sciendo

Time Time Series Forecasting of Stock Price for Maritime Shipping Company in COVID-19 Period Using Multi-Step Long Short-Term Memory (LSTM) Networks

Ahmad GHAREEB
Bucharest University of Economic Studies, Bucharest, Romania[1]
[*]*Corresponding author, ahmadghareeb18@stud.ase.ro*

Abstract. *Sea transportation is one of the most critical components of the global economy, where it is the main means of transporting huge quantities of goods across the world. The coronavirus disease (COVID-19) has caused a health-related economic crisis at the global level with widespread repercussions on sea transportation and trade. Several algorithms in machine learning and/or deep learning have been suggested as a simplification of this hard problem of stock price index forecasting, which has been the focus of much study for many years. In this study, I will propose a forecasting method based on Multi-Step Long Short-Term Memory (MS LSTM) networks to predict stock prices for three of the most important companies in the world in maritime transport. A novel optimization method for stock price prediction is proposed. It is based on a Multi-Step Long Short-Term Memory (MS LSTM) model and utilizes the Adam optimizer. Conclusions showed that using the MS LSTM algorithm, 95.44% prediction accuracy on the training data set and 95.11% prediction accuracy on the testing data set were both obtained. For the training set and testing set, respectively, it was observed that the mean absolute percentage error is 4.56% and 4.89%. By running the predictions for five days in the future, it is observed that the model has predicted that the prices would maintain their balance with little downward movement, leading to positive impact on sea shipping with lower prices expected.*

Keywords: COVID-19, Maritime transport, Long short-term memory, Predicting, Stock price, Shipping company.

Introduction

Sea transportation is one of the most critical components of the global economy, where it is the main means of transporting huge quantities of goods across the world. It supports the supply chains and economic connectivity with cargo and seaports, which are estimated to handle more than 80% of global merchandise in volume and more than 70% in value (UNCTAD, Mar 2021).

The coronavirus disease (COVID-19) has caused a global health and economic crisis with widespread repercussions on sea transportation and trade. The disruptions caused by the imposed restrictions in response to the pandemic have affected ports, shipping, and supply chains. Many industries have faced challenges along their supply chain such as raw material shortages, deadlines and time issues, sailing empty oceans, port closures, reduced port hours, equipment, labor shortages, as well as truck/transport capacity constraints.

The smooth movements of trade flows and supply chain operations were undermined by these impediments, with significant possibility to undermine transport services trade liberalization and trade facilitation gains made over the years (UNCTAD, Mar 2021). Such abrupt and widespread disruptions caused by the pandemic have also significantly increased the cost of shipping. This increase in costs was driven by a variety of factors, including supply chain

[1] *Doctoral School of Cybernetics and Economic Statistics, Faculty of Cybernetics, Statistics and Economic Informatics*

DOI: 10.2478/picbe-2023-0154
© 2023 A. Ghareeb, published by Sciendo.
This work is licensed under the Creative Commons Attribution 4.0 License.

disruptions, changing consumer behavior, and changing priorities for shipping companies (UNCTAD, Jun 2022).

Supply chain disruptions caused by the pandemic have led to a decrease in the availability of shipping containers, ships and other resources needed for transportation (UNCTAD, Apr 2021). This decrease in capacity increased demand for limited resources available, driving prices up. In addition, the pandemic has led to changes in consumer behavior, with many people choosing to buy goods online rather than in stores. This shift in purchasing behavior has increased demand for shipping services, resulting in higher prices. The growing demand for personal protective equipment (PPE) and medical supplies has led to increased demand for shipping services, which has led to increased costs.

Another factor contributing to the increase in the cost of shipping is that the COVID-19 pandemic has also increased operating costs for shipping companies. Shipping lines were forced to implement measures to protect the health and safety of crew members, resulting in increased costs of personal protective equipment, medical supplies, quarantine requirements, and additional cleaning and disinfection procedures. The pandemic has also led to a shortage of seafarers, with many crew members unable to disembark at ports due to movement restrictions and border closures. This increased the costs of shipping companies, as they had to pay higher wages to retain crew members and recruit new sailors.

Despite these challenges, the shipping industry has remained resilient in the face of the pandemic. Shipping companies have adapted to changing circumstances by applying new technologies, such as digital platforms for communication and tracking, exploring alternative modes and modes of transportation. The industry has also been supported by government initiatives, such as the creation of special corridors for essential commodities, which has helped reduce disruptions and ensure that important goods continue to reach their destinations.

Recurrent Neural Network (RNN) comes in several types; Long Short-Term Memory (LSTM) is a type that is specifically designed to handle sequential data, such as time series, speech, and text. LSTM networks demonstrate capability of learning long-term dependencies in sequential data, providing conducive contexts for tasks such as language translation, speech recognition, and time series forecasting.

Understanding the way specific metric changes over time requires deploying time series analysis in numerous contexts. As such, it is highly suitable for non-stationary data with regards to the regularly fluctuating things over a specific time period, such as prices of goods and commodities. Given that sales and currency values always vary, their time series analyses are regularly used in the fields of finance, economics, and retail industries.

Sequence prediction problems, as some of the hardest-to-solve problems in the data science domain, have been a subject of debate for a long time, ranging on a wide spectrum from predicting sales to identifying patterns of in data of stock markets, and from language translations to predicting the next possible word on the keyboards of smart phone.

The recent breakthroughs witnessed in data science have revealed that the most effective solution comes in the forms of Long short-Term Memory networks, such LSTMs for almost all of these sequence prediction problems.

Given their selectivity feature in recalling patterns for long time durations, LSTMs have, in many ways, an advantage over conventional feed-forward neural networks and RNN.

Taking a particular stock in the stock market data as an example of sequential data, any given Artificial Neural Network or simple model of machine learning has the possibility to learn predicting the prices of the stock due to several features: the stock volume, the opening value etc.

While dependent on these features, the stock price largely depends on the recent stock values as well. As far as the trader is concerned, these values, in fact, in the recent ones (or the trend) make up a crucial deciding factor for predictions.

With regards to the conventional feed-forward neural networks, all test cases in are deemed to be independent. In other words, when fitting the model for a particular day, the recent stock prices are not considered

RNN can remember things for short intervals/periods of time, i.e., if we need the information after a short interval, it could be reproducible; however, inputting a lot of information in causes this information to be lost somewhere. When a slightly tweaked version of RNNs – the Long Short-Term Memory Networks- is applied, it helps in resolving this issue. LSTMs use multiplications and additions to introduce minor modifications to the information. The cell states represent the information flows with LSTMs, demonstrating their selectivity in remembering or forgetting things.

The two states, known as the cell state and the hidden state, must be transferred to the next cell. Remembering things takes place through the memory blocks; however, the gates represent the three major mechanisms (Forget Gate, Input Gate, and Output Gate), that are responsible for manipulations to this memory.

The below sections of the research paper are arranged as follows: Section 2 illustrates literary studies. In Section 3, the concept of LSTM is introduced, and its functions are described. The proposed structure and methodology for estimating stock prices are presented in Section 4. And in section 5 conclusions and future directions are presented.

Literature review

The importance of data science has been recently increasing given the growing sophisticated relationship between the crises and the relevant data in terms of measuring, analyzing, and understanding the impacts of crises on countries' economies and peoples' lives.

Developing the appropriate models and tools for collecting the crisis-relevant data would always help identify crisis and data patterns and correlations, and provide evidence-based mechanisms for crisis predictability, leading eventually to better crisis response and management, as well as mitigation of crisis impacts on economies and communities.

The sections below sum up some important cases and examples of data generated in connection with the Covis-19 pandemic and its various impacts on several economic sectors in many parts of the world:

In their studies, Masha Menhat et al published in April 2021 the results of examining the impact of the Covid-19 pandemic on four maritime sectors: oil and gas, shipping, maritime tourism, and fisheries sectors, The key conclusion was that services sectors such as maritime tourism is the most affected, owing to the fact that it is considered a non-essential service, hence experiences total close down during the movement control order (MCO) until the conditional movement control order (CMCO) period. However, economic initiatives and domestic tourism may mediate the post-pandemic impacts. Based on the four sectors involved, shipping was considered less affected in comparison to other sectors, considering the high demands of personal protective equipment (PPE) products and test kits during the pandemic. The oil and gas sector on the other hand, received no sector-specific incentives despite being the major economic contributor.

In a relevant study published in May 2021, Prathvi Thumbe Narasimha et al examined the impacts of COVID-19 on seaport transportation and its supply chain field; and its related issues in India; the Secondary data was used to analyze the performance indicators of major seaports in India

before and during the COVID-19 crisis; the key conclusion was that there had been a decline in total cargo traffic volume and a decrease in the majority of segment-wise cargo traffic volume and a reduction in vessel call performance during the COVID-19 period compared to the period before the COVID-19. Also, it was observed that there is a negative impact of COVID-19 on the maritime-related sector with a low level of preparedness and adaptability to handle the pandemic situation.

David Guerrero et al (November 2021) have analyzed the changes on the maritime network before and after the Covid-19 outbreak. They could demonstrate a decrease in the global maritime connectivity and significant differences between ports and inter-port links by considering a significant sample of representative vessel movements between ports. Moreover, they found that the mitigation measures implemented by governments in relation to the impacts of the Covid-19 varied in relation to regional port hierarchies, leading to a drop in port concentration in Europe and Africa compared to an increase in Asia and North America. Globally, various levels of resistance-to-crisis were witnessed in both very large ports and small but densely inter-connected ones; however, the negative impact was significant on small transshipment hubs and bridges.

Alongside the sector specific analysis, a thorough analysis using "an event study approach" was conducted by Gavalas, Syriopoulos and Tsatsaronis (2022) to examine the impact of the COVID-19 pandemic, identifying how shipping freight rates and stock values could be influenced by 'a sanitary incident', investigating how fast and comprehensive shipping markets reacted upon certain latest evidence by using the approach of market-model event study. According to the key date set as the event window, the conclusions of the analysis revealed that different results appeared reflecting and proving how resistant to pandemic the dry market, the tanker market, and the shipping stock market.

Yilmazkuday (2022) investigated, using daily data on the COVID-19, cases from China and the rest of the world, and their corresponding effects on the global economic activity. The positive correlation was demonstrated by empirical results reached, based on a structural vector autoregression model using the Baltic Exchange Dry Index (BDI) and crude oil prices (COP), which were in line and in harmony with increases in COVID-19 cases acting as negative demand shocks in the global economic activity (reflected as reductions in COP) and negative supply shocks in the global commodity transportation (reflected as increases in BDI). A parallel observation revealed that the historical decomposition results also suggest that there was correlation between the effects of early COVID-19 cases and the impacts on COP and BDI.

Zhao, He, Lu, Han, Ding and Peng (2022) presented their respective conclusions on the impact of the COVID-19 pandemic outbreak on a range of shipping sectors (such as containers and bulk), where global economic growth was significantly restrained by various lockdown strategies, leading to a remarkable decline in maritime transportation. The researchers conducted a series of analyses of the Baltic Dry Index (BDI), the China Coastal Bulk Freight Index (CCBFI), and of container throughputs, in relation to the impact of COVID-19 in their study to measure the changing trends and evaluate their impacts on dry bulk and container transportation, where the lockdown policies' results revealed that global dry bulk transportation was largely affected by the second month during COVID-19; and the BDI presented a yearly decrease of approximately 35.5% from 2019 to 2020. The container throughputs at the ports of Shanghai, Hong Kong, Singapore, and Los Angeles from 2019 to 2020 presented the largest year-on-year drops of approximately 19.6%, 7.1%, 10.6% and 30.9%, respectively. Drawing on the findings of this study, another conclusion was reached indicating that policies on smart shipping and special support could be used to mitigate the negative impacts of COVID-19.

A set of key conclusions were published in November 2022 Hideaki Sakawa et al based on examining the market reaction of Japanese listed shipping stocks during the early stage of COVID-19 pandemic:

- The negative effects of two pessimistic events such as Princess Diamond (February 3, 2020) and tightened border control of Japanese government (March 9, 2020). This implies that two pessimistic events negatively affect stock returns on Japanese listed shipping firms.
- The negative announced effect of Port Operation is larger than Maritime Transportation after 30 trading days.
- The government responses such as Stringency Social Distance Policy, would positively affect the investors in the Japanese shipping industry under the tightened border control.

In this sense, government policy responses such as non-pharmaceutical interventions are expected to mitigate the negative announcement effect on pessimistic COVID-19 related news in the shipping industry.

In December 2022 Yigit Kazancoglu et al studied the causal relations of trade parameters, demographic data and logistics, supply chain parameters data on disruptions of global supply chains have been depicted for Turkey and China since China is the biggest importer of Turkey. The key conclusions were:

- For the mining and chemical export from Turkey to China, there is two decreases in export rate of Turkey caused by COVID-19 disruption firstly, and then slowdown in industrial capacity of Turkey after April. In addition, while mining sector has not been impacted by Turkey COVID-19 disruption, chemical sector has been affected.
- Besides, it can be stated that a large part of the import and export with China is planned with maritime logistics but import shipments have been decreased due to the pandemic. This has led to an increase in freight costs in Turkey due to the lack of suitable ships in Europe.

Furthermore, according to forecasting, it was foreseen that freight costs would return to their original position by the end of June 2020, indicating a potential decrease in import and export shipments.

In April 2018, Mei-Ping Chen et al examined the effect of the severe acute respiratory syndrome (S.A.R.S.) epidemic on the long-run relationship between China and four Asian stock markets (Hong Kong, Taiwan, Singapore, and Japan) in the period 1998–2008, they investigated the impact of the S.A.R.S. epidemic on stock market integration based on the extreme value theory. The key conclusion was that some investment opportunity heterogeneity was caused by the epidemic, leading to an inefficient international risk-sharing environment. Thus, investors used portfolio diversification to make arbitrage profits among these five Asian countries in consideration of detrimental epidemic diseases.

In September 2022, Meihui Zhong et al summarized studies related to the economic development during the COVID-19 epidemic. They analyzed the COVID-19 pandemic from both qualitative and quantitative perspectives. they searched in Web of Science Core Collection using the subjects, such as “economics”, “economy”, “economic”, “financial”, and then 2274 related documents were collected, which are published from 2020 to 2022; the key conclusions were:

- From the cluster analysis, it showed that the research direction of COVID-19 and the economy is roughly in the global financial crisis and the policy of adjusting the current economic situation. Among them, three countries, namely the United States, China, and Britain, were most concerned about this field.
- It was found that there is more regional cooperation. And there was a lack of knowledge sharing, but overall, there was a tendency for global cooperation.

- Through the analysis of literature content, they found that the research topics on the economy during the epidemic period focused on three aspects, namely, economic impact, the effect of policy intervention, and the response of financial markets to the epidemic.

In October 2022, Bernard NjindanIyke et al examined whether the pandemic adversely impacted stock markets and whether monetary policy – specifically, conventional monetary policy – had moderated this impact by using daily historical data. The key conclusion was that although conventional monetary policy was important, it might not be an effective tool if implemented alone during the ongoing pandemic. To steer stock markets back to their pre-crisis level, unconventional monetary policy and other policies were necessary. Looking ahead, for economies unconstrained by the zero-lower bound of policy rates, small but significant interest rate cuts to ease the cost of borrowing should be coordinated with other policies (like forward guidance and quantitative easing) to restore stock markets to at least their pre-shock levels and to prevent a full-blown financial crisis.

In a nutshell, the several studies conducted on various sectors revealed some significant conclusions based on the generated and collected data using several tools and models.

Long short-term memory (LSTM)

Being an extension of RNN, LSTM has a strong forecasting capability in time series data.

The key difference between LSTM and an RNN is that LSTM can store long-range time dependency information, and can suitably map between input and output data (Abbasimehr et al.,2020).

The LSTM network structure differs from the conventional perceptron architecture because it has a cell and gates that control the information flow. The LSTM, namely, contains an input gate, a forget gate, internal state (cell memory), and an output gate as illustrated in Fig.1.

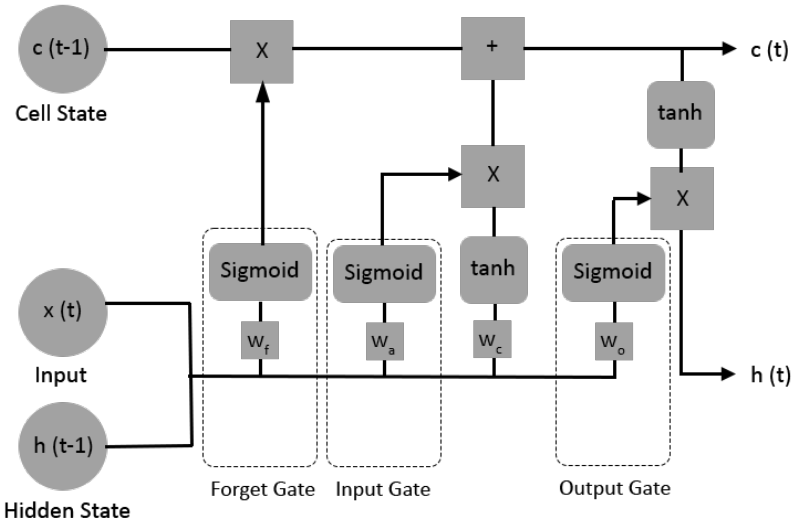


Figure 1. The structure of LSTM

Source: (Abbasimehr et al.,2020).

The notations are explained as follows:

$x(t_i)$ The input value

Cell state $c(t_i)$ - Represents the internal memory of the cell which stores both short-term memory and long-term memories.

Hidden state $h(t_i)$ - This is output state information that calculated w.r.t. current input, previous hidden state and current cell input which are used to predict the future stock market prices. Moreover, the hidden state can decide to only retrieve the short or long-term or both types of memory stored in the cell state to make the next prediction.

Forget Gate *f(ti)* - it determines which current and previous information is retained and which is discarded. including the hidden status from the previous run as well as the current status. These values are fed into a sigmoid function, which can only output values between 0 and 1. The value 0 means that all previous information is forgotten and 1 accordingly that all previous information is retained. The results from this are multiplied by the current Cell State, so that the no longer needed knowledge is forgotten, as it is multiplied by 0 and thus dropped out.

Input Gate *a(ti)* - it decides how valuable the current input is to solve the task. For this purpose, the current input is multiplied by the hidden state and the weight matrix of the last run. All information that appears to be important in the Input Gate is then added to the Cell State and forms the new Cell State *c(ti).* This new Cell State is now the current state of the short-term memory and will be used in the next run.

Output Gate *o(ti)* - the output of the LSTM model is then calculated in the Hidden State. Depending on the application, the sigmoid function decides what information can come through the output gate, and then the cell state is multiplied after it is activated with the tanh function (Abdul Quadir Md et al.,2022).

Mathematically, the forward learning of an LSTM is as follows:

$$a(t_i) = \sigma(w_a\, x(t_i) + w_{ha} h(t_{i-1}) + b_a) \quad (1)$$

$$f(t_i) = \sigma(w_f x(t_i) + w_{hf} h(t_{i-1}) + b_f) \quad (2)$$

$$c(t_i) = f_t \times c(t_{i-1}) + a_t \times tanh(w_c x(t_i) + w_{hc}(h(t_{i-1}) + b_c) \quad (3)$$

$$o(t_i) = \sigma(w_o x(t_i) + w_{ho} h(t_{i-1}) + b_o) \quad (4)$$

$$h(t_i) = o(t_i) \times tanh(c(t_i)) \quad (5)$$

where σ and $tanh$ are activation functions, $\times$ indicates point-wise multiplication, w, wh are weight matrices, and b is bias vectors.

Overall, the LSTM learns using the following steps:

(1) Using Eqs. (1) through (5) to calculate the LSTM output (forward learning).
(2) Calculate the error between the resulted data and input data of each layer
(3) The error is reversely propagated to input gate, cell, and forget gate.
(4) Regarding the error term, an optimization algorithm updated the weight for each gate.

The best weights and biases are obtained by repeating the four steps of the above-mentioned process for a given number of iterations. (Abbasimehr et al., 2020).

Methodology and model:

Dataset:

The required datasets are downloaded from Yahoo Finances. The datasets in use contain the stock information of (COSCO SHIPPING Holdings Co. Ltd. A, Evergreen Marine Corp. (Taiwan) Ltd., Yang Ming Marine Transport Corp.) dating from 2018 to 2023. The dataset contains the date, open, high, low, close, adjusted close and volume.

- The Date column indicates the date of the stock price.
- The information in the Open column refers to the price at which the stock opened for the day.
- The High column displays the stock's highest price for the day.

• The Low column displays the stock's lowest price for the day.
• The Close column reflects the information about the price when the stock was closed.
• Adj Close value adjusts for the corporate activities.
• The volume column shows how many shares of each company were traded on a certain day.

The Close and the Date columns in the aforementioned are the most interesting ones. This is because the MS LSTM model is used for performing date-wise training and prediction of the closing stock price.

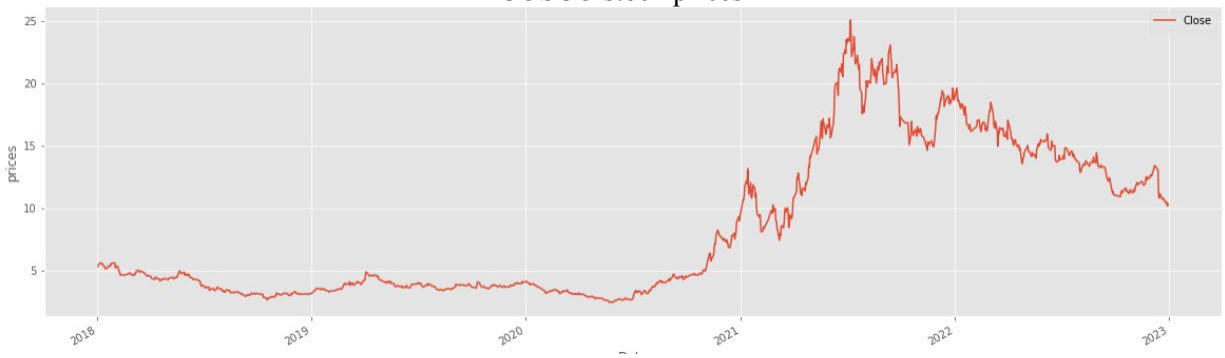
Cleaning data:

Handling missing data is the first step to start with when embarking on cleaning data, simply by scanning each row of the dataset to check for missing values, and clearly observing the absence of NA values in the columns Open, High, Low, Close, Adj and Volume.

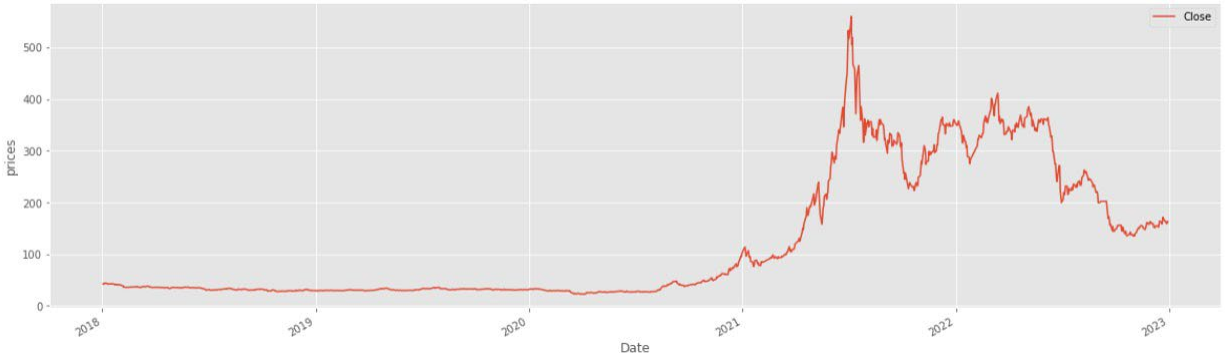
Visualizing data:

Then comes the next step of plotting the closing price of the stock over the time period of the dataset, using the x-axis to plot the date, and the y-axis to plot the closing price of the stock. Fig. 2 shows the line plot demonstrating a fair idea about the progression of the value of the COSCO, Evergreen, Yang Ming stock price over time starting from January 2018 to January 2023. Following is the extraction of the maximum and minimum closing prices of the stock along with the date of the valuation from the data and five last values are shown in Table 1.

COSCO stock prices



Evergreen stock prices



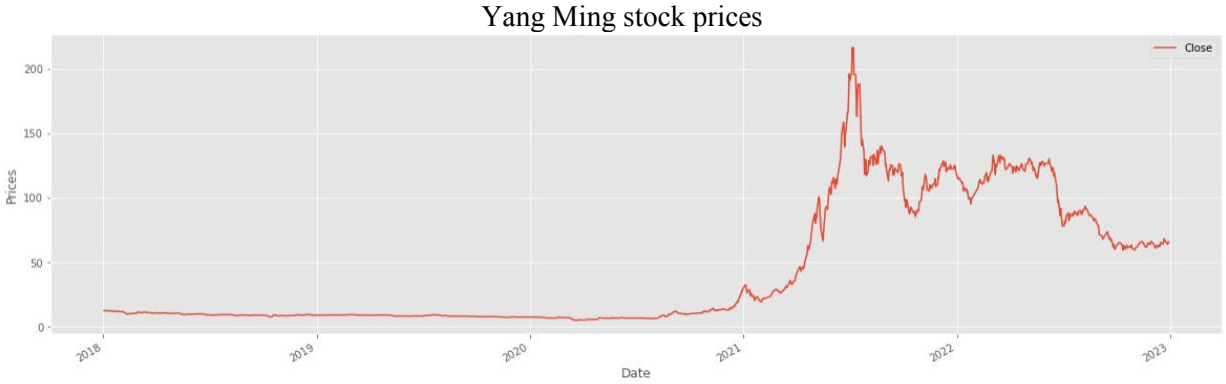


Figure 2. value of the COSCO, Evergreen, Yang Ming stocks price over 5 years

Source: Authors' own research.

Table 1. Last 5 days of the COSCO, Evergreen, Yang Ming stocks price

	COSCO	Evergreen	Yang Ming
2022-12-26	10.40	162.50	65.00
2022-12-27	10.48	162.50	65.01
2022-12-28	10.40	159.00	63.90
2022-12-29	10.14	160.00	64.80
2022-12-30	10.29	163.00	65.50

Source: Authors' own research.

It is observed that the lowest price of the COSCO stock was 2.41 on 22-05-2020, and 22.75 in 01-04-2020 for Evergreen, and 4.73 in 19-03-2020 for Yang Ming respectively in synchronization with the period when COVID-19 start to go global. The maximum valuation of the COSCO stock was observed on 07-07-2021 with a closing price of 25.07, and 560 in 06-07-2021 for Evergreen, and 216.5 in 08-07-2021 for Yang Ming, meanwhile the period in which the prices of transport containers were the highest in the time of Corona, recording as of the thirtieth of July, the estimated price of transporting a 20-foot container from Shanghai-EC America (base port) was at $9720; a 40-foot container from Shanghai-S America (Santos) was at $10067; and from Shanghai-W Africa (Lagos) a 20-foot container was at $8,102 (UNCTAD, Nov 2021).

Data normalization:

Feature scaling is performed on the dataset after visualizing it, which helps in normalizing all the data in the dataset using Eq. (6) which may fall between different ranges, converting them into a new range of any given choice, envisaged usually 0 to 1. Non-normalization of the data may lead to large gradient error values, yielding very large, and thus unstable, weight values for the LSTM model. MinMaxScaler, available in the sklearn preprocessing library, is used to normalize the data. It helps in scaling all the columns of the data between the values of 0 and 1. Preserving the shape of the dataset, without any distortion, is the reason behind choosing MinMaxScaler. The following equation yields the transformation:

$$x' = \frac{(X - Xmin)}{(Xmax - Xmin)} \quad (6)$$

where x is an original value, and x' is the normalized value

Table 2 presents a fragment result of the original data and the normalized data

Table 2. Data normalization using rescaling Max-Min

COSCO		Evergreen		Yang Ming	
original	normalized	original	normalized	original	normalized
10.40	0.35234215	162.50	0.26012099	65.00	0.28460122
10.48	0.35587236	162.50	0.26012099	65.01	0.28507342
10.40	0.35234215	159.00	0.25360633	63.90	0.27940691
10.14	0.340869	160.00	0.25546766	64.80	0.28365681
10.29	0.34748813	163.00	0.26105165	65.50	0.28696227

Source: Authors' own research.

As the process unfolds, the data is split into training and testing sets. The data is split into the above-mentioned parts for last 20 records from the data for testing data and the rest for training data.

Reshaping data:

The data is converted into time steps itself to make it suitable for the LSTM model. The step value is taken to be 30. then, the training data is reshaped to fit the intended LSTM model to be built. The reshape function combines three parameters, namely, the sample size, the time step and the number of features. For the proposed LSTM algorithm, the sample size is the number of rows in the training set, which is 1158 for COSCO dataset and 1165 for Evergreen and Yang Ming datasets. As already mentioned, the time step is 30. Finally, the feature size of the data is going to be 1, because only one feature, the closing price of the stock, is being predicted.

Building MS LSTM model:

The prediction of stock prices is realized in the next step aimed at building the MS LSTM model for this function, setting the type of model to be sequential, meaning that the data is fed to all the layers of the LSTM model in a sequential way. The intended MS LSTM model consists of 5 layers, revealed as the optimal number of layers for accurate predictions while avoiding overfitting. The first is the LSTM model to be imported from the Keras library. The input shape of the layer is defined as (30,1) because of the chosen time-step of 30, indicating that each input is of size 30 and there is 1 feature column that is set as the input, setting the number of LSTM cells as (60). The second is also an LSTM layer of shape (30); and in the same way, the third is also an LSTM layer with shape (30), the fourth is also an LSTM layer with shape (30). The final layer is the dense layer of shape (5) which is the number of days to be predicted. This indicates that all the output layers of the previous layer are densely connected (fully connected) to the output layer of the model. The optimizer used is adam, and the loss function is mean squared error.

Model Compilation and Fitting:

Once the model has been built, it is compiled with a batch size of 5 and an epoch size of 100, signifying the grouping of data together into batches of 5 and fed into the LSTM model. 100 epochs correlates that the data is fed to the model 100 times. Following the training, training data is used for making predictions. The predicted data needs to be normalized back to the normal scale, which is done using the inverse transform () function. Once the data has been inverse-scaled, the accuracy measures can be obtained by comparing the inverse-scaled data with the original data, which is, then, used to gauge the performance of the model MAPE.

The Setup of the Experiment:

The experiments were carried out on an i7-6500U Intel Core CPU with a 2.50 GHz machine with 12 GB ram and were implemented in Python 3.9.7 using Jupyter Notebook server 6.4.5. The datasets used in the experiment are that of COSCO, Evergreen, and Yang Ming stock prices recorded between 01 January 2018 and 01 January 2023, covering a range over five years. The dataset was obtained from the website of Yahoo Finances. Only the Date column and the Closing price column are selected out of the seven columns in the dataset, for further steps, where they don't have missing values. The MinMax scaling technique has been used to ensure the efficient training of the model by normalizing the used data. The training set and testing set are converted into time steps of 30. The input data is then reshaped as [samples, time steps, features].

A sequential LSTM model comprised of four stacked LSTM layers is used for the prediction. This is done with the help of Keras deep learning API that runs on top of TensorFlow. Finally, a dense layer is added to make the output a 5 value. During the compilation of the model, the mean squared error is used as the loss function, and the optimizer is adam. The training of the model is performed for 100 epochs with 5 being the batch size. The compiled model has been used for the predictions on training data. Once the predictions have been made, the original values and the predicted values are inverse transformed back to their original scale.

Results:

The training as well as the testing data are used for forecasting the stock prices and are compared with the original values. The accuracy of the proposed algorithm is measured by a set of metrics; namely: the Mean Absolute Percentage Error (MAPE), the mean absolute error (MAE), the Mean Square Error (MSE), and the Root Mean Square Error (RMSE), The first metric used (MAPE) gives information about how accurate a prediction model is. It is measured in percentages, and its value can be found by calculating the average modulus of the actual value of observation minus forecasted observation divided by the actual value of observation, as shown in Eq. (7).

$$MAPE = \frac{1}{n} \sum_{1}^{n} \left| \frac{A_t - F_t}{A_t} \right| \quad (7)$$

Where At is the actual value, Ft is the forecast value, and n is sample size.

The mean absolute error (MAE) represents the average of the absolute value of the deviation of all single observations, which can avoid the problem of mutual cancellation of errors and accurately reflect the size of the actual forecast error as shown in Eq. (8).

$$MAE = \frac{1}{n} \sum_{1}^{n} \left| X_t - \grave{X}_t \right| \quad (8)$$

Where X_t is the actual value, $\grave{X}_t$ is the forecast value, and n is sample size.

The Mean Square Error (MSE) represents the mean of the sum squares of the errors between the predicted data and the original data, as shown in Eq. (9):

$$MSE = \frac{1}{n} \sum_{1}^{n} (X_t - \grave{X}_t)^2 \quad (9)$$

The Root Mean Square Error (RMSE) defines the expected value of the square of the error caused by the predicted value and the true value, and its range is 0 to infinity. When the predicted value is completely consistent with the true value, the value of RMSE is equal to 0, which is a perfect model. On the contrary, when the prediction error is larger, the value of the RMSE will be larger, as shown in Eq. (10):

$$RMSE = \sqrt{\frac{1}{n} \sum_{1}^{n} (X_t - \acute{X}_t)^2} \quad (10)$$

The average Mean Absolute Percentage Error (MAPE) of the three stocks of shipping companies that gained from the model on training data is 4.56 and on test data is 4.89, which means that its performance accuracy is 95.44% on training data and with an accuracy of 95.11% on testing data.

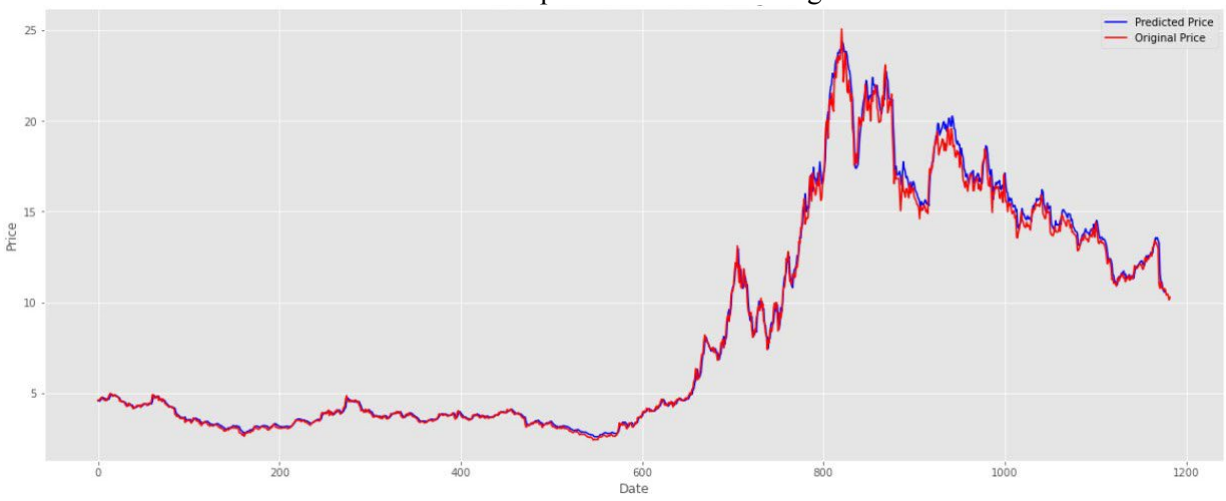
The actual stock price on a given date and the predicted stock price are tabulated in Table 3 and visualized in Fig. 3 to highlight the accuracy of forecasting by MS LSTM. It can be observed that there is a very high overlap between the actual stock price and the stock prices predicted by the model, which means that the predictions are highly accurate.

Table 3. Comparison of Original stock price with the predicted stock price for COSCO, Evergreen, Yang Ming

COSCO		Evergreen		Yang Ming	
original	predicted	original	predicted	original	predicted
15.03999996	15.33951	353.75	353.09158	123.00	121.36572
14.64000034	15.347001	358.75	355.59857	123.50	122.73707
14.77000046	15.033326	363.75	358.8626	125.00	123.43644
13.56999969	15.126878	368.75	363.70413	126.50	124.53178
13.56000042	14.167503	348.75	367.26535	121.50	125.85691

Source: Authors' own research.

COSCO stock price Prediction vs Original



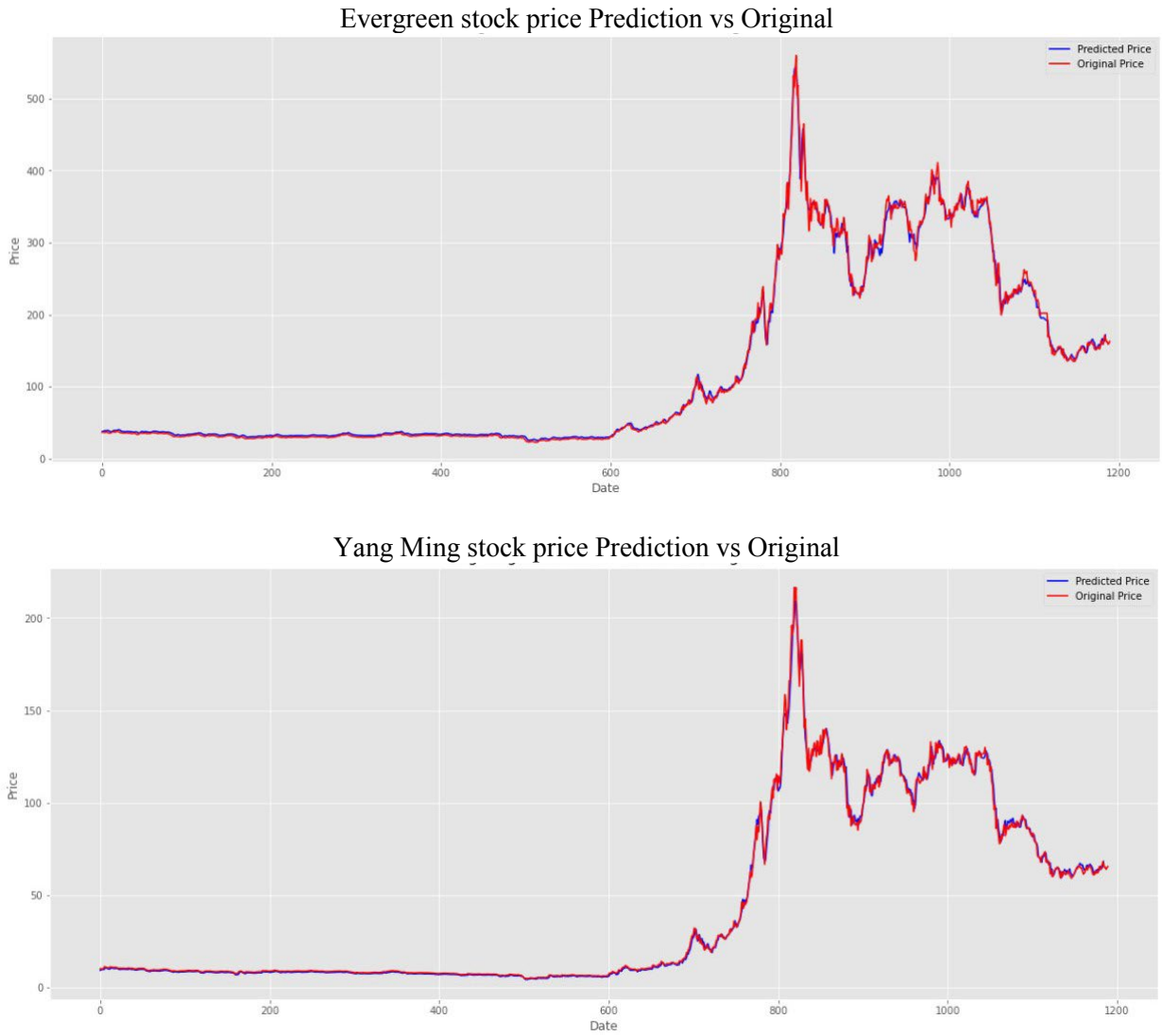


Figure 3. Original and predicted values for COSCO, Evergreen, Yang Ming stocks price

Source: Authors' own research.

Following that, I will run the predictions for five days in the future by inputting the stock prices for the last 10 days of trading into the model; it can be observed that the model yielded the prediction and calculated it the prediction presented in Fig. 4 and Table 4 for the respective companies (COSCO, Evergreen, Yang Ming). By comparison, it is observed that the model has predicted that the prices would maintain their balance with little downward movement, leading to positive impact on sea shipping with lower prices expected; however, the stock prices may not be predicted for a long time given the various variables that may affect the stock price, especially the stocks of transportation/shipping companies, simply because they are connected with some political or logistical problems, or due to turmoil and troubles inside the companies, let alone the wars and natural disasters, including pandemics, as clearly witnessed during the Corona pandemic that caused huge volatilities in transportation/shipping companies' stock prices.

Table 4. Predicted values for 5 days for COSCO, Evergreen, Yang Ming

COSCO	Evergreen	Yang Ming
10.166268	165.748779	63.051708
10.289643	162.912659	64.180527
10.135217	160.541748	64.996010
10.242349	164.822418	64.538300
10.387588	164.824203	63.632236

Source: Authors' own research.

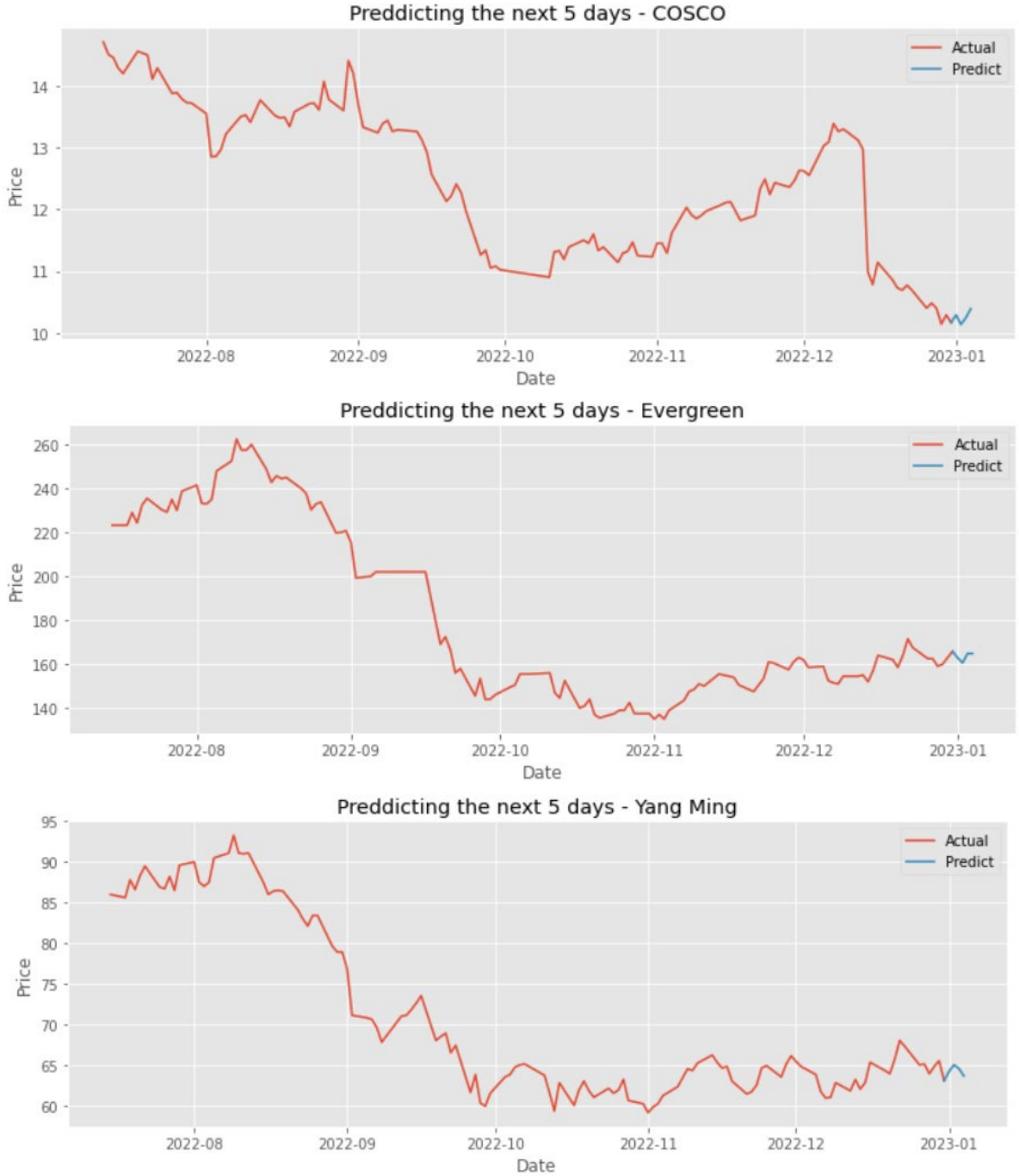
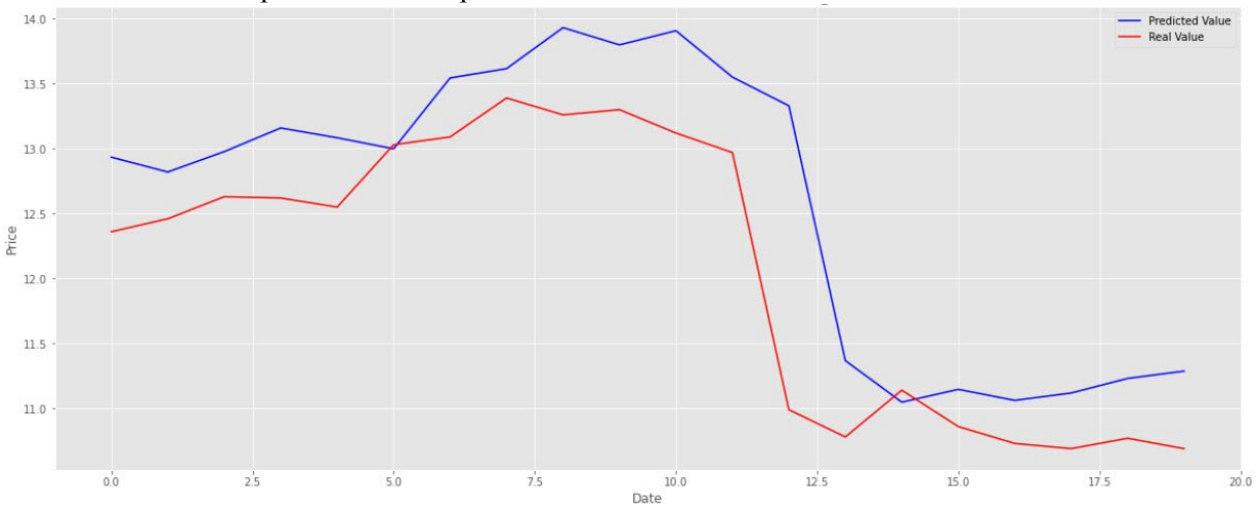


Figure 4. Predicted values for 5 days for COSCO, Evergreen, Yang Ming

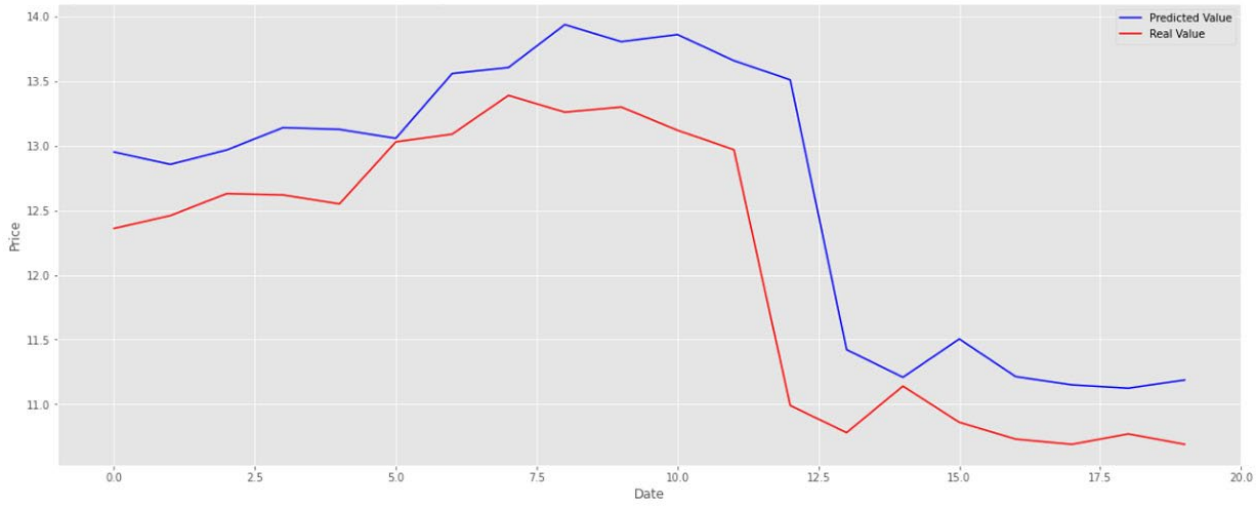
Source: Authors' own research.

Verifying the effectiveness of the proposed model requires comparing different models including RNN, GRU, and MS LSTM with regards to the three stock datasets (COSCO, Evergreen, Yang Ming). The processed training set data to train RNN, GRU, and MS LSTM, respectively were used; then, the model completed by training is used to predict the test set data, and the real value is compared with the predicted value as shown in figure 5.

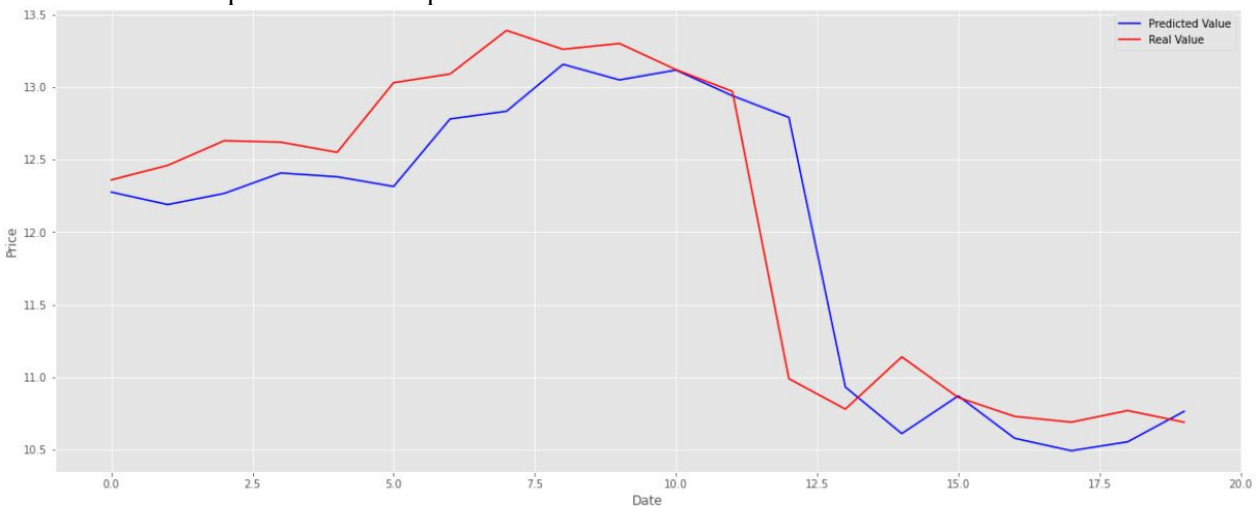
Comparison between predicted value and actual value for RNN COSCO



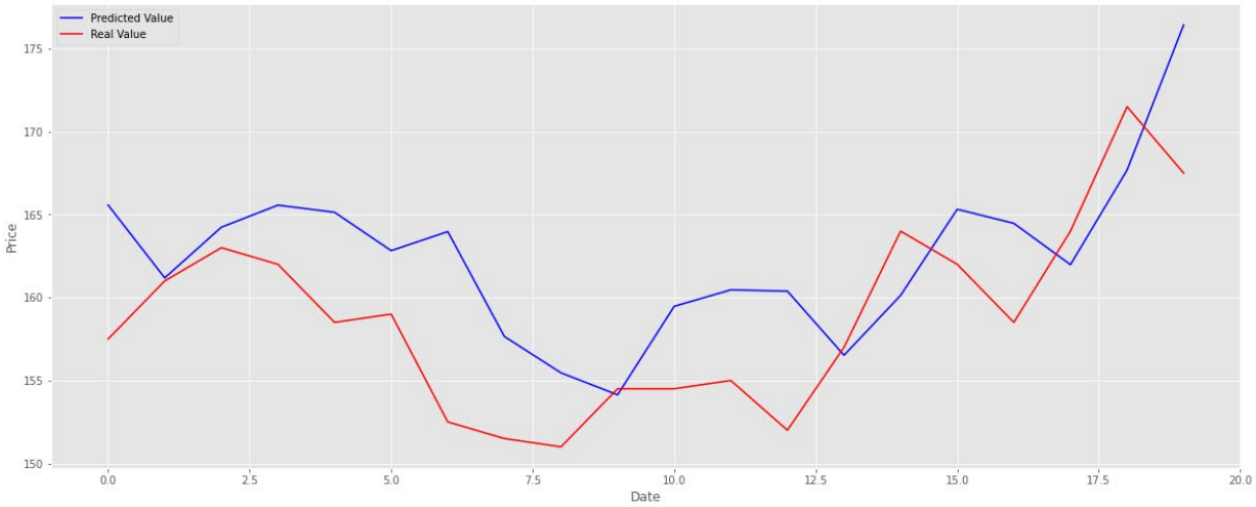
Comparison between predicted value and actual value for GRU COSCO



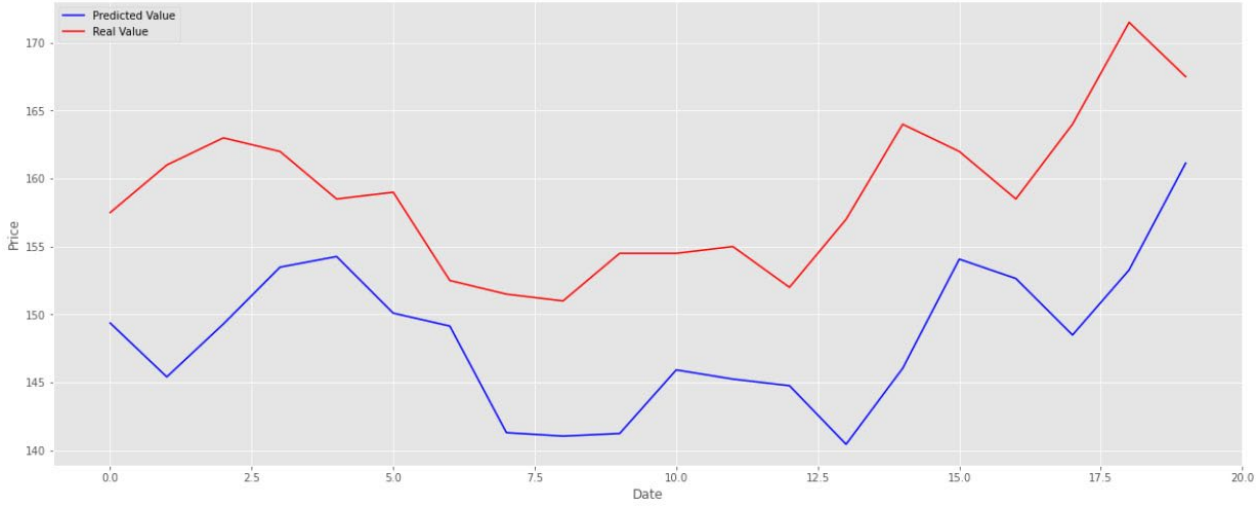
Comparison between predicted value and actual value for MS LSTM COSCO



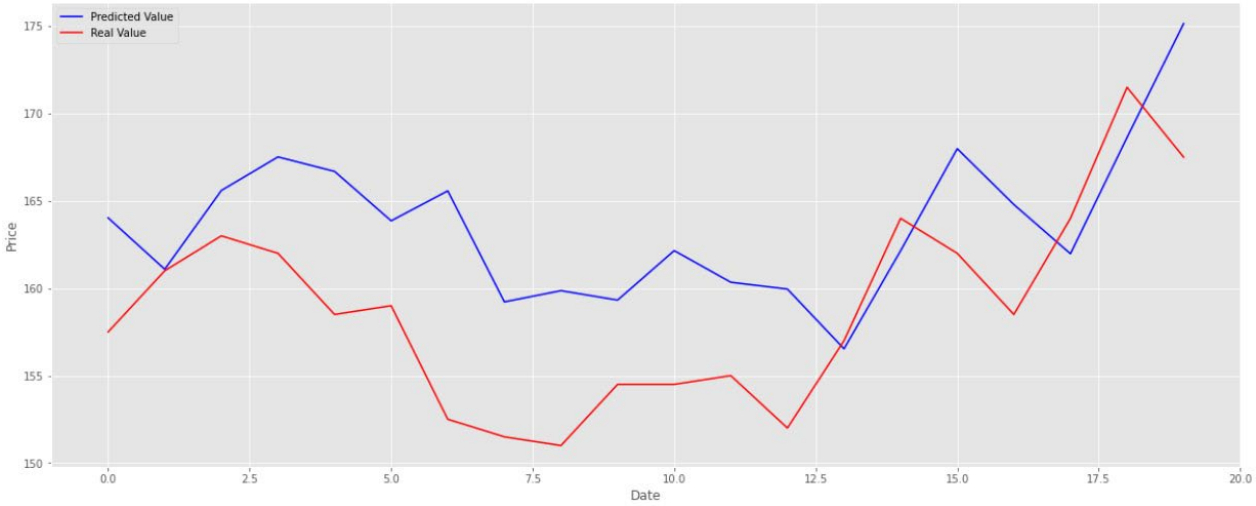
Comparison between predicted value and actual value for RNN Evergreen



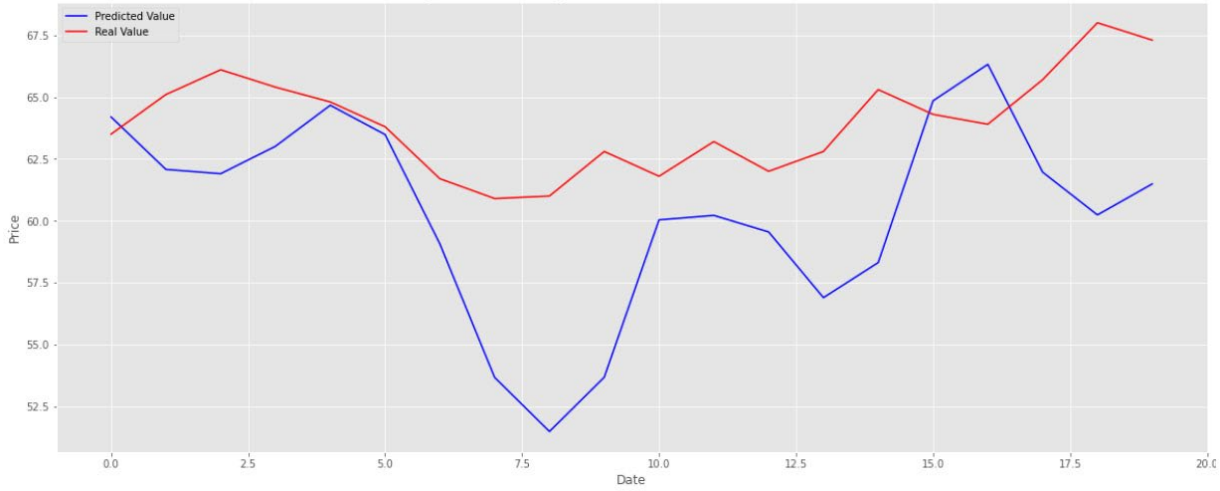
Comparison between predicted value and actual value for GRU Evergreen



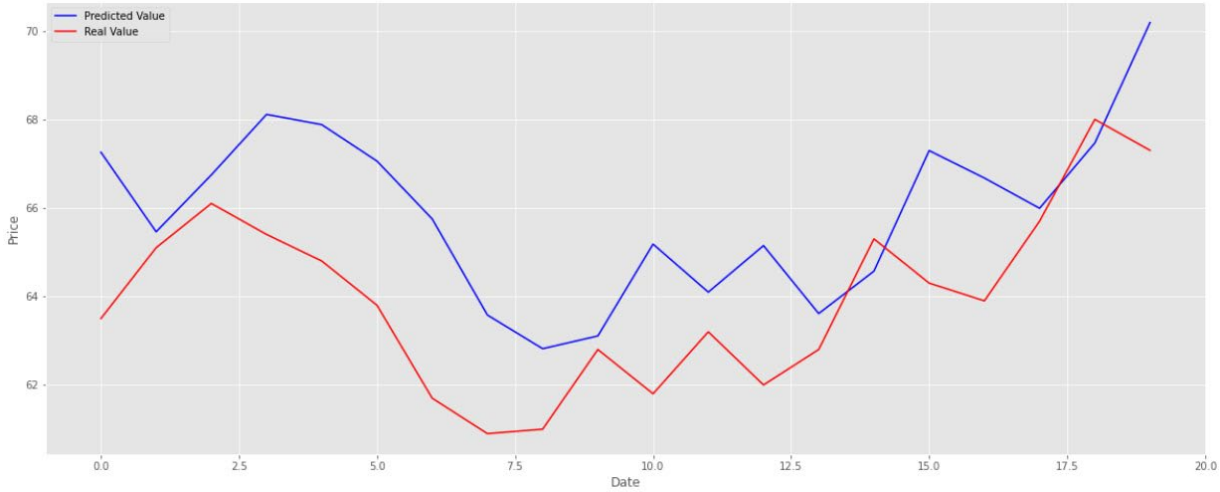
Comparison between predicted value and actual value for MS LSTM Evergreen



Comparison between predicted value and actual value for MS RNN Yang Ming



Comparison between predicted value and actual value for MS GRU Yang Ming



DOI: 10.2478/picbe-2023-0154, pp. 1728-1747, ISSN 2558-9652 |
Proceedings of the 17th International Conference on Business Excellence 2023

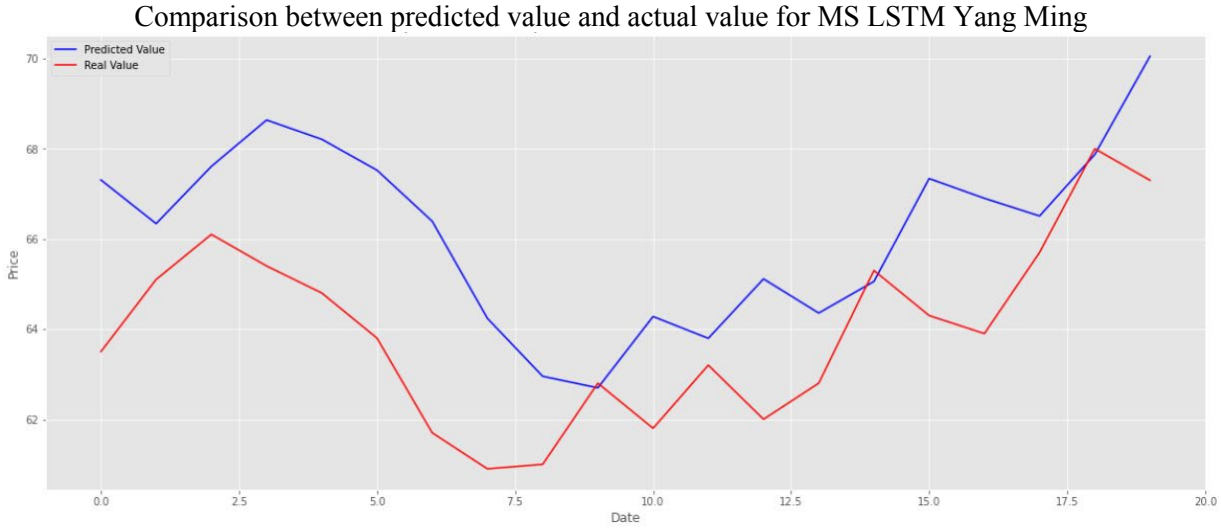


Figure 5. Comparison of the predicted value and the real value for RNN, GRU, and MS LSTM on COSCO, Evergreen, Yang Ming test data

Source: Authors' own research.

Figure 5 demonstrates that among the three predicting methods, the broken line fitting degree of real value and predicted value for RNN, GRU, MS LSTM. It can be observed that there is a very high overlap between the actual value and the predicted value by the MS LSTM model. According to the predicted value and real value of each method, the evaluation index of each method can be calculated, and the comparison results of the three methods are shown in Table 5.

Table 5. Forecast errors of different network models

Model	COSCO				Evergreen				Yang Ming			
	MAPE	MAE	MSE	RMSE	MAPE	MAE	MSE	RMSE	MAPE	MAE	MSE	RMSE
RNN	7.93	0.89	1.54	1.24	7.41	11.65	209.96	14.49	9.19	5.85	48.29	6.94
GRU	7.26	0.81	1.39	1.18	5.94	9.50	116.89	10.81	5.50	3.47	18.39	4.28
MS LSTM	5.43	0.63	0.83	0.91	4.01	6.33	56.84	7.53	5.23	3.30	17.50	4.18

Source: Authors' own research.

From Table 5, the MAPE, MAE, MSE, and RMSE of RNN are the largest, while the MAPE, MAE, MSE, and RMSE of MS LSTM are the smallest. The results show that the performance of MS LSTM is the best among the three methods. As for prediction accuracy, MAPE, MAE, MSE, and RMSE are the smallest among the three prediction models and have high prediction accuracy.

Conclusions

The researcher concludes that a new approach for stock price prediction can be applied based on a Multi-Step Long Short-Term Memory (MS LSTM) model that makes use of the Adam optimizer. As a result of the MS LSTM algorithm, normalized time series data are divided into time steps to determine the relationship between past and future values for accurate prediction. A prediction accuracy of 95.44% and 95.11% on the testing data set was attained using MS LSTM. The proposed model predicts the values with an average error percentage of 4.89% on the testing data and 4.56% on training data, which is considerably low, consequently yielding highly accurate predictions. By running the predictions for five days in the future, it is observed that the model has predicted that the prices would maintain their balance with little downward movement, leading to positive impact

on sea shipping with lower prices expected. By comparing the three methods (RNN, GRU, and MS LSTM), the results show that the performance of MS LSTM is the best among the three methods. In view of prediction accuracy, MAPE, MAE, MSE, and RMSE are the smallest among the three prediction models and have high prediction accuracy.

References

Abdul Quadir Md, Sanjit Kapoor, Chris Junni A.V., Arun Kumar Sivaraman, Kong Fah Tee, Sabireen H., Janakiraman N. (2022). Novel optimization approach for stock price forecasting using multi-layered sequential LSTM. *Applied Soft Computing, Volume 134,* https://doi.org/10.1016/j.asoc.2022.109830.

Bernard Njindan Iyke, M.M.J.D. Maheepala. (2022). Conventional monetary policy, COVID-19, and stock markets in emerging economies. *Pacific-Basin Finance Journal*, *Volume 76,* https://doi.org/10.1016/j.pacfin.2022.101883.

David Guerrero, Lucie Letrouit, Carlos Pais-Montes. (2022). The container transport system during Covid-19: An analysis through the prism of complex networks, *Transport Policy*, *Volume 115, pp. 113-125*, https://doi.org/10.1016/j.tranpol.2021.10.021.

Dimitris Gavalas, Theodoros Syriopoulos, Michael Tsatsaronis. (2022). COVID–19 impact on the shipping industry: An event study approach. *Transport Policy*, *Volume 116, 2022, pp. 157-164*, https://doi.org/10.1016/j.tranpol.2021.11.016.

Hakan Yilmazkuday. (2022). Coronavirus disease 2019 and the global economy. *Transport Policy*, *Volume 120, pp. 40-46,* https://doi.org/10.1016/j.tranpol.2022.03.003.

Hideaki Sakawa, Naoki Watanabel. (2023). The impact of the COVID-19 outbreak on Japanese shipping industry: An event study approach. *Transport Policy, Volume 130, pp. 130-140*, https://doi.org/10.1016/j.tranpol.2022.11.002.

Hong-Mei Zhao, Hong-Di He, Kai-Fa Lu, Xiao-Long Han, Yi Ding, Zhong-Ren Peng. (2022). Measuring the impact of an exogenous factor: An exponential smoothing model of the response of shipping to COVID-19. *Transport Policy*, *Volume 118, pp. 91-100,* https://doi.org/10.1016/j.tranpol.2022.01.015.

Hossein Abbasimehr, Mostafa Shabani, Mohsen Yousefi. (2020). An optimized model using LSTM network for demand forecasting. *Computers & Industrial Engineering, Volume 143*, https://doi.org/10.1016/j.cie.2020.106435.

Masha Menhat, Izyan Munirah Mohd Zaideen, Yahaya Yusuf, Nurul Haqimin Mohd Salleh, Meor Azry Zamri, Jagan Jeevan. (2021). The impact of Covid-19 pandemic: A review on maritime sectors in Malaysia. *Ocean & Coastal Management*, *Volume 209*, https://doi.org/10.1016/j.ocecoaman.2021.105638.

Meihui Zhong, Mingwei Lin. (2022). Bibliometric analysis for economy in COVID-19 pandemic. *Heliyon, Volume 8, Issue 9*, https://doi.org/10.1016/j.heliyon.2022.e10757.

Mei-Ping Chen, Chien-Chiang Lee, Yu-Hui Lin & Wen-Yi Chen. (2018). Did the S.A.R.S. epidemic weaken the integration of Asian stock markets? Evidence from smooth time-varying cointegration analysis. *Economic Research-Ekonomska Istraživanja*, *31:1, 908-926,* DOI: 10.1080/1331677X.2018.1456354.

Prathvi Thumbe Narasimha, Pradyot Ranjan Jena, Ritanjali Majhi. (2021). Impact of COVID-19 on the Indian seaport transportation and maritime supply chain. *Transport Policy*, *Volume 110, pp. 191-203*, https://doi.org/10.1016/j.tranpol.2021.05.011.

Yigit Kazancoglu, Esra Ekinci, Sachin Kumar Mangla, Muruvvet Deniz Sezer, Melisa Ozbiltekin-Pala. (2023). Impact of epidemic outbreaks (COVID-19) on global supply chains: A case of trade between Turkey and China. *Socio-Economic Planning Sciences, Volume 85*, https://doi.org/10.1016/j.seps.2022.101494.

*** UNCTAD. (April 2021). Shipping during COVID-19: Why container freight rates have surged. from https://unctad.org/news/shipping-during-covid-19-why-container-freight-rates-have-surged. Time access: November 2022.

*** UNCTAD. (March 2021). COVID-19 and Maritime Transport Impact and Responses. https://unctad.org/system/files/official-document/dtltlb2021d1_en.pdf. Time access: November 2022.

*** UNCTAD. (June 2022). COVID-19 and Maritime Transport: Navigating the Crisis and Lessons Learned. https://unctad.org/system/files/official-document/tcsdtlinf2022d1_en.pdf. Time access: December 2022.

*** UNCTAD. (November 2021). Freight rates, maritime transport costs and their impact on prices. https://unctad.org/system/files/official-document/rmt2021ch3_en.pdf. Time access: November 2022.