM model is adopted to learn the mapping between input features and port impact rankings. An important feature of the proposed method to account for resiliency to the uncertainty of hurricane trajectory forecast is that we replace the actual single trajectory with ten forecasted, possible trajectories during our model training process.

The performance of the proposed recommendation system is evaluated from precision, recall, and RBO by using the 2022 data. The results indicate that the proposed FM-based recommendation algorithm shows superior performance compared to the existing distance-based methods in terms of prediction accuracy of meaningfully advanced dates and robustness to hurricane uncertainty. Consequently, our model can provide a trustworthy and reliable port impact ranking for incoming cyclones. The significance of this study is to deliver a valuable tool for port authorities and stakeholders in port disaster management and, at the same time, enhance port resilience by providing more accurate, advanced hurricane impact information.

This study has certain limitations that may also open avenues for future research. Although a more reliable ranking model for port disaster management is proposed, the model currently only considers several critical locations of tropical cyclones (e.g., formation and landfall locations). In the future, a complete hurricane trajectory could be incorporated to extract more related features to improve the representation ability of the recommendation algorithm. Furthermore, applying more advanced recommendation algorithms, such as DeepFM (*43*) and Wide & Deep Learning (*44*), could further improve the model's performance and shed new light on the crucial factors influencing port performance under tropical cyclones.

REFERENCES

1. FHWA. National Highway Freight Network. https://ops.fhwa.dot.gov/freight/infrastructure/nfn/index.htm. Accessed May 8, 2021.
2. Mollaoglu, M., B. C. Altay, and A. Balin. Bibliometric Review of Route Optimization in Maritime Transportation: Environmental Sustainability and Operational Efficiency. *Transportation Research Record*, Vol. 2677, No. 6, 2023, pp. 879–890. https://doi.org/10.1177/03611981221150922.
3. Psaraftis, H. N., and C. A. Kontovas. Balancing the Economic and Environmental Performance of Maritime Transportation. *Transportation Research Part D: Transport and Environment*, Vol. 15, No. 8, 2010, pp. 458–462.
4. Webster, P. J., G. J. Holland, J. A. Curry, and H.-R. Chang. Changes in Tropical Cyclone Number, Duration, and Intensity in a Warming Environment. *Science*, Vol. 309, No. 5742, 2005, pp. 1844–1846.
5. Pielke Jr, R. A., C. Landsea, M. Mayfield, J. Layer, and R. Pasch. Hurricanes and Global Warming. *Bulletin of the American Meteorological Society*, Vol. 86, No. 11, 2005, pp. 1571–1576.
6. Ng, A. K., A. Becker, S. Cahoon, S.-L. Chen, P. Earl, and Z. Yang. *Climate Change and Adaptation Planning for Ports*. Routledge, 2015.
7. Verschuur, J., E. E. Koks, and J. W. Hall. Port Disruptions Due to Natural Disasters: Insights into Port and Logistics Resilience. *Transportation Research Part D: Transport and Environment*, Vol. 85, 2020, p. 102393. https://doi.org/10.1016/j.trd.2020.102393.
8. Asariotis, R., Hassiba Benamara, and V. Mohos-Naray. Port Industry Survey on Climate Change Impacts and Adaptation. 2018. https://doi.org/10.13140/RG.2.2.28176.66569.
9. Hu, P. S., R. R. Schmitt, R. Robinson, M. Chambers, and W. H. Moore. 2023 Port Performance Freight Statistics Program: Annual Report to Congress. 2023.

10. Li, L., X. Xiong, and H. Yuan. Ships’ Response Strategies to Port Disruptions Caused by Hurricanes. *Ocean & Coastal Management*, Vol. 227, 2022, p. 106275.
11. Sturgis, C. L. A., T. C. Smythe, and A. Tucci. Port Recovery in the Aftermath of Hurricane Sandy. *Center for a New American Security*, 2014.
12. Lin Moe, T., and P. Pathranarakul. An Integrated Approach to Natural Disaster Management: Public Project Management and Its Critical Success Factors. *Disaster Prevention and Management: An International Journal*, Vol. 15, No. 3, 2006, pp. 396–413.
13. Edwards, D. R., F. O. Idoko, C. Vogiatzis, L. B. Davis, and P. Mirchandani. Determining Optimal Fuel Delivery Strategies under Uncertainty. *Socio-Economic Planning Sciences*, Vol. 88, 2023, p. 101612.
14. Farhadi, N., S. A. Parr, K. N. Mitchell, and B. Wolshon. Use of Nationwide Automatic Identification System Data to Quantify Resiliency of Marine Transportation Systems. *Transportation Research Record*, Vol. 2549, No. 1, 2016, pp. 9–18.
15. Trepte, K., and J. B. Rice Jr. An Initial Exploration of Port Capacity Bottlenecks in the USA Port System and the Implications on Resilience. *International Journal of Shipping and Transport Logistics*, Vol. 6, No. 3, 2014, pp. 339–355.
16. Touzinsky, K. F., B. M. Scully, K. N. Mitchell, and M. M. Kress. Using Empirical Data to Quantify Port Resilience: Hurricane Matthew and the Southeastern Seaboard. *Journal of Waterway, Port, Coastal, and Ocean Engineering*, Vol. 144, No. 4, 2018, p. 05018003.
17. Tahri, M., M. Maanan, M. Maanan, H. Bouksim, and M. Hakdaoui. Using Fuzzy Analytic Hierarchy Process Multi-Criteria and Automatic Computation to Analyse Coastal Vulnerability. *Progress in Physical Geography*, Vol. 41, No. 3, 2017, pp. 268–285.
18. Folkman, D., A. Gharehgozli, J. Mileski, and C. B. Galvao. Port Resiliency and the Effects of Hurricanes on Port Operations. *International Journal of Advanced Operations Management*, Vol. 13, No. 4, 2021, pp. 409–430.
19. Dhanak, M., S. Parr, E. I. Kaisar, P. Goulianou, H. Russell, and F. Kristiansson. Resilience Assessment Tool for Port Planning. *Environment and Planning B: Urban Analytics and City Science*, Vol. 48, No. 5, 2021, pp. 1126–1143.
20. Wang, N., M. Wu, and K. F. Yuen. Assessment of Port Resilience Using Bayesian Network: A Study of Strategies to Enhance Readiness and Response Capacities. *Reliability Engineering & System Safety*, Vol. 237, 2023, p. 109394. https://doi.org/10.1016/j.ress.2023.109394.
21. Hossain, N. U. I., S. El Amrani, R. Jaradat, M. Marufuzzaman, R. Buchanan, C. Rinaudo, and M. Hamilton. Modeling and Assessing Interdependencies between Critical Infrastructures Using Bayesian Network: A Case Study of Inland Waterway Port and Surrounding Supply Chain Network. *Reliability Engineering & System Safety*, Vol. 198, 2020, p. 106898.
22. Filom, S., A. M. Amiri, and S. Razavi. Applications of Machine Learning Methods in Port Operations – A Systematic Literature Review. *Transportation Research Part E: Logistics and Transportation Review*, Vol. 161, 2022, p. 102722. https://doi.org/10.1016/j.tre.2022.102722.
23. Park, K., S. Sim, and H. Bae. Vessel Estimated Time of Arrival Prediction System Based on a Path-Finding Algorithm. *Maritime Transport Research*, Vol. 2, 2021, p. 100012.
24. El Mekkaoui, S., L. Benabbou, and A. Berrado. Deep Learning Models for Vessel’s ETA Prediction: Bulk Ports Perspective. *Flexible Services and Manufacturing Journal*, Vol. 35, No. 1, 2023, pp. 5–28.

25. Li, Y., T. Li, Y. Zuo, C. P. Chen, Q. Shan, Y. Xiao, and X. Fan. A Review of Research on Port Throughput Forecasting. 2019.
26. Choi, T.-M. *Risk Analysis in Logistics Systems: A Research Agenda during and after the COVID-19 Pandemic*. Elsevier, 2021.
27. Isinkaye, F. O., Y. O. Folajimi, and B. A. Ojokoh. Recommendation Systems: Principles, Methods and Evaluation. *Egyptian informatics journal*, Vol. 16, No. 3, 2015, pp. 261–273.
28. Ricci, F., L. Rokach, and B. Shapira. Recommender Systems: Techniques, Applications, and Challenges. *Recommender Systems Handbook*, 2021, pp. 1–35.
29. Roy, K. C., S. Hasan, A. Culotta, and N. Eluru. Predicting Traffic Demand during Hurricane Evacuation Using Real-Time Data from Transportation Systems and Social Media. *Transportation Research Part C: Emerging Technologies*, Vol. 131, 2021, p. 103339. https://doi.org/10.1016/j.trc.2021.103339.
30. Asgari, N., A. Hassani, D. Jones, and H. H. Nguye. Sustainability Ranking of the UK Major Ports: Methodology and Case Study. *Transportation Research Part E: Logistics and Transportation Review*, Vol. 78, 2015, pp. 19–39.
31. Hanson, S., R. Nicholls, N. Ranger, S. Hallegatte, J. Corfee-Morlot, C. Herweijer, and J. Chateau. A Global Ranking of Port Cities with High Exposure to Climate Extremes. *Climatic Change*, Vol. 104, No. 1, 2011, pp. 89–111. https://doi.org/10.1007/s10584-010-9977-4.
32. Cangialosi, J. P., E. Blake, M. DeMaria, A. Penny, A. Latto, E. Rappaport, and V. Tallapragada. Recent Progress in Tropical Cyclone Intensity Forecasting at the National Hurricane Center. *Weather and Forecasting*, Vol. 35, No. 5, 2020, pp. 1913–1922.
33. Emanuel, K. Increasing Destructiveness of Tropical Cyclones over the Past 30 Years. *Nature*, Vol. 436, No. 7051, 2005, pp. 686–688. https://doi.org/10.1038/nature03906.
34. Yang, D., L. Wu, S. Wang, H. Jia, and K. X. Li. How Big Data Enriches Maritime Research–a Critical Review of Automatic Identification System (AIS) Data Applications. *Transport Reviews*, Vol. 39, No. 6, 2019, pp. 755–773.
35. Taylor, S. J., and B. Letham. Forecasting at Scale. *The American Statistician*, Vol. 72, No. 1, 2018, pp. 37–45.
36. Ricci, F., L. Rokach, and B. Shapira. Introduction to Recommender Systems Handbook. In *Recommender systems handbook*, Springer, pp. 1–35.
37. NHC Tropical Cyclone Graphical Product Descriptions. https://www.nhc.noaa.gov/aboutnhcgraphics.shtml. Accessed Jul. 21, 2023.
38. Rendle, S. Factorization Machines. 2010.
39. Li, Z., X. Kong, and Y. Zhang. Exploring Factors Associated With Crossing Assertiveness of Pedestrians at Unsignalized Intersections. *Transportation Research Record*, 2023, p. 03611981221145140. https://doi.org/10.1177/03611981221145140.
40. Balakrishnan, S., T. Lim, and Z. Zhang. A Methodology for Evaluating the Economic Risks of Hurricane-Related Disruptions to Port Operations. *Transportation Research Part A: Policy and Practice*, Vol. 162, 2022, pp. 58–79. https://doi.org/10.1016/j.tra.2022.05.010.
41. Glossary of NHC Terms. https://www.nhc.noaa.gov/aboutgloss.shtml. Accessed Jul. 22, 2023.
42. Historical Hurricane Tracks. https://coast.noaa.gov/hurricanes/#map=4/32/-80. Accessed Jul. 24, 2023.

43. Guo, H., R. Tang, Y. Ye, Z. Li, and X. He. DeepFM: A Factorization-Machine Based Neural Network for CTR Prediction. *arXiv preprint arXiv:1703.04247*, 2017.
44. Cheng, H.-T., L. Koc, J. Harmsen, T. Shaked, T. Chandra, H. Aradhye, G. Anderson, G. Corrado, W. Chai, and M. Ispir. Wide & Deep Learning for Recommender Systems. 2016.