

# 天津市物流需求影响因素的分析和预测研究

罗成敏, 王涛\*

云南师范大学, 云南 昆明

收稿日期: 2022年9月24日; 录用日期: 2022年10月19日; 发布日期: 2022年10月26日

## 摘要

本文以物流需求理论为基础, 以天津市的物流需求为例, 运用Matlab软件和SPSS软件进行分析, 建立物流需求预测模型。首先运用SPSS软件分析了天津市1978~2019年的物流需求指数的线性相关性, 结果表明, 天津市货物运输量与各经济指标之间存在极强的相关性, 说明了选取指标的有效性。其次, 本文从天津市物流需求影响因素多元化角度出发, 建立了多元回归模型对物流需求影响因素进一步分析, 对归一化后1978~2014年的数据进行训练, 2015~2019年的数据进行预测, 得出对物流需求影响最大的三项经济指标为: 外贸进出口总额、第三产业产值、CPI。然后引用trainlm函数建立BP神经网络模型, 结合先前的多元回归预测模型得到2种预测结果, 与实际值相比较, 检验模型的有效性。最后, 根据本文研究所得到的结果, 综合众多因素为当地政府在物流需求领域提出实时建议。

## 关键词

物流需求, 多元回归模型, BP神经网络, 因素分析, 预测

## Forecast Analysis of Logistics Demand in Tianjin Based on Neural Network

Chengmin Luo, Tao Wang\*

Yunnan Normal University, Kunming Yunnan

Received: Sep. 24<sup>th</sup>, 2022; accepted: Oct. 19<sup>th</sup>, 2022; published: Oct. 26<sup>th</sup>, 2022

## Abstract

Based on the theory of logistics demand, this paper takes the logistics demand of Tianjin as an example, analyzes the Matlab software and SPSS software, and establishes the logistics demand prediction model. Firstly, the linear correlation of the logistics demand index of Tianjin from 1978

\*通讯作者。

to 2019 was analyzed by SPSS software, and the results showed that there was a strong correlation between the cargo transportation volume and various economic indicators in Tianjin, which explained the effectiveness of the selected indicators. Secondly, from the perspective of the diversification of the influencing factors of logistics demand in Tianjin, this paper establishes a multi-regression model to further analyze the influencing factors of logistics demand, trains the data from 1978 to 2014 after normalization, and forecasts the data from 2015 to 2019, and draws that the three economic indicators with the greatest impact on logistics demand are: total foreign trade import and export, tertiary industry output value, and CPI, then reference the trainlm function to establish the BP neural network model, combine the previous multiple regression prediction model to obtain two kinds of prediction results, compare with the actual value, and test the effectiveness of the model. Finally, based on the results of the research in this paper, a combination of factors provides real-time recommendations for local governments in the field of logistics needs.

## Keywords

Logistics Needs, Multiple Regression Models, BP Neural Network, Factor Analysis, Predict

Copyright © 2022 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

物流行业作为支撑国民经济发展的基础性和先导性行业, 在促进各地区经济发展、物资流通和提升人民生活质量等方面发挥着至关重要的作用[1] [2] [3]。考虑到物流需求的众多影响因素, 本文将对物流需求的众多因素展开调查, 并分析预测天津市近几年的货物运输量来反映天津市的物流需求规模。

本文结合来自《天津市统计年鉴》的货物运输量  $Y$  (万吨)作为天津市物流需求规模的衡量指标, 地区总产值(GDP)  $X_1$  (亿元)、第一产业产值  $X_2$  (亿元)、第二产业产值  $X_3$  (亿元)、第三产业产值  $X_4$  (亿元)、零售总额  $X_5$  (亿元)、进出口总额  $X_6$  (亿元)和居民消费水平(CPI)  $X_7$  (元)作为天津市物流需求模型预测的经济指标。运用 SPSS 软件对天津市物流需求指标做相关性分析, 判断出各指标之间存在着极强的线性相关性, 结合 Matlab 编程软件建立多元回归模型, 分析影响物流需求的因素, 找出主要的影响因素, 针对分析结果为当地政府给出相关应对措施。再选取前 1978~2014 年的数据作为训练样本, 选取 2015~2019 年的数据作为预测样本, 并根据所建立的多元回归模型预测, 建出 BP 神经网络预测模型, 并对传统统计模型(多元回归模型)与 BP 神经网络模型的预测能力做对比, 并对两种模型的性能做出合理解释, 为未来物流需求行业相关措施提供技术性的意见。

## 2. 天津市物流需求年度变化趋势分析与数据预处理

### 2.1. 天津市物流需求年度变化趋势分析

为了能够从多方面探究天津市物流需求的影响因素, 本文从《天津市统计年鉴》中选取了 1978~2019 年的一项衡量指标和七项经济指标, 并对其系统的分析。如图 1 和图 2 分别是 7 项经济指标之间年度增长变化图和衡量指标(货物运输量)年度变化趋势图, 图 1 中显示出地区总产值(GDP)和第三产业产值随着年份的增加而增长。第二产业产值从 1978 年以来长期产值趋于波动状态并且还呈现出下降的趋势, 商品零售总额在 1987~1992 年、2012 年和 2013~2014 年期间总额呈下降趋势, 第一产业产值、居民消费水平

和外贸进出口总额随着年度增加变化不大一直趋于平稳的状态在发展。由此可以看出, 该七项经济指标并非像普通经济指标一样随之年份的增长而增加, 那么其中隐藏着哪些隐含关系, 需要我们进一步探究。

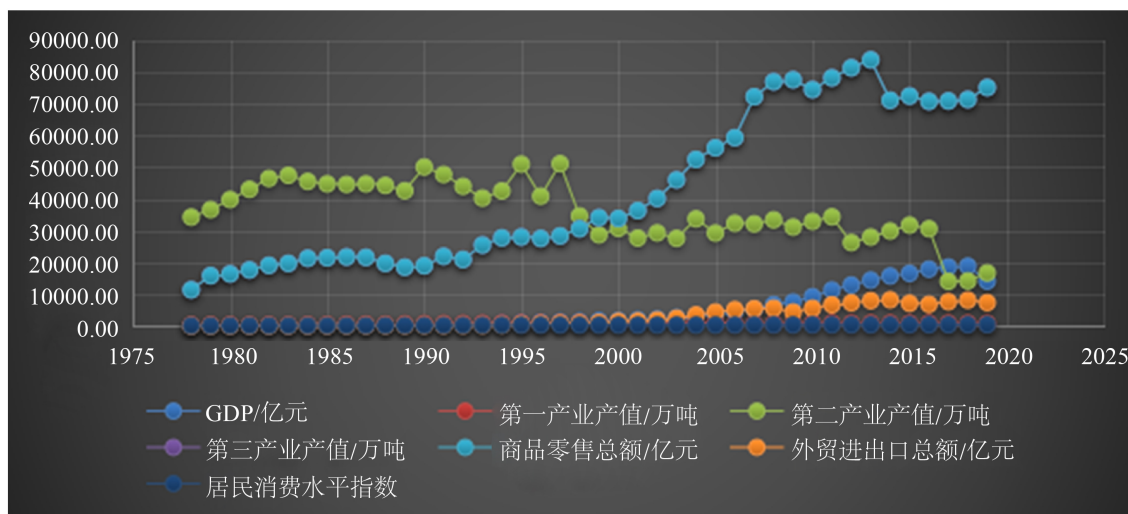


Figure 1. Annual trend chart for each economic indicator

图 1. 各经济指标年度趋势图

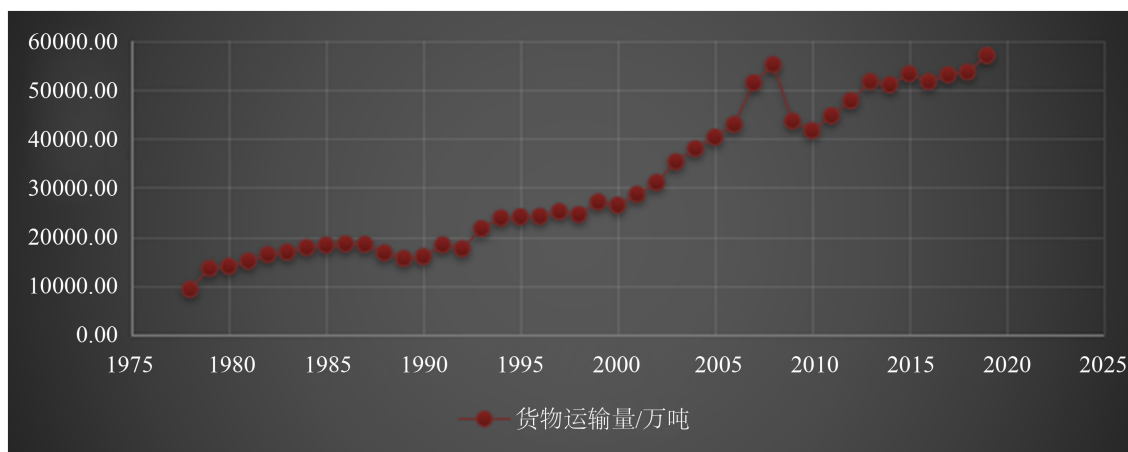


Figure 2. Annual trend chart of cargo traffic

图 2. 货物运输量年度趋势图

如图 2 所示, 随着年份的增加, 货物运输量并非跟随增长, 其中分别在 1980 年、1986~1989 年、1992 年、1998 年、2000 年、2008~2010 年、2014 年、2016 年呈现出了下降的趋势, 天津市在这几年呈现下降趋势的原因有待我们继续探究。为便于分析出其中的各因素之间所存在的关系对货物运输量会产生哪些质的影响, 我将会结合数据特征, 从多元的角度出发, 去探究其中的奥秘。

## 2.2. 数据预处理

### 2.2.1. 缺失值处理

本文根据初始指标体系选取天津市 1987~2019 年的物流需求数据, 发现地区总产值(GDP)缺失了 1981

年的数据。在处理时, 选取 1981 年前 2 年和后 2 年(即 1979~1980 年和 1982~1983 年)的数据, 以它们的平均值作为 1981 年的填充值, 第一产业产值、第三产业产值和进出口总额都有数据缺失, 均前述均值法替之。

### 2.2.2. 归一化处理

由于各数据之间存在着量纲的不同, 为消除量纲不同所带来的影响, 我们对原始数据进行数据归一化, 将数据处理为区间[-1, 1]之间的数据, 归一化公式为:

$$\bar{x}_i = 2 \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} - 1 \tag{1}$$

反归一化公式为:

$$t_i = \frac{(\bar{x}_{\max} - \bar{x}_{\min}) \cdot \bar{x}_i}{2 + \bar{x}_{\min}} \tag{2}$$

## 3. 天津市物流需求影响因素分析

### 3.1. 多元回归模型的建立

综合考虑各指标之间的相关关系, 利用 SPSS 软件, 分析各因素间的线性相关性[4], 得出天津市货物运输量与各经济指标之间存在极强的线性相关性, 而且相对于其他指标因素来说, 天津市对于经济指标的统计较为完善, 相对容易获得。因此, 以上选取的经济指标极具代表性和实际意义。

为了更进一步探索各经济指标之间所存在的关系, 设定多元回归模型[5] [6]为:

$$Y = \alpha_0 + \alpha_1 X1 + \alpha_2 X2 + \alpha_3 X3 + \alpha_4 X4 + \alpha_5 X5 + \alpha_6 X6 + \alpha_7 X7$$

其中  $\alpha_0$  表示常数,  $\alpha_0, \dots, \alpha_7$  表示回归系数。利用 Matlab 软件估计得到各系数值为-0.0341, -0.2992, -0.0504, 0.0467, 0, 0.4668, 0.2829,  $p = 0.0115 < 0.05$  故得到模型为:

$$Y = -0.0341 - 0.2992X1 - 0.0504X2 + 0.0467X3 + 0.4193X5 + 0.4668X6 + 0.2829X7 \tag{3}$$

其中 X5 的系数为 0, 说明商品零售总额对物流需求波动的影响几乎为 0, 这也验证了我们在前面对各指标年度的趋势分析。

### 3.2. 多元回归模型的分析

根据模型(3), 我们可以得出各经济指标对物流需求的影响程度为: 外贸进出口总额 > 第三产业产值 > CPI > 第二产业产值, 对物流需求产生正影响; 第一产业产值和 GDP 对物流需求产生负影响。

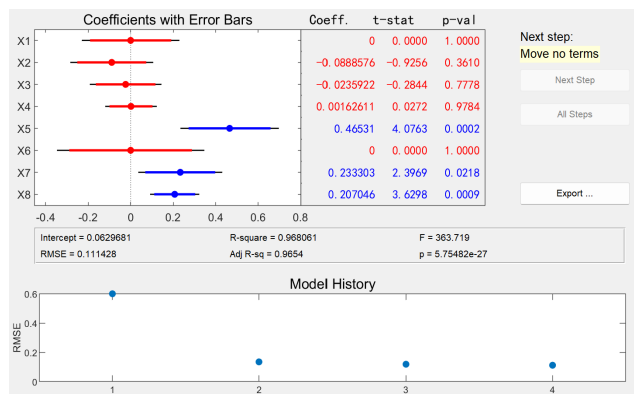


Figure 3. Annual trend chart for each economic indicator  
图 3. 各经济指标年度趋势图

由图 3 可知, 对物流需求产生影响最大的前三项经济指标为: 外贸进出口总额、第三产业产值、CPI。符合模型(3)的分析。

### 3.3. 多元回归模型的预测

根据所建立出的模型(3), 对 2015~2019 年的物流需求规模进行预测,  $Y$  表示的真实值,  $t1$  表示多元回归模型的反归一化预测值,  $r1 = |Y - t1|$  表示真实值与预测值之差的绝对值。得到以下结果:

**Table 1.** Comparison data of multiple regression model prediction results

**表 1.** 多元回归模型预测结果对比数据

| 年份   | Y        | t1       | r1       |
|------|----------|----------|----------|
| 2015 | 53179.00 | 55035.33 | -1856.33 |
| 2016 | 51580.00 | 54842.76 | -3262.76 |
| 2017 | 52992.00 | 54795.96 | -1803.96 |
| 2018 | 53548.00 | 54884.20 | -1336.20 |
| 2019 | 56941.00 | 55367.24 | 1573.76  |

由表 1 可以看出, 用多元回归预测模型得出来结果误差有点大, 那么在下文中我将会引入机器学习中的神经网络预测模型, 对本文物流需求进行预测, 并和我的多元回归预测模型结果作对比, 找出两者之间的差异。

## 4. BP 神经网络对物流需求预测分析

### 4.1. BP 神经网络模型的建立

在神经网络中, `trainlm` 函数收敛速度快, 网络的训练误差较小, 因此本文选取 `trainlm` 函数作为 BP 神经网络模型的训练函数, 建立 BP 神经网络模型。采用三层的 BP 神经网络作为天津市物流需求规模预测模型[7] [8] [9]。结合上面的指标体系以及数据的可获得性原则, 确定训练集为前 37 个年份(1978 年~2014 年)的数据, 预测集为后 5 个年份(2015 年~2019 年)的数据, 输入个数为  $p = 7$ , 输出层为货物运输量, 输出个数  $q = 1$ , 隐含层节点个数的选取原则遵循公式:  $l = \sqrt{p+q} + a$ , 其中  $a$  为 0~10 中的任意整数, 本文取  $a = 2$ , 即隐含层节点个数为 5, 由此就构成了一个 7-5-1 的 BP 神经网络, 其调用格式为:

$$\text{Net} = \text{feedforwardnet}(\text{hiddenSizes}, \text{trainFcn}) \quad (4)$$

其中, `hiddenSizes` 表示隐含层的神经元节点个数, `trainFcn` 为训练函数, 默认值为“`trainlm`”。设定训练次数为 10,000 次, 训练目标为 0.0001, 学习速率为 0.001。

### 4.2. BP 神经网络模型的预测

图 4 是神经网络运行过程的一些结果, 进度条的左边代表下限, 中间代表实际值, 右边代表上限。Epoch 表示训练次数, 14 代表实际训练次数, Time 表示运行时间, Performance 表示性能, 达到设置的上限则停止训练, Grandient 表示梯度, 进度条里的数值代表当前实际梯度, 达到上限训练停止, Validation Checks 表示如果连续训练 6 次误差都不下降则停止训练。

由于训练的初始条件不同, 每次训练的结果都不同, 经过多次的训练, 使预测数据更加接近原始数据。根据所建立的 BP 神经网络预测模型, 预测天津市 2015~2019 年的物流需求规模, 结果见表 2。其中,

Y 表示的真实值,  $t_2$  表示 BP 神经网络模型的反归一化预测值,  $r_2 = |Y - t_2|$  表示真实值与预测值之差的绝对值。

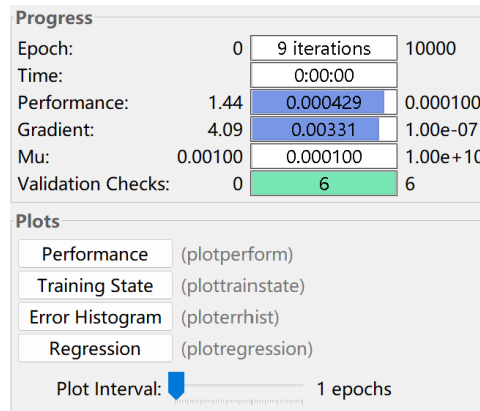


Figure 4. BP neural network training process diagram  
图 4. BP 神经网络训练过程图

Table 2. Comparison data of BP network prediction results  
表 2. BP 网络预测结果对比数据

| 年份   | Y        | $t_2$    | $r_2$  |
|------|----------|----------|--------|
| 2015 | 53179.00 | 53052.88 | 126.12 |
| 2016 | 51580.00 | 51382.24 | 197.76 |
| 2017 | 52992.00 | 52322.33 | 669.67 |
| 2018 | 53548.00 | 52904.05 | 643.95 |
| 2019 | 56941.00 | 56326.45 | 614.54 |

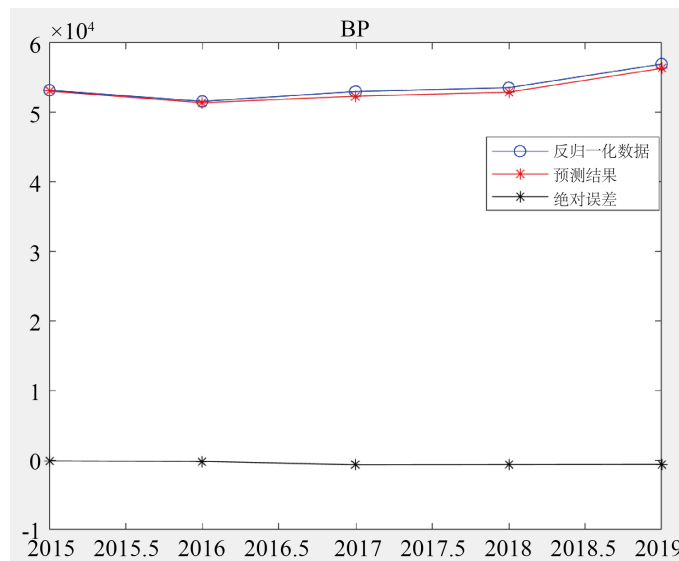


Figure 5. A plot of true values versus predicted values  
图 5. 真实值与预测值的对比图



-----误差计算-----

隐含层节点数为5时的误差结果如下:

平均绝对误差MAE为: 0.25539

均方误差MSE为: 0.04181

均方根误差RMSE为: 0.20448

**Figure 6.** BP neural network error result plot

**图 6.** BP 神经网络误差结果图

根据图 5 和图 6 结论可以得知 BP 神经网络模型的预测性能  $MSE = 0.04181$ , 要比多元回归预测模型的精度要高, 性能优于多元回归预测模型。

## 5. 总结与建议

本文收集了天津市 42 个年份(1978 年~2019 年)的物流预测指标数据: 包括货物运输量、地区总产值(GDP)、第一产业产值、第二产业产值、第三产业产值、零售总额、进出口总额和居民消费水平(CPI), 比较了多元回归模型和 BP 模型的测试结果, 结合实际情况对两种模型进行了分析和解释。得到以下结论:

1) 神经网络模型具有极强的学习能力和非线性逼近能力, 能解决一些因素较多且各因素间非线性相关的问题。

2) 针对本文预测天津市物流需求规模问题, 通过对所建立的两种模型的测试结果和均方误差值(MSE)的对比分析, 可知本文中的两个模型性能排序为: BP 神经网络模型 > 多元回归模型。

3) 研究表明, 针对本文预测天津市物流需求规模问题, BP 神经网络模型, 具有出色的非线性逼近能力, 可准确地表征变化趋势, 且精度明显优于相应的传统模型。

从以上预测结果和结论可以看出, 天津未来几年物流需求将大幅增长, 物流业将迎来良好的发展机遇。物流业不仅是工业品进出口活动与初级产品生产活动之间的桥梁和纽带, 物流业的快速发展必将推动工农业的发展。这将对天津的经济发展产生重大影响。因此, 天津市的物流规划应综合考虑, 根据天津市物流需求的特点和趋势, 进行天津市物流总体规划, 以补充和促进天津市物流业和物流经济的发展。

## 参考文献

- [1] 耿勇, 鞠颂东, 陈娅娜. 基于 BP 神经网络的物流需求分析与预测[J]. 物流技术, 2007(7): 35-37.
- [2] 潘珠. 基于 BP 神经网络的海南省农产品冷链物流需求预测分析[J]. 物流技术, 2020, 39(11): 69-72.
- [3] 黄韧, 潘立军. 基于 BP 神经网络的湘潭市物流需求预测分析[J]. 现代商业, 2011(9): 64-65.
- [4] 陈袁鑫, 张丰硕. 基于 SPSS 的高校图书馆学生满意度影响因素分析[J]. 玉溪师范学院学报, 2020, 36(6): 114-119.
- [5] 谢睿, 杜悦, 朱家明. 基于多元回归对安徽房价影响因素的计量分析[J]. 哈尔滨师范大学(自然科学学报), 2022, 38(4): 36-42.
- [6] 黄钰茜. 基于多元回归分析陕西省物流业发展的环境影响[J]. 中国集体经济, 2022(14): 112-115.
- [7] 陈敏. 基于 BP 神经网络的成都市的物流需求预测[J]. 中国储运, 2021(5): 107-108.
- [8] 刘艳利, 伍大清. 基于改进 BP 神经网络的水产品冷链物流需求预测研究——以浙江省为例[J]. 中国渔业经济, 2020, 38(5): 93-101.
- [9] 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[J]. 铁道科学与工程学报, 2008, 5(6): 62-68.