

罗慧, 刘杰, 巩在武, 等. 西安世园会客流影响及预测的气象计量经济分析[J]. 气象, 2012, 38(11): 1408-1416.

西安世园会客流影响及预测的气象计量经济分析^{* 1}

罗 慧¹ 刘 杰^{2,3} 巩在武⁴ 沈 瑾¹ 杨胜利¹ 杜岩岫⁵
肖 波¹ 邓小丽¹ 马小梅⁶ 姚东升⁷ 高武虎⁷

1 陕西省西安市气象局, 西安 710016

2 中国社会科学院研究生院, 北京 100005

3 国家气象中心, 北京 100081

4 南京信息工程大学经济管理学院, 南京 210044

5 西安世园执委会世园投资集团投资策划处, 西安 710024

6 中国气象局气象干部培训学院, 北京 100081

7 陕西省防雷中心, 西安 710015

提 要: 以 2011 年西安世界园艺博览会(简称世园会)的入园客流数量和逐日气象资料为基础, 以最高温度、平均风速和平均相对湿度为定量解释变量, 以降水、节假日、每一天和月份为虚拟解释变量, 引入随机误差项的 AR(2)结构, 通过逐步优化的建模思想, 构建了 2011 年西安世园会入园客流的气象计量经济模型。以该模型为核心, 以网络技术和数据库技术为基础, 构建世园会入园客流数量预测的气象服务业务化系统。2011 年 8 月底至 10 月 22 日, 气象部门和世园会管理层同时开展了基于网络的业务化应用, 运用逐日温度和降水等天气预报信息, 结合入园票务政策变化、优惠措施出台、指定日特别活动等信息, 联合开展对逐日、未来一段时间和整个运营期的可能入园客流数量的定量预测, 为世园执委会部署和控制客流数量、确保世园会安全运营提供重要参考依据。结论表明, 气象计量经济学模型能很好地解释气象因素、节假日、每一天以及不同月份变化对入园客流的影响, 准确模拟入园客流数量的逐日与逐月变化特征, 能较好地预测未来两周的入园客流人数。在 8 月提前两个月准确预测出西安世园会可能入园客流总人数逾 1600 万人, 与实际人数约 1573 万很接近。

关键词: 西安世界园艺博览会, 气象服务, 气象计量经济模型, 气象服务系统, 客流预测

Econometric Analysis on the Effects of Meteorological Factors on Daily Visitors to Expo 2011 Xi'an and Its Application for Predicting the Number of Visitors

LUO Hui¹ LIU Jie^{2,3} GONG Zaiwu⁴ SHEN Jin¹ YANG Shengli¹ DU Yanxiu⁵
XIAO Bo¹ DENG Xiaoli¹ MA Xiaomei⁶ YAO Dongsheng⁷ GAO Wuhu⁷

1 Xi'an Meteorological Bureau of Shaanxi Province, Xi'an 710016

2 Graduate School of the Chinese Academy of Social Sciences, Beijing 100005

3 National Meteorological Centre, Beijing 100081

4 College of Economics and Management, Nanjing University of Information Science and Technology, Nanjing 210044

5 Investment Planning Department, Xi'an International Horticultural Exposition Investment Group Co., Ltd, Xi'an 710024

6 CMA Training Centre, Beijing 100081

7 Shaanxi Provincial Lightning Protection Centre, Xi'an 710015

Abstract: An applied meteorological econometric model was built by using real-time data of daily number of visitors and meteorological factors from April 28 to August 15, 2011. The factors in the model included the number of

* 国家自然科学基金(70901043 和 71171115)、教育部人文社科基金(09YJC630130)、公益性行业(气象)科研专项(GYHY201106037)及陕西创新基金(2011M-62)共同资助

2012 年 1 月 13 日收稿; 2012 年 8 月 25 日收修定稿

第一作者: 罗慧, 主要从事气象服务效益评估与气象防灾减灾等研究. Email: luohui6812@sina.com

daily visitors, the daily weather data during this period, and the quantitative explanatory variables of the daily high temperature, mean wind speed and mean relative humidity, meanwhile, the virtual explanatory variables were considering the precipitation, holidays, daily variables, and monthly variables. Besides, a random error factor of AR (2) was introduced. With the above, the meteorological econometric model was created. Furthermore, the special meteorological service system for predicting the number of visitors of the Expo was developed and the operational tests as well as application were carried out. The system for predicting the number of daily visitors has been built and applied in predicting the number of visitors with quantitative analysis on different weather conditions, and this helped the decision-makers of the Executive Committee of Expo 2011 Xi'an control the number of individual visitors in daily and safely operating. The system has been proved to be successful in the short-range and medium-range predictions which simulated the characteristics of the daily and monthly changes of the numbers, and predicted the number of visitors in 15 days accurately, while the 1-7 d prediction was even better than that of 8-15 d. The result offered continuous support to the Executive Committee for decision-making. According to the prediction of October 2011, 10 million visitors were expected till 22 of October, and the total (including the securities, staff) would exceed 16 million, which was very close to the actual total number, roughly 15.73 million.

Key words: the 2011 Xi'an International Horticultural Exposition, meteorological services, meteorological econometric model, meteorological service system, predicting the number of visitors

引 言

我国重大社会经济活动往往由中央政府或地方政府筹备并举办,具有规格高、影响大、社会效应强等显著特点,也日益受到多变的气象因素的影响,气象保障服务成为活动的重要组成部分。开展并改进气象服务评估工作是气象服务工作的重要基础工作,也是一项复杂的工作。正是因为气象服务对象和用户的多样性、服务效益的综合性、效益评价的差异性,使得仅从一个方面、一类用户、一种方法难以获得科学、客观的气象服务效益评估结论。我国学者从奥运气象服务满意度 CSIWS 评估、条件价值评估法、模糊数学、多元回归模型、SVAR 模型和专家调查法等多种方法、多角度分析研究服务评估问题^[1-8]。国内运用计量经济学方法研究气象与社会经济的定量关系的成果并不多,但为气象学与社会科学的交叉研究提供了新思路和新方法^[9-12]。高温、雷电、暴雨和大风等高影响天气对北京奥运会、上海世博会和西安世园会等大型的或者长期活动的客流人数、交通状况和安全运行有着很大影响。根据已往综合性世界园艺博览会举办国的经验,举办国都十分关注客流量。把实际客流量作为举办成功与否的重要标志。王小平等^[8]开展了多元回归模型的上海世博会客流量预测分析,比以前常用的渗透率模型和引力模型考虑因素更为全面,但由于只从

定性角度分析预测世博参观人数,还存在一定的局限性。

2011年4月28日西安世界园艺博览会(简称世园会)盛大开园,本文以开园日到8月15日的世园会入园客流数量和气象数据为基础构建气象计量经济模型,并运用网络技术和数据库技术建立了业务化气象服务系统。2011年8月底至10月22日,气象部门和世园管理层同时开展了基于网络的业务化应用,运用逐日温度和降水等天气预报信息,结合入园票务政策变化、优惠措施出台、指定日特别活动等信息,联合开展对逐日、未来一段时间和整个运营期的可能入园客流数量的定量预测。该系统的研发应用使得在世园气象服务中不仅开展气象预报服务,还借助智力资源优势将气象信息作为重要信息融入到西安世园指挥运行体系中,从单纯气象要素预报的定性服务向入园客流数量预测的定量服务迈进,开展了世园管理决策层所高度关注的客流预测控制的气象服务工作。

1 2011 西安世园会入园客流预测气象服务系统框架

影响世园会入园客流人数的因素主要有非气象要素和气象要素两大类。世园会客流数量预测气象服务系统的设计是以气象计量经济学模型为核心,以计算机网络技术与数据库技术为基础而建立起来

的一个气象服务网络平台。计量经济学(Econometric)是综合数学、统计推断及经济理论等工具应用于定量分析的新兴学科,计量经济分析的根本目的是研究结构规律、进行较高精度的预测和实证分析。本文将用计量经济学理论建立的数学模型,对随机误差项的 AR(2)结构、DW 统计量、SBC 以及 AIC 等统计变量进行计算,进而应用它们进行入园客流定量预测及分析等。

该系统包括气象要素数据库和世园会要素数据库(如图 1)。气象要素数据库包含了基本的温度、降水、风速以及相对湿度等资料;世园会要素数据库包含了入园票检人数、门票销售、团队游客、医疗服务和停车等综合信息。利用未来一段时间的天气预报信息,该系统实现了未来可能入园人数的实时预

测。预测信息成为执委会官方了解和掌握世园会客流分布情况的重要参考依据,有利于世园会主办方合理安排交通、餐饮和世园会园区露天娱乐节目表演,以保障西安世园会的有序进行。

2 西安世园会入园客流人数和气象条件分析

从 2011 年 4 月 28 日西安世园会开园到 8 月 15 日的逐日入园客流数量、最高温度、逐 12 h 降水量、平均风速以及平均相对湿度的逐日时间序列变化如图 2 所示。入园客流表现出显著的周期变化特征,以 7 天为周期而上下波动。其中周六、周日、五一节假日以及暑假是入园客流数量的高峰期。从整体趋势来看,入园人数在 4 和 5 月较多,6 月入园人数有减少的趋势,但是随着 7 和 8 月暑假的来临,入园人数又表现出增加的趋势,尤其是在 7 月的中下旬达到了一个入园客流数量的顶峰。

逐日最高温度在波动变化中表现出上升的趋势,7—8 月高于 30℃ 的天数明显多于 4、5 和 6 月,并且最高温度波动变化幅度更小。世园会开园至 8 月 15 日这段时间一共有 10 次明显的降水过程,进入 7 月以后,降水天数有所增加且降水量有逐渐增多的趋势。4 和 5 月的平均风速波动不大;进入 6 月以后平均风速变大,波动也增大;7 和 8 月平均风速有所降低,但波动仍然较大。平均相对湿度周期

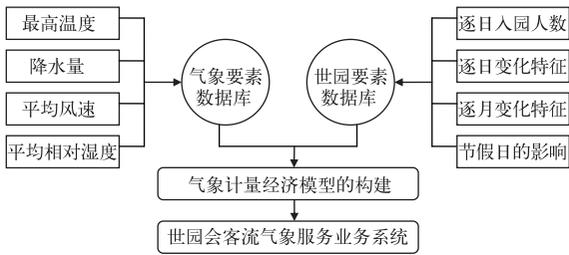


图 1 2011 年西安世园会入园客流气象服务系统的总体框架

Fig. 1 Framework of the system for meteorological service of Expo 2011 Xi'an's visitors

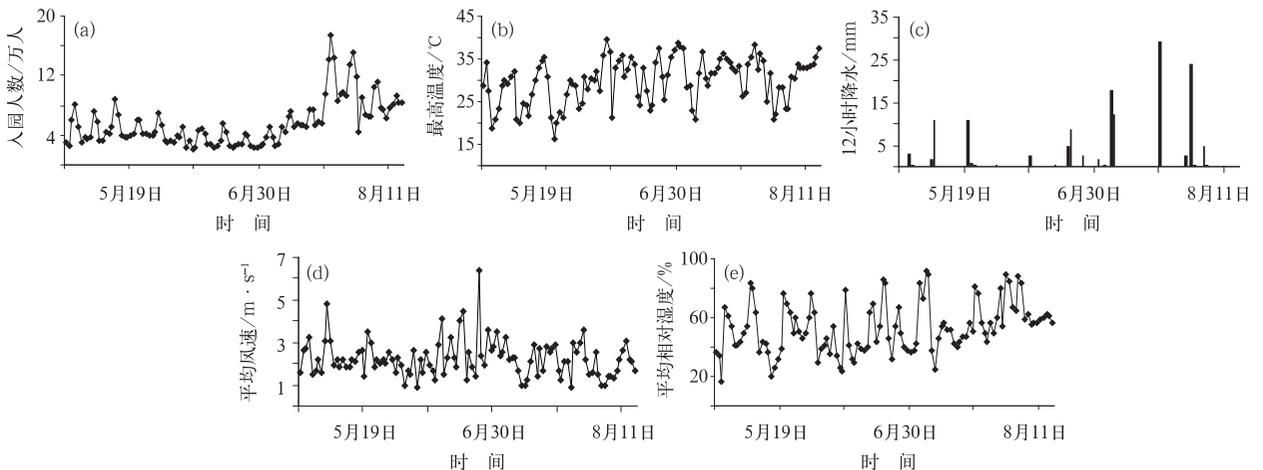


图 2 2011 年 4 月 28 日至 8 月 15 日西安世园会入园客流人数(a)及气象要素的时间序列变化 (b) 最高气温, (c) 12 小时降水, (d) 平均风速, (e) 平均相对湿度

Fig. 2 Time series of visitors (a) and meteorological factors

(b) maximum temperature, (c) 12 h precipitation, (d) mean wind speed, (e) mean relative humidity of Expo 2011 Xi'an during 28 April to 15 August 2011

性比较明显,只要出现降水过程湿度都比较高,7 和 8 月的相对湿度维持在较高的水平,波动相对较小。

3 西安世园会入园客流预测气象计量经济模型的构建

3.1 影响入园客流气象因素及非气象变量的设定

入园客流人数 P_t 。入园人数的统计按照每天检票入园的人数为准,作为模型的因变量。由于入园人数变化较大,所以出现异方差的可能性较大。异方差性违背了回归分析中方差一致性的假设条件,使得参数估计值不是无偏最优的^[13]。本文对入园客流人数取对数 $\log(P_t)$,以减少异方差性对参数估计的影响。

时间趋势项 t 。时间趋势项用以描述入园客流人数随时间的长期变化趋势,这种趋势主要受参观者心理因素以及西安世园会吸引力变化的影响。从图 2 可以看出,从 4 月 28 日世园会开园到 8 月 15 日这段时间,入园客流人数表现出波动上升趋势。

每天的虚拟变量 D_{it} 。该变量代表一周内的每天对入园客流人数的影响,其中 $i=2, 3, \dots, 7$, 分别表示周二、周三, ..., 周日。周一为基础日,当 D_{it} 全都为零时,就表示周一。虚拟变量是定性变量,每个虚拟变量的取值只能是 0 或 1。例如,当某一天为周二时,则有 $D_{2t}=1$,其余 $D_{it}=0$,以此类推。从图 2 可以看出,逐日入园客流人数的变化具有明显的周期性,周一至周五入园人数较少,而周六和周日的人园人数较多。通过引入一周之内每天的虚拟变量,可以很好地反映这种非气象要素引起的逐日变化特征。

月份的虚拟变量 M_{it} 。该变量代表不同月份对入园客流人数的影响,其中 $i=5, 6, 7, 8$, 分别表示 5、6、7 和 8 月。4 月为基础月,当 M_{it} 全都为零时,就表示 4 月。虚拟变量 M_{it} 的取值只能是 0 或 1,当某一天属于 5 月的时候, $M_{5t}=1$,其余 $D_{it}=0$,以此类推。从图 2 可以看出,不同月份的世园会入园客流人数也有明显的不同,4 和 5 月入园人数较多,而 6 月的人园人数相对有所减少,7 和 8 月人数有显著的增加。在模型中引入月份的虚拟变量,可以反映非气象要素引起的逐月变化特征。

节假日的虚拟变量 V_t 。节假日期间入园人数会有所增加,比如每年的五一国际劳动节、暑假以及

十一国庆节等时间相对固定的节假日。此外,世园会期间官方举办的一系列打折优惠活动也可以看成是该虚拟变量,在这些节假日期间 $V_t=1$,其余时间 $V_t=0$ 。节假日虚拟变量的引入,可以使模型的设定形式更加准确。

气象要素变量。天气预报已经成为人们做出出行决策的重要判断信息。每日天气状况不仅影响园区树木花卉的生长,也对人们出行有显著影响。游客一方面根据预报做出预判断,另一方面根据实况做出现实选择,这一因素会对客流量产生影响,特别是预报的准确性,会直接影响游客对是否游园的判断。在模型中的设定有 W 项(即预报信息)和 W_{t-1} 项(即前一天的天气实况),当用模型进行客流预测时,既考虑了未来天气预报的影响,又考虑了现在天气实况的影响。为分析世园会期间一些重要的气象要素影响入园客流人数的定量大小,模型引入的气象要素包括最高温度(T_t)、降水的虚拟变量(R_t)、平均风速(W_t),以及平均相对湿度(H_t)4 个变量。

3.2 入园客流预测模型形式及其参数估计

由 3.1 节设定的变量可知,影响世园会入园人数的因素主要有非气象要素和气象要素两大类。4—6 月这段时间,最高温度、降水、风速以及相对湿度都是影响入园客流的气象因素,综合考虑这些因素,建立第一个模型:

$$\log(P_t) = c + \alpha t + \beta_i \sum_{i=2}^7 D_{it} + \gamma_i \sum_{i=5}^8 M_{it} + \lambda_i \sum V_{it} + mT_t + nR_t + pW_t + qH_t + \epsilon_t$$

其中, $c, \alpha, \beta_i, \gamma_i, \lambda_i, m, n, p$ 以及 q 为需要估计的模型参数, ϵ_t 为随机误差项。第一个模型参数估计结果见表 1 所示。常数项非常显著,时间趋势项为很小的负值但不显著。周二和周三的系数为负值,而周四、周五以及周末两天的系数都为正值,但只有周五、周六和周日的系数非常显著。参数估计结果说明,相对于周一来说,周二和周三这两天的入园人数会分别减少 8.5% 和 3.9%,而周五至周日这 3 天的入园参观人数会显著地分别增加 22.9%、58.5% 和 53.8%。相对于 4 月来说,5 月的入园人数增加 45.8%,6 月的入园人数增加 14%,但 t 检验不显著。4 类气象要素对入园人数都有负面影响,其中降雨和平均风速的系数 t 检验显著,即发生降雨的天气将使入园人数减少 15.2%,平均风速每增加 $1 \text{ m} \cdot \text{s}^{-1}$ 将使入园人数减少 5.4%。模型的拟合优度

为 0.87,不是很高。DW 统计值为 2.43,说明模型的随机误差项存在负的一阶自相关性。

由以上分析可知,第一个模型的设定形式不够完整,使得模型的随机误差项存在负的一阶自相关性。自相关性的存在使得模型参数估计值不是无偏最优的,参数估计值也可能是不合理的,例如最高温度 T_t 和平均相对湿度 H_t 两个气象要素对入园人数的影响 t 检验不显著。因此,模型设定形式需要进一步完善。为了消除自相关问题,在第一个模型的基础上引入随机误差项的自回归结构(AR),考虑到人们通常根据未来气象预报信息做出出行决策,以及天气预报的准确性问题,确定了二阶自回归结构 AR(2)的形式。第二个模型的完整形式如下:

$$\log(P_t) = c + \alpha t + \beta_i \sum_{i=2}^7 D_{it} + \gamma_i \sum_{i=5}^8 M_{it} + \lambda_i \sum V_{it} + mT_t + nR_t + pW_t + qH_t + (\phi_1 \varepsilon_{t-1} + \phi_2 \varepsilon_{t-2} + \alpha_t)$$

其中, $(\phi_1 \varepsilon_{t-1} + \phi_2 \varepsilon_{t-2} + \alpha_t)$ 为随机误差项的 AR(2) 结构, ϕ_1 和 ϕ_2 为需要估计的参数, α_t 为随机白噪声过程。第二个模型参数估计结果见表 1 所示。可以看出,拟合优度 (R^2) 提高到 0.93,说明模型的拟合能力有很大的提高。DW 统计值为 2.01,自相关性被消除。节假日造成入园人数显著地增加 28.3%。与第一个模型相比,每一天和月份的虚拟变量的系数值改变不大,但影响入园人数显著的气象因子为最高气温和降水。最高气温每升高 1℃,入园人数将减少 1.2%。发生降雨的天气,入园人数将减少 14.4%。

第二个模型在对实际值的拟合能力上有较大的提高,也消除了异方差性对参数估计的影响。事实上,由于气象要素变化连续性,当天的气象要素(最高气温、降雨、风速和湿度)对次日的入园人数有较大的影响。为了让模型的设定形式更加合理,在第二个模型的基础上引入气象要素的一阶滞后项,第三个模型的完整形式如下:

$$\log(P_t) = c + \alpha t + \beta_i \sum_{i=2}^7 D_{it} + \gamma_i \sum_{i=5}^8 M_{it} + \lambda_i \sum V_{it} + mT_t + nR_t + pW_t + qH_t + m_1 T_{t-1} + n_1 R_{t-1} + p_1 W_{t-1} + q_1 H_{t-1} + (\phi_1 \varepsilon_{t-1} + \phi_2 \varepsilon_{t-2} + \alpha_t)$$

其中, T_{t-1} , R_{t-1} , W_{t-1} 和 H_{t-1} 分别为最高温度、降

雨、平均风速和平均相对湿度的一阶滞后项, m_1 , n_1 , p_1 和 q_1 为对应的需要估计的参数。模型参数估计结果见表 1。节假日期间,入园人数将增加 32.7%, t 检验十分显著。与周一相比较,周二和周三的入园人数有所减少,而周五至周日的入园人数将十分显著地分别增加 22.9%、55.8% 和 47.3%。与 4 月相比较,5 月入园人数显著地增加 47.4%,6 月系数 t 检验不显著。引入气象要素的一阶滞后项之后,当天的气象要素对世园会入园人数的影响不具有统计显著性,而前一天的最高气温 (T_{t-1}) 升高 1℃,当天入园人数将减少 1.7%,前一天的平均风速 (W_{t-1}) 增加 1 m · s⁻¹,当天的入园人数增加 6.4%。模型的拟合优度 (R^2) 为 0.93, DW 统计值为 1.94,自相关性被消除。

进入 7 月以后,西安迎来了高温多雨的天气。高温和降水是影响世园会入园客流数量的主要因子,而风速和相对湿度的影响并不显著。为感谢和回馈广大市民,西安市委、市政府决定,从 7 月开始向每户西安市民发放一份市民礼包,礼包发放和使用都有较为严格的时间限制,这为业务系统预测可能入园客流人数增加了很大的不确定性。针对 4 至 8 月的入园客流情况,本文只考虑高温天气和降水两个气象因素,并且在模型中引入“大礼包”这个虚拟变量对模型参数进行重新计算。

4—8 月这个时间段的模型参数估计与显著性检验如表 2 所示。由模型 3 的参数估计值可知,与非节假日相比,节假日的到来将会使得入园客流人数增加 50.1%;与周一相比,周六和周日的入园人数将会显著增加 47.3% 和 41.3%;高温天气和降水天气对入园客流人数的影响十分显著。当发生高温和降雨的天气时,入园客流人数将分别减少 15.7% 和 7.73%,而前一天的高温天气会使得次日的入园客流人数减少 12.7%。

4 西安世园会入园客流预测气象计量模型评估及其业务化应用

4.1 世园会入园客流预测气象计量模型的评估

通过对比表 1 中的模型 2 可以发现,模型 3 稍微有所改进。引入气象要素的一阶滞后项之后,造成当天气象变量的系数 t 检验变得不显著。为剔除大礼包派送等不确定性因素的影响,本文仅对表 1

表 1 模型参数估计及显著性检验(4—6月)

Table 1 Parameter estimation and significance test for models (April to June)

解释变量	模型 1		模型 2		模型 3	
	参数	<i>t</i> 统计量	参数	<i>t</i> 统计量	参数	<i>t</i> 统计量
<i>C</i>	10.4295	32.51**	10.6816	45.13**	10.6643	32.40**
<i>t</i>	-0.0001	-0.05	0.0007	0.40	0.0013	0.58
<i>V</i>	0.3444	3.00**	0.2833	3.28**	0.3271	3.27**
<i>D</i> ₂	-0.0850	-1.04	-0.0942	-1.02	-0.0952	-1.10
<i>D</i> ₃	-0.0390	-0.49	-0.0457	-0.52	-0.0616	-0.82
<i>D</i> ₄	0.0203	0.25	0.0172	0.22	0.0056	0.06
<i>D</i> ₅	0.2289	2.85**	0.2248	2.95**	0.2287	2.83**
<i>D</i> ₆	0.5847	7.39**	0.5684	6.45**	0.5578	7.62**
<i>D</i> ₇	0.5380	7.06**	0.5428	6.02**	0.4730	5.35**
<i>M</i> ₅	0.4583	4.28**	0.4462	5.72**	0.4741	4.20**
<i>M</i> ₆	0.1402	0.96	0.1171	1.20	0.1782	1.35
<i>T</i>	-0.0058	-0.76	-0.0119	-2.10**	-0.0011	-0.12
<i>R</i>	-0.1518	-1.76*	-0.1437	-2.04**	-0.1049	-1.22
<i>W</i>	-0.0535	-1.70*	-0.0425	-1.48	-0.0415	-1.33
<i>H</i>	-0.0009	-0.32	-0.0031	-1.50	-0.0012	-0.42
<i>T</i> _{<i>t</i>-1}					-0.0166	-2.02**
<i>R</i> _{<i>t</i>-1}					-0.0966	-1.22
<i>W</i> _{<i>t</i>-1}					0.0641	2.15**
<i>H</i> _{<i>t</i>-1}					-0.0018	-0.69
AR(1)			0.39	2.22	0.27	1.26
AR(2)			0.28	1.60	-0.05	-0.20
回归 <i>R</i> ²	0.87		0.93		0.93	
DW 统计量	2.43		2.01		1.94	
SBC	-12.51		-11.98		-7.55	
AIC	-40.89		-44.15		-46.84	

注:(1) DW 值在 2.0 附近,说明模型的残差项不存在一阶自相关,AIC 为赤池信息准则;

(2) ** 和 * 分别表示 *t* 统计量通过 0.05 和 0.1 的显著性水平检验。

表 2 模型参数估计及显著性检验(4—8月)

Table 2 Parameter estimation and significance test for models (April to August)

解释变量	模型 1		模型 2		模型 3	
	参数	<i>t</i> 统计量	参数	<i>t</i> 统计量	参数	<i>t</i> 统计量
<i>C</i>	10.0415	75.32**	10.0425	61.89**	9.9800	58.93**
<i>t</i>	0.0080	3.47**	0.0081	2.49**	0.0080	2.55**
<i>V</i>	0.5594	8.25**	0.4673	4.9**	0.5008	5.45**
<i>D</i> ₂	0.0081	0.11	0.0060	0.11	-0.0041	-0.08
<i>D</i> ₃	0.0177	0.25	0.0059	0.10	-0.0117	-0.20
<i>D</i> ₄	0.0961	1.38	0.0865	1.4	0.0817	1.31
<i>D</i> ₅	0.2103	3.01**	0.2009	3.26**	0.1902	3.10**
<i>D</i> ₆	0.4636	6.54**	0.4661	7.98**	0.4726	8.11**
<i>D</i> ₇	0.4427	6.27**	0.4381	8.01**	0.4129	7.59**
<i>M</i> ₅	0.3533	2.61**	0.3225	2.01**	0.4150	2.49**
<i>M</i> ₆	-0.1856	-1.07	-0.1402	-0.65	-0.0339	-0.16
<i>M</i> ₇	0.0276	0.13	0.0660	0.23	0.1804	0.65
<i>M</i> ₈	-0.2638	-0.98	-0.2183	-0.62	-0.1304	-0.38
<i>T</i>	-0.1875	-3.46**	-0.1531	-3.06**	-0.1573	-3.18**
<i>R</i>	-0.0891	-2.95**	-0.0752	-2.79**	-0.0773	-2.90**
<i>T</i> _{<i>t</i>-1}					-0.1267	-2.59**
<i>R</i> _{<i>t</i>-1}					-0.0365	-1.38
AR(1)			-0.36	-3.62**	-0.37	-3.68**
AR(2)			-0.19	-1.89*	-0.15	-1.44
回归 <i>R</i> ²	0.86		0.88		0.89	
DW 统计量	1.25		1.99		2.00	
SBC	14.72		0.57		1.56	
AIC	-27.22		-46.96		-51.40	

注:(1) DW 值在 2.0 附近,说明模型的残差项不存在一阶自相关,AIC 为赤池信息准则;

(2) ** 和 * 分别表示 *t* 统计量通过 0.05 和 0.1 的显著性水平检验。

中的模型 2 进行预测能力评估,并用该模型对入园客流人数开展预测,讨论其业务运用的可行性。将 4 月 28 日至 6 月 30 日的入园客流人数时间序列分为两部分。6 月 15 日之前的数据为第一部分,用来估计模型的参数,并检验模型对数据的拟合水平。6 月 16 日之后的数据为第二部分,用来检验模型的实际预测能力。

图 3 为模型的拟合值与实际值的比较,拟合值与实际值十分吻合,表明模型中所引入的变量科学合理,模型能很好地模拟世园会入园客流人数的逐日和逐月变化特征。将 6 月 15 日之后两周的气象要素信息带入模型,对未来两周的可能入园客流人数展开预测。从图 4 的结果可以看出,模型比较准确地预测了未来两周的入园客流人数,表明模型的实际预测能力较强,可以投入实际的业务化运用。

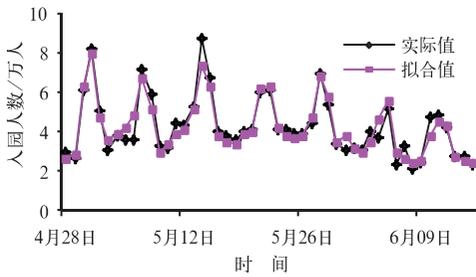


图 3 模型拟合值与实际值的比较

Fig. 3 Comparison between simulated values and observed values

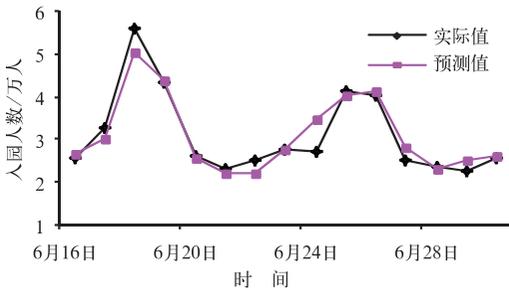


图 4 模型预测值与实际值的比较

Fig. 4 Comparison between predicted values and observed values

为了进一步研究模型的预测时效性,需要对比分析模型在不同时间长度上的预测准确性。图 5 为预测时间长度分别为 1~7 d 和 8~15 d 时,模型预测误差指标的对比分析。当预测时间长度为 1~7 d 时,预测绝对误差为 2000 人,预测相对误差为 6.05%,系统偏差比例为 4.78%。当预测时间长度

为 8~15 d 时,预测绝对误差为 2043 人,预测相对误差为 6.98%,系统偏差比例为 5.47%。可见,模型对 1~7 d 的预测准确性略高于 8~15 d 的预测结果。

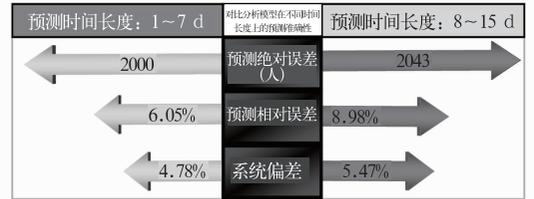


图 5 模型实际预测能力的误差指标

Fig. 5 Error indexes of prediction for the model

4.2 世园会入园客流预测模型的业务化服务系统

西安世园会入园客流人数预测系统的核心是表 2 中的模型 3。由于模型中考虑了输入较多的相关因子,在业务化系统设计中既要充分考虑系统的可维护性,又要考虑世园执委会用户群和气象部门业务人员每日业务化应用的便利,使用 ASP.NET+Access 技术设计业务系统,程序设计使用 Microsoft Visual Studio 2005,并采用 B/S 结构,链接在世园气象服务网(5 月与世园会官方网站链接)上运行、管理和日常维护(图 6)。经过 8 月底为期 10 d 的双方测试后,正式投入双方日常运行工作中,世园执委会投资策划部门用其对未来 3 日的散客、团队客等客流进行预测及控制工作。

运用已经建立好的业务系统对 2011 年 8 月 16 日至 10 月 22 日区间段的可能入园人数开展定量预测分析。西安市 1970—2009 年历史同期气候特征见表 3,依据此表对 8 月 16 日至 10 月 22 日世园会开园期间的天气状况进行经验判定。

利用西安市 1970—2009 年历史同期气候特征,将未来天气状况的经验判定结果带入业务系统中,经过计算得到 8 月 16 日至 10 月 22 日时间段的逐日可能入园散客客流数量,以及 4 月 28 日至 10 月 22 日整个世园会运营期间的入园散客总人数。由图 7 可见,8 月下旬的入园客流人数呈减少趋势;但在 9 月和 10 月上旬的入园客流将有所增加。业务系统预测结果表明 4 月 28 日至 10 月 22 日整个世园会运营期间的入园散客总人数约为 1032 万人。8 月,气象部门和世园执委会相关部门,依据开园以来的连续数据分析,散客人数所占总人数的比例约为



图 6 世园会入园客流气象服务业务系统预测流程及业务界面
 Fig. 6 Flow chart and interface of the prediction system for meteorological service of Expo's visitors

表 3 西安市 1970—2009 年 4—10 月气候特征

Table 3 Characteristics of Xian's climate in April to October from 1970 to 2009

时段	平均降水量/mm	日平均气温/℃	经验判定
春季回暖期(4月 1—30 日)	40.2	15.2	无判定
初夏少雨期(5月 1 日至 6 月 20 日)	90.6	21.8	无判定
初夏多雨期(6 月 21 日至 7 月 20 日)	102.0	25~28	无判定
盛夏伏旱期(7 月 21 日至 8 月 20 日)	73.1	24~28	多高温天气
初秋多雨期(8 月 21 日至 10 月 10 日)	151.2	15~24	多降雨天气
秋季凉爽期(10 月 11—31 日)	34.4	11~15	无高温天气

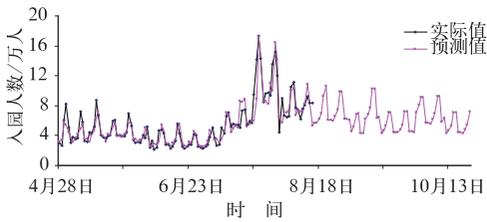


图 7 2011 年西安世园逐日入园客流人数预测

Fig. 7 Prediction for the number of daily visitors of Expo 2011 Xi'an

55%~65%，提前两个月计算得到世园会入园客流总人数逾 1600 万人。这个预测数据与 10 月 22 日官方公布的实际入园人数 1573 万人很接近。

通过对 4 月 28 日至 9 月 11 日预测值与实际值的统计计算，预测值与实际值之间的相关系数为 0.959，通过了 0.01 的显著性水平检验。将每一天的实际客流量与前一天的客流量的差值视为客流变化，其结果为增加或减少；将当天预测客流量与前一天的客流量的差值视为客流趋势预测。根据预测结

果统计，趋势预测准确率为 72.6%。

5 西安世园会入园客流预测气象服务主要结论和讨论

复杂多变的气象因素对 2011 年西安世园会的参观人数有重大的影响。开展气象因素影响入园客流人数的定量分析，有助于进一步加深认识气象与社会的关系，可以更加及时地开展气象保障服务，促进世园会的平稳持续进行。采用计量经济学方法，以最高温度、平均风速、平均相对湿度为定量解释变量，以降水、节假日、每一天、月份为虚拟解释变量，引入随机误差项的 AR(2) 结构，通过逐步优化的建模思想，构建了气象因素与入园客流人数的计量经济模型，能很好地模拟世园会入园客流人数的逐日和逐月变化特征，能比较准确地预测未来两周的可能入园人数和总人数。主要结论如下。

(1)模型很好地解释了节假日、每一天以及不同月份对入园客流人数的定量影响。相对于周一来

说,周二至周四的入园人数没有统计上的显著性差别,而周五至周日的入园人数会显著性地增加。相对于4月,5月的入园人数显著地增加41.5%;6月入园客流人数减少3.39%;7月的入园客流人数增加18.0%;而8月的人数减少13.0%。

(2)2011年4—6月期间,四类气象要素对入园客流人数都有负面影响。最高温度和降雨的系数 t 检验显著。最高气温升高 1°C ,入园人数将减少1.2%,如果发生降雨,入园人数将减少14.4%。7—8月期间,高温天气和降水天气对入园客流人数的影响十分显著。当发生高温和降雨的天气时,入园客流人数将分别减少15.7%和7.73%,而前一天的高温天气会使得次日的入园客流人数减少12.7%。

(3)利用西安市1970—2009年历史同期气候特征,将未来天气状况的经验判定结果带入该业务系统中,8月底即提前两个月,经过计算得到世园会运营期间的入园散客总人数约为1032万人,预测世园会入园客流总人数逾1600万人,与2011年10月22日官方公布的实际入园人数约1573万人很接近。

(4)客流预测气象计量经济模型和客流预测气象服务业务系统能很好地模拟世园会入园人数的逐日、逐月变化特征,能比较准确地预测未来15 d的入园人数,并且对1~7 d的预测准确性略高于8~15 d的预测结果。虽然模型和业务系统能较好地模拟入园客流人数的逐日、逐月变化特征,但对个别天数的模拟值存在一定的偏差。主要原因可能与在西安世园会举办期间,组委会和执委会官方不定期“大礼包派送”等激励政策和公益活动,或不同类门票的到期时间不同有关,这些构成客流预测不确定性影响因素。

在模型的实际运用中,模型预测的准确性依赖于温度、降水、风速和相对湿度等气象因素的预报准确性。在模型中的设定有 W_t 项(即预报信息)和 W_{t-1} 项(即前一天的天气实况),当用本文的模型进

行客流预测时,既考虑了未来天气预报的影响,又考虑了现在天气实况的影响。游客一方面根据预报做出预判断,另一方面根据实况做出现实选择,这些因素会对客流量产生影响,特别是预报的准确性,会直接影响游客对是否游园的判断,但情况会比较复杂。游客根据预测做出预判断,预报准确率越高,游客对未来预报信息越信赖。实践应用时特别强调中短期预报的准确性,进一步增加预报信息 W 这一项的置信度,提高对入园客流人数中短期预测的准确率。对复杂和不确定因素的影响还需要继续深入研究。

参考文献

- [1] 罗慧, 谢璐, 俞小鼎. 奥运气象服务社会经济效益评估个案分析[J]. 气象, 2007, 33(3): 89-94.
- [2] 姚秀萍, 吕明辉, 范晓青, 等. 气象服务效益评价研究进展[J]. 气象, 2011, 37(6): 749-755.
- [3] 王静, 孙健, 姚秀萍. 气象服务效益评估方法研究综述[J]. 气象, 2008, 34(T1): 300-304.
- [4] 王桂芝, 李廉水, 黄小蓉, 等. 条件价值评估法在公众气象效益评估中的应用研究[J]. 气象, 2011, 37(10): 1309-1313.
- [5] 罗慧, 赵奎峰, 巩在武, 等. 基于数据包络分析法的陕西气象资源效率评估[J]. 气象, 2011, 37(11): 1438-1442.
- [6] 罗慧, 苏德斌, 丁德平, 等. 对潜在气象风险源的公众支付意愿评估[J]. 气象, 2008, 34(12): 79-83.
- [7] 吴振玲, 刘爱霞, 白玉荣, 等. 花粉预报服务社会经济效益估算分析[J]. 气象, 2011, 37(5): 626-632.
- [8] 王小平, 孙彩贤. 基于多元回归模型的2010年上海世博会客流量预测分析[J]. 江汉大学学报(自然科学版), 2010, 38(2): 24-26.
- [9] 孙宁, 李廉水. 基于SVAR模型的气温变化对南京市工业经济的影响研究[J]. 气象, 2009, 35(10): 90-96.
- [10] 刘杰, 许小峰, 罗慧. 气象条件影响我国农业经济产出的计量经济分析[J]. 气象, 2010, 36(10): 46-51.
- [11] 吴向阳, 张东海. 北京市气温对电力负荷影响的计量经济分析[J]. 应用气象学报, 2008, 19(5): 531-538.
- [12] 戚岳. 基于计量经济学的中长期电力负荷分析及预测[D]. 北京: 北京交通大学, 2008.
- [13] 高铁梅. 计量经济分析方法与建模[M]. 北京: 清华大学出版社, 2009: 145-146.