

# 基于ARIMA模型的甘肃省2015年蔬菜价格预测

唐占彪<sup>1</sup>, 王炳文<sup>2</sup>

(1. 甘肃农业大学经济管理学院, 甘肃 兰州 730070; 2. 甘肃农业职业技术学院, 甘肃 兰州 730020)

**摘要:** 利用“甘肃重点农产品市场采集分析系统”采集的数据, 对马铃薯和其它被监测的22种大众蔬菜平均价格的波动特征进行描述性分析, 采用乘法原则对原始蔬菜价格序列进行季节分解, 并对季节调整后的马铃薯和蔬菜平均价格建立了ARIMA模型, 预测了马铃薯和大众蔬菜2015年各月份整体平均价格。2015年监测蔬菜整体均价大体呈“V”型。

**关键词:** 蔬菜; 价格预测; 时间序列; ARIMA模型; 甘肃省

**中图分类号:** N32; F326.13 **文献标识码:** A **文章编号:** 1001-1463(2016)02-0016-05

**doi:** 10.3969/j.issn.1001-1463.2016.02.005

农产品价格对于指导农业生产及农业产业结构调整具有重要意义, 因此, 使用蔬菜历史价格数据对未来价格趋势进行有效预测显得尤为重要。时间序列定量分析是一种动态数据处理的统计方法, 其基本思想是: 根据系统有限长度的运行记录(观察数据), 建立比较精确地反映时间序列中所包含的动态依存关系的数学模型, 这个模型一旦被识别, 就可以从时间序列的过去值及现在值来预测未来值<sup>[1]</sup>。事实上, 时间序列的任何变异都是由各种不同的冲击形成的。ARIMA模型在经济预测过程中既考虑了经济现象在时间序列上的依存性, 又考虑了随机波动的干扰性, 对于经济运行短期趋势的预测准确率较高, 是近年应用比较广泛的方法之一<sup>[2-11]</sup>, 我们选择ARIMA模型对2015年甘肃省蔬菜价格数据进行时间序列建模并分析预测<sup>[12]</sup>。

ARIMA模型全称为差分自回归移动平均模型(Autoregressive Integrated Moving Average Model, ARIMA), 是由博克斯(Box)和詹金斯(Jenkins)于20世纪70年代初提出的一种著名的时间序列预测方法, 所以又称Box-Jenkins模型。其中ARIMA(p, d, q)模型由3个过程组合而成, 即差分平稳过程、自回归过程、移动平均过程。差分过程是得到平稳的时间序列。如果原时间序列是平稳的, 则ARIMA模型变为ARMA模型(自回归移动平均模型)。AR是自回归过程, p为自回归项数, 是利

用前期若干时刻的随机变量的线性组合来描述以后某时刻随机变量的线性回归模型; MA为移动平均过程, q为移动平均项数, 表示被模拟值无法由本身过去线性预测的部分。ARIMA模型根据原序列是否平稳以及回归中所含部分的不同, 包括移动平均过程(MA)、自回归过程(AR)、自回归移动平均过程(ARMA)以及差分自回归移动平均过程(ARIMA), 因此也就有对应的五种具体模型, 但统称ARIMA模型<sup>[13]</sup>。表示为:  $\Delta d Y_t = C + \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \dots + \alpha_p Y_{t-p} + \varepsilon_t + \beta_1 \varepsilon_{t-1} + \beta_2 \varepsilon_{t-2} + \dots + \beta_q \varepsilon_{t-q}$ , 式中,  $Y_t$ 为原序列,  $Y_{t-p}$ 是原序列的p阶滞后序列,  $\varepsilon_t$ 为移动平均算子, 也称为新息(Innovation)。

## 1 数据来源与预测方法

研究分析数据来自甘肃重点农产品市场信息采集分析系统, P代表蔬菜批发市场监测价格, PF是最终预测价格, P\_SA是原始序列剔除季节因子后的价格, P\_SAF是对季节调整后价格序列的预测值, P\_SF是季节因子, C是常数项。图表横轴2010 M01……2015 M12代表2010年1月……2015年12月, 纵轴代表价格或预测价格。M01……M12代表某一年的1月份……12月份。

## 2 结果分析

### 2.1 马铃薯价格波动分析与预测

批发市场监测数据显示, 马铃薯的最高价格一般出现在5—6月份。如果以1a为一个价格周期, 除2011年之外, 2012—2014年3a的价格变

收稿日期: 2015-08-18; 修订日期: 2015-11-11

作者简介: 唐占彪(1973—), 男(东乡族), 甘肃东乡人, 在读研究生, 研究方向为农业与农村经济及农业综合开发。

E-mail: 582550434@qq.com

通讯作者: 王炳文(1964—), 男, 甘肃庄浪人, 高级工程师, 硕士生导师, 主要从事食品工程与设计研究工作。

E-mail: gszs@sina.com

动均表现1—5月逐渐上涨,7—10月逐渐下降,之后略有回升,价格变动和季节性紧密相关。监测发现,2011—2014年马铃薯4a平均批发价格分别是1.64元/kg、1.43元/kg、2.32元/kg、1.99元/kg。月度最低批发价格在2011年10月,为1.00元/kg;月度最高批发价格在2013年5月,达到2.93元/kg。马铃薯批发价格基本价格趋势是突破1.50元/kg,波动性紧逼2.00元/kg(图1, P11、P12、P13、P14分别为2011、2012、2013、2014年的价格曲线)。

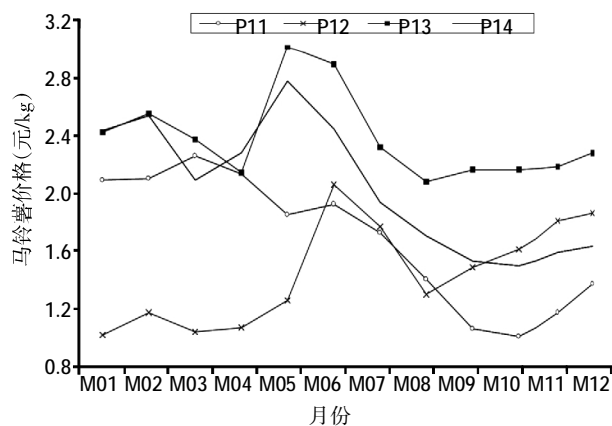


图1 2011—2014年马铃薯价格趋势

相应的12个季节因子如表1所示,对P\_SAF进行季节还原得到真实的预测值。马铃薯价格的季节因子全年4个月大于1,8个月小于1。因此,马铃薯价格受季节性影响,降价的空间更大,涨价的空间更小<sup>[14]</sup>,原因可能是容易保存,属于半吉芬商品。

表1 马铃薯季节调整后季节因子

| 月份  | 季节因子    | 月份  | 季节因子    | 月份  | 季节因子    |
|-----|---------|-----|---------|-----|---------|
| M01 | 1.078 9 | M05 | 1.269 3 | M09 | 0.860 7 |
| M02 | 1.136 1 | M06 | 1.194 2 | M10 | 0.873 3 |
| M03 | 0.975 6 | M07 | 0.964 0 | M11 | 0.920 3 |
| M04 | 0.921 2 | M08 | 0.864 3 | M12 | 0.942 2 |

剔除了季节因素外,建立带有截距项的ARMA(2,3)预测2012年1月至2015年12月的马铃薯价格(图2)。图3是图2经过季节调整后真实值与预测值结果。

根据甘肃重点农产品采集分析系统得到的数据,2015年1月份批发市场马铃薯价格为1.58元/kg,预测值是1.663 067元/kg,误差为4.99%,较好的预测了马铃薯价格变动。从图3可反映出,样本外预测即2015年马铃薯的价格整体是下降

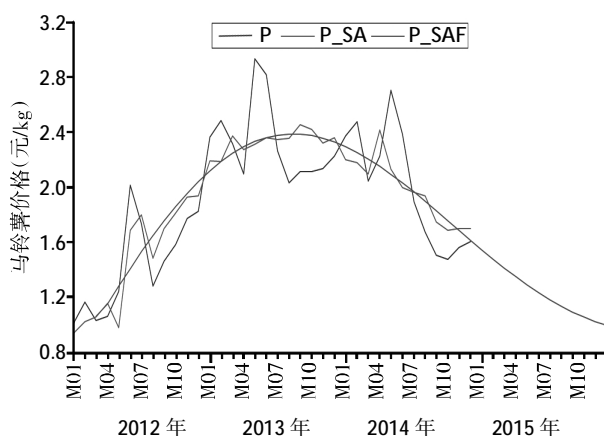


图2 P\_SA 95%置信区间预测值 P\_SAF

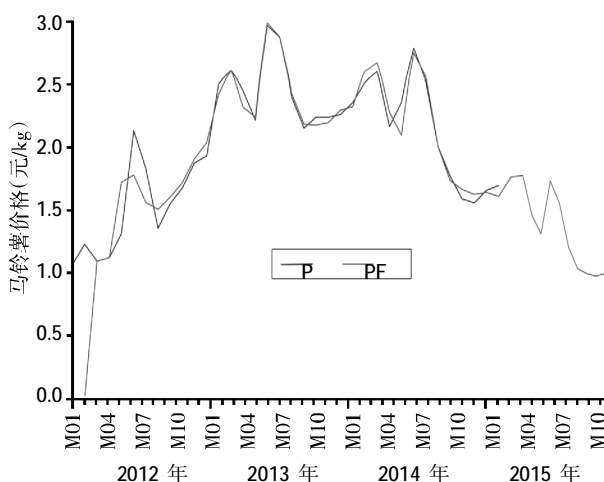


图3 ARMA(2,3)模型下马铃薯价格真实值与预测值拟合

的,而且表现出阶段性平滑趋势。

表2是采用马铃薯2012/01—2014/12为样本空间,经过季节性调整,根据季节性调整后序列的偏相关和自相关图建立的ARMA(2,3)模型,很好的模拟了马铃薯的价格变动趋势。相比采用2010/06—2014/12为样本空间建立的ARMA(2,3)模型优势明显,拟合优度更高。所以样本空间不是越大越好,选择恰当的历史数据作为样本空间,对构建合理的蔬菜价格预测模型至关重要。

$$p_t = (1.446 0 + 1.952 1 p_{t-1} - 0.962 0 p_{t-2} + \varepsilon_t + 2.766 4 \varepsilon_{t-1} + 2.566 1 \varepsilon_{t-2} - 0.798 1 \varepsilon_{t-3}) \times p\_sf_t$$

## 2.2 蔬菜整体均价波动分析与预测

在分析和预测了马铃薯价格之外,进一步分析甘肃重点农产品市场信息采集分析系统中监测的22种大众蔬菜的平均价格波动,并建立相应的价格预测模型。

表 2 ARMA (2, 3) 模型预测值与真实值比较

| 年月      | 预测值<br>(PF) | 真实值<br>(P) | 误差          | 年月      | 预测值<br>(PF) | 真实值<br>(P) | 误差          |
|---------|-------------|------------|-------------|---------|-------------|------------|-------------|
| 2012M01 | NA          | 1.01       | NA          | 2014M01 | 2.471 237 5 | 2.37       | -4.096 633  |
| 2012M02 | NA          | 1.16       | NA          | 2014M02 | 2.555 023 1 | 2.47       | -3.327 683  |
| 2012M03 | 1.032 089 7 | 1.03       | -0.202 476  | 2014M03 | 2.147 468 7 | 2.04       | -5.004 435  |
| 2012M04 | 1.057 639 4 | 1.06       | 0.223 191 8 | 2014M04 | 1.978 640 4 | 2.22       | 12.198 253  |
| 2012M05 | 1.621 891 1 | 1.24       | -23.546 04  | 2014M05 | 2.652 493 8 | 2.70       | 1.791 002 6 |
| 2012M06 | 1.676 896 1 | 2.01       | 19.864 316  | 2014M06 | 2.421 289 3 | 2.38       | -1.705 26   |
| 2012M07 | 1.471 163 3 | 1.73       | 17.594 012  | 2014M07 | 1.891 257 2 | 1.89       | -0.066 476  |
| 2012M08 | 1.419 801 7 | 1.28       | -9.846 563  | 2014M08 | 1.636 867 5 | 1.67       | 2.024 138 5 |
| 2012M09 | 1.508 844 1 | 1.46       | -3.237 185  | 2014M09 | 1.570 017   | 1.50       | -4.459 631  |
| 2012M10 | 1.620 790 3 | 1.58       | -2.516 694  | 2014M10 | 1.530 970 4 | 1.47       | -3.982 465  |
| 2012M11 | 1.795 514 2 | 1.77       | -1.420 999  | 2014M11 | 1.547 819 2 | 1.56       | 0.786 962 9 |
| 2012M12 | 1.919 738 3 | 1.82       | -5.195 411  | 2014M12 | 1.517 912 7 | 1.60       | 5.407 909 1 |
| 2013M01 | 2.282 041 1 | 2.36       | 3.416 190 6 | 2015M01 | 1.663 067   | 1.58       | -4.994 81   |
| 2013M02 | 2.480 166 8 | 2.48       | -0.006 725  | 2015M02 | 1.673 911 7 | 1.66       | -0.838 1    |
| 2013M03 | 2.186 545 8 | 2.31       | 5.646 081 5 | 2015M03 | 1.373 345 3 |            |             |
| 2013M04 | 2.109 061 2 | 2.09       | -0.903 776  | 2015M04 | 1.238 924 7 |            |             |
| 2013M05 | 2.954 397 2 | 2.93       | -0.825 794  | 2015M05 | 1.631 631 7 |            |             |
| 2013M06 | 2.813 157 8 | 2.81       | -0.112 25   | 2015M06 | 1.468 717 6 |            |             |
| 2013M07 | 2.288 199   | 2.26       | -1.232 367  | 2015M07 | 1.136 005 7 |            |             |
| 2013M08 | 2.058 833 6 | 2.03       | -1.400 484  | 2015M08 | 0.978 095 2 |            |             |
| 2013M09 | 2.049 4386  | 2.11       | 2.955 021 7 | 2015M09 | 0.937 974   |            |             |
| 2013M10 | 2.070 387 1 | 2.11       | 1.913 311   | 2015M10 | 0.919 435 9 |            |             |
| 2013M11 | 2.164 444 3 | 2.13       | -1.591 368  | 2015M11 | 0.939 799 9 |            |             |
| 2013M12 | 2.190 482 1 | 2.22       | 1.347 550 8 | 2015M12 | 0.937 387 4 |            |             |

2.2.1 2010 年 1 月至 2014 年 12 月蔬菜价格整体波动分析 如果蔬菜价格变动以一年期为周期,从图 4 曲线 P 的变动来看,价格变化基本呈“V”型,每一年的均价从 1 月、2 月的最高价波动式或连续下降,到 7 月、8 月达到最低价,然后逐渐波动式或连续上涨,并在 11 月、12 月达到一年期的另一个价格高峰。从历年价格数据看,2010—2014 年,1 月、2 月的蔬菜均价最高是 3.54 元/kg,11 月、12 月的均价是 2.93 元/kg,11 月、12 月的价格高峰期较之于 1 月、2 月的价格高峰要低 0.61 元/kg。2010—2014 年蔬菜价格最低点是 7 月、8 月,均价是 2.12 元/kg,较 1 月、2 月价格高峰 3.54 元低 1.42 元/kg。图 4 蔬菜价格曲线 P 同时反

映出,批发市场蔬菜价格 2010—2014 年没有明显的涨价趋势,但是消费却明显的感觉到了蔬菜价格有明显上涨,这是图形反映的问题,本文不做深入讨论。

季节调整后的价格序列 P\_SA 和原始价格序列相比,波动性明显趋缓,剔除季节影响后,P\_SA 价格序列均处在同一环境下,季节调整是对时间序列预测的基本要求。

图 4 是蔬菜价格波动与季节因子曲线,其基本趋势是季节因子与蔬菜价格同方向变动,蔬菜价格随着季节因子的下降而下降,随着季节因子的上升而上升。这不仅说明蔬菜价格预测中季节调整的重要性,也说明了对季节调整序列预测后

通过季节因子还原的必要性。

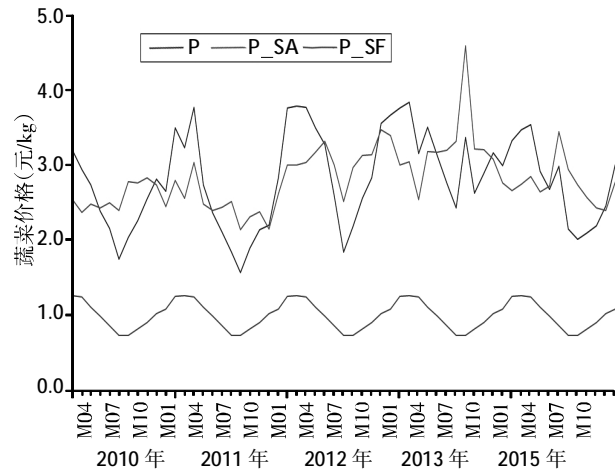


图 4 2010—2014 年蔬菜价格波动趋势

2.2.2 季节调整与平稳性检验 被监测 22 种大众蔬菜的季节因子显示，每年 1 月、2 月、3 月、4 月、11 月、12 月的季节因子大于 1，季节因素对蔬菜整体价格起到拉高作用；每年 5 月、6 月、7 月、8 月、9 月、10 月的季节因子小于 1，季节因素对蔬菜整体价格起拉低作用。其中 2 月份的季节因子最大，为 1.261 5，即 2 月份季节性因素对蔬菜价格的整体拉高作用最大，7 月份的季节因子最小是 0.731 7，即 7 月份季节因素对蔬菜价格的整体拉低效应最明显。

表 3 蔬菜价格季节因子

| 月份  | 季节因子      | 月份  | 季节因子      | 月份  | 季节因子      |
|-----|-----------|-----|-----------|-----|-----------|
| M01 | 1.253 037 | M05 | 0.984 429 | M09 | 0.817 542 |
| M02 | 1.261 541 | M06 | 0.863 970 | M10 | 0.900 843 |
| M03 | 1.241 088 | M07 | 0.731 652 | M11 | 1.025 496 |
| M04 | 1.104 711 | M08 | 0.734 490 | M12 | 1.081 201 |

表 4 是季节调整后序列 P\_SA 的平稳性检验结果，检验结果显示，P\_SA 是平稳性时间序列，严格拒绝具有单位根的原假设。因此 ARIMA 模型的 I=0，可以对 P\_SA 时间序列建立 ARMA (P, Q) 模型。

表 4 P\_SA 平稳性检验

Null Hypothesis: P\_SA has a unit root  
 Exogenous: Constant  
 Lag Length: 0 (Automatic - based on SIC, maxlag=10)

|  | t-Statistic | Prob.*  |
|--|-------------|---------|
| Augmented Dickey-Fuller test statistic | -3.798 854  | 0.004 9 |
| Test critical values: 1% level         | -3.548 208  |         |
| 5% level                               | -2.912 631  |         |
| 10% level                              | -2.594 027  |         |

### 2.3 模型建立与价格预测

根据 P\_SA 序列的自相关、偏相关检验，P\_SA 偏相关系数在 P=2 之后截尾，自相关系数在 q=2 之后结尾，所以 P\_SA 是一个 ARMA (2, 2) 过程，建立相应的 ARMA (2, 2) 模型。带有常数项的 ARMA (2, 2) 模型各参数估计结果均显著，如表 5。

表 5 ARMA (2, 2) 参数估计结果

Dependent Variable: P\_SA  
 Method: Least Squares  
 Date: 03/09/15 Time: 17: 36  
 Sample (adjusted): 2010M04 2014M12  
 Included observations: 57 after adjustments  
 Convergence achieved after 59 iterations  
 MA Backcast: 2010M02 2010M03

|                    | Coefficient | std.Nrror             | t-Statistic | Prob.   |
|--------------------|-------------|-----------------------|-------------|---------|
| C                  | 2.842 255   | 0.088 695             | 32.045 14   | 0.000 0 |
| AR (1)             | 1.321 985   | 0.202 318             | 6.534 179   | 0.000 0 |
| AR (2)             | -0.682 986  | 0.176 045             | -3.879 606  | 0.000 3 |
| MA (1)             | -0.975 524  | 0.180 565             | -5.402 620  | 0.000 0 |
| MA (2)             | 0.693 477   | 0.141 399             | 4.904 394   | 0.000 0 |
| R-squared          | 0.378 036   | Mean dependent var    | 2.834 355   |         |
| Adjusted R-squared | 0.330 193   | S.D. dependent var    | 0.412 648   |         |
| S.E. of regression | 0.337 718   | Akaike info criterion | 0.750 422   |         |
| Sum squared resid  | 5.930 794   | Schwarz criterion     | 0.929 638   |         |
| Log likelihood     | -16.387 04  | Hannan-Quinn criter.  | 0.820 072   |         |
| F-statistic        | 7.901 537   | Durbin-Watson stat    | 1.763 482   |         |
| Prob (F-statistic) | 0.000 047   |                       |             |         |

蔬菜均价预测模型是：

$$p_t = (2.842 3 + 1.321 9 p_{sa,t-1} - 0.682 9 p_{sa,t-2} + \epsilon_t - 0.975 5 \epsilon_{t-1} + 0.693 5 \epsilon_{t-2}) \times p_{s,t}$$

根据已有的样本数据得到了甘肃省 2015 年 22 种大众蔬菜的月度均价，需要注意的是，ARIMA 模型是动态的，随着新的监测数据不断更新，样本空间也在不断扩大，模型所依赖的历史数据发生变化，时间序列的稳定性和变化趋势也在发生变化，因此具体的 ARIMA (AR、ARMA、MA、ARIMA) 模型也在变化，是否带有截距项也需要重新考察，模型的 P、D、Q 都需要重新界定，参数也需要重新估计。总之，动态的时间序列对应的是一个动态的 ARIMA 模型。



图5是以2010年2月(2010年1月数据缺失)至2014年12月为样本空间对2015年蔬菜价格趋势预测,表6是2015年的预测值。从1月和2月的预测结果来看,误差保持在10%以内,预测效果较好。

从图5和表6可知,2015年甘肃省22种监测大众蔬菜价格变动基本呈现“V”型,最高价出现在3月份(3.67元/kg),最低价出现在7月份(2.15元/kg)。

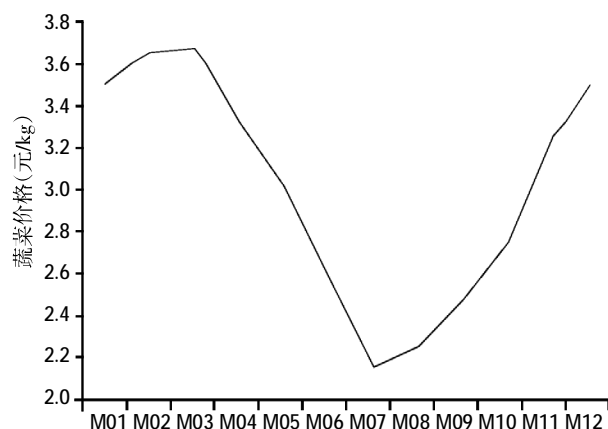


图5 2015年蔬菜价格趋势预测

表6 误差分析

| 时间      | 预测值 (PF) | 真实值 (P) | 误差     | 时间      | 预测值 (PF) |
|---------|----------|---------|--------|---------|----------|
| 2015M01 | 3.475    | 3.16    | 9.974  | 2015M07 | 2.152    |
| 2015M02 | 3.635    | 3.65    | -0.396 | 2015M08 | 2.251    |
| 2015M03 | 3.661    |         |        | 2015M09 | 2.468    |
| 2015M04 | 3.296    |         |        | 2015M10 | 2.734    |
| 2015M05 | 2.996    |         |        | 2015M11 | 3.230    |
| 2015M06 | 2.571    |         |        | 2015M12 | 3.468    |

### 3 小结与讨论

1) 季节因素对蔬菜价格波动影响显著,因此在进行蔬菜价格预测时,有必要将季节影响因子剔除,然后预测季节调整后的价格序列,得到季节调整后的预测价格并通过季节因子还原,便得到蔬菜价格真正的预测值。ARIMA模型是蔬菜价格预测较为理想的模型,经研究发现,ARIMA模型中的I,也就是价格序列平稳性对蔬菜价格预测有重要影响,通常在建立该模型之前必须要检验时间序列的平稳性,模型选择的关键还有P和Q的选择,

也就是一般意义上的ARIMA(P, Q)模型。特殊的会用到AR(P)、MA(Q)模型、ARMA(P, Q)。

2) 甘肃省蔬菜价格季节变动明显,但从2010至2014年的平均批发价格水平来看,不存在明显的涨价趋势,但是现实生活中,消费者却明显感到了菜价上涨,因此,有必要对批发价格和零售价格进行对比分析。

### 参考文献:

- [1] FRANCIS X. DIEDBOLD. 经济预测基础教程[M]. 杜江,李恒,译. 北京:机械工业出版社,2012.
- [2] 单玉隆. ARIMA模型与遗传算法优化神经网络在GDP预测中的应用[D]. 兰州:兰州大学,2014.
- [3] 赵丽平,王雅鹏. 2015年中国蛋鸭产业市场研究与预测[J]. 中国家禽,2015,37(6):33-36.
- [4] 王振寰,杨堰锬,张峰,等. ARIMA模型在城镇居民家庭人均收入的应用[J]. 内蒙古农业大学学报,2014,35(1):165-169.
- [5] 王振寰,杨婷. ARIMA模型在农村居民消费水平预测中的应用[J]. 内蒙古农业大学学报(自然科学版),2015,36(3):163-168.
- [6] 何太淑,刘乃郗. 从降雨指数看粮食产量风险管理——以河北省石家庄为例[J]. 调研世界,2015(6):58-61.
- [7] 黄华,蔡仁,努尔古丽·艾力,等. 基于ARIMA和GA-Elman神经网络的新疆年降水耦合预测研究[J]. 新疆农业科学,2015,52(6):1093-1098.
- [8] 马晓涛,温继文,陈英义. 基于ARIMA和RBF神经网络模型的溶解氧预测分析[J]. 江苏农业科学,2015,43(5):413-415.
- [9] 魏同洋,靳乐山. 基于ARIMA模型的蔬菜虚拟水含量变动分析——以北京市为例[J]. 江苏农业科学,2014,42(8):331-334.
- [10] 田苗,王鹏新,张树誉,等. 基于条件植被温度指数的冬小麦产量预测[J]. 农业机械学报,2014,45(2):239-245.
- [11] 麦尔旦·吐尔孙,王雅鹏,闫建伟. 我国肉鸭产业经济发展与预测研究[J]. 中国家禽,2015,37(11):37-42.
- [12] 李嫣怡,刘荣,丁维岱. EVIEWS统计分析与应用[M]. 北京:电子工业出版社,2013.
- [13] JEFFREY M. WOOLDRIDGE. 计量经济学导论[M]. 费剑平,译. 北京:中国人民大学出版社,2010.
- [14] 韩雯. ARIMA模型在贵州省农产品价格预测中的应用——以辣椒为例[J]. 安徽农业科学,2011,39(21):13226-13227;13229.

(本文责编:郑立龙)