文章编号：2097-5899(2025)09-0058-07　　**中图分类号**：U294.1　　**文献标识码**：A
DOI：10.16669/j.cnki.issn.2097-5899.202409030002
引用格式：江路路，高天泽. 基于组合模型的黄大铁路货运量预测研究[J]. 铁路物流，2025，43(9)：58-64.
JIANG Lulu，GAO Tianze. Research on Freight Volume Forecasting of Huanghua-Dajiawa Railway Based on a Combined Model[J]. Railway Logistics，2025，43(9)：58-64.

基于组合模型的黄大铁路货运量预测研究

江路路[1]，高天泽[2]

(1. 国能黄大铁路有限责任公司 运营管理中心，山东 东营 257091；2. 北京交通大学 交通运输学院，北京 100044)

摘　要：山东省对煤炭的需求主要依赖胶济铁路运输，但运能不足的弊端逐渐凸显，黄大铁路建设项目纵跨河北、山东两省，极大缓解了朔黄铁路煤炭疏运通道压力，是解决煤运能力不足问题的关键点。通过搜集黄大铁路近年的经济、社会、政策等历史数据，选择了与货运量相关度高的6项影响因素，运用主成分分析提取主要影响因子，结合回归模型、时间序列模型与神经网络模型，构建了综合权重的货运量组合预测模型，结合历年数据，完成了货运量预测及结果分析。相比传统单一预测模型，组合模型能够通过权重组合来缩小平均绝对误差，有效提高预测准确性，为铁路运输方案优化和运输组织提供科学依据。

关键词：黄大铁路；铁路货运；运量预测；组合模型；主成分分析

0　引言

为应对经济要素加速流动带来的铁路运力需求增加，国家推进了多项跨世纪重大铁路工程建设，特别强调提升铁路货物运输比例。黄大铁路(黄骅南—大家洼)作为朔黄铁路(神池南—黄骅港)的重要延伸线，是解决山东省煤炭运输瓶颈的关键工程。山东省作为能源消费大省，煤炭需求量巨大，但省内煤炭资源匮乏，90%以上依赖外省调入，主要通过胶济铁路(青岛—济南)运输，运能不足问题日益突出。黄大铁路全长216.8 km，北起朔黄铁路黄骅南站，南接益羊铁路(益都—大家洼)大家洼站，纵跨河北、山东两省，不仅缓解了朔黄铁路的疏运压力，更为山东省煤炭供应提供了新的运输通道。准确预测黄大铁路货运量对于优化运输组织、提高运营效率具有重要意义。

国内外学者在多种交通方式货运量预测的相关研究较为丰富。在时间序列模型方面，Brockwell等[1]认为ARIMA模型对于短期趋势预测的准确率高；燕学博等[2]研究发现ARIMA模型的预测精度较好且泛化能力较强；陆豪冬[3]运用时间序列模型进行铁路货运量预测，取得了理想的效果；郭洪鹏等[4]建立基于双向长短时记忆网络(Bi-LSTM)的短期货运量预测模型，验证了其在货运量预测方面的精确度与泛化能力。在回归模型方面，宋光平[5]通过比较发现，二次指数平滑模型更适合铁路货运量预测；贾春苗等[6]等选取铁路货运量影响因素，利用多元线性回归模型对甘宁地区铁路货运量的影响因素进行实证分析，得出甘宁地区铁路货运量未来5年的预测值，并分析其变化趋势。人工神经网络，尤其是BP神经网络也是铁路货运量预测的常用方法。BP神经网络包含输入层、隐藏层、输出层，其参数设置与数据格式会影响到整个模型的效果。薛锋等[7]根据铁路货运量不同影响因素，使用粒子群算法配合神经网络进行货运量预测，探究得出人均GDP是影响铁路客货运量的重要因素。组合模型预测方面，孟建军等[8]对常见的铁路货运量预测模型进行了对比，指出单一模型的局限性，凸显了组合预测模型发展的重

收稿日期：2024-09-03　**修订日期**：2025-08-10
基金项目：中央高校基本科研业务费专项资金项目(2025YJS065)

要性；韩纯良等[9]综合考虑GDP、铁路营业里程、原煤产量、钢材产量、第二产业增加值以及新冠疫情等因素对铁路货运量的影响，使用SVR组合模型进行货运量预测；张海山等[10]基于异质时空图注意力网络对铁路车站货运量进行预测，利用图注意力机制捕捉车站与其相邻车站之间的空间关联性，有助于提升货运量预测结果的精准度。

本研究通过比较不同的货运量预测模型，提出一种基于组合模型的方法来实现对黄大铁路近、中、远期货运量的预测，为未来的运输方案优化和运输组织提供参考依据。

1　黄大铁路建设背景与运营现状

黄大铁路北起朔黄铁路黄骅南站，纵跨河北、山东2省，经沧州、滨州、东营、潍坊4市11个区县(市)，南接益羊铁路大家洼车站，全长216.8 km，黄大铁路车站分布如图1所示。线路主要技术条件为国家Ⅰ级干线，单线电气化(预留双线条件)铁路，自动站间闭塞配合调度集中(CTC)指挥，限制坡度4‰，最小曲线半径一般地段1 200 km，困难地段800 m。黄骅南至东营西到发线有效长度为1 800 m，东营西至大家洼为1 050 m。由于山东省煤炭需求缺口较大、朔黄铁路煤炭疏运通道不足，因此将黄大铁路作为重要的运输通道，于2015年开工建设。

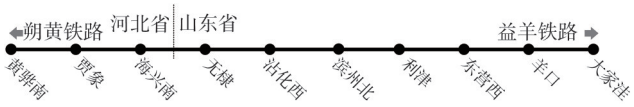


图1　黄大铁路车站分布

Fig.1　Distribution of Huangda Railway Station

目前，黄大铁路面临以下2方面问题。

(1) 山东省煤炭需求缺口大。山东省是经济和人口大省，能源需求随着经济发展逐年扩大，煤炭短缺最严重的是鲁西北和胶济铁路沿线地区。2017年，煤炭需求约为6 500万t，电厂发电需求3 500万t，省外调入5 800万t，其中90%来自晋中地区。因此，即使近年来山东省持续推进清洁煤炭改善能源供应结构，但庞大的煤炭需求在短期内依然难以改变。

(2) 朔黄铁路煤炭疏运通道不足。朔黄铁路是西煤东运的第二大通道，由于各地出台了严格的环保措施，煤炭需求逐年放缓，朔黄扩建新增的运输能力将超过下游需求，天津港和黄骅港在2014年出现存煤攀升问题，朔黄铁路下游的疏运通道略显不足。黄大铁路作为朔黄铁路煤炭运输的后方疏运通道，成为了国家煤炭运输网络的重要组成部分。

针对黄大铁路面临的煤炭需求缺口大、朔黄铁路疏运通道不足等局限性，提出货运量组合预测模型尤为必要。该模型能精准预测货运量，助力铁路部门对煤炭运输组织的科学规划与管理，确保煤炭等关键物资稳定供应。模型考虑多因素协同作用，缓解朔黄铁路运输压力，优化整体煤炭运输网络，为清洁能源转型提供数据支撑，促进可持续发展。面对运量波动等突发事件，模型能提前识别风险，制定应急预案，保障运输安全与稳定。

2　黄大铁路货运量组合预测模型

2.1　预测依据

黄大铁路作为新建线路，开展运量预测时，历史数据支撑不足。考虑到黄大铁路建设主要是为了解决山东省的煤炭运输问题，因此货运量预测分为煤炭和其他货物2类。参考《铁路线路设计规范》《铁路车站及枢纽设计规范》《国家发展改革委关于新建黄骅南至大家洼铁路项目核准的批复》(发改基础〔2013〕1659号)、《中国神华能源股份有限公司总裁常务会议纪要》关于黄大铁路的批复意见，《关于新建铁路黄大线初步设计评审的专家组意见》等文件，黄大铁路货运量预测的基年为2015年，近、中、远期分别为2025—2030年、2030—2035年、2035—2040年。

2.2　组合模型货运量预测技术路线

(1) 货运量调查。交通调查的任务是通过对规划区域社会经济系统、交通运输系统历史和现状的调查，为交通需求预测模型构建提供直接的依据。运输量调查是该阶段的核心工作之一。主要内容是调研吸引区内综合运输方式的OD运输量，尤其是道路网的基年OD运输量，以及铁路、公路、水运、航空和管道5种运输方式历年完成的货运量。

(2) 货运量影响因素分析。铁路货运量主要影响因素包括国家宏观经济因素、铁路所在区域市场需求、其他运输方式的竞争和其他因素等。国家宏观经济因素包括经济发展阶段、经济体制结构和经济政策等。铁路货物运输重点关注区域内的市场需求，一条铁路辐射范围内的企业数量和规模直接决定了该条铁路的货物类型和运输量。一条铁路的长度和地理辐射范围是固定的，该范围内企业数量越多，企业规模越大，货物的需求量就越大。研究聚焦的影响因素包括现有各种交通方式的货运量、国民经济发展指标、投资状况、人均收入水平等。

（3）组合模型的构建。组合模型的任务是将调研的货运量时间序列数据映射成未来数据，组合模型融入时间序列模型、回归模型与神经网络模型，在OD矩阵调查基础上，筛选关键因素进行预测分析，标定各预测模型的参数，在各模型预测结果基础上采用加权法预测货运量分布，货运量组合预测框架如图2所示。

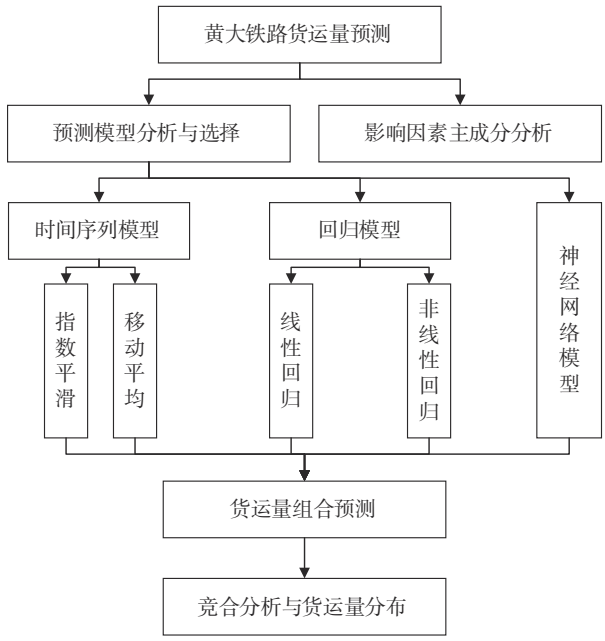


图2　货运量组合预测框架

Fig.2　Framework for freight volume composite forecasting

2.3　组合模型货运量预测方法

目前铁路货运量的预测模型主要有时间序列、回归分析、灰色预测、神经网络、支持向量机、分形理论、粗糙集理论等。基于选择回归方法、时间序列方法、神经网络模型，形成组合模型对黄大铁路货运量进行预测，以预测过程中预测值的绝对误差与相对误差均值最小为原则，求得各单一模型预测结果的权重值。

假设本次预测过程中共采用了$n(n\geqslant 2)$个不同的单一方法预测模型，为将其进行组合，首先按照实际观测值计算各单一方法预测模型结果，即$y_i(i=1,2,\cdots,n)$,各单一方法预测模型结果在组合预测模型中的相对权重值为$w_i(i=1,2,\cdots,n)$，各单一方法预测模型结果在组合预测模型中的绝对权重值为$m_i(i=1,2,\cdots,n)$。

$$y=\sum_{i=1}^{n}\left(y_i\times\frac{w_i+m_i}{2}\right) \tag{1}$$

$$w_i=\frac{1}{Se_i}\Big/\sum_{j=1}^{n}\frac{1}{Se_j} \tag{2}$$

$$e_i=\left|x_i-y_i\right| \tag{3}$$

$$m_i=\frac{1}{e_i}\Big/\sum_{j=1}^{n}\frac{1}{e_j} \tag{4}$$

式中：y表示铁路货运总量的预测值，y_1到y_n是第n种模型的预测值,万t；n为所采用预测模型的数量；w_i为各模型的预测相对权重值；Se_i为平均相对误差；x_i为当年度的实际观测取值；e_i为当年度的绝对误差计算值；m_i为各模型预测的绝对权重值。

3　黄大铁路货运量预测分析

根据已构建的铁路货运量组合预测模型所需数据要求，根据黄大铁路区域基年经济、社会、政策、资源、环境等历史数据，利用组合模型对货运量进行预测，得出黄大铁路货运量预测值，另考虑公路、铁路、水路、航空之间的竞合关系，确定黄大铁路货运量分担情况，进而确定货量分布，为经济评价奠定基础。

根据文献统计，黄大铁路货运量与煤炭供需量、地区生产总值(GDP)的相关系数较高，此外，人口、地区面积、固定资产投资等与货运量的关系系数可作为辅助参考。因此，主要提取与货运量相关的煤炭产量(x_1)、煤炭需求量(x_2)、地区生产总值(x_3)、人口(x_4)、地区面积(x_5)、固定资产投资总额(x_6)6项因素进行分析，其中前3项为与货运量高度相关的主要因素，后3项则为次要因素，起到辅助修正作用。

3.1　主成分分析

主成分分析(PCA)是一种统计方法，通过正交变换将一组可能存在相关性的变量转换为一组线性不相关的变量，转换后的这组变量称为主成分。PCA旨在利用降维的思想，把多指标转化为少数几个综合指标，以便于描述和解释数据。这种方法通过保留低阶主成分，忽略高阶主成分，从而减少数据集的维数，同时保持数据集对方差贡献最大的特征。以沧州市为例，采用主成分分析方法，主成分分析结果如表1所示。

提取的主成分为

$$f_1=0.551\,8x_1+0.465\,1x_2+0.431\,2x_3+0.211\,1x_4+0.307\,1x_5-0.384\,1x_6 \tag{5}$$

$$f_2=0.394\,7x_1+0.350\,9x_2+0.034\,4x_3+0.491\,4x_4-0.114\,7x_5+0.705\,0x_6 \tag{6}$$

式中：f_1表示由第一主成分(PC1)构成的线性组合，为数据中第一主要变化方向；f_2表示由第二主成分

表1　主成分分析结果

Tab.1　Results of principal component analysis

变量	PC1	PC2	PC3	PC4
x_1	0.551 8	0.394 7	−0.648 1	0.601 9
x_2	0.465 1	0.350 9	0.165 5	0.141 1
x_3	0.431 2	0.034 4	−0.124 1	−0.300 2
x_4	0.211 1	0.491 4	0.614 2	0.746 5
x_5	0.307 1	−0.114 7	−0.031 4	0.860 1
x_6	−0.384 1	0.705 0	0.384 7	−0.210 0

(PC2)构成的线性组合，为数据中第二主要变化方向。

PC1为所选取6个特征的线性组合，其解释数据中最大的方差。PC2是与PC1正交的另一组线性组合，其解释剩余数据中最大的方差。主成分PC1与PC2在PCA中起到了简化数据结构、提取关键特征、去除噪声以及增强可解释性的作用。

3.2　基于混合模型的货运量预测

3.2.1　基于回归模型的货运量预测

回归模型通过分析货运量与一个或多个影响因素之间的统计关系，建立数学模型来预测未来货运量，适用于当货运量受多个可量化因素影响时。黄大铁路主要服务于工业原材料运输，其运量受地区GDP、供需关系等宏观经济指标影响明显，适合采用多元回归进行建模。

应用3.1节提取的主成分f_1, f_2进行回归计算，回归模型为

$$y=b_0+b_1f_1+b_2f_2 \tag{7}$$

式中：y表示模型输出的铁路货运量预测值，万t；b_0回归常数项，表示当所有自变量（主成分f_1和f_2）均为0时，货运量的基准值；b_1表示第一主成分f_1的回归系数；b_2表示第二主成分f_2的回归系数；f_1表示第一主成分，根据式(5)得出；f_2表示第二主成分，根据式(6)得出。

线性模型回归系数如表2所示。回归模型预测结果如表3所示。

3.2.2　基于时间序列模型的货运量预测

时间序列模型基于以往数据的变化趋势来预测未来情况，其输入输出值均为预测变量，特别适用于当货运量数据呈现出明显的时间相关性和周期性时。例如，当黄大铁路货运量受季节性、周期性经济波动或节假日等因素影响时，时间序列模型能够较好地捕捉这些规律并用于预测，其中ARIMA模型擅长处理具

表2　线性模型回归系数

Tab.2　Linear model regression coefficients

需求地	滨州	沧州	东营	潍坊
常数	−3.410	7.143	7.141	2.86
f_1	1.113	0.615	−0.304	1.291
f_2	1.980	2.841	5.148	−0.412
N	10	10	10	10
R^2	0.931 4	0.862 5	0.907 1	0.922 4
F	59.15	13.40	43.28	81.39
$RMSE$	0.122	0.157	0.138	0.119

表3　回归模型预测结果　　万t

Tab.3　Regression model forecasting results

年份	2025	2030	2035	2040
铁路货运量	1 412	2 158	2 607	4 174

有趋势性、季节性的时间序列数据，可以有效捕捉货运量的发展规律，而双重指数平滑模型能够依据货运量季节性与趋势特征，进行中长期具有线性趋势数据的预测，能够良好地刻画数据的稳定增长与下降趋势。

（1）自回归综合移动平均(ARIMA)模型。为了检验时间序列数据的平稳性，对货运量数据进行了增强型Dickey-Fuller(ADF)检验。增强型Dickey-Fuller检验结果如表4所示。原假设为差分后的时间序列$D(DY1)$存在单位根，即该序列是非平稳的。检验中采用了基于Schwarz信息准则(SIC)自动选择的1阶滞后长度。由表4可见，ADF检验统计量为−3.733。在5%的显著性水平下，临界值为−3.403。由于−3.733小于−3.403，表明在5%的显著性水平下可拒绝原假设，即差分后的时间序列$D(DY1)$不存在单位根。此外，p值为0.033 8，小于0.05，同样支持拒绝原假设的结论。通过增强型Dickey-Fuller检验，差分后的时间序列$D(DY1)$是平稳的，不存在单位根。

一阶差分以后平稳，产生新的序列y，模型$ARIMA(p,d,q)$中$d=1$。经过对不同p和q值的拟合，选择$p=2$，$q=0$，时间序列模型预测结果如表5所示，其$RMSE$值为0.192，MAE为0.15，$MeanAPE$为139.07。

（2）双重指数平滑模型。双重指数平滑模型的结构如下。

$$L_t=\alpha Y_t+(1-\alpha)(L_{t-1}+T_{t-1}) \tag{8}$$

表4　增强型Dickey-Fuller检验结果

Tab.4　Augmented Dickey-Fuller test results

统计量类型	数值
ADF检验统计量	−3.733
1%显著性水平临界值	−4.803
5%显著性水平临界值	−3.403
10%显著性水平临界值	−2.842
p值	0.033 8

$$T_t=\beta(L_t-L_{t-1})+(1-\beta)T_{t-1} \tag{9}$$

$$\hat{Y}_{t+h}=L_t+hT_t \tag{10}$$

式中：Y_t为第t时期的实际货运量，万t；L_t为第t时期的水平分量，表示数据当前的水平；T_t为第t时期的趋势分量，表示数据的增长或下降趋势；$\hat{Y}_{t+h}$表示在第t时期对第$t+h$时期的预测值；α为水平平滑系数，控制当前观测值对水平分量的影响程度；β为趋势平滑系数，控制趋势变化的平滑程度；h为预测步长，表示向前预测的期数。

对双重指数平滑模型中的参数进行拟合，拟合的水平平滑系数α为0.214 5，趋势平滑系数β为0.130 7。构建的水平方程与趋势方程为

$$L_t=0.214\,5Y_t+0.785\,5(L_{t-1}+T_{t-1}) \tag{11}$$

$$T_t=0.130\,7(L_t-L_{t-1})+0.869\,3T_{t-1} \tag{12}$$

式中：Y_t为第t时期的实际货运量，万t；L_t为第t时期的水平分量，表示数据当前的水平；T_t为第t时期的趋势分量，表示数据的增长或下降趋势。

以4 a为预测步长，将对应年份的Y_t，T_t，L_t代入公式(11)—(12)，预测结果$RMSE$值为0.148，MAE为0.132，$MeanAPE$为4.305，根据时间参数，时间序列模型预测结果如表5所示。

表5　时间序列模型预测结果　　万t

Tab.5　Time series model forecasting results

模型	2025年	2030年	2035年	2040年
ARIMA	1 716	2 289	2 603	3 492
双重指数平滑	1 924	2 407	2 583	3 015

3.2.3　基于神经网络的货运量预测

神经网络模型特别适用于处理复杂、非线性且难以用传统统计方法描述的关系。当黄大铁路货运量受众多难以量化的模糊因素影响时，如供需关系、政策导向等，神经网络模型能够通过学习和训练，自动捕捉这些模糊因素与货运量之间的复杂关系，并进行预测。

应用R软件包进行铁路货运量BP神经网络预测。应用命令library(neuralnet)加载程序包；读入吸引区2005—2023年人口、地区面积、国内生产总值、固定资产投资总额等数据，将数据规范化到[0，1]，模型时间窗口设置为2005—2040年，铁路货运量规范后的最小值设为0，最大值设置为20 000万t。

采用BP神经网络作为铁路运量的预测模型，神经网络连接权值如图3所示。相应的$RMSE$值为0.845，MAE为0.16，$MeanAPE$为23.135。神经网络模型预测结果如表6所示。

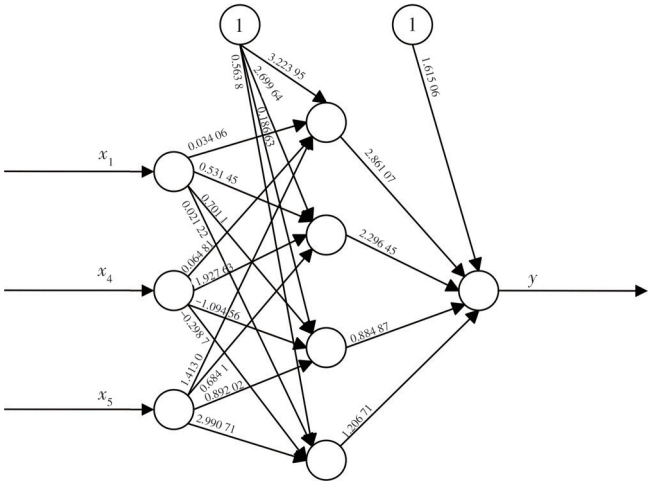


图3　神经网络连接权值

Fig.3　Neural network connection weights

表6　神经网络模型预测结果　　万t

Tab.6　Neural network model forecasting results

年份	2025	2030	2035	2040
铁路货运量	2 858	3 097	3 594	4 250

3.3　组合模型货运量预测结果分析

将上述的回归模型、时间序列模型和人工神经网络模型预测结果汇总，按照权重计算公式求得绝对权重。模型预测结果汇总如表7所示。

表7　模型预测结果汇总　　万t

Tab.7　Summary of model forecasting results

模型	2025年	2030年	2035年	2040年	MAE	绝对权重
回归模型	1 412	2 158	2 607	4 174	0.10	0.28
ARIMA	1 716	2 289	2 603	3 492	0.15	0.18
双重指数平滑	1 924	2 407	2 583	3 015	0.13	0.21
神经网络模型	2 858	3 097	3 594	4 250	0.06	0.33
组合运量预测值	2 051.42	2 543.74	2 926.95	3 832.93	—	

根据表7的预测结果可以看出，2025年黄大铁路货运量预测结果为2 051.42 万t，2030年为2 543.74 万t，2035年为2 926.95 万t，2040年为3 832.93 万t。

从数据集与整体趋势分析可知，2025—2040年的整体预测结果呈现持续增长的趋势，且整体增速呈现快速增长再过渡到平稳增长的趋势。神经网络模型预测值增速相对激进，货运量从2025年的2 858 万t增长至2040年的4 250 万t，15年间增长约0.49倍，呈现较为平缓的增长态势；回归模型和ARIMA模型的预测结果相对接近，增长趋势较为一致，货运量分别从2025年的1 412 万t和1 716 万t增长至2040年的4 174万t和3 492 万t；双重指数平滑模型的预测增长趋势则呈现由急至缓的变化趋势，2025年与2030年的预测结果相较回归模型与ARIMA模型偏高，而长期预测结果则偏低，这是由于模型预测侧重点为中期所致。从模型预测精度MAE指标对比分析，神经网络模型的预测精度最高，误差率仅为0.06；回归模型和双重指数平滑模型的*MAE*分别为0.10和0.13；ARIMA模型的*MAE*为0.15。从综合趋势判断分析，近期预测方面(2025—2030年)各模型预测结果较为统一，差异相对较小，预测可靠性较高；中期预测(2030—2035年)开始出现分化，但整体仍保持上升趋势；远期预测(2035—2040年)各模型差异明显加大，回归模型与神经网络模型所得预测结果高于其他2个预测模型，这提示在使用远期预测结果时需要更加谨慎，需结合多因素进行综合分析，使用其他指标进行修正与优化。

从整体预测结果的增速变化情况来看，2025年度的货运量预测增速在30%左右，符合黄大铁路开行后货运量迅速增长情况；中远期预测结果显示，黄大铁路煤炭运量呈现较为平稳的增长速度和效率，预测结果符合未来发展趋势。

4　结束语

针对黄大铁路货运量预测问题，通过主成分分析法对影响货运量的6项因素进行降维处理，提取主要影响因子，构建了基于回归模型、时间序列模型和神经网络模型的综合权重组合预测模型。实证结果表明，组合模型通过权重优化有效降低了平均绝对误差，相比单一预测模型具有更高的预测精度和稳定性。预测结果显示，黄大铁路货运量在短期内将保持稳定增长，中长期受煤炭需求和区域经济发展影响呈现波动上升趋势，这为铁路运输部门制定运营策略、优化运力配置提供了科学依据。

参考文献：

[1] BROCKWELL P J，DAVIS R A. Time Series：Theory and Methods[M]. New York：Springer New York，1991.

[2] 燕学博，曹　雨. 基于多模型的铁路货运量预测对比[J]. 物流科技，2023，46(21)：74-78.
YAN Xuebo，CAO Yu. Comparison of Railway Freight Volume Forecast Based on Multiple Models[J]. Logistics Sci-Tech，2023，46(21)：74-78.

[3] 陆豪冬. 基于时间序列数据的货物运输方式与CO_2排放关系研究[J]. 物流工程与管理，2023，45(5)：86-91，99.
LU Haodong. Study on the Relationship between Freight Transport Mode and CO_2 Emission Based on Time Series Data[J]. Logistics Engineering and Management，2023，45(5)：86-91，99.

[4] 郭洪鹏，刘　斌，肖　尧. 基于Bi-LSTM网络的铁路短期货运量预测研究[J]. 铁道货运，2022，40(2)：52-58.
GUO Hongpeng，LIU Bin，XIAO Yao. Short-Term Railway Freight Volume Prediction Based on Bi-LSTM Network[J]. Railway Freight Transport，2022，40(2)：52-58.

[5] 宋光平. 铁路货运量预测方法研究[D]. 北京：北京交通大学，2007.

[6] 贾春苗，傅忠宁，马亚玲，等. 基于ARMA模型与多元回归的甘宁区铁路货运量预测[J]. 综合运输，2022，44(9)：147-154.
JIA Chunmiao，FU Zhongning，MA Yaling，et al. Freight Volume Forecast of Ganning District Railway Based on ARMA Model and Multiple Regression[J]. China Transportation Review，2022，44(9)：147-154.

[7] 薛　锋，吴林鸿，汪雯文，等. 基于MI-PSO-RBF神经网络的铁路客货运量预测研究[J]. 铁道运输与经济，2024，46(9)：123-135.
XUE Feng，WU Linhong，WANG Wenwen，et al. Railway Passenger and Freight Volume Prediction Based on MI-PSO-RBF Neural Network[J]. Railway Transport and Economy，2024，46(9)：123-135.

[8] 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.
MENG Jianjun，CHEN Pengfang，LI Decang，et al. Summary of Research on Railway Freight Volume Forecast[J]. Railway Standard Design，2022，66(10)：18-26.

[9] 韩纯良，李默涵，洒　雨，等. 基于EMD-APSO-SVR组合模型的铁路货运量预测[J]. 综合运输，2024，46(4)：132-140，180.
HAN Chunliang，LI Mohan，SA Yu，et al. Rail Freight Volume Forecasting Based on Combined EMD-APSO-SVR Model[J]. China Transportation Review，2024，46(4)：132-140，180.

[10] 张海山，王文斌，周　瑾. 基于异质时空图注意力网络的铁路车站货运量预测[J]. 铁道货运，2024，42(6)：52-59.
ZHANG Haishan，WANG Wenbin，ZHOU Jin. Freight Volume Prediction of Railway Station Based on Heterogeneous Spatial-Temporal Graph Attention Network[J]. Railway Freight Transport，2024，42(6)：52-59.

（责任编辑：金　晶）

Research on Freight Volume Forecasting of Huanghua-Dajiawa Railway Based on a Combined Model

JIANG Lulu[1], GAO Tianze[2]

(1. Operation Management Center, CHN Energy Huangda Railway Co., Ltd., Dongying 257091, Shandong, China; 2. School of Traffic and Transportation, Beijing Jiaotong University, Beijing 100044, China)

Abstract: Coal demand in Shandong Province primarily relies on Qingdao-Ji'nan Railway transportation. However, inadequate transportation capacity has become increasingly apparent. Construction of the Huanghua-Dajiawa Railway, spanning Hebei and Shandong provinces, significantly alleviates pressure on coal transportation channels along the Shuozhou-Huanghua Railway and provides a crucial solution to insufficient coal transportation capacity. This study collected historical economic, social, and policy data related to the Huanghua-Dajiawa Railway, selected six highly relevant factors influencing freight volume, and used principal component analysis to extract the main influencing factors. By integrating regression models, time series models, and neural network models, a comprehensive weighted freight volume forecasting model was constructed. Combined with historical data, the study completed freight volume prediction and result analysis. Compared to traditional single forecasting models, the composite model reduces mean absolute error through weight combination, effectively improving prediction accuracy and providing a scientific basis for railway transportation plan optimization and organization.

Keywords: Huanghua-Dajiawa Railway; Railway Freight; Freight Volume Forecasting; Combined Model; Principal Component Analysis

(上接第57页)

Research on Fault Emergency Response of Marshalling Station Control System Based on Knowledge Graph

LI Taoyu[1,2], ZHANG Zhenhai[1]

(1. School of Automation and Electrical Engineering, Lanzhou Jiaotong University, Lanzhou 730070, Gansu, China; 2. Wuhanbei Station, China Railway Wuhan Group Co., Ltd., Wuhan 430012, Hubei, China)

Abstract: Taking the emergency response of the railway marshalling station control system as the background, this paper constructed a knowledge graph for the safety emergency response of the marshalling station control system. In the work of obtaining the entities and relationships of the safety emergency response of the marshalling station control system, for unstructured data, a technology based on regular expressions and natural language processing was adopted to extract entities and relationships from specific format .docx files. In the implementation of the knowledge graph for the safety emergency response of the marshalling station control system, a knowledge graph for the safety emergency response of the marshalling station control system was constructed based on Neo4j, achieving the visual expression of the fault data set during the operation of the marshalling station control system. The Cypher language was used to query the emergency handling process when a fault occurs. The research results can provide a reference for the emergency response of the marshalling station control system under fault conditions.

Keywords: Knowledge Graph; Marshalling Station Control System; Knowledge Storage; Neo4j; Auxiliary Decision