

# 基于组合方法的三亚机场客流量预测<sup>①</sup>

刘夏<sup>1</sup>, 陈磊<sup>1</sup>, 李苑辉<sup>1</sup>, 杨萍<sup>2</sup>, 陈明锐<sup>3</sup>

<sup>1</sup>(三亚航空旅游职业学院, 三亚 572000)

<sup>2</sup>(三亚凤凰国际机场有限公司 综合管理部, 三亚 572000)

<sup>3</sup>(海南大学 信息科学技术学院, 海口 570228)

**摘要:** 准确地预测机场客流量对机场的建设和发展有着至关重要的作用. 在三亚机场 2005-2015 年客流量数据的基础上, 分别采用 Holt-Winter 季节模型、ARMA 和线性回归模型分别对三亚机场 2016-2017 年的客流量进行了预测. 为了减少预测误差提高预测精度, 采用了组合加权的方法对数据进行了组合预测. 经验证, 该方法可以作为有效预测机场客流量的一种方法.

**关键词:** 机场; 客流量预测; 季节模型; 自回归平滑模型; 线性回归; 组合

## Sanya Airport Passenger Flow Forecast Based on Combination Forecast Method

LIU Xia<sup>1</sup>, CHEN Lei<sup>1</sup>, LI Yuan-Hui<sup>1</sup>, YANG Ping<sup>2</sup>, CHEN Ming-Rui<sup>3</sup>

<sup>1</sup>(Sanya Aviation Tourism College, Sanya 572000, China)

<sup>2</sup>(Comprehensive Management Department, Sanya Phoenix International Airport Company Limited, Sanya 572000, China)

<sup>3</sup>(College of Information Science & Technology, Hainan University, Haikou 570228, China)

**Abstract:** Accurate passenger-flow forecast has always been playing the crucial role on the construction and development of the airport. This passage based on the data of passenger flow from 2005 to 2015 at Sanya Phoenix International Airport is proposed to predict the flow from 2016 to 2017 by adopting Holt-Winter seasonal model, ARMA model and liner regression model. In order to promote the forecasting accuracy, the portfolio weight method is used to forecast the result and reduce the error. According to the verification, this measure is one of the effective ways to predict the passenger flow.

**Key words:** airport; passenger flow; holt-winter; ARMA; liner regression; combination

### 1 引言

准确地预测机场客流量数据, 对机场运力安排、未来的发展和建设、功能的规划都有着至关重要的作用. 因此, 准确预测机场客流量就成为了机场经营管理中的一个重要课题, 是机场资源有效配置的基本根据. 预测客流量的方法很多, 据不完全统计, 世界上大约有约 300 种方法, 其中有 150 多种比较成熟, 30 多种比较常用, 10 多种使用比较普遍, 但根据不同的标准, 大体分为两类: 一类是线性理论和非线性理论, 另一类是定性预测法和定量预测法, 但总体都包括: 时间序列模型、灰色预测模型、专家预测模型、指数平滑法、神经网络、支持向量机、趋势外推法、回归分

析法等.

国内的学者开展过关于机场客流量预测的研究, 取得了相应的成果. 南京航空航天大学的演克武、朱金福提出了支持向量机的回归模型, 并与 BPANN 和线性回归算法这两种预测方法进行了对比, 证明支持向量机回归算法能获得最小的相对误差, 是有效的一种航空客流量预测方法<sup>[1]</sup>. 西安交通大学的屈拓将灰色模型和 BP 神经网络相结合, 利用灰色模型对线性变化部分进行预测, 然后采用 BP 神经网络对非线性变化部分进行预测, 并对预测误差进行补偿, 解决了单一预测模型存在的缺陷, 提高了机场旅客吞吐量预测精度<sup>[2]</sup>. 合肥工业大学的陈荣、梁昌勇、陆文星等人

① 基金项目: 2014 年三亚市院地合作科技项目(2014YD52)

收稿时间: 2015-12-02; 收到修改稿时间: 2016-01-14 [doi:10.15888/j.cnki.csa.005268]

提出一种季节支持向量回归(SSVR)和粒子群算法(PSO)结合模型,实现对旅游客流量的预测,该模型预测精度明显高于SVR-PSO、SVR-GA、BPNN、ARIMA等方法<sup>[3]</sup>.中国民航飞行学院的黄邦菊、林俊松、郑潇雨等人建立了机场旅客吞吐量的多元线性回归预测模型,并利用时间序列法对所得的预测值进行验证,准确的预测出机场未来年的旅客吞吐量<sup>[4]</sup>.中国民航大学的关静依据灰色预测和支持向量机的特点,提出了一种将两种预测方法相结合的灰色支持向量机,结合旅客吞吐量的预测结果,对比了灰色预测模型、支持向量机和灰色支持向量机的预测结果,验证了灰色支持向量机的预测精度高和预测结果准确可靠<sup>[5]</sup>.中国民航飞行学院的景崇毅提出了基于迭加趋势的航线季节客运需求分析方法,根据航线季节客运量历史数据,构建航线季节客运量趋势变动模型,通过OLS方法估计直线趋势方程参数,进一步考虑季节变化对客运需求的影响,建立航线季节客运需求分析模型<sup>[6]</sup>.中国民航大学的陈玉宝、曾刚采用多元线性回归模型和时间序列趋势外推模型,对首都机场2012-2016年的客流量进行了预测,为提高精度,又采用了组合加权方法对预测结果进行组合预测,提高了准确性,减少了预测误差<sup>[7]</sup>.沈阳工业大学的田中大、李树江、王艳红等人提出了一种基于经验模式分解和时间序列分析的网络流量预测方案,具有更好的预测效果和精度<sup>[8]</sup>.华东师范大学的杜刚、刘娅楠处理了港口集装箱吞吐量月度数据中的季节性波动,用季节时间序列模型对集装箱的吞吐量进行了季节性处理,提高了预测精度<sup>[9]</sup>.

Holt-Winter 将具有线性趋势、季节变动和随即波动的时序进行分解研究,结合指数平滑法,分别对长期趋势、趋势外增量和季节波动进行估计,建立预测模型并外推预测值. ARMA 模型(Auto-Regressive and Moving Average Model)是研究时间序列的重要方法,由自回归模型(简称 AR 模型)与滑动平均模型(简称 MA 模型)为基础“混合”构成,主要基于对平稳随即时间序列进行分析来建立模型,其形式简单,对数据进行拟合较为方便,便于分析数据的结构和内在性质,在最小方差的意义下进行最佳预报和控制,它是一种精确度较高的短期预测模型,因为阶数可以依据情况调整,也显得比较灵活,但是所需要的历史数据量较大(一般在 50 个以上).线性回归是利用数理统计中回

归分析,来确定两种或两种以上变量间相互依赖的定量关系的一种统计分析方法,运用十分广泛.回归模型中,一元回归是最简单且稳健的,但描述复杂系统的行为时往往乏力,因此基于多元回归的预测技术更为常见.传统的多元回归模型一般是线性的.由于可能存在的非显著变量以及各自变量之间的相关关系,会导致回归的正规方程组出现严重的病态,影响到回归方程的稳定性,所以多元线性回归面临的一个基本问题是寻找“最优”回归方程.神经网络(Neural Networks, NN)是由大量的、简单的处理单元(称为神经元)广泛地互相连接而形成的复杂网络系统,它反映了人脑功能的许多基本特征,是一个高度复杂的非线性动力学习系统,其逼近效果好,计算速度快,不需要建立数学模型,精度高.具有强非线性拟合能力.但是无法表达和分析被预测系统的输入和输出间的关系,预测人员无法参与预测过程,收敛速度慢,难以处理海量数据,得到的网络容错能力差,算法不完备<sup>[10]</sup>.

## 2 数据来源及描述性分析

### 2.1 数据的选取

本文选取由三亚凤凰国际机场提供的从 2005 年 1 月到 2015 年 10 月的月度客流量共 130 个数据,分别用 Holt-Winters 季节性预测模型, ARMA 模型,一元线性回归进行建模预测,最后利用这三种方法进行组合预测,得出预测效果,对未来两年三亚机场的客流量进行预测.

### 2.2 描述性分析

利用 Eviews6.0 软件做出三亚机场客流量的时序图,观察客流量的变化趋势,得到的结果如图 1 所示.

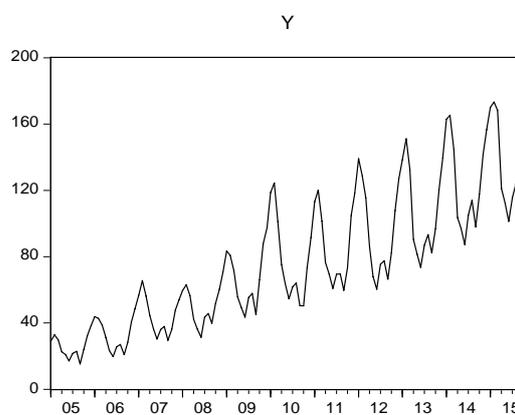


图 1 客流量时序图

### 3 基于Holt-Winter季节模型预测

Holt-Winter 季节模型主要把长期趋势, 季节变动和随机变动的时间序列进行分解, 然后和指数平滑法进行结合, 分别对长期趋势, 季节变动进行估计, 然后建立模型, 进行样本外预测. 本文主要选用 Holt-Winter 季节乘法模型, 该方法适用具有长期趋势和乘法季节变化的序列. 其平滑序列的计算公式:

$$\hat{y}_{t+k} = (a_t + b_t k) S_{t+k}, \quad t = s+1, s+2, \dots, T \quad (1)$$

其中  $a_t$  表示截距,  $b_t$  表示斜率,  $a_t + b_t k$  表示长期趋势,  $S_t$  表示乘法模型的季节因子,  $s$  表示季节周期长度(本文月度  $s = 12$ ).

该模型需要三个系数来给出季节因子第一年的初值, 截距和斜率的初值, 这三个系数的定义如下:

$$\begin{aligned} a_t &= \alpha \frac{y_t}{S_{t-s}} + (1-\alpha)(a_{t-1} + b_{t-1}) \\ b_t &= \beta(a_t - a_{t-1}) + (1-\beta)b_{t-1} \\ S_t &= \gamma \frac{y_t}{a_t} + (1-\gamma)S_{t-s} \end{aligned} \quad (2)$$

其中  $k > 0, \alpha, \beta, \gamma$  为阻尼系数在  $0 \sim 1$  之间, 如果  $t = T$ , 其预测值计算公式为:

$$\hat{y}_{T+k} = (a_T + b_T k) S_{T+k-s} \quad (3)$$

其中,  $S_{T+k-s}$  为样本数据最后一年的季节因子. 利用 Eviews6.0 软件对上述数据进行预测得到的预测参数估计结果如表 1 所示.

表 1 预测参数估计结果

Parameters:	Alpha	0.9000
	Beta	0.0000
	Gamma	0.0000
Sum of Squared Residuals		2860.531
Root Mean Squared Error		4.690851

根据表 1 得到样本预测的结果, 得到的结果如图 2 所示.

根据图 2 可得基于 Holt-Winter 季节乘法模型三亚机场的实际客流量也预测客流量几乎是重叠, 计算预测的平均绝对误差其计算公式:

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{\hat{y}_i - y_i}{y_i} \right| \times 100 \quad (4)$$

计算得到  $MAPE = 4.54$  说明 Holt-Winter 乘法模型预测的效果比较理想.

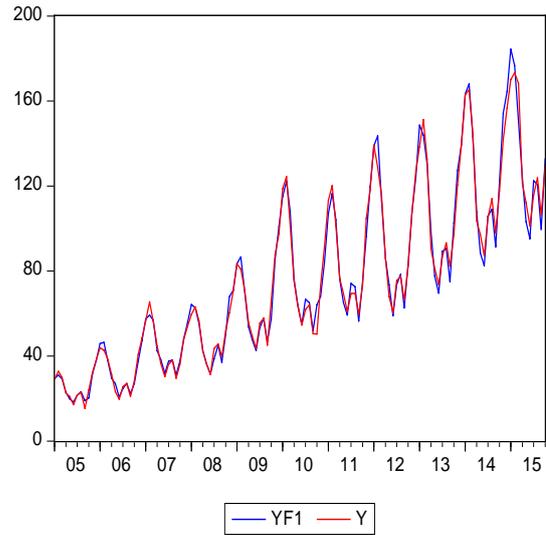


图 2 原序列与预测序列的时序图

### 4 基于ARMA模型预测

根据图 1 可以看出样本内的机场客流量具有明显的季节变动, 在建立 ARMA 模型之前需要对机场客流量  $Y$  进行季节调整, 本文选用 X12 季节调整方法<sup>[11]</sup>, 利用 Eviews6.0 进行季节调整后的机场客流量的时序图如图 3 所示.

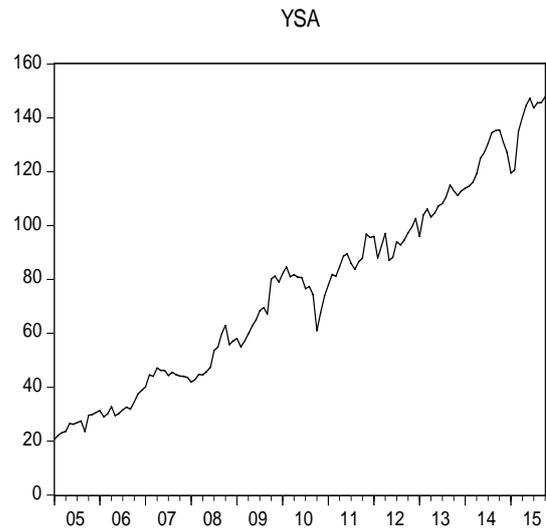


图 3 消除季节变动客流量时序图

根据图 3 可以得到经过季节调整之后, 机场客流量不受季节变动因素影响, 下面对调整后的客流量序列建立 ARMA 模型.

### 4.1 序列的平稳性检验

采用单位根检验来检验时间序列平稳性, 选用 ADF 检验, 运用 Eviews6.0 软件对 YSA 进行 ADF 单位根检验的结果如下表 2 所示.

表 2 单位根的 ADF 检验结果

变量	ADF 检验值	P 值	1% 临界值	5% 临界值	10% 临界值	检验结果
YSA	0.0826	0.9633	-3.4816	-2.8839	-2.5788	非平稳
D(YSA)	-11.3324	0.0000	-3.4820	-2.8841	-2.5789	平稳

根据表 2 可得调整后的机场客流量序列 YSA 原序列存在单位根, 即非平稳序列, 一阶差分之后不存在单位根即是平稳序列. 下面对 YSA 序列进行一阶差分之后建立 ARMA 模型.

### 4.2 ARMA(p,q)模型的建立与识别

#### 4.2.1 建立 ARMA(p,q)模型的代数表达式

$$y_t = \alpha_0 + \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \dots + \alpha_p y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q} \quad (5)$$

#### 4.2.2 对模型中参数 p 和 q 进行识别

利用 Eviews 6.0 得出一阶差分后 YSA 序列的自相关函数 ACF 和偏自相关函数 PACF 图, 对 p 和 q 进行初步的判断, 对 D(YSA)建立的模型为 ARMA(1,1), 其表达式为:

$$D(YSA)_t = c + \alpha_1 D(YSA)_{t-1} + \varepsilon_t + \beta_1 \varepsilon_{t-1} \quad (6)$$

得到下面 ARMA(1.1)建模估计结果.

表 3 ARMA(1.1)建模的回归结果

变量	回归系数	标准差	T 统计量	P 值
C	0.970238	0.064923	14.94442	0.0000
AR(1)	0.817874	0.057371	14.25589	0.0000
MA(1)	-0.983407	0.016533	-59.48068	0.0000
$R_a^2$		0.059414		
F 统计量		5.011116		
		(0.008064)		

根据表 3 可得在 0.05 的显著水平下, ARMA(1.1)的各项系数均是显著的, 且总体方程的概率 P 值为 0.008064 小于 0.05 的显著水平, 也是显著的. 得到具体的方程为:

$$D(YSA)_t = 0.176567 + 0.818053 D(YSA)_{t-1} - 0.983371 \varepsilon_{t-1} \quad (7)$$

利用上述建立的模型对样本数据进行预测. 得到的结果如图 4 所示.

根据图 4 可得基于 ARMA(1,1)三亚机场的实际客流量也预测客流量几乎是重叠, 计算得到预测的平均绝对误差  $MAPE = 3.59$  说明 ARMA(1,1)模型预测的

效果比较理想.

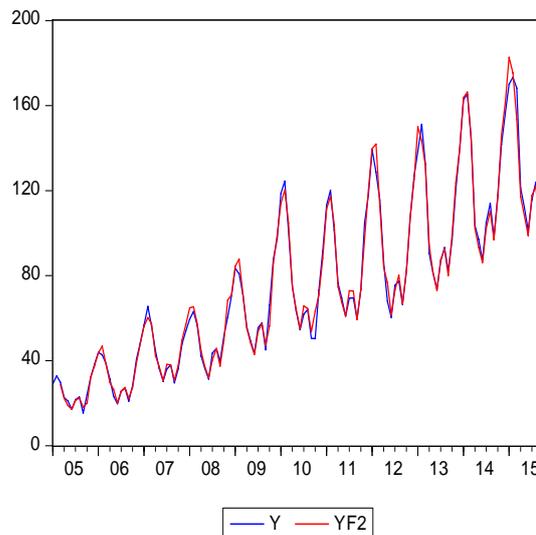


图 4 原序列与预测序列的时序图

## 5 基于线性回归模型预测

根据图 3 可以得到调整后的机场客流量几乎呈随时间呈线性增长趋势, 下面就建立客流量随时间的一元线性回归模型, 进行拟合预测. 建立的一元线性回归模型为:

$$YSA_t = \beta_0 + \beta_1 T + \varepsilon_t \quad T = 1, 2, \dots \quad (8)$$

其中  $\varepsilon_t$  表示随机误差项,  $T$  表示为时间 2005 年 1 月为 1, 后面依次类推. 利用 Eviews6.0 得到的估计结果如表 4 所示.

表 4 模型估计结果

变量	回归系数	标准差	T 统计量	P 值
C	15.15110	1.081433	14.01020	0.0000
T	0.935417	0.014326	65.29591	0.0000
$R_a^2$		0.970625		
F 统计量		4263.556		
		(0.0000)		

根据表 4 可得在 0.05 的显著水平下, 回归系数均是显著的, 且总体方程 F 统计量的概率 P 值为 0.0000 小于 0.05 的显著水平, 说明总体方程是显著的. 另一方面调整后的拟合优度值  $R_a^2 = 0.970625$ , 说明上述建立的回归方程的拟合效果非常好. 得到回归方程为:

$$YSA = 15.1511 + 0.935417T \quad (9)$$

利用建立的回归方程进行预测得到的预测结果如图 5 所示.

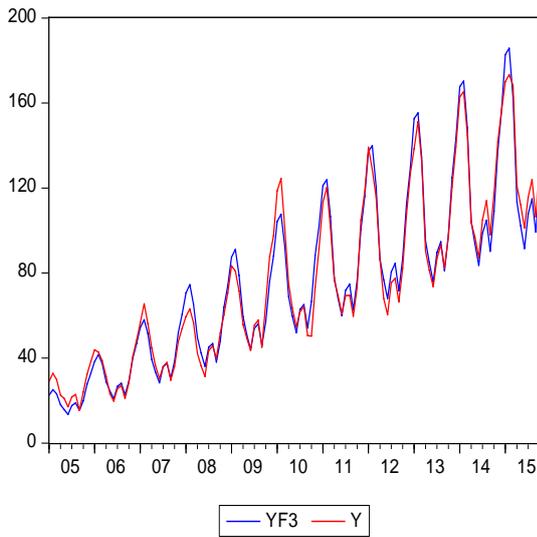


图 5 原序列与预测序列的时序图

根据图 5 可得基于回归模型三亚机场的实际客流量也预测客流量几乎是重叠, 计算得到预测的平均绝对误差  $MAPE = 7.62$  说明回归模型预测的效果比较理想.

### 6 基于组合方法预测

经过上述三种方法预测可以看出 ARMA 模型的预测误差最小, 而线性回归模型的预测误差最大, Holt-Winter 季节乘法模型的预测误差介于二者之间. 且 Holt-Winter 季节乘法模型的预测值偏小而 ARMA 模型和线性回归模型的预测误差偏大. 采用一种方法可能存在偏差, 本文最后采用组合模型, 因为回归的方法预测的误差较大, 所以最后只选择 Holt-Winter 季节乘法模型和 ARMA 模型进行组合预测, 即是取其上述二个模型平均绝对误差的比值作为权重进行预测, 其中 Holt-Winter 季节乘法模型的权重为 0.5584, ARMA 模型的权重为 0.4416. 可以得到组合预测的公式:

$$Y_4 = \partial_1 Y_1 + \partial_2 Y_2 \tag{10}$$

$\partial_1, \partial_2$  表示权重.

利用上述二种方法的组合来对三亚机场客流量进行预测得到的结果如图 6 所示.

根据图 6 可以看出组合预测的预测值介于上述二种预测方法得到的值之间, 其预测的平均绝对误差

$MAPE = 3.98$  下面给出四种方法预测的平均绝对误差. 得到结果如表 5 所示.

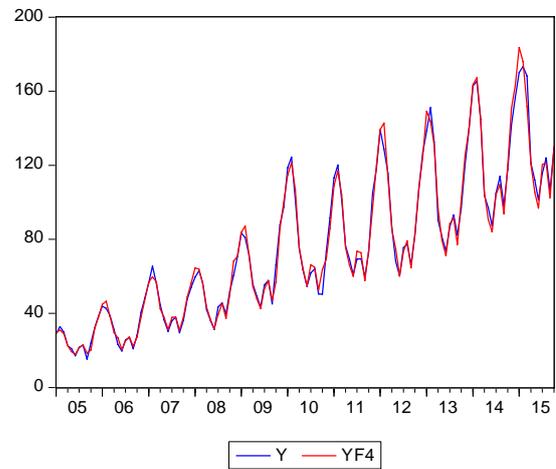


图 6 三种方法与组合预测的比较

表 5 各方法预测的平均绝对误差

预测方法	Holt-Winter 季节模型	ARMA 模型	回归模型	组合预测模型
MAPE	4.54	3.59	7.62	3.98

根据表 5 可以看出, ARMA 模型的预测效果最好, 其实是组合预测模型, 然后是 Holt-Winter 季节模型, 回归模型的预测效果相对较差. 但是四种预测方法算出的平均绝对误差均小于 10, 说明四种方法总体的预测效果比较理想. 四种方法预测未来机场两年的客流量的预测值, 结果如表 6 所示, 折线图如图 7 所示.

表 6 未来机场客流量预测

时间	Holt-Winter 季节模型	ARMA 模型	回归模型	组合预测模型
Nov-15	169.19	154.22	149.13	162.58
Dec-15	194.19	177.52	171.66	186.83
Jan-16	227.13	206.16	199.36	217.87
Feb-16	233.57	213.06	206.02	224.51
Mar-16	201.73	184.41	178.32	194.08
Apr-16	150.00	135.30	130.83	143.51
May-16	127.29	116.42	112.57	122.49
Jun-16	108.93	99.96	96.65	104.97
Jul-16	132.54	121.32	117.31	127.59
Aug-16	137.42	126.01	121.83	132.38
Sep-16	110.54	103.63	100.19	107.49
Oct-16	138.85	127.43	123.20	133.81
Nov-16	181.07	166.83	161.29	174.78

Dec-16	207.75	191.93	185.56	200.76
Jan-17	242.89	222.80	215.39	234.02
Feb-17	249.69	230.13	222.48	241.05
Mar-17	215.57	199.09	192.47	208.29
Apr-17	160.23	146.00	141.14	153.95
May-17	135.92	125.56	121.38	131.35
Jun-17	116.28	107.76	104.17	112.52
Jul-17	141.43	130.73	126.38	136.71
Aug-17	146.59	135.72	131.19	141.79
Sep-17	117.87	111.56	107.84	115.08
Oct-17	148.01	137.12	132.54	143.20

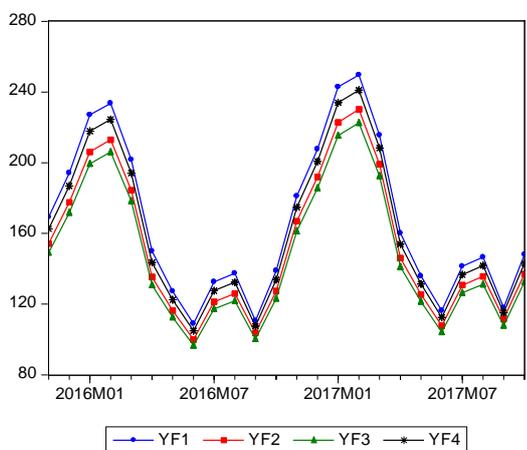


图 7 未来机场客流量预测

### 7 结语

Holt-Winter 季节乘法模型的预测值相对于其他方法, 预测值偏高, ARMA 模型预测值次之, 而回归模型的预测值相对于其他方法偏低, 组合预测方法介于 Holt-Winter 季节乘法模型和 ARMA 模型之间。

在“一带一路”、国家海洋强国、海南国际旅游岛三大战略的前提下, 三亚作为海南省南部的重要城市, 承载着实现国家三大战略的历史使命。三亚凤凰国际

机场作为南中国重要的交通港, 客流量总体呈上升的趋势, 根据预测, 旅客吞吐量将会在两年内突破 2000 万人次大关。因此, 在三大战略的背景下, 机场改、扩建迫在眉睫, 建设新机场也势在必行, 新机场的总体规划也应长远考虑, 加快项目落地和推进。

### 参考文献

- 1 演克武,朱金福.基于支持向量机回归算法的航空公司客流量预测研究.企业经济,2010,23:88-90.
- 2 屈拓.组合模型在机场旅客吞吐量预测中的应用.计算机仿真,2014,29(4):108-111.
- 3 陈荣,梁昌勇,陆文星,等.基于季节 SVR-PSO 的旅客客流量预测模型研究.系统工程理论与实践,2013,33(1):1-7.
- 4 黄邦菊,林俊松,郑潇雨,等.基于多元线性回归分析的民用运输机场旅客吞吐量预测.数学的实践与认识,2013,43(4):172-178.
- 5 关静.基于灰色支持向量机的民航旅客吞吐量预测.大连交通大学学报,2013,34(6):41-43.
- 6 景崇毅.基于迭加趋势的航线季节客运需求分析方法.中国民航飞行学院学报,2014,25(2):5-7,11.
- 7 陈玉宝,曾刚.基于组合预测方法的民航旅客吞吐量预测研究.中国民航大学学报,2014,32(2):59-64.
- 8 田中大,李树江,王艳红等.经验模式分解与时间序列分析在网络流量预测中的应用.控制与决策,2015,30(5):905-910.
- 9 杜刚,刘娅楠.季节性变动影响下的上海港即装下吞吐量预测.华东师范大学学报(自然科学版),2015(1):234-239.
- 10 方英国,王芬.时间序列预测方法综述.浙江树人大学学报,2006,6(2):61-65.
- 11 陈飞,高铁梅.结构时间序列模型在季节调整方面的应用——与 X-12 季节调整方法比较分析.系统工程理论与实践,2007,(11):7-14.