

基于优化的灰色马尔可夫模型对船舶流的预测

张晓雷, 黄洪琼

(上海海事大学 信息工程学院, 上海 201306)

摘要: 为了提高船舶交通流预测的效率和准确率, 综合考虑季节、气候等因素, 通过分析历史船舶流量数据, 在传统的灰色模型基础上构建了基于季节指数的灰色-马尔可夫预测模型。模型将纵向与横向分析相结合的方法运用到船舶预测中, 通过季节指数修正船舶的横向季节性变化, 再用灰色模型进行预测, 最后通过马尔可夫进行误差修正。利用武汉大桥断面的船舶流量数据对该模型进行了实例分析, 使用 MATLAB 将 BP 神经网络模型、GM(1,1) 模型与优化后的灰色马尔可夫模型进行仿真预测, 结果表明, Markov-GM(1,1) 模型具有更高的预测精度和效率, 从而能够相对准确、高效地对船舶交通流量进行预测。

关键词: 船舶交通流量; 季节指数; 灰色预测; 马尔可夫模型; 预测

中图分类号: TP391.9

文献标识码: A

文章编号: 1673-629X(2018)10-0101-04

doi: 10.3969/j.issn.1673-629X.2018.10.021

Prediction of Ship Traffic Flow Based on Optimized Grey Markov Prediction Model

ZHANG Xiao-lei, HUANG Hong-qiong

(School of Information Engineering, Shanghai Maritime University, Shanghai 201306, China)

Abstract: In order to improve the efficiency and accuracy of ship traffic flow forecast, on the basis of the traditional grey model, we build a grey Markov model based on seasonal index by analyzing historical traffic data, comprehensively considering seasons, climates and so on. This model applies a combination of longitudinal and lateral analysis method to ship traffic prediction. Firstly through seasonal index the change of ship transversal seasons is revised. Then the grey model is used for prediction. Finally, the Markov model is applied to correct deviations. The model is verified by the example analysis by ship traffic data of Wuhan bridge section, and BP neural network model, GM(1,1) model and the optimized grey Markov model are simulated and predicted by MATLAB. The results show that Markov-GM(1,1) model has higher prediction accuracy and efficiency, which can predict ship traffic flow accurately and efficiently.

Key words: ship traffic flow; seasonal index; gray prediction; Markov model; prediction

0 引言

近年来,随着社会经济和对外贸易的迅速发展,各种水域的船舶流量也在不断增加,从而导致先进的航行技术和落后的港口运输管理之间的矛盾已经成为现阶段中国水路运输发展的突出矛盾。因此,提升航行水域船舶流量预测的准确性和高效性变得尤为重要^[1]。

同时,现实中的多种因素,如季节、气候、经济、地域文化甚至行业的周期性变化等,都会对船舶流量带来直接或间接的影响,所以拥有一种高效、准确的预测模型是十分必要的。

1 理论基础

目前,用于船舶流预测的方法主要包括灰色预测模型^[2]、线性回归分析^[3]、灰色神经网络^[4]、极大似然估计^[5]、支持向量机^[6]等。这些预测方法具有各自的特点,但有个共同的特点就是要求预测的历史数据有较好的光滑度^[7]。故文中采用一种新型的预测方法,即基于季节指数的灰色-马尔可夫预测模型。

在同一年时间内船舶流量的变化存在明显的季节变化,故引入季节指数来对季节变化进行修正。将修正后的数据作为灰色预测模型的原始数据用于建模预测,然后引入 Markov 模型来修正和弥补灰色模型的

收稿日期:2017-11-11

修回日期:2018-03-14

网络出版时间:2018-05-28

基金项目:国家自然科学基金面上项目(61673260)

作者简介:张晓雷(1993-),男,硕士,研究方向为智能交通系统及航运智能运输系统;黄洪琼,副教授,研究方向为智能信息处理、智能交通系统及航运智能运输系统。

网络出版地址: <http://cnki.net/kcms/detail/61.1450.TP.20180525.1616.082.html>

预测误差。通过构建实际值与灰色模型预测残差的时间序列,研究误差的变化规律与趋势。

2 模型的建立

2.1 数据的预处理

为了进一步寻找船舶流量在一定时间跨度上的变化规律,在模型中引入功率谱。功率谱即功率密度谱,表征了单位频带内的功率大小,通过功率谱分析可以提取出各个频率分量。以各个月的船舶总流量为采样点,由于一年有 12 个月,因此采样间隔为 12 Hz,得到离散序列 $\{x(n)\}$ 。然后,对序列 $\{x(n)\}$ 去均值消除零频道影响^[8]。同时,采用带阻滤波的方法过滤去 1 Hz 左右的频率分量,根据滤波后的船舶数据序列功率谱研究确定船舶的年际变化周期。

2.2 季节指数

季节指数是反映事物季节性变动规律的一套指数,表示每年反复出现有规律的周期性变动,且每年上下变动的幅度大体相似。船舶的月流量季节指数可以用来描述一个年度内各个月的船舶流量特征,反映各月的船舶流量占全年的船舶平均数值大小。文中采用平均法^[9]测定各月船舶流季节指数,计算步骤如下:

(1) 将原始序列 $\{x(n)\}$ 按周期分组,划分为 N 年,每年 $M = 12$ 个数据,即 x_{NM} 。

(2) 计算过去 N 年中,各月的船舶流的平均值 $\bar{x}_i (i=1, 2, \dots, M)$ 为:

$$\bar{x}_i = \frac{1}{N} \sum_{j=1}^N x_{ij} \quad (1)$$

(3) 计算 N 年月平均船舶流的总平均数 \bar{x}_i :

$$\bar{x}_i = \frac{1}{M} \sum_{i=1}^M \bar{x}_i \quad (2)$$

(4) 计算各月季节指数 τ_i :

$$\tau_i = \frac{\bar{x}_i}{\bar{x}_i} \quad (3)$$

(5) 测定季节指数后,利用 τ_i 对原始数据进行修正。为了满足灰色模型对历史数据的要求,将原始数据进行取对数处理。

2.3 GM(1,1) 模型

灰色理论是基于数学理论的系统工程科学,通过少量的、不完全的信息,建立灰色微分预测模型^[10],故可以运用灰色理论的方法解决船舶流的预测问题,而 GM(1,1) 模型是最常用的灰色模型。

设经过季节指数修正和对数变换后的数据列为:

$$x^{(0)} = \{x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n)\} \quad (4)$$

其中, n 为序列长度。

将式 4 进行累加,得:

$$x^{(1)} = \{x^{(1)}(1), x^{(1)}(2), \dots, x^{(1)}(n)\} \quad (5)$$

其中:

$$x^{(1)}(n) = \sum_{t=1}^n x^{(0)}(t) \quad (6)$$

$$\frac{dX^{(1)}}{dt} + \alpha X^{(1)} = b \quad (7)$$

其中, α 为参数,记为 $A = [ab]^T$,并用最小二乘法确定参数:

$$A = (B^T B)^{(-1)} B^T X_{N-1} = [ab]^T \quad (8)$$

其中:

$$B = \begin{bmatrix} -\frac{1}{2}(X^{(1)}(1) + X^{(2)}(2)) & 1 \\ -\frac{1}{2}(X^{(1)}(2) + X^{(2)}(3)) & 2 \\ \vdots & \vdots \\ -\frac{1}{2}(X^{(1)}(n-1) + X^{(2)}(n)) & 1 \end{bmatrix} \quad (9)$$

$$x_{N-1} = [x^{(0)}(2) x^{(0)}(3) \dots x^{(0)}(n)]^T \quad (10)$$

得到微分方程:

$$\hat{x}^{(1)}(t+1) = (x^{(0)}(1) - \frac{b}{a})e^{-at} + \frac{b}{a} \quad (11)$$

经过累减还原得到:

$$\hat{x}^{(0)}(t+1) = \hat{x}^{(1)}(t+1) - \hat{x}^{(1)}(t), k=1, 2, \dots, s \quad (12)$$

将上式数据通过 $e^{\hat{x}^{(0)}(k)}$ 进行还原^[11],然后再利用季节指数 τ_i 进行修正还原得到 $x_i^{(0)}$ 。

模型误差分析时,常用相对误差进行检验,即:

$$\delta_k = \frac{x^{(0)}(k) - \hat{x}_i^{(0)}(k)}{\hat{x}_i^{(0)}(k)} \times 100\% \quad (13)$$

平均相对误差为:

$$\bar{\delta} = \frac{1}{n} \sum_{k=1}^n |\delta_k| \quad (14)$$

2.4 马尔可夫修正模型

灰色预测依旧存在一定误差^[12],故引入马尔可夫模型进行修正。马尔可夫模型的预测原理是根据某些变量的显著状态及其变化趋势,预测其在未来某一特定期间内可能出现的状态^[13],步骤为:

(1) 计算波动指数序列:由式 13 可得 δ_k 。

(2) 划分 Markov 状态。

$$\delta_k \in [\alpha_{1i}, \alpha_{2i}], i=1, 2, \dots, s \quad (15)$$

上式表示第 k 对象的波动指数处于第 i 种状态 E_i , α_{1i} 、 α_{2i} 分别表示 E_i 的上、下界。因此,总的状态集合表示为 $E = (E_1, E_2, \dots, E_n)$ 。

(3) 构造状态转移矩阵 P ,其元素 $p_{ij}^{(n)}$ 为:

$$p_{ij}^{(n)} = \frac{M_{ij}^{(n)}}{M_i}, i, j=1, 2, \dots, s \quad (16)$$

其中, $M_{ij}^{(n)}$ 为随机过程经过 n 次转移由状态 E_i 到状态 E_j 的原始样本数据; M_i 为处于状态 E_i 的原始样

本数。

(4) 构造未来状态矩阵 p_{ij} 。

根据相对于预测目标的距离选取 s 个预测对象,按照从近到远的顺序排序,分别以各个对象所对应的状态为矩阵的初始状态,在 n 步转移概率矩阵 p 中选取各自所对应的行向量^[14],构造出新的概率矩阵:

$$P_{ij} = \begin{bmatrix} P_{i1}^{(1)} & P_{i2}^{(1)} & \cdots & P_{is}^{(1)} \\ P_{i1}^{(2)} & P_{i2}^{(2)} & \cdots & P_{is}^{(2)} \\ \vdots & \vdots & \ddots & \vdots \\ P_{i1}^{(s)} & P_{i2}^{(s)} & \cdots & P_{is}^{(s)} \end{bmatrix} \quad (17)$$

(5)取 $\max\{\sum_{n=1}^s p_{si}^{(n)}, i \in E\}$ 所对应的状态作为预测对象的状态,则预测方程为:

$$\hat{Y}(k) = \hat{X}^{(0)}(k) [1 + 0.5(\alpha_{1i} + \alpha_{2i})],$$
$$k = 1, 2, \cdots, n \quad (18)$$

3 实验仿真及结果分析

以 MATLAB 为实验平台,测试数据选取武汉大桥断面 2007 年 1 月—2016 年 12 月的船舶交通流量,其中以 2007 年 1 月—2015 年 12 月的 108 个数据作为学习样本来建立模型,以 2016 年 12 个月的船舶流量数据作为测试样本来检测模型。模型原始数据 $x^{(0)}$ 为:

表 1 2007 年—2015 年各月的船舶误差表									
月份	2007 年	2008 年	2009 年	2010 年	2011 年	2012 年	2013 年	2014 年	2015 年
1 月	10.22	-7.69	-13.62	8.15	-5.57	3.04	0.89	4.02	6.82
2 月	-5.57	-2.9	-3.43	-1.08	-2.29	-0.59	-0.96	1.55	8.55
3 月	3.64	2.56	5.68	3.42	4.63	1.02	2.33	-0.44	2.43
4 月	6.47	5.41	4.78	5.54	-3.01	0.78	4.12	3.72	2.28
5 月	4.96	-3.01	6.04	5.06	3.42	2.11	9.56	6.28	-3.68
6 月	-3.95	-2.45	-2.74	-3.91	1.91	6.89	11.58	5.44	2.66
7 月	6.28	5.78	6.89	6.31	3.11	-0.56	9.56	-1.1	4.58
8 月	-1.26	2.31	3.14	2.52	-2.12	-3.5	4.68	-7.04	8.45
9 月	-0.44	6.48	5.09	4.38	0.89	0.77	8.79	6.09	1.02
10 月	3.28	2.84	2.45	1.49	0.22	3.66	6.36	-1.01	0.66
11 月	-2.29	-3.59	-1.57	-3.75	-0.78	2.93	7.11	-2.95	3.89
12 月	6.95	7.06	5.42	4.44	1.01	3.53	2.33	4.65	-0.88

通过用上述方法对数据进行处理后,分别用 GM(1,1)模型和马氏修正后的 GM(1,1)模型对 2016 年各个月的船舶流进行预测,并与实际值进行对比。同时,为了更好地对实验结果进行对比,用 BP 神经网络^[15]对同一时间段的船舶流量进行了预测,结果分别如图 2~4 所示。

为了对实验结果进行比较,分别记录 BP 模型、GM(1,1)模型、马氏修正后的 GM(1,1)三种算法的平均绝对误差(MAE)、平均绝对百分比误差

$$x^{(0)} = \begin{bmatrix} 12\ 537 & 9\ 683 & \cdots & 11\ 684 \\ 12\ 764 & 9\ 374 & \cdots & 12\ 101 \\ \vdots & \vdots & \ddots & \vdots \\ 11\ 067 & 8\ 124 & \cdots & 11\ 749 \end{bmatrix} \quad (19)$$

2007 年—2015 年的船舶交通流量数据序列经过过滤以后的功率谱如图 1 所示。

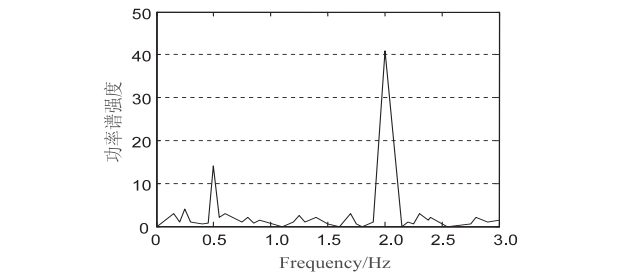


图 1 滤去 1 Hz 分量后的月船舶流量功率谱
由图 1 可知,功率谱中有 2 个比较明显的峰值:

(1)频率为 2 Hz 时,对应 0.5 年的周期分量,其功率谱强度和一年中季节和生活规律的升降规律一致;

(2)频率为 0.5 Hz 时,对应 2 年的周期分量,表明船舶流的年际变化以 2 年为周期,因此,文中确定武汉大桥的船舶流量的季节性变化周期为 2 年。

综上,用 GM(1,1)预测模型进行逐年预测,得出了 2007 年—2015 年各月船舶流量的误差计算结果,如表 1 所示。

(MAPE)、均方根误差(RMSE)^[16]。令 p_t 为实际测值, \hat{p}_t 为预测值, n 为观测序列总数,预测指标如下:

$$MAE = \frac{1}{n} \sum_{t=1}^n |p_t - \hat{p}_t| \quad (20)$$

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{p_t - \hat{p}_t}{p_t} \right| \quad (21)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (p_t - \hat{p}_t)^2} \quad (22)$$

将三种模型的 MAE、MAPE、RMSE 进行计算比

较,其结果如表 2 所示。

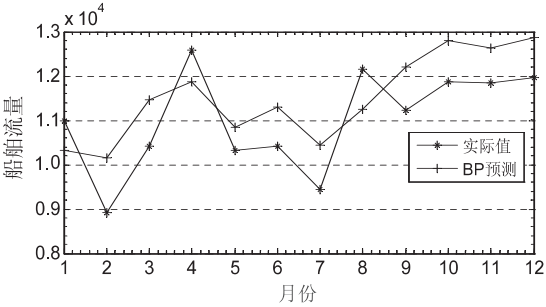


图 2 BP 神经网络模型预测结果

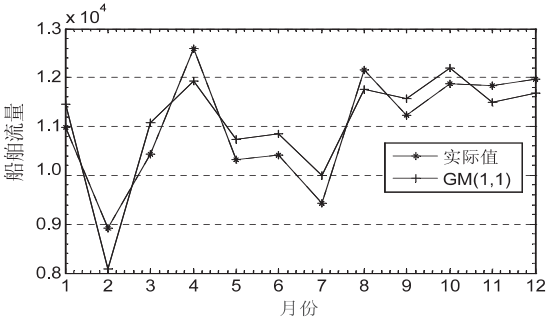


图 3 GM(1,1) 模型预测结果

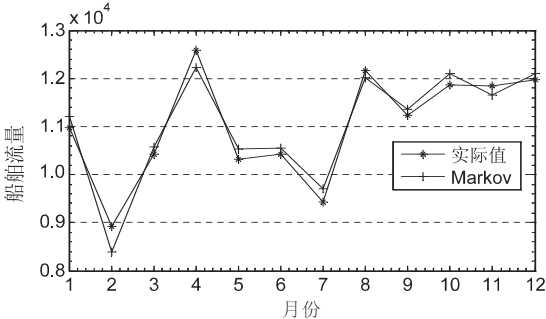


图 4 马氏修正后的预测结果

表 2 3 种预测模型性能指标比较

性能指数	BP	GM(1,1)	Markov-GM(1,1)
MAE	878.42	473.43	188.67
MAPE	8.14%	4.43%	2.11%
RMSE	896.92	499.15	249.64

由表 2 可知, Markov - GM (1 , 1) 预测模型的 MAE、MAPE、RMSE 都明显小于 BP 模型和 GM (1 , 1) 模型。因此, Markov - GM (1 , 1) 模型比 BP 神经网络模型和 GM (1 , 1) 模型在船舶预测的准确性和高效性等方面优势显著。

4 结束语

在实际生活中, 船舶交通流量会受到季节、气候、人为活动等因素的影响而形成季节性周期波动。文中利用周期波动性因素, 在传统灰色预测模型的基础上, 引入马尔可夫模型对其进行修正, 构建了一种新型的船舶预测模型。实验结果表明, 改进的预测模型可以更好地反映船舶流量的总体变化趋势和数据波动特

征, 并且与传统的预测模型相比, Markov - GM (1 , 1) 模型的预测精度更高, 稳定性更好。研究表明, 该方法适用于中长期船舶交通流量的预测。

参考文献:

[1] YANG Liang,HAO Yong,LIU Qing,et al. Ship traffic volume forecast in bridge area based on enhanced hybrid radial basis function neural networks[C]//International conference on transportation information and safety. Wuhan, China: IEEE,2015:38-43.

[2] 张树奎,肖英杰. 船舶交通流量预测的灰色神经网络模型[J]. 上海海事大学学报,2015,36(1):46-49.

[3] 王 东,熊锡龙. 基于影响因素分析的船舶交通流量预测多元线性回归模型[J]. 船海工程,2010,39(3):178-180.

[4] 李 俊,徐志京,唐贝贝. 基于 GA 优化的灰色神经网络船舶交通流量预测方法研究[J]. 船海工程, 2013,42(5):135-137.

[5] 谷耀辉. 船舶交通流量及几种预测方法分析[J]. 科技视界,2016(12):147.

[6] 崔翔鹏,黄洪琼. 基于 GA 优化 ELM 的船舶交通流预测模型[J]. 微型机与应用,2017,36(9):15-17.

[7] 何思兰,孙红兵. 基于灰色预测和 BP 神经网络模型的云南省人口总量预测研究[J]. 计算机与数字工程,2016,44(2):193-196.

[8] 周华任,李浩然,孙学金,等. 一种基于季节指数的灰色马尔可夫气温预测模型[J]. 数学的实践与认识,2016,46(4):167-173.

[9] 贺 辉,刘清良. 电力负荷季节指数的测定和分析[J]. 电力需求侧管理,2010,13(1):33-36.

[10] 李梦婉,沙秀艳. 基于 GM(1,1) 灰色预测模型的改进与应用[J]. 计算机工程与应用,2016,52(4):24-30.

[11] CAO W,LIU Z,ZHENG L. Application of gray neural network in the prediction of slope deformation[J]. International Journal of Applied Mathematics & Statistics,2013,51(22):196-204.

[12] WANG Yuhong,DANG Yaoguo,LI Yueqing,et al. An approach to increase prediction precision of GM(1,1) model based on optimization of the initial condition[J]. Expert Systems with Applications,2010,37(8):5640-5644.

[13] 李克昭,李志伟,赵磊杰. 马尔可夫理论的优化灰色模型预测建模[J]. 测绘科学,2016,41(8):1-5.

[14] SUN Xu,SUN Wangshu,WANG Jianzhou,et al. Using a grey-Markov model optimized by Cuckoo search algorithm to forecast the annual foreign tourist arrivals to China[J]. Tourism Management,2016,52:369-379.

[15] 翟久刚,田延飞,严新平. 基于 BP 神经网络与残差分析的船舶交通流量预测[J]. 上海海事大学学报,2013,34(1):19-22.

[16] 贾朝龙. 铁路轨道不平顺数据挖掘及其时间序列趋势预测研究[D]. 北京:北京交通大学,2013.