

# Throughput Prediction of Fishing Goods Based on the Grey Multiple Linear Regression Method

Changping Chen<sup>1, #</sup>, Changlu Zhou<sup>1</sup>, Xueda Zhao<sup>2, #</sup>, Yanna Zheng<sup>1</sup>, Xianying Shi<sup>1</sup>

1. School of Marine and Civil Engineering, Dalian Ocean University, Dalian, Liaoning 116023, China

2. School of Science, Dalian Ocean University, Dalian, Liaoning 116023, China

#Email: zxueda@dlou.edu.cn

## Abstract

Based on the grey prediction method and multiple linear regression method, the grey multiple linear regression method was presented. This method was applied to the throughput prediction of fishing goods according to five fishing ports' actual throughput data. The result of comparing the calculating conclusion to the time series one-dimensional linear regression method and grey prediction method proved that the method of calculation and analyzing was more effective and the forecasting precision was higher.

**Keywords:** *Throughput of Fishing Goods; Prediction Method; Multiple Linear Regression Method*

## 基于灰色多元线性回归方法的渔货卸港量预测\*

陈昌平<sup>1</sup>, 周长禄<sup>1</sup>, 赵学达<sup>2</sup>, 郑艳娜<sup>1</sup>, 史宪莹<sup>1</sup>

1. 大连海洋大学海洋与土木工程学院, 大连 辽宁 116023

2. 大连海洋大学理学院, 大连 辽宁 116023

**摘要:** 灰色多元线性回归方法是灰色预测法与多元线性回归方法相结合的一种预测方法, 依据五个渔港实际渔货卸港量资料, 利用该方法对渔货卸港量进行预测, 并与灰色预测法、时间序列的一元线性回归法预测结果相比, 结果表明采用灰色多元线性回归方法的预测精度要优于其他两种方法。

**关键词:** 渔货卸港量; 预测方法; 多元线性回归方法

## 前言

渔港是供渔船和渔业辅助船停泊、使用的港口, 主要用于船舶停靠、锚泊、避风、装卸渔货和补充渔需及生活物资, 并具备渔货冷冻、加工、储运、渔船维修、渔具制造、通信联络、及船员休息、娱乐、医疗等功能。衡量渔港发展水平的因素较多, 如渔货卸港量、水产养殖面积、水产品产量、机动渔船数量、渔业人口数量和渔业经济总产值等<sup>[1]</sup>, 其中渔货卸港量不论是在渔港的前期规划建设中还是在后期的扩建升级中, 都是一项主要决策参数, 科学地对渔货卸港量发展水平进行预测, 对于合理确定渔港建设规模, 并进行渔港规划以及生产运营具有重要的现实意义<sup>[2]</sup>。

目前国内渔港规划中对渔货卸港量的预测方法主要分为时间序列法和概率分析法<sup>[3]</sup>, 其中时间序列法需要较多年份资料, 且计算结果容易出现偏差, 但是由于计算简便, 在各个渔港的工程可行性研究报告中采用较多; 概率分析法是将卸港量理解为随机事件, 综合各项因素做出对未来卸港量的推测, 此方法多用于商港吞吐量的预测, 在渔港卸港量的预测中较少采用。灰色系统理论分析方法由邓聚龙教授于 1982 年在国际经济学会议上提出, 主要是针对系统模型不明确性, 信息不完整情况下, 进行关于系统的关联分析、

\*基金资助: 受国家自然科学基金(51109022)、辽宁省自然科学基金(201202020)、大连理工大学重点实验室开放基金(LP1205)支持资助。

模型建构，并借预测及决策方法来探讨及了解系统。灰色预测模型所需建模信息少，运算方便，在各种预测领域都有着广泛的应用，是处理小样本预测问题的有效工具，该理论不足之处在于针对短期预测的精度高，对于波动较大的曲线进行拟合时精度较低。线性回归分析是研究变量间相关关系的一个统计分支，具有模型精简、预测结果准确、模型解释能力强的优点。在灰色系统理论的基础上结合多元线性回归分析方法目前在粮食产量预测<sup>[4]</sup>、用电量预测<sup>[5]</sup>、就业人数预测<sup>[6]</sup>等方面应用较广。本文采用灰色多元线性回归预测方法对渔港的渔货卸港量进行预测，并与灰色预测法、时间序列的一元线性回归预测方法进行比较，结果表明采用灰色多元线性回归方法的预测精度要优于其他两种方法的预测精度。

## 1 预测方法的基本思路

将渔港统计资料中的各相关因素与渔货卸港量进行相关性分析，选取其中相关性较好的因素与渔货卸港量之间建立多元线性回归模型，然后利用灰色预测理论中的 GM(1.1)模型分别对各个相关因素的数值进行预测，最后将各预测结果带入多元线性回归方程中，得到渔货卸港量的预测值。

## 2 计算实例

### 2.1 多元线性回归方法

石城一级渔港位于福建省莆田市埭头镇石城村北，是莆田市重要的渔港，表 1 为石城一级渔港 1999-2006 年渔业生产统计资料。本算例选取 1999-2005 年共 7 年相关资料，对 2006 年的渔货卸港量进行预测。

表 1 石城一级渔港 1999 年-2006 年渔业生产统计资料

年份	卸港量（吨）	水产品年产量 （吨）	渔船数（艘）	渔船总吨位 （吨）	渔船总功率 （千瓦）
1999	39705	76356	968	6464	31512
2000	58532	104523	1026	7056	33862
2001	71576	122986	1042	7304	34588
2002	88953	142325	1066	7512	35656
2003	89868	152355	1086	7592	35956
2004	90979	157451	1089	7622	36206
2005	100665	159972	1103	7716	36542
2006	101830	160776	1118	7820	36722

以石城渔港 1999-2005 年的水产品年产量、渔船数、渔船总吨位以及总功率与卸港量做相关性分析，根据多元线性回归方法，取显著性水平系数  $\alpha = 0.05$ ，对各影响因素做 Pearson 相关性分析（即选取 Pearson 相关系数作为相关性分析的相关系数），结果如表 2 所示，表中 “\*\*” 表示在 0.01 水平（双侧）上显著相关，N 为用于预测年份的总数。

表 2 Pearson 相关性分析结果

预测对象	参数	年产量	渔船数	渔船总吨位	总功率
卸港量	Pearson 相关性	0.991**	0.985**	0.982**	0.986**
	显著性（双侧）	0	0	0	0
	N	7	7	7	7

由表 2 可知年产量、渔船数、渔船总吨位、总功率与卸港量的相关系数分别为 0.991、0.985、0.982、0.986，各影响因素与卸港量之间的相关系数均较高，且均在 0.01 水平（双侧）上显著相关，可见回归效果较好。运用 Statistical Product and Service Solutions 软件（以下简称 “SPSS”）进行 T 检验，各变量相应的显著性概率值 Sig.均为 0，即 T 检验通过。表 3 为使用 SPSS 软件计算后的模型相关系数汇总，表中 a 为预

测变量；R 为整体相关系数； $\hat{\delta}$  为标准估计的误差。

表 3 模型相关系数汇总

模型	R	R <sup>2</sup>	调整 R <sup>2</sup>	$\hat{\delta}$
1	0.992 <sup>a</sup>	0.984	0.951	4814.891

由表 3 可知，模型中的 4 个自变量与因变量的整体相关系数  $R=0.992$ ，经调整后的整体相关系数  $R^2=0.951$ ，表明所建立的回归模型具有良好的适应性。采用 SPSS 软件计算的变量方差分析结果如表 4 所示，变量非标准化系数结果如表 5 所示。表中 Q 为线性分析平方和，df 为线性分析自由度，MS 为均方，F 为均方比值，Sig 为显著性检验的值，B 为对应多元线性回归方程中各系数， $\delta$  为标准误差。

表 4 多元线性回归方差分析

参数	Q	df	MS	F	Sig.
回归	2.78E+09	4	6.95E+08	29.965	0.033 <sup>a</sup>
残差	4.64E+07	2	2.32E+07		
总计	2.83E+09	6			

表 5 多元线性回归非标准化系数

参数	非标准化系数	
	B	$\delta$
(常量)	-83624.738	297890.167
年产量	0.531	0.498
渔船数	-96.948	494.888
总吨位	-5.02	78.549
总功率	6.597	27.553

根据显著性水平系数  $\alpha=0.05$ ，查 F 检验的临界值表<sup>[7]</sup>，得  $F_{\alpha}(m,n-m-1)=F_{0.05}(4,4)=6.39 < F=29.965$ ，所以检验效果显著。由于各自变量之间存在相关影响，所以对于显著性概率值采用各自变量分别与因变量进行独立对比分析，得出各自变量与因变量显著性概率值均满足要求。根据表 5 中各变量的标准化系数值，可得到预测渔货卸港量的多元线性回归方程如下：

$$Y = -83624.738 + 0.531X_1 - 96.948X_2 - 5.02X_3 + 6.597X_4 \tag{1}$$

式中：Y 为卸港量； $X_1$  为年产量； $X_2$  为渔船数； $X_3$  为总吨位； $X_4$  为总功率。

2.2 灰色预测法

以石城渔港 1999 年至 2005 年渔船总吨位为依据，采用灰色预测法中的 GM(1,1)模型对渔船总吨位进行预测，具体过程如下：

- (1) 累加序列：结果如表 6 所示，表中  $x^{(0)}$  为 1999-2005 年各年渔船总吨位值； $x^{(1)}$  为 1999-2005 年渔船总吨位的累加序列值。

表 6 渔船总吨位累加序列结果

年份	1999	2000	2001	2002	2003	2004	2005
序号	1	2	3	4	5	6	7
$x^{(0)}$	6464	7056	7304	7512	7592	7622	7716
$x^{(1)}$	6464	13520	20824	28336	35928	43550	51266

- (2) 建立矩阵 B,y。

$$B = \begin{bmatrix} -\frac{1}{2}[x^{(1)}(2) + x^{(1)}(1)] & 1 \\ -\frac{1}{2}[x^{(1)}(3) + x^{(1)}(2)] & 1 \\ -\frac{1}{2}[x^{(1)}(4) + x^{(1)}(3)] & 1 \\ -\frac{1}{2}[x^{(1)}(5) + x^{(1)}(4)] & 1 \\ -\frac{1}{2}[x^{(1)}(6) + x^{(1)}(5)] & 1 \\ -\frac{1}{2}[x^{(1)}(7) + x^{(1)}(6)] & 1 \end{bmatrix} = \begin{bmatrix} -99921 & 1 \\ -171721 & 1 \\ -245801 & 1 \\ -321321 & 1 \\ -397391 & 1 \\ -474081 & 1 \end{bmatrix} \quad y = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ x^{(0)}(4) \\ x^{(0)}(5) \\ x^{(0)}(6) \\ x^{(0)}(7) \end{bmatrix} = \begin{bmatrix} 7056 \\ 7304 \\ 7512 \\ 7592 \\ 7622 \\ 7716 \end{bmatrix}$$

(3) 计算逆矩阵  $(B^T B)^{-1} =$

$$\begin{bmatrix} 1.01703020243725e-009 & 2.89892593852377e-005 \\ 2.89892593852377e-005 & 0.992971684640? \end{bmatrix}$$

(4) 根据  $\hat{U} = (B^T B)^{-1} B^T y$  计算估计值  $\hat{a}$  和  $\hat{u}$ 。将  $\hat{a}$  和  $\hat{u}$  的值带入时间响应方程，得时间响应方程为：

$$\hat{x}^{(1)}(k+1) = 431700.5e^{0.01645655k} - 425236.5 \quad (2)$$

式中： $\hat{x}^{(1)}(k+1)$  为拟合值， $k$  为与计算年份相关的系数(0,1,2,3...)。

(5) 计算拟合值  $\hat{x}^{(1)}(i)$ ，根据  $\hat{x}^{(1)}(1) = \hat{x}^{(0)}(1)$ ， $\hat{x}^{(1)}(2) = \hat{x}^{(0)}(2) + \hat{x}^{(0)}(1) \dots$ ，进行后减运算还原，可依次得到  $\hat{x}^{(0)}(i)$  值，并使用残差检验对结果进行检验，相关结算结果如表 7 所示。

表 7 模拟计算值统计表

i	拟合值 $\hat{x}^{(1)}(i)$	模拟计算值 $\hat{x}^{(0)}(i)$	实际值 $x^{(0)}(i)$	残差 E(i)	相对残差 e(i)
1	6464	6464	6464	0	0
2	13627.1	7163.1	7056	-107.1	-1.50%
3	14445	7281.9	7304	22.1	0.30%
4	14684.7	7402.8	7512	109.2	1.50%
5	14928.4	7525.6	7592	66.4	0.90%
6	15176.1	7650.5	7622	-28.5	-0.40%
7	15427.9	7777.4	7716	-61.4	-0.80%

(6) 精度验算与预测。

由表 7 可知，相对残差的数值均小于 10%，可知预测方程式 (2) 可用。进行外推预测：依次令  $k=6,7$  带入预测方程，可得  $\hat{x}^{(1)}(7) = 51265.23$ ， $\hat{x}^{(1)}(8) = 59171.69$ 。因此石城渔港 2006 年的总吨位预测值  $\hat{x}^{(0)}(8) = \hat{x}^{(1)}(8) - \hat{x}^{(1)}(7) = 7906.46$  吨。

## 2.3 灰色多元线性回归方法

根据灰色预测法，可依次推求得到 2006 年石城渔港年产量、渔船数以及渔船总功率的预测值，预测结果如表 8 所示。

表 8 2006 年石城渔港其他参数灰色预测法计算结果

预测参数	年产量 (吨)	渔船数 (艘)	总功率 (千瓦)
预测值	181477.97	1124	37349.1

根据灰色多元线性回归方法原理，将年产量、渔船数、渔船总吨位、总功率 4 个预测结果带入所建立的多元线性回归方程式 (1) 中，可得到石城渔港 2006 年渔货卸港量的预测值为 110472.1 吨。

## 2.4 计算方法比较

分别采用灰色预测法、时间序列一元线性回归方法及灰色多元线性回归方法对石城渔港 2006 年的渔货卸港量进行预测值，3 种方法预测结果比较如表 9 所示。

由表 9 可知，采用灰色多元线性回归方法预测的石城渔港 2006 年渔货卸港量与实际值相比，相对误差

最小。

表 9 石城渔港渔货卸港量预测结果比较

参数	2006 年卸港量数值 (吨)	误差	相对误差
灰色多元线性回归方法	110472.1	8642.1	8.49%
灰色预测法	112564	10734	10.54%
时间序列一元线性回归法	115192.1	13362.1	13.12%
实际卸港量	101830		

## 2.5 灰色多元线性回归方法的进一步验证

采用灰色多元线性回归方法分别对沙埕中心渔港、东埔一级渔港、港尾一级渔港、澳角一级渔港的渔货卸港量进行预测，并与单独使用灰色预测法与时间序列法的预测结果进行比较，比较结果见表 10。

表 10 灰色多元线性回归方法对各渔港渔货卸港量预测结果比较

渔港	项目	实际值 (吨)	灰色多元线性回归方法 (吨)	灰色预测法 (吨)	时间序列一元线性回归法 (吨)
沙埕中心渔港	2011 年卸港量		169186	207850	192307.09
	误差	148300	20886	59550	44007.09
	相对误差		14.08%	40.16%	29.67%
东埔一级渔港	2005 年卸港量		50650.04	52706.3	53427.65
	误差	51000	-349.96	1706.3	2427.65
	相对误差		-0.69%	3.35%	4.76%
港尾一级渔港	2005 年卸港量		74200.8	74026	72362.5
	误差	74638	-437.2	-612	-2275.5
	相对误差		-0.59%	-0.82%	-3.05%
澳角一级渔港	2004 年卸港量		31966.23	30194.83	30937.3
	误差	33800	-1833.77	-3605.17	-2862.7
	相对误差		-5.43%	-10.67%	-8.47%

通过以上五个渔港卸港量预测结果的对比可以发现，采用灰色多元线性回归方法比单一使用灰色预测法及时间序列法的预测结果误差更小，具有较好的实际应用价值。

## 3 结论

本文将灰色预测理论和多元线性回归理论相结合，提出了一种渔货卸港量新的预测方法，并通过与其他预测方法进行比较，表明采用该方法进行渔港渔货卸港量的预测，具有较好的稳定性，预测结果更接近实际值。采用灰色多元线性回归方法并借用 SPSS、MTALAB 等软件进行渔货卸港量的预测，可避免以往对于卸港量粗略估算模式，将对于优化渔港的布局扩建，进一步增强渔港建设科学性以及选取优势渔港带动当地渔业生产发展具有积极的实践意义。

## REFERENCES

- [1] Huihui Sun. Study on Coastal Fishing Ports Distribution of Shandong Province [D]. Ocean University of China. 2009
- [2] Jinsong Gui. Prediction Methods Inquire of Fish-unloading Capacity Developing Level in a Fishing Port [J]. Journal of Dalian Fisheries College, 1995, 10(3): 75-78
- [3] Zijian Guo. Port Planning and Layout [M]. Beijing: China Communications Press, 2011.6

- [4] Juncheng Li, Guohua Chen, Xiaofang Shi. Grain Yield Forecasting Based on Grey Multivariable Linear Regression[J].Journal of Anhui AgriSci, 2010, 38(16): 8281-8282
- [5] Hao Ding, Rong Rong. Electricity Consumption Forecasting of Shandong Based on Multivariate Linear Regression Model and Gray Theory [J]. HENAN SCIENCE,2013,31(9):1535-1539
- [6] Bianping Su, Yanping Cao. Analysis of Multiple Linear Regression Based on Grey Systems Theory [J]. MATHEMATICS IN PRACTICE AND THEORY, 2006,36(8): 219-222
- [7] Shucai Tai, Wenyu Sun, Juanjuan He, eds. Application of Mathematical Statistics (2nd edition)[M]. Wuhan: Wuhan University Press, 2005.7

#### 【作者简介】



陈昌平（1971- ），男，汉族，博士，副教授，研究方向为港口、海岸及近海工程，1991-1995 年就读于大连水产学院，获港口航道与治河工程专业学士学位；1999-2002 年就读于大连理工大学，获水利工程硕士学位，2003-2010 年就读于大连理工大学，获港口、海岸及近海工程博士学位。

Email: ccp@dlou.edu.cn。