

ARMA Model-Based Prediction of the Number of Vessels Navigating the Istanbul Strait Unassisted by Maritime Pilots

Pelin Bolat, Gizem Kayisoglu

The Istanbul Strait is one of the busiest and riskiest trade routes, with the annual traffic of 50,000 ships. Such high traffic density is managed by the enforcement of a passage regimen by the Vessel Traffic Service (VTS) and maritime pilots of the Directorate General of Coastal Safety of the Republic of Turkey. VTS operations and maritime pilot actions are assumed to complement each other. Accordingly, a vessel unaccompanied by a maritime pilot is expected to interact with the VTS to a greater extent than a vessel assisted by a maritime pilot. Thus, estimating the number of ships that pass through the Istanbul Strait, especially those that do not use maritime pilot assistance, will be an effective tool for the Istanbul Strait traffic scheme management, as it will allow the authorities to balance and integrate VTS and maritime pilot operations. The predictive model based on Autoregressive Moving Average (ARMA) described in

this paper has been developed to estimate the number of ships that navigate through the Istanbul Strait without pilot assistance. The best ARMA model was identified through the use of historical data on 100-150 meter and 150-200-meter-long ships that passed through the Istanbul Strait unaccompanied by pilots in 2012-2019. The ARMA model obtained has also been validated through the comparison of real and estimated data.

1. INTRODUCTION

The Istanbul Strait, which is one of the narrowest international shipping straits in the world, is approximately 31 km long, with the average width of 1.5 km (Ulusu et al, 2009). It is the world's most important and hazardous natural water surface characterized by complex geographical structure, strong currents, sharp turns, variable weather conditions, traversed by approximately 140 ships, 25 dangerous cargo ship and 2,500 regional vessels each day. Given these intense and complex conditions and the limited functionality of the Istanbul Strait, the volume of traffic in the Istanbul Strait may safely be assumed to have reached very critical and dangerous levels. According to information obtained from the Akten study (2004), only 4,700 ships passed through the Istanbul Strait in 1936, compared to the present 50,000. In other words, after the Montreux Convention, important for the Istanbul Strait, was signed in 1936, the number of ships passing through the Istanbul Strait has increased approximately by a factor of 8. Apart from this dramatic increase in the number of ships in the Istanbul Strait, the 2019 maritime sector report by the Maritime Chamber of Shipping clearly

KEY WORDS

~ Istanbul Strait
~ ARIMA
~ Forecasting
~ Time Series
~ VTS

Istanbul Technical University, Maritime Faculty, Istanbul, Turkey

e-mail: yilmazp@itu.edu.tr

doi: 10.7225/toms.v10.n01.001

This work is licensed under



Received on: Apr 7, 2020 / Revised on: Dec 19, 2020 / Accepted on: Feb 15, 2021 /
Published: Apr 20, 2021

shows a 25 % decrease in the number of ships passing through the Istanbul Strait in 2006-2018 and a 29 % increase in the gross tonnage of ships passing through the Istanbul Strait (Table 1 - Istanbul & Marmara, Aegean, Mediterranean, Black Sea Region,

Turkish Chamber of Shipping, 2012-2019). In spite of the relative decline in the number of ships, once the tonnage is fixed, vessel numbers will clearly increase to reflect commercial volume growth going hand in hand with global developments.

Table 1.
Number of vessels passing through the Istanbul Strait.

The Istanbul Strait			
Years	Number of vessels	Gross tonnage (GT)	GT changes
2006	54,880	475,796,880	
2007	56,606	484,867,696	%2
2008	54,239	515,639,614	%6
2009	51,422	514,656,446	%0
2010	50,872	505,615,881	%-2
2011	49,798	523,543,509	%4
2012	48,328	550,526,579	%6
2013	46,532	551,771,780	%0
2014	45,529	582,468,334	%6
2015	43,544	565,216,784	%-4
F2016	42,553	565,282,287	%0
2017	42,978	599,324,748	%7
2018	41,103	613,088,166	%3

At the moment, Vessel Traffic Services (VTS) and pilot services, affiliated to the Turkish Directorate General of Coastal Safety, are management tools ensuring safe navigation to huge numbers of vessels passing through the Istanbul Strait. They are important control mechanisms in the Turkish straits given the high density of sea traffic, increasing vessel length and tonnage, growing number of passages by vessels carrying dangerous cargo, complex and difficult traffic structure. All activities of these services are oriented on guiding the ships in compliance with national and international developments, the Montreux contract, rules and recommendations of the International Maritime Organization (IMO) and other similar systems, all in cooperation with other maritime activities in the region. The VTS is of utmost importance as it monitors maritime traffic, reacts to changing conditions and actively organizes and plans maritime traffic to increase maritime safety, taking into account maritime traffic efficiency in its area of responsibility. Basically, the IMO approved three VTS services: Information Service (IS), Traffic Organization Service (TOS) and Navigation Aid Service (NAS) (Vessel Traffic Services Manual, 2012; Mısı, 2008). Within the scope of activity of the IS, TOS and NAS, VTS operators provide general information to ships in the region and neighboring areas. The information generally include (Yahsi and Ugurlu, 2014)

sea traffic information, information on a vessel's position with respect to the position of other ships in the VTS area, information on speed relative to the position, route and location of other ship, weather and hydrological conditions, warnings to seafarers issued or requested to be published by the relevant authorities, information on pilot operations, status and suitability of safe waterways, status of defective navigational aids, any danger threatening ships and other issues that may affect the safety of maritime traffic. Similarly, marine pilots have knowledge of all the local factors capable of disturbing the safety of navigation, and cooperate with the VTS, port authorities, tugs, and mooring teams during berthing operations or while the ship is passing through the strait or navigating local waters. In addition, they greatly contribute to the communication between vessels and the VTS as they use their native language. In addition, since marine pilots are also seafarers, they are familiar with the professional terminology and hierarchical language onboard ships, enabling them to give navigational advice properly and easily (Kuronen & Tapaninen 2010; Lappalainen et al. 2014). Though marine companies see pilotage services as a separate activity, they should be considered a public service owing to their ability to prevent potentially huge damage in any maritime accidents.

According to the Maritime Statistics of the Ministry of Transport and Infrastructure of the Republic of Turkey (accessed: 18 March 2020), a total of 41,112 vessels passed through the Istanbul Strait in 2019, of which only 26,632 used pilotage service in the Istanbul Strait. The remaining 14,480 vessels passed the Istanbul Strait with the assistance of the VTS, which is responsible for sea traffic management. Taking efficiency into account, the VTS transmits information to the ships where necessary, measures and records navigational information to be used for this purpose. The VTS undoubtedly has to invest more effort into ships passing through the Istanbul Strait without pilot assistance, than in ships that use pilot services, as pilots are familiar with the region and the features of ports in the Istanbul Strait, the local legislation and rules to which the Istanbul Strait is subject, as well as with the currents and winds in the region. On the other hand, the VTS provides the same information to any ships trying to pass through the Strait unaccompanied by a pilot. In the Istanbul Strait, annually navigated by approx. 50,000 vessels, almost half of which do not use pilotage services, the VTS has to ensure the safety of navigation, life, property and environment.

The research question sprang from the above circumstances, as the conclusion was reached that the capability to predict the number of ships that will pass through the Istanbul Strait unaccompanied by pilots is important to the VTS. For this purpose, a predictive model based on Autoregressive Moving Average (ARMA) was created by using historical data on 100-150 meter and 150-200 meter-long ships that passed through the Istanbul Strait without pilots in 2012 and 2019. The prediction based on the best ARMA model coincides with the actual data for 2019. Accordingly, future predictions that would be useful to the VTS can be obtained by shifting historical data forwards. The goal is to assess whether a change of the vessel scheme planning will become necessary in the near future to ensure more efficient marine traffic management in the Istanbul Strait.

2. OVERVIEW OF THE RELEVANT LITERATURE

VTS and marine pilotage services in the Istanbul Strait are the most important contributing factors to the safety of navigation, clean and stable environment, safety of property and life. Several studies have dealt with this topic. According to Oral and Ozturk's study (2006), ships accompanied by a pilot are less likely to suffer accidents in the Istanbul Strait. In addition, Akten (2004) stated in his study that pilotage was the service that minimized the number of shipping accidents in the Istanbul Strait. Similarly, various studies have found that pilotage services decrease the risk of accident (Ugurlu et al, 2016; Simsir et al, 2014; Arslan and Turan, 2009; Simsir and Ertugrul, 2009; Malekipirbazari et al, 2015; Birpinar et al, 2009; Köse et al, 2003). On the other hand, the VTS is the biggest decision support system available to ships passing through the strait without a pilot, as the VTS system

has at its disposal techniques and skills that help prevent vessel collisions and provides both general assistance, and assistance in conditions where the crew's view is limited.

In this context, since the VTS has onshore operators at disposal, they use a variety of decision support systems to provide safe, effective, and efficient services to the maritime community. The VTS, strengthened with the above predictive capabilities, will thus greatly benefit the maritime community.

The important factors affecting the effective decision making by the VTS have already been researched. For example, Ozbas (2005) proposed the "Istanbul Strait Non-Stop over Transition Simulation Model" that encompasses the Istanbul Strait traffic rules and regulations, traffic density, characteristics of cargo transported, weather and geographical conditions, services such as pilots and tugboats. Having examined the effect of six different factors on twenty dependent variables, the study found the number of pilots and tug services to be the most important factor, followed by the rate of ships passing through the Istanbul Strait. These results suggest that there is a need to predict the number of ships passing through the Istanbul Strait.

There is a limited number of studies that directly address this topic. Only two studies close to our research subject deal with the Istanbul Strait. One is Bulut's study (2014) and the other Malekipirbazari et al. (2015) (Bulut, 2014; Malekipirbazari et al, 2015). Bulut (2014) developed the fuzzy time series forecasting (FTSF) method. The proposed model was applied to the volume of vessel traffic in the Istanbul Strait in order to compare the accuracy of the proposed model with benchmark methods in 2014. Bulut also stated that vessel traffic control planning was a critical issue in the organization of the vessel traffic service. He found that the data set relating to the vessel traffic in the Istanbul Strait is used for the implementation of the proposed method. Bulut study and this study differ in the method used and year. In this study, the general ARMA model, incorporating the Box-Jenkins method, was used in 2012-2019. In contrast, the Bulut study used the seasonal time series FTSF model that involved the application of the bivariate fuzzy time series method in 2005-2011. In addition, this study contributes to the Bulut study in terms of timeliness. The second similar study was conducted by Malekipirbazari et al. They applied the ARIMA statistical forecasting methodology to period 2006-2013 to predict vessel type traffic in the Istanbul Strait anchorages over the next three years, believing that a good estimate of future vessel type traffic was the key parameter for effective anchorage area management. The only similarity between the Malekipirbazari et al. and this study is the use of the ARIMA method, as the research area is quite different. While the research subject of the Malekipirbazari study is the prediction of vessel types berthing at anchorage zones in the Istanbul Strait, the research subject of this study is predicting the number of vessels passing through the Istanbul Strait unaccompanied by a pilot. Apart from the above studies, Yoo et al (2013), though

not related to the Istanbul Strait predicted future marine traffic volume based on the data on the number of vessels entering Incheon port between January 1996 and June 2013, using the ARIMA model. They estimated future monthly traffic volumes with the time series model. Their results show that in the period commencing in January 2014 and ending in December 2015, the lowest traffic volume was recorded in February and the highest in January. They also stated that future marine traffic volume prediction based on time series analysis is more suitable than the estimation of marine traffic volume through regression analysis.

There are also some analyses of maritime traffic flows based on other methods. For example, Mostafa (2004) estimated maritime traffic flows in the Suez Canal by combining the ARIMA and Neural network models. He explained the importance of the ARIMA and neural network models as follows. The ARIMA is a basic univariate model that takes into consideration outliers, interventions, calendar variations, or other real factors of a frequently observed time series. Similarly, neural network models are also used for forecasting in classification areas where statistical techniques such as regression models have been used conservatively. He analyzed the behavior of the Canal's maritime traffic flows from June 1975 to June 1998. Xiao et al (2017) proposed a knowledge-assisted maritime traffic prediction methodology based on a vessel's waterway pattern and motion behavior. By incorporating traffic density, waterway pattern, water point pattern, vessel motion pattern and vessel classification, they introduced a set of probable estimates with corresponding probabilities for route planning, traffic scheduling, vessel coordinates, traffic optimization, and traffic safety management. They achieved this through the integration of the route trace discovery process of a long term traffic prediction algorithm that includes both knowledge mining and forecasting components, as well as through the use of other prediction methods, such as ARIMA and neural networks (for sailing distance forecasting) to support the maritime traffic prediction. This concept is also a beneficial and practical way to obtain forecast results and identify the mathematical model, that simultaneously facilitates maritime traffic planning and coordination. In their report, Hara and Nakamura (1995) introduced the evaluation of the safety of the maritime environment for shipping-traffic by means of a shipping traffic flow simulation capable of reproducing ship movements. They used subjective judgement values, obtained from simulator experiments using elements that show vessels in relation to other vessel navigating in the same area. In this way, they established the safety assessment system flow for the shipping-traffic environment using simulation techniques. Consequently, they identified three safety assessment areas applicable to any shipping- traffic environment: "assessment of the rationales for a marine transport plan", "assessment of the degree of danger of collisions in a specific sea area, and assessment of the traffic volume on a traffic route", and

"assessment of the ship handling environment and port and harbor facilities". Accordingly, as long as vessels navigate straits and bays with high traffic density, traffic volume assessment and collision risk assessment are mandatory requirements for any new marine transport plan. They therefore carried out simulations to estimate future shipping-traffic volume. Wen et al (2015) stated that maritime traffic situations are difficult to manage due to the rapid increase in traffic density and increase in the probability that ships will come into close proximity. From this point of view, they presented a marine traffic complexity model to assess the traffic situation and support traffic controllers and mariners to raise their awareness of the traffic situation. For this purpose, they took pair-wise ship traffic characteristics, such as relative speed, relative distance, and intersecting trajectory and ran a simulation for the Shenzhen Wet Sea area. They constructed a model that showed complexity in different traffic situations. Zhou et al (2020) proposed deep learning-based solutions to estimate the flow of vessels in a selected rectangle marine area in the Singapore Strait. They used the AIS information of vessels in the region, such as vessel ID, time stamp, and coordinates, to predict the inflow and outflow of vessels in each grid in the rectangle area, by dividing the area with a matrix. The data were collected for 31 days, from 1 October 2013 to 31 October 2013. They established the advantages and disadvantages of convolutional neural networks, long short-term memory networks, and their hybrid versions. They found that the hybrid of a bidirectional long short-term memory network and a convolutional neural network was the best model for this purpose. Finally, Praetorius et al (2015) emphasized that the responsibility for the safe entrance of ships into ports and maritime traffic fluency lies primarily with the VTS. Their study therefore focused on the way in which the VTS ensured traffic efficiency and safety in its area. Everyday operations of VTS systems have been described using the functional resonance analysis method. Their results suggest that if changes are planned and implemented by the VTS, and everyday operations are properly analyzed and understood, we can understand how technological and organizational changes will affect overall system performance.

3. THE ARMA MODEL

The aim of this study is to identify a model capable of predicting the number of vessels passing through the Istanbul Strait unaccompanied by a pilot by analyzing the data series collected between January 2006 and March 2019. In this context, the ARMA model based on the Box-Jenkins (BJ) method that makes predictions based on previous period values and previous error terms was used.

The Box-Jenkins method is used for the forward estimation of univariate time series. It systematically establishes forward prediction models for discrete and stationary time series

consisting of observed values obtained in equivalent time intervals and makes predictions (Mabert and Radeliffe, 1974). Discrete and stationary series of observed values obtained in equivalent time intervals are important assumptions in the B.J method.

The Box-Jenkins prediction method differs from other prediction methods in that it does not require any prior knowledge about time series structure or the general development trend. In addition, while other methods require the series to have a certain tendency, the Box-Jenkins method can be applied to complex time series as it is free of this restriction (Ozcelik, 1980). An important advantage of this method is that it uses past observed values as an explanatory variable.

Rather than being a method expressed with a single model, the Box-Jenkins prediction techniques, which are an experimental process, can choose an appropriate model among various model options and check the eligibility of the selected model at every stage (Duru, 2007).

In our paper, there is one explanatory variable (monthly number of ships). In addition, our data series has a seasonal effect and when a test was conducted that canceled out the seasonality effect, the data series became stationary. Furthermore, the data are approximately linear. Therefore, a basic method, such as ARMA, was sufficient to predict the number of ships passing through the Istanbul Strait without pilots and be serviceable for marine traffic management in the Istanbul Strait. On the other hand, unlike our paper, the ARIMAX is suitable for analysis in the case of additional explanatory variables (multivariate) in categorical and/or numeric format (Smarten 2018). Literature suggests that the Fuzzy ARIMA model is used where data availability is inadequate (Wang, 2011; Torbat et al, 2018; Kannan et al, 2019; Mehdi et al, 2019). However, in our paper, the real value sufficed for ARMA model predictions. In general, the Neuro-ARIMA is used for nonlinear modelling (Zhang, 2003; McDonald et al, 2013; Adebisi et al, 2014).

All steps of the analysis have been clarified in Section 4 in detail. To sum up, in the first step, seasonality of the data series is examined, since the seasonality effect may cause false regression of the frequency time series. The seasonality effect in the data series used had the significance level of 1 %. Therefore, the data series was cleared of the seasonality effect through the application of the X12-Censun Method. Then, the stationary part of the data series was analyzed, as data series is required to be stationary for the purposes of econometric analyses that include time series. On the other hand, when nonstationary time series is used in time series analyses, there is the risk of false regression. For this reason, the Augmented Dickey Fuller (ADF) unit root test (Dickey-Fuller, 1979) and Phillips-Perron (PP) unit root test (Perron, 1989) have been performed on the "Ships" variable that indicates the number of vessels passing through the Istanbul

Strait unaccompanied by a pilot, and is the sole variable used in the analysis. When the time path graph of the "Ships" variable was examined in the findings section, some periods of increase and decrease were identified and interpreted as a structural break. The unit root test with structural break was therefore also performed on the time series data, as well as classic unit root tests.

In unit root tests, it is important whether a series has characteristics such as trend and intercept, intercept and none. In a unit root test, if all of the specifications with intercept, none and trend indicating that the series is stationary or not stationary are present, the unit root can be determined regardless of the characteristics of the series. On the other hand, if there is a discrepancy in the specifications, unit root test findings that reflect series characteristics should be taken into account. Consequently, series graphics in the findings section were examined, and the series was found to have decrease trends. Then, unit root tests were conducted, and the findings of both the ADF unit root test and the Phillips-Perron unit root test for trend and intercept with trend specification indicate the stationarity of level values. On the other hand, the results of the unit root test with a structural break suggest that all specifications of the series are stationary. In the light of these findings the series was determined to be stationary at the level, and the ARMA modeling phase was initiated in keeping with the Box-Jenkins approach.

The ARMA model can be expressed as (1):

$$Y_t = \delta + \sum_{i=1}^p \Phi_i Y_{t-i} + \sum_{j=1}^q \epsilon_j \Phi_{t-j} + \epsilon_t \quad (1)$$

In the equation, parameter " Φ " indicates autoregressive (AR) root that includes previous period values. Parameter " θ " is the moving average (MA) root that involves previous period error terms. " δ " is constant term and " ϵ " is error term. " p " and " q " state optimal lag length for the roots of AR and MA, respectively. Optimal lag lengths are determined by considering performance and criteria simultaneously, such as Akaike Information Criteria (AIC), Schwarz Information Criteria (SIC), adjusted determination coefficient (R^2), Root mean square error (RMSE), and mean absolute error (MAE). On the other hand, in this study, we wanted to establish whether the roots of AR and MA take place in the unit circle, that is to say, whether the ARMA system is stationary or steady. Once the estimated ARMA modelling was evaluated by looking into all criteria and performances, and it provided features on demand, the predictive graph was plotted. The prediction was verified by using historical values of the series itself. Finally, the prediction for April 2019 was obtained and the conclusion presented.

4. ANALYSIS AND MODELING

4.1. Data Description

The data on the number of ships passing through the Istanbul Strait unaccompanied by a pilot were obtained from maritime sector reports published annually from 2012 to 2019 by the Turkish Chamber of Shipping for Istanbul & Marmara, Aegean, Mediterranean, Black Sea Region (2012-2019). While the total number of ships passing through the Strait and the number of ships passing through the Istanbul Strait accompanied by a pilot are given in the report, the number of unaccompanied ships was obtained by subtracting these two data. The data obtained relate to vessels between 100-150m and 150-200m in length. The vessels of these lengths were included in the analysis because maritime sector reports by the Turkish Chamber of Shipping for Istanbul & Marmara, Aegean, Mediterranean, Black Sea Region (2012-2019), the majority of vessels passing through the Istanbul Strait are of the stated length. Besides, the reports have shown that almost all vessels over 200 meters in length have taken a pilot, which makes them irrelevant for this analysis. Table 2 gives the number of vessels that passed through the Istanbul Strait

obtained from the maritime sector report by the Turkish Chamber of Shipping for Istanbul & Marmara, Aegean, Mediterranean, Black Sea Region (2018) in 2018. It shows the figures for vessels of all lengths passing through the Istanbul Strait. In addition, according to the "Vessels subject to additional Passage Rules through Turkish Straits and Principles of Implementation", tankers and ships carrying dangerous cargoes having the overall length (LOA) of 200 meters and more (except Roro and container ships) and vessels having LOA over 250 meters (except tankers and ships carrying dangerous goods) and/or deep draft ships are strongly recommended to effect their passage through the strait assisted by pilots and tugboats (TSVTS Center, 2019). Consequently, vessels over 200 meters in length generally hire a pilot. Vessels under 100 meters in length are known to generally operate regionally and pose a lesser risk. As seen from Table 2, these ships hired pilots less frequently precisely for this reason. However, this study looked into vessels having the length between 100-150m and 150-200m, as they are not required to hire a pilot, yet are considered high risk. Thus, prediction of the number of unaccompanied ships from this group is relevant and risk profiles for the following years can form the basis of other studies.

Table 2.

The monthly statistics of vessels passing through the Istanbul Strait by length and request for pilot assistance (2018).

Months	Over 300 m		250-300 m		200-250 m		150-200 m		100-150 m		Under 100 m		Total	
	Total vessels	With pilot	Total vessels	With pilot	Total vessels	With pilot	Total vessels	With pilot	Total vessels	With pilot	Total vessels	With pilot	Total vessels	With pilot
January	0	0	104	104	198	198	954	827	1.217	533	901	256	3.374	1.918
February	0	0	103	103	186	186	935	806	1.160	550	853	225	3.237	1.870
March	0	0	106	106	223	221	1.025	883	1.363	626	989	267	3.706	2.103
April	0	0	123	123	229	229	973	871	1.263	564	1.007	284	3.595	2.071
May	1	1	122	122	219	216	924	796	1.254	502	1.048	288	3.568	1.925
June	1	1	118	118	205	203	854	716	1.194	479	857	240	3.229	1.757
July	0	0	112	112	229	227	994	869	1.198	495	1.075	281	3.608	1.984
August	0	0	116	114	278	269	1.020	860	1.110	467	920	231	3.444	1.941
September	0	0	117	117	238	238	952	812	1.050	402	800	225	3.157	1.794
October	0	0	123	123	232	232	1.058	961	1.229	478	876	270	3.518	2.064
November	1	1	119	119	238	238	957	913	1.126	456	752	238	3.193	1.965
December	0	0	114	114	251	251	994	979	1.302	576	813	253	3.474	2.173
Total	3	3	1.377	1.375	2.726	2.708	11.640	10.293	14.466	6.128	10.891	3.058	41.103	23.565

In the analyzed data, the total number of ships passing through the Istanbul Strait unaccompanied by a pilot is available for all months from 2012 to 2018; only the first quarter of 2019 is considered. The analysis based on ARMA model was conducted on an EViews software.

4.2. Findings

The descriptive statistics of the variable are calculated in Table 3. The "Ships" factor is a normally distributed variable, between the maximum of 1394 and minimum of 648, with the standard deviation of 157.001 and the mean value of 998.239.

Table 3.

Descriptive Statistics of Variable.

Variable	Mean	Max.	Min	Std. Dev	S	K	J.B
Ships	998.239	1394	648	157.001	0.284	2.398	4.545 (p=0.103)

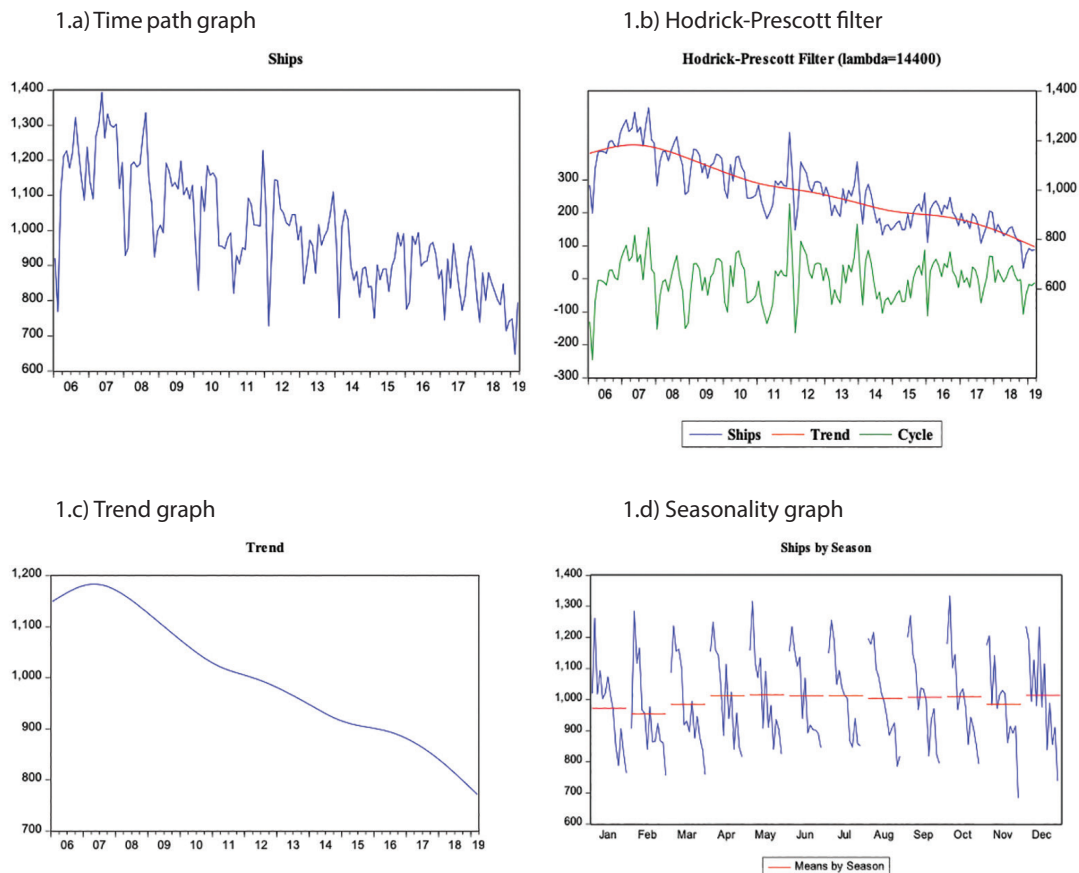


Figure 1.

Graphs of "Ships" Value.

The graphics of the series are presented in Graph 1. The time path graph of the series is shown in Graph 1a. The series is seen to progress with a nearly constant variance in a generally decreasing trend. When the trend variable obtained using the Hodrick-Prescott Filter system is analyzed, the decreasing trend of the series is clearly seen. On the other hand, some structural breaks are observed in the open series obtained by subtracting the trend from the series in the Hodrick-Prescott Filter system. Finally, when the seasonal averages of the series are analyzed, the differences between month averages cause suspicion of seasonality.

The seasonality of the series was determined by performing the seasonal F test, the findings of which are presented in Table 4. When Table 4 is examined, the series is seen to have a statistically significant seasonal effect of 1 % significance level. The series

was cleared of seasonality using the X-12 Census method, which prevented the seasonality observed from causing spurious regression.

Table 4.
Seasonal F test.

Variable		Sum of Squares	Dgrs. of Freedom	Mean Square	F
Ships	Between months	5617.7029	11	510.70027	
	Residual	4983.5124	147	33.90144	15.064***
	Total	10601.2153	158		

*** shows the availability of the seasonality effect of a (%1) significant level.

In addition, the Augmented Dickey Fuller (ADF) test and the Phillips-Perron (PP) test, that test the stationarity of the time series analysis, have been performed as classic unit root tests that were used as statistical evidence. Table 5 clearly shows that according to ADF test statistics, the series is not stationary at the level suggested by test statistics calculated for unit root test regression specifications with and without intercept. Yet, test statistics calculated using the regression specification of the unit root test with trend and intercept, indicate that the significance level of

its stationarity is 1 %. Similarly, the PP test shows that the series is stationary at 1 % significance level in unit root test containing trend and intercept, stationary at 10 % significance level in unit root test containing only intercept, and not stationary when non-intercept and non-trend specification is used. In this case, given the decreasing trend of the series, unit root test findings containing the trend should be taken into account and the series should be considered stationary in level values.

Table 5.
Unit root test findings.

Variable	Augmented Dickey-Fuller test			Phillips-Perron test		
	None	Intercept	Trend and intercept	None	Intercept	Trend and intercept
Ships p values	-1.101686 (0.2447)	-0.697724 (0.8431)	-6.842787*** (0.0000)	-0.924085 (0.3148)	-2.590240* (0.0971)	-6.621643*** (0.0000)

*(10), ***(%1) significance level. Unit Root Test Regression Delay Lengths have been determined using the SIC criteria.
(Max. Lag=12)

Table 6.
Unit root test with structural breaks.

Variable	Trend specification			
	Intercept	Intercept	Trend	Trend and intercept
Ships	-5.109782*** (<0.01)	-7.202718*** (<0.01)	-7.560766*** (<0.01)	-7.837742*** (<0.01)

***shows the availability of stationarity at (%1) significance level

In addition, unit root test with structural break was performed to account for data series structural breaks. The findings of the unit root test with structural breaks are presented in Table 6. Table 6 shows that the series stationarity has a 1 % significance level in all specifications according to the statistical significance values of the unit root test with structural break. When the results of the ADF and the PP unit root test with trend and intercept specifications and the results of the unit root test with structural break were evaluated together, the series was determined to be stationary. Consequently, series differencing for purposes of obtainment of a stationary series was considered unnecessary.

The appropriate ARMA model can be determined due to series stationarity. Parsimony, which is the fundamental idea behind the Box-Jenkins methodology, is the keyword at this stage. Parsimonious models give better forecasts than over-parameterized models (Vandekerckhove et al, 2015). The most appropriate model by parsimonious standards should meet the following criteria:

- The most significant coefficients
- The lowest volatility
- The highest adjusted R²
- The lowest AIC and SIC

For this purpose, in the framework of the Box-Jenkins approach, all models up to maximum ARMA (4, 4) are presented in Table 7 in order to contrast parameter significance, determination coefficients, regression performances (AIC, SIC) and estimation performances (RMSEA, MAE). On the other hand, in order to ensure that the ARMA models meet the requirement of being stationary and steady, statistics on whether the roots of AR and MA take place in the unit circle are also given in Table 7 ($|AR|$ and $|MA| < 1$). Table 7 shows that many models, starting from ARMA (1, 1), have both parameter significance and meet the requirement of being stationary and steady. In this case, diagnostic values need to be inspected to identify the best-performing model. The examination performed by determination coefficient adjustment has shown ARMA (3, 2) to be the best performing model. ARMA (3, 2) was also found to be the optimal model in terms of AIC. When other regression and prediction criteria were examined, these values have been shown to be very close to each other in many models and are not indicative of a single model. In the light of these findings, ARMA (3, 2) was the only model that could be considered optimal for a number of criteria. The ARMA (3, 2) model was chosen for ARMA estimation.

Table 7.

Estimation of ARMA (p, q) Modelling based on the Box-Jenkins approach.

	ARIMA (1,1)	ARIMA (2,1)	ARIMA (1,2)	ARIMA (2,2)	ARIMA (3,1)	ARIMA (1,3)	ARIMA (2,3)	ARIMA (3,2)
δ	968.229***	942.167***	942.873***	941.992***	942.423***	941.611***	941.771***	950.514***
Φ_1	0.96572***	1.441***	0.994***	1.114***	1.357***	0.996***	0.849	1.997***
Φ_2	-	-0.444***	-	-0.118	-0.571***	-	0.146	-1.561***
Φ_3	-	-	-	-	0.208**	-	-	0.554***
Φ_4	-	-	-	-0.440*	-	-	-	-
θ_1	-0.392***	-0.817***	-0.332***	-0.247*	-0.674***	-0.315***	-0.171	-1.338***
θ_2	-	-	-0.296***	-	-	-0.274***	-0.321	0.618***
θ_3	-	-	-	-	-	-0.062	-0.106	-
θ_4	-	-	-	-	-	-	-	-
$ AR $	<1	<1	<1	<1	<1	<1	<1	<1
$ MA $	<1	<1	<1	<1	<1	<1	<1	<1
$\overline{R^2}$	0.7861	0.799	0.803	0.802	0.804	0.801	0.800	0.805
AIC	11.199	11.152	11.132	11.144	11.133	11.144	11.156	11.129
SIC	11.258	11.229	11.209	11.240	11.229	11.241	11.272	11.245
RMSE	133.022	145.248	129.042	145.388	125.047	131.033	145.361	122.034
MAE	110.080	117.721	107.608	117.610	104.995	109.928	117.626	101.091

	ARIMA (3,3)	ARIMA (4,1)	ARIMA (1,4)	ARIMA (4,2)	ARIMA (2,4)	ARIMA (4,3)	ARIMA (3,4)	ARIMA (4,4)
δ	950.458***	942.975***	943.493***	950.448***	953.185***	950.461***	942.323***	943.549***
Φ_1	1.984***	1.195***	0.993***	2.002***	1.551***	1.051***	1.401***	0.456*
Φ_2	-1.531***	-0.423**	-	-1.566***	-0.559	0.342**	-1.394***	0.205
Φ_3	0.538***	0.079	-	0.565**	-0.899**	-0.921***	0.985***	-0.304
Φ_4	-	0.138	-	-0.009	-0.084	0.509***	-	0.621***
θ_1	-1.328***	-0.525**	-0.333***	-1.341***	0.082	-0.374	-0.736	0.215
θ_2	0.589**	-	-0.283***	0.614***	0.143	-0.699***	0.896	-0.113
θ_3	0.012	-	-0.059	-	-	0.611***	-0.314	0.266
θ_3	-	-	0.063	-	-	-	-0.238	-0.328**
AR	<1	<1	<1	<1	<1	<1	<1	<1
MA	<1	<1	<1	<1	<1	<1	<1	<1
$\overline{R^2}$	0.804	0.804	0.801	0.804	0.802	0.805	0.802	0.804
AIC	11.142	11.136	11.151	11.142	11.150	11.146	11.146	11.157
SIC	11.2767	11.252	11.267	11.277	11.285	11.301	11.301	11.331
RMSE	121.966	118.739	128.389	113.238	147.832	113.286	126.379	116.117
MAE	101.046	100.166	106.722	95.034	119.731	95.018	106.641	99.578
***(%1),**(%5),*(%10) show parameter significance level.								

The findings of AR and MA roots of the ARMA (3, 2) model are presented in Graph 2. As seen in Graph 2, all AR and MA roots of the ARMA (3, 2) model are part of the unit circle. In this case, the ARMA (3, 2) model, which is optimal for ARMA estimation, can be said to meet the requirement of being stationary and steady.

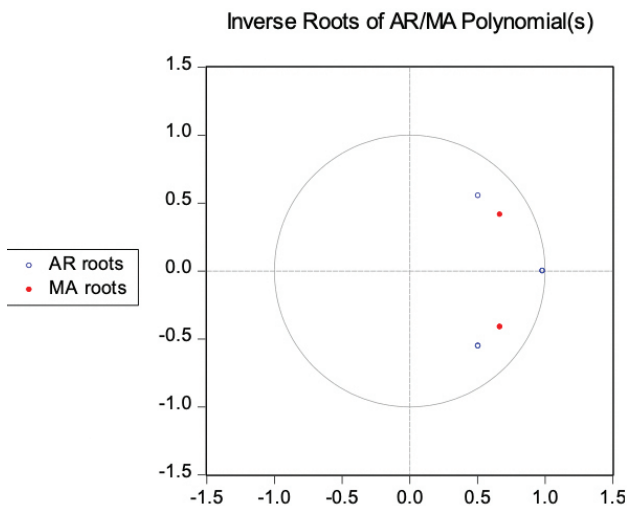


Figure 2.
Graph of the unit root of the ARMA (3, 2) model.

Other estimation performances of the ARMA (3, 2) model, proven to be a stationary and steady system with other conditions, are illustrated in Graph 3, showing that the number of ships navigating the Istanbul Strait without pilot assistance (Forecast-Ships), statically estimated with the ARMA (3,2) model, fluctuates in the ± 2 standard error range without significant deviation.

Another prediction performance is the comparison of an actual past series value with the model's estimate for that period. In Graph 4, the time path graph of both series is presented together. As seen in Graph 4, the time path of both series progresses together in the same direction and with small deviations. Similarly, since the calculated correlation coefficient between the series is 0.91, the conclusion was reached that the series' progression was very close.

The difference series, depicting the difference between the estimation series and the original series, follows the time path illustrated in Graph 5. When the difference series is examined in Graph 5, it is found to follow the time path around the zero mean. The maximum level of the difference corresponds to the highest structural break in the series in December 2011. (Difference: 220).

In the light of all these findings, the ARMA (3, 2) model was found to be the appropriate prediction model. The estimation made for the period after the end of the data set (April 2019) is 752.

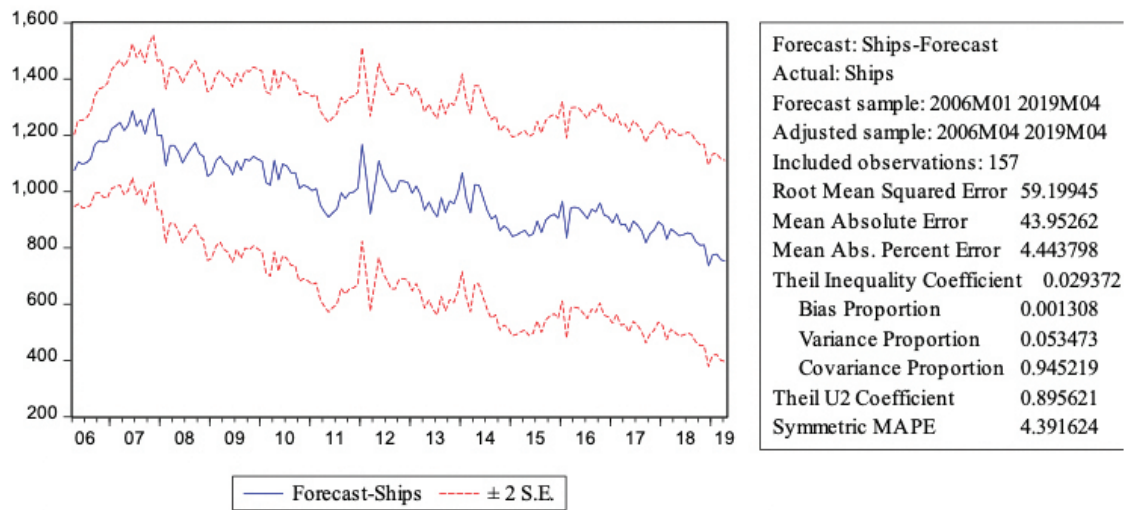


Figure 3.
 Estimation performances of the ARMA (3, 2) model.

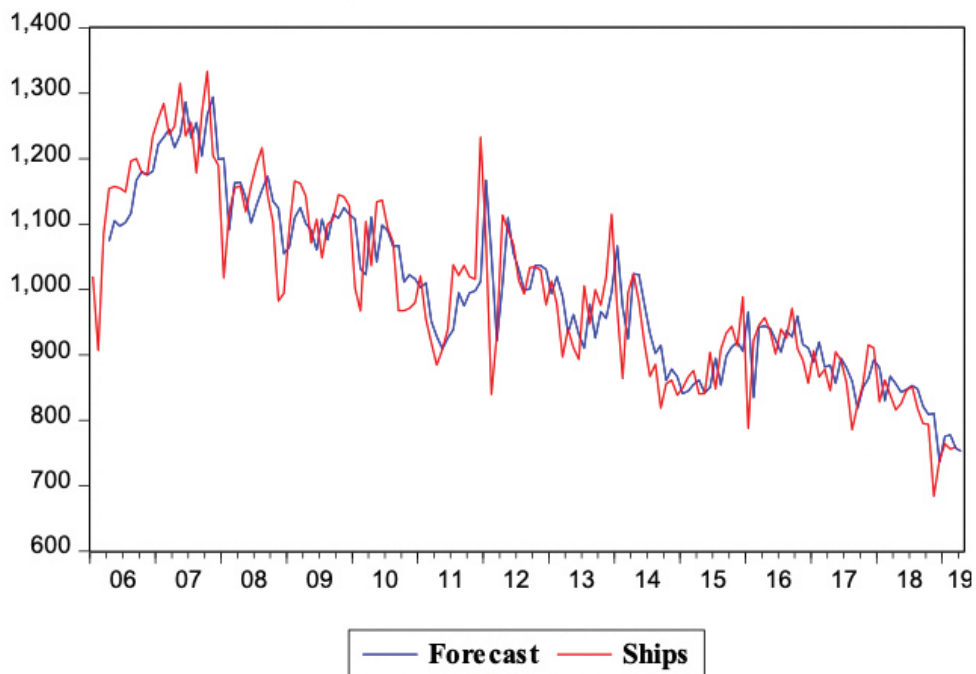


Figure 4.
 Estimation series of the ARMA (3, 2) model and past period value graph.

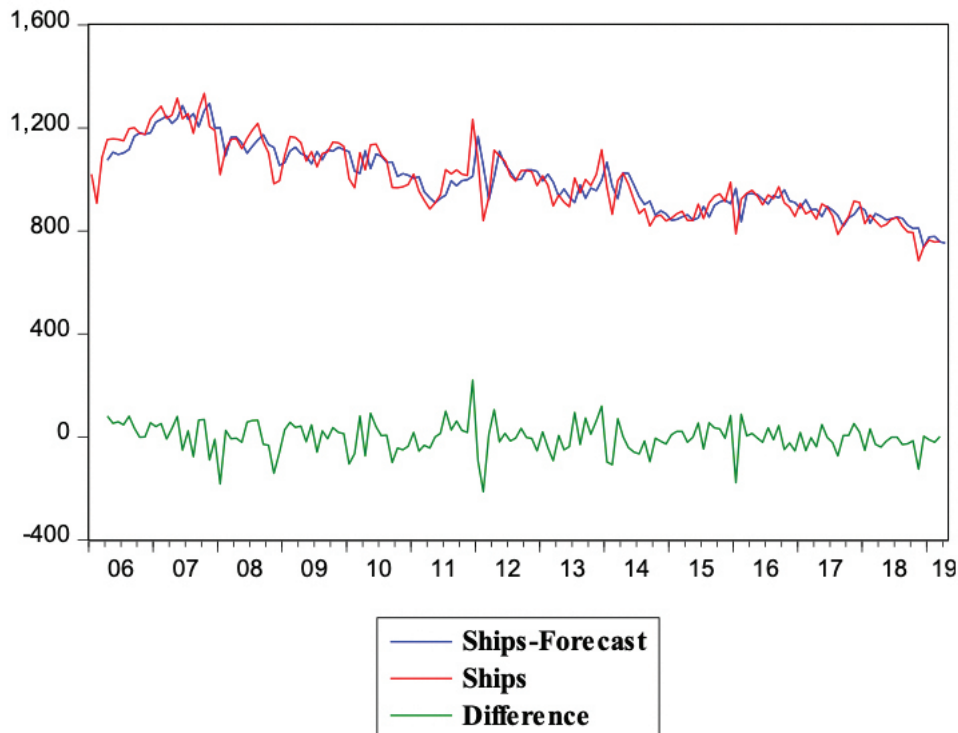


Figure 5.

Graph showing the difference between the ARMA (3, 2) estimation series and the original series.

5. RESULTS AND DISCUSSION

In short, the best ARMA model is an appropriate tool for predicting the number of ships that navigate the Istanbul Strait without pilot assistance. The ARMA (3, 2) was found to be the best model for this study. The estimated number of ships navigating the Istanbul Strait without pilot assistance for April 2019 was 752. According to the Maritime Statistics of the Ministry of Transport and Infrastructure (accessed: 14 October 2020), the real figure registered in April 2019 is 706. The difference between the actual and estimated data is acceptable and allows stakeholders, such as the VTS, pilot & tugboat center in the Istanbul Strait, to plan and coordinate the vessel schedule.

6. CONCLUSION

This study presents ARMA-based time series modeling and prediction of the number of ships navigating the Istanbul Strait unaccompanied by a pilot through the use of the EViews software. This paper can be very helpful to the VTS, acting under the authority of the Directorate General of Coastal Safety of the Republic of Turkey. Given that the VTS is a very significant decision maker in the maritime community, it is vital for the

maritime industry to systematically study many parameters analyzed by the VTS and make predictions for the future. The VTS should strengthen the decision supporting system that assists ships with their navigation decisions, with the help of new technologies and predictions. Predictions are key for effective decision-making in any organization not opting to deal with the future by insurance and risk protection. One of the advantages of forecasting is that decision-makers that do not have predictions at disposal have problems with any factors dependent upon future developments contemplated by an organization, that increase complexity and require technological structure. Another advantage of forecasting is that it allows organizations to choose a more systematic decision-making process, with a clear rationale for individual actions. Forecast information are clearly only useful when utilized in an organization's decision-making and planning processes. In this context, though practical application can be derived from theory, predictive methods require significant modification before use. The predictions relating to the number of ships navigating the Istanbul Strait without pilot assistance were analyzed in the framework of this study. The proposed ARMA model was found to be suitable for short-term predictions for the needs of the VTS, by shifting data and providing diagnostic controls and prediction validity.

Predictions based on the ARMA model are a comparatively improved time series estimation method. The ARMA model can factually clarify the dynamic change rules. It can be used for statistical analysis and time series predictions under certain conditions. The model is particularly suitable for short-term time range predictions, otherwise large deviations occur. Another reason for the deviations is that the ARMA model estimation relies solely on current values and historical data, without taking into account a multitude of explanatory factors.

Future studies could contribute to the plan and schedule of vessels navigating the Istanbul Strait by adding real and available dynamics and elements to the predicted number of vessels through the use of a simulation based method or a system dynamics method. This would allow the VTS to develop a meaningful and effective decision support system for the future, make investments and carry out maintenance at a more optimum level.

REFERENCES

- Akten, N., 2004. Analysis of Shipping Casualties in the Bosphorus. *Journal of Navigation*, 57(3), pp.345–356. Available at: <http://dx.doi.org/10.1017/s0373463304002826>.
- Arsilan, O. and Turan, O., 2009. Analytical investigation of marine casualties at the Strait of Istanbul with SWOT-AHP method. *Maritime Policy & Management*, 36(2), pp.131–145. Available at: <http://dx.doi.org/10.1080/03088830902868081>.
- Birpınar, M.E., Talu, G.F. and Gönençgil, B., 2008. Environmental effects of maritime traffic on the İstanbul Strait. *Environmental Monitoring and Assessment*, 152(1-4), pp.13–23. Available at: <http://dx.doi.org/10.1007/s10661-008-0292-8>.
- Bulut, E., 2014. Modeling seasonality using the fuzzy integrated logical forecasting (FILF) approach. *Expert Systems with Applications*, 41(4), pp.1806–1812. Available at: <http://dx.doi.org/10.1016/j.eswa.2013.08.079>.
- Dickey, D.A. and Fuller, W.A., 1979. Distribution of the Estimators for Autoregressive Time Series with a Unit Root. *Journal of the American Statistical Association*, 74(366a), pp.427–431. Available at: <http://dx.doi.org/10.1080/01621459.1979.10482531>.
- Duru, O., 2007. Zaman Serileri Analizinde ARIMA Modelleri ve Bir Uygulama, Yüksek Lisans Tezi, İstanbul Üniversitesi Sosyal Bilimler Enstitüsü Ekonometri Anabilim Dalı.
- Hara, K. and Nakamura, S., 1995. A comprehensive assessment system for the maritime traffic environment. *Safety Science*, 19(2-3), pp.203–215. Available at: [http://dx.doi.org/10.1016/0925-7535\(94\)00021-t](http://dx.doi.org/10.1016/0925-7535(94)00021-t).
- Istanbul and Marmara, Aegean, Mediterranean, Black sea Region Turkish Chamber of Shipping, 2012-2019. Maritime sector report. Available at: <https://www.denizticaretodasi.org.tr/yayinlarimiz/sektorraporu>.
- Köse, E. et al., 2003. Simulation of marine traffic in Istanbul Strait. *Simulation Modelling Practice and Theory*, 11(7-8), pp.597–608. Available at: <http://dx.doi.org/10.1016/j.simpat.2003.10.001>.
- Kuronen, J. and Tapaninen, U., 2010. Evaluation of maritime safety policy instruments. *WMU Journal of Maritime Affairs*, 9(1), pp.45–61. Available at: <http://dx.doi.org/10.1007/bf03195165>.
- Lappalainen, J., Kunnaala, V. and Tapaninen, U., 2013. Present pilotage practices in Finland. *WMU Journal of Maritime Affairs*, 13(1), pp.77–99. Available at: <http://dx.doi.org/10.1007/s13437-013-0055-4>.
- Mabert, V. A., and Radeliffe, R. C., 1974. A Forecasting Methodology as Applied to Financial Series, *The Accounting Review*, C. 49, January, p. 61.
- Malekipirbazari, M., Aksakalli, V. and Aydogdu, V., 2015. A temporal analysis of vessel type traffic in Istanbul Strait anchorages. 2015 International Conference on Industrial Engineering and Operations Management (IEOM). Available at: <http://dx.doi.org/10.1109/ieom.2015.7093878>.
- Ministry of Transport and Infrastructure Maritime Statistics, Data Statistics regarding Vessel Passing Through Istanbul Strait, 2020. Available at: https://atlantis.udhb.gov.tr/istatistik/gemi_gecis.aspx, accessed on:18 Oct 2020.
- Mısır, D. S., 2008. Gemi Trafik Hizmetleri-VTS, Sümae Yunus Araştırma Bülteni, 8, pp. 11-15.
- Mostafa, M.M., 2004. Forecasting the Suez Canal traffic: a neural network analysis. *Maritime Policy & Management*, 31(2), pp.139–156. Available at: <http://dx.doi.org/10.1080/0308883032000174463>.
- Oral, N., and Ozturk, B., 2006. The Turkish Straits, Maritime Safety, Legal and Environmental Aspects, Turkish Marine Research Foundation, Publication Number 25.
- Ozbaz, B., 2005. Simulation of Maritime Transit Traffic in the Istanbul Channel, Department of Industrial Engineering, Boğaziçi University.
- Ozcelik, S., 1980. İktisadi Zaman Serilerinde Tahmin Yöntemleri İstanbul Ticaret Odası Toptan Esya Fiyatları İndeksi Üzerine Bir Uygulama, Doçentlik Tezi, Erzurum, sf. 55.
- Perron, P., 1989. The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis. *Econometrica*, 57(6), p.1361. Available at: <http://dx.doi.org/10.2307/1913712>.
- Praetorius, G., Hollnagel, E. and Dahlman, J., 2015. Modelling Vessel Traffic Service to understand resilience in everyday operations. *Reliability Engineering & System Safety*, 141, pp.10–21. Available at: <http://dx.doi.org/10.1016/j.res.2015.03.020>.
- Simsir, U. and Ertugrul, S., 2009. Prediction of manually controlled vessels' position and course navigating in narrow waterways using Artificial Neural Networks. *Applied Soft Computing*, 9(4), pp.1217–1224. Available at: <http://dx.doi.org/10.1016/j.asoc.2009.03.002>.
- Simsir, U. et al., 2014. Decision support system for collision avoidance of vessels. *Applied Soft Computing*, 25, pp.369–378. Available at: <http://dx.doi.org/10.1016/j.asoc.2014.08.067>.
- Smarten, 2018. What is ARIMAX Forecasting and How is it Used for Enterprise Analysis, *ElegantJ BI*. Available at: <https://www.elegantjbi.com/blog/what-is-arimax-forecasting-and-how-is-it-used-for-enterprise-analysis.htm>.
- TSVTS Center, 2019. Instructions for the Implementation of the Rules Regulating the Turkish Straits Maritime Traffic Scheme.
- Uğurlu, Ö., Erol, S. and Başar, E., 2015. The analysis of life safety and economic loss in marine accidents occurring in the Turkish Straits. *Maritime Policy & Management*, 43(3), pp.356–370. Available at: <http://dx.doi.org/10.1080/03088839.2014.1000992>.
- Uluşçu, Ö.S. et al., 2008. Transit Vessel Scheduling in the Strait of Istanbul. *Journal of Navigation*, 62(1), pp.59–77. Available at: <http://dx.doi.org/10.1017/s0373463308005092>.

Vandekerckhove, J., Matzke, D. & Wagenmakers, E.-J., 2015. Model Comparison and the Principle of Parsimony J. R. Busemeyer et al., eds. Oxford Handbooks Online. Available at: <http://dx.doi.org/10.1093/oxfordhb/9780199957996.013.14>.

Vessel Traffic Services Manual, 2012. International Association of Marine Aids to Navigation and Lighthouse Authorities, 5. Ed.

Wen, Y. et al., 2015. Modelling of marine traffic flow complexity. Ocean Engineering, 104, pp.500–510. Available at: <http://dx.doi.org/10.1016/j.oceaneng.2015.04.051>.

Xiao, Z. et al., 2017. Maritime Traffic Probabilistic Forecasting Based on Vessels' Waterway Patterns and Motion Behaviors. IEEE Transactions on Intelligent

Transportation Systems, 18(11), pp.3122–3134. Available at: <http://dx.doi.org/10.1109/tits.2017.2681810>.

Yahsi, M., and Ugurlu, O., 2014. Gemi Trafik Hizmetleri ile Hava Trafik Hizmetlerinin Karşılaştırılması, I. Ulusal Gemi Trafik Hizmetleri Kongresi, December.

Yoo, S.-R. et al., 2013. Forecast of marine traffic volume using time series model. 2013 International Conference on Fuzzy Theory and Its Applications (IFUZZY). Available at: <http://dx.doi.org/10.1109/ifuzzy.2013.6825421>.

Zhou, X. et al., 2020. Using Deep Learning to Forecast Maritime Vessel Flows. Sensors, 20(6), p.1761. Available at: <http://dx.doi.org/10.3390/s20061761>.