

基于PMGM(1,N)灰色马尔可夫模型的 货运量需求预测*

刘星辉,任建伟,李雪琦

(内蒙古大学交通学院,呼和浩特 010070)

摘要:为提升城市货运量需求预测精度,构建了PMGM(1,N)灰色马尔可夫预测模型,并以宁波市和芜湖市为例进行实证研究。首先,建立多因素指标体系,并通过灰色关联分析厘清核心变量;然后,采用粒子群算法(Particle swarm optimization, PSO)优化MGM(1,N)模型的背景值插值系数,提升模型拟合能力;最后,引入马尔可夫状态转移矩阵修正残差,提高模型对数据波动的适应性。实证结果表明,PMGM(1,N)灰色马尔可夫模型在拟合精度和预测精度上均优于GM(1,1)、多元线性回归、MGM(1,N)及PMGM(1,N)模型,充分验证了其在城市货运需求预测中的有效性与优越性。

关键词:城市货运量预测;灰色预测;PSO算法;马尔可夫理论

中图分类号:U121 **文献标志码:**A

在经济全球化与区域一体化加速推进的背景下,物流需求发展呈快速增长趋势,而货运量需求预测作为现代物流体系和城市交通规划的核心环节,对优化资源配置、提升运输效率以及制定运输政策具有重要意义。

目前,在货运量需求预测研究方面,回归模型、神经网络模型、灰色预测模型和混合预测模型被广泛使用。例如,Sultanbek等^[1]使用差分自回归移动平均(Autoregressive integrated moving average, ARIMA)模型对哈萨克斯坦的铁路货运需求进行了精确预测;Peng等^[2]将自回归移动平均(Autoregressive moving average, ARMA)过滤器和门控循环单元(Gate recurrent unit, GRU)相结合,提出了新的神经网络模型,并对铁路冷链货运需求进行了短期预测;Van等^[3]使用基于外生因素及季节性的差分自回归移动平均(Seasonal autoregressive integrated moving average with exogenous regressors, SARIMAX)方法对莱茵河集装箱吞吐量进行了预测;Zhang等^[4]建立了反向传播(Back propagation, BP)神经网络,对我国公路货运量进行了预测;Yang等^[5]将支持向量回归(Support vector regression, SVR)与模糊集相结合,实现了对航空货运量的预测;龙宇等^[6]通过构建组合预测模型,对我国铁路货运量进行了预测;柯桥等^[7]对灰色预测模型进行了改进,并与神经网络相结合,基于该模型预测了三峡枢纽过坝货运量;钟蒙等^[8]将灰色关联度分析与BP神经网络相结合,提高了模型非线性学习能力,并对江西省公路货运量进行了预测;谭雪等^[9]构建了GRU深度网络预测模型,并对铁路月度货运量进行了预测;张昭俊等^[10]利用校正的Logit模型对三峡枢纽五条路径的未来过坝货运量进行了预测。然而,城市货运需求系统具有非线性、时变性和小样本的特点,这些预测模型虽各有特点,但在

* 收稿日期:2025-05-02;修回日期:2025-07-10

基金项目:国家自然科学基金项目(72262024)

作者简介:刘星辉(2000—),男,陕西神木人,2023级硕士研究生。E-mail:1309397067@qq.com

通信作者:任建伟(1982—),男,山西寿阳人,教授,博士。主要从事物流规划与设计、DEA效率评价研究。

E-mail:renjianwei309@126.com

城市货运量需求预测上存在应用局限。其中,时间序列模型没有考虑外界因素的影响;回归模型虽然考虑了外部影响因素,但对数据质量要求高且非线性关系处理能力有限;神经网络模型能够很好地处理非线性关系,但易受训练数据规模和噪声的干扰;而基于灰色系统理论的灰色预测模型,能对小样本、非线性、不确定性高和信息不完整的系统建模,其中MGM(1,N)灰色预测模型充分考虑了多因素耦合的预测场景,很好地切合了城市货运量需求预测特征。因此,可以使用MGM(1,N)灰色预测模型对城市货运量进行预测。

许多学者对MGM(1,N)模型进行了不同类型的研究,以提高其预测精度。Li等^[11]将MGM(1,N)模型的插值系数进行了优化,改进后的模型预测精度得到了提高;Shan等^[12]考虑了新信息优先原则,提高了MGM(1,m)模型的稳定性;陈先飞等^[13]利用遗传算法搜索灰色模型的最优参数,建立了MGM(1,n,r)模型,结果表明模型精度得到了提高;朱文静等^[14]对MGM(1,n)模型的初始值和背景值进行了优化,通过实例验证了模型的准确性和可靠性;李树等^[15]通过蚁群算法确定组合预测模型权重,将MGM(1,N)模型、新息MGM(1,N)模型和初值优化MGM(1,N)模型相结合,测试结果表明模型拟合精度得到了较大提升;胡涛等^[16]优化了MGM(1,N)模型的背景值,提高了模型的精度;姜晋云等^[17]对MGM(1,N)模型进行了残差修正,避免了因数据跳动较大而引起的预测风险,提高了模型的可靠性。

综上所述,多变量灰色预测模型所呈现的特性能够很好地反映城市货运需求非线性、时变性和小样本的灰色特性。鉴于此,本文选择MGM(1,N)模型作为基础模型,同时针对模型在背景值插值系数设定方面的局限,以及对系统随机波动和突发变化刻画能力的不足,采用粒子群算法(Particle swarm optimization, PSO)对背景值插值系数进行优化,并引入马尔可夫理论对残差序列进行修正,以增强灰色模型对系统随机性与突变因素的刻画能力,最终构建出适用于城市货运量需求预测的PMGM(1,N)灰色马尔可夫模型。

1 模型方法

1.1 指标体系构建

城市货运量是一个复杂的社会经济现象,科学、全面地选取影响货运量的相关因素是建立可靠预测模型的前提。为全面反映影响城市货运需求的主要因素,本文从宏观经济运行、产业结构、消费能力、对外贸易、人口规模以及基础设施投资等多个维度出发,遵循系统性、相关性、代表性、可量化性和数据可获得性等原则,并结合相关文献与城市实际发展特征,构建出多因素指标体系,如表1所示。

表1 城市货运需求影响因素指标体系

Table 1 Indicator system for influencing factors of urban freight demand

指标名称	单位	变量类型	指标含义
货运量	万 t	因变量(x_1)	反映城市整体货运需求的总量与变化,是研究的预测目标
GDP	亿元	自变量(x_2)	代表城市区域经济总量,是货运需求的宏观基础驱动力
第一产业增加值	亿元	自变量(x_3)	代表农业和林业等基础产业,反映城市农产品运输情况
第二产业增加值	亿元	自变量(x_4)	代表工业与建筑业,是大宗物资等货运需求的主要来源
第三产业增加值	亿元	自变量(x_5)	代表服务业,体现服务型经济对物流的间接推动作用
社会消费品零售总额	亿元	自变量(x_6)	代表城市消费水平,体现消费品流通对物流的直接影响作用
进出口总额	亿元	自变量(x_7)	代表城市对外贸易程度,反映外贸货物运输需求
常住人口数	万人	自变量(x_8)	代表城市基本规模,是城市货运需求增长的长期支撑因素
固定资产投资总额	亿元	自变量(x_9)	代表城市基础设施发展水平,反映城市未来产能和运输能力

1.2 灰色关联分析

在初步构建多因素指标体系后,本文采用灰色关联分析方法,定量评估各影响因素指标与城市货运量之间的关联程度,筛选出与货运量关联度最高、关系最紧密的核心指标进行模型构建,从而有效减少冗余特征,提高模型效率。通过均值化对数据进行无量纲处理后,灰色关联度计算步骤如下所示。

① 关联系数 δ_i 计算。

$$\delta_i(k) = \frac{\min_i \min_k \Delta_i(k) + \rho \max_i \max_k \Delta_i(k)}{\Delta_i(k) + \rho \max_i \max_k \Delta_i(k)} \quad (1)$$

$$\Delta_i(k) = |y(k) - x_i(k)| \quad (2)$$

其中, ρ 为分辨系数,取值区间为 $(0, 1)$,通常取 $\rho = 0.5$; $y(k)$ 为参考数列; $x_i(k)$ 为比较序列。

② 关联度 r_i 计算。

$$r_i = \frac{1}{n} \sum_{k=1}^n \delta_i(k), k = 1, 2, \dots, n \quad (3)$$

③ 关联度排序。按照关联度大小进行排序,如果 $r_1 < r_2$,则代表参考数列 $y(k)$ 与比较数列 $x_2(k)$ 关联度更高。

1.3 PMGM(1,N)灰色马尔可夫模型构建

PMGM(1,N)灰色马尔可夫模型的预测流程如图1所示,首先构建传统MGM(1,N)模型,然后引入PSO算法对模型背景值插值系数进行全局寻优,以降低固定系数所带来的预测误差,最后通过引入马尔可夫状态转移矩阵对残差进行修正,增强模型对数据波动的适应能力,最终形成PMGM(1,N)灰色马尔可夫模型。

1.3.1 传统MGM(1,N)模型构建

准备 n 个以 m 为周期的灰时间序列,记为 $x_i^{(0)}(k) (i = 1, 2, \dots, n; k = 1, 2, \dots, m)$,生成 $x_i^{(0)}(k)$ 的一次累加序列 $x_i^{(1)}(k)$,即

$$x_i^{(1)}(k) = \sum_{j=1}^k x_i^{(0)}(j) \quad (4)$$

构建 n 元一阶常微分方程组,即

$$\begin{cases} \frac{dx_1^{(1)}}{dt} = a_{11}x_1^{(1)} + a_{12}x_2^{(1)} + \dots + a_{1n}x_n^{(1)} + b_1 \\ \frac{dx_2^{(1)}}{dt} = a_{21}x_1^{(1)} + a_{22}x_2^{(1)} + \dots + a_{2n}x_n^{(1)} + b_2 \\ \vdots \\ \frac{dx_n^{(1)}}{dt} = a_{n1}x_1^{(1)} + a_{n2}x_2^{(1)} + \dots + a_{nn}x_n^{(1)} + b_n \end{cases} \quad (5)$$

记

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}, B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix},$$

$$X^{(0)}(k) = (x^{(0)}_1(k), x^{(0)}_2(k), \dots, x^{(0)}_n(k))^T,$$

$$X^{(1)}(k) = (x^{(1)}_1(k), x^{(1)}_2(k), \dots, x^{(1)}_n(k))^T,$$

则动态微分方程组可表示为

$$\frac{dX}{dt} = AX^{(1)} + B \quad (6)$$

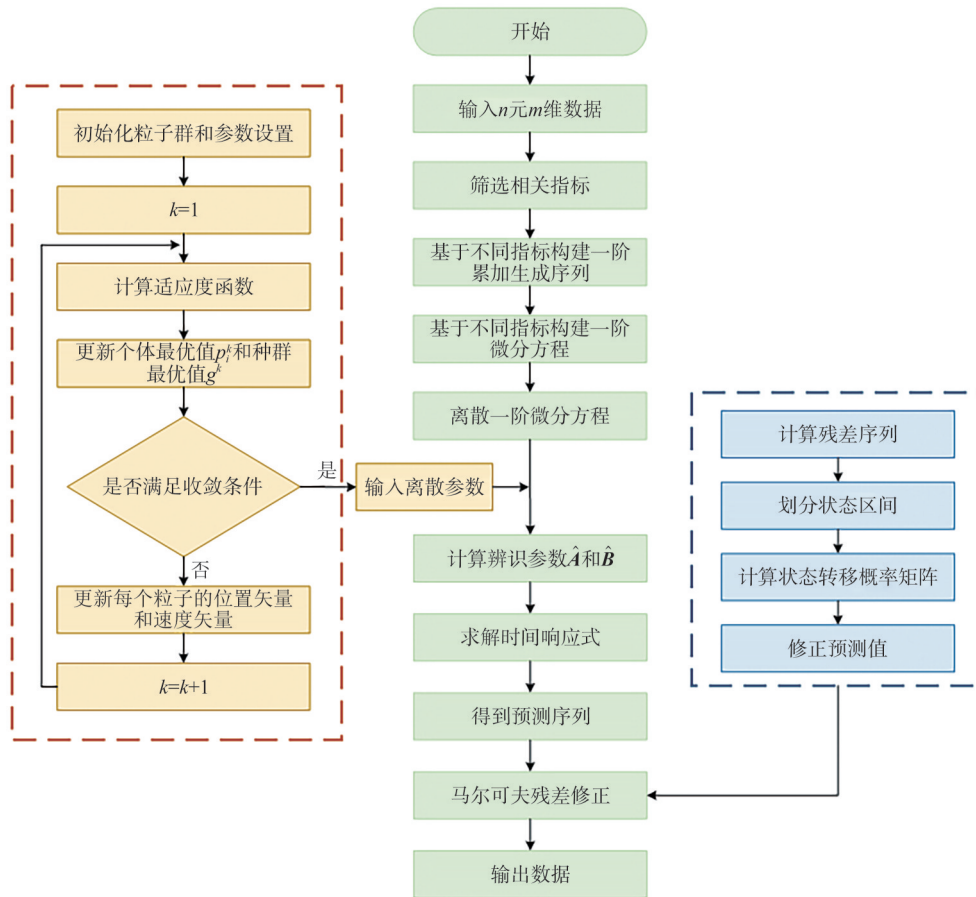


图 1 PMGM(1,N)灰色马尔可夫模型预测流程

Fig. 1 PMGM(1,N) grey Markov model prediction process

通过积分生成变换,得到连续时间响应函数为

$$X^{(1)}(t) = e^{At} X^{(1)}(1) + A^{-1}(e^{At} - I)B \tag{7}$$

其中, $e^{At} = I + \sum_{k=1}^{\infty} \frac{A^k}{k!} t^k$, I 为单位矩阵。

为求辨识参数,将式(5)离散化可得

$$x_i^{(1)}(k) = \sum_{j=1}^n \frac{a_{ij}}{2} [x_j^{(1)}(k) + x_j^{(1)}(k-1)] + b_i \tag{8}$$

记 $a_i = (a_{i1}, a_{i2}, \dots, a_{in}, b_i)^T, i = 1, 2, \dots, n$ 。由最小二乘法可得 a_i 的辨识值 \hat{a}_i 为

$$\hat{a}_i = \begin{bmatrix} \hat{a}_{i1} \\ \hat{a}_{i2} \\ \vdots \\ \hat{a}_{in} \\ \hat{b}_i \end{bmatrix} = (L^T L)^{-1} L^T Y_i, i = 1, 2, \dots, n \tag{9}$$

其中,

$$L = \begin{bmatrix} Z_1^{(1)}(2) & Z_2^{(1)}(2) & \cdots & Z_n^{(1)}(2) \\ Z_1^{(1)}(3) & Z_2^{(1)}(3) & \cdots & Z_n^{(1)}(3) \\ \vdots & \vdots & \vdots & \vdots \\ Z_1^{(1)}(m) & Z_2^{(1)}(m) & \cdots & Z_n^{(1)}(m) \end{bmatrix}, Y_i = \begin{bmatrix} x_i^{(0)}(2) \\ x_i^{(0)}(3) \\ \vdots \\ x_i^{(0)}(m) \end{bmatrix}$$

$$Z_i^{(1)}(k) = \frac{1}{2} [x_i^{(1)}(k) + x_i^{(1)}(k-1)] \quad (10)$$

则辨识参数 \hat{A} 和 \hat{B} 为

$$\hat{A} = \begin{bmatrix} \hat{a}_{11} & \hat{a}_{12} & \cdots & \hat{a}_{1n} \\ \hat{a}_{21} & \hat{a}_{22} & \cdots & \hat{a}_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ \hat{a}_{n1} & \hat{a}_{n2} & \cdots & \hat{a}_{nn} \end{bmatrix}, \hat{B} = \begin{bmatrix} \hat{b}_1 \\ \hat{b}_2 \\ \vdots \\ \hat{b}_n \end{bmatrix}.$$

参数估计离散化后的时间响应函数为

$$\hat{X}^{(1)}(k) = e^{\hat{A}(k-1)} X^{(1)}(1) + \hat{A}^{-1} (e^{\hat{A}(k-1)} - I) \hat{B} \quad (11)$$

$$\hat{X}^{(0)}(1) = \hat{X}^{(1)}(1) = X^{(1)}(1) \quad (12)$$

$$\hat{X}^{(0)}(k) = \hat{X}^{(1)}(k) - \hat{X}^{(1)}(k-1) \quad (13)$$

1.3.2 基于PSO算法的插值优化

PSO算法的核心理念是通过模拟生物群体的社会行为,利用个体间的信息交流与协作机制来实现对问题解空间的探索。相较于传统优化算法,PSO具有实现简便和收敛效率高等特点,同时在复杂问题求解中表现出较强的适应性和鲁棒性。通过式(10)可知,传统MGM(1,N)模型中各个变量的背景值插值系数默认为0.5,这也是模型进行预测时导致误差较大的主要原因。因此,本文通过粒子群算法寻找插值系数的最优解,设置插值系数为 λ_i ,则式(10)更新为

$$Z_i^{(1)}(k) = \lambda_i x_i^{(1)}(k) + (1 - \lambda_i) x_i^{(1)}(k-1) \quad (14)$$

PSO算法的速度更新公式为

$$v_i^{k+1}(k) = \omega_i v_i^k + c_1 r_1 (p_i^k - x_i^k) + c_2 r_2 (g^k - x_i^k) \quad (15)$$

其中, v_i^k 为粒子 i 在第 k 次迭代中的速度; ω_i 为惯性权重; c_1 为个体学习因子; c_2 为群体学习因子; p_i^k 为第 i 个粒子 k 次迭代后的最优解; g^k 为种群 k 次迭代后的最优解; r_1, r_2 是区间 $[0, 1]$ 的随机数。

采用线性递减权重策略获取动态惯性权重,公式为

$$\omega_i = (\omega_{\text{ini}} - \omega_{\text{min}})(G - g_i)/G + \omega_{\text{min}} \quad (16)$$

其中, ω_{ini} 为初始权重; ω_{min} 为最小权重; G 为最大迭代次数; g_i 为当前迭代次数。

PSO算法的位置更新公式为

$$x_i^{k+1}(k) = x_i^k + v_i^{k+1} \quad (17)$$

本文以平均绝对百分比误差(Mean absolute percentage error, MAPE)作为PSO算法的适应度函数,误差越低,预测值与真实值拟合效果越好,PSO算法的适应度函数 E_{MAPE} 公式为

$$E_{\text{MAPE}} = \frac{1}{k} \sum_{i=1}^k \frac{|x_i^{(0)} - \hat{x}_i^{(0)}|}{x_i^{(0)}} \times 100\% \quad (18)$$

其中, $x_i^{(0)}$ 和 $\hat{x}_i^{(0)}$ 分别为实际值和预测值; k 为时间序列的个数。

1.3.3 马尔可夫残差修正

灰色预测模型能够很好地解决城市货运需求非线性和小样本的问题,但处理数据波动性的能力较差,可以通过马尔可夫理论有效处理数据波动性大的问题,以弥补灰色预测模型的局限性。马尔可夫残差修正计算步骤如下。

① 划分状态区间。根据原始数据序列和PMGM(1,N)模型预测序列,计算残差序列,根据残差序列变化范围,按照等分位点将其划分成若干个状态区间。

② 建立状态转移概率矩阵。计算变量从状态 $S_i(i=1, 2, \dots, n)$ 转移到 $S_j(j=1, 2, \dots, n)$ 的概率 p_{ij} ,最终得到一步状态转移概率矩阵 P 为

$$P = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{bmatrix} \quad (19)$$

式中, $p_{ij} = m_{ij}/M_i$, m_{ij} 为状态 S_i ($i = 1, 2, \dots, n$) 转移到 S_j ($j = 1, 2, \dots, n$) 的次数; M_i 为状态 S_i 的总出现次数。

③ 修正预测值。通过期望值修正法修正预测值, 公式为

$$\hat{x}_{(t+1)} = x_{(t+1)} + \sum_{j=1}^n p_{ij} u_j \quad (20)$$

其中, $\hat{x}_{(t+1)}$ 为修正预测值; $x_{(t+1)}$ 为灰色预测值; $\sum_{j=1}^n p_{ij} u_j$ 为期望残差; u_j 为状态 S_j ($j = 1, 2, \dots, n$) 的残差中值。

1.4 模型适用性条件

PMGM(1,N)灰色马尔可夫预测模型在处理小样本、不完全信息及非线性系统等方面具备独特优势,但其建模效果仍受到灰色模型理论假设和系统特征的制约。因此,模型本身在应用中存在一定的适用范围和条件限制,若脱离这些前提条件,预测结果的科学性与可靠性将难以保障。为确保模型在实际应用中具有科学性与可靠性,本文提出以下适用性条件。

1) 变量间具有较强的关联性

经灰色关联分析筛选出的核心影响因素与城市货运量之间应存在显著的灰色关联关系(关联度大于0.7),确保多变量建模具有合理的因果性和相关性基础。

2) 数据变化趋势相对平稳且波动幅度在可建模范围内

虽然PMGM(1,N)灰色马尔可夫模型通过马尔可夫理论增强了对系统短期波动与非线性扰动的描述能力,但灰色模型依赖系统行为的可拟合性,若时间序列趋势存在明显不稳定性或城市货运系统在预测期内受到严重干扰(如极端突发事件或重大政策变动),将严重影响模型拟合与预测精度。因此,要求所预测城市的货运量数据波动幅度适中且发展趋势相对稳定。

3) 数据具备一定连续性和时效性

尽管灰色模型适用于小样本建模,但仍需保障时间序列完整、连续。本文要求建模数据至少包含10年连续观测值,以确保参数估计的稳定性与预测的可行性。

4) 系统具备灰色建模特征

所选城市货运系统应呈现“部分信息已知、部分未知”的灰色特性,即既非完全随机系统,也非完全确定系统,从而保证灰色建模方法的理论适用性。

综上所述,本文虽然通过背景值参数优化与马尔可夫残差修正增强了预测模型对复杂系统行为的适应能力,但为提升模型在实际应用中的稳定性与推广性,建模前应对数据特性与城市发展背景进行适用性检验,仅当所需预测的城市满足上述前提条件时,方可采用PMGM(1,N)灰色马尔可夫模型进行预测分析,从而提高模型应用的准确性和实用性。

2 实证分析

为验证PMGM(1,N)灰色马尔可夫预测模型的有效性与普适性,本文选取宁波市和芜湖市进行实证分析。两个城市产业结构完备,物流需求稳定,相关指标数据连续完整(包含10年及以上连续观测值),核心影响因素与货运量间的灰色关联度均大于0.7,符合本文提出的模型适用性条件。

2.1 宁波市PMGM(1,4)灰色马尔可夫货运量预测

数据选取时间跨度为2012—2024年。其中,2012—2022年数据用于数据建模,2023—2024年数据用于预测验证。各指标含义详见表1,具体数据如表2所示。

表2 2012—2024年宁波市货运量及相关指标

Table 2 Freight volume and related indicators in Ningbo City from 2012 to 2024

年份	x_1 /万t	x_2 /亿元	x_3 /亿元	x_4 /亿元	x_5 /亿元	x_6 /亿元	x_7 /亿元	x_8 /万人	x_9 /亿元
2012	32615.72	6862.02	263.99	3572.82	3025.21	2190.45	965.73	804.10	2901.42
2013	35409.00	7432.09	265.15	3794.14	3372.80	2433.72	1003.29	821.60	3422.95
2014	40406.50	7904.81	267.60	4108.26	3528.95	2738.31	1047.04	841.60	3989.46
2015	42083.20	8295.35	275.13	4210.91	3809.31	3050.18	1004.66	858.20	4506.58
2016	46258.00	8972.83	290.72	4463.78	4218.33	3337.61	949.23	872.80	4961.39
2017	52519.70	10146.55	297.05	5119.01	4730.49	3672.91	1121.97	890.30	5009.58
2018	61454.00	11193.14	302.33	5507.15	5383.66	3964.39	1301.04	911.50	5189.92
2019	68407.00	12035.11	323.14	5686.34	6025.63	4268.56	1330.76	929.40	5610.31
2020	71898.10	12599.22	338.52	5793.11	6467.59	4238.26	1412.77	942.00	5918.88
2021	78747.10	14703.20	359.28	6867.26	7476.66	4649.10	1845.28	954.40	6569.95
2022	80100.40	15704.30	381.99	7413.47	7908.84	4896.72	1904.98	961.80	7253.23
2023	87000.00	16452.80	383.80	7540.50	8528.50	5212.60	1817.18	969.70	7797.22
2024	95000.00	18147.70	451.10	7882.70	9813.90	5605.10	1952.74	977.70	7688.06

注:2012—2023年数据来源于《宁波市统计年鉴》,数据访问地址为<https://tjj.ningbo.gov.cn/col/col1229042824/index.html>;2024年数据来源于《2024年宁波市国民经济和社会发展统计公报》,数据访问地址为https://tjj.ningbo.gov.cn/art/2025/3/20/art_1229042825_58921061.html。

根据表2中的数据,通过式(1)—(3)进行灰色关联分析,得到货运量与各影响因子之间的关联度如表3所示。

为提高模型计算效率,选择关联度排名前四的指标因素构建PMGM(1,4)灰色马尔可夫模型,即GDP、第二产业增加值、第三产业增加值和社会消费品零售总额。

表3 宁波市货运量与各影响因素之间的关联度

Table 3 Correlation degree between freight volume and influencing factors in Ningbo City

影响因子	关联度	影响因子	关联度
x_2	0.81	x_6	0.76
x_3	0.55	x_7	0.67
x_4	0.74	x_8	0.49
x_5	0.86	x_9	0.73

2.1.1 PSO参数优化

基于式(4)—(13)建立传统MGM(1,N)模型,并结合式(14)—(18),利用Matlab实现PSO背景值插值系数优化,参数设置如表4所示。

模型经过500次迭代达到收敛,收敛过程如图2所示,最终得到优化后的背景值插值系数 $\lambda_i=(0.4604,0.4650,0.4770,0.4525,0.4955)$ 。

表4 PSO算法参数设置

Table 4 PSO algorithm parameter settings

参数设置	数值
种群数量	500
迭代次数	500
初始惯性权重	0.9
最小惯性权重	0.4
个体学习因子	1.5
群体学习因子	1.5

2.1.2 PMGM(1,4)模型预测

将传统 MGM(1,4)模型中各个变量的默认背景值插值系数(0.5,0.5,0.5,0.5,0.5)修改为优化后的背景值插值系数 $\lambda_i=(0.4604, 0.4650, 0.4770, 0.4525, 0.4955)$,即可得到改进后的 PMGM(1,4)预测模型。根据表 2 中宁波市 2012—2022 年的相关数据进行建模预测,得到 PMGM(1,4)模型货运量预测结果如表 5 所示。

2.1.3 马尔可夫残差修正

根据表 5 可知,2012—2022 年的货运量残差变换范围在区间 $[-1675.24, 1635.13]$ 内,本文按照等分位点法将残差划分为 5 种状态,具体状态划分情况如表 6 所示。根据表 5 残差情况,可得 2012—2022 年货运量残差值对应的状态分别为(3,1,5,3,2,3,5,4,1,4,2)。进一步,依据式(19)的马尔可夫状态转移统计原理,得到一步状态转移概率矩阵 P 为

$$P = \begin{bmatrix} 0 & 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 1 & 0 & 0 \\ 1/3 & 1/3 & 0 & 0 & 1/3 \\ 1/2 & 1/2 & 0 & 0 & 0 \\ 0 & 0 & 1/2 & 1/2 & 0 \end{bmatrix}。$$

表 5 PMGM(1,4)模型货运量预测值

Table 5 Freight volume forecast values of the PMGM(1,4) model

年份	原始值/万 t	预测值/万 t	残差/万 t	相对误差/%
2012	32615.72	32615.72	0	0
2013	35409.00	37084.24	-1675.24	4.73
2014	40406.50	38771.37	1635.13	4.05
2015	42083.20	41812.50	270.70	0.64
2016	46258.00	46449.56	-191.56	0.41
2017	52519.70	52672.79	-153.09	0.29
2018	61454.00	59970.95	1483.05	2.41
2019	68407.00	67294.23	1112.77	1.63
2020	71898.10	73411.73	-1513.63	2.11
2021	78747.10	77671.40	1075.70	1.37
2022	80100.40	80866.25	-765.85	0.96
2023	87000.00	85615.43	1384.57	1.59
2024	95000.00	95617.70	-617.70	0.65

根据一步状态转移概率矩阵 P ,可对 PMGM(1,4)货运量预测数据进行马尔可夫修正。以 2013 年货运量预测数据为例,2012 年宁波市货运量处于状态 3,通过一步状态转移概率矩阵 P 可知,2013 年货运量状态行向量为(1/3,1/3,0,0,1/3),则根据式(20)可得

$$\hat{x}_{(2013)} = 37084.24 + (1/3)(-1333) + (1/3)(-573.5) + (1/3)1430 = 36925.40。$$

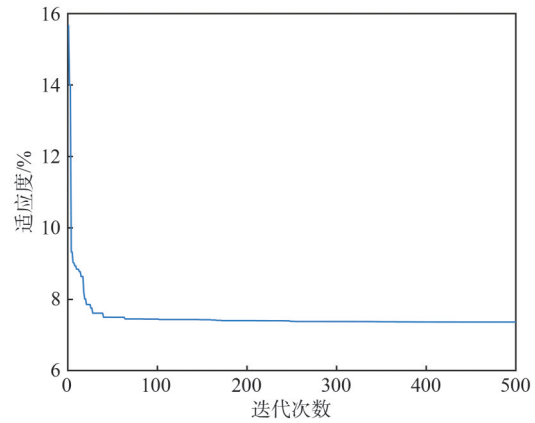


图 2 迭代过程

Fig. 2 Iteration process

同理,对其余年份进行马尔可夫残差修正,最终得到宁波市货运量PMGM(1,4)灰色马尔可夫预测结果如表7所示,不同模型预测相对误差结果如表8所示。

根据表8可知,在模型拟合方面,PMGM(1,4)灰色马尔可夫模型拟合精度最优,平均相对误差为1.02%,表现出更强的建模能力。在模型预测方面,MGM(1,4)模型在预测时误差最大,平均相对误差为11.82%,主要原因

在于其背景值插值系数选取不当,未能有效拟合系统动态特征。相比之下,PMGM(1,4)模型通过改进背景值插值系数,显著改善了模型预测性能,将平均相对误差降至1.12%,而PMGM(1,4)灰色马尔可夫模型进一步结合马尔可夫残差修正,在预测阶段表现最优,平均相对误差仅为0.98%,在5种预测模型中精度最高。此外,从图3可以看出,PMGM(1,4)灰色马尔可夫模型的预测数据与实际值曲线拟合效果最好,表明PMGM(1,4)灰色马尔可夫模型的预测值更具参考价值。

表6 状态划分情况

Table 6 Status division details

状态	区间	区间中值
1	(-1676.00, -990.00)	-1333.00
2	(-990.00, -157.00)	-573.50
3	(-157.00, 351.00)	97.00
4	(351.00, 1224.00)	787.50
5	(1224.00, 1636.00)	1430.00

表7 宁波市货运量不同模型预测值

Table 7 Forecast values of freight volume in Ningbo City using different models

年份	实际值/万t	模型预测值/万t				
		GM(1,1)	多元线性回归	MGM(1,4)	PMGM(1,4)	PMGM(1,4)灰色马尔可夫
2012	32615.72	32615.72	31988.76	32615.72	32615.72	32615.72
2013	35409.00	36695.20	36295.92	36080.16	37084.24	36925.40
2014	40406.50	40277.17	38516.83	37839.61	38771.37	39880.12
2015	42083.20	44208.78	43492.47	41751.17	41812.50	42254.75
2016	46258.00	48524.17	48620.61	48044.72	46449.56	46290.72
2017	52519.70	53260.81	52606.07	55974.57	52672.79	52769.79
2018	61454.00	58459.81	59444.10	63573.57	59970.95	59812.12
2019	68407.00	64166.31	67646.14	68320.48	67294.23	67736.48
2020	71898.10	70429.83	71321.20	68864.03	73411.73	72458.48
2021	78747.10	77304.77	78428.41	67114.09	77671.40	78780.15
2022	80100.40	84850.79	81528.93	69118.40	80866.25	79913.00
2023	87000.00	93133.42	81288.99	82945.00	85615.43	85712.43
2024	95000.00	102224.54	98464.34	113031.61	95617.70	95458.87

表8 宁波市货运量不同模型预测相对误差

Table 8 Relative error of Ningbo City freight volume forecast by different models

模型	GM(1,1)	多元线性回归	MGM(1,4)	PMGM(1,4)	PMGM(1,4)灰色马尔可夫
拟合平均相对误差/%	3.29	2.28	5.07	1.69	1.02
预测平均相对误差/%	7.33	5.11	11.82	1.12	0.98

2.2 芜湖市PMGM(1,4)灰色马尔可夫货运量预测

数据选取时间跨度为2013—2024年。其中,2013—2022年数据用于数据建模,2023—2024年数据用于预测验证。具体数据如表9所示。

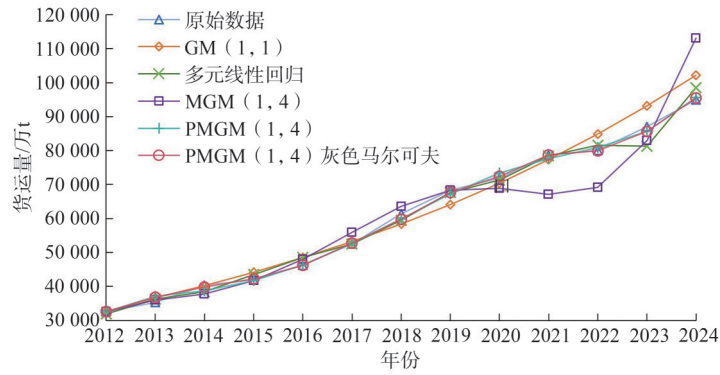


图 3 宁波市货运量实测值与 5 种模型预测值对比

Fig. 3 Comparison between measured freight volume and predicted values of five models in Ningbo City

表 9 2013—2024 年芜湖市货运量及相关指标

Table 9 Freight volume and related indicators in Wuhu City from 2013 to 2024

年份	x_1 /万 t	x_2 /亿元	x_3 /亿元	x_4 /亿元	x_5 /亿元	x_6 /亿元	x_7 /亿元	x_8 /万人	x_9 /亿元
2013	25461.24	2099.53	128.60	1388.22	582.70	559.95	336.50	359.60	2040.65
2014	31239.66	2307.90	136.20	1516.02	655.68	632.71	395.97	361.70	2392.64
2015	26202.40	2457.32	120.02	1540.60	796.70	733.04	424.71	365.40	2709.19
2016	24959.53	2699.44	126.90	1506.32	1066.22	828.17	379.21	367.00	3006.90
2017	28449.82	2963.26	129.67	1623.40	1210.19	930.86	430.69	369.60	3342.24
2018	29244.30	3278.53	133.03	1710.64	1434.86	1028.26	454.95	374.80	3666.44
2019	31628.52	3618.26	146.57	1756.87	1714.82	1149.61	497.05	377.80	4018.42
2020	29871.29	3753.02	161.56	1787.28	1804.18	1584.35	581.42	364.40	4267.56
2021	30461.98	4302.60	169.58	2048.96	2084.09	1972.95	743.82	367.20	4937.56
2022	31504.53	4502.13	179.25	2135.58	2187.30	1995.12	915.32	373.10	5204.19
2023	38747.50	4741.07	182.99	2181.22	2376.86	2091.77	1275.17	375.60	5167.76
2024	38688.43	5120.50	188.60	2295.80	2636.10	2169.04	1469.16	379.70	5529.51

注：2013—2023 年数据来源于《芜湖市统计年鉴》，数据访问地址为 <https://tjj.wuhu.gov.cn/zhsj/tjnj/index.html>；2024 年数据来源于《芜湖市 2024 年国民经济和社会发展统计公报》，数据访问地址为 <https://tjj.wuhu.gov.cn/tjxx/tjgb/8768514.html>。

通过灰色关联分析，得到芜湖市货运量与各影响因子之间的关联度如表 10 所示。

表 10 芜湖市货运量与各影响因素之间的关联度

Table 10 Correlation degree between freight volume and influencing factors in Wuhu City

影响因子	关联度	影响因子	关联度
x_2	0.73	x_6	0.55
x_3	0.81	x_7	0.64
x_4	0.85	x_8	0.84
x_5	0.58	x_9	0.71

选择 GDP、第一产业增加值、第二产业增加值和常住人口数作为核心影响因素构建 PMGM(1, 4) 灰色马尔可夫模型，优化后的背景值插值系数为 $\lambda_i = (0.1982, 0.2163, 0.5808, 0.5473, 0.7159)$ ，最终得到芜湖市货运量 PMGM(1, 4) 灰色马尔可夫预测结果如表 11 所示，相对误差如表 12 所示。

表 11 芜湖市货运量不同模型预测值

Table 11 Forecast values of freight volume in Wuhu City using different models

年份	实际值/万 t	模型预测值/万 t				
		GM(1,1)	多元线性回归	MGM(1,4)	PMGM(1,4)	PMGM(1,4)灰色马尔可夫
2013	25461.24	25461.24	27203.77	25461.24	25461.24	25461.24
2014	31239.66	27491.61	29104.78	30472.99	31239.91	30401.97
2015	26202.40	27923.28	26438.73	25690.63	25913.51	26497.55
2016	24959.53	28361.72	26828.67	26676.00	26435.35	25597.40
2017	28449.82	28807.05	27548.48	28729.73	28011.34	28997.64
2018	29244.30	29259.38	28831.97	29987.05	29244.82	28721.64
2019	31628.52	29718.80	31015.51	30038.92	30019.61	30603.65
2020	29871.29	30185.44	28763.96	29466.51	30661.76	30138.58
2021	30461.98	30659.41	29907.16	29385.03	31589.39	30426.07
2022	31504.53	31140.82	33363.79	30847.39	33159.07	31995.75
2023	38747.50	31629.79	40386.06	34279.20	35600.52	36586.82
2024	38688.43	32126.44	49940.08	39199.08	38997.68	38474.50

表 12 芜湖市货运量不同模型预测相对误差

Table 12 Relative error of Wuhu City freight volume forecast by different models

模型	GM(1,1)	多元线性回归	MGM(1,4)	PMGM(1,4)	PMGM(1,4)灰色马尔可夫
拟合平均相对误差/%	4.24	4.00	2.68	2.52	1.59
预测平均相对误差/%	17.67	16.66	6.43	4.46	3.06

根据表 12 可知,GM(1,1)和多元线性回归预测模型虽然在拟合阶段表现良好,但在预测阶段的平均相对误差显著高于多变量灰色预测模型,表明了两种模型在预测精度上的不足;同时,对比 MGM(1,4)、PMGM(1,4)及 PMGM(1,4)灰色马尔可夫 3 种多变量灰色预测模型可知,通过对传统 MGM(1,4)模型进行背景值插值系数优化和残差修正后,得到的 PMGM(1,4)灰色马尔可夫模型在模型拟合与预测方面均表现出了最优的建模能力和预测性能,拟合和预测平均相对误差分别为 1.59% 和 3.06%。此外,根据图 4 可知,该模型预测曲线与实际货运量数据吻合最高,验证了模型在复杂系统建模中的显著优势。

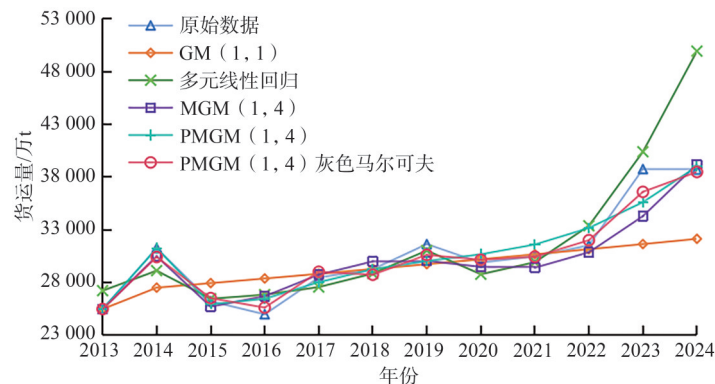


图 4 芜湖市货运量实测值与 5 种模型预测值对比

Fig. 4 Comparison between measured freight volume and predicted values of five models in Wuhu City

通过宁波市与芜湖市的实证分析可知,多参数模型的预测效果优于单参数模型,具备更强的结构适应性和预测能力;同时,多元线性回归预测模型虽在拟合阶段表现尚可,但对非线性特征与时变行为的刻画能力有限,在预测精度上远不及改进后的灰色模型。由此验证,本文构建的PMGM(1,N)灰色马尔可夫模型能更有效地捕捉城市货运系统中的耦合关系与随机扰动,在多变量小样本场景下具有良好的适应性和实用价值。

3 结论

本文通过融合灰色预测、粒子群算法与马尔可夫理论,提出了适用于城市货运量需求预测的PMGM(1,N)灰色马尔可夫模型,并以宁波市和芜湖市为研究对象开展实证分析,主要研究结论如下。

1) 建立了涵盖宏观经济运行、产业结构、消费能力、对外贸易、人口规模及基础设施投资等多维度的城市货运量影响因素体系,并结合灰色关联分析识别城市货运量的关键影响因子,从而保证了指标选择的科学性与针对性,为预测模型的有效构建奠定了基础。

2) 在MGM(1,N)模型的基础上,采用PSO算法优化背景值插值系数以提升模型拟合能力,并引入马尔可夫状态转移矩阵修正残差,从而显著增强模型对随机波动与突发变化的适应性。该改进策略突破了传统灰色模型在复杂城市系统预测中的局限,使模型在精度与稳健性方面均实现实质性提升,并在刻画城市货运需求的复杂性与不确定性方面表现出更优性能。

3) 宁波市和芜湖市的典型案例实证结果表明,PMGM(1,N)灰色马尔可夫模型在拟合精度和预测精度上均显著优于GM(1,1)、多元线性回归、MGM(1,N)和PMGM(1,N)方法,充分验证了本文模型指标体系构建与改进策略的科学性与有效性。同时,在满足变量间具有较强关联性、数据变化趋势相对平稳、数据具备连续性和时效性以及系统符合灰色建模特征等多个适用性条件下,本文模型可推广用于其他城市的货运量需求预测,具备良好的通用性与指导价值,可为城市物流规划、交通基础设施投资及区域运输管理提供科学的决策支持。

参考文献:

- [1] SULTANBEK M, ADILOVA N, ŚŁADKOWSKI A, et al. Forecasting the demand for railway freight transportation in Kazakhstan: A case study[J]. *Transportation Research Interdisciplinary Perspectives*, 2024, 23: 101028.
- [2] PENG T, GAN M, OU Q, et al. Railway cold chain freight demand forecasting with graph neural networks: A novel GraphARMA-GRU model[J]. *Expert Systems with Applications*, 2024, 255: 124693.
- [3] VAN M N, RASHED Y, STORMS K, et al. The future container throughput for inland shipping on the traditional Rhine: A SARIMAX approach[J]. *European Journal of Transport and Infrastructure Research*, 2022, 22(4): 25-50.
- [4] ZHANG Y S, TIAN C X, GUO B H, et al. Multi-factor highway freight volume prediction based on backpropagation neural network[J]. *Applied Sciences*, 2024, 14(13): 5948.
- [5] YANG C H, SHAO J C, LIU Y H, et al. Application of fuzzy-based support vector regression to forecast of international airport freight volumes[J]. *Mathematics*, 2022, 10(14): 2399.
- [6] 龙宇, 许浩然, 余华云, 等. 基于ARIMA-LSTM-XGBoost组合模型的铁路货运量预测[J]. *科学技术与工程*, 2023, 23(25): 10879-10886.
- [7] 柯桥, 邓萍. 基于改进灰色神经网络模型的三峡枢纽过坝货运量预测[J]. *上海海事大学学报*, 2021, 42(1): 82-87.
- [8] 钟蒙, 薛运强, 周珣, 等. 基于灰色-反向传播神经网络的江西省公路货运量预测[J]. *科学技术与工程*, 2021, 21(24): 10478-10484.
- [9] 谭雪, 张小强. 基于GRU深度网络的铁路短期货运量预测[J]. *铁道学报*, 2020, 42(12): 28-35.

- [10] 张昭俊,孙洁,王秀丽. 三峡枢纽过坝货运需求预测模型及其应用[J]. 内蒙古大学学报(自然科学版),2021,52(5):551-560.
- [11] LI Z H, LU J H. An improved MGM(1, n) model for predicting urban electricity consumption[J]. *Energies*, 2024, 17(16):3872.
- [12] SHAN X, CAO Y. Forecasting Guangdong's marine science and technology, marine economy, and employed persons by coastal regions-based on rolling grey MGM(1, m) model[J]. *Water*, 2022, 14(5):824.
- [13] 陈先飞,何山,王杰,等. 基于MGM(1, n, r)的主动配电网中长期负荷预测[J]. 太阳能学报, 2020, 41(2):188-193.
- [14] 朱文静,张映雪,武焱,等. 优化MGM(1, n)模型在边坡沉降预测中的应用研究[J]. 交通科学与工程, 2023, 39(2):105-113.
- [15] 李树,王丰效. 基于蚁群算法的多变量MGM(1, N)组合预测模型[J]. 数学的实践与认识, 2021, 51(14):41-47.
- [16] 胡涛,王栋,姜龙亭,等. 基于改进MGM(1, N)轨迹预测的空战态势评估[J]. 电光与控制, 2020, 27(11):39-44,96.
- [17] 姜晋云,李文海,陈天驰,等. 基于残差修正MGM(1, N)的轨迹模拟方法[J]. 电子测量技术, 2019, 42(17):83-87.

(责任编辑 李成兵)

Freight Volume Demand Forecasting Based on PMGM(1, N) Grey Markov Model

LIU Xinghui, REN Jianwei, LI Xueqi

(*Institute of Transportation, Inner Mongolia University, Hohhot 010070, China*)

Abstract: To improve the accuracy of urban freight demand forecasting, a PMGM(1, N) grey Markov prediction model was developed, with Ningbo City and Wuhu City as empirical study cases. First, a multi-factor indicator system was established, and grey relational analysis was employed to identify the core variables. Then, the particle swarm optimization (PSO) algorithm was applied to optimize the background value interpolation coefficient of the MGM(1, N) model, enhancing its fitting capability. Finally, a Markov state transition matrix was introduced to correct the residuals, thereby improving the model's adaptability to data fluctuations. The empirical results show that the PMGM(1, N) grey Markov model outperforms GM(1, 1), multiple linear regression, MGM(1, N), and PMGM(1, N) models in both fitting and forecasting accuracy, fully verifying its effectiveness and superiority in urban freight demand forecasting.

Key words: urban freight volume forecasting; grey prediction; PSO algorithm; Markov theory