

# *E-commerce Logistics Transportation Prediction Problem Based on ARMA and LSTM Neural Networks*

Hantao Zhang<sup>1,a,\*</sup>, Xiaoxuan Xie<sup>1,b</sup>

<sup>1</sup>*School of Economics and Modern Finance, Gannan University of Science and Technology,  
Ganzhou, China*

<sup>a</sup>*z18017817980@163.com*, <sup>b</sup>*x1021194282@163.com*

*\*Corresponding author*

**Keywords:** E-Commerce Logistics, ARMA, LSTM, Neural Network, Logistics Network

**Abstract:** Today in the Internet era, online shopping has become an indispensable part of life, then the transportation of e-commerce logistics has become a major problem, if the logistics site is out of service, it will inevitably lead to problems in processing and transportation, at this time, it is necessary to predict the processing and transportation capacity of each logistics site, to ensure that the logistics of the normal operation of the logistics, and at the same time, designing alternatives can greatly reduce the impact of the out-of-service. This paper establishes a prediction model combining ARMA and LSTM to carry out an in-depth study on the emergency call of logistics and logistics network. In this paper, we first pre-processed the data, made a pivot table based on the existing data, which is convenient for observation and application, and then established an ARMA model, and found that the prediction results were inaccurate, and then combined with the LSTM neural network to weight the value of the prediction, and finally obtained the DC14→DC10, DC20→DC35, DC25→DC62 three lines from January 1, 2023 to January 31, 2023 daily cargo volume.

## 1. Introduction

The e-commerce logistics network is composed of individual logistics sites. During holidays such as the Spring Festival and the National Day, as well as promotional activities such as the Double Eleven and Double Twelve, users tend to consume a large amount, resulting in a surge in the number of orders, and when the logistics site cannot be used due to natural and man-made disasters, the parcels to be processed will be diverted from the logistics site to other logistics sites. Whether it is an increase in logistics due to people's consumption or a diversion of logistics due to warehouse outage, it will affect the number of transportation on the logistics routes as well as the burden on the logistics sites to handle the parcels. Therefore, if we can predict the number of parcels in each place and route, the manager can make preparations in advance to reduce costs and improve efficiency.

## 2. Related Works

In recent years, teaching and research in the logistics program has received a lot of attention, and scholars have devoted themselves to exploring different pedagogical approaches to improve graduate students' application of quantitative analysis methods. First, Ding, Yang, and Huang, Liang (2019) conducted a case study on the teaching of quantitative analysis methods in a graduate logistics program [1], highlighting the key role of real-world examples in helping students understand and apply quantitative methods. Woschank and Pacher (2020) presented a study on teaching and learning methods in industrial logistics engineering education, with a paper focusing on the holistic teaching methods in LOGILAB at Montanuniversitaet Leoben [2-3]. Senna et al. (2013) focused on the challenges of teaching business logistics to international students, emphasizing the impact of cultural differences on teaching and learning [4]. For their part, Qiang Zhang, Yonggang and Min Zhang (2021) wrote Logistics Systems Engineering - Theory, Methods and Case Studies [5], which provide an in-depth introduction to the theory and methods of logistics systems engineering. In terms of involving prediction and analysis, Mgandu et al. (2020) used exponential smoothing for trend analysis and prediction of water levels in Mtera Reservoir [6], while Sakpere et al. (2021) investigated the effect of COVID-19 on students' academic performance through correlation and regression modeling [7]. Padmaja and Haritha (2017) and Padmaja and Haritha (2015), on the other hand, investigated the application of grey correlation analysis in software workload estimation from a software engineering perspective [8][11]. Chen et al. (2021) investigated the prediction modeling of the dehumidification wheel air outlet states through multiple regression and artificial neural network methods [9]. Luu et al. (2021) tested a multiple linear regression system using a variance testing method [10]. In addition, Yichung (2020) constructed a gray prediction model for magnesium material demand through gray correlation analysis and neural networks [12], while Zeng and Liu (2012) proposed a prediction model based on amplitude compression of stochastic oscillatory sequences [13]. Together, this body of literature constitutes a broad and deep research area for teaching and research in logistics, covering a wide range of topics from case-based teaching to gray correlation analysis, predictive modeling, and international student teaching challenges [14].

This paper explores the problem of e-commerce logistics movement prediction using both ARMA (AutoRegressive Moving Average) and LSTM (Long Short-Term Memory) neural networks. The ARMA model is used to capture autoregressive and sliding average relationships in time series data, while the LSTM neural network improves the prediction performance by learning the long-term dependencies to improve the prediction performance. By comparing the performance of the two methods, the study aims to find an effective model that is most suitable for e-commerce logistics movement forecasting, which will provide a reference for improving the efficiency of logistics operation and service level. The results of this study will help optimize e-commerce logistics planning and resource allocation to cope with the increasingly complex market demand and dynamic supply chain environment.

## 3. Theory and Method

### 3.1 Data Pre-processing

The data was first organized using a pivot table to obtain the annual and monthly volumes for each route, which was used to predict the trend of the overall volume using SPSS as shown in Figure 1 below.

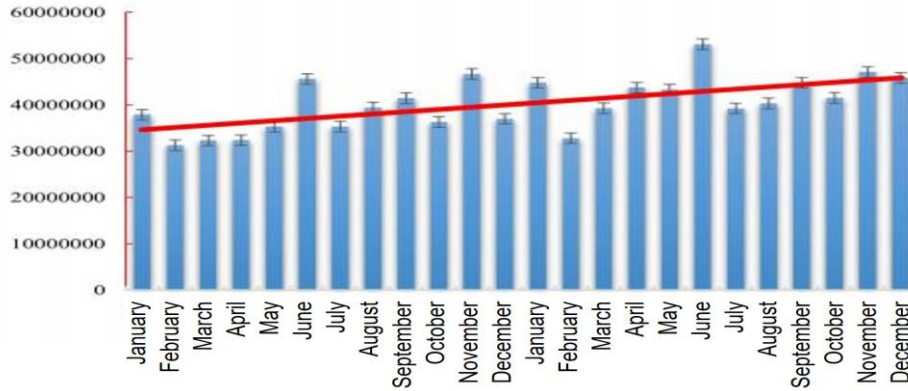


Figure 1: Trend chart of cargo volume of all logistics nodes from 2021 to 2022

As shown in Fig.1, the change trend of cargo volume of all logistics nodes in 2021-2022 is showing an upward trend, so it can be assumed that the cargo volume of most routes is showing an upward trend.

### 3.2 Time Series Model

Time series analysis is the theory and method of establishing mathematical models through curve fitting and parameter estimation based on systematically obtained time series data. Generally, curve fitting and parameter estimation methods, such as nonlinear least squares, are used. The smoothness of the smooth time series model refers to the wide range of smoothness, which is characterized by the statistical properties of the level does not change over time, that is, the mean and covariance do not change over time.

The general formula for the general moving average model MA(q) is as follows.

$$Y_t = e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \dots - \theta_q e_{t-q} \quad (1)$$

The upper equation of theq A moving average process of order, recorded as MA (q)

The general autoregressive model AR (p) has the following general formula.

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + e_t \quad (2)$$

The above equation is an autoregressive process ofp autoregressive process of order,recorded asAR(p).Therefore, the autoregressive moving average mixed model ARMA (p, q) has the following general formula.

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \dots - \theta_q e_{t-q} \quad (3)$$

The above equation is an autoregressive moving average mixed process of order p andq ,recorded as ARMA(p ,q),from the general formula,it can be seen that the model ARMA is a combination of model AR and model MA, which has both properties.

For the normal model ARMA (p, q), the general prediction form is.

$$\begin{aligned} \hat{Y}_t(\zeta) = & \phi_1 \hat{Y}_t(\zeta - 1) + \phi_2 \hat{Y}_t(\zeta - 2) + \dots + \phi_p \hat{Y}_t(\zeta - p) + \theta_0 \\ & - \theta_1 E(e_{t+\zeta-1} | Y_1, Y_2, \dots, Y_t) - \theta_2 E(e_{t+\zeta-2} | Y_1, Y_2, \dots, Y_t) \\ & - \theta_q E(e_{t+\zeta-q} | Y_1, Y_2, \dots, Y_t) \end{aligned} \quad (4)$$

In the above equation, the  $\zeta$  denotes the lag time unit, i.e.

$$\hat{Y}_t(0) = Y_t, E(e_{t+j} | Y_1, Y_2, \dots, Y_t) = \begin{cases} 0 & j > 0 \\ e_{t+j} & j < 0 \end{cases} \quad (5)$$

The variables involved in this problem and the data collected are more in line with this model. Time is a linear variable, so this traditional model was chosen.

### 3.3 LSTM Neural Network

LSTM neural network is a complex network system formed by extensive interconnection of a large number of simple processing units. This neural network has massively parallel, distributed storage and processing, self-organization, self-adaptation and self-learning capabilities. Combining AMRA time series with LSTM neural network can make the derived data more accurate.

As shown in Figure 2, the neural network is divided into input, hidden and output layers. The point of the hidden layer is to abstract the features of the input data into another dimension to show its more abstract features. These features allow for better linear discretization and can be changed as needed.

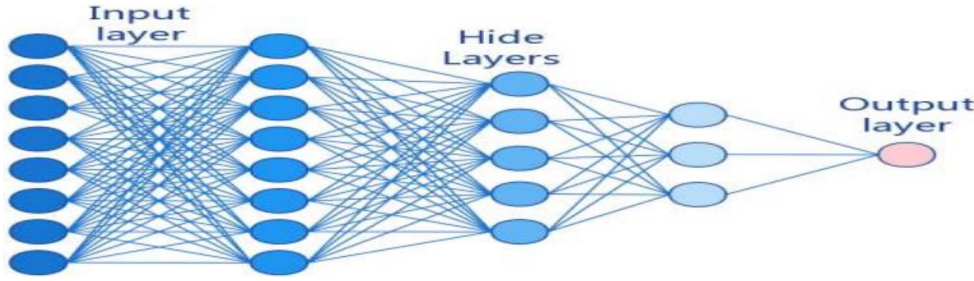


Figure 2: The neural network model

The input layer is usually a  $n$  dimensional vector, denoted as

$$(X_1 \ X_2 \ X_3) \quad (6)$$

The connection between the input layer and the hidden layer, the hidden layer, and the output layer are the weights, denoted by

$$W = \begin{vmatrix} w_{11} & w_{12} & w_{13} \\ w_{21} & w_{22} & w_{23} \\ w_{31} & w_{32} & w_{33} \end{vmatrix} \quad H = \begin{vmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{vmatrix} \quad (7)$$

The output layer is usually a  $n$  dimensional vector or a value, denoted as

$$(y_1 \ y_2 \ y_3) \quad (8)$$

If no activation function is used, the output of each layer is a linear function of the input of the previous layer. No matter how many layers the neural network has, the output is a linear combination of the inputs. By using an activation function, it is possible to introduce a nonlinear factor into the neurons, allowing the neural network to approximate any nonlinear function at will. The variables in this paper meet the nonlinear fitting predictions, thus allowing the neural network to be used in a nonlinear model of the problem. The common activation functions are as follows. In this paper, we propose to use the Relu function to output the hidden layer neurons, and the corresponding equations are as follows.

$$f_{relu}(u) = \begin{cases} 0 & u < 0 \\ u & u \geq 0 \end{cases} \quad (9)$$

Taking a simple three-layer neural network structure as an example, the initialization of weights

$W$  and  $H$  are randomly initialized. If no activation function is added, the computation process is as follows.

$$Y = X * W * H$$

$$(y_1 \ y_2 \ y_3) = (X_1 \ X_2 \ X_3) \begin{pmatrix} w_{11} & w_{12} & w_{13} \\ w_{21} & w_{22} & w_{23} \\ w_{31} & w_{32} & w_{33} \end{pmatrix} \begin{pmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{pmatrix} \quad (10)$$

In general, an offset term  $b$  is added, so that the linear process is.

$$(y_1 \ y_2 \ y_3) = (X_1 \ X_2 \ X_3) \begin{pmatrix} w_{11} & w_{12} & w_{13} \\ w_{21} & w_{22} & w_{23} \\ w_{31} & w_{32} & w_{33} \end{pmatrix} + b_1 \begin{pmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{pmatrix} + b_2 \quad (11)$$

For ease of calculation and documentation, the above equation is calculated and documented as follows.

$$(y_1 \ y_2 \ y_3) = (\beta_1 \ \beta_2 \ \beta_3) \quad (12)$$

Since the problem to be solved in this paper is nonlinear, a nonlinear activation function can be added, and a nonlinear activation function can be added according to the purpose of the study. The purpose of this paper is to add an output layer. The aim of this paper is to add an output layer: the

$$Y = f[(X * W + b_1) * H + b_2] \quad (13)$$

## 4. Results and Discussions

### 4.1 Time Series Modeling Results

In this paper, we intend to build an ARMA time series model to forecast the cargo volume in January 2023 in the future. For this purpose, we need to test whether the data satisfy smoothness, and considering that the difference between the before and after data of 2021-2022 is too large, the data of the last three months of 2022 is selected for forecasting in this paper.

Table 1: ADF test table

Variant	Difference in order	t	P	AIC	1%	5%	10%
DC14-DC10	0	-4.608	0.000***	1040.596	-3.513	-2.897	-2.586
	1	-7.042	0.000***	1029.944	-3.519	-2.9	-2.587
	2	-6.35	0.000***	1034.019	-3.525	-2.903	-2.589
Variant	Difference in order	t	P	AIC	1%	5%	10%
DC25-DC62	0	-2.681	0.077*	1823.35	-3.506	-2.895	-2.584
	1	-10.739	0.000***	1807.695	-3.507	-2.895	-2.585
	2	-6.491	0.000***	1798.766	-3.517	-2.899	-2.587
Variant	Difference in order	t	P	AIC	1%	5%	10%
DC20-DC35	0	-8.237	0.000***	1591.641	-3.504	-2.894	-2.584
	1	-7.403	0.000***	1580.6	-3.509	-2.896	-2.585
	2	-6.793	0.000***	1579.056	-3.514	-2.898	-2.586

From the above Table 1, it can be seen that all lines are smooth time series at all differences except DC25-DC62 which is an unsteady time series at the time of 0th order difference, which is next tested for white noise.

Figure 3 show the partial autocorrelation plot (PACF) of the model residuals, including the coefficients, upper confidence limits, and lower confidence limits. The correlation coefficients are all within the dashed line, the residuals of the sliding average model (MA) are white noise series, and the time series requires the model residuals to be white noise series. Therefore, the ARMA time

series model can be used for this data.

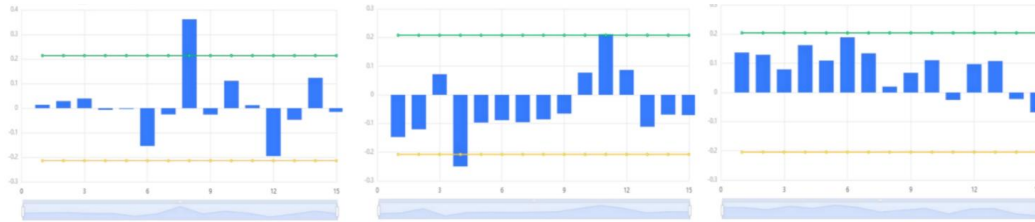


Figure 3: Model residual partial autocorrelation plot

## 4.2 Modeling Results for Combining Time Series and LSTM

Based on the processed data from the annexes, time series prediction of DC14-DC10, DC25-DC62, DC20-DC35, and LSTM neural network using SPSS, in which the weight of ARMA is 0.25 and the weight of LSTM neural network is 0.75, the following results are obtained in Table 2:

Table 2: Prediction Results

Time	DC14-DC10	DC20-DC35	DC25-DC62
2023.1.1	27479	102	15729
2023.1.2	26683	104	13060
2023.1.3	26624	111	10084
2023.1.4	27186	115	11654
2023.1.5	28018	119	6841
2023.1.6	28715	123	8300
2023.1.7	28801	128	13041
2023.1.8	28229	135	13846
2023.1.9	27452	156	11645
2023.1.10	26831	165	7462
2023.1.11	26406	185	7514
2023.1.12	26037	149	8431
2023.1.13	25578	191	10800
2023.1.14	25057	183	13040
2023.1.15	25177	165	8624
2023.1.16	28357	185	11505
2023.1.17	40174	198	11555
2023.1.18	65243	206	8800
2023.1.19	67847	217	7934
2023.1.20	41639	227	15844
2023.1.21	34165	233	9605
2023.1.22	34615	237	5740
2023.1.23	31294	235	8355
2023.1.24	30054	220	10268
2023.1.25	30935	225	9707
2023.1.26	35216	264	8355
2023.1.27	42516	365	10268
2023.1.28	48249	784	7307
2023.1.29	48628	1492	10851
2023.1.30	59946	800	10648
2023.1.31	56946	1000	11431

## 5. Conclusion

In this study, we applied two methods, ARMA and LSTM neural networks, to the e-commerce logistics transportation prediction problem for in-depth research. Through empirical analysis and performance comparison, we found that the LSTM neural network exhibits higher accuracy and robustness in dealing with the e-commerce logistics shifting prediction task. Its learning of long-term dependencies enables it to better capture the underlying patterns in the time series, which improves the prediction performance. In contrast, ARMA models may be limited by data nonlinearity and long-term dependencies in some contexts. This finding emphasizes the importance of employing deep learning techniques when dealing with the e-commerce logistics movement forecasting problem. LSTM neural networks are not only better able to adapt to changing market conditions, but are also able to deal with logistics time-series data with complex, nonlinear structures. This is important for e-commerce companies to improve their logistics operation efficiency and service level in the face of increasingly complex market demands and dynamic supply chain environments. However, we also recognize that the issue of model selection depends on the specific application context and data characteristics. In some contexts, the ARMA model may still be an appropriate choice, especially when the data are relatively simple and linear relationships are obvious. Therefore, future research could further explore the applicability of the model in different contexts and how to combine multiple models to improve the robustness of predictions. Overall, this study provides a feasible approach for e-commerce logistics movement forecasting and provides useful insights for future research and practical applications.

## References

- [1] Slawomir Konrad Tadeja, Yupu Lu, Maciej Rydlewicz, Wojciech Rydlewicz, Tomasz Bubas, Per Ola Kristensson. *Exploring Gestural Input for Engineering Surveys of Real-Life Structures in Virtual Reality Using Photogrammetric 3D Models. Multim, Tools Appl.* 80(20): 31039-31058 (2021)
- [2] Ding, Y., Huang, L. *Case teaching research of quantitative analysis method in graduate logistics major. New Finan. econ.* 1, 114-115 (2019). (in Chinese).
- [3] Woschank, M., Pacher, C. *Teaching and learning methods in the context of industrial logistics engineering education. Procedia Manuf.* 51, 1709-1716 (2020)
- [4] Woschank, M., Pacher, C. *A holistic didactical approach for industrial logistics engineering education in the LOGILAB at the Montanuniversitaet Leoben. Procedia Manuf.* 51, 1814-1818 (2020)
- [5] Senna, E.T.P., dos Santos Senna, L.A., da Silva, R.M. *The challenge of teaching business logistics to international students. IFAC Proc.* 46(24), 463-470 (2013)
- [6] Zhang, Q., Yong, G., Zhang, M. *Logistics System Engineering - Theory, Method and Case Study (3rd Edition). Publishing House of Electronics Industry, Beijing (2021). Publishing House of Electronics Industry, Beijing (2021).*
- [7] Mgandu, F.A., Mkandawile, M., Rashid, M. *Trend analysis and forecasting of water level in Mtera dam using exponential smoothing. int. j. math. sci. Comput. (IJMSC)* 6(4), 26-34 (2020)
- [8] Sakpere, A.B., Oluwadebi, A.G., Ajilore, O.H., Malaka, L.E. *The impact of COVID-19 on the academic performance of students: a psychosocial study using association and regression model. Int. J. Educ. Manage. Eng. (IJEME)* 11(5), 32-45 (2021)
- [9] Padmaja, M., Haritha, D. *Software effort estimation using grey relational analysis. Int. J. Inf. Technol. Comput. Sci. (IJITCS)* 9(5), 52-60 (2017)
- [10] Chen, X., Tu, R., Li, M., Yang, X. *Prediction models of air outlet states of desiccant wheels using multiple regression and artificial neural network methods based on criterion numbers. Appl. Therm. Eng.* 204, 117940 (2021)
- [11] Lu, Q. H., Lau, M. F., Ng, S.P.H., Chen, T. Y. *Testing multiple linear regression systems with metamorphic testing. j. syst. softw.* 182, 111062 (2021)
- [12] Padmaja, M., Haritha, D.: *software effort estimation using grey relational analysis. Int. J. Intell. Syst. Appl.* 7(2), 27-33 (2015)
- [13] Yichung, H. *Constructing grey prediction models using grey relational analysis and neural networks for magnesium material demand forecasting. appl. Soft Comput.* 93, 106398 (2020)
- [14] Zeng, B., Liu, S. *Prediction model of random oscillation sequence based on amplitude compression. Syst. Eng. Theor. Pract.* 32(11), 2493-2497 (2012)