

文章编号:2095-0365(2015)02-0010-06

基于灰色 GM(1,n)模型的河北物流货运量预测

赵莉琴, 刘敬严

(石家庄铁道大学 经济管理学院, 河北 石家庄 050043)

摘 要:在分析河北省货运量影响因素的基础上,提出了基于灰色关联分析的灰色 GM(1,n)预测模型。在国家统计局指标分类基础上,将河北省货运量影响因素分为国民经济、固定资产投资和房地产、对外经济贸易、能源、运输和邮电、社会消费六类 18 个指标,采用灰色关联分析法对 18 项指标与河北省货运量进行相关性分析,根据灰色关联系数与排序,筛选灰色 GM(1,n)预测变量,减少预测模型输入工作量,计算模型参数。通过对河北省 1993—2012 年货运量实例分析表明,该预测方法具有运算快、精度高的优点。

关键词:货运量;预测;灰色关联分析;GM(1,n)模型;河北省

中图分类号:F713 **文献标识码:**A **DOI:**10.13319/j.cnki.sjztdxbskb.2015.02.02

2009 年以来,我国政府把物流产业作为十大重点产业之一,提出了一系列振兴规划,制定了区域物流协调和持续发展目标^[1]。因此,基于区域经济发展的物流需求预测对区域物流规划设计、道路水路发展、配送节点设立、信息系统及网络配置等方面具有积极的现实意义。区域物流需求预测包括货运量、运输量、道路需求、物流企业及节点需求、车辆需求量等方面的预测^[2],其中货运量指标具有综合性和辐射广泛性特点,可以作为区域物流需求预测的代表性指标。

目前预测研究中经常使用的方法主要集中在时间序列预测模型、回归预测模型和神经网络预测模型三类。时间序列模型在传统的指数平滑法和静态时间序列模型后,现在又根据非线性状态,提出混沌时序分析法,比如基于混沌预测的非线性自适应滤波方法、遗传算法快速混沌预测法、小波网络预测模型等^[3-7]。回归预测模型包括在传统的线性基础上发展起来的一元回归、多元回归、比例估算法和朗格系数法,到近期发展的非线

性回归预测法,比如曲线直线化方法、非线性最小二乘法和近似非线性法等^[8-10]。时间序列预测模型和回归预测模型一般都是基于数学理论和基本假设为前提进行数学建模,对于动态、非规律性、难以描述数据内在影响的事件很难演绎建立数学模型。神经网络法作为典型的非线性预测方法在近期应用比较广泛,但是神经网络预测模型采用误差导数梯度下降的方法进行迭代,难以避免出现局部误差较大的情况^[11-12]。在近期研究中发现,将神经网络模型与其他方法的结合,比如回归神经网络、模糊算法、遗传算法等,可弥补局部极差问题的存在。

灰色 GM(1,n)模型能够从非线性模型中找出规律,寻找数据的整体功能。灰色理论是建立在生产数据的建模上,而非原始数据模型,所以其预测结果精度相对比较高^[13]。灰色 GM(1,n)模型可以依靠小样本进行原始数据重新生成数据计算,然后根据参数对非线性、复杂的数据进行预测。所以采用 GM(1,n)模型能够适应区域货运

收稿日期:2014-12-25

作者简介:赵莉琴(1973—),女,副教授,研究方向:物流管理。

基金项目:河北省教育厅重点项目(SD141049),河北省社科基金项目(HB14GL022),河北省教育厅人文社科重点基地。

本文信息:赵莉琴,刘敬严.基于灰色 GM(1,n)模型的河北物流货运量预测[J].石家庄铁道大学学报:社会科学版,2015,9(2):10-15.

量与影响因素复杂、非线性关系,能够依据比较少的数据进行更好的预测。

由于影响区域货运量因素比较多,如果直接把强相关和弱相关的全部影响因素数据输入 GM(1,n)模型进行测算,不但导致输入工作量大,而且会造成模型结构复杂、运算量大、模拟时间长、精度不够等结果^[14-16]。所以使用灰色关联分析法,对全部影响区域货运量因素进行筛选分析,对定量研究系统内多因素之间相互作用、相互影响进行发展态势的量化比较,有效减少 GM(1,n)模型的变量数量。

一、基本原理

(一)灰色关联分析

灰色关联分析法是对一个系统发展变化态势的定量描述和比较的方法,通过确定参考数据列与若干个比较数据列的几何形状相似度来判断其联系是否紧密,反映了曲线之间的关联程度。与参考数列关联度越大的比较数列,其发展方向和

速率与参考数列越接近,与参考数列的关系越紧密^[17]。具体分析过程如下:

1. 确定分析数列

确定反映系统行为特征的参考数列和影响系统行为的比较数列,参考数列为:

$$Y_0 = \{Y_{0(k)} | k=1, 2, \dots, n\}$$

比较数列为:

$$Y_i = \{Y_{i(k)} | k=1, 2, \dots, n\}, i=1, 2, \dots, n。$$

2. 变量的无量纲化

由于系统中各因素数列的数据可能因纲量不同,不便于进行比较,或者不能得到正确结论,因此在进行灰色关联分析时先进行无量纲化。参考数列进行无量纲化后得到 $X_0 = \{X_{0(k)} = 1, 2, \dots, n\}$;比较数列进行无量纲化得到 $X_i = \{X_{i(k)} | k=1, 2, \dots, n\}; i=1, 2, \dots, n。$

3. 求灰色关联系数

计算参考数列 X_0 与比较数列 $X_1, X_2, \dots, X_n$ 的灰色关联系数。计算公式如下:

$$x_{0i}(k) = \frac{\min\{\min\{|X_{0k} - X_{ik}|\}\} + \xi \max\{\max\{|X_{0k} - X_{ik}|\}\}}{|X_{0k} - X_{ik}| + \xi \max\{\max\{|X_{0k} - X_{ik}|\}\}}$$

式中 $\xi \in [0, 1]$ 为分辨系数,一般取值为 $\xi = 0.5$ 。

4. 求灰色关联度

计算参考数列 Y_0 与比较数列 Y_i 的灰色关联度。因为参考数列与比较数列在曲线各个点的关联程度值不是一个,而且过于分散,所以计算在曲线各个点的平均值,作为参考数列与比较数列的关联程度的表示,其计算公式如下:

$$r_i = \frac{1}{n} \sum \xi_i(k), k=1, 2, \dots, n$$

式中, r_i 代表参考数列 Y_0 与比较数列 Y_i 的灰色关联度。

5. 排关联序

关联度排序是指各因素之间的关联度,主要使用关联度大小次序进行描述,而不仅仅是关联度的大小。将 k 个比较因素与参考因素的关联度按大小进行排序,便组成了关联序,反映了参考因

素对各比较因素的“优劣”关系。

(二)灰色 GM(1,n)模型测算分析

灰色 GM(1,n)模型反映了 $n-1$ 个变量对一个变量的一阶导数的影响,通过拟合效果的检验建立最优模型。

1. 对原始数据做一次累加形成新生数列

$$x_i^{(1)}(k) = \sum_{i=1}^k x_i^{(0)}(t) \quad i=1, 2, 3, \dots, n; k=1, 2,$$

$\dots, m。$

2. 求参数向量

$$\hat{\alpha} = [a, b_1, b_2, \dots, b_{n-1}]^T$$

根据 $y_n = B\hat{\alpha} = (B^T B)^{-1} B^T y_n$, 式中 α 为 GM(1,n)的发展系数; b_i 为 x_i 的协调系数。

$$y_n = [y(2), y(3), \dots, y(m)]^T$$

B 值计算公式为:

$$B = \begin{bmatrix} -[y^{(1)}(1) + y^{(1)}(2)]/2 & x_1^{(1)}(2) & x_2^{(1)}(2) & \dots & x_{n-1}^{(1)}(2) \\ -[y^{(1)}(2) + y^{(1)}(3)]/2 & x_1^{(1)}(3) & x_2^{(1)}(3) & \dots & x_{n-1}^{(1)}(3) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -[y^{(1)}(m-1) + y^{(1)}(m)]/2 & x_1^{(1)}(m) & x_2^{(1)}(m) & \dots & x_{n-1}^{(1)}(m) \end{bmatrix}$$

3. 建立时间响应函数

$$\hat{y}^{(1)}(k) = y_1^{(0)}(1) - \frac{1}{a} \sum_{i=2}^n b_i x_i^{(1)}(k+1) e^{-ak} + \frac{1}{a} \sum_{i=2}^n b_i x_i^{(1)}(k+1) \quad k=2,3,\dots,m$$

4. 建立还原函数

$$\hat{y}_i^{(0)}(k+1) = \hat{x}_i^{(1)} - \hat{x}_i^{(1)}(k)$$

二、河北物流货运量影响因素灰色关联分析

区域物流货运量受多种因素影响,同时也带有明显的地域特征。相对于河北省来说,影响区域货运量的因素主要集中在6个方面^[18],从国家统计局统计标准进行分类,可以分为国民经济、固定资产投资和房地产、对外经济贸易、能源、运输和邮电、社会消费。相对其他省份来说,河北省预测物流货运量必须考虑能源消费和社会消费两个指标。因为河北省能源结构相对来说集中化程度非常高,主要依托煤炭作为能源来源,但是河北省不是煤炭储存大省,需要从山西、内蒙古进行调运,无形中增加了货运量的值。同时,河北省与北京、天津的地缘关系,成为北方小商品集散地,社会消费品的批发功能对河北省货运量也有非常大的影响。

结合河北省特点,选取6个方面作为比较因素。其中国民经济指标中选取了GDP(X_1)、第一产业增加值(X_2)、第二产业增加值(X_3)、第三产业增加值(X_4)四个指标作为比较因素;在固定资产投资和房地产指标中选取了全社会固定资

产额(X_5)作为比较因素;在对外经济贸易指标中选取了经营单位所在地进出口额(X_6)作为比较因素;在能源指标中选取了煤炭消耗量(X_7)、原油消耗量(X_8)作为比较因素;在社会消费品指标中采用批发业主营收入(X_9)、零售业主营收入(X_{10})、社会零售品消费总额(X_{11})作为比较因素;在运输和邮电指标中选取交通运输、仓储和邮电业增加值(X_{12})、铁路营运里程(X_{13})、公路里程(X_{14})、高速等级里程(X_{15})、公路营运车辆拥有量(X_{16})、民用机动船净载重量(X_{17})、快递量(X_{18})七个指标作为比较因素。

在灰色关联分析中,采用河北省1993—2012年统计数据进行分析。因为影响河北省货运量各因素指标统计标准不一,数量级别差异很大,所以首先对20年数据 Y_0 与 $Y_1, Y_2, Y_3, \dots, Y_n$ 进行无量纲化处理,把预处理后的数据转变成级别差异不大的无量纲化数据 $X_0, X_1, X_2, \dots, X_n$ 备用。然后根据灰色关联基本原理中的第三、第四步骤计算关联系数和关联度,然后根据关联度大小和关联程度进行关联序排序。

在计算过程中,设定分辨系数 $\rho=0.5$,然后在关联系数和关联度的基础上导出灰色关联序。最后多次改变分辨系数值,分别设为0.6、0.7、0.8、0.9,其计算结果表明, Y_0 与 $Y_1, Y_2, \dots, Y_n$ 的关联系数和关联度数值有所差异,而灰色关联序的排序没有变化。下面以 $\xi=0.9$ 的值,计算 Y_0 与 $Y_1, Y_2, \dots, Y_n$ 的关联度与关联序的数值,见表1。

表1 影响河北省货运量各因素的灰色关联分析结果

关联度	关联序	关联度	关联序
0.816 5	GDP	0.823 7	零售业主营收入
0.856	第一产业增加值	0.834 1	批发业主营收入
0.805 2	第二产业增加值	0.803 5	交通运输、仓储和邮电业增加值
0.818 4	第三产业增加值	0.591 3	铁路营运里程
0.743 1	全社会固定资产投资	0.814 3	公路营运车辆拥有量
0.745 5	经营单位所在地进出口额	0.818 7	民用机动船净载重量
0.811 6	煤炭消耗量	0.834 1	公路里程
0.772 2	原油消耗量	0.825 1	高速等级里程
0.818 1	社会零售品消费总额	0.817 2	快递量

根据选取影响主因素 $\geq 65\%$ 的原则,计划从18项指标中选取12个作为影响河北省物流货运量的主要指标。从表2子类划分结果中可以看出,与货运量指标灰色关联度指标从大到小排序中可知,12个主要影响指标包括GDP、第一产业增加值、第三产业增加值、煤炭消耗量、批发业主营收入、零售业主营收入、社会零售品消费总额、公路里程、高速等级公路里程、民用机动船净载重量、公路营运车辆拥有量、快递量。说明以上12个指标与河北省货运量指标曲线变化趋势明显吻合,灰色关联度比较紧密。下面把12指标分别进行子类归类。

表2 影响河北省货运量各因素子类划分结果

子类名称	影响因素
国民经济	GDP,第一产业增加值,第三产业增加值
能源	煤炭消耗量
社会消费品零售总额	社会零售品消费总额、批发业主营收入、零售业主营收入
运输和邮电	公路营运车辆拥有量,民用机动船净载重量,公路里程,高速等级里程、快递量

从表1和表2中可以看出,在国民经济子类中,河北省第一产业和第三产业对货运量影响都比较大,原因是河北省第一产业产值在国民经济中占有近30%份额,货运量与第一产业关系也比较大,同时河北省第三产业发展较快,反观第二产业优势并不明显,所以在国民经济子类中,选取河北省GDP、第一产业产值、第三产业产值作为影响因素。固定资产投资和房地产、对外经济贸易各子类指标对河北省货运量影响不大,直接扣除。在能源子类中,煤炭消耗量指标其关联度较紧密,选取煤炭消耗量作为影响因素。在运输和邮电子类中,公路里程、高速等级公路里程、民用机动船净载重量、公路营运车辆拥有量、快递量的关联度较大,选取该五项指标为影响因素。

三、河北省货运量灰色预测模型的建立与预测

(一)统计数据

根据河北省1993—2012年经济统计数据,分

别得出预测指标货运量 x_1 ,影响因素GDP(亿元) x_2 、第一产业增加值(亿元) x_3 、第三产业增加值(亿元) x_4 、煤炭消耗量(万吨) x_5 、批发业主营收入(亿元) x_6 、零售业主营业务收入(亿元) x_7 、社会零售品消费总额(亿元) x_8 、公路里程(万公里) x_9 、高速等级里程(万公里) x_{10} 、民用驳船净载重量(万吨) x_{11} 、公路营运载货汽车拥有量(万辆) x_{12} 、快递量(件) x_{13} 。其中快递量数据(1993—2000年),是根据2000—2012年数据的线性规律进行推算得出。

(二)计算过程

根据河北省1993—2012年货运量及各经济指标统计数据对各数列一次累加,形成货运量的相关因素数据列和另外13个影响因素的系统特征数据列。在此基础上求得参数:

$$\hat{a} = (38\ 832.890\ 0, \ 61\ 287.900\ 0, \ 80\ 102.770\ 0, \ 96\ 726.495\ 0, \ 111\ 536.140\ 0, \ 124\ 073.600\ 0, \ 134\ 813.455\ 0, \ 144\ 058.325\ 0, \ 151\ 757.785\ 0, \ 158\ 227.570\ 0, \ 163\ 995.090\ 0, \ 169\ 275.450\ 0, \ 174\ 054.525\ 0, \ 178\ 439.625\ 0, \ 182\ 544.520\ 0, \ 186\ 247.895\ 0, \ 189\ 399.140\ 0, \ 191\ 917.645\ 0, \ 193\ 856.810\ 0)^T$$

根据参数计算结果构建 y_n 和 B 矩阵,构建GM(1,n)河北省货运量预测模型。其中系统发展系数 a 和驱动项 b 分别为: $a = 1.777\ 7$ 。 $b = 6.564\ 9; -0.037\ 6; -0.048\ 5; 0.822\ 9; 7.998\ 9; -1.8175; -139.269\ 3; 14\ 484.161\ 2; 0.015\ 3; 53.974\ 7; -0.002\ 6; -0.001\ 1$ 。

表3为河北省货运量GM(1,12)预测结果,通过建模分析可知,其准确率为97%以上。

(三)结果分析

通过建模分析,河北省1993—2012年货运量预测值准确率非常高,平均相对误差只有2.924%,相对误差除了2012年、2011年差距比较大以外,其他10个年份都在3%以下,还有7个年份在1%以下。从历史实际数据来看,预测值的准确度比较高,预测结果有较强的可信度。

表 3 河北省货运量 GM(1,12)预测结果

年份	X_1 实际值	X_1 预测值	残差	残差率
2012	219 130	219 130	0	0
2011	189 799	192 723.949 4	2 924.949 4	0.015 4
2010	156 596	153 640.873	-2 955.127 0	-0.018 9
2009	123 065	122 284.726 4	-780.273 6	-0.006 3
2008	106 922	106 867.009 7	-54.990 3	-0.000 5
2007	96 891	96 870.051 2	-20.948 8	-0.000 2
2006	90 831	90 928.891 7	97.891 7	0.001 1
2005	88 342	88 577.047 9	235.047 9	0.002 7
2004	83 672	83 837.863	165.863 0	0.002 0
2003	77 089	77 080.415 6	-8.584 4	-0.000 1
2002	82 226	82 162.925 2	-63.074 8	-0.000 8
2001	78 886	78 817.627 8	-68.372 2	-0.000 9
2000	75 604	75 746.436 6	142.436 6	0.001 9
1999	74 708	75 035.473 6	327.473 6	0.004 4
1998	73 836	73 633.612 5	-202.387 5	-0.002 7
1997	74 516	74 477.309	-38.691 0	-0.000 5
1996	74 475	74 558.662 8	83.662 8	0.001 1
1995	72 535	72 612.476 5	77.476 5	0.001 1
1994	71 719	71 740.692 9	21.692 9	0.000 3
1993	65 949	65 903.813 3	-45.186 7	-0.000 7

四、结论

使用灰色关联分析,确定影响河北省货运量的主要因素,然后根据灰色 GM(1, n)预测模型构建新生数据列矩阵,对河北省货运量进行预测。

实例表明,基于灰色关联分析确定的影响变量合理可行,减少灰色 GM(1, n)预测模型输入变量数目,可以简化预测模型结构,相对提高预测准确度,同时可以简化日常工作量,有利于方法的推广。

参考文献:

[1]赵莉琴,郭跃显.城市物流竞争力评价方法研究[J].地域研究与开发,2011(2):78-79.

[2]赵莉琴,刘敬严,菅红兵.京津冀区域物流与城市群物流功能建设分析[J].石家庄铁道大学学报:社会科学版,2014,8(3):13-15.

[3]Aloísio Carlos de pina, Bruno da Fonseca Monteiro. ANN and Wavelet network meta-models for the coupled analysis of floating production systems[J]. Applied Ocean Research,2014,48(1):21-23.

[4]Takashi Kuremoto, Shinsuke Kimura. Time series forecasting using a deep belief network with restricted Boltzamn Machines[J]. Neurocomputing, 2014, 137 (11):47-54.

[5]Kulikova M V. Maximum likelihood estimation via the extend covariance and combined square-root filters[J]. Mathematics and computers in Simulation,2009, 79(2):1641-1657.

[6]王大鹏.灰色预测模型及中长期电力负荷预测应用[D].武汉:华中科技大学,2013.

[7]Vahid Faghihi, Kenneth F, Reinschmidt, Julian H kang. Construction scheduling using genetic Algorithm based on Building information mode[J]. Expert

- Systems with Applications, 2014, 41(12): 7565-7578.
- [8] 刘昆, 刘向东. 我国铁路集装箱专用平车需求量预测[J]. 铁道车辆, 2014(1): 28-31.
- [9] 邵峰晶, 于忠清. 数据挖掘原理与算法[M]. 北京: 中国水利水电出版社, 2003.
- [10] 段晓晨, 郭兰英, 张新宁. 新建高铁工程非线性造价估算方法研究[J]. 铁道学报, 2013(10): 115-116.
- [11] 刘志杰, 季令, 叶玉玲. 基于径向基神经网络的铁路货运量预测[J]. 铁道学报, 2006, 28(5): 1-5.
- [12] 苏博, 刘鲁, 杨方廷. 基于灰色关联分析的神经网络模型[J]. 系统工程理论与实践, 2008, 28(9): 98-104.
- [13] 周伟, 方志耕. 非线性优化 GM(1, N)模型及其应用研究[J]. 系统工程与电子技术, 2010 (2): 34-39.
- [14] 张永军, 石辉, 梁锦照. 基于 AHP 灰关联及 GM(1, N)建模的静态网损因素分析[J]. 华南理工大学学报, 2011(3): 63-65.
- [15] 樊爱宛, 潘中强, 王巍, 等. 灰色 GM(1, N)模型在河南省煤炭需求预测中的应用[J]. 煤炭技术, 2011 (10): 7-10.
- [16] 鲁亚运, 原峰, 谭枝登. 时滞 GM(1, N)模型及其应用[J]. 统计与决策, 2014(14): 72-73.
- [17] 耿立艳, 张天伟, 赵鹏. 基于灰色关联分析的 LS-SVM 铁路货运量预测[J]. 铁道学报, 2012(3): 2-3.
- [18] 刘敬严, 陈国勋. 京津冀交通运输与区域经济复合系统协同发展分析[J]. 石家庄铁道大学学报: 社会科学版, 2014, 8(4): 23-27.

Forecast of Logistics Freight Volume in Hebei Province

Based on Grey GM (1, n) Model

ZHAO Li-qin, LIU Jing-yan

(School of Economics and Management, Shijiazhuang Tiedao University, Shijiazhuang 050043, China)

Abstract: The prediction model is proposed to improve forecast level of the logistics demand in Hebei Province. The grey prediction GM (1, n) model based on grey relational analysis is constructed by analyzing the influencing factors on traffic volume in Hebei province. Firstly, according to the classification of the national bureau of statistics indexes, influencing factors on freight volume in Hebei province are divided into 6 categories (18 indexes): the national economy, investment in fixed assets and real estate, foreign economic and trade, energy, transportation and post and telecommunications and social spending. Secondly, a correlation analysis is made on 18 indexed and freight volume in Hebei province using grey relational analysis method. Next, According to grey relation coefficient and sorting theory, the grey GM (1, n) prediction variables is screened to reduce the input work load of the forecasting model so as to calculate the model parameters efficiently. Finally, the freight volume data (1993—2012) in Hebei province are input into the model to make an example analysis, and it is shown that the forecasting method has some advantages such as fast operation, high accuracy and so on.

Key words: freight volume; forecast; grey relational analysis; GM (1, n) model; Hebei province