Forecasting Port Container Throughput with Deep Learning Approach

Fuxin Jiang
Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China and University of Chinese Academy of Sciences, Beijing 100049, China
jiangfuxin17@mails.ucas.ac.cn

Gang Xie*
Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China
gxie@amss.ac.cn

Shouyang Wang
Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China
sywang@amss.ac.cn

ABSTRACT

Due to the international transfer of manufacturing industry, the change of trade policy and frequent irregular events in the global trade, it becomes more difficult to predict port container throughput accurately. In order to improve the predictive accuracy, we develop a bidirectional long short-term memory network model to forecast the throughput. Using the data of port in Qingdao, this study investigates for the first time how to use the deep learning approach to predict port container throughput. The empirical results show that the proposed model can achieve highest average predictive accuracy, which indicates that the approach is effective in the increasingly complex trade situation.

CCS CONCEPTS

• **Applied computing**; • **Operations research**; • **Forecasting**;

KEYWORDS

Container throughput forecasting, bidirectional long short-term memory network model

1 INTRODUCTION

Port container transportation is an important guarantee of international trade, which is closely related to the core interests of many enterprises and even the economy of relevant countries [1]. Container throughput is one of the most important indicators to measure the development of port economy. Accurate prediction of port container throughput can provide necessary information for port management, port related enterprises, city and national economic development decision-making [2].

In recent years, a large number of forecasting methods have been proposed to forecast port containers, mainly divided into univariate methods and multivariable methods. For the univariate methods, most of the researches on container throughput forecasting use univariate time series forecasting models, including econometric models, artificial intelligence (AI) models, hybrid models and decomposition-ensemble models. The econometric models include linear regression [3], the autoregressive integrated moving average (ARIMA) model [4], the seasonal autoregressive integrated moving average (SARIMA) model [5], the Holt-Winters exponential smoothing [6], and vector error correction model [7]. AI models are grey Verhulst model [8], multilayer perceptron neural network (MLP) [3], radial basis function (RBF) neural networks [9], genetic programming [10], back propagation neural network (BPNN) [11], support vector regression (SVR), and least squares support vector regression (LSSVR) [12]. Generally, AI models can deal with the nonlinear information of data, so their forecasting performance is better than econometric models.

Besides single forecasting models, hybrid models are developed to effectively utilize the advantages of different models. Combining the grey model with the classical decomposition model, Peng and Chu [5] proposed a hybrid grey model, which extended the analysis by changing the size of the initial sequence in grey forecast and searched for the one that has the lowest prediction errors. Tian et al. [13] presented an integrated forecasting model, which includes four steps: an econometric model for the linear component, a RBF neural network for the nonlinear component, impact assessment of economic and other events, integrating predictions. The decomposition-ensemble methodology has also been used in the prediction of container throughput at ports. Extending the classical decomposition model in Peng and Chu [5], Xie et al. [14] proposed three hybrid approaches based on seasonal decomposition and LSSVR model. Continuously, Xie et al. [14] proposed a novel decomposition-ensemble methodology, which includes four steps: data decomposition, data characteristic analysis, individual forecasts, ensemble forecast. Consequently, there is the need for an alternative forecasting model.

For the multivariable methods, most researchers utilized economic indexes as explanatory variables [15, 16]. Chou et al. [17] predicted Taiwan's import container volume with economic indicators, including export container volume, import container volume, population, industrial production index, GNP, GNP per capita, wholesale price, GDP, agricultural GDP, industrial GDP, and service GDP. Using the port throughput data and socio-economic

index data of Shanghai city, Geng et al. [18] evaluate the feasibility and effectiveness of the proposed scheme. Intihar et al. [19] examined the impact of integration of macroeconomic indicators on the predictive accuracy of forecasting model. Using cross-correlation function to test relationships between different economic variables and container throughput, Rashed et al. [20] found that industrial confidence indicator in the Euro Area (ICI_EA) with two lags shows the best fit as a leading indicator.

In practice, many studies use both economic and trade indexes in the prediction of container throughput at ports. Dagenais and Martin [21] proposed a long-term prediction method for the overseas container traffic volume of Montreal port. The forecasting model considers the expected changes of three factors: the volume of international trade between Canada and the USA, the rate of containerization and the growth of Canada's domestic region.

Besides economic and trade indexes, some correlated time series data of container throughput at other ports are used to assist in modeling the time series at target ports [22]. In addition to the total trade volume and international transportation flow, de Langen [23] also used three variables related to the containerization proportion of transportation flow, including the share of transportation flow that can be containerized, the containerization rate and the share of shipping in international trade. Many econometric models have been used for forecasting container throughput with explanatory variables. Chou et al. [17] proposed a modified regression model with economic data for forecasting the volumes of Taiwan's import containers. van Dorsser et al. [15] developed a regression model with GDP to predict a very long term of the total port throughput in the Le Havre - Hamburg range up to 2100. Another econometric model is ARIMAX, which has been used for the prediction of container throughput at the Port of Koper [19] and the port of Antwerp [20], respectively. These studies showed that more accurate future throughput forecasts can be achieved by applying explanatory variables to the models. Simultaneously, AI models are also used as multivariate methods for the prediction of container throughput at ports. Using transfer learning technique and a pattern matching method, Xiao et al. [22] proposed a TF-DPSO model. Geng et al. [18] developed a port throughput forecasting scheme by hybridizing the RSVR, chaotic simulated annealing particle swarm optimization (CSAPSO) algorithm and multivariable adaptive regression splines (MARS). Here, MARS is adopted to select the final inputs for RSVR model. Tsai & Huang [24] developed the MLPNN model to predict container flows between the major ports of Asia.

In this study, in order to enhance predictive accuracy, we develop a BLSTM model to forecast container throughput.Additionally, an empirical study is conducted to illustrate the proposed methodology. Using data of Qingdao port, this study is the first to investigate how search big data is used in a deep learning approach for container throughput forecasting.

The remainder of this study proceeds as follows. Section 2 introduces the proposed framework and method in detail. Section 3 illustrates the proposed methodology by empirical study.

2 METHODOLOGY

In order to solve the problems mentioned above, we propose a novel framework, where BLSTM model is used to enable more accurate forecasts of container throughput at ports.

2.1 BLSTM Neural Network

Long short-term memory (LSTM) network is a deep learning technique that can be used to model nonlinear and non-stationary data [25]. It is difficult for the standard recurrent neural network (RNN) to learn the long-term dependence relationship because of the gradient vanishing or gradient explosion in the training process. In order to solve this problem, Hochreiter and Schmidhuber [26] proposed LSTM, which is also a special RNN. The architecture of the LSTM is shown in Figure 1(a), where the LSTM unit is composed of cell, forget gate, input gate and output gate. LSTM can remove or add the information to cell state through a gate, which is a way for information to pass selectively. The gate structure is implemented by sigmoid function. The output of sigmoid function is a value between 0 and 1, describing how much information can pass. 0 means that no information is allowed to pass, while 1 means that any information is allowed to pass.

The working principle of LSTM can be described as follows:

First, LSTM needs to decide which information to be discarded from the memory state, which is realized by forget gate.

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f)$$

After the forget gate, the LSTM needs to decide what cell states should be added. This process is mainly divided into two steps: first, input gate determines which information need to be updated, and then tanh generates a new candidate vector $\tilde{C}_t$.

$$i_i = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i)$$

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C)$$

The LSTM unit then multiplies i_t and $\tilde{C}_t$ to update the memory state as follows.

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t$$

Finally, output gate is used to output memory state

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o)$$

$$h_t = o_t * \tanh(C_t)$$

As we have mentioned above, LSTM can only process forward information. In contrast to LSTM, BLSTM includes forward and backward LSTM layers, as shown in Figure 1(b). Therefore, BLSTM can process both forward and backward information simultaneously [27].

2.2 Model Evaluation

As ARIMA, BPNN and LSSVR models have been used for container throughput forecasting at ports, they are used as benchmarks. In addition, the LSTM model is also used for comparison with BLSTM. In order to realize the comprehensive evaluation of models, both point and interval forecasts are generated for each model. Root mean square error (MSE), mean absolute error (MAE), mean absolute percentage error (MAPE) are used as the evaluation criteria
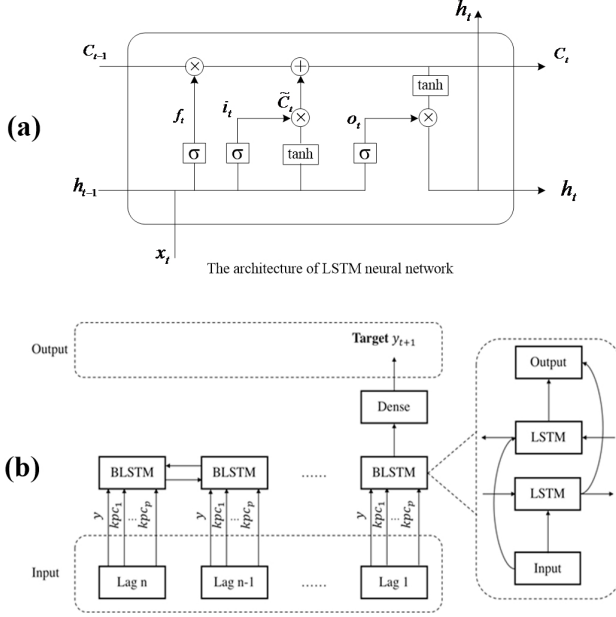
Figure 1: BLSTM for Container Throughput Forecasting.

Figure 2: Monthly Container Throughput at Qingdao Port.

Table 1: Parameters of Forecasting Models

Model	ARIMA	LSTM	BLSTM
Parameter	(1,0,1)	(16,4,0,0)	(16,4,0,0)

and directional accuracy (DA), respectively, as follows:

$$RMSE = \sqrt{\frac{1}{N}\sum_{t=1}^{N}(y_t - \hat{y}_t)^2}$$

$$MAE = \frac{1}{N}\sum_{t=1}^{N}|y_t - \hat{y}_t|$$

$$MAPE = \frac{1}{N}\sum_{t=1}^{N}(|y_t - \hat{y}_t|/|y_t|)$$

$$DA = \frac{1}{N}\sum_{t=1}^{N}d_t, \quad d_t = \begin{cases} 1 & if\ (y_{t+1} - y_t)(\hat{y}_{t+1} - y_t) \geq \\ 0 & otherwise \end{cases}$$

where N is the size of the testing set, y_t and $\hat{y}_t$ are actual and predicted values at period t.

3 EMPIRICAL STUDY

This section takes the container throughput data of Qingdao port in China as an example to verify the method proposed in this paper. The data and predictive results are introduced as follows.

3.1 The Data

The monthly data of container throughput at ports were collected from the WIND Database [28], as shown in Figure 2. The sample data for each port includes 105 observations covering the period from January 2011 to April 2021. In experiments, the first 99 and the last 25 observations were used as training and testing sets, respectively.
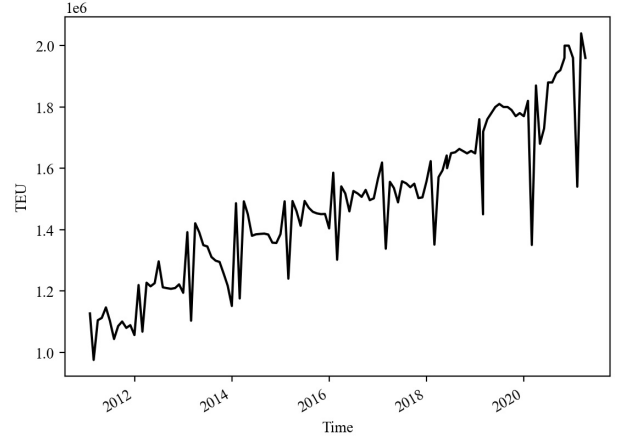
3.2 Predictive Result

Using ARIMA, LSTM and BLSTM models with different independent variables, the static forecast of port container throughput was implemented. According to Schwarz criterion (SC) and Akaike information criterion (AIC), the parameters of ARIMA models were determined. In the LSTM and BLSTM models, the time step is set to 12. Due to the limited data length of training set, both models adopted the regularization technique. Dropout and L1 regularization were presented in the model training phase. The last 20% of the training set is set as the validation set. The parameters of BLSTM was obtained by using the grid search. The LSTM hidden unit $p \in \{8, 16, 32, 64, 128\}$, the dropout rate $d \in \{0, 0.1, 0.2, 0.3\}$, L1 regularization coefficient $l_1 \in \{0, 0.1, 0.2, 0.3, 0.5, 1, 1.5, 2, 2.5, 3\}$. Similarly, the grid search shows that the benchmark method LSTM has the same hidden unit, the dropout rate, and L1 regularization coefficient as BLSTM. The learning rate is set to 0.01, the number of training epochs is set to 100, the training batch size $b \in \{2, 4, 8, 16, 32\}$.

The parameters of forecasting models are shown in Table 1, where (16, 16, 0, 0) in LSTM/BLSTM means that the number of hidden cells is 16, batch size is 16, dropout is 0 and the regularization parameter is 0. The actual value and predicted values of container throughput by different models for Qingdao are shown in Figure 3

As shown in Table 2, BLSTM model prediction is better than LSTM and ARIMA in various evaluation criteria. Compared with LSTM (ARIMA), BLSTM improves 5.3% (14.7%) in RMSE, 6.7% (21.5%) in MAE, 6.7% (26.7%) in MAPE, and 25% (25%) in DA.

4 CONCLUSION

This study investigates how to forecast container throughput by using a deep learning approach. Using the data of Qingdao port

Table 2: Performance Comparison of Different Methods

Model	RMSE	MAE	MAPE	DA
ARIMA	83142.6	66121.6	0.038	0.64
LSTM	74839.3	55609.6	0.032	0.64
BLSTM	70897.5	51888.2	0.03	0.8
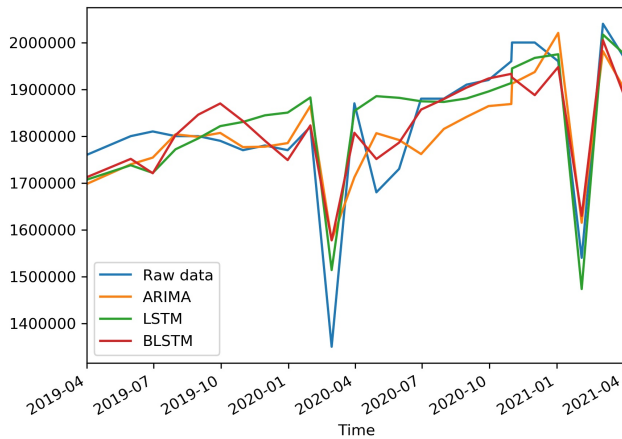
Figure 3: Monthly Container Throughput at Qingdao Port.

in China, empirical analysis shows that the proposed model can achieve higher predictive accuracy than benchmark models in point forecasting. This study only forecasts the port containers in Qingdao and can be extended to other port container forecasts in the future.

ACKNOWLEDGMENTS

REFERENCES

[1] Cheon, S.H., C.Y. Lee, Y. Wang. (2017). Processing time ambiguity and port competitiveness. Production and Operations Management, 26(12), 2187-2206.

[2] Xiao, Y., S. Wang, J.J. Liu, J. Xiao, Y. Hu. (2016). Throughput estimation based port development and management policies analysis. Maritime Policy & Management, 43(1), 84-97.

[3] Gosasang, V., W. Chandraprakaikul, S. Kiattisin. (2011). A comparison of traditional and neural networks forecasting techniques for container throughput at Bangkok port. The Asian Journal of Shipping and Logistics, 27(3), 463-482.

[4] Lee, S.-Y., H. Lim, H.-J. Kim. (2017). Forecasting container port volume: Implications for dredging. Maritime Economics & Logistics, 19(2), 296-314.

[5] Peng, W.-Y., C.-W. Chu. (2009). A comparison of univariate methods for forecasting container throughput volumes. Mathematical and Computer Modelling, 50, 1045-1057.

[6] Schulze, P.M., A. Prinz. (2009). Forecasting container transshipment in Germany. Applied Economics, 41, 2809-2815.

[7] Pang, G., B. Gebka. (2017). Forecasting container throughput using aggregate or terminal-specific data? The case of Tanjung Priok Port, Indonesia. International Journal of Production Research, 55(9), 2454-2469.

[8] Guo, Z., X. Song, J. Ye. (2005). A Verhulst model on time series error corrected for port throughput forecasting. Journal of the Eastern Asia society for transportation studies, 6, 881-891.

[9] Goulielmos, A.M., E. Kaselimi. (2011). A non-linear forecasting of container traffic: the case-study of the Port of Piraeus, 1973-2008. International Journal of Shipping and Transport Logistics, 3(1), 72-99.

[10] Chen, S.H., J.N. Chen. (2010). Forecasting container throughputs at ports using genetic programming. Expert Systems with Applications, 37, 2054-2058.

[11] Huang, A., K.K. Lai, Y. Li, S. Wang. (2015). Forecasting container throughput of Qingdao port with a hybrid model. Journal of Systems Science and Complexity, 28 (1), 105-121.

[12] Xie, G., S. Wang, Y. Zhao, K.K. Lai. (2013). Hybrid approaches based on LSSVR model for container throughput forecasting: A comparative study. Applied Soft Computing, 13, 2232-2241.

[13] Tian, X., L. Liu, K.K. Lai, S. Wang. (2013). Analysis and forecasting of port logistics using TEI@I methodology. Transportation Planning and Technology, 36(8), 685-702.

[14] Xie, G., N. Zhang, S. Wang. (2017). Data characteristic analysis and model selection for container throughput forecasting within a decomposition-ensemble methodology. Transportation Research Part E: Logistics and Transportation Review, 108, 160-178.

[15] van Dorsser, C., M. Wolters, B. van Wee. (2012). A very long term forecast of the port throughput in the Le Havre - Hamburg Range up to 2100. European Journal of Transport and Infrastructure Research, 12(1), 88-110.

[16] Li, M.-W., J. Geng, W.-C. Hong, Z.-Y. Chen. (2017). A novel approach based on the Gauss-vSVR with a new hybrid evolutionary algorithm and input vector decision method for port throughput forecasting. Neural Computing and Applications, 28, 621-640.

[17] Chou, C.-C., C.-W. Chu, G.-S. Liang. (2008). A modified regression model for forecasting the volumes of Taiwan's import containers. Mathematical and Computer Modelling, 47, 797-807.

[18] Geng, J., M.W. Li, Z.H. Dong, Y.S. Liao. (2015). Port throughput forecasting by MARS-RSVR with chaotic simulated annealing particle swarm optimization algorithm. Neurocomputing, 147, 239-250.

[19] Intihar, M., T. Kramberger, D. Dragan. (2017). Container throughput forecasting using dynamic factor analysis and ARIMAX model. Promet - Traffic&Transportation, 29(5), 529-542.

[20] Rashed, Y., H. Meersman, E. Van de Voorde, T. Vanelslander. (2017). Short-term forecast of container throughout: An ARIMA-intervention model for the port of Antwerp. Maritime Economics & Logistics, 19, 749-764.

[21] Dagenais, M.G., F. Martin. (1987). Forecasting containerized traffic for the port of Montreal (1981-1995). Transportation Research Part A: General, 21(1), 1-16.

[22] Xiao, J., Y. Xiao, J. Fu, K.K. Lai. (2014). A transfer forecasting model for container throughput guided by discrete PSO. Journal of Systems Science and Complexity, 27, 181-192.

[23] de Langen, P.W. (2003). Forecasting container throughput: A method and implications for port planning. Journal of International Logistics and Trade, 1(1), 29-39.

[24] Tsai, F.-M., L. J. W. Huang. (2017). Using artificial neural networks to predict container flows between the major ports of Asia. International Journal of Production Research, 55(17), 5001-5010.

[25] Law, R., G. Li, D.K.C. Fong, X. Han. (2019). Tourism demand forecasting: A deep learning approach. Annals of Tourism Research, 75, 410-423.

[26] Hochreiter, S., J. Schmidhuber. (1997). Long short-term memory. Neural computation, 9(8), 1735-1780.

[27] Yildirim, Ö. (2018). A novel wavelet sequence based on deep bidirectional LSTM network model for ECG signal classification. Computers in Biology and Medicine, 96, 189-202.

[28] Wind Info. 2019. http://www.wind.com.cn. (Dec 20, 2019).