Journal of Applied Research and Technology

www.jart.icat.unam.mx

Journal of Applied Research and Technology 22 (2024) 494-509

Original

A predictive methodology for vessel travel times: An application on the Gulf Intracoastal Waterway (GIWW)

M. Monsreal[a] • C. J. Kruse[b] • D. H. Kang[a] • R. B. Carmona-Benítez[c]*

[a]*Texas A&M Transportation Institute, Bryan, Texas, United States of America*
[b]*Texas A&M Transportation Institute, Houston, Texas, United States of America*
[c]*Universidad Anáhuac México, Facultad de Economía y Negocios, Estado de México, México*

Received 02 16 2024; accepted 05 24 2024
Available 08 31 2024

Abstract: The Gulf Intracoastal Waterway (GIWW) is one of the most used corridors in the U.S. inland waterway commerce network, necessitating accurate travel time estimation for operational planning (departure times and on-time arrivals). This paper addresses the significant need to assess GIWW travel times and proposes a two-phase approach using Automatic Identification System data and other data sets.
In the initial phase, forecasting models and event evaluation methods were applied to predict travel times based on specific events, and in the second phase, the impact of different variables on system performance was investigated.
The results indicate that sample count (completed trips through a link) does not significantly influence travel time across any link. The statistical analysis highlights two critical conditions affecting travel time: dredging and shoaling. Furthermore, the analysis presented in this paper estimates the expected magnitude of these events and their probability of occurrence.
By applying the proposed methodology to estimate travel times of the GIWW, this paper contributes to enhancing travel time estimation tools, offering valuable information for decision-makers, operators, and users navigating this crucial waterway.

Keywords: Forecasting, event evaluation, vessel travel time, automatic identification system (AIS)

*Corresponding author.
E-mail address: rafael.carmona@anahuac.mx (R. B. Carmona-Benítez).
Peer Review under the responsibility of Universidad Nacional Autónoma de México.

1. Introduction

Transportation infrastructure connects various kinds of facilities (suppliers, production, distribution, warehouses, points of sale, and customers) through a supply chain network. Transportation infrastructure is important for the development of regional and national economies because it connects them to international markets (Welch et al., 2023). Supply chain transportation is composed of four transportations modes: sea, air, train, and road. Maritime is the most significant mode of transportation because over 80 percent of cargo is transported through this transportation mode worldwide (Korkmaz et al., 2023). Maritime transport is the transportation of humans and goods through waterways. This mode of transportation is accomplished over oceans, lakes, rivers, and canals; and it includes ports, terminals, locks and dams, navigation channels, and shipping vessels (Mahmoudzadeh et al., 2021). Although maritime transport is the most used cargo transportation mode over oceans, inland waterways contribute less to national economies than rail and road do, due to underinvestment (Tan et al., 2022). Despite recognizing the inherent advantages of inland waterways in supply chain coordination and their potential to impact consumer prices positively (Wohlgemuth et al., 2020) by enabling the flow of goods and services at low costs (Brum et al., 2023), there is a growing need to explore their untapped potential.

The study of the transportation of cargo through inland waterways is highly relevant because it can help to maintain the flow of goods and even alleviate disruptions caused by the interruption of other transportation modes, which can have a serious impact on businesses and the economy of regions and countries (Chen & Li, 2021). Although historically the inland waterway modal system has been the subject of research, the number of papers studying these organization systems in supply chains is limited (de Barros et al., 2022). Moreover, these articles focus on analyzing the level of impact that waterway infrastructure has in socioeconomic systems (Mehan & Casey, 2023), and a few of them focus on the estimation of vessel travel or traversal times in inland waterways. So far, few models have been developed to estimate vessel travel times, and these models estimate travel times based on pilots' experience and navigation reports (Wu et al., 2020). These estimates are not fully accurate because varied factors, such as a captain's experience or environmental uncertainties, cause discrepancies between the actual and reported travel times (Fan et al., 2023).

The study of vessel travel times is important for decision-makers to evaluate the state of the transportation system, determine baseline measures, quantify the effects of factors that affect reliability, quantify impacts of operations or maintenance decisions, measure capacity and congestion, and provide useful information to managers, planners, traffic control, inland waterway users, and researchers (Choe et al., 2001). The outcome of these studies can be used to develop strategies, methodologies, and mathematical formulations to minimize costs and maximize the efficiencies of inland waterway infrastructures, such as the optimization of port operations and port safety (Fan et al., 2023). However, there is currently a lack of standard performance metrics for marine traffic. To address this gap, this paper proposes a methodology for evaluating the performance of waterways to understand events and/or conditions, focusing specifically on the Gulf Intercoastal Waterway (GIWW) (Figure 1). The aim is to develop a forecasting and event evaluation methodology to estimate vessel waterway travel times and identify key influencing factors or vessel performance metrics.

The forecasting and event evaluation methodology developed in this paper uses different databases from the U.S. Army Corps of Engineers (USACE) Lock Performance Monitoring System (U.S. Army Corps of Engineers, 2023) and the Local Notice to Mariners. Data also come from the Automatic Identification System (AIS) (U.S. Coast Guard Navigation Center, 2023), which is a maritime navigation safety communications system standardized by the International Telecommunication Union and adopted by the International Maritime Organization. The AIS sends data from vessels regarding vessel information (e.g., identity, type, position, course, speed, navigational status, maritime mobile service identity [MMSI], motion trajectory, and direct travel times); safety-related information from shore stations, other ships, satellites, and aircraft; and automated information from similarly fitted ships. The AIS also monitors and tracks ships, and exchanges data with shore-based facilities. Therefore, AIS records are important data related to a vessel's itinerary, such as speed, motion trajectory, and direct travel times (Fan et al., 2023).

This paper is organized as follows: Section 2 provides a review of waterway performance metrics, waterway AIS data analysis, and waterway travel time modeling studies. Section 3 describes a forecasting and event evaluation methodology to estimate vessel waterway travel times from origin-destination (O-D) locations in the mainstream of the GIWW. Section 4 describes the case study and results. Section 5 provides conclusions, limitations, and future research topics.

Figure 1. Texas GIWW map.

2. Literature review

The study of maritime traffic flow complexity touches on macroscopic and microscopic models. Macroscopic models have limitations in inland waterways, necessitating a focus on microscopic models to understand irregularity and unpredictability in ship travel time sequences. The assessment of fluidity and performance metrics in waterways has been an important subject of research in maritime transportation because these metrics provide valuable insights to help understand the dynamics and challenges associated with maritime traffic. To overcome these challenges, AIS data have played a significant role in creating decisions and simulation models to perform error analysis, ship behavior clustering, and microcosmic simulation models for ship traffic (Li et al., 2018; Xin et al., 2019; Zhao et al., 2018; Zhou et al., 2019) and for inland waterway travel time estimation.

Yang et al. (2019) published a literature review on the various uses of AIS data in the shipping industry. Although their study did not focus on travel time forecasting, their study outlined different applications, including ship behavior analysis relevant to waterway traffic. Their paper highlights the use and utility of AIS data to determine traffic flows in restricted waterways, providing a valuable resource for researchers seeking to understand waterway dynamics and traffic flow complexity.

In the literature, different models have been developed to evaluate traffic complexity (Wen et al., 2015), analyze ship traffic demand (Zhang et al., 2017), estimate speed–density relations (Kang et al., 2018), and calculate geometric probabilities of ship collision (Kujala et al., 2009). Near-miss risk assessment in ice-covered waters has also been explored (Zhang et al., 2016; 2018; 2019), and correlations between ship accidents and traffic conditions have been investigated (Bye & Aalberg, 2018; Mazaheri et al., 2015). All these models provide valuable information about the factors that influence vessel waterway travel times.

Researchers have studied waterway fluidity. Mitchell et al. (2019) conducted an analysis using 2017 AIS data to evaluate waterway fluidity for the Ohio River, the Mississippi River, and the GIWW. Mitchell et al. studied vessel travel times and added reliability as a key metric. The authors performed summary statistics monthly, calculating vessel trips and speeds in the 25th percentile by filtering data with a week upper bound. The comparison of down and upper bound transport speeds revealed significant differences in the Mississippi River, where down bound currents were 50 percent faster. This study emphasized the impact of current flow on travel speeds, providing a foundation for understanding waterway dynamics. Wu et al. (2020) expanded on this work by developing a model using AIS data to estimate travel time and density in the Houston Ship Channel. The authors focused on determining vessel travel times in inland channels, emphasizing the identification of distances and time stamps based on the AIS data. Their methodology identifies destination docks, arrival times, and departure times as main factors to calculate total travel times. The study found that travel times between dock lines and the Beltway 8 Bridge followed a lognormal distribution, while travel times between the destination and the bridge exhibited a normal distribution. Their results indicate that larger ships and tanker ships exhibit slower travel times and showed a relation between congestion in adjacent areas. Although this study did not predict future travel times, it did develop a methodology to estimate travel speeds, which could improve estimation of future projections.

Other researchers have studied waterway travel times. Asamer and Prandtstetter (2015) applied a nonlinear modeling approach to estimate ship travel times. They considered AIS data. Their model differentiates travel times between locks and the actual section between them since both are affected by different variables. A support vector machine measures the travel time when passing a lock and sections, and a linear regression is developed to better show the relationship between predicted speeds and observed ones. Their proposed regression considers ship properties and weather conditions because their results show that both influence travel times. Like in this paper, their linear regression predicts travel times at different sections of the waterway. Alessandrini et al. (2019) and Park et al. (2021) researched travel time estimation algorithms to estimate travel times, dividing an estimated path length by an estimated speed. In one instance, Alessandrini et al. (2019) proposed a path selection approach using a grid structure network considering a Dijkstra strategy to find the best navigation grid based on direction and density features. In another instance, Park et al. (2021) employed reinforcement learning to predict paths and estimate average speed. The main disadvantage of their methods is relying on the linear relationship between distance and speed because it invalidates assumptions for long paths with increased uncertainties. Finally, DiJoseph et al. (2019) developed a methodology to estimate O-D waterway travel times from vessel transit data. Their methodology needs historic AIS vessel transit data. The methodology does not require transit data from all vessels in the whole waterway between the origin and destination and works with a sample of the vessel population, considering that vessels might navigate a small section of the waterway and not necessarily the total from origin to destination.

Lately, researchers have been applying machine learning (ML) methods to estimate ship waterway travel times because these methods allow recognition of the nonlinear relationship between multiple factors and waterway travel times. Yu et al. (2018) used an ML method to map vessel arrival times with navigation day, month, route type, and vessel length. Xu et al. (2022) applied a clustering algorithm to group motion patterns and employed support vector regression models for each pattern, considering factors like latitude, longitude, speed, course, navigation status, and remaining distance.

Unlike road traffic literature, which extensively considers congestion as a major factor that influences travel times, waterway travel time studies often lack this factor. Few studies explore the impact of traffic congestion on waterway travel times. Sui et al. (2020) and Fan et al. (2023) considered congestion as a factor to estimate waterway travel times. Sui et al. (2020) studied congestion states and interactions between vessels using complex network theory and network dynamics. Conversely, Fan et al. (2023) considered the impact of traffic congestion on vessel travel times. They developed a complex vessel interaction network to capture vessel-to-vessel interactions as input data to feed a convolutional neural network, which is a deep learning model, to get spatial trajectory information.

Finally, factors that ensure maritime safety affect waterway travel times. Tirunagari et al. (2012) identified the factors that cause maritime accidents based on 135 papers produced by the Marine Accident Investigation Branch in the United Kingdom. These factors are traffic density, ship speed, confusion, equipment, severe weather, fatigue, and health.

In conclusion, the reviewed papers highlight the complexity of estimating inland waterway travel times. AIS data have a critical role in the development of decision-making and simulation models and methodologies to evaluate traffic complexity, analyze maritime traffic, estimate speed–density relations, calculate the probability of ship collisions, evaluate the effect of congestion in waterways, and estimate travel times. The literature review contributes significantly to the understanding of maritime traffic dynamics and indicates the factors for maritime safety and congestion that are relevant to this paper since some of these factors affect waterway travel times. Finally, the literature review highlights the scarcity of studies tackling the problem of estimating inland waterway travel times. Only a few papers address this problem but do not focus on estimating travel times in the GIWW. Therefore, the aims of this paper provide valuable insights into the inland waterway transportation system and for regional and U.S. national economies.

3. Methodology

This paper proposes a new methodology to estimate vessel waterway travel times from origins to destinations for the main channel of the GIWW, which is the focus of this paper. The designated O-Ds for the GIWW are port facilities, the Louisiana state boundary, the western terminus at the Port of Brownsville, and intersections with ship channels or tributaries since vessels enter and exit the GIWW there.

Specifically, the methodology has four activities (Figure 2) that include several methods and corresponding steps.

3.1. Activity 1: Data collection and cleansing

The data collection activity gathers AIS data from vessels: speed, motion trajectory, and direct travel times. This activity builds on prior work undertaken by USACE (U.S. Army Corps of Engineers, 2023), but in this paper, the data acquisition activity developed by USACE is modified to account for differences in the GIWW from the rest of the inland waterway system. The main changes are:

The GIWW crosses several deep-draft ship channels.

This activity is intense in freight clusters along the waterway.

When collecting AIS data, many vessels are included that are not inland towing vessels. Therefore, the data are filtered to include only inland towing vessels. Therefore, the following vessel types are required to have an AIS transponder: a self-propelled vessel of 65 feet or more in length engaged in commercial service, a towing vessel of 26 feet or more in length and more than 600 horsepower engaged in commercial service, a self-propelled vessel certificated to carry more than 150 passengers, and a self-propelled vessel engaged in dredging operations in or near a commercial channel or shipping fairway. The U.S. Coast Guard provides details on its website (Navcen.uscg.gov) and AIS page on how AIS transmitters and receivers operate (U.S. Coast Guard Navigation Center, 2023).

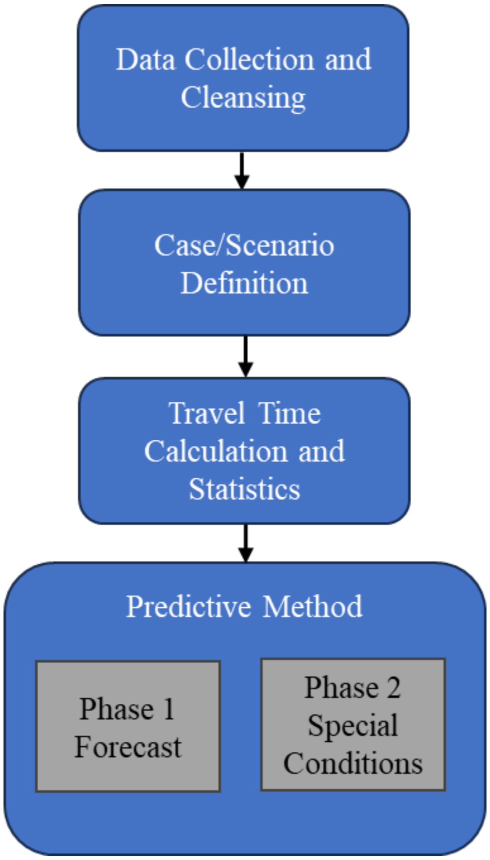

Figure 2. Methodology activities.

AIS data are not always complete and continuous for a given vessel. This is due to atmospheric conditions, physical obstructions, or equipment malfunctions. The proposed methodology developed in this study overcomes this problem by allowing for the fact that transit data are not available for the whole of the waterway between O-D pairs. Relying on a robust sample of the population means the methodology does not require transit data to be available for the entire population of vessels on the waterway. The methodology also considers that each vessel may not transit the entire distance between O-D pairs, instead making shorter transits.

In this paper, for the main channel of the GIWW, the acquired AIS data covers the whole years of 2018 and 2019. The USACE Engineering Research and Development Center provided 93 million raw data points, which were filtered for inland towing vessels along the GIWW geofences between Beaumont and Brownsville, Texas. The sampling interval is 5 minutes.

The cleansing data activity involves the following steps:

Step 1: Remove all records where the stated ship type is not an inland towing vessel.

Step 2: Remove all remaining records where the MMSI is not a valid U.S. MMSI.

Step 3: Remove all remaining records where the MMSI is clearly invalid (not enough digits).

Step 4: Examine all remaining records using multiple public sources to determine the vessel type and remove those that are not inland towing vessels. For the main channel of the GIWW, the main public sources are the U.S. Coast Guard Port Information Exchange, the Federal Communications Commission wireless license search, the USACE Waterborne Transportation Lines of the United States, the National Oceanic and Atmospheric Administration vessel search, the USACE listing of tugs and towboats, and a general Google search.

Step 5: Remove all records that do not have information about the vessel type.

Step 6: Use the remaining records as the inland towing dataset.

3.2. Activity 2: Case/scenario definition

The second main activity includes:

Step 1: Determine O-D points and routes. The entrance to each deep-sea port is designated as a destination.

Step 2: Segment the inland waterway into links with lengths that vary considerably because of specific features along the coast, intersections with ship channels, and floodgates/locks including mooring areas. However, the lengths of links should be as identical as possible to prevent one location from unduly influencing the analysis of traffic behavior. Because floodgates and locks can be contained within a single link for their respective rivers, these links isolate the effects of the structures on vessel behavior and thus travel times. These travel-time-related behaviors include deceleration time approaching the lock, queuing time to enter the lock, passage time through the lock, and acceleration time away from the lock.

3.3. Activity 3: Travel time calculation and statistics

The third main activity includes:

Step 1: Divide the waterway between the O-D pairs into shorter, consecutive sections, called links, so that each link has homogeneous vessel travel behavior.

Step 2: Estimate the travel time for each link by calculating the vessel's movements from one boundary of a link to the other.

Step 3: Identify and remove travel time outliers defined as travel times that exceed the following cutoffs:

- For links containing floodgates/locks, the cutoff is 48 hours (2 days).
- For the link containing the Port of Houston, the cutoff is 12 hours.
- For links without floodgates/locks, the cutoff is the amount of time it takes to travel the entire link at 2.6 knots (3 mph).

Step 4. Calculate link travel time performance measures such as average travel times and standard deviations.

Step 5. Calculate the O-D travel time performance measures from the link travel time performance measure results.

3.4. Activity 4: Predictive method

This paper proposes a two-phase predictive method that delivers travel time forecast projections and assesses how special conditions affect travel times. Phase 1 focuses on the forecast projections, and Phase 2 focuses on evaluating the impacts of special conditions on travel times.

Phase 1: Forecasting method

Phase 1 includes:

Step 1: Data preliminaries: The goal is to look at assessing the data for features needed to identify the proper forecast model. These features are data visualization to determine stationarity (cyclicity and seasonality); application of the Dickey-Fuller test to determine stationarity; and application of the Box-Pierce's Q statistic tests, the Schwarz's Bayesian info criteria, the Akaike's information criteria, and the Hannan and Quin information criterion procedures to evaluate correlation and partial autocorrelation.

Step 2: Lag selection: Analyze past data for forecasting and capturing dynamic effects. The selection of lags identifies which of the previous periods are included as moving average, integrated, and autoregressive components. This is fundamental in autoregressive moving average models since too many lags could increase the error in the forecast whereas too few lags could leave out relevant information. Autocorrelation identifies the moving average order, and partial autocorrelation helps in determining the autoregressive order. The integration component, which denotes taking differences from specific lags, is only necessary when the time series show evidence of being non-stationary.

Step 3: Forecasts: Separate the data sample into three data groups (training data, ex-post data, and ex-ante data), and select the most appropriate forecasting model according to the analysis of the data series features. Training data or an estimation sample is a subsample of the data used as input for the forecasting model; ex-post data are a subsample of the data used as a benchmark for model comparison; and ex-ante data are a projection calculated with the selected forecasting model.

Phase 2: Special conditions

Phase 2 includes:

Step 1: Gather information about the following special conditions if available:

- Major weather threats such as hurricanes that halt navigation and, consequently, cause trip data, such as AIS data, to be absent.
- Dredging: The presence of dredging equipment in or adjacent to the channel.
- Shoaling: A reduction in available draft large enough to warrant a notice to mariners.
- Bridge closure: The closure of the channel due to bridge construction activity.
- Submerged vessel: The presence of a submerged vessel in the vicinity of the channel.
- Construction: The placement of riprap along the banks.
- Lock closure: The closure of a floodgate or lock to all traffic for a defined period.
- Submerged pipe: The presence of a submerged pipe in or near the channel.
- Bridge clearance: A reduction in the clearance in terms of width and/or height at a bridge crossing.
- Submerged obstruction: The presence of an unidentified submerged object in or near the channel.
- Regatta: A recreational event requiring the use of the channel.
- The number of complete trips through the link (part of Step 3).

Step 2: The information gathered in Step 1 as special conditions (all but the number of complete trips through the link) must be coded binary (i.e., 1 denoting occurrence and 0 denoting no presence of a corresponding event or condition) and assigned to each link and time (i.e., week) to match the travel time data used in Phase 1.

Step 3: Include the number of complete trips through the link in the database to assess their relevance and impact on travel times.

Step 4: Check the linearity between the sample count and travel time to determine the general type of models to use for assessing the effects of the number of trips on transit time.

Step 5: Check the collinearity/independence because different special conditions may exist at the same time in the same link. Collinearity assessment evaluates the redundancy of special conditions. If redundancy in the form of collinearity is found, then special conditions must be evaluated in different models—not combined in a single model—to estimate their effects more accurately.

Step 6: Check the homogeneity and normality. Homogeneity focuses on the variability through different values of the explanatory variable. Normality validates whether the shape of the distribution is normal or not to apply the proper model.

Step 7: Run linear or nonlinear models when the relation between sample counts and travel times is not linear, such as exponential trend, logarithmic, power curve, reciprocal, log reciprocal, modified exponential, Gompertz, and logistic.

Step 8: Select the best model comparing the mean square error, the mean absolute error, and the mean average percentage error statistical measures.

4. Travel time estimation results

The GIWW begins at the Louisiana border and ends at the Brazos Island Harbor Ship Channel near Brownsville, Texas, with a length of 379 miles. The GIWW links 11 deep-draft ports (25 feet or deeper) and 13 shallow-draft channels (Texas Department of Transportation, 2014). The deep-draft ports manage both shallow- and deep-draft vessels, so the two systems are intertwined. Figure 1 provides a map of the Texas GIWW.

Activity 1: Data collection and cleansing

The data gathered for this paper are archived by various public and private entities as indicated in the methodology in Activity 1 (collection) and Step 4 of cleansing.

Activity 2: Case/scenario definition

The entrance to each deep-sea port is designated as a destination along the Texas GIWW. Additionally, the petrochemical complex at Chocolate Bayou just west of Houston/Galveston, Texas, the entrance to the channel leading to the Port of Victoria, and the entrance to the Arroyo Colorado, which leads to the Port of Harlingen, are designated as O-Ds. Because the ship channels for Houston, Galveston, and Texas City, Texas, all intersect the GIWW near each other, these three ports are considered one O-D. Finally, the Louisiana–Texas border at Mile Marker 262 is designated as an O-D since it marks the eastern terminus of the Texas GIWW, while the Port of Brownsville is the western terminus of the GIWW. Table 1 presents the O-D points used for this study. Figure 3 shows the locations of port and GIWW intersections.

The Texas GIWW inland waterway is segmented into links first defined for ship channel intersections and for the location of the Brazos River Floodgates and the Colorado River Locks. The reaches between these links are defined as additional links. The lengths of these links are primarily determined by geographical features with homogeneous operational characteristics. The longest link (50 miles) is between Port Arthur and the Bolivar Peninsula. The next longest links (44 miles and 49 miles) are in the Laguna Madre south of Corpus Christi. These three links are in areas with no development.

The links for the floodgates/locks include the mooring areas on either side of each river where two vessels wait for the opportunity to cross the river. The lengths of these links are made identical. Figure 4, Figure 5, and Figure 6 depict the locations of all the links.

Figure 5 shows that link 22 splits into two sub-links: main and alternative routes, both of which converge at the Corpus Christi Ship Channel. The Lydia Ann Channel sub-link (Link 22B) is the primary route in this analysis.

Table 1. Origin-destination points.

Origin-destination	Eastern/ northern mile marker	Western/ southern mile marker
Port of Beaumont/ Port of Port Arthur	275.5	288
Ports of Houston/ Galveston/Texas City	348.5	352
Chocolate Bayou	374.5	392
Port Freeport	392	398
Calhoun Port Authority	456	473
Port of Victoria	487	493
Port of Corpus Christi	539	549
Port of Harlingen	642	645
Port of Port Isabel	665	668
Port of Brownsville	677	682

Figure 3. Locations of Port and GIWW intersections.

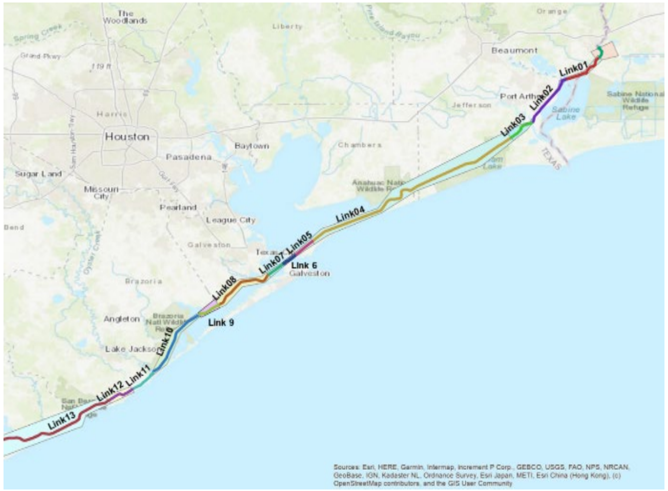

Figure 4. GIWW links 1–13.

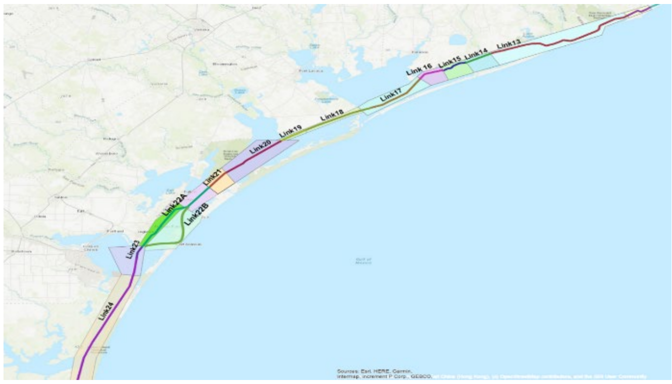

Figure 5. GIWW links 14–24 (including 22A and 22B).

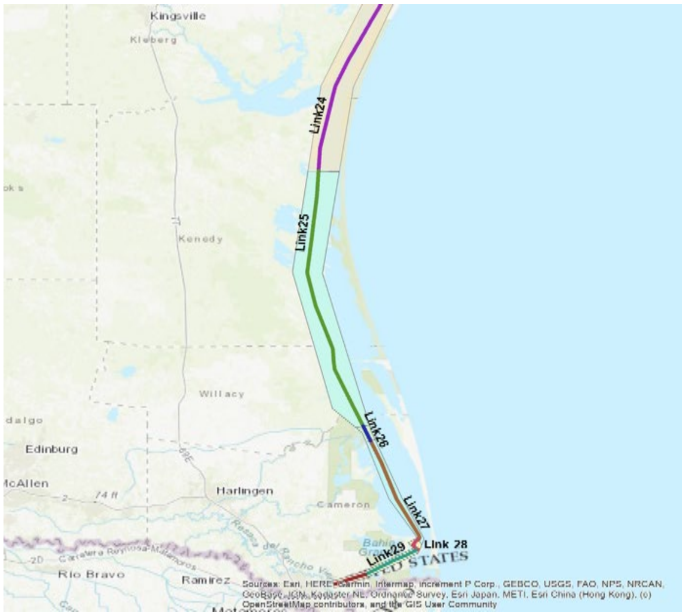

Figure 6. GIWW links 25–29.

Activity 3: Travel time calculation and statistics

Table 2 shows the average transit time (ATT), standard deviation (SD), and coefficient of variation (CV) by direction for 2018 and 2019. Conditional color formatting by variable in each column highlights higher values in red and lower values in green. In Table 2, the four groups of direction and year combinations (i.e., westbound–2018, westbound–2019, eastbound–2018, and eastbound–2019) have similar color patterns, indicating there is no noticeable change in the performance measures.

The travel times are consistent across both years and directions with low variability for most links. This is not surprising given that the GIWW has no river current and is not subject to flooding or drought conditions. Seasonality is also not a factor.

Activity 4: Predictive method

Phase 1: Forecasting method

The input data consist of 104 data points representing 2 years (2018 and 2019) of weekly travel time averages for 31 different links. The GIWW splits in the vicinity of Port Aransas, with the two branches coming together at the Corpus Christi Ship Channel. The branch known as the Lydia Ann Channel is the more heavily transited of the two. The two branches are links 22A and 22B and are evaluated independently. No other data are available that could help explain travel time behavior, and therefore there are no explanatory or dependent variables. Because of this, the dataset is formed of a single series per link.

Two forecasting methods are selected after applying Activity 4 Phase 1 of the proposed methodology: The autoregressive integrated moving average (ARIMA) and exponential smoothing (Hyndman & Athanasopoulos, 2021; Sadeghi Gargari et al., 2019). These two methods are selected to estimate tow vessel waterway travel times between O-Ds for the main channel of the GIWW.

Table 2. GIWW link average and standard deviation of travel time, 2018 and 2019.

Link	Westbound/southbound trips						Eastbound/northbound trips					
	2018			2019			2018			2019		
	ATT (Hour)	SD (Hour)	CV	ATT (Hour)	SD (Hour)	CV	ATT (Hour)	SD (Hour)	CV	ATT (Hour)	SD (Hour)	CV
1	1.39	0.37	0.26	1.34	0.33	0.25	1.45	0.31	0.22	1.46	0.32	0.22
2	1.92	0.55	0.29	1.88	0.53	0.28	2.01	0.55	0.27	1.99	0.54	0.27
3	0.99	0.23	0.23	0.97	0.23	0.23	1.03	0.25	0.24	1.01	0.25	0.24
4	9.22	2.09	0.23	9.05	2.11	0.23	9.32	2.06	0.22	9.17	2.03	0.22
5	0.89	0.27	0.30	0.89	0.27	0.30	0.83	0.21	0.25	0.83	0.21	0.25
6	0.61	0.69	1.13	0.58	0.53	0.93	0.77	0.95	1.23	0.80	1.06	1.32
7	0.62	0.13	0.21	0.61	0.14	0.22	0.60	0.11	0.19	0.60	0.12	0.20
8	2.26	0.50	0.22	2.25	0.52	0.23	2.29	0.51	0.22	2.30	0.50	0.22
9	0.74	0.17	0.23	0.74	0.18	0.24	0.75	0.17	0.22	0.75	0.16	0.21
10	2.90	0.64	0.22	2.88	0.68	0.24	3.02	0.66	0.22	3.03	0.68	0.22
11	1.05	0.24	0.23	1.08	0.27	0.25	1.07	0.23	0.22	1.10	0.25	0.23
12	4.82	5.07	1.05	6.52	6.81	1.04	5.40	5.68	1.05	7.65	7.86	1.03
13	6.06	1.38	0.23	6.14	1.50	0.24	6.45	1.33	0.21	6.57	1.47	0.22
14	5.26	6.62	1.26	6.55	7.34	1.12	4.66	5.56	1.19	5.82	6.08	1.04
15	1.01	0.21	0.21	1.01	0.21	0.21	1.09	0.21	0.20	1.09	0.23	0.21
16	1.08	0.28	0.26	1.09	0.28	0.26	1.08	0.21	0.20	1.09	0.22	0.20
17	2.58	0.42	0.16	2.58	0.44	0.17	2.65	0.43	0.16	2.63	0.43	0.16
18	2.47	0.54	0.22	2.45	0.54	0.22	2.65	0.71	0.27	2.67	0.71	0.27
19	1.10	0.26	0.24	1.08	0.26	0.24	1.05	0.21	0.20	1.02	0.20	0.20
20	2.60	0.44	0.17	2.60	0.46	0.18	2.84	0.62	0.22	2.88	0.66	0.23
21	0.99	0.24	0.24	1.00	0.26	0.26	1.02	0.21	0.20	1.02	0.21	0.21
22	4.00	0.70	0.17	4.07	0.80	0.20	4.08	0.72	0.18	4.09	0.74	0.18
22A	4.32	0.79	0.18	4.43	0.86	0.19	4.35	0.78	0.18	4.33	0.72	0.17
22B	3.96	0.67	0.17	4.02	0.78	0.20	4.04	0.70	0.17	4.06	0.74	0.18
23	2.02	0.39	0.19	2.05	0.38	0.19	1.90	0.43	0.23	1.89	0.41	0.22
24	8.32	1.44	0.17	8.05	1.33	0.17	7.40	1.58	0.21	7.51	2.05	0.27
25	9.27	1.54	0.17	8.98	1.53	0.17	8.11	1.80	0.22	8.15	1.95	0.24
26	0.56	0.10	0.18	0.54	0.10	0.19	0.48	0.09	0.19	0.47	0.09	0.19
27	3.61	0.51	0.14	3.47	0.51	0.15	3.18	0.48	0.15	3.14	0.51	0.16
28	0.45	0.13	0.30	0.42	0.12	0.27	0.39	0.13	0.33	0.38	0.12	0.30
29	1.25	0.25	0.20	1.32	0.25	0.19	1.21	0.18	0.15	1.19	0.19	0.16

One hundred observations are used as training data, and four observations are retained to evaluate forecasts for the ARIMA models. In the case of exponential smoothing, because there is no need to select the best model lags, the whole sample is used for forecasting. The mean square error, the mean absolute error, and the mean average percentage error indicate that ARIMA is the best model to estimate vessel travel times in the main channel of the GIWW. Although 52 weeks are projected into the future using the selected model, the analysis of the results recommend considering only 12 weeks as forecasts because ARIMA tends to converge to single values the farther into the future the forecast goes.

Table 3 shows the forecasted values in hours for the first 12 weeks (i.e., weeks 105–116) per link with total travel time (TTT) at the bottom (considering traveling throughout all links). O-D pairs and links are not correspondent one to one. Therefore, some O-D pairs cover multiple links, which explains the lack of cell borders in the O-D pair column of the table.

Figure 7 shows the behavior of travel times for the last 12 weeks of data and the forecasted weeks for all links, with total travel time at the top (red dotted line). The black dotted vertical line separates actual data (the last 12 weeks on the left) from forecasted values.

Table 3 and Figure 7 show that the forecasted values behave more smoothly (i.e., have less variably) than actual travel times. This is expected because there is only a single data time series available, and thus the necessary use of ARIMA and exponential smoothing yields projections that tend to converge to specific values overall. Some of the variability in the actual values may be explained by special conditions.

Phase 2: Special conditions

The data obtained to run this phase of the proposed methodology have been gathered as indicated in the proposed methodology (Activity 4, Phase 2: Special conditions). In this study, 361 occurrences are considered in the analysis.

The statistical analysis suggests that sample counts (i.e., completed trips through a link) do not influence trave l time significantly in any link. These results could be because the period used as the unit of measurement is far larger than the average travel time for each link. In other words, one would expect that the amount of traffic in each link would affect travel time in the same link. This is analogous to the effect of road congestion in road travel time (Yeon et al., 2008). However, to capture this effect, the time unit of measurement should be at an adequate resolution to the actual travel time. For instance, if travel time is a few hours but the time unit of measurement is a week—as in this case—then the variations in travel time counted in hours due to the number of vessels traveling in the same link would dilute and/or net out when averaged over a week. For a model to capture such effects, the unit of measurement should be hours or days since travel time averages for a single link are between 0.41 hours (i.e., 24.6 minutes) and 9.25 hours for links 28 and 4. Unfortunately, data do not allow the necessary resolution to adequately capture effects due to changes in the number of trips (sample counts) on travel time. (If the number of vessels per hour are counted, even the highest number is two vessels per hour [at link 5].)

Nevertheless, the effects of special conditions on travel time are obtained with the application of the proposed methodology. Specifically, the statistical analysis deemed two special conditions relevant for travel time: dredging and shoaling. The effects yielded by the analysis are in the form of magnitude and probability of occurrence. On average, when dredging is present in a link, there is a probability of 78.9 percent that travel time for that link increases by 0.38 hours (i.e., 23 minutes). Similarly, when shoaling is present on any link, the probability is 64.7 percent that travel time for that link increases by 0.35 hours (i.e., 21 minutes). These probabilities provide an idea of the general effects of these two special conditions; however, specific effects vary from link to link. Figure 8 and Figure 9 show the individual effects per link. The blue columns represent the magnitude of the effect in hours, and the orange dots represent the probability of that effect occurring. Several links do not show effects mostly due to the lack of special conditions on those links.

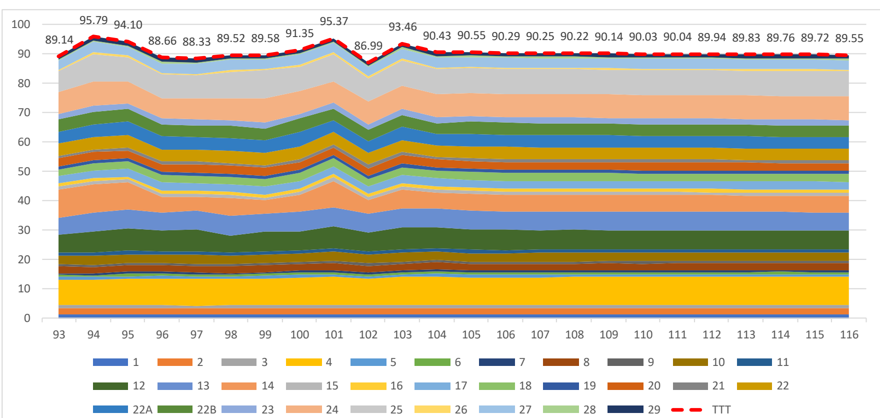

Figure 7. Travel times for last 12 weeks of data and forecasted ravel times.

Table 3. Twelve-week forecasts per link (hours).

Link	O-D pair	105	106	107	108	109	110	111	112	113	114	115	116
1	Louisiana border to Port Beaumont/Port Arthur upstream boundary	1.43	1.42	1.42	1.42	1.42	1.42	1.42	1.42	1.42	1.42	1.42	1.42
2	Port Beaumont/Port Arthur upstream boundary to Port Arthur downstream boundary	1.96	1.95	1.93	1.96	1.98	1.93	1.96	1.94	1.94	1.94	1.94	1.96
3	Port Arthur downstream boundary to Port Houston/Pelican Island mooring	1.04	1.01	1.00	1.02	1.01	1.03	1.00	1.01	1.01	1.02	1.02	0.99
4		9.52	9.55	9.58	9.60	9.62	9.64	9.65	9.66	9.66	9.65	9.65	9.63
5		0.87	0.87	0.87	0.87	0.87	0.87	0.86	0.86	0.86	0.86	0.86	0.86
6	Port of Houston/ Galveston/Texas City	0.73	0.76	0.78	0.80	0.82	0.83	0.84	0.85	0.85	0.86	0.86	0.87
7	Port Houston/Pelican Island mooring to Chocolate Bayou	0.62	0.62	0.62	0.62	0.62	0.61	0.61	0.61	0.61	0.61	0.61	0.61
8		2.27	2.27	2.28	2.28	2.28	2.27	2.27	2.27	2.28	2.28	2.27	2.27
9		0.75	0.75	0.75	0.74	0.74	0.74	0.74	0.74	0.74	0.74	0.74	0.74
10	Chocolate Bayou to Port Freeport upstream boundary	2.97	2.97	2.96	2.96	2.96	2.96	2.96	2.96	2.96	2.96	2.96	2.96
11	Port Freeport upstream boundary to Port Freeport downstream boundary	1.11	1.09	1.10	1.11	1.09	1.09	1.10	1.08	1.08	1.10	1.08	1.08
12	Port Freeport downstream boundary to Colorado River	6.92	6.82	6.74	6.66	6.59	6.54	6.49	6.44	6.40	6.37	6.33	6.31
13		6.41	6.32	6.32	6.32	6.32	6.32	6.32	6.32	6.32	6.32	6.32	6.32
14	Colorado River industry	5.79	5.69	5.67	5.67	5.67	5.67	5.67	5.67	5.67	5.67	5.67	5.67
15	Colorado River to Calhoun	1.06	1.06	1.06	1.05	1.05	1.05	1.05	1.05	1.05	1.05	1.05	1.05
16		1.09	1.09	1.09	1.09	1.09	1.09	1.09	1.09	1.09	1.09	1.09	1.09
17	Port Lavaca (Calhoun Port Authority)	2.63	2.60	2.62	2.61	2.61	2.61	2.61	2.61	2.61	2.61	2.61	2.61
18	Calhoun to Victoria	2.61	2.60	2.60	2.60	2.60	2.59	2.59	2.59	2.59	2.58	2.58	2.58
19	Port of Victoria	1.06	1.09	1.05	1.07	1.08	1.05	1.07	1.07	1.05	1.07	1.06	1.06
20	Victoria to Corpus Christi upstream boundary	2.67	2.71	2.77	2.71	2.69	2.69	2.75	2.73	2.71	2.70	2.76	2.69
21		1.01	1.01	1.01	1.01	1.01	1.01	1.01	1.01	1.01	1.01	1.01	1.01
22		4.08	4.06	4.04	4.03	4.01	4.00	3.99	3.98	3.97	3.96	3.96	3.95
22A	Aransas Pass	4.28	4.23	4.17	4.13	4.09	4.05	4.02	3.99	3.96	3.94	3.92	3.90
22B	Lydia Ann Channel	4.04	4.03	4.02	4.01	4.00	3.99	3.98	3.97	3.97	3.96	3.96	3.95
23	Corpus Christi upstream boundary to Corpus Christi downstream boundary	1.95	1.98	2.00	2.02	2.03	2.03	2.02	2.01	1.99	1.98	1.96	1.96
24	Corpus Christi downstream boundary to Arroyo Colorado upstream boundary	7.74	7.80	7.84	7.88	7.91	7.94	7.95	7.97	7.98	7.99	8.00	8.01
25		8.51	8.51	8.51	8.51	8.51	8.51	8.51	8.51	8.50	8.50	8.50	8.50
26	Arroyo Colorado upstream boundary to downstream boundary (Port of Harlingen)	0.53	0.53	0.50	0.51	0.51	0.52	0.51	0.52	0.53	0.52	0.52	0.51
27	Arroyo Colorado to Port Isabel upstream boundary	3.32	3.32	3.33	3.33	3.33	3.34	3.34	3.34	3.35	3.35	3.35	3.35
28	Port Isabel upstream boundary to Port Isabel downstream boundary	0.41	0.43	0.44	0.44	0.45	0.45	0.45	0.46	0.46	0.46	0.46	0.46
29	Port Isabel downstream boundary to Port Brownsville upstream boundary	1.19	1.19	1.19	1.19	1.20	1.20	1.20	1.20	1.20	1.20	1.20	1.20
TTT		90.55	90.29	90.25	90.22	90.14	90.03	90.04	89.94	89.83	89.76	89.72	89.55

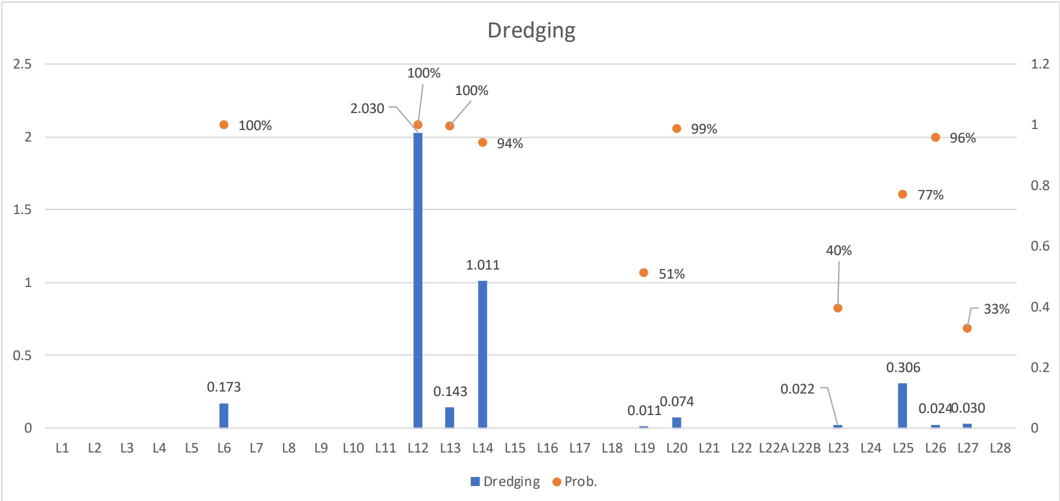

Figure 8. Individual effects (magnitude in hours and probability of occurrence as percentage) per link—dredging.

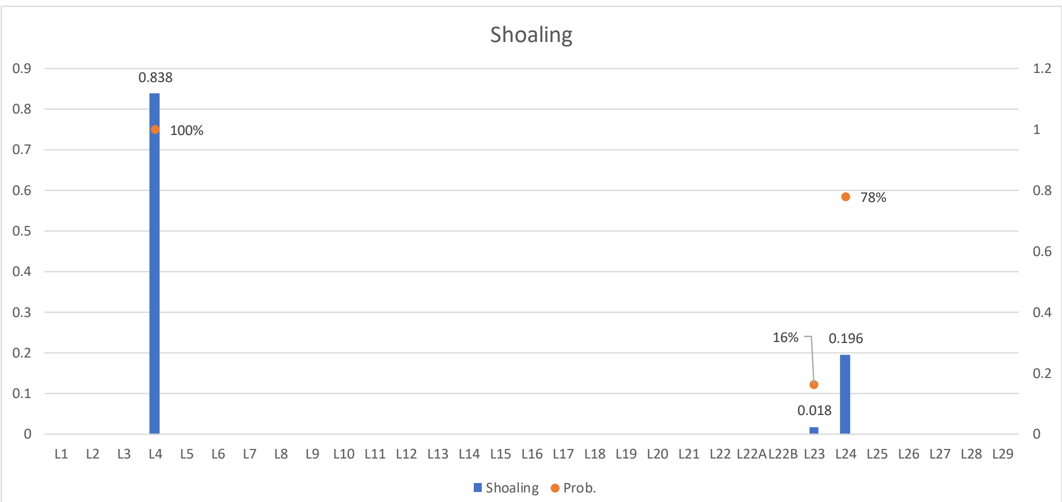

Figure 9. Individual effects (magnitude in hours and probability of occurrence as percentage) per link— shoaling.

5. Conclusions

AIS data provide a robust sample for the calculation of performance measures. Much of the effort involved in using AIS data focuses on cleaning the data and reducing it to the inland towing traffic that uses the GIWW.

The GIWW is a complex waterway. Because it crosses 11 ship channels and connects to numerous shallow-draft channels, the waterway must be segmented into small links to evaluate performance. Future research may want to investigate the interactions with these connections.

Since there is no current on the GIWW (as there would be on a river), no significant variations in travel times by direction were expected, and this turned out to be the case. The links containing the Brazos River Floodgates and the Colorado River Locks showed a high degree of variability in travel times. This could be because of congestion at the structures or a hesitancy to cross the rivers when conditions are suboptimal.

Ideally, a path-based approach is preferable to a link-based approach for calculating travel times between O-D pairs, but some practical considerations such as sample size led the authors to opt for the use of link-based travel times. A path approach requires a separate sample to be developed and maintained for each O-D pair, and each pair must be evaluated independently. Given the complexity of the GIWW, this was not a viable approach for this research.

The predictive analysis was divided into two phases. The first phase looked at developing the forecast projections for travel time based on historical data. The second phase focused on assessing the effects of special conditions on travel time. The authors also explored the relation of the number of trips (i.e., sample count) with travel time in this last phase of the analysis.

Results from Phase 1 show that forecasted projections smooth out the deeper they go into the future. This is because, due to data availability, projections were based exclusively on the historical data of the same target variable; or, in other words, projections used historical information—past values—of travel time to predict the future. This type of data determined the type of statistical tools the authors implemented and resulted in the selection of a smoothing type of model. Therefore, the authors recommend the use of only the first 12 weeks of projections.

Phase 2 analysis found no significant relation between sample count and travel time. This is likely because of an imbalance in magnitude between the time used as a unit of measurement (week) and the average travel time for each link (hours). Analogous to the effect of road congestion in road travel time, to capture such effects, the magnitude of the unit of measurement should be in accordance with average travel time; more specifically, for this type of research, the unit of measurement should be in hours or days since travel time averages for a single link are between 0.41 hours (i.e., 24.6 minutes) and 9.25 hours. Unfortunately, the data did not allow the necessary resolution to adequately capture said effects.

The authors found two special conditions relevant for travel time: Dredging and shoaling. These effects were in the form of magnitude of impact and the probability of that impact occurring when the special condition is present. On average, when dredging is present, there is a probability of 78.9 percent that travel time increases by 0.38 hours (i.e., 23 minutes). Similarly, when shoaling is present on any link, there is a probability of 64.7 percent that travel time increases by 0.35 hours (i.e., 21 minutes). These probabilities provide an idea of the general effects of these two special conditions; however, specific effects vary from link to link, as shown in Figure 8 and Figure 9.

The methodology developed provides quantitative results that predict, describe, and validate future travel time behaviors based on specific factors. It illustrates the need to clean AIS data before they are used in any analysis. In the case of the Texas GIWW, it is also important to filter out vessels that have been recorded in the GIWW but are not using the GIWW. Users of the GIWW can use statistics such as those provided by this study to have a sense of estimated travel time and potential effects of special conditions in a link they may need to traverse. The methodology also enables an analysis of the effect of special conditions, such as those announced in the U.S. Coast Guard's notice to mariners or USACE's local notices to mariners. If this type of study is performed regularly, it would highlight significant changes in links and allow analysts to focus on trouble spots along the waterway. Such data will also aid in planning the timing and magnitude of maintenance activities on the GIWW.

The limitations of this work lie in the data available for the analysis. Therefore, future research should focus on obtaining additional data that enable more robust projections by not relying on a single time series. Also, a higher resolution should be considered to balance the time unit of measurement with average travel time, and to be able to assess traffic effects on travel time. In addition, expanding the analysis to include additional years (e.g., 2020) could provide more data to increase the sample size and make more accurate predictions and assessments that could be tested comparing forecasted projections with actual data; for instance, obtaining data for years 2018–2020, a researcher could forecast 1 year of behavior using 2018 and 2019 data and then compare it with 2020. This could be informative for the special conditions considered in this analysis, but also useful to investigate the effects of additional conditions such as pandemic impacts.

Conflict of interest

The authors have no conflict of interest to declare.

Acknowledgments

The authors acknowledge the invaluable assistance of Patricia DiJoseph of USACE in acquiring data and providing background on previous USACE research.

Funding

This work was funded by the Maritime Transportation Research and Education Center, a U.S. Department of Transportation Tier 1 University Transportation Center.

References

Alessandrini, A., Mazzarella, F., & Vespe, M. (2019). Estimated time of arrival using historical vessel tracking data. *IEEE Transactions on Intelligent Transportation Systems, 20*, 7–15. https://doi.org/10.1109/TITS.2017.2789279

Asamer, J., & Prandtstetter, M. (2015). Improving the Estimation of Ship Travel Times on Inland Waterways. In 94rd Annual Meeting of the Transportation Research Board.

Brum, C. P., de Faria Corrêa, R. G., de J. Pacheco, D. A., Nepomuceno, T. C. C., de Oliveira, A. L. R., & Marsola, K. B. (2023). Efficiency analysis of inland waterway locks in maritime transportation systems: practical, economic and policy implications. *Maritime Policy & Management*, 1-24. https://doi.org/10.1080/03088839.2023.2252431

Bye, R. J., & Aalberg, A. L. (2018). Maritime navigation accidents and risk indicators: An exploratory statistical analysis using AIS data and accident reports. *Reliability Engineering and System Safety, 176*, 174–186. https://doi.org/10.1016/j.ress.2018.03.033

Chen, Z., & Li, X. (2021). Economic impact of transportation infrastructure investment under the Belt and Road Initiative. *Asia Europe Journal, 19*(1), 131–159. https://doi.org/10.1007/s10308-021-00617-3

Choe, T., Skabardonis, A., & Varaiya, P. (2001). Freeway Performance Measurement System (PeMS): An operational analysis tool. In Transportation Research Board 81st Annual Meeting. National Research Council.

de Barros, B. R. C., de Carvalho, E. B., & Junior, A. C. P. B. (2022). Inland waterway transport and the 2030 agenda: Taxonomy of sustainability issues. *Cleaner Engineering and Technology*, 8, 100462. https://doi.org/10.1016/j.clet.2022.100462

DiJoseph, P. K., Mitchell, K. N., Tetreault, B. J., & Marshall, J. K. (2019). Inland marine transportation system travel time atlas via automatic identification system (AIS) data: Ohio River, Upper Mississippi River, and Illinois River. U.S. Army Corps of Engineers, Engineering Research and Development Center.

Fan, T., Chen, D., Huang, C., Tian, C., & Yan, X. (2023). Inland Vessel Travel Time Prediction via a Context-Aware Deep Learning Model. *Journal of Marine Science and Engineering*, *11*(6), 1146. https://doi.org/10.3390/jmse11061146

Hyndman, R. J., & Athanasopoulos, G. (2021). *Forecasting: Principles and practice* (2nd ed.). OTexts. https://otexts.com/fpp2/

Kang, L., Meng, Q., & Liu, Q. (2018). Fundamental diagram of ship traffic in the Singapore Strait. *Ocean Engineering, 147*, 340–354. https://doi.org/10.1016/j.oceaneng.2017.10.051

Korkmaz, S. A., Erginer, K. E., & Onur, Y. (2023). Environmental and economic analyses of fuel cell and battery-based hybrid systems utilized as auxiliary power units on a chemical tanker vessel. *International Journal of Hydrogen Energy, 48*(69), 23279–23295. https://doi.org/10.1016/j.ijhydene.2023.01.320

Kujala, P., Hänninen, M., Arola, T., & Ylitalo, J. (2009). Analysis of the marine traffic safety in the Gulf of Finland. *Reliability Engineering & System Safety*, *94*(8), 1349-1357. https://doi.org/10.1016/j.ress.2009.02.028

Li, L., Lu, W., Niu, J., Liu, J., & Liu, D. (2018). AIS data-based decision model for navigation risk in sea areas. The Journal of Navigation, 71(3), 664-678.

Mahmoudzadeh, A., Khodakarami, M., Ma, C., Mitchell, K. N., Wang, X. B., & Zhang, Y. (2021). Waterway maintenance budget allocation in a multimodal network. *Transportation Research Part E: Logistics and Transportation Review*, *146*, 102215. https://doi.org/10.1016/j.tre.2020.102215

Mazaheri, A., Montewka, J., Kotilainen, P., Sormunen, O. V. E., & Kujala, P. (2015). Assessing grounding frequency using ship traffic and waterway complexity. *Journal of Navigation, 68*(1), 89–106. https://doi.org/10.1017/S0373463314000502

Mehan, A., & Casey, Z. S. (2023). Blue Infrastructures: An Exploration of Oceanic Networks and Urban–Industrial–Energy Interactions in the Gulf of Mexico. *Sustainability*, *15*(18), 13699. https://doi.org/10.3390/su151813699

Mitchell, K. N., DiJoseph, P. K., Chambers, M., & Kress, M. M. (2019). Inland Marine Transportation System Fluidity: Case Studies from the Ohio River, Lower Mississippi River, and Gulf Intracoastal Waterway. Engineer Research and Development Center. https://hdl.handle.net/11681/34693

Park, K., Sim, S., & Bae, H. (2021). Vessel estimated time of arrival prediction system based on a path-finding algorithm. *Maritime Transport Research, 2*, 100012. https://doi.org/10.1016/j.martra.2021.100012

Sadeghi Gargari, N., Akbari, H., & Panahi, R. (2019). Forecasting short-term container vessel traffic volume using hybrid ARIMA-NN model. *International Journal of Coastal and Offshore Engineering, 4*(3), 47–52. https://doi.org/10.29252/ijcoe.3.3.47

Sui, Z., Wen, Y., Huang, Y., Zhou, C., Xiao, C., & Chen, H. (2020). Empirical analysis of complex network for marine traffic situation. *Ocean Engineering, 214*, 107848. https://doi.org/10.1016/j.oceaneng.2020.107848

Tan, Z., Zeng, X., Shao, S., Chen, J., & Wang, H. (2022). Scrubber installation and green fuel for inland river ships with non-identical streamflow. *Transportation Research Part E: Logistics and Transportation Review, 161*, 102677. https://doi.org/10.1016/j.tre.2022.102677

Texas Department of Transportation. (2014). *Master plan for the Gulf Intracoastal Waterway in Texas.* https://ftp.dot.state.tx.us/pub/txdot-info/tpp/giww/master-plan-0814.pdf

Tirunagari, S., Hanninen, M., Guggilla, A., Stahlberg, K., & Kujala, P. (2012). Impact of similarity measures on causal relation based feature selection method for clustering maritime accident reports. Journal of Global Research in Computer Science, 3(8), 46-50.

U.S. Army Corps of Engineers. (2023). *Waterborne Commerce Statistics Center.* https://ndc.ops.usace.army.mil/wcsc/webpub/#/

U.S. Coast Guard Navigation Center. (2023). *Automatic Identification System (AIS) overview*. https://www.navcen.uscg.gov/automatic-identification-system-overview

Welch, K. L., Lambert, L. H., Lambert, D. M., & Shideler, D. (2023). Multi-state economic contribution and multi-congressional district impact analysis of an inland waterway disruption. *Case Studies on Transport Policy, 13*, 101043. https://doi.org/10.1016/j.cstp.2023.101043

Wen, Y., Huang, Y., Zhou, C., Yang, J., Xiao, C., & Wu, X. (2015). Modelling of marine traffic flow complexity. *Ocean Engineering, 104*, 500–510. https://doi.org/10.1016/j.oceaneng.2015.04.051

Wohlgemuth, M., Fries, C. E., Sant Anna, A. M. O., Giglio, R., & Fettermann, D. C. (2020). Assessment of the technical efficiency of Brazilian logistic operators using data envelopment analysis and one inflated beta regression. *Annals of Operations Research, 286*(1–2), 703–717. https://doi.org/10.1007/s10479-018-3105-7

Wu, X., Roy, U., Hamidi, M., & Craig, B. N. (2020). Estimate travel time of ships in narrow channel based on AIS data. *Ocean Engineering, 202*, 106790. https://doi.org/10.1016/j.oceaneng.2019.106790

Xin, X., Liu, K., Yang, X., Yuan, Z., & Zhang, J. (2019). A simulation model for ship navigation in the "Xiazhimen" waterway based on statistical analysis of AIS data. *Ocean Engineering, 180,* 279–289. https://doi.org/10.1016/j.oceaneng.2019.03.052

Xu, X., Liu, C., Li, J., & Miao, Y. (2022). Trajectory clustering for SVR-based time of arrival estimation. *Ocean Engineering, 259,* 111930. https://doi.org/10.1016/j.oceaneng.2022.111930

Yang, D., Wu, L., Wang, S., Jia, H., & Li, K. X. (2019). How big data enriches maritime research—A critical review of Automatic Identification System (AIS) data applications. *Transport Reviews, 39*(6), 755–773. https://doi.org/10.1080/01441647.2019.1649315

Yeon, J., Elefteriadou, L., & Lawphongpanich, S. (2008). Travel time estimation on a freeway using discrete time Markov chains. *Transportation Research Part B: Methodological, 42*(4), 325–338. https://doi.org/10.1016/j.trb.2007.08.005

Yu, J. J., Tang, G. L., Song, X. Q., Yu, X., Qi, Y., Li, D., & Zhang, Y. (2018). Ship arrival prediction and its value on daily container terminal operation. *Ocean Engineering, 157,* 73–86. https://doi.org/10.1016/j.oceaneng.2018.03.038

Zhang, W., Goerlandt, F., Kujala, P., & Qi, Y. (2018). A coupled kinematics model for icebreaker escort operations in ice-covered waters. *Ocean Engineering, 167,* 317–333. https://doi.org/10.1016/j.oceaneng.2018.08.035

Zhang, W., Goerlandt, F., Kujala, P., & Wang, Y. (2016). An advanced method for detecting possible near miss ship collisions from AIS data. *Ocean Engineering, 124,* 141–156. https://doi.org/10.1016/j.oceaneng.2016.07.059

Zhang, L., Meng, Q., & Fang Fwa, T. (2017). Big AIS data based spatial-temporal analyses of ship traffic in Singapore port waters. *Transportation Research Part E Logistics and Transportation Review, 129,* 287–304. https://doi.org/10.1016/j.tre.2017.07.011

Zhang, W., Zou, Z., Goerlandt, F., Qi, Y., & Kujala, P. (2019). A multi-ship following model for icebreaker convoy operations in ice-covered waters. *Ocean Engineering, 180,* 238–253. https://doi.org/10.1016/j.oceaneng.2019.03.057

Zhao, L., Shi, G., & Yang, J. (2018). Ship trajectories pre-processing based on AIS data. *Journal of Navigation, 75*(5), 1210–1230.
https://doi.org/10.1017/S0373463318000188

Zhou, Y., Daamen, W., Vellinga, T., & Hoogendoorn, S. P. (2019). Ship classification based on ship behavior clustering from AIS data. *Ocean Engineering, 175*, 176–187.
https://doi.org/10.1016/j.oceaneng.2019.02.005