

基于 ARIMA 模型的物流网络中心货物量预测研究

张天宇^{1,*}, 王淼馨¹, 郭利伟¹, 于小淼¹, 梁 瑞²

(1. 辽宁工程技术大学 安全科学与工程学院 125105, * 1159192693@qq.com;

2. 辽宁工程技术大学 软件学院 125105)

摘 要: 随着电商物流行业的兴起, 物流网络中心的货物量预测对其发展至关重要。为了研究电商物流网络中心的未来货物量预测问题, 建立 AIRMA (p, d, q) 货物量预测模型。以收集到的某物流网络中心的 57 个分拣中心, 在 2023 年 8 月 1 日—2023 年 11 月 30 日的货量, 以及这 57 个分拣中心 2023 年 11 月 1 日—2023 年 11 月 30 日 23 点的货量, 建立 ARIMA 货物量预测模型, 确定模型参数 p, d, q 的分别取值后, 对序列数据进行白噪声检验, 得到了未来 30 天每天及每小时的货物量, 并应用 ARIMA (p, d, q) 模型得到了其他 56 个分拣中心的预测结果。结果表明: 基于上述建立的 ARIMA (p, d, q) 模型, 发现对于分拣节点的货物量的预测较为适用。

关键词: 物流中心货物量; 自回归积分滑动平均模型; 白噪声检验; 预测

引言

自从电商风靡全国, 物流行业的发展就开始不断跃升。2022 年, 中国社会物流总额达到 347.6 万亿元, 物流业总收入 12.7 万亿元, 已经连续七年成为全球最大规模的物流市场。公路货运是最主要运输方式, 占总运量比重的 75% 左右。如何大规模的货物源完整地由商户配送到个人成了物流公司需要解决的难题, 分拣中心作为一个物流网络的中转站, 完美地解决了物流混乱的问题。数学算法在此方面具有重要作用, 主要应用于建立模型预测未来物流网络中心货物量。统计模型 (如 ARIMA、VAR) 或机器学习算法 (如神经网络、支持向量机) 可利用历史数据预测未来物流中心货物量[1-2]。数学算法还可应用于大规模数据挖掘和特征提取, 从多个方面提取相关特征, 建立预测模型。本研究旨在以某物流网络中心的分拣中心的历史货运量数据为研究对象, 建立 ARIMA 的物流网络中心货物量预测模型。

1 ARIMA 模型介绍

ARIMA(p, d, q) 模型是一种时间序列分析和预测的方法该模型结合了自回归 (AR)、差分 (I) 和移动平均 (MA) 这三个组成部分, ARIMA (p, d, q) 模型可以根据历史物流站点的货运量数据来预测未来货运量, 模型中的参数 q 表示自回归系数; p 表示偏自回归系数; d 表示差分系数[3-4]。建模基本流程如下: 首先对时间序列平稳性进行检验, 利用差分将非平稳序列转化成平稳序列, 然后确定差分参数 (d), 接着通过观察偏自相关 (PACE) 以及自相关图 (ACF) 来确定自回归参数 (p) 和移动平均参数 (q)。具体如下式:

$$\begin{aligned}\phi(B)\nabla^d x_t &= \varphi(B)\varepsilon_t \\ E(\varepsilon_t) &= 0 \\ V_{ar}(\varepsilon_t) &= \sigma_\varepsilon^2 \\ E(\varepsilon_t \cdot \varepsilon_s) &= 0, s \neq t \\ E(x_t \cdot \varepsilon_t) &= 0, \forall s < t\end{aligned}\tag{1}$$

式中: $\phi(\square)$ 、 $\varphi(\square)$ 分别表示 p 阶自回归系数多项式、q 阶移动平均系数多项式, $\nabla^d = (1-B)^d$ 表示差分计算, B

表示延迟算子，且 $B^n x_t = x_{t-n}$ ， ε_t 表示随机误差序列， $E(\cdot)$ 、 $V_{ar}(\cdot)$ 分布表示 $\cdot$ 的期望与方差。

2 数据收集与分析

通过收集到的某物流网络中心的 57 个分拣中心，在 2023 年 8 月 1 日—2023 年 11 月 30 日的货量，以及这 57 个分拣中心 2023 年 11 月 1 日—2023 年 11 月 30 日 23 点的货量进行分析研究。以 SC1 分拣节点为例，通过收集到的 SC 分拣点的历史货物量，建立 ARIMA 货物量预测模型，确定模型参数 p 、 d 、 q 的分别取值后，对序列数据进行白噪声检验，得到了其分拣点未来 30 天每天及每小时的货物量。

表 1 SC1 分拣中心历史货物量

分拣中心	时间	货物量
SC1	2023/8/1	39517
SC1	2023/8/2	40008
SC1	2023/8/3	38477
SC1	2023/8/4	35773
SC1	2023/8/5	31876
SC1	2023/8/6	31319
SC1	2023/8/7	29584
SC1		
SC1	2023/11/24	44719
SC1	2023/11/25	44496
SC1	2023/11/26	42879
SC1	2023/11/27	41085
SC1	2023/11/28	44200
SC1	2023/11/29	45217
SC1	2023/11/30	45707

3 ARIMA 货物量预测模型的建立——以 SC1 分拣节点为例

ARIMA (p, d, q) 模型是一种时间序列分析和预测的方法，该模型结合了自回归 (AR)、差分 (I) 和移动平均 (MA) 这三个组成部分，可以通过对序列本身的概率分析来捕捉时间序列中的标准时间结构[5]。
“ d ”为差值的阶数，“ p ”为自回归系数，“ q ”为移动平均项。具体表达式如下所示：

$$(1 - \sum_{i=1}^p \alpha_i L^i)(1 - L)^{d_1} / t = \alpha_0 + (1 + \sum_{i=1}^q \beta_i L^i) \varepsilon_t$$

(2)

STEP 1: 平稳性测试—— d 的选择

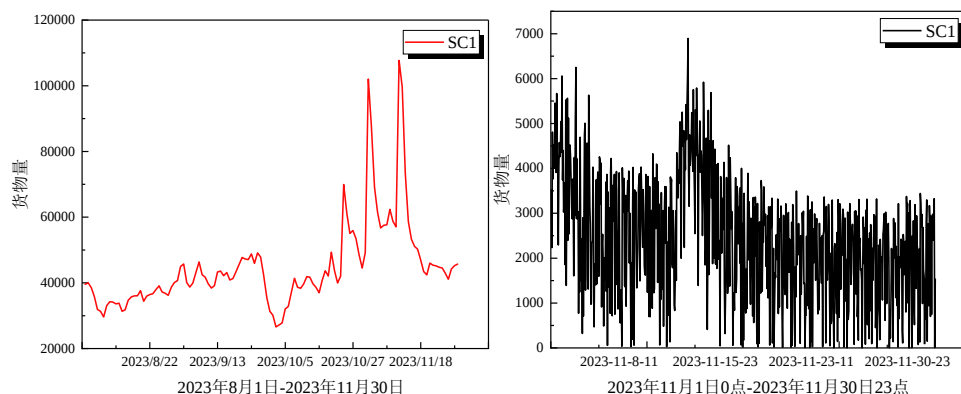


图1 时间序列图

建立 ARIMA 模型的先决条件之一是进行平稳性测试，由上图 1 初步发现该序列数据不符合平稳要求，因此需要对序列进行平稳性测试，并且需要满足以下三个条件来确定是否具有协方差平稳性：

$$\begin{cases} \phi(B)\nabla^d x_t = \theta(B)\varepsilon_t \\ E(\varepsilon_t) = 0, \text{Var}(\varepsilon_t) = \sigma_\varepsilon^2, E(\varepsilon_t, \varepsilon_s) = 0 \\ E(\varepsilon_t, \varepsilon_s) = 0, \forall s < t \end{cases} \quad (3)$$

基于上图并结合上式（3）对 SC1 历史货物量数据序列进行一阶差分，并进行 ADF 检验，测试结果如下表 2 所示：

表 2 ADF 检验结果

Variable	t	1%	5%	10%	p	测试结果
Hum	-2.1471	-3.1724	-3.2491	-3.1132	0.1379	不平稳
□ hum	-11.8127	-1.9573	-1.6787	-1.4097	0.0022	平稳

从上表中可以看出，对于 SC1 历史数据序列的一阶差值的检验统计量在 1%的水平上时显著的，所以序列是平稳的，确定 d 为 1。

STEP2: AIC 准则——p, d 的确定

使用 AIC 准则，不仅考虑了模型对 SC1 分拣节点货物量的预测精度，还考虑了预测模型中的参数，非常适用于 ARIMA 的检验，因此，为了确定 p, q 的值，我们选择采用 AIC 准则。AIC 准则是一种用来比较不同 ARIMA 模型拟合优度的方法，一般认为，AIC 准则越小，模型的效果越好，AIC 准则函数表达如下式（4）所示：

$$AIC = \ln(\hat{\sigma}^2) + \frac{2}{n}(p+q+1) \quad (4)$$

式中： $\hat{\sigma}^2$ 为模型误差的方差，n 为样本量，p, d 为模型的阶数，p, q 的确定需要通过 AIC 判断函数和相关系数法来确定，其 AIC 准则结果如下图 2 所示：

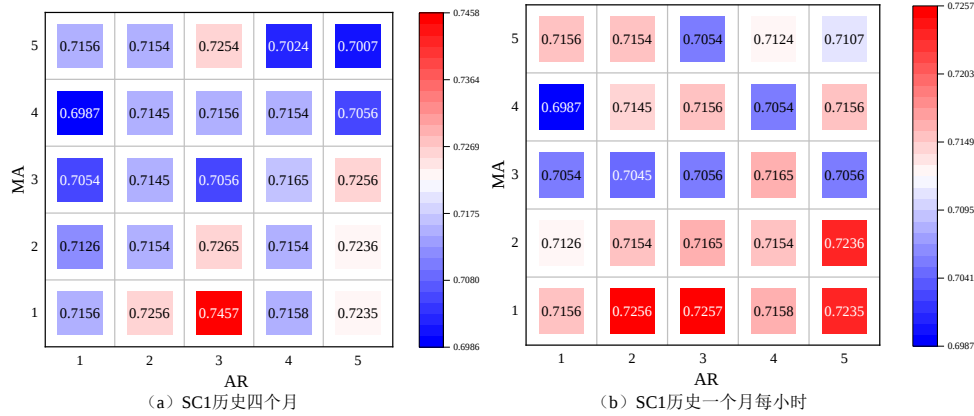


图 2 AIC 准则结果

为了避免过拟合及模型复杂度过高，需要选择合适的 ARIMA 模型的 p, q 参数，经过 AIC 准则和显著性测试的方法，最终确定当 $p=2, q=4$ 时，AIC 的值相对较小，因此我们选择 ARIMA (2, 1, 4)，作为 SC1 分拣节点货物量的预测模型。

STEP3: 白噪声检验

然后，对残差序列进行白噪声测试，其目的是避免数据过于随机。其本质是确定数据中是否存在自相关。我们通过控制一个自相关图来检验序列数据是否有效，测试结果如下图 3 所示：

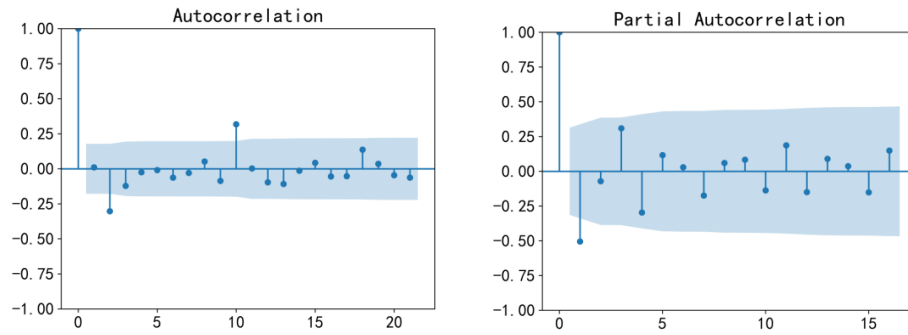


图 3 自相关检验图

从上图 3 观察发现，以上的自相关和偏自相关图显示出现的明显的自相关。综上所述，发现许多数据特征与白噪声不一致，因此，该序列数据通过了白噪声测试，并可以建立 ARIMA (2, 1, 4) 模型，如下式 (5) 所示：

$$(1 - 0.3231L - 0.2174L^2)(1 - L)(1/t - \mu) = (1 + 0.2127L + 0.5332L^2 + 0.4874L^2 + 0.1787L^4)\varepsilon_t \quad (5)$$

式中， y_t 表示 SC1 分拣节点的货物量预测值， μ 为序列均值， ε_t 为白噪声误差， L 为滞后运算符。

4 ARIMA 货物量预测模型的求解——以 SC1 分拣节点为例

通过将 SC1 分拣节点的历史四个月每天的货物量数据以及历史一个月每天每小时的货物量数据代入到建立的 ARIMA (2, 1, 4) 模型中，分别预测出未来一个月每天的货物量和未来一个月每天每小时的货物量，得到的预测结果如下表 3 所示：

表 3 SC1 分拣中心未来三十天每天的货物量

分拣中心	时间	货物量预测	分拣中心	货物量预测	货物量预测
SC1	2023/12/1	45459.24814	SC1	2023/12/16	44980.13871
SC1	2023/12/2	45358.63518	SC1	2023/12/17	44976.79386
SC1	2023/12/3	45278.44851	SC1	2023/12/18	44974.12807
SC1	2023/12/4	45214.54123	SC1	2023/12/19	44972.00349
SC1	2023/12/5	45163.60832	SC1	2023/12/20	44970.31024
SC1	2023/12/6	45123.01573	SC1	2023/12/21	44968.96075
SC1	2023/12/7	45090.66419	SC1	2023/12/22	44967.88523
SC1	2023/12/8	45064.88061	SC1	2023/12/23	44967.02806
SC1	2023/12/9	45044.33158	SC1	2023/12/24	44966.34491
SC1	2023/12/10	45027.95438	SC1	2023/12/25	44965.80046
SC1	2023/12/11	45014.90206	SC1	2023/12/26	44965.36654
SC1	2023/12/12	45004.4996	SC1	2023/12/27	44965.02071
SC1	2023/12/13	44996.20904	SC1	2023/12/28	44964.74509
SC1	2023/12/14	44989.60161	SC1	2023/12/29	44964.52543
SC1	2023/12/15	44984.33561	SC1	2023/12/30	44964.35036

4 预测结果的分析

通过建立的 ARIMA（2，1，4）模型，得到 SC 未来一个月每天的货物量和未来一个月每天每小时的货物量，得到结果如下图 4 所示：

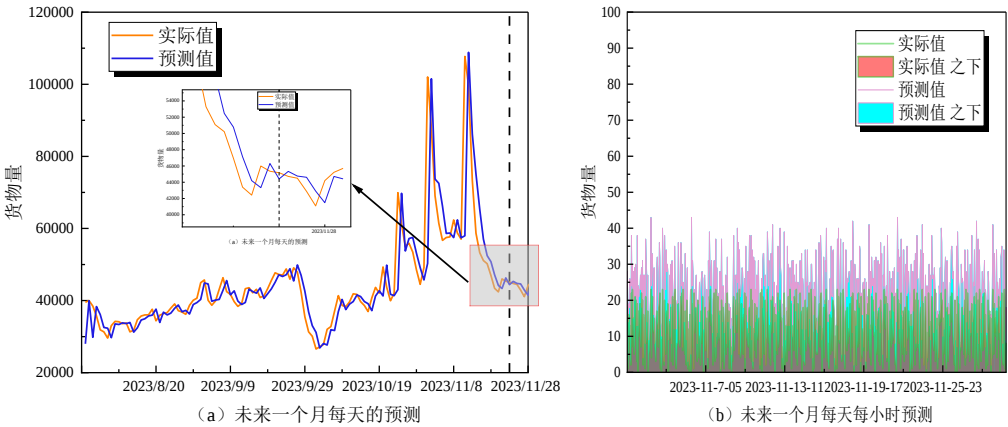


图 4 SCI 分拣节点预测结果

图 4（a）为 ARIMA 模型预测未来一个月每天的预测结果，“黄线”表示实际值，“蓝色”表示预测值，观察历史数据的四个月的预测结果，发现预测曲线与实际曲线几乎贴近变化趋势保持一致，说明预测效果较好，适用于该分拣节点的货物量预测，图 4（b）为 ARIMA（2，1，4）模型预测未来一个月每天每小时的预测结果，选用面积填补发现，预测趋势几乎保持一致，也取得了较好的效果，因此，可以认为 ARIMA 模型适用于分拣节点的货物量预测情况。

5 结论

ARIMA 模型应用方便简单，只需要考虑内生变量，从而避免需要借助其他外生变量，大大缩短计算时间。基于本文建立的 ARIMA（p，d，q）物流网络中心货物量预测模型，发现对于分拣节点的货物量的预测

较为适用。

参考文献

- [1] 陈燕峰,王贺,李岩等.组合两步分解和 ARIMA-LSTM 的短期风速预测研究[J].太阳能学报,2024,45(02):164-171.
- [2] 郭龙,郭文文.基于 SVR 和随机森林模型的动力煤高位发热量预测研究[J].能源工程,2024,44(01):35-42.
- [3] Siarni-Namini S, Tavakoli N, Namin A S. A comparison of ARIMA and LSTM in forecasting time series[C]//2018 17th IEEE international conference on machine learning and applications (ICMLA). IEEE, 2018: 1394-1401.
- [4] Dey B, Roy B, Datta S, et al. Forecasting ethanol demand in India to meet future blending targets: A comparison of ARIMA and various regression models[J]. Energy Reports, 2023, 9: 411-418.
- [5] 黄玲,任苏灵.基于 ACO 优化的 ARIMA-ES-RF 布伦特原油价格预测研究[J].科技和产业,2024,24(05):111-119.