

基于指数平滑 - 回归模型及灰色 - 回归模型的长三角物流需求预测分析

李洋洲, 李程*

上海工程技术大学航空运输学院, 上海

收稿日期: 2023年7月5日; 录用日期: 2023年8月13日; 发布日期: 2023年8月21日

摘要

随着全国经济运行整体好转, 长三角作为我国经济水平最高的地区之一, 科学预测长三角区域物流需求有助于把握物流业未来的发展变化, 为政府部门制定区域经济发展政策以增强地区经济活力提供重要参考依据。文章从宏观经济、对外开放程度、人口发展水平和人民生活水平四个维度选取6个典型指标作为多元回归因子, 并通过灰色关联分析法分析各指标与长三角物流需求之间关联的程度, 采用指数平滑 - 回归模型与灰色 - 回归模型分别对长三角区域的物流需求进行预测。其预测精度分析表明: 指数平滑 - 回归模型在MAPE、MSE、MAE三个精度指标上均优于灰色 - 回归模型, 指数平滑 - 回归预测模型效果优于灰色 - 回归预测模型, 更适合对长三角物流需求进行中短期预测。预测结果显示未来五年长三角区域物流需求平均增速约为2.2%。结合预测结果, 文章提出了相关管理建议以更好地促进长三角区域经济与物流业高质量发展。

关键词

区域物流, 灰色预测, 指数平滑预测, 多元线性回归

Logistics Demand Forecasting in Yangtze River Delta Based on Exponential Smoothing-Regression Model and Grey-Regression Model

Yangzhou Li, Cheng Li*

College of Air Transportation, Shanghai University of Engineering Science, Shanghai

Received: Jul. 5th, 2023; accepted: Aug. 13th, 2023; published: Aug. 21st, 2023

*通讯作者。

文章引用: 李洋洲, 李程. 基于指数平滑-回归模型及灰色-回归模型的长三角物流需求预测分析[J]. 运筹与模糊学, 2023, 13(4): 4025-4040. DOI: 10.12677/orf.2023.134403

Abstract

As a result of the overall improvement in national economic functioning as one of the areas with the highest economic level in the Yangtze River Delta, scientific forecasting of regional demand for logistics in the Yangtze River Delta is helpful for capturing the future development and change of the logistics industry and serves as an essential reference point for government departments in formulating regional economic development policies to enhance regional economic vitality. Six typical indicators are chosen here as multiple regressors from four dimensions of macroeconomics, the openness to the outside world, the level of population development and people's standard of living, and the degree of correlation between each indicator and the logistic demand of the Yangtze River Delta are analyzed by grey correlation analysis method. Both the exponential smoothing regression model and the gray regression model are used to predict the logistics demand of the Yangtze River Delta. The prediction accuracy analysis shows that the exponential smooth-regression model is better than the gray-regression model in MAPE, MSE, and MAE, and the effect of the exponential smooth-regression model is better than the gray-regression model, which is more appropriate for short- and medium-term forecasting of Yangtze River Delta logistics demand. Based on the forecast results, the average rate of growth of logistics demand in the Yangtze River Delta over the next 5 years is approximately 2.2%. Taken together with the results of the forecasts, the paper advances some management suggestions to promote the high-quality development of the Yangtze River Delta's regional economy and logistics industry.

Keywords

Regional Logistics, Gray Forecast, Exponential Smoothing Forecast, Multiple Linear Regression

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

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



Open Access

1. 引言

长江三角洲地区(以下简称长三角区域)作为我国经济最活跃的区域,也是我国综合实力最雄厚的地区,其经济发展水平高于其他地区,同时由于其人均 GDP 的快速增长,催生出巨大的物流需求。从 2005 年到 2021 年,长三角区域的货运量增长了约 3 倍,其年均增长率约 7.74%。文章从区域物流需求端出发结合科学的分析手段,研究影响长三角区域物流需求的主要经济与社会因素,从而构建长三角区域物流需求的影响因素指标体系,进而构建精确与合理的预测模型,对长三角区域物流需求进行预测,可以避免区域物流基础设施建设与物流需求不匹配而产生资源浪费,有助于科学制定区域物流发展规划促进长三角物流业降本增效与行业健康发展。精确与合理的物流需求预测对于保证区域物流服务的供给与需求满足动态平衡,使长三角区域物流资源配置更加完善,实现区域经济与物流协调发展具有重要现实意义。

在对区域物流需求预测的研究方面:过秀成[1]等指出区域经济与社 会因素对区域物流需求有很大影响,区域物流需求的研究应着重分析区域社会经济和物流供应条件的影响。丁红等[2]主要运用了灰色 GM(1,1)模型对江西省物流发展情况进行了预测,并阐述了物流需求预测对于物流规划的作用。赵文等[3]构建了以物流消费及其基础能力与信息能力为影响因素指标的预测指标体系;主要运用 BP 神经网络模型对广东省物流能力展开预测与分析。黄勇福等[4]运用灰色 GM(1,N)模型预测了广东省“十四五”期间

的货运量, 并分析了广东省物流需求量的发展规律。在预测方法方面: 雷斌等[5]通过改进粒子群算法的灰色神经网络预测模型对铁路货运量展开预测结果表明该模型预测精度优于传统灰色神经网络; 周文杰等[6]基于 2001~2012 年西北五省统计数据, 选择货运量为指标, 采用一元线性回归模型对 2002~2020 年间西北五省货运量和周转量规模进行了预测。龙忠芬[7]通过多元线性回归模型分析和检验, 对影响广东省货运量的因素进行分析, 得出地区生产总值等对广东省货运量有着显著影响; 张明[8]构建多个时间序列预测模型并且选择平均相对误差最小的四次多项式回归时间序列预测模型作为预测模型。通过最佳时间序列预测模型预测未来五年的铁路货运量。赵美珍等[9]指出单一的预测方法, 包括: 移动平均、灰色模型和随机时间序列等已经无法满足人们对于预测精度的要求。姜金德等[10]通过趋势外推法(TE)对各相关经济指标进行预测, 将预测数据代入 PCR 模型得到了 2020~2024 年江苏省货运量数据。李思聪等[11]运用灰色 - 回归组合模型对我国农产品冷链物流市场需求进行预测, 其研究结果表明: 相比单一灰色预测和多元线性回归, 灰色 - 回归预测组合模型预测精度较好。

综上所述, 现有研究尚存在以下不足: 一是从研究对象来看长三角作为我国经济发展最活跃的区域之一, 但是对长三角的物流需求预测的研究较少, 随着长三角一体化战略的深入推进, 为更好地服务于长三角经济与物流高质量协调发展, 应该重视对于长三角物流需求量的预测。二是对于区域物流需求影响因素方面, 普遍从供给端出发结合水路、铁路、航空等不同的运输方式作为区域物流需求量的影响因素来开展物流预测。三是从方法来看, 多数对于区域物流需求的预测多采用单一模型进行预测, 存在一定局限性。如普遍采用的灰色 GM(1,1)模型、多元回归预测模型、指数平滑预测模型在预测时往往只把自变量作为被预测主体, 忽略物流需求量相关影响因素; 其中 LSTM 等模型非线性拟性好, 但需要大量的数据训练才能得到较为精确的预测结果[12]。本研究主要采用指数平滑 - 回归预测模型对长三角物流需求进行预测, 同时引入灰色 - 回归模型进行精度对比分析, 结合影响因素指标数据, 在提高预测精度与考虑影响因素的同时, 选择预测效果较好的指数平滑 - 回归模型对未来五年长三角区域物流需求进行预测。

2. 研究方法

2.1. 灰色关联分析

为了量化评价各指标对长三角物流需求量的影响程度, 并识别最关键因素, 文章通过构建灰色关联度模型计算长三角物流需求量与影响长三角物流需求量的指标的关联度, 关联度越大, 表示该指标对长三角物流需求量的影响程度越高。

灰色关联系数计算步骤如下:

1) 首先采用初值化处理方法, 对变量进行无量纲化处理, 处理方法为: 用各数列的每一个数除以各列第一个。

2) 关联系数计算公式为:

$$\xi_i(k) = \frac{\min \min |x_0(k) - x_i(k)| + \rho \cdot \max_i \max_k |x_0(k) - x_i(k)|}{|x_0(k) - x_i(k)| + \rho \cdot \max_i \max_k |x_0(k) - x_i(k)|} \quad (1)$$

(k = 1, \dots, n; i = 1, \dots, m)

其中 $\xi_i(k)$ 为关联系数; $x_0(k)$ 表示母序列; $x_i(k)$ 表示子数列; ρ 表示分辨系数, 取值范围为[0,1], 常取 0.5; $k = 1, 2, \dots, 11$ 。

3) 关联度 θ 计算公式为:

$$\theta_i = \frac{1}{n} \sum_{k=1}^n \xi_i(k) \quad (k = 1, 2, \dots, n) \quad (2)$$

2.2. 多元回归

多元线性回归分析主要用于研究 X (自变量)对 Y (因变量)的影响关系, 其中通过 R^2 值判断模型拟合情况, 通过 VIF 值判断模型是否存在共线性问题, 最后得出多元回归模型的方程。

多元回归的模型可表示为:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_m X_m + \varepsilon \quad (3)$$

其中: β_0 是常量, 当 X_1, X_2, \dots, X_m 均为 0 时, 这个常量表示 Y 的值。 $\beta_1, \beta_2, \dots, \beta_m$ 为 Y 对应于每个自变量的偏回归系, ε 为残差。

2.3. 指数平滑

指数平滑由布朗提出常用于中短期经济发展趋势预测, 其基本思想是不同的数据会被赋予不同的权数, 新数据会比旧数据被赋予更大的权数, 再通过逐层平滑计算消除掉随机因素造成的影响进而得出预测的基本变化趋势。 当时间序列呈平滑趋势时可采用一次平滑法进行预测; 当序列具有一定线性趋势往往需要对第一次平滑数据再次平滑提高预测精度这也被称为二次平滑法; 而三次平滑法适用于曲线数据序列。

三类指数平滑公式为:

$$S_k^{(1)} = ay_k + (1-a)S_{k-1}^{(1)}; \quad (4)$$

$$S_k^{(2)} = aS_k^{(1)} + (1-a)S_{k-1}^{(2)}; \quad (5)$$

$$S_k^{(3)} = aS_k^{(2)} + (1-a)S_{k-1}^{(3)}. \quad (6)$$

三类指数平滑式子中: S_k 为第 k 期的预测值; y_k 为第 k 期的实际值; S_{k-1} 为第 $k-1$ 期的预测值; a 为平滑常数, 其取值范围为 $[0,1]$ 。

2.4. 灰色预测

灰色预测 GM(1,1)模型的预测过程是将原始无规律数据通过累加而形成的一组有规律的累加数据序列进行建模, 再将通过模型生成的数据通过累减还原得到预测值。

假设原始数列为 $X^{(0)}$:

$$X^{(0)} = (X^{(0)}(1), X^{(0)}(2), \dots, X^{(0)}(n)) \quad (7)$$

$X^{(0)}$ 进行累加, 生成累加序列为 $X^{(1)}$:

$$X^{(1)} = (X^{(1)}(1), X^{(1)}(2), \dots, X^{(1)}(n)) \quad (8)$$

$$X^{(1)}(i) = \sum_{k=1}^i X^{(0)}(k), \quad i = 1, 2, \dots, n \quad (9)$$

构造序列级比检验:

$$\sigma(k) = \frac{X^{(0)}(k)}{X^{(0)}(k-1)} \in \left(e^{-\frac{2}{n+1}}, e^{\frac{2}{n+1}} \right) \quad (10)$$

如 $X^{(1)}$ 序列通过级比检验, 可构建一阶线性微分方程:

$$\frac{dX^{(1)}}{dt} + aX^{(1)} = u \quad (11)$$

令

$$Z^{(1)}(k) = \mu X^{(1)}(k) + (1 - \mu) X^{(1)}(k-1), k = 2, 3, \dots, n \quad (12)$$

其中, $Z^{(1)}(k)$ 称为一阶线性微分方程的背景值, $\mu \in [0, 1]$, μ 为权重系数。假定 μ 取值为 0.5, 则有:

$$Z^{(1)}(k) = \frac{X^{(1)}(k) + X^{(1)}(k-1)}{2} \quad (13)$$

在一阶线性微分方程公式(11)中, a 为发展系数; u 为灰色作用量。对累加生成数据做均值生成 B 与常数项向量 Y , 通过计算参数 a, u , 就能求出 $X^{(1)}(t)$, 进而求出 $X^{(0)}$ 预测值。

$$Y = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(n) \end{bmatrix}, B = \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \vdots & \vdots \\ -z^{(1)}(n) & 1 \end{bmatrix} \quad (14)$$

由最小二乘法求解灰参数 $\hat{a} = [a, u]^T$ 得到:

$$\hat{a} = (B^T B)^{-1} B^T Y_n \quad (15)$$

将灰参数 \hat{a} 代入公式(11)并对方程求解, 可得:

$$\hat{x}^{(1)}(t+1) = \left(x^{(1)}(1) - \frac{u}{a} \right) e^{-at} + \frac{u}{a} \quad (16)$$

将上述结果累减还原得到预测值:

$$\hat{x}^{(0)} = \{ \hat{x}^{(0)}(1), \hat{x}^{(0)}(2), \dots, \hat{x}^{(0)}(n) \} \quad (17)$$

$$\hat{x}^{(0)} \{ (n+1), \dots, \hat{x}^{(0)}(n+m) \} \quad (18)$$

2.5. 预测精度评价体系

研究从平均绝对百分比误差(MAPE)、均方根误差(MSE)和平均绝对误差(MAE)三方面评价长三角物流需求量预测精度的好坏, 其中 y_i 为第 i 年的物流需求量实际值; \hat{y}_i 为第 i 年的物流需求量的预测值; M 为预测的样本长度。

$$MAPE = \frac{1}{M} \sum_{i=1}^M \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100\% \quad (19)$$

$$MSE = \frac{1}{M} \sum_{i=1}^M (y_i - \hat{y}_i)^2 \quad (20)$$

$$MAE = \frac{1}{m} \sum_{i=1}^M |y_i - \hat{y}_i| \quad (21)$$

3. 指标选取及指标数据

3.1. 因变量与自变量指标选取

长三角地区物流需求量受很多因素的影响, 主要包括宏观经济影响因素, 对外开放程度, 人民生活水平, 人口发展水平这四个维度。通过查找大量国内外学者的文献, 遵循重要与简洁性原则、经济与可

得性原则、目标与适用性原则文章选取与长三角物流需求量相关性较高的指标如下:

选取货运量作为长三角的物流需求量的量化指标。货运量是最主要、最客观、最直接的反映物流需求量的指标。并通过对大量文献进行研究分析,研究者通常都将货运量作为物流需求量的衡量指标。所以选取货运量作为长三角物流需求量的量化指标[13]。选取生产总值与第三产业生产总值反映宏观经济水平,生产总值主要反映区域经济发展规模,第三产业生产总值侧重反映产业结构,经济的快速发展会带来更多的物流需求,以满足企业的生产需求和市场的供需平衡。而物流业属于第三产业,物流的需求规模、层次、结构与区域产业结构密切相关[13]。选取进出口额反映对外开放程度,2022年长三角进出口总值达15.07万亿元比2021年增长6.9%,占全国进出口总值比重为35.8%。而进出口贸易会在促进商品流通进而带动区域物流需求。选取常住人口与旅客发送量反映区域人口发展水平,一方面,常住人口数量增加,意味着消费市场和劳动力市场的扩大,从而催生了更多的经济活动和物流需求。同时旅客发送量较大的地区,区域的旅游产业发展态势较好,旅游产业带来的物流需求也在逐渐增加,这也带动了物流企业 and 供应链的发展。选取人均可支配收入反映区域人民生活水平,人均可支配收入是反映人民生活水平的重要指标之一,人均可支配收入对物流需求量产生了重要的影响。人均可支配收入的增加也会导致人们的购买能力和消费水平也随之提高,从而导致物流需求量的增加。

研究选取我国2005~2021年的相关指标数据作为长三角物流需求影响因素指标。预测指标为长三角区域物流需求总量 Y_1 ,影响因素 X_i 为自变量($i=1,2,\dots,6$)指标如表1所示。

Table 1. Indicators of factors influencing logistics demand

表 1. 物流需求影响因素指标

目标层	一级指标	影响因素指标	选取说明
物流需求	宏观经济	生产总值 X1/亿元	反映经济发展规模
		第三产业生产总值 X2/亿元	反映产业结构
	对外开放程度	进出口总额 X3/亿美元	反映对外贸易规模
	人口发展水平	常住人口 X4/万人	反映区域人口规模
		旅客发送量 X5/万人	反映区域人口流动性
人民生活水平	居民人均可支配收入 X6/元	反映居民消费水平	

3.2. 指标数据收集

通过查找2006~2022年《上海市统计年鉴》《安徽省统计年鉴》《江苏省统计年鉴》《浙江省统计年鉴》,分别收集到2005~2021年长三角三省一市的进出口总额、常住人口、居民人均可支配收入等七个指标数据,最后通过计算求得2006~2022年长三角区域的生产总值、第三生产总值、进出口总额、常住人口、旅客发送量、居民人均可支配收入指标数据如表2所示。

Table 2. Statistics on indicators of logistics demand and influencing factors in the Yangtze River Delta, 2005~2021

表 2. 2005~2021年长三角物流需求及影响因素指标统计数据

年份	Y1/万吨	X1/亿元	X2/亿元	X3/亿美元	X4/万人	X5/万人	X6/元
2005	374,970	46,023	18,996	11,655,237	20,589	388,231	44,388
2006	411,985	53,643	22,290	15,144,661	20,802	424,815	49,537
2007	458,611	65,449	27,879	19,284,937	21,060	474,421	56,012
2008	577,493	76,285	32,827	23,161,600	21,251	565,861	62,963
2009	585,868	83,913	37,486	20,343,174	21,427	576,022	68,857

Continued

2010	668,250	99,949	44,676	27,789,335	21,576	627,697	76,657
2011	76,0045	116,989	52,690	34,081,331	21,917	678,613	88,360
2012	829,197	127,732	59,118	35,182,652	22,191	730,955	97,758
2013	830,209	140,472	67,169	38,152,168	22,455	550,245	109,035
2014	928,182	152,643	74,637	40,442,477	22,635	444,885	118,921
2015	849,354	165,482	84,031	220,512,431	22,769	372,936	126,430
2016	883,907	180,800	95,565	226,468,231	22,954	342,972	136,441
2017	976,763	200,874	107,624	261,427,433	23,117	325,529	147,799
2018	1,029,933	221,233	120,480	291,424,814	23,271	308,556	160,861
2019	1,047,297	235,952	131,574	315,265,957	23,417	307,887	174,765
2020	1,101,976	244,522	138,057	346,253,662	23,539	190,170	179,922
2021	1,191,028	276,054	153,635	424,902,397	23,648	156,773	198,622

3.3. 基于灰色关联分析的影响因素分析

由于各指标数据项在单位上不一致, 选用初值法对 2005~2021 年长三角物流需求总量及影响因素指标统计数据进行无量纲化处理, 以保证预测体系数量级与影响因素的数量级一致。再针对 6 个评价项(X1, X2, X3, X4, X5, X6), 以及 2005~2021 年 17 项数据进行灰色关联度分析, 并且以 Y1 作母序列, 研究 6 个子序列与 Y1 的灰色关联关系, 使用灰色关联度分析时, 分辨系数取 0.50, 结合关联系数计算公式可得 2005~2021 年生产总值 X1、第三产业生产总值 X2、进出口总额 X3、常住人口 X4、旅客发送量 X5、居民人均可支配收入 X6 与物流需求量 Y1 的灰色关联系数如图 1 所示。

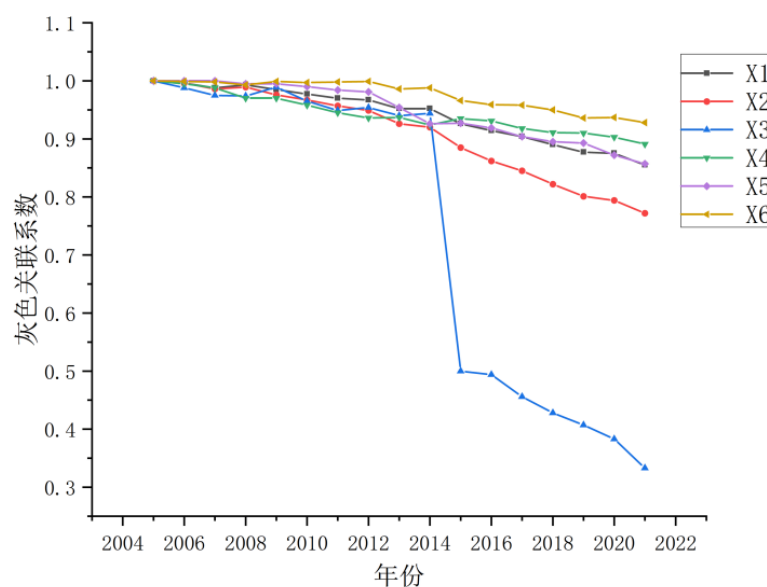


Figure 1. Gray correlation coefficients of influencing factors

图 1. 影响因素灰色关联系数

根据上图的灰色关联系数进一步求出 2005~2021 年生产总值 X1、第三产业生产总值 X2、进出口总额 X3、常住人口 X4、旅客发送量 X5、居民人均可支配收入 X6 与物流需求量 Y1 的灰色关联度并根据

灰色关联度值的大小对其进行排序如表 3, 以关联度大于 0.6 作为较强关联关系的接受原则, 文章选取的 6 个影响因素指标与长三角物流需求均存在较强的关联关系。

Table 3. Gray correlation of impact factor indicators
表 3. 影响因素指标灰色关联度

评价项	关联度	排名
居民人均可支配收入 X6/元	0.976	1
旅客发送量 X5/万人	0.947	2
常住人口 X4/万人	0.942	3
生产总值 X1/亿元	0.942	4
第三产业生产总值 X2/亿元	0.909	5
进出口总额 X3/亿美元	0.746	6

4. 长三角物流需求量预测

4.1. 多元回归模型

通过上述灰色关联度分析(GRA)得出所选取的六个影响因素指标 X1, X2, X3, X4, X5, X6 与物流需求量均有较强的关联关系。因此将六个影响因素指标作为回归因子纳入多元回归方程, 建立长三角物流需求多元回归初始模型, 相关求解系数如表 4 所示。

Table 4. Table of coefficients for multiple linear regression analysis
表 4. 多元线性回归分析系数表

影响因素指标	B	t	p	VIF
X1	9.081	2.661	0.024*	2082.038
X2	-5.947	-1.321	0.216	1326.308
X3	-0.001	-2.786	0.019*	36.184
X4	24.113	0.446	0.665	103.399
X5	0.037	0.333	0.746	11.327
X6	-1.801	-0.341	0.74	2339.014
R ²			0.995	
调整 R ²			0.992	
F		F(6,10) = 353.678, p = 0.000		
D-W 值		2.957		

* p < 0.05, ** p < 0.01.

从上表可知, 将 X1, X2, X3, X4, X5, X6 作为自变量, 而将 Y1 作为因变量进行线性回归分析, 从上表可以看出, 模型公式为:

$$Y1 = -357160.972 + 9.081 \times X1 - 5.947 \times X2 - 0.001 \times X3 + 24.113 \times X4 + 0.037 \times X5 - 1.801 \times X6 \quad (22)$$

模型 R² 值为 0.995, 意味着 X1, X2, X3, X4, X5, X6 可以解释 Y1 的 99.5% 变化原因, 表明该模型对于因变量有较好的解释度。但模型自变量的显著性检验值部分大于 0.05, 且模型中 VIF 值均大于 10, 意味着该多元线性回归模型存在着共线性问题。

为消除多元回归中存在的共线性问题选取逐步回归的处理方法对自变量进行筛选, 通过逐步回归模

型逐步 stepwise 法自动识别出有显著性的自变量, 将 X1, X2, X3, X4, X5, X6 作为自变量, 而将 Y1 作为因变量进行逐步回归分析, 最终结果如表 5 所示, 最终余下 X4 一共 1 项在模型中, X4 的回归系数值为 241.929 ($t = 22.438, p = 0.000 < 0.01$), 意味着 X4 会对 Y1 产生显著的正向影响关系。 $R^2 = 0.971$ 意味着 X4 可以解释 Y1 的 97.1% 变化原因。调整 $R^2 = 0.969$, 说明模型能较好拟合, $VIF = 1$ 小于 10 说明逐步回归后成功消除了共线性问题, 而且模型通过 F 检验 ($F = 503.466, p = 0.000 < 0.05$), 说明模型有效。

Table 5. Table of coefficients of results of stepwise regression analysis

表 5. 逐步回归分析结果系数表

模型系数	B	t	p	VIF
β_0	-4593710.345	-19.111	0.000**	-
X4	241.929	22.438	0.000**	1
R^2		0.971		
调整 R^2		0.969		
F		F(1,15) = 503.466, p = 0.000		
D-W 值		1.29		

* $p < 0.05$, ** $p < 0.01$.

由此建立一元线性回归模型为:

$$Y1 = -4593710.345 + 241.929 \times X4 \quad (23)$$

4.2. 自变量预测

4.2.1. 自变量灰预测

通过 GRA 以及多元回归模型建立, 对常住人口 X4 建立 GM(1,1) 灰色模型进行预测。如表 6 所示原始序列级比值均处于标准范围区间 [0.895, 1.118] 内, 意味原始值序列通过级比检验适合构建 GM(1,1) 模型 [14]。

Table 6. Gray GM(1,1) model level ratios

表 6. 灰色 GM(1,1) 模型级比值

序号	原始值	级比值 λ	序号	原始值	级比值 λ
2005	20589.4	-	2013	22454.87	0.988
2006	20801.57	0.99	2014	22635.15	0.992
2007	21059.61	0.988	2015	22768.7	0.994
2008	21250.53	0.991	2016	22953.84	0.992
2009	21427.05	0.992	2017	23116.78	0.993
2010	21575.5	0.993	2018	23270.58	0.993
2011	21916.52	0.984	2019	23417.43	0.994
2012	22191.31	0.988	2020	23538.62	0.995
	→		2021	23647.83	0.995

构建模型后通过最小二乘法分别求得发展系数 a、灰色作用量 u; 进一步计算的后验比 C 值和小误差概率 p 值如表 7 所示。

Table 7. GM(1,1) modeling results

表 7. GM(1,1)模型构建结果

发展系数 a	灰色作用量 u	后验差比 C 值	小误差概率 p 值
-0.0087	20671.8907	0.0140	1.000

从上表可知, 模后验差比 C 值 $0.014 < 0.35$, 小误差概率 p 值为 $1.000 < 1.0$, 结合表 8, 模型精度判定指标表判断该 GM(1,1)灰色模型精度很好。

Table 8. GM(1,1) model accuracy determination index table

表 8. GM(1,1)模型精度判定指标表

模型精度	后验差比 C	小误差概率 p 值
好	$C < 0.35$	$p > 0.95$
合格	$C < 0.5$	$p > 0.8$
基本合格	$C < 0.65$	$p > 0.7$
不合格	$C \geq 0.65$	$p \leq 0.7$

将灰参数 a 代入公式(11)并求解, 可得:

$$\hat{x}^{(1)}(t+1) = (x^{(1)}(1) + 2376079.4)e^{0.0087t} - 2376079.4 \tag{24}$$

再对计算结果进行累减还原可得 2005~2021 年常住人口 X4 拟合值如图 2 所示。

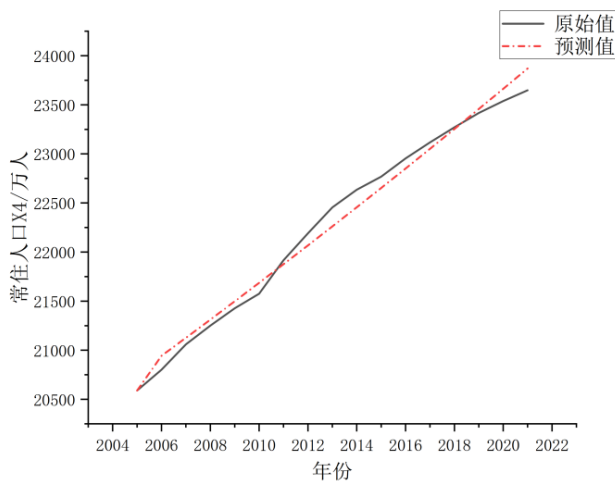


Figure 2. 2005~2021 resident population raw vs. GM(1,1) comparison of projected values

图 2. 2005~2021 常住人口原始值与 GM(1,1)预测值对比

从以上计算结果可知, GM(1,1)灰色模型预测常住人口 X4 的拟合值与原始值的相对误差值最大值 $0.009 < 0.1$, 平均相对误差值为 4.39%, 级比偏差最大值 $0.007 < 0.1$, 意味着模型拟合效果较优。使用 GM(1,1)灰色模型预测 2022~2026 年常住人口 X4 预测值如表 9 所示。

Table 9. GM(1,1) model X4 projections for 2022~2026

表 9. 2022~2026 年 GM(1,1)模型 X4 预测值

年份	2022	2023	2024	2025	2026
预测值 X4/万人	24079.25	24290.21	24503.01	24717.67	24934.21

4.2.2. 自变量指数平滑预测

指数平滑法适用于数据中短期预测；由于数据序列为 17 个(介于 10~20 个之间)，因此选取自变量常住人口 X4 前 2 期数据的平均值 20695.485 作为指数平滑法的初始值。平滑系数 α 值介于 0~1 之间，为了提高模型拟合效果，平滑系数 α 值选取了 0.05~0.95 共 11 个值，并针对 11 个平滑系数 α 值分别选用三种平滑类型最终进行了 33 组测试，选取 33 组测试中 RMSE 值最小的一组作为当前筛选条件下的最佳模型参数。

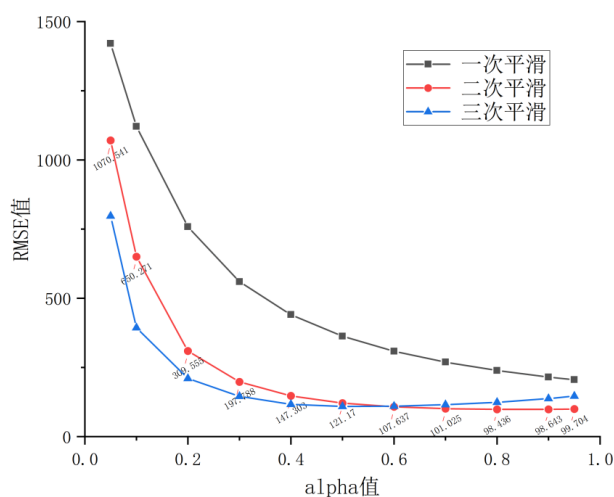


Figure 3. Test results of exponential smoothing parameter selection

图 3. 指数平滑参数选取测试结果

如图 3 所示，测试结果表明当初始值为 20695.485， α 值为 0.800，平滑类型为二次平滑，此时 RMSE 值为 98.436。此为当前条件下最优的模型参数，并以此参数进行模型构建从而得到自变量 X4 拟合值。拟合值与原始值平均相对误差为 3.01%最大相对误差为 0.0094 < 0.1，说明该模型拟合效果好，因此使用指数平滑法得到 2005~2021 年常住人口 X4 拟合值如图 4 所示。

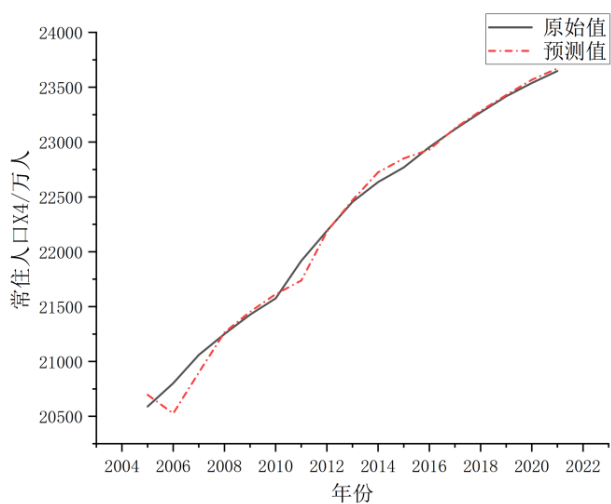


Figure 4. Comparison of raw and index-smoothed projections of resident population in 2005~2021

图 4. 2005~2021 常住人口原始值与指数平滑预测值对比

使用指数平滑法预测 2022~2026 年常住人口 X4 预测值如表 10 所示。

Table 10. Exponential smoothing model X4 Forecasts, 2022~2026

表 10. 2022~2026 年指数平滑模型 X4 预测值

年份	2022	2023	2024	2025	2026
X4 预测值/万人	23765.23	23881.70	23998.16	24114.63	24231.09

4.3. 因变量预测

4.3.1. 因变量灰色 - 回归预测

将 GM(1,1)灰色模型自变量 X4 预测数据带入一元线性回归模型公式(23)中, 对 2022~2026 年长三角物流需求总量进行预测, 并引入 GM(1,1)模型对因变量 Y1 进行单一预测作为对照组, 预测结果如表 11 所示, 灰色 - 回归模型(GM(1,1)-回归模型)预测精度优于单一的 GM(1,1)模型其平均相对误差为 5.24%。

Table 11. GM (1,1)-regression model predictions

表 11. GM(1,1)-回归模型预测值

序号	年份	原始数据	GM(1,1)模型		灰色 - 回归模型	
			拟合数据	相对误差	拟合数据	相对误差
1	2005	374,970	374,970	0.00%	387462.6	3.33%
2	2006	411,985	489452.365	18.80%	472,933	14.79%
3	2007	458,611	528960.749	15.34%	517320.4	12.80%
4	2008	577,493	569397.982	1.40%	562096.9	2.67%
5	2009	585,868	610785.902	4.25%	607265.5	3.65%
6	2010	668,250	653146.859	2.26%	652829.7	2.31%
7	2011	760,045	696503.729	8.36%	698793.3	8.06%
8	2012	829,197	740879.927	10.65%	745159.2	10.14%
9	2013	830,209	786299.417	5.29%	791931.6	4.61%
10	2014	928,182	832786.726	10.28%	839113.8	9.60%
11	2015	849,354	880366.961	3.65%	886709.2	4.40%
12	2016	883,907	929065.815	5.11%	934721.7	5.75%
13	2017	976,763	978909.588	0.22%	983154.7	0.65%
14	2018	1,029,933	1029925.196	0.00%	1032012	0.20%
15	2019	1,047,297	1082140.191	3.33%	1081297.6	3.25%
16	2020	1,101,976	1135582.768	3.05%	1131014.7	2.64%
17	2021	1,191,028	1190281.79	0.06%	1181167.3	0.83%
对未来五年进行预测			预测数据	相对误差	预测数据	相对误差
18	2022		1246266.796		1231759.496	
19	2023		1303568.018		1282794.66	
20	2024		1362216.403		1334277.152	
21	2025		1422243.621		1386210.599	
22	2026		1483682.089		1438599.114	
平均相对误差				5.42%	5.27%	

4.3.2. 因变量指数平滑 - 回归预测

将指数平滑模型自变量 X4 预测值带入一元线性回归模型公式(23), 得到 2022~2026 年长三角物流需求量, 并引入指数平滑模型(初始值为 393477.500, alpha 值为 0.500, 平滑类型为二次平滑)对因变量 Y1 进行单一预测作为对照组, 指数平滑 - 回归模型预测值如表 12 所示, 指数平滑 - 回归模型预测精度优于单一指数平滑模型其平均相对误差为 4.99%。

Table 12. Exponential smoothing-regression model predictions

表 12. 指数平滑 - 回归模型预测值

序号	年份	原始数据	指数平滑模型		指数平滑 - 回归模型	
			拟合数据	相对误差	拟合数据	相对误差
1	2005	374,970	393477.5	4.94%	413127.6	10.18%
2	2006	411,985	374,970	8.99%	372063.6	9.69%
3	2007	458,611	407358.125	11.18%	462404.5	0.83%
4	2008	577,493	463237.875	19.79%	550790.3	4.62%
5	2009	585,868	594933.094	1.55%	596503.3	1.82%
6	2010	668,250	631871.875	5.44%	635,240	4.94%
7	2011	760,045	711987.602	6.32%	665,372	12.46%
8	2012	829,197	812877.133	1.97%	773401.8	6.73%
9	2013	830,209	894043.482	7.69%	842573.4	1.49%
10	2014	928,182	899135.449	3.13%	904120.9	2.59%
11	2015	849,354	981149.829	15.52%	934544.6	10.03%
12	2016	883,907	909583.466	2.91%	954077.3	7.94%
13	2017	976,763	911187.509	6.71%	1001321.4	2.51%
14	2018	1,029,933	997624.393	3.14%	1039510.4	0.93%
15	2019	1,047,297	1067188.265	1.90%	1074586.9	2.61%
16	2020	1,101,976	1092629.417	0.85%	1108213.4	0.57%
17	2021	1,191,028	1142335.601	4.09%	1133065.8	4.87%
对未来五年进行预测			预测数据	相对误差	预测数据	相对误差
18	2022		1233724.247		1155788.7	
19	2023		1288593.593		1183964.7	
20	2024		1343462.939		1212140.7	
21	2025		1398332.286		1240316.8	
22	2026		1453201.632		1268492.5	
平均相对误差				6.24%	4.99%	

4.4. 精度分析

为了验证模型的精度, 选取精度较高的模型对长三角物流需求量进行预测, 文章分别通过 GM(1,1) 灰色模型、指数平滑模型、灰色 - 回归模型(GM(1,1)-回归模型)、指数平滑 - 回归模型对长三角物流需求量进行预测, 并将 2005~2021 年预测值与原始值进行比较, 分别从平均绝对百分比误差(MAPE)、均方根误差(MSE)和平均绝对误差(MAE)三方面评价长三角物流需求量预测精度的好坏, 精度分析结果如表 13 所示。

Table 13. Precision analysis error value
表 13. 精度分析误差值

	GM(1,1)	指数平滑	灰色 - 回归模型	指数平滑 - 回归模型
MAPE	5.42%	6.24%	5.27%	4.99%
MAE	37330.51253	44530.50924	36855.27059	36476.49412
RMSE	48555.77387	55583.03305	45102.40845	45049.12915

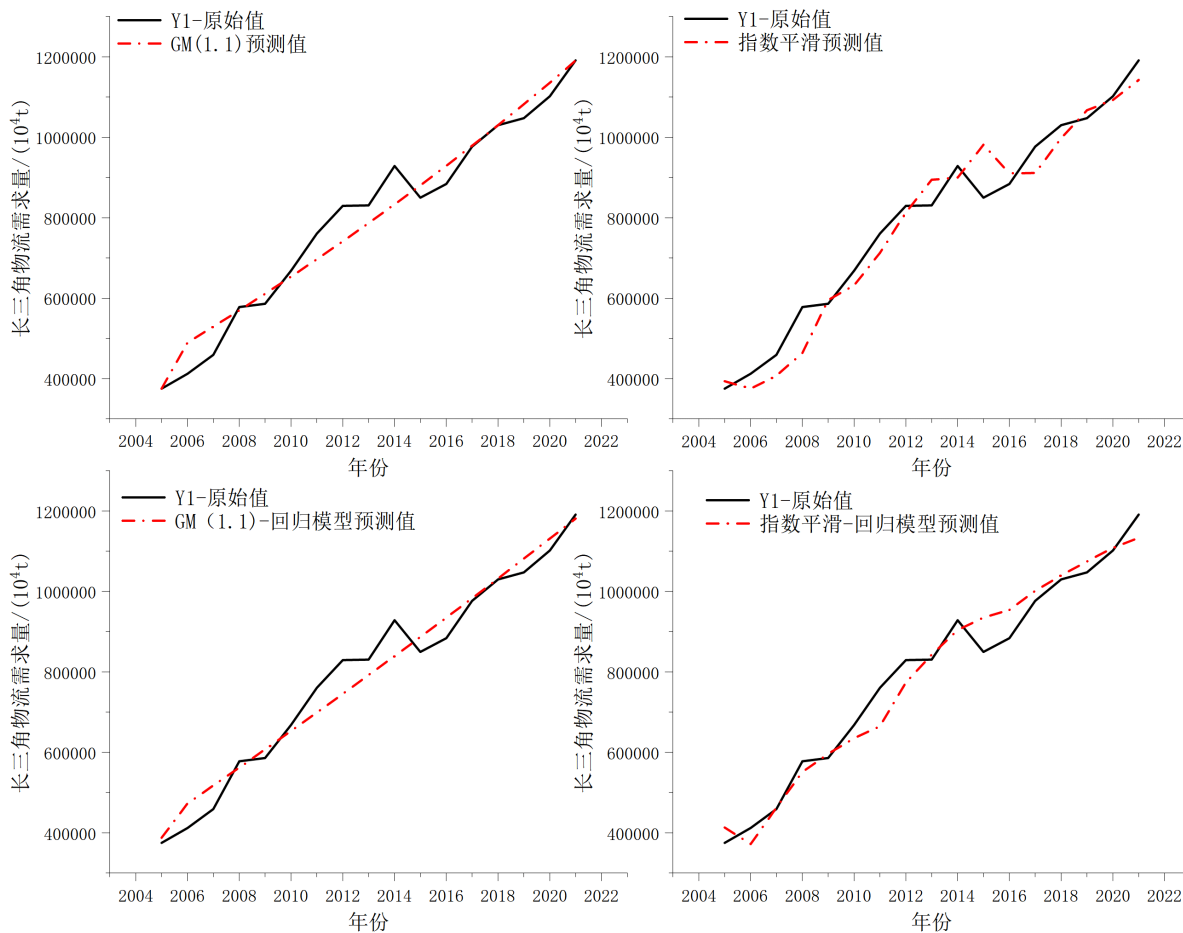


Figure 5. Comparison of predicted and original values of four types of models
图 5. 四类模型预测值与原始值对比

通过对表 13 与图 5 分析可得：研究借助回归模型有效提高了 GM(1,1)模型与指数平滑模型的预测精度，在平均绝对百分比误差(MAPE)层面回归模型对 GM(1,1)模型与指数回归模型的精度提升率分别了 2.77%和 20.03%，回归模型对于指数平滑模型精度提升较大，且指数平滑 - 回归模型在 MAPE、MSE、MAE 三个精度指标层面均优于其他三类模型对于长三角物流需求量的预测其精度最高。

因此选取精度最高的指数平滑 - 回归预测模型作为长三角物流需求量预测模型，最终指数平滑 - 回归模型预测结果如表 14 所示，根据预测结果显示未来 2023~2026 年长三角区域的物流需求量增速维持在 2.2%左右，物流需求量增速较为稳定，模型对长三角物流需求量的中短期预测也比较符合实际经济情况。

Table 14. Logistics demand in Yangtze River Delta, 2022~2026
表 14. 2022~2026 年长三角物流需求量

年份	2022	2023	2024	2025	2026
长三角物流需求总量	1155788.7	1183964.7	1212140.7	1240316.8	1268492.5
增速	-2.96%	2.44%	2.38%	2.32%	2.27%

5. 结论

5.1. 政策建议

结合对长三角经济发展状况的分析及其物流需求预测研究, 为更好地促进长三角经济与物流业高质量协调发展, 文章提出以下三点建议。

1) 构造新形势下的长三角区域物流业发展战略

在当前世界经济增长动力略显不足, 增速出现下降的背景下, 长三角区域作为我国经济发展的重要驱动力量, 促进其区域物流产业健康与可持续发展对区域的经济发展起到重要作用, 地方政府应充分考虑当前的国内外经济形势, 参考长三角物流需求预测值, 并结合长三角内经济发展状况与国内外经济形势, 合理规划长三角区域物流交通基础设施, 避免区域物流基础设施建设与物流需求不匹配而产生资源浪费。调控区域物流服务的供给与需求以实现供需动态平衡[15]。

2) 成立长三角物流一体化发展组织

从 2005 年到 2021 年, 长三角区域的货运量增长了约 3 倍, 其年均增长率约 7.74%, 快速的物流需求增长下长三角物流一体化发展水平仍待提升, 主要表现在: 局部地区交通网络覆盖率不足、交通物流运行效率不佳, 长三角区域物流一体化水平有待提升等方面。建议各地区政府应协同组建与支持长三角物流一体化发展组织, 推进物流一体化发展、因地制宜制定长三角区域内物流相关的政策和法规实现制度层面上对物流业发展的促进和保障, 着力推进物流信息技术标准制定, 并构建长三角物流一体化大数据共享平台, 打破技术壁垒保证物流信息安全、平稳运行。

3) 发挥长三角 G60 科创走廊人才优势

物流产业目前还没有完全摆脱传统的低附加值, 劳动密集型产业, 其中的一个主要原因就是缺乏物流规划和管理的高级人才, 物流技术创新型人才。这一定程度上制约了部分地区物流业的发展。长三角 G60 科创走廊作为长三角一体化发展国家战略的重要平台之一, 覆盖沪苏浙皖三省一市的九个主要城市。为解决长三角物流人才问题, 地方政府可以充分发挥长三角 G60 科创走廊的人才优势, 依托长三角 G60 科创走廊引进与培育物流技术创新型人才, 提升区域物流产业附加值。

5.2. 创新与展望

研究对象方面本文提出的基于指数平滑 - 回归模型的长三角物流需求预测填补了国内对于长三角物流需求预测的空白。区别于传统的区域物流需求影响因素一般选取不同的运输方式作为影响因素, 本文侧重从更宏观的维度选取典型指标作为多元回归因子。最后从预测方法来看, 多数对于区域物流需求的预测多采用单一模型进行预测, 存在一定局限性, 并且容易忽略物流需求量相关影响因素; 而精度较高的智能算法模型, 需要大量的数据训练才能得到较为精确的预测结果; 本研究从原始数据量较少的实际情况出发, 借组多元回归考虑区域物流需求影响因素实现了对指数平滑预测模型与灰色预测模型的精度提升, 并取得了良好的预测精度。

研究虽然考虑了区域物流需求影响因素, 但是为了消除多元回归中多重共线性的干扰, 通过逐步回归消除共线性提高了模型拟合效果, 也剔除了大部分影响因素, 只保留了主要影响因素。未来的研究需

要着眼于在保留更多影响因素的前提下消除多重共线性, 并进一步提高预测精度。

基金项目

上海高校市级重点课程项目(s202108002)。

参考文献

- [1] 过秀成, 谢实海, 胡斌. 区域物流需求分析模型及其算法[J]. 东南大学学报(自然科学版), 2001(3): 24-28.
- [2] 丁红, 王剑斌. 区域物流规划中的物流需求预测研究[J]. 科技广场, 2009(2): 76-77.
- [3] 赵文德, 黄丽娟, 胡子瑜. 基于 BP 神经网络组合模型的广东省区域物流需求能力预测[J]. 物流技术, 2020, 39(12): 49-56+66.
- [4] 黄永福. 高质量发展视角下广东省“十四五”物流需求量预测研究——基于 GM(1, N)灰色模型[J]. 经营与管理, 2022(10): 174-179.
- [5] 雷斌, 陶海龙, 徐晓光. 基于改进粒子群优化算法的灰色神经网络的铁路货运量预测[J]. 计算机应用, 2012, 32(10): 2948-2951+2962.
- [6] 周文杰, 黄慧淼, 黄雪茜. 基于回归模型的西北五省区货运量预测[J]. 物流科技, 2014, 37(10): 142-145.
- [7] 龙忠芬. 货运量回归模型分析——以广东省为例[J]. 物流技术, 2015, 34(22): 130-132+136.
- [8] 张明. 铁路货运量预测及影响因素研究[D]: [硕士学位论文]. 北京: 北京交通大学, 2016.
- [9] 赵美珍. 山西省工业物流需求预测[D]: [硕士学位论文]. 太原: 山西大学, 2016.
- [10] 姜金德, 周海花. 基于区域经济指标的区域物流需求 PCR 预测研究——以江苏省为例[J]. 济南大学学报(社会科学版), 2021, 31(4): 124-132+160.
- [11] 李思聪, 叶静. 基于灰色回归模型的农产品冷链物流需求分析及预测[J]. 公路交通科技, 2022, 39(5): 166-174.
- [12] 孟建军, 陈鹏芳, 李德仓, 等. 铁路货运量预测研究综述[J]. 铁道标准设计, 2022, 66(10): 18-26.
- [13] 高爱霞, 满广富, 姚兴华. 基于主成分回归的区域物流发展预测模型构建与实证分析——以泰安市为例[J]. 山东工会论坛, 2018, 24(2): 68-76.
- [14] 史红伟, 陈家亮, 张继群, 等. 基于灰色系统理论的高校用水量预测[J]. 河北工程大学学报(自然科学版), 2022, 39(1): 78-83.
- [15] 孙战旗. 基于系统动力学的厦漳泉区域物流协同发展研究[D]: [硕士学位论文]. 厦门: 集美大学, 2015.