

文章编号:2095-2716(2018)03-0120-09

ARMA-BP 物流需求预测模型及应用

刘凤春¹, 赵亚宁², 董新雁², 刘源铄², 乔鹏², 谢志远², 王立亚², 张春英²

(1. 华北理工大学 迁安学院, 河北 唐山 063000; 2. 华北理工大学 理学院, 河北 唐山 063210)

关键词: 时间序列; 神经网络; 残差序列; 物流需求预测

摘要: 基于物流需求的实时性和不确定性, 提出融合时间序列自回归-滑动平均模型 ARMA 和 BP 神经网络, 构建了物流需求预测 ARMA-BP 模型, 提出预测货物运输的 ARMA-BP 结合预测算法。以唐山市近几年物流运输数据为研究对象, 分别运用 ARMA 模型、BP 模型和 ARMA-BP 模型对物流数据进行预测分析, 结果表明, 与传统预测模型相比, ARMA-BP 模型预测精度更高, 具有一定的实用价值。

中图分类号: TP391 **文献标识码:** A

随着电子商务的进一步发展, 预测货运量并高效地将货物配送至客户手中成了各大电商企业的研究热点^[1]。对物流需求的预测大多采用时间序列、粗糙神经网络或者改进神经网络等单一模型。文献^[2]将时间序列的 ARMA 模型应用在物流需求的预测中; 文献^[3]提出广义自回归条件异方差模型进行拟合, 预测效果较自回归模型稍好; 文献^[4]构建了基于 3 层 BP 人工神经网络的物流需求规模预测模型。但单一模型存在着预测结果不稳定的弊端, 为此, 文献^[5]提出了结合时间序列法、灰色预测法以及相关性分析法的组合预测方法, 应用在码头吞吐量的预测中, 解决了以往预测过程中因素复杂且可用数据较少的难题; 施泽军等综合了灰色模型和指数平滑法, 提出了新的组合预测方法^[6]; Ling Zhou 等根据选择方法的适用性和区分性原则, 选择指数平滑法、ARIMA 法和神经网络法, 形成组合方法, 研究了一种 LTL 载体的短期需求预测^[7], 证明了时间序列和神经网络组合预测方法的优越性。

该项目采用 ARMA-BP 结合的方法, 对于货运量数据的单一线性序列, 应用 ARMA 模型方法, 而对于多个变量的非线性序列, 引入 BP 人工神经网络方法, 研究建立一种通用预测模型, 通过 2 种方法进行结合的预测模型, 最终的预测误差进一步减小。

1 基本知识

1.1 时间序列自回归-滑动平均模型 ARMA

根据时间序列的特征, 可将其分为平稳序列与非平稳序列。平稳时间序列模型包括自回归模型(AR)、滑动平均模型(MA)和自回归-滑动平均模型(ARMA)3 类^[8]。

AR 模型原理如式(1):

$$y_t = \mu + \frac{1}{1 - AR(1) * BAR(p) * B^p} \epsilon_t \quad (1)$$

MA 模型原理如式(2):

$$y_t = \mu + (1 - MA(1) * BMA(q) * B^q) \epsilon_t \quad (2)$$

收稿日期: 2018-03-24 修回日期: 2018-05-27

基金项目: 河北省自然科学基金资助(F2016209344, F2018209374)。

ARMA 模型原理如式(3):

$$y_t = \mu + \frac{1 - MA(1) * BMA(q) * B^q}{1 - AR(1) * BAR(p) * B^p} \epsilon_t \tag{3}$$

其中, B 为滞后算子, $B y_t = y_{t-1}$, p 、 q 为模型的阶数, y_t 为时间序列当前值; ϵ_t 为随机干扰, μ 为常数项^[9]。

而对于非平稳时间序列,一般先进行差分处理,使其变为平稳时间序列^[10],再运用平稳时间序列模型方法进行处理。为此引入有序差分算子 $\nabla = 1 - B$,且 $\nabla^d = (1 - B)^d$,应用一阶有序差分,对非平稳时间序列进行变换^[9],得式(4):

$$\nabla y_t = (1 - B) y_t = y_t - y_{t-1} \tag{4}$$

此差分变换需要不断重复,直至成为平稳序列。经过 d 阶差分后的序列满足式(5):

$$\nabla^d y_t = (1 - B)^d y_t \tag{5}$$

如何选择模型以及确定阶数,需要根据自相关系数和偏相关系数共同决策,一般可按表 1 来确定。

表 1 模型类型及阶数确定

自相关系数	偏相关系数	模型定阶
拖尾	P 阶截尾	AR(p)模型
q 阶截尾	拖尾	MA(q)模型
拖尾	拖尾	ARMA(p, q)模型

1.2 BP 神经网络基本原理

BP 神经网络运用多层前向神经网络学习获知系统差分方程中的非线性函数,采用梯度搜索技术,使得网络输出误差与网络期望输出的均方值达到最小。BP 神经网络结构如图 1 所示,由输入层、隐含层及输出层构成^[12]。

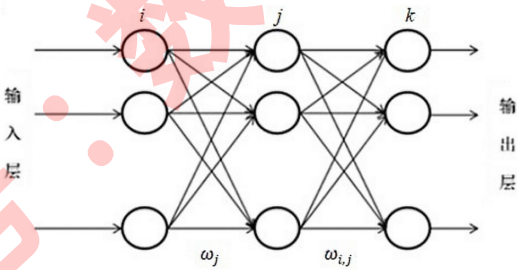


图 1 BP 神经网络结构

输入与输出间的关系如式(6):

$$x_t = \omega_0 + \sum_{j=1}^q \omega_j g(\omega_{0,j} + \sum_{i=1}^p \omega_{i,j} x_{t-i}) e_t \tag{6}$$

其中:

$\omega_{i,j}$ ($i = 0, 1, 2, \dots, p, j = 1, 2, \dots, q$) 和 ω_j ($j = 0, 1, 2, \dots, q$) 表示连接权重; p 为输入层节点的个数, q 为隐含层节点的个数。

2 ARMA-BP 模型

2.1 组合方法

在数据拟合过程中,不同年份的货运量数据看成一个时间序列,分为线性部分和非线性部分,采用 AR-

MA 方法拟合线性部分,并得到残差序列,然后应用 BP 神经网络处理非线性的残差部分^[13],具体方法如下:

第 1 步,用 ARMA 方法拟合时间序列并计算残差

$$e_t = \hat{L}_t - y_t \quad (7)$$

式(7)中, \hat{L}_t 为 ARMA 方法的预测值, y_t 为实际值。

第 2 步,对第一步计算出的 e_t ,使用 BP 神经网络方法进行一系列的训练学习,预测出非线性残差部分。

2.2 输入指标与输出指标的选取

物流需求与多种因素有关,选择出了下列因素,然后进行指标综合及定量筛选,最终确定 6 个因素,具体见图 2。

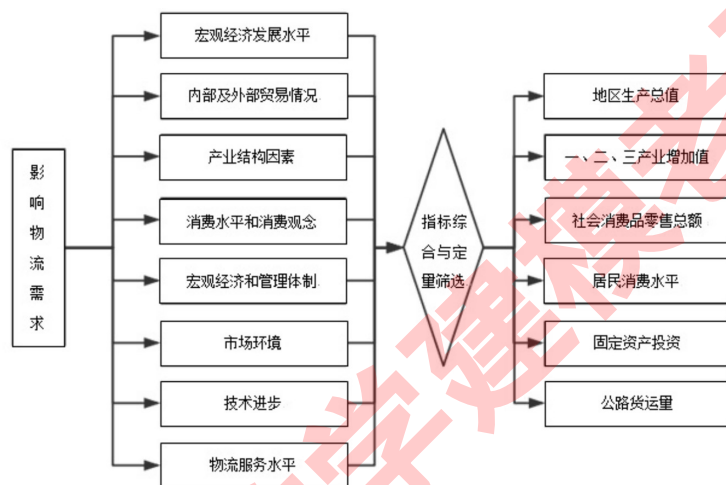


图 2 输入指标的确定

选择公路货运量作为物流需求量模型的输出指标。在货运量这个输入指标上,取用时间序列处理得到货运量的残差值作为模型的输入。

2.3 算法设计与分析

图 3 所示为过程进行时间序列处理,主要包括平稳性检验和处理以及白噪声检验等主要过程。

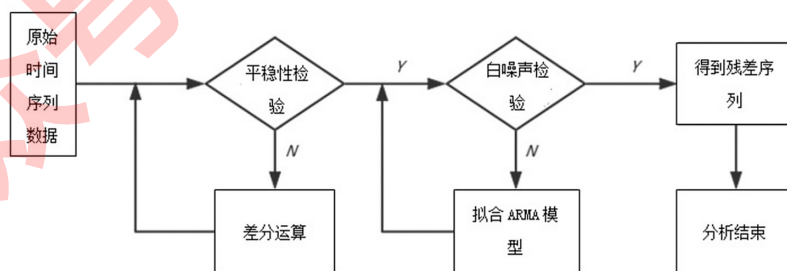


图 3 时间序列处理流程

(1) 平稳性检验

检验方法有时序图检验,自相关图检验 ADF 检验等,结合物流数据的特点,应用时序图检验最为方便快捷。平稳序列的时序图波动范围有显著的界限、没有明显的走势变化及周期性。

(2) 数据平稳化处理

应用差分原理进行 n 阶差分,差分后得到的序列进行步骤 1 的平稳性检验。

(3)模型阶数的确定

得到经过步骤 2 处理后序列的自相关与偏自相关系数图,通过图像及表 1 方法初步确定模型类型及阶数。

(4)白噪声检验

根据残差序列自相关系数与偏自相关系数的数学特征判定残差序列是否为白噪声序列,若通过白噪声检验则确定模型,得到预测值与残差序列的值,分析结束^[11]。

之后采用 BP-ARMA 模型处理对数据进行仿真预测,如图 4 所示,主要步骤如下。

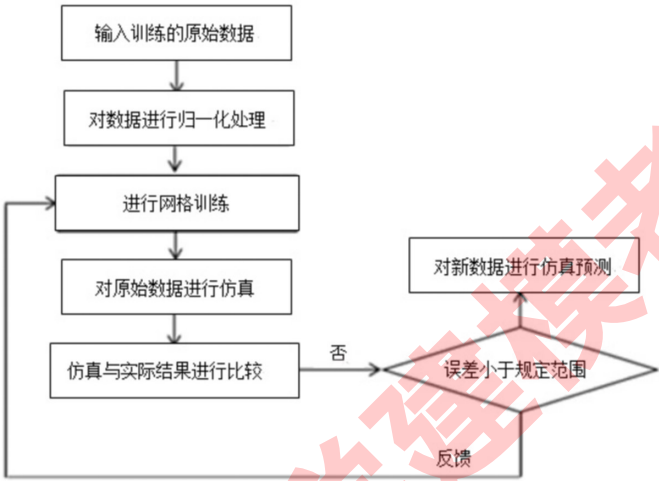


图 4 BP-ARMA 处理流程

(1)选取时间序列处理得到的残差序列和已确认好的指标作为训练神经网络的原始数据。

(2)选择 min-max 归一化函数对数据进行标准化处理,经过处理后的原始数据,各指标的数量级处于同一水平。

(3)进行第 1 次网络训练,得到一个训练结果,与实际值比较,若误差值大于规定值,则反馈调节权值和阈值,继续训练,直到误差值小到一定范围。

(4)网络训练结束,输入预测数据进行仿真预测,从而得到预测结果。

3 实验分析

3.1 实验环境

获取唐山统计年鉴数据,采用 1978~2015 年唐山市公路货运量数据作为原始数据,原始数据如表 2 所示,应用时间序列法对数据进行初步建模。

表 2 1978~2015 年唐山市公路货运量

年份	货运量	年份	货运量
1978	2 444	2011	33 005
1979	2 692	2012	37 625
1980	2 228	2013	43 043
1981	1 885	2014	33 479
.....	2015	36 358

在模型的建立与预测时,时间序列的处理及残差序列的获取利用 Eviews 9.0 数据处理软件来完成。BP 神经网络利用 python 语言搭建模型并进行分析预测。

3.2 ARMA 实验

将货运量数据导入软件中,命名为 x 序列,绘制序列 x 的趋势图及相关图如图 5 和图 6,观察后可以发现序列明显的上升趋势,且在相关图中观察 P 值,发现样本数据是非平稳的,故对原序列进行一阶差分。

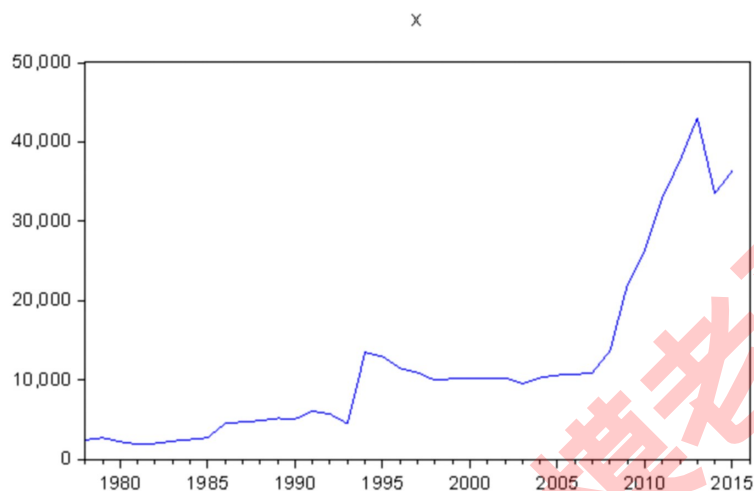


图 5 样本数据趋势图

Date: 12/08/17 Time: 17:40
Sample: 1978 2016
Included observations: 38

Autocorrelation	Partial Correlation	AC	PAC	Q-Sta...	Prob
		1 0.884	0.884	32.091	0.000
		2 0.772	-0.04...	57.262	0.000
		3 0.591	-0.38...	72.424	0.000
		4 0.429	-0.04...	80.645	0.000
		5 0.285	0.080	84.373	0.000
		6 0.184	0.080	85.985	0.000
		7 0.108	-0.03...	86.556	0.000
		8 0.074	0.044	86.837	0.000
		9 0.056	0.013	87.001	0.000
		1... 0.041	-0.08...	87.091	0.000
		1... 0.028	-0.02...	87.134	0.000
		1... 0.016	0.038	87.149	0.000
		1... 0.011	0.053	87.156	0.000
		1... 0.009	-0.00...	87.162	0.000
		1... 0.007	-0.04...	87.165	0.000
		1... -0.01...	-0.08...	87.172	0.000

图 6 样本数据相关图

应用差分功能,输入 $\text{genr } x_t = d(x)$ 指令,得到新序列 x_t ,绘制新序列 x_t 的趋势图,通过图 7 结合时序图检验方法,判断 x_t 为平稳序列。

得到序列 x_t 的相关图,如图 8 所示。由图 8 可以看出,自相关系数和偏自相关系数始终在 2 倍标准差范围内,所以将 $\{x_t\}$ 识别为 $\text{ARMA}(p, q)$ 模型。

进行多次实验,结合 AIC(赤池准则)最小化原理,比较每次的 n, m 值,最后确定结果为 $\text{ARMA}(20, 2)$ 模型。绘制预测值与实际值的趋势图,拟合结果如图 9 所示。

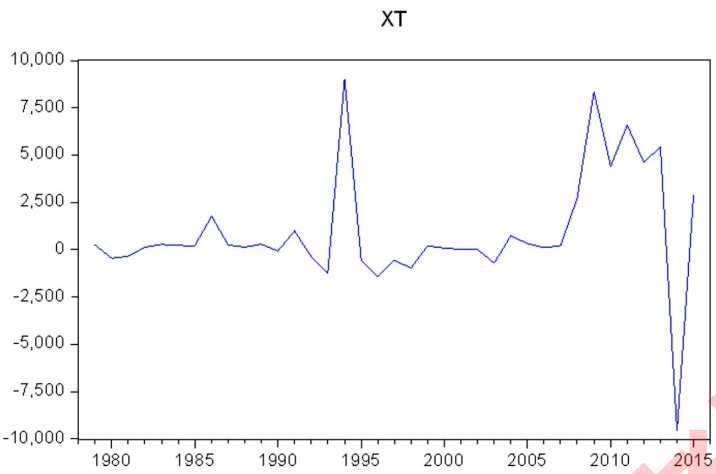


图 7 一阶差分序列的趋势图

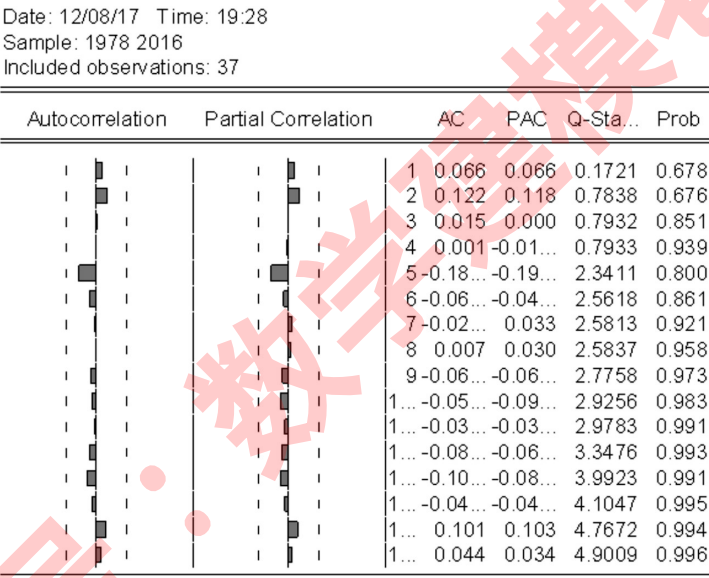


图 8 序列 x_t 相关图

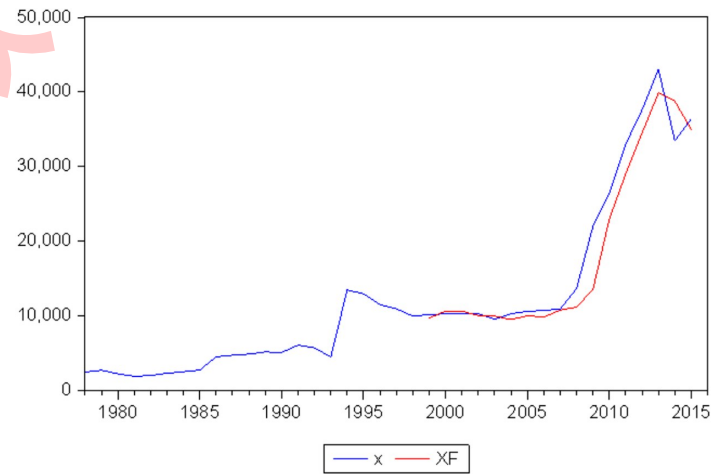


图 9 预测值与实际值趋势图

图中 XF 为预测值序列, X 为实际值序列。从图 9 可以看出, 两者的趋势图基本吻合, 但数据结果差别较大。

3.3 ARMA-BP 实验

输入 2001~2014 年的输入指标; 选择输入数据后一年的输出指标作为学习模型的输出。为使预测结果更为准确, 对输入数据进行数次单独训练, 训练过程中误差变化见图 10。

同时单独运用 BP 神经网络, 为误差的对比分析提供参考。3 种方法预测值与实际值对比见图 11。

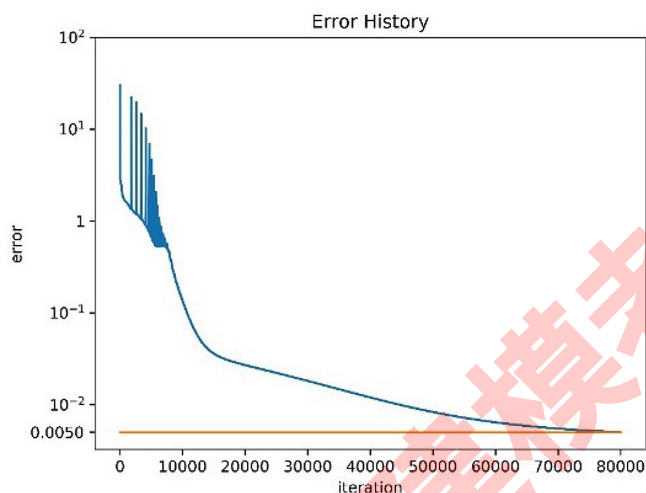


图 10 BP-ARMA 模型训练误差图

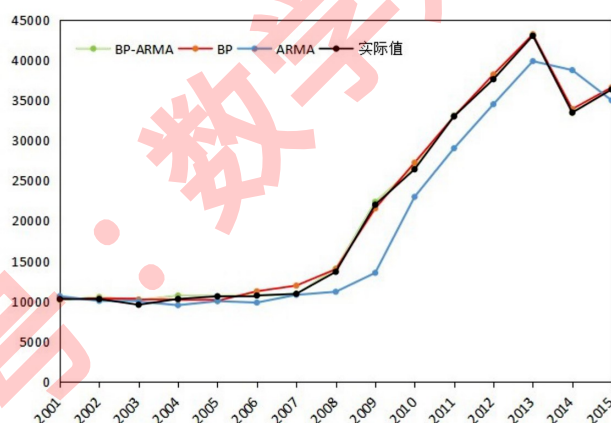


图 11 预测值与实际值对比图

3.4 结果对比分析

从图 11 中可以看出, 时间序列与神经网络结合的方法的预测值与真实值相吻合, 预测精度显著好于单一时间序列模型, 误差值达到 0.005, 在与单一的 BP 模型对比分析时, 仅通过趋势图很难看出优越性, 故绘制单一 BP 模型的误差图^[14], 如图 12 所示。

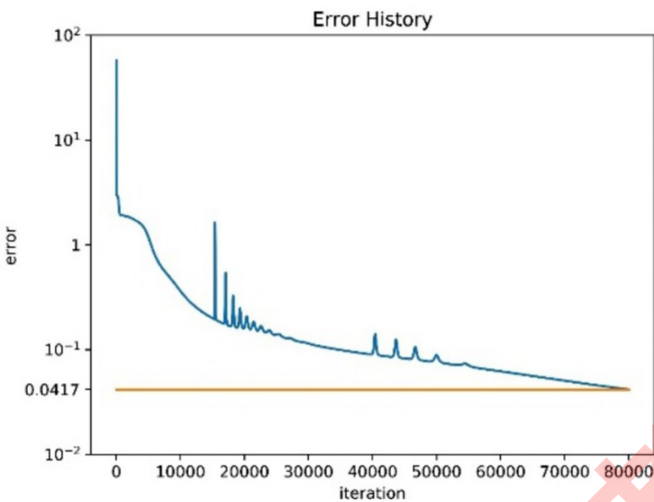


图 12 BP 模型训练误差图

结合图 12 与图 10,可以看出 BP-ARMA 模型的误差小于单一 BP 模型,预测精度进一步提高。

综上所述,ARMA-BP 结合的模型较单一模型方法预测精度有所提高,建立的模型是较为精准可靠的。最后,基于此模型,预测 2016 年货运量,预测结果汇总到表格中,如表 3 所示。

表 3 货运量真实值与预测值对比

年份	真实值	预测值	年份	真实值	预测值
2011	33 005	33312.10	2014	33479	33 439.75
2012	37 625	37688.73	2015	36358	36 437.77
2013	43 043	43221.51	2016	-	31 279.28

4 结论与展望

(1)将时间序列与 BP 神经网络结合,综合时间序列与 BP 神经网络的特点,建立时间序列-神经网络组合的方法,得出更加可靠的 ARMA-BP 模型,给电商企业足够的时间进行货物调整,增强电商企业的优势。

(2)此模型较之前的模型相比准确率进一步提高,但在预测未来若干时刻的物流需求时,目前多次运用一步预测的方法会误差积累,直接多步预测的效果也不理想。所以未来的研究方向应着手于神经网络多步预测方法的改进方面,从而使其满足当下及未来市场对物流需求预测的要求。

(3)在实际应用过程中也不能教条地认为长期趋势稳定不变,比如上游卖家的运营策略发生改变,增加了双十一这样的活动,很大程度上影响了趋势,所以经营者应结合实际情况,掌握最新的数据,对预测方程及时进行修正,并确定最佳定货量,从而最大化的获得收益。

参考文献:

- [1] 李壮阔,吕恒. 基于改进的 PSO-BP 神经网络的桂林市物流预测研究[J]. 大众科技, 2017,19(06):5-8.
- [2] 赵彦军,陈玉. 时间序列分析方法在物流需求预测中的应用[J]. 物流科技, 2017,40(06):12-14.
- [3] Bahram Adrangi, Arjun Chatrath, Kambiz Raei. The demand for US air transport service; a chaos and nonlinearity investigation[J]. Transportation Research Part E, 2001(37),337-353.
- [4] Jian Feng Luo, Tian Shan Ma. An Integrated Predicting Model of New-Built Regional Logistics Center's Demand Based on the Artificial Neural Network[J]. Applied Mechanics and Materials,2013,2155(253).
- [5] 莫艳芳,赵一飞.基于组合预测模型的滚装码头吞吐量预测[J].重庆交通大学学报:社会科学版,2014,14(2): 25-30.
- [6] 施泽军,李凯.基于灰色模型和指数平滑法的集装箱吞吐量预测[J]. 重庆交通大学学报:自然科学版,2008,27(2):302-304+332.

- [7] Ling Zhou. Prediction of a service demand using combined forecasting approach[J]. Journal of Physics: Conference Series, 2017, 887(1).
- [8] 田根平, 曾应昆. 基于时间序列模型在物流需求预测中的应用[J]. 物流科技, 2007, 30(9): 96-99.
- [9] 蔡凯, 谭伦农, 李春林, 等. 时间序列与神经网络法相结合的短期风速预测[J]. 电网技术, 2008, (08): 82-85+90.
- [10] 杨蕾, 张苗苗. 时间序列模型在物流需求预测中的应用[J]. 商业时代, 2013, (13): 26-27.
- [11] 方应国, 王芬. 时间序列预测方法综述[J]. 浙江树人大学学报(自然科学版), 2006, 6 (02): 61-65.
- [12] 付海兵, 曾黄麟. BP 神经网络的算法及改进[J]. 中国西部科技, 2012, 11(08): 23-24.
- [13] 宋玉强. 人工神经网络在时间序列预测中的应用研究[D]. 西安: 西安建筑科技大学, 2005.
- [14] 姜春雷, 张树清, 张策, 等. 基于 SARIMA-BP 神经网络组合方法的 MODIS 叶面积指数时间序列建模与预测[J]. 光谱学与光谱分析, 2017, 37(01): 189-193.

ARMA-BP Logistics Demand Forecasting Model and Its Application

LIU Feng-chun¹, ZHAO Ya-ning², DONG Xin-yan², LIU Yuan-shuo²,

QIAO Peng², XIE Zhi-yuan², WANG Li-ya², ZHANG Chun-ying²

(1. Qian College, North China University of Science and Technology, Tangshan Hebei 063000, China;

2. College of Sciences, North China University of Science and Technology, Tangshan Hebei 063210, China)

Key words: time series; neural network; residual error sequence; logistics demand forecasting

Abstract: Based on the real time and uncertainty of logistics demand, the fusion time series autoregressive-moving average model ARMA and BP Neural network were proposed, and the logistics demand forecasting ARMA-BP model was modelled, An ARMA-BP combined prediction algorithm for predicting freight transport was proposed. Based on the logistics transportation data of Tangshan in recent years, using ARMA model, BP model and ARMA-BP model, the logistics data were predicted. The results show that the ARMA-BP model has higher precision and practical value than the traditional prediction model.