Logistics transportation model based on ARIMA and linear programming

Shudan Chen #, Xiaotian Zhu *, #

School of Science, Jimei University, Xiamen, China, 361021

* Corresponding Author Email: 18921188402@163.com

#These authors contributed equally.

Abstract. With the popularization of the Internet, the e-commerce logistics network has seen rapid growth. However, factors such as holidays, epidemics, and natural disasters like earthquakes can cause fluctuations in order volumes from e-commerce users, leading to changes in parcel quantities at each logistics site. To ensure the smooth operation of the logistics network, this paper conducts an in-depth analysis of the circulation relationships between different routes and establishes an ARIMA model and linear programming model. By utilizing existing data, this article study various scenarios that may occur at logistics sites. Based on these models, this article obtain line prediction results using the available data and construct a line adjustment model in a dynamic environment. This model effectively addresses sudden situations, enhancing the flexibility and responsiveness of the logistics system, thereby optimizing the operational efficiency and service quality of the entire logistics network. It provides robust support for the development of e-commerce logistics.

Keywords: ARIMA, Linear Programming, Logistics Management.

1. Introduction

Affected by the promotional activities of merchants, orders will be in a short period of time to achieve explosive growth, and in the event of unforeseen circumstances leading to the logistics site out of service and other situations, has been parceled out of the express delivery will face the risk of being stranded or even lost.

In this paper, this article will develop a diversion scheme after a missing line DC5, based on the available data, so that the total cumulative daily amount of parcels that fail to flow normally at a certain time is as small as possible, and give various scenarios for normal and abnormal flow situations. After completing the above work, the optimization model makes its network structure adjustable every day and the missing line is DC9.

In recent years, the logistics distribution problem of express stations has attracted significant attention from scholars both domestically and internationally. In China, Wang wei [1] proposed an optimization model based on genetic algorithms to reduce distribution time and cost. Li Hua et al. [2] developed a multi-objective optimization model that considered distribution efficiency, customer satisfaction, and environmental impact. Chen Jing [3] used big data technology to propose a dynamic scheduling method, significantly improving distribution efficiency.

Zhao Min et al. [4] addressed the inventory management problem by proposing an optimization strategy based on prediction models, reducing inventory backlog and out-of-stock issues. Liu Tao [5] explored supply chain collaboration, proposing a collaborative optimization model to enhance overall supply chain efficiency. Lei Zhang [6] applied machine learning to forecast customer demand, providing data support for logistics allocation.

Internationally, Smith and Johnson [7] proposed a mixed integer linear programming model for optimizing vehicle routes, which performed well in practical applications. Brown et al. [8] studied green logistics, proposing a low-carbon distribution scheme. Garcia and Martinez [9] used simulation technology to improve distribution efficiency by adjusting time windows. Lee and Kim [10] developed an IoT-based smart warehousing system to enhance warehouse automation.

Although previous researchers have carried out very in-depth studies in relation to logistics management, most of the studies on inventory management problems have focused on static forecasting models, whereas the changes in inventory demand in actual operation are highly uncertain. How to make real-time adjustments in the dynamically changing market environment still needs to be further explored. In this paper, a dynamic adjustment model is added to the static forecasting model in order to be applied to general occasions.

2. Model Introduction

2.1. Introduction to the ARIMA model

ARIMA model, also known as autoregressive moving average model, is a common statistical model applied in the field of time series, which firstly performs stemmed differencing to change the non-smooth time series into a smooth series, and linearly fits the smooth time series using the ARIMA model to get the predicted value of the future trend of the original series. In this paper, the ARIMA model will be used to establish the prediction model of line cargo volume. The specific modeling steps are as follows:

(1) Firstly, the time series graph was plotted and the data was tested for smoothness by calculating its Spearman correlation coefficient:

$$q_{XY}=\frac{\sum_{i=1}^{n}\left(R_i-\bar{R}\right)\left(S_i-\bar{S}\right)}{\sqrt{\sum_{i=1}^{n}\left(R_i-\bar{R}\right)^2}\sqrt{\sum_{i=1}^{n}\left(S_i-\bar{S}\right)^2}} \tag{1}$$

thereby demonstrating that

$$q_{XY}=1-\frac{6}{n(n^2-1)}\sum_{i=1}^{n}(R_i-S_i)^2 \tag{2}$$

where when the series is non-stationary, it can be d-differenced to form a stable time series data.

(2) Identification and ordering of the model: parameter identification is carried out using the Adaptive Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC), from which the optimal p and q values are selected. The value of d is determined according to the smoothness of the sequence after differencing;

$$AIC(p,q)=\ln\hat{\sigma}_{\varepsilon}{}^2+\frac{2(p+q)}{N} \tag{3}$$

$$BIC(p,q)=\ln\hat{\sigma}_{\varepsilon}{}^2+\frac{(p+q)\ln N}{N} \tag{4}$$

where $\hat{\sigma}_{\varepsilon}{}^2$ is the great likelihood estimate. N is the number of samples.

(3) Model Determination

A complete expression for the model ARIMA(p,d,q) can be derived.

$$\begin{cases} \Phi(B)\nabla^d x_t=\Theta(B)\varepsilon_t \\ E(\varepsilon_t)=0, Var(\varepsilon_t)=\sigma_{\varepsilon}{}^2, \\ E(\varepsilon_i\varepsilon_s)=0, s\neq t; E(x_s\varepsilon_t)=0, \forall s<t. \end{cases} \tag{5}$$

where

$$\nabla^d x_t=(1-B)^d x_t=\sum_{i=0}^{d}(-1)^i C_d^i x_{t-i}; \tag{6}$$

$$\Phi(B)=1-\phi_1 B-\phi_2 B^2-\ldots-\phi_p B^p \tag{7}$$

denotes the autoregressive polynomial of the smooth reversible ARIMA(p,q) model,

$$\Theta(B)=1-\theta_1 B-\theta_2 B^2-\ldots-\theta_p B^p \tag{8}$$

denotes the moving smoothing coefficient polynomial of the smooth reversible ARIMA(p,q) model.

(4) Parameter estimation: model parameter estimation is estimated using the great likelihood method, and model testing: test whether the residual series is a white noise series to ensure the validity of the model.

(5) Model prediction.

2.2. Basic theory of linear programming models

Linear programming (referred to as LP), is an important branch of operations research earlier, faster development, wide application, more mature methods, is to assist people in scientific management of a mathematical method, is the study of linear constraints under the linear objective function of the extreme value of the problem of mathematical theory and methods.

The three elements of linear programming model include: (1) decision variables; (2) objective function; (3) constraints

The mathematical model of a general linear programming problem can be expressed in the following forms:

$$\max(or\ \min) z=c_1 x_1+c_2 x_2+\ldots+c_n x_n$$

$$s.t.\begin{cases} a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n \leqslant (or\ =,\geqslant) b_1 \\ a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n \leqslant (or\ =,\geqslant) b_2 \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1+a_{m2}x_2+\ldots+a_{mn}x_n \leqslant (or\ =,\geqslant) b_m \\ x_1,x_2,\ldots x_n \geqslant 0 \end{cases} \tag{9}$$

A shorthand form of the above model is:

$$\max(or\ \min) z=\sum_{j=1}^{n} c_j x_j$$

$$s.t.\begin{cases} \sum_{j=1}^{n} a_{ij}x_j \leqslant (or\ =,\geqslant) b_i (i=1,2,\ldots,m) \\ x_j \geqslant 0 (j=1,2,\ldots,n) \end{cases} \tag{10}$$

where c_i, b_i, a_{ij} are the parameters of the model, where c_j is called the value coefficient, b_i is the right end term of the constraint, which indicates the availability of the ith resource, a_{ij} is the process or technology coefficient, which indicates the jth consumption of the ith resource per unit of product produced.

When solving a linear programming problem, the model is first standardized:

$$\max\ z=\sum_{j=1}^{n} c_j x_j$$

$$s.t.\begin{cases} \sum_{j=1}^{n} a_{ij}x_j = b_i (i=1,2,\ldots,m) \\ x_j \geqslant 0 (j=1,2,\ldots,n) \end{cases} \tag{11}$$

and then apply the graphical or simplex method to solve the problem.

3. Results

3.1. The results of ARIMA model

Based on the above theory, take some of the routes as an example, this article first drew a time series plot and tested the data for smoothness. The results of the time series plot and ADF test are shown in figure1:

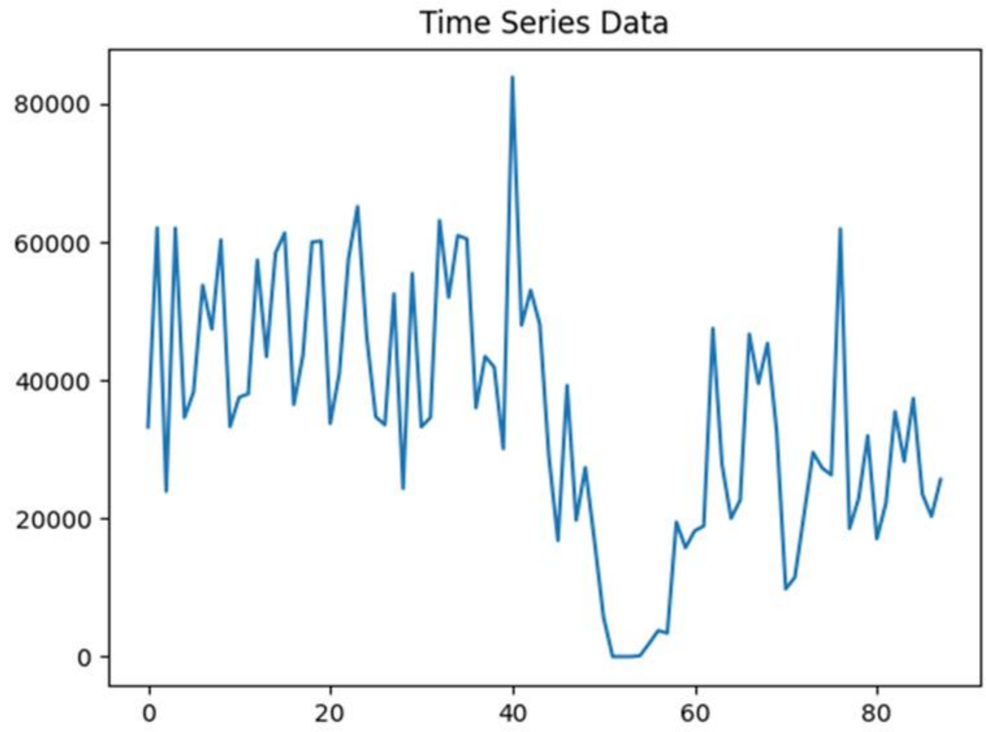

Figure 1. DC14-DC10 time series plot

The ADF test was then performed and the results are shown in Table 1.

Table 1. The ADF test

variable	Differential order	t	p	AIC	Critical value		
14-10					1%	5%	10%
	0	-4.547	0.000****	11101.038	-3.443	-2.867	-2.57
	1	-11	0.000****	11087.567	-3.443	-2.867	-2.57
	2	-10.273	0.000****	11124.83	-3.444	-2.867	-2.57

From the above table, this article can get $p < 0.05$, the series is smooth. Therefore, the rest of the sequence is also smooth.

The model is calculated after the test and based on the available data, this article give the prediction results as shown in Table 2.

Table 2. LSTM prediction table

time	DC14-DC10	DC20-DC35	DC25-DC62
2023.01.01	27854	135	17670
2023.01.02	26871	115	14250
2023.01.03	26739	144	10515
2023.01.04	27491	146	11888
2023.01.05	28510	157	5868
2023.01.06	29423	176	7771
2023..01.07	29497	192	13740
2023.01.08	28723	145	15229
2023.01.09	27779	173	11910
2023.01.10	26938	202	6625
2023.01.11	26298	215	6675
2023.01.12	25830	188	7808
2023.01.13	25171	208	11006
2023.01.14	24529	240	13742
2023.01.15	24672	249	8039
2023.01.16	28821	208	11838
2023.01.17	44344	264	11889
2023.01.18	77323	288	8244
2023.01.19	80780	269	7105
2023.01.20	46232	294	17489
2023.01.21	37039	247	8564
2023.01.22	36660	310	2850
2023.01.23	33295	255	12583
2023.01.24	30969	322	7098
2023.01.25	31947	287	4050
2023.01.26	37565	300	7549
2023.01.27	47166	427	10066
2023.01.28	54798	505	6138
2023.01.29	55339	587	10688
2023.01.30	69798	582	10749
2023.01.31	73843	675	9824

3.2. The analysis and results of linear programming model

Based on the above problems and background, this article constructed a multi-objective optimization model to make the cargo volume of each route under this logistics network as balanced as possible while ensuring that the amount of cargo that cannot flow normally is as little as possible.

First determine the decision variables

$h_{ijt} = 0$ indicates that no new routes were added on day t

$h_{ijt} = 1$ Indicates new routes added on day t

$d_{ijt} = 0$ Indicates that the original line is not closed on day t

$d_{ijt} = 1$ Indicates that the original line is closed on day t

V denotes all nodes in this logistics network

The objective function is then established:

After the closure of DC9, in order to reduce its impact on the entire logistics network, when developing the daily adjustment programme, this article need to make the number of routes that change during the adjustment process the least, and the routes that change can be expressed as follows

$$\min\left[\sum_{i\in V}\sum_{j\in V}\sum_{t}^{T}h_{ijt}+\sum_{i\in V}\sum_{j\in V}\sum_{t}^{T}d_{ijt}\right] \tag{12}$$

After the addition or closure of the original logistics sites, this article have to ensure a balanced cargo volume for each route in the new logistics network, set up ω_{ij} is the sum of the shipments transported from route i to route j. this article measure the balance of shipments across routes by the standard deviation

$$\sqrt{\frac{\sum_{i\in V}\sum_{j\in V}(w_{ij}-\bar{w})^{2}}{n}} \tag{13}$$

If it is not possible to make all the relevant routes flow normally, then the minimum number of parcels that cannot flow should be ensured, let the number of parcels that cannot flow normally be m_{ij} , then the objective is to

$$\min\sum_{i\in V}\sum_{j\in V}m_{ij}(t) \tag{14}$$

Then, based on the actual problem, set the constraints, and the conditions should satisfy the following points:

1) The total cargo volume of a route is the inflow minus the outflow, and the total cargo volume should be equal to the volume of new logistics sites plus the volume of unclosed logistics sites, and the subconstraints can try to ensure the normal flow of goods.

2) The total value of all inflows cannot exceed the maximum capacity of all previously existing sites.

3) If a node is closed, the amount of goods flowing out of that node is 0.

4) Ensure that all lines are balanced.

5) The volume of cargo per route goes to a positive integer.

Finally, this article construct the planning model shown below

$$\min\sum_{i\in V}\sum_{j\in V}m_{ij}(t)$$

$$s.t.\begin{cases} x_{ijt}-x_{jit}=(1-h_{ijt})\times\sum_{i\in V}x_{ij(t-1)}+(1-d_{ijt})\times\sum_{i\in V}x_{ij(t-1)} \\ x_{ijt}\leqslant h_{ijt}\times c_{ij} \\ \sum_{i\in V}d_{ijt}\times x_{ijt}=0 \\ \max\left(x_{ij}^{(2)}-\frac{\sum_{i=1}^{n}\sum_{j=1}^{n}x_{ij}^{(2)}}{n}\right)-\min\left(x_{ij}^{(2)}-\frac{\sum_{i=1}^{n}\sum_{j=1}^{n}x_{ij}^{(2)}}{n}\right)\leqslant\partial \end{cases} \tag{15}$$

The solution is then solved using the simulated annealing algorithm and the results are obtained as shown in Table 3.

Table 3. Excess/remaining stock and load factor

route	Excess/remaining stock	load factor
69→5	0	100%
69→8	0	100%
69→14	0	100%
69→62	0	100%
69→10	542.7419	94.86%
60→5	1130.29	79.23%
23→32	254.2581	57.90%
64→8	776.0645	54.24%
60→8	10207.9	53.17%
60→10	3552.774	52.20%
27→10	7829.839	51.47%
8→14	15699.1	51.38%
52→8	1979.806	50.48%
26→64	5839.065	49.19%
4→23	1806.226	48.67%
70→4	198.4194	48.46%
38→64	7565.29	47.00%
30→5	2112.935	46.99%
27→8	30669.13	46.25%
47→14	34615.94	45.92%

4. Conclusion

This paper addresses the issue of explosive order growth due to merchant promotions and the risk of parcel stagnation or loss caused by unforeseen logistics site interruptions. A diversion plan is proposed for scenarios like the failure of line DC5, aiming to minimize the total daily amount of parcels that fail to flow normally. An optimization model with an adjustable network structure is developed to allow daily adjustments, such as in the case of DC9's closure.

The paper reviews existing research on logistics distribution, highlighting the limitations of static forecasting models and introducing a dynamic adjustment model to handle uncertain inventory demand. The ARIMA model is used to predict cargo volumes, and a multi-objective linear programming model is constructed to minimize non-flowing parcels and balance cargo distribution across routes. Decision variables, objective functions, and constraints are defined to ensure the model's effectiveness and practicality.

Empirical analysis shows that the ARIMA model accurately predicts cargo volumes, and the linear programming model effectively minimizes non-flowing parcels while maintaining a balanced distribution. These findings provide valuable insights for optimizing logistics management and enhancing the efficiency and stability of the logistics system.

References

[1] Wang wei. An optimization model based on genetic algorithms for vehicle scheduling in express stations [J]. Logistics Technology, 2019, 38 (12): 123-128.

[2] Li Hua, Zhang Qiang, Liu Yang. Multi-objective optimization model for express station logistics distribution [J]. Systems Engineering - Theory & Practice, 2020, 40 (5): 1234-1241.

[3] Chen Jing. Dynamic scheduling method for express stations using big data technology [J]. Computer Engineering and Applications, 2021, 57 (15): 156-162.

[4] Zhao Min, Wang Li, Li Na. Inventory optimization strategy for express stations based on prediction models [J]. Management Science, 2022, 25 (3): 345-352.

[5] Liu Tao. Collaborative optimization model for supply chain collaboration in express stations [J]. Chinese Journal of Management Science, 2023, 31 (6): 678-685.

[6] Lei Zhang. Machine learning approach for customer demand forecasting in express stations [J]. Management Review, 2024, 36 (1): 102-109.

[7] Smith, J., & Johnson, K. A mixed integer linear programming model for vehicle routing in parcel delivery stations [J]. Transportation Research Part E: Logistics and Transportation Review, 2019, 123: 123-138.

[8] Brown, R., Davis, S., & Green, T. Green logistics strategies for parcel delivery stations [J]. Journal of Cleaner Production, 2020, 245: 118897.

[9] Garcia, F., & Martinez, L. Simulation-based optimization of parcel delivery station operations [J]. Simulation Modelling Practice and Theory, 2021, 104: 102115.

[10] Lee, H., & Kim, J. IoT-based intelligent warehousing systems for parcel delivery stations [J]. IEEE Transactions on Industrial Informatics, 2022, 18 (3): 1892-1901.