Journal of Marine Science and Technology

Volume 32 | Issue 1 Article 3

A new-type deep learning model based on Shapley regulation for containerized freight index prediction

Yen-Chang Shih
Department of Shipping and Transportation Management, National Taiwan Ocean University, Keelung 202301, Taiwan

Ming-Shue Lin
Department of Electrical Engineering, National Taiwan Ocean University, Keelung 202301, Taiwan

Taih-Cherng Lirn
Department of Shipping and Transportation Management, National Taiwan Ocean University, Keelung 202301, Taiwan

Jih-Gau Juang
Department of Communications, Navigation and Control Engineering, National Taiwan Ocean University, Keelung 202301, TAIWAN, jgjuang@mail.ntou.edu.tw

Recommended Citation

Shih, Yen-Chang; Lin, Ming-Shue; Lirn, Taih-Cherng; and Juang, Jih-Gau (2024) "A new-type deep learning model based on Shapley regulation for containerized freight index prediction," *Journal of Marine Science and Technology*: Vol. 32: Iss. 1, Article 3.
DOI: 10.51400/2709-6998.2729
Available at: https://jmstt.ntou.edu.tw/journal/vol32/iss1/3

RESEARCH ARTICLE

A New-type Deep Learning Model Based on Shapley Regulation for Containerized Freight Index Prediction

Yen-Chang Shih [a], Ming-Shue Lin [b], Taih-Cherng Lirn [a], Jih-Gau Juang [c,*]

[a] Department of Shipping and Transportation Management, National Taiwan Ocean University, Keelung 202301, Taiwan
[b] Department of Electrical Engineering, National Taiwan Ocean University, Keelung 202301, Taiwan
[c] Department of Communications, Navigation and Control Engineering, National Taiwan Ocean University, Keelung 202301, Taiwan

Abstract

In this study, we have crafted an innovative methodology that represents a groundbreaking synthesis of deep learning techniques with cooperative game theory. In this study, we use the accuracy of data prediction by different LSTM models as a measurement index and assign different LSTM models corresponding weights through the Shapley value calculation method to construct a more accurate predictive analysis model. We use this improved Shapley regulation model to calibrate a long short-term memory (LSTM) neural network by using historical freight data to predict the China Container Freight Index (CCFI), the leading export container freight index commonly used in China. Afterward, it is found that the neural networks calibrated in this way reduce their prediction bias in terms of mean absolute percentage error (MAPE), mean absolute error (MAE), root mean square error (RMSE), and mean square error (MSE) to improve prediction accuracy.

Keywords: Neural network, Deep learning, Long Short Term memory (LSTM), Shapley value, Containerized freight index, Cooperative game

1. Introduction

The China Containerized Freight Index (CCFI) was first compiled and released by the Shanghai Shipping Exchange on April 23, 1998. It employs 12 major international trading routes as samples. The index is derived from the freight rates provided by 22 renowned domestic and international ocean container carriers and is updated weekly. As China is the world's leading exporting nation, the CCFI serves as an objective and timely barometer of the shipping freight rates from Chinese ports. It is evident that Chinese export volumes are influenced by global economic activity and domestic production capacity. During periods of robust global trade demand, Chinese exports tend to increase, leading to a tighter container supply and subsequently higher freight rates. This results in elevated indices such as the CCFI and the Shanghai Containerized Freight Index (SCFI), and the converse is true when global demand wanes.

The COVID-19 pandemic has severely impacted the global economy over the past three years. With the progressive rollout of vaccines and advances in medical technology, pandemic-related lockdowns around the world have been steadily lifted. China's expeditious post-lockdown recovery has markedly enhanced its production efficiency, attracting a resurgence of international orders to its factories. By analyzing the CCFI, a leading indicator of ocean freight rates, we can gain insights into China's economic dynamism and the revival of global trade. In this study, we have utilized historical CCFI data prior to the COVID-19 outbreak to predict future indices using a newly proposed Shapley-Like Neural Correction Model, an innovative approach in time series analysis.

Received 2 August 2023; revised 14 November 2023; accepted 6 December 2023.
Available online 11 March 2024

* Corresponding author.
E-mail addresses: dadashih671111@gmail.com (Y.-C. Shih), mingshitlin@gmail.com (M.-S. Lin), tedlirn@mail.ntou.edu.tw (T.-C. Lirn), jgjuang@mail.ntou.edu.tw (J.-G. Juang).

https://doi.org/10.51400/2709-6998.2729

The contributions of this work are as follows:

1. **Methodology Development:** Hirata and Matsuda [1] utilized both LSTM and SARIMA methods to forecast the comprehensive and route-specific SCFI, demonstrating that LSTM deep learning models typically outperformed SARIMA models across the majority of datasets. These results highlight the potential advantages of employing advanced LSTM models. To train an effective deep learning model, it is essential to have sufficient data for training, construction, and substantial computation with a computer. However, this conventional process may not be directly applicable in analyzing the CCFI and other freight indices due to limitations in data size and resources. Building on these insights and inspired by the principles of the Shapley value, we have crafted an innovative methodology that represents a groundbreaking synthesis of deep learning techniques with cooperative game theory. In this study, we use the accuracy of data prediction by different LSTM models as a measurement index and assign different LSTM models corresponding weights through the Shapley value calculation method to construct a more accurate predictive analysis model. This endeavor marks a successful first in the field, bridging gaps not previously addressed in the existing literature.
2. **Impact and Implications:** Our research focuses on improving the predictive accuracy of the China Container Freight Index (CCFI), a key barometer of the volatility in the container shipping market. While our study primarily addresses the application of sophisticated deep learning techniques for the predictive analysis of freight rates, the methodologies and insights derived could indirectly support the enhancement of shipping route optimization by contributing to a more detailed understanding of market dynamics.

The structure of this paper is organized as follows: Section II reviews the existing forecasting methods, providing background and a survey of the related literature. Section III introduces our modified Long Short-Term Memory (LSTM) model and highlights the innovation presented in this study. Our proposed deep learning model, the Shapley Regulation-LSTM (SR-LSTM), represents a significant innovation and contribution of this paper. We compare the predictive accuracy of the SR-LSTM against traditional LSTM models with benchmarks such as Mean Absolute Percentage Error (MAPE), Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Mean Square Error (MSE) detailed in Section IV. Finally, Section V discusses the conclusions and future research directions derived from this study.

Notation

N	Cooperative game (TU-game) represents all participants, and the number is n
S	All sub-sets of N, named and defined as unions, have a total of $\{2^N \setminus \phi\}$, and a single S element is denoted by $\|S\|$
$i(j)$	Each independent member of N in the action decides whether to participate or not $\{0,1\}$
$v(u)$	Each alliance in this game (S) is mapped to a real number
$\phi_i(N:v)$	Each independent member i (or j) in this cooperative game can also be simplified as $\{\phi_i\}_{i\in N}$, the Shapley value of each member

2. Advances in forecasting methods: A literature review

Currently, three primary methodologies dominate the field of data forecasting: (1) econometric models, (2) artificial neural networks, and (3) deep learning models. This section provides an overview of these techniques, accompanied by a discussion of the relevant literature.

2.1. Econometric models

Econometric models employ statistical tools for predicting desired variables and assessing the impact of explanatory economic factors. Among these, the regression model stands out as particularly prevalent, with the ARIMA (Autoregressive Integrated Moving Average) model being chiefly utilized for univariate time series forecasting [2]. Prior research has demonstrated that the LSTM (Long Short-Term Memory) technique within neural network modeling often surpasses the performance of naive forecasts and traditional time series methodologies, which include the Moving Average (MA), Simple Exponential Smoothing (SES), Holt−Winters (HW), and Seasonal Autoregressive Integrated Moving Average (SARIMA) models [3].

2.2. Artificial intelligence (artificial neural network)

Box and Jenkins pioneered the development of time series forecasting methodologies,

incorporating autoregressive (AR) and moving average (MA) processes tailored for linear time series analysis. However, subsequent research has identified that most time series data exhibit nonlinear characteristics. Consequently, the linear framework established by Box and Jenkins often fails to effectively model the complex dynamics inherent in these nonlinear time series [4]. The recent advancements in artificial intelligence, coupled with substantial improvements in computational power, have propelled the practical application of AI in forecasting beyond mere theoretical constructs. Researchers have increasingly turned to AI techniques to address the challenges associated with nonlinear time series [5]. Consequently, various configurations of Artificial Neural Networks (ANNs) have been developed, leveraging their prowess in capturing nonlinearity and learning sequence patterns. Significant advances have been reported; for instance, ANNs with exogenous inputs [6] and quasi-periodic ANN [7] have been utilized to forecast industrial average returns. In addition, adaptive artificial neural networks [8] and hybrid adaptive neuro-fuzzy inference systems have demonstrated superior performance [5]. Moreover, empirical studies on container throughput forecasting have established ARIMA combined with ANNs (ARIMA + ANN) as a robust and widely recognized approach [9−12].

2.3. Deep learning

The recent advancements in deep learning have been marked by the introduction of several effective methodologies, with the Long Short-Term Memory (LSTM) network being one of the most noteworthy. As a specialized form of Recurrent Neural Networks (RNNs), LSTMs offer a solution to the limitations of traditional RNNs, particularly their difficulty in preserving long-term dependencies due to the vanishing gradient problem, a phenomenon extensively explored by Bengio et al. [13]. LSTMs, designed with ingenuity to counteract this problem, are further elucidated in seminal works [14,15]. Initially proposed by Sepp Hochreiter and Jürgen Schmidhuber between 1995 and 1997, the LSTM architecture introduced the Constant Error Carousel (CEC) unit, effectively addressing the vanishing gradient issue [15,16]. This foundational architecture, comprising cell states and input and output gates [17], was subsequently refined in 1999 with the introduction of the forget gate, or "keep gate," by Felix Gers, Schmidhuber, and Fred Cummins, enabling the network to selectively reset its internal state [18].

By 2000, Gers and colleagues had integrated peephole connections into the design, which allowed for more nuanced state management, and even opted to remove the output activation function for greater efficiency [17,19]. This evolving architecture propelled an LSTM-based model to victory in the ICDAR Connected Handwriting Recognition Competition in 2009, thanks to the precision and speed of the models developed by a team whose efforts were recognized in [20,21]. The prowess of LSTM was further cemented in 2013 when it achieved a groundbreaking 17.7 % phoneme error rate on the TIMIT natural speech dataset [22], and by 2014, the LSTM landscape was diversified with the introduction of the Gated Recurrent Unit (GRU) by Kyunghyun Cho [23]. LSTMs continued to demonstrate commercial viability, with Google implementing them in 2015 for speech recognition in Google Voice, slashing transcription errors by nearly half [23,24]. The technology's application expanded in 2016 across various Google services, including the Allo app and the Google Translate Neural Machine Translation system, which saw a significant reduction in translation errors [25−27]. The same year, Apple and Amazon also integrated LSTMs for iPhone QuickType and Alexa's speech generation, respectively [28,29], signaling a robust and precise commercial utilization of LSTM.

The LSTM narrative took another turn in 2017 when a collaborative study from Michigan State University, IBM Research, and Cornell University introduced a neural network at the KDD conference that outperformed established LSTM networks on select datasets [30]. Microsoft, too, leveraged "dialog session-based LSTM" to achieve an impressive 94.9 % accuracy on a large vocabulary speech recognition task. The advancements did not stall, as evidenced in 2019 when the University of Waterloo researchers proposed an RNN architecture, underpinned by Legendre polynomials that surpassed traditional LSTMs in some tests [31]. That year also saw an LSTM-based model rank third in a large-scale text compression challenge, further underscoring the model's accuracy and relevance [32].

This comprehensive overview of LSTM's development, from its inception to its significant strides in both academic and commercial realms, showcases its revolutionary impact on deep learning and its vital role in advancing the field of artificial intelligence.

2.4. Shapley value

Shapley values are often used to solve problems of cost allocation, benefit sharing, distribution, and

partnership dissolution based on participants' contributions to the coalition (Moulin, 1992; Pérez-Castrillo and Wettstein, 2001; Petrosjan and Zaccour, 2003; Macho- Stadler, 2007) [30−33]. It has been widely used in various industries. Dubey (1982) saw an example that airport runways are suitable for aircraft of different sizes, and the Shapley value can be used to plan the allocation of the number of various types of airplanes to use various runways [34], Tan and Lie (2002) used the Shapley value for the cost allocation to various power system users [35], Bartholdi and Ziya (2005) use the Shapley value to distribute profits among all coalition participants [36], Narayanam and Narahari (2011) used this Shapley value technique for efficient communication in social networks [37], and Yu et al. (2014) used this method to discuss carbon emission reduction allowances in industrial areas [38].

In the field of supply chain, Kemahlõoğlu-Ziya and Bartholdi [39] and Zhang and Liu [40] demonstrated that the allocation of Shapley values can guarantee that participants are incentivized to respond positively to coordinate supply, and Raghunathan [41] used the concept of Shapley values to analyze the expected shares of manufacturers and retailers in the surplus generated by data sharing. Rosenthal developed a model to fairly quantify transaction prices in vertically integrated organizations, and the results show that Shapley value can help allocate entire supply chain benefits in the upstream and the downstream [42]. Leng and Parlar analyzed the allocation and cost savings from sharing demand data among supply chain participants [43,44]. Gordan (2017) proposed the change of Sharpley value as a method to solve the uncertain alliance game. In this game, the player's profit is derived from the extension of the prerequisite variables. In addition, Wang (2018) used Shapley value to explore China's cruise supply chain (cruise port-travel agency-shore service), and proposed incentive policies for fare subsidies distribution [45], Li (2019) included the influential factor of online shopping in southwest China in terms of weight, which contributes to the practical use of the Shapley value [46], Liang (2019) used an uncertain cooperative game to allocate public resources among vulnerable groups, and proposed a linearly added superimposed Shapley value, and named it α-value best solution [47]. Li and Wang (2019) used the Shapley value to consider the asymmetric contribution of partners in the proposed profit distribution scheme of express companies in the logistics service market, and then showed more rationality for profit distribution [48]. Otero and Amaya (2020) employed a simulation technique to allocate the profits according to the investments by using the Shapley value function and demonstrated expected cost reduction in inventory management [49]. Jang and Jeong (2021) [50] developed three LSTM-RNN-based models to calculate their Shapley values to explain the accuracy of each model and the effectiveness of the model, and applied these models to bankruptcy prediction. Shalit (2020) attempted to quantify the relative risk of securities in an optimal portfolio [51].

2.5. CCFI's related forecast research

Despite the growing aptitude of neural networks for many time series predictions, their application in forecasting container freight rates, particularly the China Container Freight Index (CCFI), remains comparatively unexplored. The field recognizes two principal container freight indices: the CCFI, first published in 1998, and the Shanghai Containerized Freight Index (SCFI), introduced in 2005. A comprehensive review of the Web of Science database yielded six pertinent studies on container freight index forecasting. Previous studies on the CCFI have utilized SARIMA, Empirical Mode Decomposition (EMD), and Grey Wave Forecasting models for predictions. Chen et al. [52] enhanced the scope of grey system theory and the graphic forecasting method by applying a decomposition−ensemble method based on EMD to forecast the CCFI, finding superior performance over the ARMA model and random walk for multi-step-ahead predictions. Munim [53] demonstrated that either the TBATS model alone or combined with SARIMA outperforms SARIMA and SNNAR, as well as their combinations, in both training and test samples for CCFI forecasting. Turning to the SCFI, Munim and Schramm [54] utilized an autoregressive conditional heteroscedasticity (ARIMARCH) model to predict container freight rates along Asia−North Europe routes, showing that the ARIMARCH model yielded better short-term forecasts compared to existing models on weekly and monthly data. In a similar vein, Munim and Schramm [55] assessed the efficacy of ARIMA, VAR/VEC, and ANN models using SCFI data for four major trade routes and concluded that VAR/VEC models excel in training-sample forecasts over ARIMA and ANN. Lastly, Koyuncu and Tavacõoğlu [56] found that the SARIMA model outperformed other freight-rate forecasting models, including Holt−Winters Methods, when applied to short-term monthly forecasts of the SCFI.

Hirata and Matsuda [1] utilized LSTM and SARIMA methodologies to forecast the comprehensive and route-specific SCFI, finding that LSTM deep

learning models typically outperform SARIMA models in the majority of datasets. These findings underscore the potential benefits of adopting advanced LSTM models for enhanced neural forecasting. The recent years have witnessed a surge in the application of deep learning techniques across various domains, as evidenced by emerging studies such as those referenced in [57]. This growing body of work contributes to the expanding utilization of sophisticated neural models like the SR-LSTM, which are proving to be vital in the progression of predictive analytics.

To summarize, there are primarily three techniques utilized in data prediction. Econometric models are adept at forecasting and estimating quantities to gauge the impact of economic changes. Neural networks harness their unique ability to capture non-linearities and learn sequential behavior for prediction purposes. Deep learning builds upon the evolution of mathematical functions, demonstrates the enhanced performance of composite neural networks, and is adept at handling long-term data dependencies. Compared to other methods, LSTM exhibits clear superiority in sequential forecasting and commercial applications. By integrating a Shapley value-based weight selection technique into LSTM's time series forecasting, we expect to further enhance training outcomes. These advancements will be discussed in subsequent sections of this paper.

3. Methodology

3.1. LSTM network

In recent years, applications related to machine learning have garnered significant attention, with some becoming particularly popular. Firstly, there is image recognition, which translates any problem into a comparable form; secondly, there is a sequence to sequence translation, encompassing speech-to-text, translation between languages, or prediction of time series data. Most of the former applications are executed using Convolutional Neural Networks (CNN), while the latter are primarily handled using Recurrent Neural Networks (RNN), especially Long Short-Term Memory (LSTM) networks. The original LSTM model possesses two notable capabilities: one is to discern "what to remember from past data" and "what requires additional learning from new information"; the second is to utilize the aforementioned understanding to make predictions with interpretations (see [14,15,58,59]). For ease of explanation, Fig. 1 illustrates the basic architecture of the RNN model, and Fig. 2 depicts the architecture of the LSTM model.

The following statements briefly introduce the process and mechanisms of LSTM.

Step 1: The operation mode of the forget gate vector is as formula (1):

$$f_t = \sigma\left(W_f \cdot [X_t, Y_{t-1}] + b_f\right). \tag{1}$$

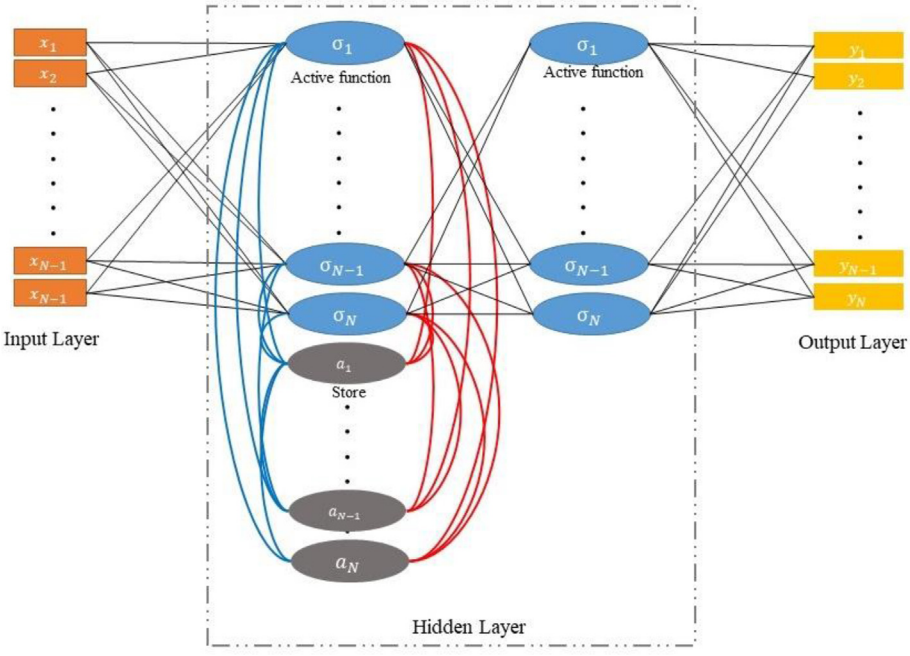

Fig. 1. RNN model architecture diagram (modified from [59]).

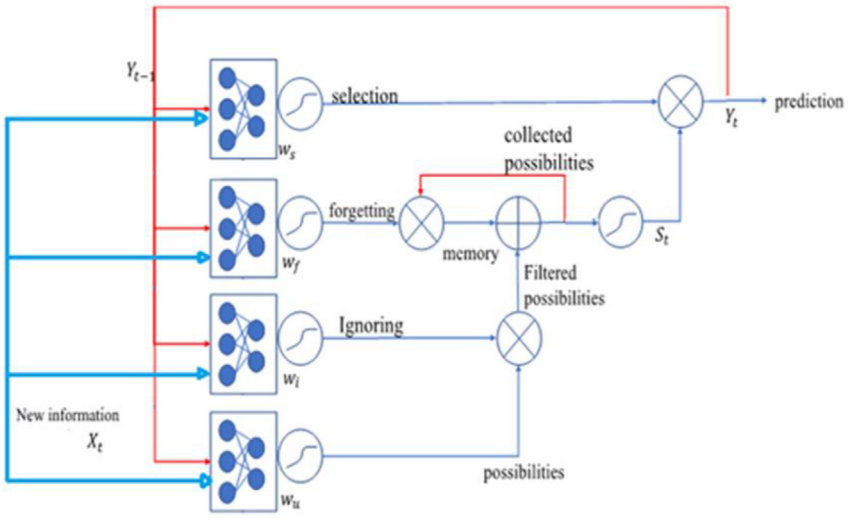

Fig. 2. LSTM model architecture diagram (modified from [15]).

Where W_f is the vector of the connection weights, b_f is the bias, X_t are the current inputs, and Y_{t-1} are the previously predicted outputs. The activation function $\sigma(x)$ is the famous sigmoid function as follows (2). This layer is mainly used to deal with whether the information needs to be remembered or forgotten. If it is to be forgotten, the output value is close to zero; otherwise, it is close to one.

$$\sigma(x)=\frac{1}{1+e^{-x}}. \tag{2}$$

Step 2: The selection information can be divided into two parts: (3) and (4). The sigmoid function helps to select the information that must be preserved from the time step.

$$I_t=\sigma(W_i\cdot[X_t,Y_{t-1}]+b_i) \tag{3}$$

$$\widehat{S}_t=\tanh(W_s\cdot[X_t,Y_{t-1}]+b_s). \tag{4}$$

The next step is to process the old cell state S_{t-1} to the new unit state S_t. The product (dot product [*]) is used to accomplish this task. Here, the forgetting layer output is combined with the selection layer output to update the cell state, as shown in the following equation:

$$S_t=f_t*S_{t-1}+I_t*\widehat{S}_t \tag{5}$$

Step 3: The final output is its relational formula, such as Y_t in (6). W_u is the weight at this time.

$$U_t=\sigma(W_u\cdot[X_t,Y_{t-1}]+b_o). \tag{6}$$

$$Y_t=U_t\cdot\tanh(S_t).$$

3.2. Shapley regulation model

1. Shapley Regulation method: First, the definition of cooperative game with transferable utility (referred to as *TU* game) is expressed as $(N:\upsilon)$. For the set of participants, the name N of each subset in this game is coalition S (coalition). The feature of the *TU* game is to give each coalition S (coalition) a reward feature function ν as long as the coalition contains all members i. By unanimously agreeing (to form this alliance instead of joining other alliances, as well as the distribution plan of this reward), members can get the pre-agreed reward.

The eigenvalue function is expressed in the following (7):

$$\nu:(0,\infty)\to R \tag{7}$$

2. Shapley value of traditional single-choice cooperative game: the Shapley value refers to the N cooperative game against one person $(N:\upsilon)$. Each member, $i\in N$, should expect to be paid $\phi_i(N:\upsilon)$ and must possess the following axioms:
   (1) Efficiency: The rewards after cooperation need to be fully distributed to the contestants of the game, that is:

$$N\sum\nolimits_{i\in N}\phi_I(N:v)=v() \tag{8}$$

   (2) Symmetry and Dummy player property:
   i. Each member (player) $\phi_i(N:\upsilon)$ is completely determined by their marginal contribution to the coalition S (coalition). So, if there are two stakeholders, *i and j, their marginal* contribution to the coalition is

always equal, then we get $\phi_i(N:v) = \phi_j(N:v)$.

ii. On the other hand, if the marginal contribution of $\phi_i(N:v) = 0$ from a person i to the alliance S (coalition) is equal to zero, we define such a member as a dummy player; otherwise, they are called a carrier.

iii. Linear function on the vector space formed by all N-player cooperative games in the real number domain $\{\phi_i\}_{i\in N}$. Take any two cooperative games, the $(N:v)$ sum $(N:u)$, and the real numbers *a and b* for any $i\in N$:

$$\phi_i(N:av+bu) = a\phi_i(N:v)+b\phi_i(N:u) \tag{9}$$

3. Let the $\phi_i(N:v)$ be defined as follows:

$$\phi_i(N:v) = v(N\cap[1,i] - v(N\cap[1,i-1])) \tag{10}$$

Taking the marginal contribution as a solution satisfies all conditions except symmetry; that is, giving the number of all participants an operation (permutation), we may get another set of solutions. If all operations are regarded as equivalent, redefining the expected value can have symmetry on the basis of ensuring that the original conditions are satisfied.

4. Shapley not only proved the uniqueness of the existence of the Shapley value but also gave the calculation formula for calculating the Shapley value. $C(S)$ is defined as the set value of the reward, the reward of the $W(|S|)$ alliance, and the weight factor formed by all the subset alliances, as follows:

$$\phi_i = \sum\nolimits_{|s|i\in N} W(|S|)\times[C(S\setminus i)], i = 1,2,\ldots..,n \tag{11}$$

$$W(|S|) = \frac{(n-|S|)!\times(|S|-1)!}{n!}, in\ \Phi(v) = \{\phi_1,\phi_2,\phi_3,\ldots..\phi_n\}\in R^n \tag{12}$$

3.3. Shapley regulation neural correction mode flow chart

In our study, Fig. 3 illustrates the workflow of the Shapley Regulation model, which is primarily divided into two main components: the construction of the Shapley value-weighted network is one of these. Initially, data is inputted into several pre-trained LSTM neural networks (three were selected for this study) to generate forecast data. Subsequently, the Shapley Region algorithm is employed to calculate the Shapley weights among the different

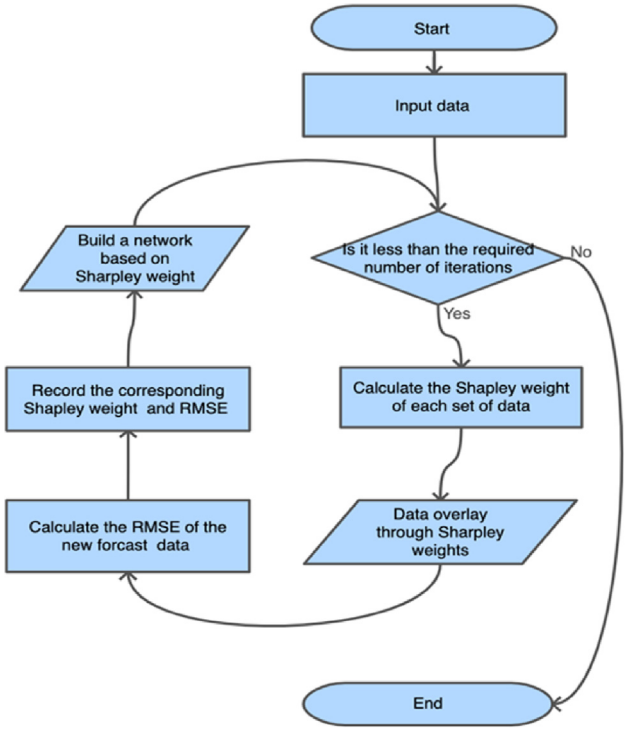

Fig. 3. Shapley regulation neural correction mode flow chart.

LSTM models. We take the forecast data generated by each LSTM model and perform an arithmetic mean of the predictions from different models. If the average forecast data exhibits a lower RMSE value than the original data, it is considered that the group of LSTM models forms an effective coalition. Based on this, we iterate and update the forecast data, adjusting and consolidating iteratively. After several iterations, we select the forecast with the lowest RMSE value as the optimal result and record the Shapley weights for each data set in every iteration. This process ultimately constructs the required network architecture for Shapley calibration. In the end, we use this network structure and Shapley weights to predict time-series data, serving not only as a mechanism to adjust the predictive trends generated by multiple LSTM networks but also as a validation of its effectiveness in practical applications.

3.4. Data collection

When discussing the CCFI freight index, the following three steps are used to calculate the index.

3.4.1. Selection of CCFI sample routes

According to the three basic principles of typicality, regional distribution and correlation, 14 routes were screened out, and the freight index was calculated based on the freight rates and container volumes of these routes, namely Hong Kong, South

Korea, Japan, Southeast Asia, Australia and New Zealand, the Mediterranean, Europe, East and West Africa, West America, East America, South Africa, South America, and Bohong and Taiwan. Its domestic departure ports include Dalian, Tianjin, Qingdao, Shanghai, Nanjing, Ningbo, Xiamen, Fuzhou, Shenzhen, Guangzhou and other ten ports. Because the service relationship on the container liner route is relatively small, and the prices of each route are relatively independent, the result of the calculation can more representatively reflect the price fluctuations of the sub-routes.

3.4.2. Selection of CCFI sample shipping companies

The basic principles adopted are as follows.

(1). Select 3 to 5 Chinese and foreign shipping companies for each route to improve the representativeness.
(2). Characteristics of the selected shipping companies: these shipping companies have independent legal personalities in China, good business reputation, wide route distribution and large market shares.
(3). The selected shipping companies all voluntarily participate in the freight index compilation committee and can immediately disclose accurate information about freight rates in accordance with the requirements of the Shanghai Shipping Exchange.
(4). Currently selected companies: there are 22 Chinese and foreign shipping companies with outstanding reputations and large market share in the route, which provide the freight rate information required for CCFI compilation according to the voluntary principle. They are CMA CGM (China) Co., Ltd. (CMA-CGM), COSCO Container Lines Co., Ltd. (COSCO), China Shipping Container Lines Co., Ltd. (CSCL), Evergreen Shipping Co., Ltd. (EMC), Hanjin Shipping (China) Co., Ltd. (HANJIN), Shanghai Haihua Shipping Co., Ltd. (HASCO), Hapag-Lloyd Shipping (China) Co., Ltd. (HLAG), Hamburg Süd (China) Shipping Co., Ltd. (HSDG), Shanghai Jinjiang Shipping Co., Ltd. (JINJIANG), Korea Shipping (Shanghai) Co., Ltd. (KMTC), Kawasaki Steamship (China) Co., Ltd. (K-LINE), Maersk (China) Shipping Co., Ltd. (MAERSK), Merchant Shipping Mitsui (China) Co., Ltd. (MOL), Lisheng Mediterranean Shipping (Shanghai) Co., Ltd. (MSC), Nippon Yusen (China) Co., Ltd. (NYK), Orient Overseas Container Lines (China) Co., Ltd. (OOCL), Pacific Shipping Lines (China) Co., Ltd. (PIL), Honghai Container Lines Co., Ltd. (RCL), Sinotrans Container Lines Co., Ltd. (SINOTRANS), Xinhaifeng Container Lines Co., Ltd. (SITC), Wanhai Shipping Co., Ltd. (WANHAI), and Yang Ming Shipping Co., Ltd. Company (Yangming).

3.4.3. Summation of CCFI calculation formula (Shanghai Shipping Exchange):

I. The route calculation formula adopts the Laplace index calculation method, and the calculation formula is as follows:

$$L_i = \frac{P_i}{P_{io}} \times L_{io} \tag{13}$$

(1) L_i is the ith route index for the reporting period of the route.
(2) L_{iO} is the ith base period route index of the route.
(3) P_{iO} is the ith base period average freight rate of the route.
(4) P_i is the ith average freight rate of the route in the reporting period, which is calculated by a fixed weighted average. The calculation formula is as follows:

$$P_i = \sum_{j=1}^{n} \left(P_{ij} \times W_{ij}\right) \tag{14}$$

where n is the sample number of the *ith* route, W_{ij} is the j th fixed weight of the sample company on the$j$-th route, and P_{ij} is the j th freight income per container of the sample company on the i-th route. Further, freight composition includes ocean freight (O/F), emergency fuel surcharge (EBS), container shipping surcharges such as Imbalance Charge (CIC), as well as the Terminal Handling Charge (OTHC) at the Loading Port and the Terminal Handling Charge (DTHC) at the Discharging Port.

II. The component index adopts the calculation method of weighted average, and the calculation formula is as follows:

$$L_{cf} = \sum_{i=1}^{m} \left(\frac{P_i}{P_{iO}} \times W_i \times L_{cfO}\right) \tag{15}$$

where m is the number of routes for the constituent index.

(1) W_i is the weight of the i-th route.

(2) L_{cfO} is the base period component index.

III. The composite index adopts the calculation method of the weighted average of the component indices, and the calculation formula is as follows:

$$L_{zh} = L_{jkcf} \times W_{jkcf} + L_{ckcf} \times W_{ckcf} \quad (16)$$

The variables are defined as follows:

(1) L_{jkcf} is the import component index.
(2) W_{jkcf} is the import component index weight.
(3) L_{ckcf} is the export component index.
(4) W_{ckcf} is the weight of the export component index.

As the base period of the price index, the choice must not only reflect the normal distribution of freight rates, but also consider the stability, availability, and comparability of price information. The comparability includes the changes of the sample routes, such as the increase or decrease of the number of routes, the change of the weight of the routes, etc., and the rules about the replacement of routes and weight changes must be standardized. The composition usually fluctuates greatly in the long term, so the selection of the base period should not be too far from the reporting period. The base period of CCFI is set as January 1, 1998, and the base period index is 1000. The CCFI is published weekly and compiled and published every Friday.

IV. Determination of CCFI tariff type:

The formulation of container freight rates is not only affected by the value of transported products, the supply and demand relationship in the transport market and the behavior of market players. Thus, the freight rates on the same route for different cargoes may vary due to the difference of delivery methods, trade terms, and services provided by shipping companies. Theoretically, the freight rate index reflects the freight rate under the action of internal factors such as the value of transportation products, market supply and demand, and the behavior of market players, excluding the influence of external factors such as handover methods and trade terms. However, it is difficult to calculate the freight index accordingly during operation. Based on theoretical induction and proof, it is feasible to use the comprehensive freight rate affected by various factors to compile the freight index, which can reflect the fluctuation of the container freight rate under the influence of intrinsic factors. The freight rate adopted by CCFI is the comprehensive freight rate under the influence of various factors, that is, the overall weighted average of the freight rates of major shipping companies in all ports. The freight rate of the shipping company refers to the freight rate using CY-CY terms.

V. Selection of CCFI Base Period and Frequency of Release:

Comparability of price information: in the long run, the distribution of liner routes and the composition of freight rates will undergo major changes. Therefore, the selection of the base period should not be too far from the reporting period. The base period of CCFI is set as January 1, 1998, and the base period index is 1000. The CCFI is published weekly and compiled and published every Friday.

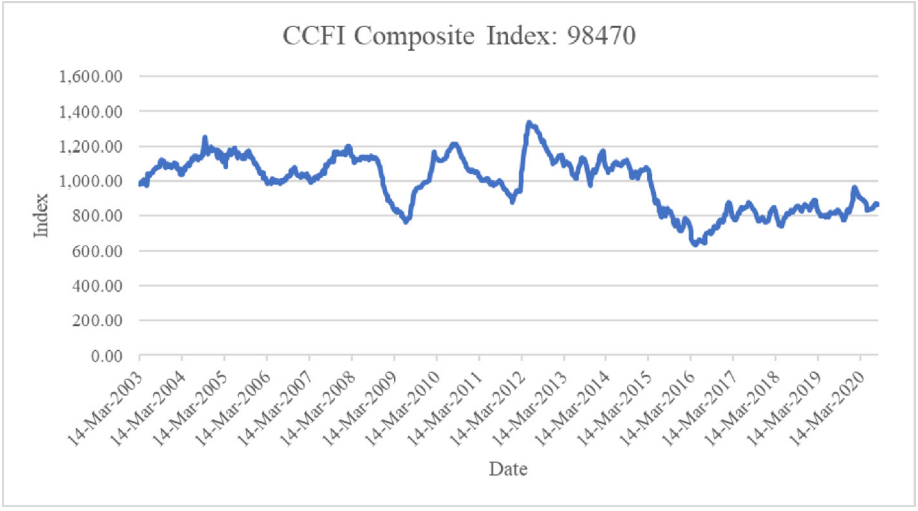

Fig. 4. CCFI composite index: 98470.

Table 1. Forecasting error matrix for traditional LSTM and Shapley Regulation modified LSTM (SR-LSTM).

	Bias	MAPE	MAE	MAE%	RMSE	RMSE%	MSE
LSTM_100	−7.21271	12.66451	0.015367	0.015945	16.03446	0.019455	257.1039
LSTM_300	−9.66284	13.88462	0.016847	0.017491	17.36751	0.021073	301.6304
LSTM_300_2	−5.76557	12.37378	0.015014	0.015691	16.11064	0.019548	259.5528
LSTM450	−8.83531	13.67791	0.016596	0.017384	17.52009	0.021258	306.9534
SR-LSTM 1	−6.81121	12.66528	0.015367	0.016036	16.28316	0.019757	265.1414
SR-LSTM 2	−6.89094	12.70034	0.01541	0.016082	16.32285	0.019805	266.4354
SR-LSTM 3	−5.94742	12.34343	0.014977	0.015621	15.96206	0.019368	254.7873
SR-LSTM 4	−6.35974	12.47882	0.015141	0.015791	16.07951	0.01951	258.5506

VI. Modification of CCFI Compilation Method and Maintenance of Base Period:

Over time, the method of compiling the freight index may change accordingly due to changes in the actual situation. Therefore, it is important to ensure the comparability of the index before and after adjustment. The main factors affecting the comparability of the index are the changes of the sample routes, such as the increase or decrease of the number of routes and the change of the weight of the routes. CCFI has compiled rules on route change, weight change, etc. The time-series data for CCFI Composite Index: 98470 are shown in Fig. 4.

3.5. Commonly used forecast indicators the forecast KPI

This paper utilizes Key Performance Indicators (KPIs) for numerical prediction, referencing the algorithm presented in [60]. Formulas (17) to (24) are employed as indicators for model predictions. The calculations are as follows:

1. Error is defined as the difference between the real data (f_t) and forecast data (d_t)

$$e_t = f_t - d_t \tag{17}$$

where f_t represents real data and d_t denotes forecast data.

2. Bias is calculated as the average error across all observations:

$$bias = \frac{1}{n}\sum_{1}^{n} e_t \tag{18}$$

where n is the number of historical periods where you have both a forecast and a demand.

3. Mean Absolute Percentage Error (MAPE) is the mean of the absolute percentage errors:

$$MAPE = \frac{1}{n}\sum_{1}^{n}\frac{|e_t|}{d_t} \tag{19}$$

4. Mean Absolute Error (MAE) is the average of the absolute errors:

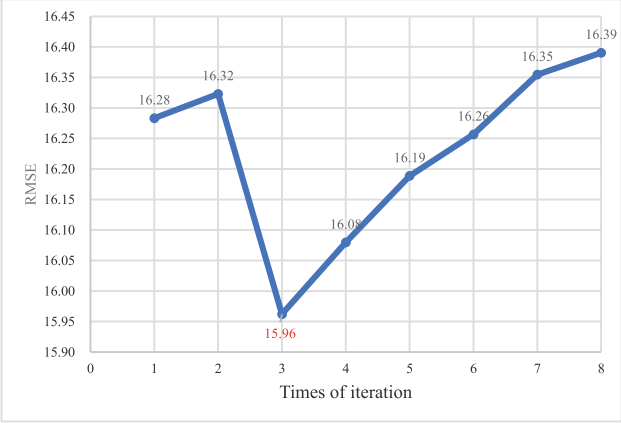

Fig. 5. The number of each iteration and the corresponding RMSE value.

Table 2. The relationship between the Shapley weight and RMSE across different LSTM models.

	MAPE	MAE	MAE%	RMSE	Shapley weight
LSTM_100	12.66451	0.015367	0.015945	16.03446	0.417
LSTM_300	13.88462	0.016847	0.017491	17.36751	−0.083
LSTM_300_2	12.37378	0.015014	0.015691	16.11064	0.417
LSTM450	13.67791	0.016596	0.017384	17.52009	0.25
Correlation Coefficient with Shapley weight	−0.86656	−0.86659	−0.83780	−0.76855	1

Table 3. Shapley weight of each Shapley Regression.

	Iteration			
Model	iter1	iter2	iter3	iter4
LSTM_100	0.417	0.283	0.283	0.267
LSTM_300	−0.083	−0.050	−0.217	−0.150
LSTM_300_2	0.417	0.283	0.333	0.133
LSTM450	0.250	0.200	−0.017	0.050
SR-LSTM 1	0.000	0.283	0.333	0.283
SR-LSTM 2	0.000	0.000	0.283	0.283
SR-LSTM 3	0.000	0.000	0.000	0.133

$$MAE = \frac{1}{n}\sum_{1}^{n}|e_t| \tag{20}$$

To express MAE as a percentage, divide it by the average demand:

$$MAPE\% = \frac{\frac{1}{n}\sum_{1}^{n}|e_t|}{\frac{1}{n}\sum_{1}^{n}d_t} = \frac{\sum_{1}^{n}|e_t|}{\sum_{1}^{n}d_t} \tag{21}$$

5. Root Mean Squared Error (RMSE) is the square root of the average of the squared errors:

$$RMSE = \sqrt{\frac{1}{n}\sum_{1}^{n}e_t} \tag{22}$$

where n is the number of historical periods where you have both a forecast and a demand.

RMSE% is the RMSE scaled to the demand:

$$RMSE\% = \frac{\sqrt{\frac{1}{n}\sum_{1}^{n}e_t}}{\sum_{1}^{n}d_t}. \tag{23}$$

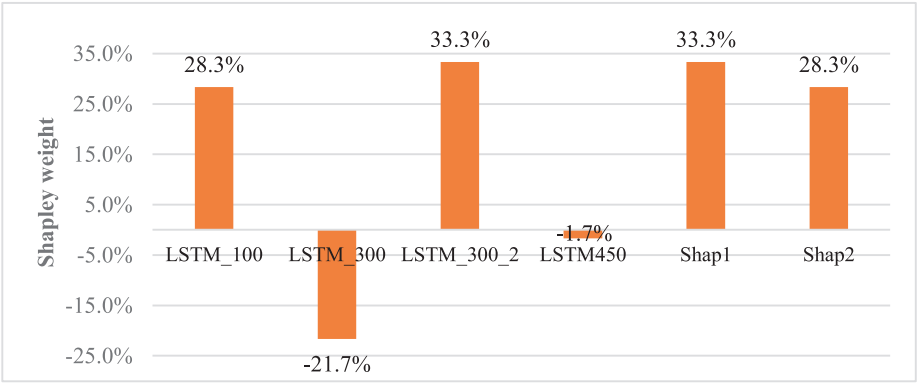

Fig. 6. Shapley weights of the best predicted data for the third iteration.

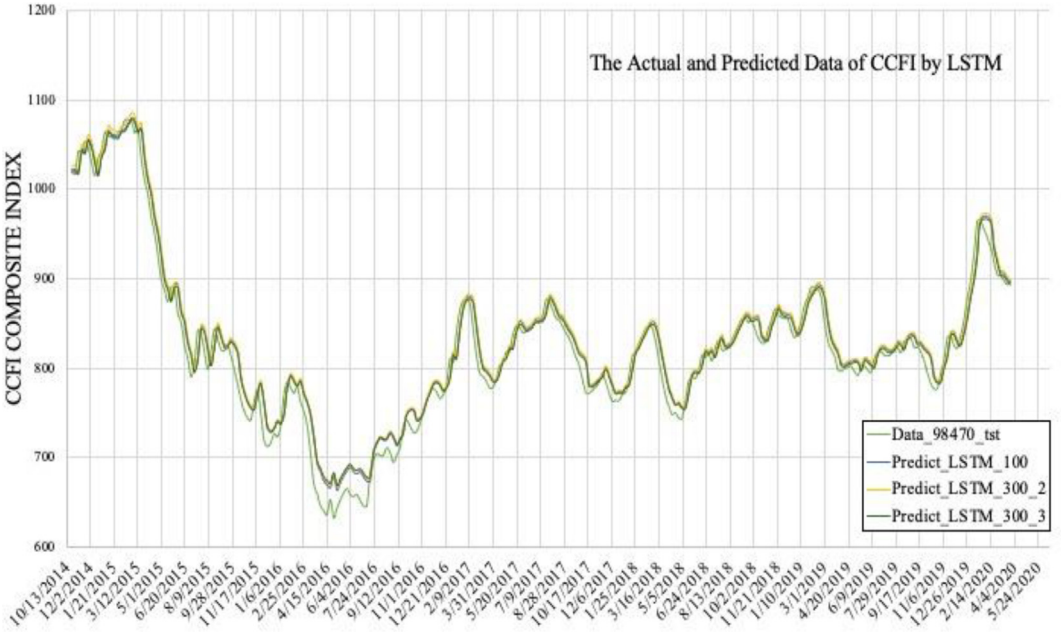

Fig. 7. The actual and predicted data of CCFI by traditional LSTM

Fig. 8. The actual and predicted data of CCFI by shapley regression- LSTM (SR-LSTM).

6. Mean Squared Error (MSE) is the average of the squared errors:

$$MSE = \frac{1}{n}\sum_{1}^{n} e_t^2 \tag{24}$$

where n is the number of historical periods where you have both a forecast and a demand.

4. Data analysis results

We generate prediction data using LSTM models with 200, 250, and 150 neurons—denoted as LSTM1,

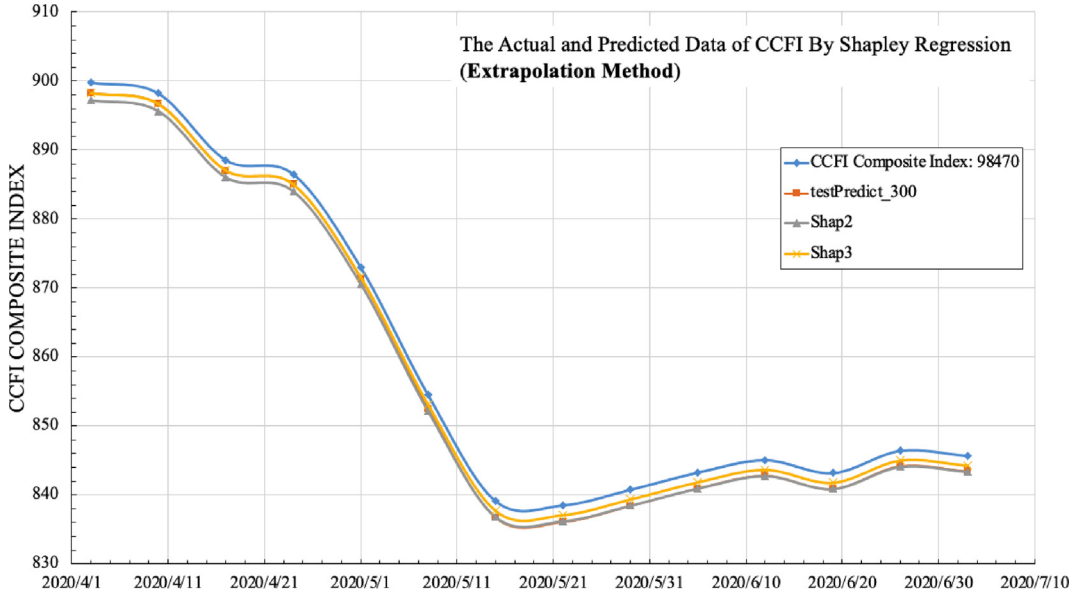

Fig. 9. The actual and predicted data of CCFI by shapley regression (extrapolation method).

LSTM2, and LSTM3, respectively. This data is then input into the Shapley regression algorithm for iterative processing. The outcome of the first iteration is designated as 'Shapley1', with subsequent iterations labeled sequentially up to 'Shapley8', corresponding to the eighth iteration. As indicated by Table 1, the Shapley regression algorithm significantly enhances the accuracy of the data. Furthermore, Fig. 5 reveals that the algorithm, akin to deep learning algorithms, is susceptible to overfitting after numerous iterations. Nonetheless, the algorithm's optimal prediction accuracy for the CCFI Composite Index was attained during the third iteration, within the timeframe of October 13, 2014, to May 24, 2020.

Table 3 and Fig. 6 underscore the efficacy of the Shapley weight calculation algorithm in enhancing predictive performance. Table 2 reveals a positive trend: as the prediction accuracy of LSTM models increases, so do their corresponding Shapley weights. This direct relationship is further confirmed by the negative correlation coefficients between the Shapley weights and various performance metrics—MAPE, MAE, MAE%, and RMSE—indicating that higher accuracy correlates with higher weights. The Correlation Coefficient with Shapley weight particularly highlights this relationship, offering a statistical measure of the strength and direction of the link between model accuracy and the allocation of Shapley weights. In our study, we define an effective coalition to be when the RMSE for the averaged predictions of two distinct LSTM models is lower than the RMSE of each individual model's prediction. This effective coalition, as depicted in Table 2, demonstrates that models yielding more accurate predictions receive higher Shapley weights, thereby validating the association between the precision of predictions and the Shapley weight distribution.

Figure 7 illustrates the actual versus predicted data for the China Containerized Freight Index (CCFI) as forecasted by the traditional LSTM model. The enhancement of predictive performance through the application of the Shapley regulation-LSTM is depicted in Fig. 8, where the integration of appropriate adjustments has improved the model's predictive reliability. Additionally, this approach allows for a more accurate estimation of the CCFI's behavioral trends.

From Fig. 9, we can compare the predicted data, generated by different iterative processes of this method, with the actual data. It appears that the predicted data progressively approaches the actual data more accurately, capturing and forecasting data trends effectively. In Fig. 9, 'CCFI' represents the actual CCFI index data; 'testPredict_300' denotes data produced by Traditional LSTM; 'Shap2' corresponds to data from SR-LSTM 2; and 'Shap 3' represents data from SR-LSTM 3.

5. Conclusion and suggestions

The results of this study reveal that the SR-LSTM neural network, developed under the Shapley mathematical framework, boasts robust mathematical logic, comprehensive data integration, high fault tolerance, automatic association adjustment, and effective noise filtering capabilities. This network is adept at constructing nonlinear models and surmounting the limitations inherent to traditional statistical methods, which often require extensive assumptions during model development. Advancements in computing power and efficiency greatly benefit the application of these findings to trend forecasting for indices with complete datasets, such as the CCFI, thereby enhancing forecasting precision.

Time series analysis typically encompasses five elements: Trend, Cycle, Seasonality, Random, and Irregular variations, with the latter often posing the greatest challenge to accurate predictions due to its erratic nature. To mitigate the effects of irregular variations, one might consider the fractal prediction approach pioneered by Benoît B. Mandelbrot, which trains on the data's self-similarity though the presence of Fat-tailed distributions, should not be disregarded. Future research may involve bifurcating big data into pre- and post-emergency states, such as pandemics, to refine noise filtration and boost predictive accuracy, thereby fine-tuning the parameter adjustment process. Going forward, the model's efficacy can be further enhanced by redefining the criteria for a winning coalition, thus progressively improving the model's predictive accuracy.

Conflict of interest

There is no conflict of interest.

References

[1] Hirata E, Matsuda T. Forecasting Shanghai Container freight index: a deep-learning-based model experiment. J Mar Sci Eng 2022;10(5):593.
[2] Makridakis S, Hibon M. The M3-Competition: results, conclusions and implications. Int J Forecast 2000;16(4):451−76.
[3] Shankar S, Ilavarasan PV, Punia S, Singh SP. Forecasting container throughput with long short-term memory networks. Industrial management & data systems; 2019.
[4] Clements MP, Franses PH, Swanson NR. Forecasting economic and financial time-series with non-linear models. Int J Forecast 2004;20(2):169−83.

[5] Alizadeh M, Gharakhani M, Fotoohi E, Rada R. Design and analysis of experiments in ANFIS modeling for stock price prediction. Int J Ind Eng Comput 2011;2(2):409−18.
[6] Kummong R, Supratid S. Thailand tourism forecasting based on a hybrid of discrete wavelet decomposition and NARX neural network. Industrial Management & Data Systems; 2016.
[7] Bodyanskiy Y, Popov S. Neural network approach to forecasting of quasiperiodic financial time series. Eur J Oper Res 2006;175(3):1357−66.
[8] Wong WK, Xia M, Chu W. Adaptive neural network model for time-series forecasting. Eur J Oper Res 2010;207(2): 807−16.
[9] Tseng F-M, Yu H-C, Tzeng G-H. Combining neural network model with seasonal time series ARIMA model. Technol Forecast Soc Change 2002;69(1):71−87.
[10] Wang K, Qi X, Liu H. Photovoltaic power forecasting based LSTM-Convolutional Network. Energy 2019;189:116225.
[11] Xiao Y, Xiao J, Wang S. A hybrid model for time series forecasting. Hum Syst Manag 2012;31(2):133−43.
[12] Zhang GP. Time series forecasting using a hybrid ARIMA and neural network model. Neurocomputing 2003;50: 159−75.
[13] Bengio Y, Simard P, Frasconi P. Learning long-term dependencies with gradient descent is difficult. IEEE Trans Neural Network 1994;5(2):157−66.
[14] Fischer T, Krauss C. Deep learning with long short-term memory networks for financial market predictions. Eur J Oper Res 2018;270(2):654−69.
[15] Hochreiter S, Schmidhuber J. Long short-term memory. Neural Comput 1997;9(8):1735−80.
[16] Hochreiter S, Schmidhuber J. LSTM can solve hard long time lag problems. Adv Neural Inf Process Syst 1997:473−9.
[17] Greff K, Srivastava RK, Koutník J, Steunebrink BR, Schmidhuber J. LSTM: a search space odyssey. IEEE Transact Neural Networks Learn Syst 2016;28(10):2222−32.
[18] Sutskever I, Vinyals O, Le QV. Sequence to sequence learning with neural networks. 2014. *arXiv preprint arXiv: 1409.3215*.
[19] Gers FA, Schmidhuber J, Cummins F. Learning to forget: continual prediction with LSTM. Neural Comput 2000;12(10): 2451−71.
[20] Graves A, Liwicki M, Fernández S, Bertolami R, Bunke H, Schmidhuber J. A novel connectionist system for unconstrained handwriting recognition. IEEE Trans Pattern Anal Mach Intell 2008;31(5):855−68.
[21] Märgner V, El Abed H. ICDAR 2009 Arabic handwriting recognition competition. In: 2009 10th international conference on document analysis and recognition. IEEE; 2009. p. 1383−7.
[22] Graves A, Mohamed A-r, Hinton G. Speech recognition with deep recurrent neural networks. 2013 IEEE international conference on acoustics, speech and signal processing. Ieee; 2013. p. 6645−9.
[23] Cho K, Merrienboer B, Gulcehre C, Bahdanau D, Bougares F, Schwenk H, et al. Learning phrase representations using RNN encoder-decoder for statistical machine translation. arXiv preprint arXiv:1406.1078. 2014.
[24] Beaufays F. The neural networks behind Google Voice transcription. Google Research blog; 2015.
[25] Adolfsson L. How will Artificial Intelligence impact the labour market, which jobs will be replaced and what will it mean for society, within the next decade?. ed, 2020.
[26] Metz C. An infusion of AI makes google translate more powerful than ever. Wired. com 2016;27.
[27] Wu Y, Schuster M, Chen Z, Le OV, Norouzi M, Macherey W, et al. Google's neural machine translation system: bridging the gap between human and machine translation. arXiv preprint arXiv:1609.08144. 2016.
[28] Vogels W. Bringing the magic of Amazon AI and Alexa to apps on AWS. All things distributed. 2016. https://www.allthingsdistributed.com/2016/11/amazon-ai-and-alexa-for-all-awsapps.html. 2016.
[29] Capes T, Coles P, Conkie A, Golipour L, Hadjitarkhani A, Hu Q, et al. Siri on-device deep learning-guided unit selection text-to-speech system. INTERSPEECH; 2017. p. 4011−5.
[30] Macho-Stadler I, Pérez-Castrillo D, Wettstein D. Sharing the surplus: an extension of the Shapley value for environments with externalities. J Econ Theor 2007;135(1):339−56.
[31] Moulin H. An application of the Shapley value to fair division with money. Econometrica. Journal of the Econometric Society; 1992. p. 1331−49.
[32] Pérez-Castrillo D, Wettstein D. Bidding for the surplus: a non-cooperative approach to the Shapley value. J Econ Theor 2001;100(2):274−94.
[33] Petrosjan L, Zaccour G. Time-consistent Shapley value allocation of pollution cost reduction. J Econ Dynam Control 2003;27(3):381−98.
[34] Dubey P. The shapley value as aircraft landing fees−revisited. Manag Sci 1982;28(8):869−74.
[35] Tan X, Lie T. Application of the Shapley value on transmission cost allocation in the competitive power market environment. IEE Proc Generat Transm Distrib 2002;149(1): 15−20.
[36] Bartholdi JJ, Kemahlioğlu-Ziya E. Using Shapley value to allocate savings in a supply chain. In: Supply chain optimization. Springer; 2005. p. 169−208.
[37] Narayanam R, Narahari Y. Topologies of strategically formed social networks based on a generic value function−allocation rule model. Soc Network 2011;33(1):56−69.
[38] Fu L, Lai G, Mahon PJ, Wang J, Zhu D, Jia B, et al. Carbon nanotube and graphene oxide directed electrochemical synthesis of silver dendrites. RSC Adv 2014;4(75):39645−50.
[39] Kemahlõoğlu-Ziya E, Bartholdi III JJ. Centralizing inventory in supply chains by using Shapley value to allocate the profits. Manuf Serv Oper Manag 2011;13(2):146−62.
[40] Wu L, Yang L, Liu H, Zhang Y. Shapely value-based traffic zone game partition in DRG systems. Traffic Eng Control 2013;54(3).
[41] Raghunathan S. Impact of demand correlation on the value of and incentives for information sharing in a supply chain. Eur J Oper Res 2003;146(3):634−49.
[42] Rosenthal EC. A game-theoretic approach to transfer pricing in a vertically integrated supply chain. Int J Prod Econ 2008; 115(2):542−52.
[43] Leng M, Parlar M. Allocation of cost savings in a three-level supply chain with demand information sharing: a cooperative-game approach. Oper Res 2009;57(1):200−13.
[44] Leng M, Parlar M. Lead-time reduction in a two-level supply chain: non-cooperative equilibria vs. coordination with a profit-sharing contract. Int J Prod Econ 2009;118(2):521−44.
[45] Liu X, Wang X, Qu Q, Zhang L. Double hierarchy hesitant fuzzy linguistic mathematical programming method for MAGDM based on Shapley values and incomplete preference information. IEEE Access 2018;6:74162−79.
[46] Li L, Wang X, Lin Y, Zhou F, Chen S. Cooperative game-based profit allocation for joint distribution alliance under online shopping environment. Asia Pac J Mark Logist 2019; 31(2):302−26.
[47] Liang P, Hu J, Liu Y, Chen X. Public resources allocation using an uncertain cooperative game among vulnerable groups. Kybernetes; 2019.
[48] Han T, Chen J, Wang L, Cai Y, Wang C. Interpretation of stability assessment machine learning models based on shapley value. 2019 IEEE 3rd conference on energy internet and energy system integration (EI2). IEEE; 2019. p. 243−7.
[49] Otero-Palencia C, Amaya−Mier R, Yie-Pinedo R. A stochastic joint replenishment problem considering transportation and warehouse constraints with gainsharing by Shapley Value allocation. Int J Prod Res 2019;57(10):3036−59.
[50] Jang Y, Jeong I, Cho YK. Identifying impact of variables in deep learning models on bankruptcy prediction of construction contractors. Engineering, Construction and Architectural Management; 2021.
[51] Shalit H. Using the Shapley value of stocks as systematic risk. J Risk Finance 2020;21(4):459−68.

[52] Chen Y, Liu B, Wang T. Analysing and forecasting China containerized freight index with a hybrid decomposition−ensemble method based on EMD, grey wave and ARMA. Grey Syst Theor Appl 2021;11(3):358−71.
[53] Munim ZH. State-space TBATS model for container freight rate forecasting with improved accuracy, vol. 3. Maritime Transport Research; 2022. p. 100057.
[54] Munim ZH, Schramm H-J. Forecasting container shipping freight rates for the Far East−Northern Europe trade lane. Marit Econ Logist 2017;19:106−25.
[55] Munim ZH, Schramm H-J. Forecasting container freight rates for major trade routes: a comparison of artificial neural networks and conventional models. Marit Econ Logist 2021; 23:310−27.
[56] Koyuncu K, Tavacioğlu L. Forecasting Shanghai containerized freight index by using time series models. Marine Science and Technology Bulletin 2021;10(4):426−34.
[57] Lin L, Wanwu L, Hang L, Yi S. Deep learning−based, OceanTDLx sea ice detection model for SAR image. J Mar Sci Technol 2023;31(1):3.
[58] Goodfellow I, Bengio Y, Courville A. Deep learning. MIT press; 2016.
[59] Salehinejad H, Sankar S, Barfett J, Colak E, Valaee S. Recent advances in recurrent neural networks. arXiv preprint arXiv: 1801.01078. 2017.
[60] Vandeput N. Data science for supply chain forecasting. In: Data science for supply chain forecasting. De Gruyter; 2021.