[文章编号] 1007 - 7405(2019)01 - 0031 - 06　　DOI:10.19715/j.jmuzr.2019.01.05

基于组合模型的厦门港集装箱吞吐量预测

汪　强，刘晓佳，闫长健，张　荀

（集美大学航海学院，福建 厦门 361021）

[摘要] 为预测厦门港未来集装箱吞吐量，运用灰色 Verhulst 与马尔科夫组合模型建立厦门港集装箱吞吐量的预测模型，得出厦门港 2018—2022 年的集装箱吞吐量数据。研究结果表明，组合模型将平均绝对误差由 3.74% 降低至 1.65%，预测精度为一级。预测结果具有较高的可信度，可为厦门港的集装箱未来发展规划提供参考依据。

[关键词] 集装箱吞吐量；灰色 Verhulst；马尔科夫模型；厦门港
[中图分类号] U 695.2 +2

Container Throughput Prediction of Xiamen Port Based on Combinatorial Model

WANG Qiang, LIU Xiaojia, YAN Changjian, ZHANG Xun

(Navigation College, Jimei University, Xiamen 361021, China)

Abstract: In order to predict the future development of container throughput of Xiamen Port with high precision, a model combining the Grey Verhulst with Markov is used to establish the forecasting model, and the container throughput data of Xiamen Port from 2018 to 2022 are obtained. The results show that the combined model reduces the average relative error from 3.74% to 1.65% and the prediction accuracy is Level one. The prediction results have higher reliability and can be used as a reference for the planning of container development in Xiamen Port.

Keywords: containert throughput; grey verhulst model; markov model; Xiamen Port

0　引言

近几年来，厦门港集装箱运输快速发展，港口集装箱运输总体呈上升趋势，2017 年，厦门港集装箱年吞吐量突破 10^7TEU。随着海上丝绸之路的开辟，中欧班列的开通，全国第一个全自动化码头的建立，厦门港集装箱运输面临新一轮的机遇与挑战。

集装箱运输是一个地区经济发展的重要组成部分，准确预测未来集装箱吞吐量不仅能为该地区经济发展规划提供有力的数据支撑，而且能够为集装箱运输所带来的环境影响提供可靠的数据分析。厦门港的集装箱吞吐量仍处于中等水平，历年集装箱吞吐量数据量少，影响数据发展要素权重不明确。目前对集装箱吞吐量的预测方法有：时间序列法、灰色理论和 BP 神经网络。时间序列预测虽然能消除原始时间序列的波动性问题，但是未考虑到因偶然因素而产生的随机性问题[1]。灰色理论包括

[收稿日期] 2018 - 09 - 12
[基金项目] 福建省教育厅项目（JT180260）
[作者简介] 汪强（1993—），男，硕士生，主要研究方向为交通运输规划与管理。通信作者：刘晓佳（1979—），女，副教授，博士，从事交通运输规划与管理研究。E-mail:ehappylxj1314@163.com

GM(1,1)、灰色 Verhulst 和灰关联分析等，主要用于研究数据少、信息贫的不确定性问题，但对于波动性较大的振荡序列往往存在精度不高，预测结果存在偏差较大等问题[2]。BP 神经网络模型对初始网络权重非常敏感，作为预测模型时，需要大量的原始数据进行训练[3]。由上述分析可知，单独使用一种方法并不能很好地解决“样本小、信息贫”的时间序列预测问题。因此，本文将灰色理论与马尔科夫状态转移预测方法相结合，先利用灰色 Verhulst 模型对原始时间序列进行预测，再利用相对误差修正方法对模型进行有效改进。最后，通过马尔科夫模型判断相对误差状态，对预测值进行修正。

1 灰色 Verhulst – 马尔科夫模型的建立

1.1 灰色 Verhulst 模型

灰色 Verhulst 对原始数据呈“S”型发展，先稳定上涨，随后下降，最后保持稳定增长水平的序列预测有较高的模拟精度[4]。

设一段原始非负时间序列为：

$$X^{(0)} = \{x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(n)\}, \tag{1}$$

$$X^{(1)} = \{x^{(1)}(1), x^{(1)}(2), \cdots, x^{(1)}(n)\}, \tag{2}$$

$X^{(1)}$ 为 $X^{(0)}$ 的 1 – AGO 序列，即

$$x^{(1)}(t) = \sum_{i=1}^{t} x^{(0)}(i), t = 1,2,\cdots n。\tag{3}$$

$Z^{(1)}$ 为 $X^{(1)}$ 的紧邻均值生成序列：

$$Z^{(1)} = (z^{(1)}(2), z^{(1)}(3), \cdots, z^{(1)}(n)), \tag{4}$$

$$Z^{(1)}(t) = (x^{(1)}(t) + x^{(1)}(t-1))/2, t = 2,3\cdots n, \tag{5}$$

称 $X^{(0)} + aZ^{(1)} = b\,(Z^{(1)})^2$ 为灰色 Verhulst 模型，其中 a 和 b 为参数。

$$\mathrm{d}(x^{(1)})/\mathrm{d}t + ax^{(1)} = b(x^{(1)})^2 \tag{6}$$

为灰色 Verhulst 模型的白化方程，其中 $\hat{a} = (a,b)^{\mathrm{T}} = (\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B})^{-1}\boldsymbol{B}^{\mathrm{T}}\boldsymbol{Y}$ 为参数列，且

$$\boldsymbol{Y} = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(n) \end{bmatrix}, \boldsymbol{B} = \begin{bmatrix} -z^{(1)}(2) & (z^{(1)}(2))^2 \\ -z^{(1)}(3) & (z^{(1)}(3))^2 \\ \vdots & \vdots \\ -z^{(1)}(n) & (z^{(1)}(n))^2 \end{bmatrix}。\tag{7}$$

时间响应函数为：

$$x^{(1)}(t) = ax^{(1)}(1)/(bx^{(1)}(1) + (a - bx^{(1)}(1))\mathrm{e}^{at})。\tag{8}$$

原始序列的预测值为：

$$\hat{x}^{(0)}(t+1) = \hat{x}^{(1)}(t+1) - \hat{x}^{(1)}(t)。\tag{9}$$

1.2 马尔科夫预测模型

马尔科夫预测模型对某个系统今后所处的状态仅与目前所处的状态有关，与过去系统所处的状态无关，即为无后性。并且马尔科夫预测模型对一段波动性明显的无后效时间序列进行预测时，优势明显，主要可划分为以下三步[5]。

1）状态划分 利用灰色 Verhulst 预测结果与实际数据之间的相对误差，根据其相对误差大小将其平均分成若干个状态区间[6]。

2）计算概率转移矩阵 根据各年预测的相对误差的大小，分别落入不同的状态区间的结果，计算步长为 n 的状态转移概率矩阵 $\boldsymbol{R}^{(n)}$。

3）改进预测值 通过灰色 Verhulst 预测结果与实际数据之间的相对误差大小，将其平均划分为 j 状态区间 $[w_{j-}, w_{j+}]$，w_{j-}, w_{j+} 分别表示 j 状态区间的上下确界，取相对误差状态区间的中值作为灰色

预测的修正值，

$$\hat{y} = \hat{x}^{(0)}(k)/[1 \pm (w_{j-} + w_{j+})/2]。 \tag{10}$$

其中：当预测值比实际值高估时，分母中的正负号取正值；当预测值比实际值低时，分母中正负号取负值；当预测值与实际值比较相近时，不用修正；无法比较时，取负值[7]。

1.3　组合模型

灰色 Verhulst 模型可对中长期的时间序列进行预测，但存在精度不高问题。马尔科夫模型可针对模型相对误差波动性明显的序列进行改进，提高预测精度。

本文提出先使用灰色 Verhulst 模型对厦门港集装箱序列进行预测，然后采用马尔科夫模型对其预测结果的相对误差进行修正，完成两种模型的组合。

在完成对模型的建立后，仍需要对模型的适用性进行检验。对于预测结果的有效度检验，可根据后验差比值 $c = S_2^2/S_1^2$（S_1^2 为原始数据方差，S_2^2 为预测残差方差）和小误差概率 $p = \{|q(t) - \bar{q}| < 0.6745S\}$（$S$ 为预测数据标准差；$\bar{q}$ 为残差均值）这两个指标，综合评定预测模型的精度[8]。

当所建模型的分级标准为四级（$p < 0.75, c > 0.65$）时，一般不能用该模型进行预测。

2　厦门港集装箱吞吐量预测

2.1　灰色 Verhulst 预测

厦门港全年集装箱吞吐量 2000 年首次突破 10^6TEU，2017 年，突破 10^7TEU。选用（2006—2017 年）厦门港集装箱吞吐量数据①，建立灰色 Verhulst 初始时间序列，如图 1 所示。

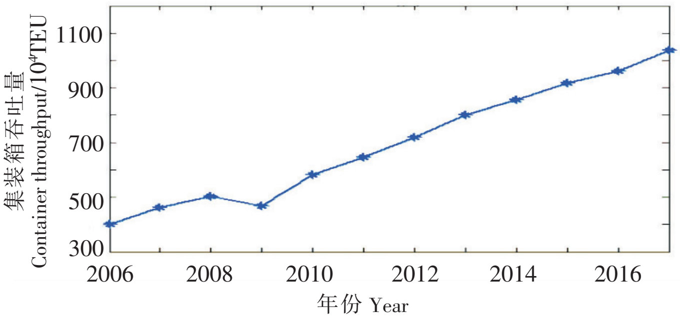

图 1　厦门港历年集装箱吞吐量发展

Fig.1　Annual container throughput of Xiamen Port

通过图 1 可发现，厦门港集装箱历年增长呈“S”型趋势，先增长后下降再保持持续增长，总体呈上升趋势，因此选用灰色 Verhulst 模型进行预测。

根据式（1）～式（10）计算得灰色 Verhulst 相应数据如下：

$$\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B} = \begin{bmatrix} -432.00 & -483.10 & -485.80 & -525.25 & -614.45 & -683.35 & -760.5 & -829.00 & -887.60 & -939.50 & -999.50 \\ 186\,624 & 233\,385 & 236\,001 & 275\,887 & 377\,548 & 466\,967 & 578\,360 & 687\,241 & 787\,833 & 88\,260 & 999\,000 \end{bmatrix},$$

$$\boldsymbol{B}^{\mathrm{T}}\boldsymbol{Y} = \begin{bmatrix} -458 & 201.155 \\ 350\,930 & 630.522 \end{bmatrix}。$$

由于 $\hat{a} = (a,b)^{\mathrm{T}} = (\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B})^{-1}\boldsymbol{B}^{\mathrm{T}}\boldsymbol{Y}$，因此得出：$a = -0.140\,271$，$b = -0.000\,076$。

灰微分方程：$\mathrm{d}(x^{(1)})/\mathrm{d}t - 0.140\,271x^{(1)} = -0.000\,076\,(x^{(1)})^2$。

时间响应式最终整理得：$\hat{x}^{(1)}(t+1) = -56.296\,358/(-0.030\,315 - 0.109\,956\mathrm{e}^{-0.140\,271t})$。

根据式（1）～式（9）计算得灰色 Verhulst 预测结果如表 1 所示，相对误差波动如图 2 所示。由表 1 可看出灰色 Verhulst 预测实际值与模拟值的残差序列仍较大。计算得灰色 Verhulst 预测平均相对误差为 3.74%，精度不高，需要改进。因此引入马尔科夫模型修正相对误差序列，改进预测结果[9]。

① 数据来源：http://www.chinaports.com/ports

表 1 模型数据对比

Tab. 1 Model data comparison

年份 Year	实际数据 The actual data /10^4TEU	灰色 Verhulst 预测 Grey Verhulst forecast /10^4TEU	残差 Residual	相对误差 Relative error δ_{1i}	年份 Year	实际数据 The actual data /10^4TEU	灰色 Verhulst 预测 Grey Verhulst forecast /10^4TEU	残差 Residual	相对误差 Relative error δ_{1i}
2006	401.30	401.34	0	0	2012	720.20	724.47	4.30	0.0060
2007	462.70	447.22	-15.49	-0.0335	2013	800.80	787.31	-13.49	-0.0168
2008	503.50	496.56	6.90	-0.0137	2014	857.20	851.51	-5.73	-0.0067
2009	468.10	549.22	81.19	0.1735	2015	918.00	916.46	-1.54	-0.0017
2010	582.40	646.49	22.55	0.1101	2016	961.00	981.52	20.52	0.0214
2011	646.50	683.53	37.03	0.0572	2017	1038.00	1047.03	9.03	0.0087

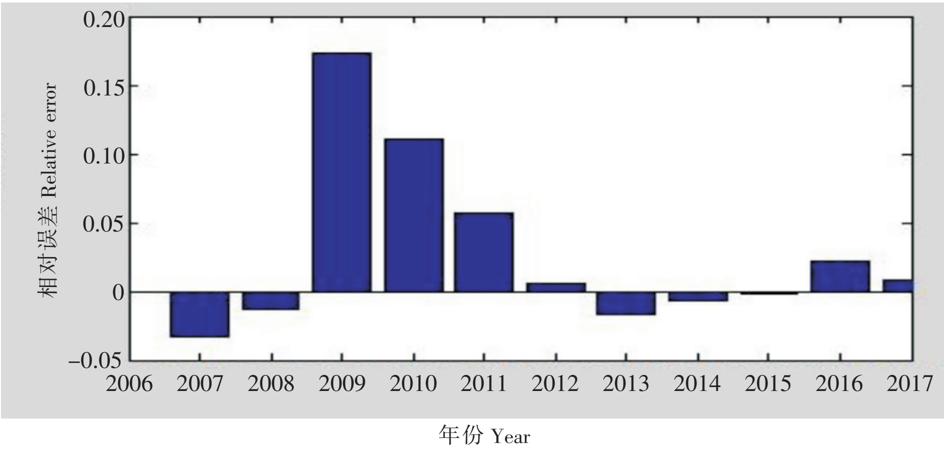

图 2 模型相对误差对比

Fig.2 Model relative error comparison

2.2 马尔科夫修正

根据预测的相对误差大小划分状态区间，因为预测值与实际值的相对误差波动较大，因此在状态划分时，应尽量考虑状态多一些[10]。这里划分为 5 个状态 $w_j(j=1,2,\cdots,5)$，结果见表 2。

表 2 相对误差状态划分区间

Tab. 2 Relative Error state partition

w_1	w_2	w_3	w_4	w_5
[-0.0335,0.0079)	[0.0079,0.0493)	[0.0493,0.0907)	[0.0907,0.1321)	[0.1321,0.1735]

由相对误差状态区间及初始状态划分计算马尔科夫模型的一步转移概率矩阵 $\boldsymbol{R}^{(1)}$ 和 $\boldsymbol{R}^{(2)}$。

$$\boldsymbol{R}^{(1)}=\begin{bmatrix}5/7 & 1/7 & 0 & 0 & 1/7\\ 0 & 1 & 0 & 0 & 0\\ 1 & 0 & 0 & 0 & 0\\ 0 & 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 1 & 0\end{bmatrix}=\begin{bmatrix}0.72 & 0.14 & 0 & 0 & 0.14\\ 0 & 1 & 0 & 0 & 0\\ 1 & 0 & 0 & 0 & 0\\ 0 & 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 1 & 0\end{bmatrix},$$

$$\boldsymbol{R}^{(2)}=(\boldsymbol{R}^{(1)})^2=\begin{bmatrix}0.52 & 0.24 & 0.14 & 0.1 & 0\\ 0 & 1 & 0 & 0 & 0\\ 0.72 & 0.14 & 0 & 0 & 0.14\\ 1 & 0 & 0 & 0 & 0\\ 0 & 0 & 1 & 0 & 0\end{bmatrix}。$$

2.3　组合预测

根据上述组合模型计算得到的预测结果见表 3，相对误差波动如图 3 所示。

由于 2017 年处于第 2 种状态，所以考虑矩阵 $\boldsymbol{R}^{(1)}$ 的第 2 行中的最大值，确定 2018 年厦门港集装箱吞吐量仍处于第 2 种状态，由 $\boldsymbol{R}^{(2)}$ 确定 2019 年也处于第 2 种状态，取第 2 种状态区间的相对误差的中值，对灰色 Verhulst 预测值进行修正。根据式（10）得到 2018、2019 年厦门港集装箱吞吐量马尔科夫修正值分别为 $1.142\,15\times10^7$TEU、$1.205\,66\times10^7$TEU。同理可以预测 2020—2022 年数据如表 4 所示。

表 3　组合预测模型数据

Tab. 3　Combined forecast

年份 Year	预测结果 Forecast results	残差 Residual	相对误差 Relative error	年份 Year	预测结果 Forecast results	残差 Residual	相对误差 Relative error	年份 Year	预测结果 Forecast results	残差 Residual	相对误差 Relative error
2006	401.3	0	0	2010	581.7	−0.71	−0.0012	2014	840.7	−16.54	−0.0192
2007	441.6	−21.11	−0.0450	2011	638.8	−7.69	−0.0120	2015	916.5	−1.50	−0.0016
2008	490.3	−13.16	−0.0261	2012	733.8	13.63	0.0189	2016	954.2	−6.80	−0.0071
2009	476.4	8.37	0.0178	2013	777.4	−23.40	−0.0292	2017	1017.1	−20.90	−0.0200

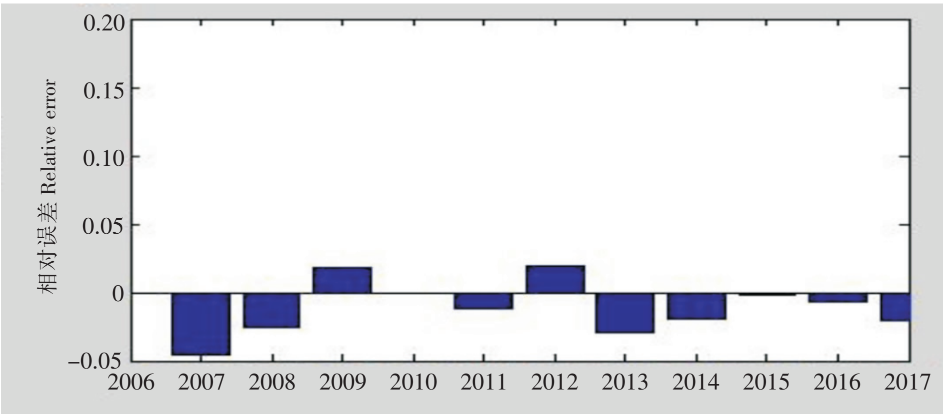

图 3　模型数据对比

Fig.3　Model data comparison

表 4　厦门港 2018—2022 年集装箱吞吐量

Tab. 4　container throughput of Xiamen Port 2018—2022

10^4TEU

年份 Year	2018	2019	2020	2021	2022
集装箱吞吐量 Container throughput	1142.15	1205.66	1266.90	1325.35	1380.77

1）平均相对误差

灰色 Verhulst 预测：$\overline{\delta_{1i}} = \sum_{i=1}^{12} \delta_{1i}/12 = 3.74\%$。

组合预测：$\overline{\delta_{2i}} = \sum_{i=1}^{12} \delta_{2i}/12 = 1.65\%$。

2）模型精度检验

后验差值比：$c = S_2^2/S_1^2 = 0.037 < 0.35$。

小误差概率：$p = \{|q(t) - \bar{q}| < 0.6\,745S\} = 1 > 0.95$。

该组合模型平均相对误差为 1.65%，低于灰色 Verhulst 预测的相对误差 3.74%，精度达到一级

($p \geqslant 0.95, c \leqslant 0.35$)[8]，对厦门港集装箱吞吐量预测数据可信度高。其中各模型拟合情况，如图 4 所示。

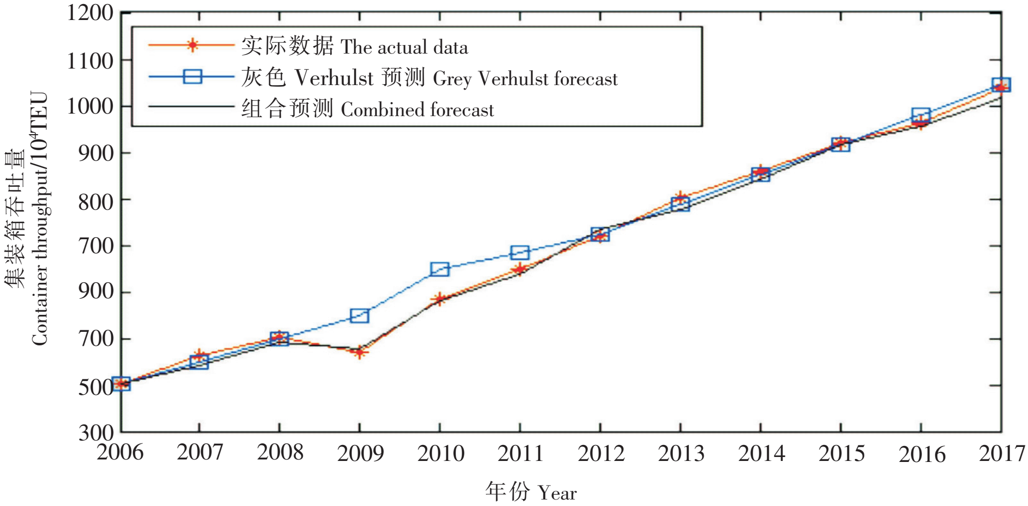

图 4　模型数据对比

Fig.4　Model data comparison

3　结论

本文先利用灰色 Verhulst 预测模型对厦门港历年集装箱吞吐量进行预测，通过比较发现，预测数据与原始数据之间相对误差波动较大，然后引入马尔科夫模型修正相对误差序列。结果表明预测精度提高了，达到一级水平，进一步可预测未来 5 年厦门港集装箱吞吐量，希望对厦门港未来建设规划有一定的参考作用。

［参考文献］

［1］余珍．基于时间序列分析的航道水位预测研究［J］．中国水运，2018，18(10)：147-149.

［2］王厦．基于灰色理论的南通港集装箱吞吐量预测分析［J］．商场现代化，2017，844(7)：18-20.

［3］范莹莹，余思勤．基于 NARX 神经网络的港口集装箱吞吐量预测［J］．上海海事大学学报，2015，36(4)：1-5.

［4］王福建．道路交通事故灰色 Verhulst 预测模型［J］．交通运输工程学报，2006，6(1)：122-126.

［5］同济大学应用数学系．工程数学：数理统计与随机过程区［M］．上海：同济大学出版社，2002.

［6］陈焕珍．基于灰色马尔科夫模型的青岛市粮食产量预测［J］．计算机仿真，2013，30(5)：429-432.

［7］王正新，党耀国，刘思峰．无偏灰色 Verhulst 模型及其应用［J］．系统工程理论与实践，2009，29(10)：138-144.

［8］林云，张荣飞，王岳能．现代统计分析方法［M］．杭州：浙江大学出版社，1991：180-181.

［9］虞盈，兰培真．基于灰色马尔科夫模型的福建辖区船舶交通事故预测［J］．中国航海，2017，40(3)：69-72.

［10］韩婷婷，吴世跃，王鹏军．基于马尔科夫残差修正的瓦斯浓度预测［J］．工矿自动化，2014，40(3)：28-31.

［11］唐普明．基于马尔科夫链状态转移概率矩阵的商品市场状况预测［J］．统计与决策，2015(2)：97-99．DOI:10.13546/j.cnki.tjyic.2015.02.28.

（责任编辑　陈　敏　英文审校　周云龙）