

# 中国西瓜价格短期预测分析

赵 姜<sup>1</sup>, 吴 芮<sup>2</sup>, 吴敬学<sup>3</sup>

(1. 北京市农林科学院 农业科技信息研究所, 北京 100097; 2. California State University(East Bay), California, United States of America 94542; 3. 中国农业科学院 农业经济与发展研究所, 北京 100081)

**摘 要:**通过对 2000 年 1 月至 2014 年 10 月的中国西瓜月度批发价格数据进行分析,建立了基于时间序列的价格预测模型。结果表明:SARIMA 模型和季节因子分离模型都能较好地模拟国内西瓜价格波动,而组合模型对单一模型的预测精度又有一定提高;在此基础上对 2014 年 11 月至 2015 年 12 月的西瓜月度批发价格进行了预测分析。

**关键词:**西瓜;价格;预测

**中图分类号:**S 651 **文献标识码:**A **文章编号:**1001-0009(2015)23-0213-04

西瓜是我国重要的园艺类作物。自 20 世纪 80 年代以来,中国西瓜产业快速发展,已成为世界西瓜生产第一大国,近 10 年西瓜产量始终占据世界总产量的 60% 以上。同时,作为我国夏季主要的鲜食水果,西瓜消费约占全国 6—8 月上市水果的 60% 左右,消费需求持续增长。然而,近年来“瓜贱伤农”的新闻频频爆出,2013 年甚至被称为“史上最难卖瓜季”。西瓜价格过高会影响城乡居民的切身利益,价格过低又会严重损害瓜农的收益,因此,在当前西瓜产量及需求量均增长的情况下,掌握西瓜价格变动规律,开展市场短期预测研究,对保障农民利益、稳定居民日常生活、优化资源配置、促进中国西瓜产业健康稳定发展具有重要的理论和现实意义,同时也为其它鲜活农产品价格预测提供有价值的参考借鉴。

农产品价格预测最早始于 HENRY<sup>[1]</sup>对棉花产量及价格建立的回归预测方程。之后 SARLE<sup>[2]</sup>、CLIFTON 等<sup>[3]</sup>、WILBUR<sup>[4]</sup>利用多元回归方法分别对生猪、产仔季生猪、牛肉和猪肉的价格进行了预测。随着时间序列分析技术的发展,指数平滑法<sup>[5]</sup>、频谱分析<sup>[6]</sup>、神经网络<sup>[7]</sup>、VAR 模型<sup>[8]</sup>等被频繁应用到农产品价格预测。国内关于农产品价格的研究也延续了多元回归、时间序列分析、智能预测的研究方法。程贤禄<sup>[9]</sup>利用多元回归方法分析了北京市蔬菜批发价格。许彪等<sup>[10]</sup>建立了 5 因素模型的猪价分析框架,并预测了未来 15 个月的生猪价格趋势。张瑞荣等<sup>[11]</sup>对肉鸡产品价格进行了模拟分析和预测评价,指出 ARDL 模型预测效果好于 ARIMA 模型,但 ARIMA 模型更便于预测。方燕等<sup>[12]</sup>利用灰色

预测模型对我国大豆进行了预测。徐明凡等<sup>[13]</sup>认为采用灰色理论模型对我国鸡蛋价格进行预测比神经网络模型具有更有效的预测效果。孙红敏等<sup>[14]</sup>运用改进的神经网络算法对东北地区畜产品价格进行预测。随着预测研究的不断发展,组合预测方法被越来越广泛运用到农产品价格预测中,李干琼等<sup>[15]</sup>、王会娟等<sup>[16]</sup>、张立杰等<sup>[17]</sup>、许杞刚等<sup>[18]</sup>的研究均表明,组合预测模型对单项模型有一定改进,不仅预测精度更高,而且预测结果更加稳定。组合预测法现已被广泛接受并成为国内外预测研究和应用的热点。

综合国内外研究发现,回归分析法需要尽可能收集所有的影响因素,成本较大且不太现实,因此对于农产品价格特别是鲜活农产品价格预测多采用时间序列模型进行预测,预测效果也比较好。现有对西瓜产业所做出的市场分析主要是一些定性研究,或以数据罗列形式显示的当期价格<sup>[19-20]</sup>,对其进行短期价格预测的文献则几乎没有。现以西瓜月度批发价格为基础数据,通过构建计量经济模型预测西瓜价格未来走势,增强其短期价格的预见性,以期为生生产者、经营者及政府相关部门提供决策参考。

## 1 我国西瓜价格波动分析

所用数据来源于农业部信息中心提供的 2000 年 1 月至 2014 年 10 月国内西瓜批发价格,能够反映全国批发市场西瓜价格的月度波动情况。由图 1 可知,中国西瓜价格总体上处于缓慢上升趋势。2014 年 2 月西瓜价格达到样本内最高值 6.53 元/kg,是 2000 年 2 月价格的 1.61 倍、2010 年 2 月价格的 1.54 倍。

西瓜生产与消费的季节性很强,从图 1 也可以看出,西瓜价格具有较大的季节波动。运用 X12 季节调整法中的乘法模型对西瓜月度价格进行分解,得到价格波

**第一作者简介:**赵姜(1985-),女,博士,助理研究员,研究方向为农业经济管理。E-mail:zhaojiang821@163.com.

**收稿日期:**2015-07-30

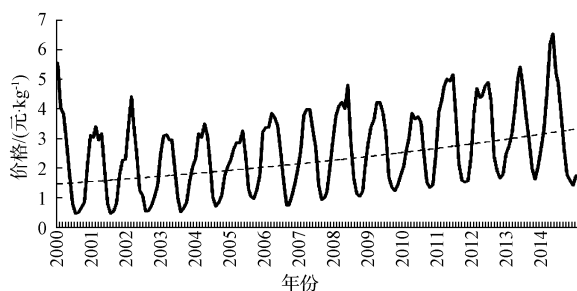


图1 中国西瓜批发价格波动趋势

动的季节成分,由图2可知,西瓜价格的季节波动性十分明显,在每年集中上市的7—8月份价格降到最低点,在1—2月份到达价格最高点季节性波动周期比较稳定,基本上每年经历从1—2月(峰)—7—8月(谷)—1—2月(峰)为12个月的季节性波动。这种显著稳定的季节因子决定了西瓜价格预测模型应该体现价格波动的季节性。

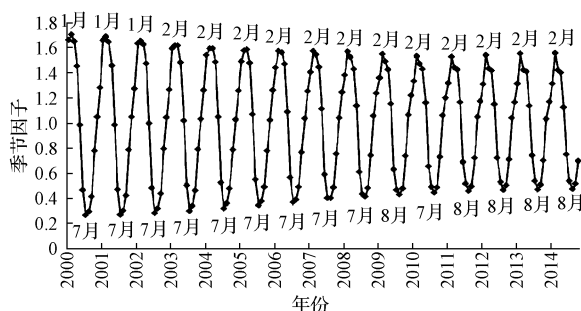


图2 中国西瓜价格波动的季节因子序列

## 2 西瓜价格预测模型设定

### 2.1 SARIMA 模型

ARIMA 模型是一种精度较高的经典时序短期预测方法,它可以通过多次差分使非平稳的时间序列转化为平稳序列,再将此序列表示成关于序列过去某一点的自回归和关于白噪声的移动平均组合。若时间序列中有明显的时间趋势和季节效应时,简单的 ARIMA 模型就不能充分描述其中的相关关系。由于西瓜价格显著的季节性波动特征,该研究采用乘积季节模型 SARIMA 用于价格预测。

SARIMA 模型用符号表示为  $SARIMA(p, d, q)(P, D, Q)$ 。其中,  $S$  代表周期,  $AR$  代表自回归,  $MA$  代表移动平均,  $I$  代表差分,  $P, Q, p, q$  分别表示季节与非季节自回归、移动平均算子的最大滞后阶数,  $d, D$  分别表示非季节和季节差分次数。其表达式为:

$$\Phi_p(L)A_p(L^s)(\Delta^d \Delta_s^D y_t) = \Theta_q(L)B_q(L^s)v_t$$

式中,  $A_p(L^s)$  和  $B_q(L^s)$  分别为季节  $P$  阶自回归算子和  $Q$  阶移动平均算子,  $v_t$  为随机扰动项。模型具体构建过程如下。

#### 2.1.1 数据处理及稳定性检验

为了减弱序列的波动、

消除异方差的影响,对西瓜的实际价格数据取自然对数。现采用 ADF 检验法对西瓜价格时间序列变量进行平稳性检验。从表1可以看出,西瓜价格原序列是非平稳时间序列,但对其一阶差分后再进行 ADF 检验,在 1% 的显著水平条件下拒绝单位根假设,表示序列是平稳的,即西瓜价格的对数时间序列数据是一阶单整序列。

表1 变量序列的单位根检验结果

变量	ADF 统计值	(c,t,k)	显著性	结论
lnw	-0.64	(c,0,10)	0.856 8	不平稳
D(lnw)	-14.34	(c,0,9)	0.000 0*	平稳

注:①c 为常数项, t 为趋势值, k 为滞后阶数;②滞后阶数 K 以 SIC 准则为选择标准;③“\*”代表 1% 的统计显著水平;④lnw 是西瓜价格的自然对数序列, D 表示变量的一阶差分。

2.1.2 模型识别 西瓜价格对数时间序列进行一阶差分后为平稳序列,但自相关图和偏相关图在  $k=12$  或 24 时显著不为 0,表明季节性存在,因此进行一阶季节差分,得到新序列,继续观察新序列的偏相关图和自相关图,结合可决系数  $R^2$ 、AIC 和 SC 准则等,考虑模型的整体拟合效果,反复试验比较,最终对西瓜月度价格序列建立如(1)式中的 SARIMA 模型,即  $ARIMA(1,1,2) \times (1,1,1)^{12}$ :

$$(1+0.477\ 4L)(1-0.216\ 7L^{12})D_{12}D\log(w) = (1-0.448)(-2.37)(-7.06)(-2.22) \\ 0.733\ 8L-0.231\ 4L^2)(1-0.673\ 1L^{12})v_t \quad (1)$$

式(1)中,  $w$  为西瓜月度价格序列,  $D_{12}D\log(w)$  表示进行了一次差分和一次季节差分后的西瓜月度价格对数序列,  $L$  是滞后算子。式中各变量系数均通过了 5% 的显著性水平检验,并对模型的残差序列进行白噪声检验,根据残差序列自相关图直观判断,残差序列的自相关系数都落入随机区间中,残差序列是纯随机序列,检验通过。同时,模型自回归 AR 部分和移动平均 MA 部分的根均在单位圆内,表明所建立的 SARIMA 模型平稳可逆。

### 2.2 季节因子分离模型

季节因子分离模型是指对剔除季节因子后的时间序列建立 ARIMA 模型。对西瓜月度价格  $w$  进行 X12 季节调整,得到季节因子  $w_{sf}$  和季节因子分离后的西瓜月度价格  $w_{sa}$ ,其中  $w=w_{sf} \times w_{sa}$ 。对  $w_{sa}$  进行平稳性检验发现其为一阶单整序列,一次差分后观察自相关图与偏相关图,反复尝试建立(2)式方程,即  $ARIMA(12,1,1)$ :

$$(1+0.467\ 6L-0.239\ 5L^{12})D(w_{sa}) = (1-0.465)(-3.37) \\ 0.833\ 482L)v_t \quad (2)$$

式(2)中,  $D(w_{sa})$  表示剔除季节因子后的西瓜价格一阶差分序列,  $L$  为滞后算子,其中 1 阶和 12 阶自回归算子、1 阶移动算子进入表达式。式中各变量系数均通

过了 1% 的显著性水平检验,并对模型的残差序列进行白噪声检验,模型残差不存在自相关,检验通过。同时,模型自回归 AR 部分和移动平均 MA 部分的根均在单位圆内,表明所建立的季节因子分离模型平稳可逆。

### 2.3 模型拟合分析

分别利用 SARIMA 模型和季节因子分离模型对 2007—2014 年的西瓜月度价格进行预测,预测值与实际值的拟合效果如图 3 所示。重合度高的曲线段表示拟合误差小,重合度低的曲线段表示拟合误差大,从图 3 可以看出,2 个模型得到的预测价格曲线与实际价格曲线非常接近,大部分曲线段都有重合,说明 2 个模型的预测精度较高。

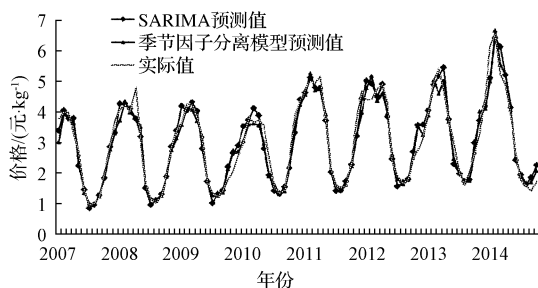


图 3 SARIMA 模型与季节因子分离模型预测值拟合效果

进一步用平均绝对百分比误差 MAPE (Mean Absolute Percentage Error) 评价模型预测效果,  $MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{\hat{y}_i - y_i}{y_i} \right|$ , 其中  $n$  为样本容量,  $\hat{y}$  为预测值,  $y$  为实际值。SARIMA 模型预测结果的 MAPE 值为 9.18%, 季节因子分离模型预测结果的 MAPE 值为 7.42%, 都小于 10%, 说明模型的预测精度较高, 具有可信度, 并且季节因子分离模型的预测精度比 SARIMA 模型略高。

### 2.4 组合预测模型设定

利用单模型预测西瓜价格波动显然存在一定的不确定性和信息损失, 为了改进预测效果, 减少预测系统误差, 该文拟综合利用各种预测信息所提供的信息, 合理选取各单个预测模型的权重, 构建价格组合预测模型。在选择优化目标时, YAGER<sup>[21]</sup> 曾选用残差平方和作为优化目标, 王会娟等<sup>[16]</sup> 则选用平均绝对百分比误差的绝对值之和为优化目标, MAPE 值与残差平方和都是衡量模型拟合效果的指标, 但残差平方和是误差的平方, 极易受到异常值影响, 而 MAPE 衡量了预测值和实际值的绝对差距, 能够较为准确地反映模型的预测效果。

因此, 为保证良好的预测效果, 选择 MAPE 值最小作为优化目标, 利用 2000 年 1 月至 2014 年 10 月的西瓜

价格作为样本, 通过非线性优化方法获得各模型的权重参数, 构建西瓜价格的组合预测模型。确定组合预测模型权重的非线性优化模型如下所示:

$$\min MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{\sum_{i=1}^k x_i y_i^t - y_0^t}{y_0^t} \right|,$$

$$\text{s. t. } \begin{cases} \sum_{i=1}^k x_i = 1 \\ x_i \geq 0, i = 1, 2, \dots, k. \end{cases}$$

式中,  $y_i^t$  表示第  $i$  个模型对第  $t$  期数据的预测值,  $y_0^t$  表示第  $t$  期数据的真实值,  $x_i$  表示第  $i$  个预测模型的权重。求解非线性优化模型, 得到组合预测模型中 SARIMA 模型的权重为 21.29%, 季节因子分离模型的权重为 78.71%。在预测效果上, 组合预测模型的预测精度相对于 SARIMA 模型提高了 1.82%, 相对于季节因子分离模型提高了 0.04%, 预测效果总体上有了一定的改善。进一步分析发现, 单个预测模型权重越大, 组合模型预测精度的优越性就越小, 在具体应用中要根据实际情况合理设立预测模型。

### 3 西瓜月度价格预测

利用季节因子分离模型对西瓜价格进行预测时, 需要通过季节变动规律, 计算预测期内的季节因子指数。观察图 2 可发现近几年同月份的季节比率都较为相近, 因此选取 2011—2014 年各月份的季节比率平均值作为预测期的季节因子指数。分别利用所建立的 SARIMA 模型、季节因子分离模型和组合预测模型对 2014 年 11 月至 2015 年 12 月的西瓜月度价格进行预测。

从表 2 可以看出, 2014 年 11 月至 2015 年 12 月西瓜价格依然会在小范围内波动, 从 2014 年 11 月开始呈上升趋势, 至 2015 年 4 月左右开始回落, 8 月降至谷底。具体而言, SARIMA 模型的预测结果显示, 西瓜价格峰值会在 2015 年 3 月出现, 然后价格逐渐回落, 9 月价格则再次回升; 季节因子分离模型的预测结果显示, 2015 年 4 月西瓜价格将达到峰值, 5 月之后逐渐下降, 一直到 9 月价格将显现回升势头。组合预测模型与单个模型预测结果有微小差距, 显示西瓜价格将在 2015 年 2 月达到峰值, 随后回落至 8 月达到最低值, 然后逐渐上升。

表 2 2014 年 11 月至 2015 年 12 月

西瓜价格预测值								
元/kg								
时间	SARIMA	季节因子	组合	时间	SARIMA	季节因子	组合	
/(年-月)	模型	分离模型	模型	年-月	模型	分离模型	模型	
2014-11	2.78	3.04	2.98	2015-06	2.83	2.26	2.38	
2014-12	3.86	3.46	3.55	2015-07	1.99	1.69	1.75	
2015-01	4.77	3.69	3.92	2015-08	1.76	1.51	1.56	
2015-02	5.56	4.37	4.62	2015-09	1.98	1.70	1.76	
2015-03	5.64	4.28	4.57	2015-10	2.76	2.50	2.56	
2015-04	5.45	4.38	4.61	2015-11	3.70	3.62	3.64	
2015-05	4.36	3.64	3.79	2015-12	4.65	4.05	4.18	



在此基础上,具体对预测期内的西瓜价格进行年度比较分析,可以发现单个模型与组合模型预测的价格变动趋势较为一致,与往年相比,2014年11—12月的西瓜价格低于2013年同期价格,而与2012年同期价格基本持平;2015年1—3月的西瓜价格与2013、2014年相比,存在小幅度下降趋势;2015年4—10月价格与前2年持平,但2015年11—12月的西瓜价格则要比2013年的价格略高。

#### 4 结论

该研究基于国家农业部信息中心提供的西瓜月度价格时间序列,构建了拟合度较高的时间序列模型和组合预测模型,对中国西瓜价格进行了预测。结果显示,2014年11月至2015年12月,西瓜价格仍呈现较为明显的季节特征,夏秋期间西瓜价格与前2年相比较为平稳,冬春西瓜价格则存在一定的波动性。在价格预测模型设定方面:一是西瓜价格具有明显的季节性和周期性,如果不考虑这些因素影响,做出的预测则可能不太准确。因此选取能够体现季节特征的SARIMA模型和季节因子分离模型进行预测分析,整体效果较好,这对存在季节性、周期性波动的其它农产品价格预测具有一定的参考意义。二是该文构建的时间序列模型是一种短期价格预测模型,对比其它的数据预测方法,虽然预测精度较高,但只能用于短期预测,对于长期价格的预测应用价值略低;同时,所建立的模型不能作为永久不变的预测工具,需要不断加入新的实际值以修正或重新拟合更有效的模型。三是在组合模型预测方法上,除了单一的时间序列模型组合,把时间序列预测、智能预测等多种方法相结合,如将神经网络模型和时间序列模型组合以提高西瓜价格预测精度的研究有待于进一步深入。

#### 参考文献

- [1] HENRY L M. Forecasting the yield and the price of cotton[M]. New York: The Macmillan Company, 1917: 100-113.
- [2] SARLE C F. The forecasting of the price of hogs [J]. American Economic Review, 1925, 15(3): 1-22.
- [3] CLIFTON B C, PATRICK J. Predicting hog prices[J]. American Journal of Agricultural Economics, 1956, 38(4): 931-939.

- [4] WILBUR R M. Forecasting livestock supplies and prices with an econometric model[J]. American Journal of Agricultural Economics, 1963, 45(3): 612-624.
- [5] JAIRETT F G. Short term forecasting of Australian wool prices[J]. Australian Economic Paper, 1965, 4(1/2): 93-102.
- [6] CAIGILL T F, RAUSSER G C. Time and frequency domain representations of futures price as a stochastic process[J]. Journal of the American Statistical Association, 1972, 67(337): 23-30.
- [7] TAYEBI S K, BAYARI L. A prediction of the Iran's chicken price by the ANN and time series methods[J]. American-Eurasian J Agri and Environ Sci, 2008, 2(1): 1-5.
- [8] TAYLOR J. Volatility forecasting with smooth transition exponential smoothing[J]. International Journal of Forecasting, 2004, 2(20): 273-286.
- [9] 程贤禄. 北京市农产品批发市场蔬菜价格预测预报体系研究[J]. 北京农业科学, 2002(2): 1-10.
- [10] 许彪, 施亮, 刘洋. 我国生猪价格预测及实证研究[J]. 农业经济问题, 2014(8): 25-32.
- [11] 张瑞荣, 王济民, 申向明. 肉鸡产品价格预测模型分析[J]. 农业技术经济, 2013(8): 23-31.
- [12] 方燕, 马艳. 我国大豆价格波动及其未来走势预测[J]. 价格理论与实践, 2014(6): 67-69.
- [13] 徐明凡, 刘合光. 关于我国鸡蛋价格的预测与分析[J]. 统计与决策, 2014(6): 104-107.
- [14] 孙红敏, 吴静婷, 李晓明. 基于改进BP神经网络的价格预测模型研究[J]. 东北农业大学学报, 2013, 44(8): 133-137.
- [15] 李干琼, 许世卫, 李哲敏, 等. 农产品市场价格短期预测方法与模型研究: 基于时间序列模型的预测[J]. 中国农业大学学报, 2011, 16(2): 172-178.
- [16] 王会娟, 肖佳宁, 曲双石. 中国玉米批发价格的短期预测及预警[J]. 中国农村经济, 2013(9): 44-53.
- [17] 张立杰, 寇纪淞, 李敏强, 等. 基于自回归移动平均及支持向量机的中国棉花价格预测, 2013(6): 30-33.
- [18] 许杞刚, 刘明军, 李海. 基于改进KNN算法的农产品价格预测模型[J]. 济南大学学报(自然科学版), 2014, 28(2): 114-117.
- [19] 杨艳涛, 吴敬学. 2013年我国西甜瓜市场贸易分析与趋势展望[J]. 长江蔬菜, 2014(17): 1-6.
- [20] 张琳, 杨艳涛, 吴敬学. 新形势下中国西瓜甜瓜产业发展的战略思考[J]. 北方园艺, 2014(19): 187-190.
- [21] YAGER R R. Induced aggregation operators[J]. Fuzzy Sets and System, 2003, 137(1): 59-69.

### Short-term Forecasting of Watermelon Price in China

ZHAO Jiang<sup>1</sup>, WU Rui<sup>2</sup>, WU Jingxue<sup>3</sup>

(1. Institute of Agricultural Sciencetech Information, Beijing Academy of Agriculture and Forestry Sciences, Beijing 100097; 2. California State University(East Bay), California, United States of America 94542; 3. Institute of Agricultural Economics and Development, Chinese Academy of Agricultural Sciences, Beijing 100081)

**Abstract:** The paper established a price forecasting model based on time series through the analysis on monthly wholesale prices of watermelon in China from January 2000 to October 2014. The results indicated that both ARIMA models and seasonal factors separate models could simulate the watermelon price effectively, and the combined model had a higher precision of prediction than a single model. Then, the monthly wholesale price of watermelon from November 2014 to December 2015 was forecasted on the base of the combined model.

**Keywords:** watermelon; price; forecasting