

郑秀国,杨娟,钱婷婷,等.近十年上海蔬菜价格的SARIMA分析与预测[J].上海农业学报,2020,36(6):138-141

近十年上海蔬菜价格的SARIMA分析与预测

郑秀国,杨娟*,钱婷婷,许叶颖

(上海市农业科学院农业科技信息研究所,上海数字农业工程技术研究中心,上海201403)

摘要:为更好地分析预测蔬菜价格的波动趋势,为蔬菜生产管理和经营提供科学决策参考,采用2010年1月至2019年12月的上海蔬菜批发价数据,研究建立最佳的上海蔬菜SARIMA模型。结果表明:SARIMA(3,1,3)×(2,1,0)₁₂模型效果最佳,利用该模型对2020年1—6月上海蔬菜价格进行预测,对比分析预测数据(经春节调整后)和实际的上海蔬菜批发价数据,得出该模型的静态预测平均误差率为7.3%,动态预测平均误差率为4.6%,表明SARIMA(3,1,3)×(2,1,0)₁₂模型能较好地表现上海蔬菜价格的波动特征,即上海蔬菜价格与前3个月价格以及前2年同期价格的紧密关联关系。SARIMA模型可用于分析预测上海蔬菜价格未来走势,指导产业发展。

关键词:蔬菜;价格;SARIMA模型;预测;时间序列

中图分类号:S126 文献标志码:A 文章编号:1000-3924(2020)06-138-04

Analysis and forecast of Shanghai vegetable price based on SARIMA model in recent ten years

ZHENG Xiuguo, YANG Juan*, QIAN Tingting, XU Yeying

(Institute of Agricultural Science and Technology Information, Shanghai Academy of Agricultural Sciences; Shanghai Engineering Research Center of Information Technology in Agriculture, Shanghai 201403, China)

Abstract: In order to analyze and predict the fluctuation trend of vegetable price, and provide decision-making reference for vegetable production management and operation, this paper used the wholesale price data of vegetables in Shanghai from January 2010 to December 2019 to study and establish the best SARIMA model of vegetables in Shanghai. The results showed that SARIMA(3,1,3)×(2,1,0)₁₂ model had the best effect. The model was used to predict the vegetable price in Shanghai from January to June 2020. The comparison between the predicted data and the actual wholesale vegetable price data in Shanghai showed that the average static prediction error rate of the model was 7.3%, the average dynamic prediction error rate was 4.6%, which showed that SARIMA(3,1,3)×(2,1,0)₁₂ model could better show the fluctuation characteristics of Shanghai vegetable price, that is, the close relationship between Shanghai vegetable price and the price in the last three months and the same period in the last two years. SARIMA model could be used to analyze and predict the future trend of vegetable price in Shanghai and guide the development of the industry.

Key words: Vegetable; Price; SARIMA model; Forecast; Time series

蔬菜是“菜篮子工程”的主要产品之一,蔬菜产业也是城市近郊农业生产者的重要经济来源。然而,蔬菜价格不仅受到市场供求、自然气候、生产成本、市场流通等规律性因素的影响,还会受到突发的气候变化、供求变化等影响,尤其是近年来大市场大流通下蔬菜供应的季节性或种类性不均衡时有发生,导致蔬菜价格波动频繁。

收稿日期:2020-04-08

基金项目:上海市科学技术委员会技术标准专项课题(17DZ2204400)

作者简介:郑秀国(1983—),男,硕士,助理研究员,主要从事农产品市场信息监测预警和智慧农业研究。E-mail:zhengxiuguo@saas.sh.cn

*通信作者,E-mail:yangjuan@saas.sh.cn

作为生物性产品,蔬菜的供应受季节气候的影响最大,其价格随着供应的周期性波动而波动,在时间序列上表现出明显的季节性和周期性特征,时间序列模型可较好地反映这种变化特征。常用于农产品价格分析的时间序列模型有自回归移动平均模型(ARIMA模型)、季节自回归移动平均模型(SARIMA模型)和Holt-Winters指数平滑模型(Holt-Winters模型)等,其中SARIMA模型、Holt-Winters模型均考虑到季节性因素的影响^[1]。SARIMA模型在农产品价格分析预测中应用十分广泛,如陈灿煌等^[2]、陈兆荣等^[3]运用该模型预测了中国农产品价格指数,王云鹏^[4]用其预测农业生产资料价格指数。应用该模型预测农产品价格、蔬菜价格和单种类蔬菜价格也都取得了较好的效果^[5-7],也是鸡蛋价格、肉鸡价格、生猪价格、猪肉价格常用的分析预测方法^[8-11]。李干琼等^[12]采用5种短期预测模型预测2009年西红柿价格,表明SARIMA模型明显优于其他4种模型。王美凤等^[13]对4种上海青菜价格预测模型进行了比较研究,表明SARIMA模型的预测效果最好。蔬菜价格是同期多种单种类菜价的汇总,开展蔬菜价格波动特征分析和预测研究,可为管理部门的政策调控和生产者的经营决策提供参考,对于减少蔬菜价格大幅波动对市民生活和蔬菜生产者的影响具有重要意义,然而尚未见蔬菜总体价格的波动特征与预测分析研究的报道。为此,本研究采用2010年1月—2019年12月上海蔬菜批发价数据,建立上海蔬菜SARIMA模型,进而对2020年1—6月上海蔬菜价格进行预测,通过模型预测效果分析进一步验证模型效果,最后根据SARIMA模型参数值分析上海蔬菜价格波动特征。

1 数据来源与模型方法

1.1 数据来源

本研究所使用的数据序列为2010年1月至2020年6月上海蔬菜批发市场平均价格(元人民币/kg,下同),记为PR,来源于上海市农产品价格监测与分析预测系统。其中2010年1月—2019年12月数据用于上海蔬菜SARIMA模型建立,2020年1—6月数据用于进一步验证分析模型预测效果。

1.2 建模方法

SARIMA(p,d,q)(P,D,Q)模型结构中,p为自回归项数,q为移动平均项数,d为差分次数,P是季节自回归阶数,Q是季节移动平均阶数,D为季节差分次数。具有季节性、趋势性和周期性的平稳数据序列 $\{B_t, t=1, 2, \dots\}$,可以采用SARIMA(p,d,q)(P,D,Q)模型结构建立模型。

首先对上海蔬菜批发价数据序列进行季节性、趋势性检验。季节性检验通常采用时间序列趋势图来判断,若趋势图是以年为周期变动,则时间序列具有季节变动特征。趋势性检验是采用CensusX-12乘法方法对数据序列进行季节调整分解,获得季节调整后的趋势循环序列,再运用H-P滤波法对获得的趋势循环序列进行趋势波动与循环波动序列分离,分离出的趋势波动图可观察出序列的变化趋势。

SARIMA模型的建立与应用可通过Eviews软件实现,主要过程如下:

1) 序列平稳性检验。可通过ADF单位根检验来判断,对于非平稳时间序列,可以通过数据变换和差分实现序列的平稳化,并采用时间序列的自相关函数(ACF)和偏相关函数(PACF)是否截尾来判定序列的平稳性。

2) 模型的识别。主要通过观察序列的自相关函数(ACF)和偏相关函数(PACF)对可能的p,d,q和P,D,Q的参数值进行估计,由此构成不同的SARIMA模型结构形式。

3) 模型的参数估计。运用最小二乘法估计模型的系数,并对其显著性进行检验。对于不同参数模型的选优可采用BIC(贝叶斯信息)准则等综合判断。最后,对选出的模型检验其残差是否为白噪声,一个适合的模型的残差序列应是白噪声过程,其ACF和PACF不应与0有显著性差异。

4) 模型预测。由上述过程确定模型的最终参数后,即可得到模型,使用模型进行预测。

2 结果与分析

2.1 序列平稳化

从图1可以看出,上海蔬菜价格表现出较明显的季节性,每年3月至次年3月蔬菜价格呈“W”型波动规律,价格高点多出现在3月和9月,价格低点多出现在6月和11月,数据在周期内呈现整体递增趋势,为非平稳序列。

首先,为消除序列可能存在的异方差,将原始序列进行对数化处理;其次,为消除趋势性,对原始序列进行季节性差分和普通一阶差分。由图2可以看出,处理后序列 $\Delta\Delta_{12}\text{LnPR}$ 的自相关函数 ACF 和偏相关函数 PACF 整体呈现截尾态势,由此认为 $\Delta\Delta_{12}\text{LnPR}$ 是平稳的序列。

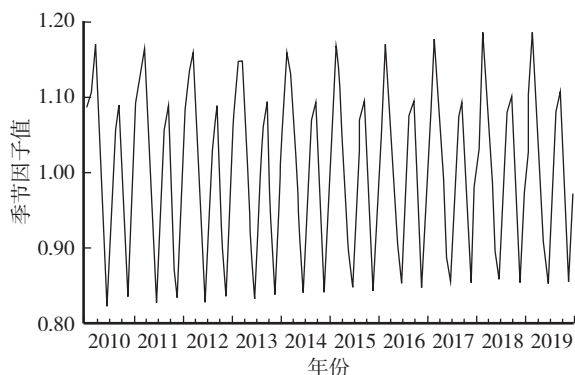


图1 上海蔬菜月均价季节因子序列

Fig. 1 Seasonal factor sequence of monthly average vegetable price in Shanghai

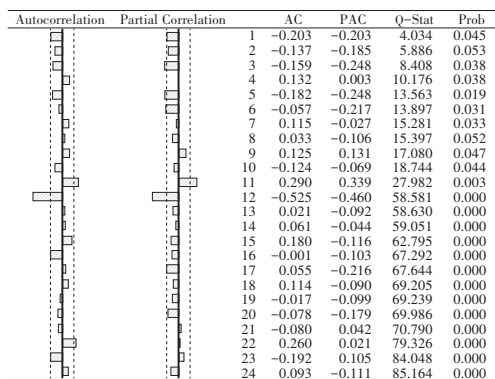


图2 对数化处理后的上海蔬菜价格时间序列 $\Delta\Delta_{12}\text{LnPR}$ 的自相关与偏相关

Fig. 2 Autocorrelation and partial correlation of sequence $\Delta\Delta_{12}\text{LnPR}$ after logarithmic treatment

2.2 模型建立与参数识别

本研究选择多种不同的模型进行拟合,如 $\text{SARIMA}(0,1,1) \times (1,1,1)_{12}$ 、 $\text{SARIMA}(2,1,1) \times (1,1,1)_{12}$ 、 $\text{SARIMA}(2,1,1) \times (1,1,2)_{12}$ 、 $\text{SARIMA}(2,1,2) \times (2,1,1)_{12}$ 、 $\text{SARIMA}(2,1,2) \times (2,1,2)_{12}$ 等 p 、 q 和 P 、 Q 为 3 以内的各种取值组合,发现 $\text{SARIMA}(3,1,3) \times (2,1,0)_{12}$ 参数显著性 t 检验结果最好。从表 1 可以看出,该模型结构各参数的相伴概率均小于 0.15,确定上海蔬菜模型结构为 $\text{SARIMA}(3,1,3) \times (2,1,0)_{12}$ 。模型结构为: $(1 - \phi_1 B - \phi_2 B^2 - \phi_3 B^3)(1 - \Phi_1 B^{12} - \Phi_2 B^{24})\Delta^1 \Delta_{12}^1 \log Pr = (1 - \theta_1 B - \theta_2 B^2 - \theta_3 B^3)\varepsilon_t$, $\varepsilon_t \sim N(0, 0.100)$ 。

回归方程的决定系统为 0.71,误差项的方差估计值为 0.100,模型中各参数的取值见表 1。模型的残差序列经检验为平稳的白噪声序列。

表 1 SARIMA 模型参数估计与相关检验结果

Table 1 Model parameter estimation and test results based on SARIMA model

变量	系数取值	标准误差	t 值	相伴概率 p
C	-0.001	0.000	-4.204	0.000
AR(1)	0.711	0.121	5.877	0.000
AR(2)	0.515	0.097	5.303	0.000
AR(3)	-0.489	0.102	-4.776	0.000
SAR(12)	-0.766	0.098	-7.845	0.000
SAR(24)	-0.246	0.088	-2.797	0.007
MA(1)	-0.977	0.060	-16.241	0.000
MA(2)	-0.956	0.017	-56.833	0.000
MA(3)	0.937	0.060	15.498	0.000

2.3 模型预测

根据建立的 $\text{SARIMA}(3,1,3) \times (2,1,0)_{12}$ 模型对 2020 年 1—6 月的上海蔬菜价格进行静态预测和动态预测,通过实际价格和预测价格的比较,判断模型的预测精度。

市场调查发现,假期对蔬菜价格有一定影响,进一步分析了春节假期对蔬菜价格的影响。从 2010—2019 年上海蔬菜价格数据来看,含春节期间价格的 1 月蔬菜月平均价格比不含春节期间价格的 1 月上海蔬菜月平均价格低 2.0%,含春节期间价格的 2 月蔬菜月平均价格比剔除春节期间价格的 2 月上海蔬菜月平均价格高 12.3%,可见,春节假期对蔬菜价格影响效果不定。但是春节当月的蔬菜价格整体高于其前后月份,如春节在 1 月,1 月蔬菜价格比其前一个月和后一个月的价格平均高 2.9%;春节在 2 月,2 月蔬菜价格比其前一个月和后一个月的价格平均高 7.7%。

为此,本研究在进行预测时考虑了假期效应,如果春节所在月份预测值高于前一个月预测值和后一个月预测值的平均值历史水平则不进行春节调整,反之则对春节所在月份的预测值进行相应调整——若

春节在1月,则1月下调2.0%;若春节在2月,则2月上调12.3%。如表2所示,静态预测结果的平均误差为7.3%,动态预测结果的平均误差率为4.6%,上海蔬菜价格预测数据与实际数据基本一致。

表2 上海蔬菜价格静态预测结果和动态预测结果
Table 2 Static and dynamic prediction results of vegetable price in Shanghai

时间	实际值/(元·kg ⁻¹)	静态预测值/(元·kg ⁻¹)	静态误差率/%	动态预测值/(元·kg ⁻¹)	动态误差率/%
2020年1月	4.04	3.94	2.4	4.06	0.5
2020年2月	4.39	4.65	6.0	4.67*	6.4
2020年3月	4.22	4.40	4.2	4.25	0.8
2020年4月	4.25	3.87	9.0	4.11	3.4
2020年5月	3.54	3.36	5.1	3.52	0.7
2020年6月	3.85	3.19	17.0	3.24	15.7
平均值			7.3		4.6

*数据经春节调整

3 结论与讨论

研究表明,上海蔬菜价格具有较为明显的季节性和周期性特征,可用SARIMA模型对上海蔬菜价格进行分析和预测。根据模型的显著性检验和模型参数的显著性检验结果,SARIMA(3,1,3)×(2,1,0)₁₂模型效果最佳。从模型参数值可以看出,上海蔬菜价格为非平稳序列,一阶差分序列为平稳序列;根据参数取值,可看出上海当月蔬菜价格不仅与其前3个月的价格密切相关,而且受前2年同月价格的影响。

上海蔬菜价格SARIMA(3,1,3)×(2,1,0)₁₂模型预测的上海蔬菜2020年1—6月批发价(经春节调整后)与实际价格的对比分析结果表明,模型静态预测的平均误差为7.3%,动态预测的平均误差为4.6%,实际应用中可结合2种预测方式,达到更好的预测效果。SARIMA(3,1,3)×(2,1,0)₁₂模型能较好地表现上海蔬菜价格的波动特征,可用于分析和预测上海蔬菜价格未来走势,指导产业发展。但还要注意,时间序列模型建模主要依据历史价格所包含的信息,而实际影响市场价格波动的还有一些突发性因素,如台风暴雨、低温、暖冬等恶劣天气或者交通、疫情等突发事件,所以在实际价格的走势分析预测中,可在模型分析预测出规律性影响因素结果的基础上,再叠加考虑这些突发因素的影响,从而科学判断出各种影响因素对蔬菜价格影响的程度。

参 考 文 献

- [1] 杨娟. 蔬菜价格大数据研究[M]. 上海:上海科学技术出版社,2020.
- [2] 陈灿煌. 基于SARIMA模型的中国农产品价格指数短期预测[J]. 湖南商学院学报,2016,23(2):16-20.
- [3] 陈兆荣,雷勋平,王亮基. ARIMA-SVM组合模型的我国农产品价格预测研究[J]. 财经理论研究,2013(2):103-107.
- [4] 王云鹏. 基于SARIMA模型的农业生产资料价格指数预测[J]. 农村经济与科技,2019,30(9):144-147.
- [5] 袁点,张世敬,戴小勇,等. 基于SARIMA模型的大连黄瓜月度零售价格预测[J]. 中国经贸导刊,2012(20):70-72.
- [6] 罗长寿,周丽英. 季节时间序列模型在平菇价格预测中的应用[J]. 贵州农业科学,2013,41(11):202-204.
- [7] 史冠宇,李玫瑰. 基于ARIMA模型的青岛市大蒜价格时间序列预测分析[J]. 山东农业科学,2017,49(5):168-172.
- [8] 唐江桥,雷娜. 中国鸡蛋价格波动预警研究[J]. 西部论坛,2011,21(6):44-49.
- [9] 刘春,辛翔飞,王济民. 我国肉鸡价格波动趋势分析[J]. 中国家禽,2019,41(20):69-74.
- [10] 李子涵,张玉梅,马佳. 基于ARIMA模型的上海市生猪价格预测[J]. 农业展望,2019(4):8-11.
- [11] 范青青,袁艳红. 基于X12-ARIMA模型的猪肉价格波动规律研究[J]. 中国畜牧杂志,2018,54(6):138-142.
- [12] 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.
- [13] 王美凤,赵京音,邵芳,等. 上海青菜价格短期预测模型比较研究[J]. 上海农业学报,2017,33(1):160-165.

(责任编辑:闫其涛)