

Research and Application of Logistics Enterprise Freight Demand Forecasting Based on Seasonal Time Series

Wei Luo, Xiaoping Fang

School of Traffic and Transportation Engineering, Central South University, Changsha Hunan
Email: fangxp@csu.edu.cn

Received: Feb. 19th, 2016; accepted: Mar. 5th, 2016; published: Mar. 9th, 2016

Copyright © 2016 by authors and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Freight volume forecasting is an important part of the reasonable logistics planning. In this paper, a historical freight volume of A company is regarded as research data, mainly basing on the trend extrapolation method and the seasonal decomposition method, establishing a single model and a combination model. Discussion on the time series of short-term forecasting is a direct curve or “remove” seasonal before fitting. For the trend model, there are two choices: One is polynomial model; another is Logistic model. While assuming the influence of monthly seasonal factors is same. We could build four predictive models ultimately, compare the goodness of fitting, then chose the best one. The result demonstrates that the combination model is better than the single model.

Keywords

Trend Extrapolation Method, Seasonal Decomposition Method, Logistic Model, Freight Volume Forecasting

基于季节性时间序列的物流企业货运需求预测 研究与应用

罗 维, 方晓平

中南大学交通运输工程学院, 湖南 长沙

Email: fangxp@csu.edu.cn

收稿日期: 2016年2月19日; 录用日期: 2016年3月5日; 发布日期: 2016年3月9日

摘要

货运量预测是制定合理物流计划的重要内容。本文以A物流公司的历史货运量为例, 主要基于趋势外推法和季节分解法来进行研究, 建立趋势预测单项模型和季节分解-趋势外推的组合模型。探讨对时间序列的短期趋势预测是直接曲线拟合还是“剔除”季节性后再拟合, 对于趋势模型的选择, 对比多项式拟合和Logistic曲线拟合的效果, 同时假设季节因素对每个月的影响是不变的, 最终构建四个预测模型, 通过比较四种模型的拟合优度来评判出合适的预测模型。预测结果证明组合模型的拟合效果要比单项模型的要好。

关键词

趋势外推法, 季节分解法, Logistic曲线, 货运量预测

1. 引言

货运量是指在一定时期内运输企业实际产生的货物的数量。对物流企业来说, 一定时期内的货运量的大小能反映企业的运营的好坏。对货运量进行预测是物流企业经营决策的重要前提, 通过合理预测, 物流企业还可以有效利用各种资源来降低成本, 在满足顾客需求的前提下获得最大利润。所以, 货运量预测对物流企业是十分重要的。

据统计, 到目前为止, 国内外应用于货运量预测的方法多达 300 多种。主要有时间序列预测法、灰色预测法、回归分析法、神经网络法和组合预测法。从相关性研究方面来看, Russo F 在考虑了供需关系、经济政策等多种影响因素的情况下, 结合宏观经济以及运输模式建立了意大利综合货运需求模型[1]。王红, 宋凤杰运用指数平滑法和灰色预测法对港口吞吐量进行建模预测, 并提出了拟合精度更高、更适合短期预测的修正的灰色模型[2]。Bashi: Z.和 EL. Hawary M.E.提出运用一种动态神经网络模型预测交通流量, 并且得到了较好的预测结果[3]。在时间序列预测方面, Box 和 Jenkins 为预测交通流量构建了 ARIMA 模型, 第二年 Cullinane 构造了预测短期生成货运量的 Box_Jenkins [4]。白世贞, 刘莉等人以哈尔滨市物流需求数据为例, 构造了基于季节指数和灰色模型的物流需求预测方法[5]。关宏志等提出了货物运输需求量的预测在理论和方法上和客流需求量的预测有所不同的依据, 并结合我国物流业的发展现状, 对货物运输量预测方法研究的发展方向进行了展望[6]。陈治亚等运用 BP 神经网络理论建立了用于物流需求规模预测的模型, 为物流需求规模的预测提供了新的思路[7]。陈实运用移动平均、指数平滑、回归和组合模型等对山东的货运量进行了预测, 并提出了一个可靠的组合预测模型[8]。

研究货运量预测的文献较多的是从国家、区域和交通通道等宏观层面, 针对物流企业进行货运量预测的较少[9] [10]。物流企业货运量需求预测是其运营决策的基础, 研究物流企业货运量随季节性波动的时序需求预测, 对这一类企业的经营决策有重要意义。本文以某物流公司的历史货运量数据为基础, 研究基于趋势外推法和季节分解法探讨物流企业货运量需求预测问题。

2. 理论基础

本文主要关注时间序列预测法中的趋势外推模型和季节性波动分解模型的组合模型。时间序列是指将某一事物发展的数值按其发生的时间的先后顺序排列形成的数列。时间序列预测法即根据历史数据对

未来进行预测。任何时间序列从要素上可以被划分为四个部分：趋势(T)、周期(C)、季节性(S)和随机元素(R)。这四种成分的随意组合，可以将季节性时间序列分为加法模型和乘法模型[11]：

$$X_{t_m} = T + C + S + R \quad (1)$$

$$X_{t_m} = T \times C \times S \times R \quad (2)$$

本文不考虑经济长周期，而只考虑年为周期的正常经济活动，故季节性也体现在该周期内。公式(1)和(2)里面的 C 与 S 就合二为一了。

2.1. 趋势外推模型

在对社会统计资料进行分析时可以发现，经济现象总是随着时间的变化呈现出某种上升或是下降的趋势，并且没有明显的季节波动，那么我们可以通过一条曲线来描述这种变化。再对这条曲线建立相应的数学模型，即趋势预测模型。在本文中主要讨论多项式模型和 Logistic 模型。

(1) 多项式模型：

$$y = a_0 + a_1x + a_2x^2 + a_3x^3 + \varepsilon \quad (3)$$

(2) Logistic 模型：

$$y = \frac{K}{1 + \alpha e^{-\beta t}}, \text{ 其中 } K \text{ 为饱和值} \quad (4)$$

2.2. 季节分解法

季节分解预测法又称为季节因素预测法或是季节指数预测法。系统内部的周期性运动、变化规律和系统外部的环境的季节性作用共同导致了季节性的波动。季节分解预测法能够既考虑到季节指数的趋势变化，还能充分发挥已知数据的信息价值。季节分解预测法的思路分为三个步骤：第一步，“剔除”季节性并预测长期趋势值；第二步，计算季节指数；第三步，通过方法让季节性在趋势上反应出来，达到预测的效果。

首先“剔除”趋势，在选择究竟是加法还是乘法模型时，高孝伟认为加法模型比乘法模型更加合理[12]。故在本文中采用加法模型。

$$X'_{t_m} = X_{t_m} - T = C + S + R \quad (5)$$

考虑到货运量与经济之间的密切关系，而且依据数据的周期特性，确定周期 $s = 12$ ，在这里我们采取季节性的 12 个月周期而忽略周期性成分。故有：

$$X'_{t_m} = X_{t_m} - T = S + R \quad (6)$$

得到季节性与随机元素的组合 $S + R$ ，然后用 $S + R$ 除以 T 即为每个月的季节因子 s_c 。

$$s_c = \frac{X_{t_m} - T}{T} = \frac{S + R}{T}, \quad c = 1, 2, \dots, 36 \quad (7)$$

假设不同年度的相同月份的季节影响是一样的，即 $s_i = s_{i+12} = s_{i+24}$, $i = 1, 2, \dots, 12$ ，则我们可以通过求三组季节因素的平均值来求得每个月的季节因子[13]：

$$\bar{s}_i = \frac{s_i + s_{i+12} + s_{i+24}}{3} = \frac{1}{3} \left(\frac{X_{t_m} - T}{T} \Big|_{m=i} + \frac{X_{t_m} - T}{T} \Big|_{m=i+12} + \frac{X_{t_m} - T}{T} \Big|_{m=i+24} \right) \quad (8)$$

即当 $i = 1$ 时， \bar{s}_1 为 1 月份的季节因子：

$$\bar{s}_1 = \frac{1}{3}(s_1 + s_{13} + s_{25}) = \frac{1}{3} \left(\frac{X_{t_m}}{T} \Big|_{m=1} + \frac{X_{t_m}}{T} \Big|_{m=13} + \frac{X_{t_m}}{T} \Big|_{m=25} \right) \quad (9)$$

按照以上方法, 我们求出了趋势模型和季节因子, 故我们建立预测模型:

$$y = T \times (1 + \bar{s}_t), \quad t = 1, 2, \dots, 12 \quad (10)$$

3. 模型应用

3.1. 数据描述

A 物流公司主要经营国内的公路零担货物运输业务。目前运输服务的线路基本上覆盖了全国各个省、市、自治区, 服务网络遍及全国, 已开设直营网点 5000 多家, 自有的营运车辆近 9000 台, 分布在全国的转运中心的总面积超过 100 万平方米。A 公司在 2011 年的货物吞吐量达到近 500 万吨。本文数据是 A 公司自 2009 年 1 月至 2011 年 12 月共 36 个月的货运量记录。时间序列记为:

$$Y^{(1)} = \{y_i^{(1)}\} = \{y_1^{(1)}, y_2^{(1)}, \dots, y_{36}^{(1)}\}, \quad i = 1, 2, \dots, 36。$$

样本的时间序列图如图 1 所示。从图可以看出货运量随着时间的推移有明显的不断上升的趋势, 而且每年的 1~2 月和 6~7 月的货运量都是全年最低的, 4~5 月和 10~11 月的货运量是一年中最多的, 故序列存在着明显的季节波动现象。

3.2. 模型建立

3.2.1. 趋势模型建立

本节内容分为两部分, 第一部分是建立趋势模型, 第二部分是计算季节因子。其中, 趋势模型的建立运用趋势外推单项模型和外推模型与季节分解模型的组合模型。单项模型基于多项式和 Logistic 曲线的直接趋势外推模型, 组合模型先运用季节分解模型得到“剔除”了季节性的趋势序列, 再进行趋势拟合。

(1) 直接趋势外推建模

A) 用多项式模型分别进行拟合, 拟合效果如图 2 所示。对拟合模型进行参数检验, 其中二次、三次模型的拟合度相等, 而且拟合优度要高于线性模型。如表 1 所示。



Figure 1. Sequence chart

图 1. 时序图

进一步发现二次模型和三次模型的系数 t 值和相应 p 值均大于临界值, 说明当使用二次和三次模型时系数是不显著的, 即统计没有意义。故选用线性方程来建立趋势模型。

模型一: $y_1 = 902948.933 + 62328.423t$

B) Logistic 曲线拟合要输入上限值, 故令 K 取不同值。输出结果如表 2 所示。

当 $K = 3700000$ 时 R^2 是最大的, 建立 Logistic 曲线模型为。

模型二: $y_2 = \frac{3700000}{1 + 3.15e^{t \ln 0.926}}$, 当 $K = 3700000$

(2) 季节分解模型与趋势外推模型组合建模

运用 Eviews8.0 的 X-12 法进行季节分解。SER01_SA 表示季节调整后的序列, SER01_SF 表示季节因子序列, SER01_TC 表示趋势与循环波动叠加序列, SER01_IR 表示不规则的波动。如图 3 所示。

A) 对 SER01_TC 进行多项式趋势拟合, 结果如表 3。

其中二次和三次模型的系数 t 值和相应 p 值均大于临界值, 故选用线性模型来建立趋势模型。

模型三: $y_3 = 1018216.912 + 54055.968t$

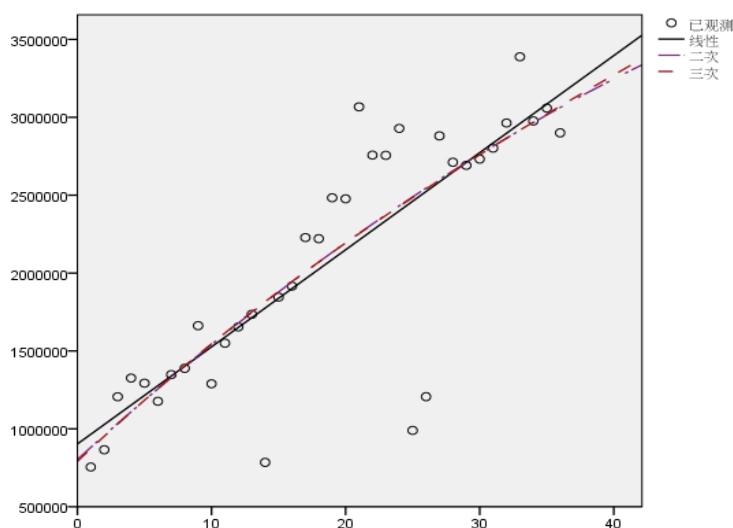


Figure 2. Polynomial model fitting
图 2. 多项式模型拟合图

Table 1. Polynomial model aggregation and parameter test
表 1. 多项式模型汇总和参数检验

方程	模型汇总					参数估计值			
	R 方	F	df1	df2	Sig.	常数	b1	b2	b3
线性	0.676	70.780	1	34	0.000	902,948.933	62,328.423		
二次	0.678	34.775	2	33	0.000	803,570.406	78,019.769	-424.090	
三次	0.678	22.483	3	32	0.000	791,283.313	81,751.564	-672.817	4.482

Table 2. Model goodness of different K
表 2. 不同 K 值的模型拟合度

K	3,500,000	3,600,000	3,700,000	3,750,000	3,800,000	3,900,000
R^2	0.66	0.669	0.671	0.67	0.67	0.669

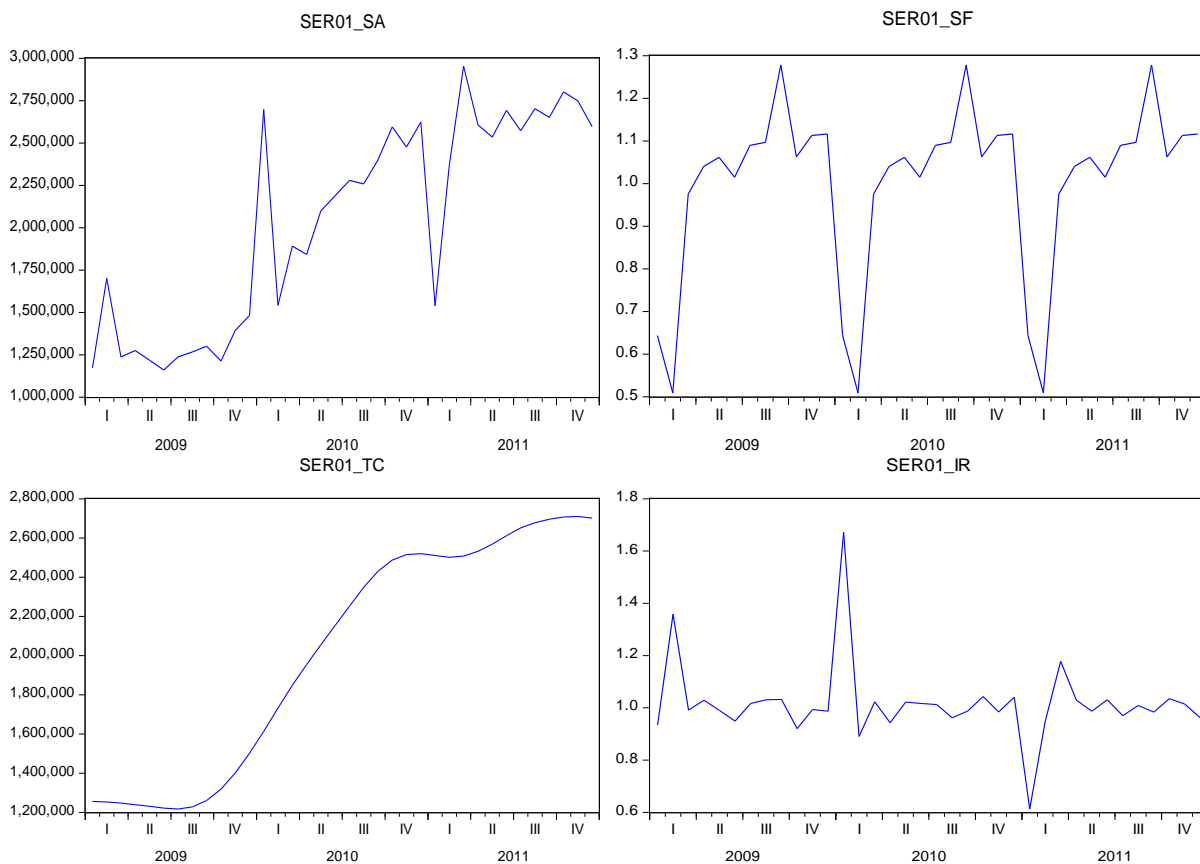


Figure 3. Sequence after seasonal decomposition of freight volume

图 3. 货运量序列季节分解后得到的各个序列

Table 3. Polynomial model aggregation and parameter test of SER01_TC

表 3. SER01_TC 序列多项式模型汇总和参数估计

方程	模型汇总					参数估计值			
	R 方	F	df1	df2	Sig.	常数	b1	b2	b3
线性	0.923	408.095	1	34	0.000	1,018,216.812	54,055.968		
二次	0.935	236.900	2	33	0.000	863,953.195	78,413.382	-658.308	
三次	0.976	434.120	3	32	0.000	1,234,551.298	-34,143.441	6843.677	-135.171

B) 对 SER01_TC 序列进行 Logistic 模型拟合, 令 K 取如下值, 输出不同 K 值时的 R^2 。见表 4。

可以发现当 $K = 2790000, 2800000, 2810000$ 时 R^2 是最大的, 且 R^2 都为 0.960。在这里综合实际情况选择建立 Logistic 曲线模型。

$$\text{模型四: } y_4 = \frac{280000}{1 + 2.56e^{t \ln 0.866}}, K = 2800000$$

3.2.2. 季节因子

根据求季节因子的公式(8), 分别求得各个模型的季节因子如表 5 所示。

根据所求得趋势模型和季节因子, 由公式(10)得到预测模型分别为:

$$\text{模型五: } y_5 = 902948.933 + 62328.423t(1 + \bar{S}_c)$$

Table 4. Model goodness of different K of SER01_TC
表 4. SER01_TC 序列不同 K 值模型的拟合度

K	2730000	2750000	2770000	2780000	2790000	2800000	2810000
R 方	0.938	0.953	0.958	0.959	0.960	0.960	0.960

Table 5. Seasonal factors of each model
表 5. 各模型的季节因子

	一月	二月	三月	四月	五月	六月	七月	八月	九月	十月	十一月	十二月
模型一	-0.268	-0.412	0.075	0.061	0.064	0.001	0.062	0.055	0.222	0.014	0.049	0.048
模型二	-0.266	-0.41	0.073	0.061	0.062	-0.003	0.057	0.05	0.216	0.007	0.044	0.043
模型三	-0.29	-0.43	0.064	0.054	0.065	0.009	0.073	0.072	0.246	0.039	0.077	0.079
模型四	-0.258	-0.393	0.086	0.069	0.065	0.001	0.061	0.059	0.229	0.026	0.067	0.069

Table 6. R^2 of forecasting model
表 6. 预测模型的 R^2

	直接趋势外推法		组合模型		文献[14]	
	多项式模型 (模型五)	Logistic 模型 (模型六)	多项式模型 (模型七)	Logistic 模型 (模型八)	GM(1,1)周期外延 组合模型	SARIMA 模型
R^2	0.892	0.899	0.900	0.904	0.960	0.445

$$\text{模型六: } y_6 = \frac{3700000}{1 + 3.15e^{t \ln 0.926}} (1 + \bar{S}_c)$$

$$\text{模型七: } y_7 = 1018216.912 + 54055.968t (1 + \bar{S}_c)$$

$$\text{模型八: } y_8 = \frac{280000}{1 + 2.56e^{t \ln 0.866}} (1 + \bar{S}_c)$$

3.3 模型拟合优度

本文采用 R^2 检验预测模型的拟合优度。假设 Y 为观测值, \hat{Y} 为模型拟合值, \bar{Y} 为样本均值, 则

$$R^2 = 1 - \frac{\sum_{i=1}^n (Y - \hat{Y})^2}{\sum_{i=1}^n (Y - \bar{Y})^2} \quad (11)$$

对不同预测模型运用拟合模型利用式进行拟合优度对比分析, 计算结果如下: [14].

由表 6 知, 组合模型的拟合优度要高于直接趋势外推模型和 SARIMA 模型, 但不如 GM(1,1)周期外延组合模型。且 Logistic 趋势模型的拟合度要比多项式模型高。

4. 结论

本文基于趋势外推法和季节分解法建立了货运量单项预测模型和组合预测模型, 并对四种预测模型拟合优度进行了横向和纵向比较分析。从纵向看, Logistic 曲线对序列趋势的拟合效果要好于多项式拟合, 说明本案例中货运量时间序列的发展变化与 Logistic 曲线变化更相似; 从横向看, 组合模型的拟合优度要高于单项预测模型, 表明在对序列趋势进行拟合之前, “剔除” 序列的季节性是有必要的。

基金项目

湖南省科技计划项目, 项目编号 2015ZK3001。

参考文献 (References)

- [1] Russo, F. and Conigliaro, G. (1997) Integrated Macroeconomic and Transport Models for Freight Demand. *Transportation Systems*, **3**, 16-18.
- [2] 王红, 宋风杰. 预测技术在港口吞吐量预测中的应用[J]. 实用物流技术, 1995(3): 8-11.
- [3] Bashir, Z. and El-Hawary, M.E. (2000) Short-term Load Forecasting by Using Wavelet Neural Networks. *Canadian Conference on Electrical and Computer Engineering*, **1**, 163-166.
- [4] Kavassanos, M.G. and Nomikos, N.K. (2000) Constant vs. Time-Varying Hedge Ratios and Hedging Efficiency in the BIFFEX Market. *Transportation Research Part E: Logistics and Transportation Review*, **36**, 229-248. [http://dx.doi.org/10.1016/S1366-5545\(99\)00029-0](http://dx.doi.org/10.1016/S1366-5545(99)00029-0)
- [5] 白世贞, 刘莉, 杨艳玲. 基于灰色模型与季节指数的物流需求预测研究——以哈尔滨市为例[J]. 物流工程与管理, 2010(6): 8-10.
- [6] 关宏志, 陈艳艳. 地区间货物运输量预测方法谱系[J]. 土木工程学报, 2003(7): 47-52.
- [7] 陈治亚, 周艾飞, 谭钦之, 方晓平. 基于改进的 BP 人工神经网络的物流需求规模预测. 铁道科学与工程学报, 2008(6): 62-68.
- [8] 陈实. 货运量预测方法及应用研究[D]. 武汉: 武汉理工大学, 2008.
- [9] 包训艳. 物流企业需求预测[J]. 中国电子商务, 2013(23): 258-259.
- [10] 王平. 青岛港物流公司物流量预测分析[J]. 青岛远洋船员学院学报, 2008, 29(2): 53-57.
- [11] Barrow, M. (2009) *Statistics for Economics, Accounting and Business Studies*. 5th Edition, Pearson Education, London.
- [12] 高孝伟, 许涛, 郑林昌, 刘静. 对季节指数计算方法的思考[J]. 统计与决策, 2006(9): 155.
- [13] Fang, X., Ansell, J. and Chen, W. (2013) Modeling of a Small Transportation Company's Start-Up with Limited Data during Economic Recession. *Discrete Dynamics in Nature & Society*, **2013**, 1-10.
- [14] 石宏. 第三方物流企业的运输需求分析模型研究与应用[D]. 长沙: 中南大学, 2013.