

Analysis and Forecast of Stock Price Law Based on ARIMA Model

Jia Shi, Wei Liu, Zhichao Feng, Chunhui Zhang, Yanxiang Zhou

College of Mechanical and Automotive Engineering, Qingdao University of Technology, Qingdao Shandong
Email: 965398928@qq.com

Received: Feb. 3rd, 2020; accepted: Feb. 17th, 2020; published: Feb. 24th, 2020

Abstract

The price of stock market involves many uncontrollable factors, and the relationship among them is complex. This paper takes the four transaction prices of Hanwang Technology Stock (OP, FP, SP, LP) from June 1 to September 21, 2019 as example to analyze and forecast. Firstly, this paper described the law and characteristics of price change, and then quantitatively analyzed the dependence of four kinds of transaction prices on time by using the least square method. Secondly, using d-order difference operation to make the non-stationary time series transformed into stationary time series. After the calculation of the autocorrelation coefficient ACF and partial correlation coefficient PACF of stationary time series, through the analysis of autocorrelation graph and partial correlation graph, we get the best order q and level p, and construct ARIMA model. Finally, use the ARIMA (p, d, q) model which is passed the test for prediction. The results show that the model is simple, accurate and the fitting effect is good.

Keywords

Stock Price, Least Squares Method, ARIMA, Forecast, Difference Calculation

基于ARIMA模型的股市价格规律分析与预测

石佳, 刘威, 冯智超, 张春晖, 周雁祥

青岛理工大学机械与汽车工程学院, 山东 青岛
Email: 965398928@qq.com

收稿日期: 2020年2月3日; 录用日期: 2020年2月17日; 发布日期: 2020年2月24日

摘要

股市价格涉及诸多不可控因素, 各个因素之间关系错综复杂。本文以汉王科技股票2019年6月1日~9月

21日的4种交易价格OP, FP, SP, LP为例进行分析与预测。首先对其变化规律和特征进行了描述, 然后利用最小二乘法, 定量分析了4种交易价格与时间 t 的依赖关系。其次对非平稳时间序列进行 d 阶差分运算, 化为平稳时间序列, 再分别求出平稳时间序列的自相关系数ACF和偏相关系数PACF, 通过对自相关图和偏相关图的分析得到最佳阶数 q 和阶层 p , 构建了ARIMA模型。最后对所建的ARIMA (p, d, q)模型进行残差检验, 并利用检验通过的模型进行预测, 结果显示所建模型较为简单、精确且拟合效果较好。

关键词

股票价格, 最小二乘法, ARIMA, 预测, 差分运算

Copyright © 2020 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

股市是掌握经济整体运行情况的重要参考条件之一, 一直以来都是经济学研究的热点话题。尽管股票价格在短期内看似无序, 但从长期时间序列来分析却有一定的自然规律。由于股票价格不仅受限于国家政策和金融状况, 还在一定程度上受大众心理等多种因素的交互影响。尽管股价不能为人们准确地预测, 但可以用数学建模的方法来加以描述和刻画, 国内外相关学者对此做过诸多研究。

苑小康[1]在几何 Brown 运动模型中加入 Poisson 跳跃参数来构建股票价格动态模型, 通过计算机编程实现仿真过程并进行短期预测; 荣馨兵[2]提出了基于模式搜索的时间序列预测方法, 可以适应复杂变化的股票数据, 能够为多维数据做出预测并给出置信区间; 卢昊宽[3]提出一种隐马尔科夫模型及预测策略, 该模型使用遗传算法将隐马尔科夫模型的初试状态参数进行优化。Zhou 等[4]使用多种异构数据源进行融合, 建立了支持向量机对单个股票在不同活动水平下的预测模型; Manas 等[5]利用 ANN 神经网络和改进粒子群优化(PSO)混合技术, 提出了智能股票价格预测模型; Wang [6]利用模糊集理论和经典时间序列预测方法构建了一种利用模糊时间序列进行股票价格预测的大数据框架, 进行股票价格的预测; Umoh 提出了反向传播和监督学习的人工神经网络预测模型, 以梯度隶属方法和 Sugeno 型模糊推理引擎来优化预测结果。

上述学者提出的预测模型大多十分复杂, 求解需要耗费较多时间。本文针对这个缺点, 使用差分整合移动平均自回归模型(Auto Regressive Integrated Moving Average Model)简称 ARIMA 模型来预测, 这对毫无规律的股票价格数据给出时间序列分析, 能够从一个侧面给出较好的预测。

2. 股票价格规律分析

2.1. 股票交易价格的变化规律和特征

选定汉王科技[7] 2019年6月1日到2019年9月21日的四种交易价格: 开盘价(OP)、收盘价(FP)、最高价(SP)、最低价(LP)为例, 部分日期原始数据如图1所示, 以其原始数据为基础做出114天变化曲线图, 如图2所示。四种交易数据在总体趋势上保持一致, 大致在前4天保持线性增长, 于第4天达到小范围极值。在第4到第31天这段时间里, 交易价格持续走低, 于第31天价格跌入谷底。在30~45天呈现增长趋势, 紧接着在45~58天期间, 价格波动不大。在59天到103天这段时间价格持续走低, 最后在104天到114天价格慢慢上涨。对于SP和LP数据, 可以清楚地看出SP在每一天的交易价格都大于LP,

且通过图 2 也可以判断 SP 为交易价格的最高价，而 LP 为交易价格的最低价。其数据特征见表 1。

2.2. 股票交易价格与时间 t 的变化依赖关系

为了量化描述股票交易价格与时间 t 的变化依赖关系，假设 2019.6.1 为 $t = 1$ 时刻，此时 $x_1 = 1$ ，2019.9.21 为 $t = 114$ 时刻，此时 $x_{114} = 114$ ，以此类推。已知函数 $f(x_i)$ 在若干点 $x_i (i = 1, 2, \dots, 114)$ 处的值为 y_i ，构造一个曲线拟合函数 $\psi(x)$ ，用来反映数据的基本趋势。



	t	VO交易量	OP开盘价	FP收盘价	SP最高价	LP最低价
1	19/06/01	217.46	14.04	15.00	15.00	14.04
2	19/06/02	122.37	14.60	15.09	15.16	14.60
3	19/06/03	99.77	15.10	15.65	15.68	15.02
4	19/06/04	66.32	15.66	15.65	15.87	15.55
5	19/06/05	45.77	15.60	15.28	15.64	15.24
6	19/06/06	20.60	15.20	15.22	15.60	15.11
7	19/06/07	23.28	15.21	15.09	15.25	15.04
8	19/06/08	19.29	15.05	15.12	15.22	15.05
9	19/06/09	35.81	15.12	15.20	15.40	14.97
10	19/06/10	22.94	15.11	15.11	15.22	15.10
11	19/06/11	21.40	15.07	15.17	15.24	15.07
12	19/06/12	14.54	15.18	15.16	15.20	15.12
13	19/06/13	21.77	15.16	15.15	15.21	15.12
14	19/06/14	19.89	15.16	15.09	15.17	15.03
15	19/06/15	24.93	15.07	14.75	15.09	14.73
16	19/06/16	17.66	14.73	14.81	14.85	14.66

Figure 1. Raw data of Hanwang Technology stocks at four trading prices during the period of 2019.6.1-9.21 OP, FP, SP, LP

图 1. 汉王科技股票 2019.6.1~9.21 期间四种交易价格 OP、FP、SP、LP 原始数据

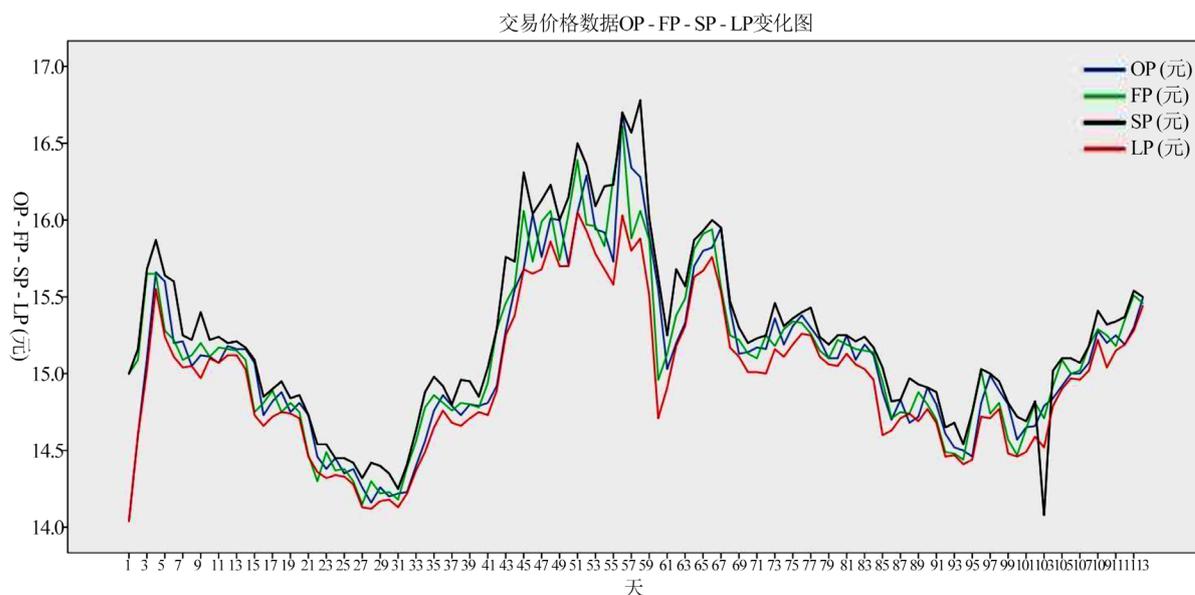


Figure 2. Data of the four trading prices OP, FP, SP, LP

图 2. 四种交易价格 OP、FP、SP、LP 数据变化图

Table 1. Data characteristics of OP, FP, SP, LP
表 1. OP、FP、SP、LP 的数据特征

交易价格	OP/元	FP/元	SP/元	LP/元
有效数量	114	114	114	114
平均值	15.0906	15.1118	15.2424	14.9641
中位数	15.0800	15.1050	15.2000	14.9850
标准差	0.52786	0.51867	0.56571	0.47366
范围	2.66	2.46	2.70	2.01
最小值	14.04	14.15	14.08	14.04
最大值	16.70	16.61	16.78	16.05

2.2.1. 曲线拟合的最小二乘法

对于曲线拟合函数 $\psi(x)$ ，不要求其严格的通过所有数据点，也就是说拟合函数 $\psi(x)$ 在 x_t 处的偏差(或残差)不都严格的等于零，设偏差为 ε_t ：

$$\varepsilon_t = \psi(x_t) - f(x_t), \quad (t=1, 2, \dots, 114) \quad (1)$$

为了使近似曲线能最大精度反映所给数据点的变化趋势，要求总偏差 ε_t 平方和最小，即要求下式具有最小值：

$$\sum_{t=1}^{114} \varepsilon_t^2 = \sum_{t=1}^{114} [\psi(x_t) - f(x_t)]^2 \quad (2)$$

假设拟合方程为二次曲线：

$$y = ax^2 + bx + c \quad (3)$$

已知数据点 (x_t, y_t) ， $t=1, 2, \dots, 114$ ，该近似拟合曲线的均方误差为：

$$Q(a, b, c) = \sum_{t=1}^{114} ax_t^2 + bx_t + c - y_t^2 \quad (4)$$

要求式 4 的极小值，可以归结为多元函数求极值的问题，则有：

$$\begin{cases} \frac{\partial Q}{\partial a} = 0 \\ \frac{\partial Q}{\partial b} = 0 \\ \frac{\partial Q}{\partial c} = 0 \end{cases} \quad (5)$$

求导过程如下：

$$\frac{\partial Q}{\partial a} = 2 \sum_{t=1}^{114} (ax_t^2 + bx_t + c - y_t) \cdot x_t^2 = 2 \sum_{t=1}^{114} (ax_t^4 + bx_t^3 + cx_t^2 - y_t x_t^2) = 0$$

$$\Rightarrow a \sum_{t=1}^{114} x_t^4 + b \sum_{t=1}^{114} x_t^3 + c \sum_{t=1}^{114} x_t^2 = \sum_{t=1}^{114} y_t x_t^2$$

$$\frac{\partial Q}{\partial b} = 2 \sum_{t=1}^{114} (ax_t^2 + bx_t + c - y_t) \cdot x_t = 2 \sum_{t=1}^{114} (ax_t^3 + bx_t^2 + cx_t - y_t x_t) = 0$$

$$\Rightarrow a \sum_{t=1}^{114} x_t^3 + b \sum_{t=1}^{114} x_t^2 + c \sum_{t=1}^{114} x_t = \sum_{t=1}^{114} y_t x_t$$

$$\frac{\partial Q}{\partial c} = 2 \sum_{t=1}^{114} (ax_t^2 + bx_t + c - y_t) = 0 \Rightarrow a \sum_{t=1}^{114} x_t^2 + b \sum_{t=1}^{114} x_t + 114c = \sum_{t=1}^{114} y_t$$

可将上述线性方程组转换为矩阵方程，即：

$$\begin{pmatrix} \sum_{t=1}^{114} x_t^2 & \sum_{t=1}^{114} x_t & 114 \\ \sum_{t=1}^{114} x_t^3 & \sum_{t=1}^{114} x_t^2 & \sum_{t=1}^{114} x_t \\ \sum_{t=1}^{114} x_t^4 & \sum_{t=1}^{114} x_t^3 & \sum_{t=1}^{114} x_t^2 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} \sum_{t=1}^{114} y_t \\ \sum_{t=1}^{114} y_t x_t \\ \sum_{t=1}^{114} y_t x_t^2 \end{pmatrix} \quad (6)$$

其系数行列式为：

$$D = \begin{vmatrix} \sum_{t=1}^{114} x_t^2 & \sum_{t=1}^{114} x_t & 114 \\ \sum_{t=1}^{114} x_t^3 & \sum_{t=1}^{114} x_t^2 & \sum_{t=1}^{114} x_t \\ \sum_{t=1}^{114} x_t^4 & \sum_{t=1}^{114} x_t^3 & \sum_{t=1}^{114} x_t^2 \end{vmatrix} \quad (7)$$

若 D 值不为 0，则方程组有解。其中：

$$D_a = \begin{vmatrix} \sum_{t=1}^{114} y_t & \sum_{t=1}^{114} x_t & 114 \\ \sum_{t=1}^{114} y_t x_t & \sum_{t=1}^{114} x_t^2 & \sum_{t=1}^{114} x_t \\ \sum_{t=1}^{114} y_t x_t^2 & \sum_{t=1}^{114} x_t^3 & \sum_{t=1}^{114} x_t^2 \end{vmatrix}, D_b = \begin{vmatrix} \sum_{t=1}^{114} x_t^2 & \sum_{t=1}^{114} y_t & 114 \\ \sum_{t=1}^{114} x_t^3 & \sum_{t=1}^{114} y_t x_t & \sum_{t=1}^{114} x_t \\ \sum_{t=1}^{114} x_t^4 & \sum_{t=1}^{114} y_t x_t^2 & \sum_{t=1}^{114} x_t^2 \end{vmatrix}, D_c = \begin{vmatrix} \sum_{t=1}^{114} x_t^2 & \sum_{t=1}^{114} x_t & \sum_{t=1}^{114} y_t \\ \sum_{t=1}^{114} x_t^3 & \sum_{t=1}^{114} x_t^2 & \sum_{t=1}^{114} y_t x_t \\ \sum_{t=1}^{114} x_t^4 & \sum_{t=1}^{114} x_t^3 & \sum_{t=1}^{114} y_t x_t^2 \end{vmatrix}$$

那么由此可以解得二次曲线拟合方程的各系数：

$$a = \frac{D_a}{D}, b = \frac{D_b}{D}, c = \frac{D_c}{D} \quad (8)$$

当拟合方程为三次曲线或四次曲线时同理。

2.2.2. 依赖关系定量化求解结果

观察图 2：四种交易价格变化图可以发现 OP, FP, SP, LP 在 114 天内的变化趋势与四次函数图像“W”的形状及其相似，因此利用上述曲线拟合的最小二乘法对 OP- t 数据做四次函数拟合，得到股票交易价格 OP 与时间 t 的变化依赖关系为：

$$y_t^{\text{OP}} = 4 \times 10^{-7} x_t^{\text{OP}4} - 9.3 \times 10^{-5} x_t^{\text{OP}3} + 6.9 \times 10^{-3} x_t^{\text{OP}2} - 0.16 x_t^{\text{OP}} + 16 \quad (9)$$

其中： y_t^{OP} 为 t 时刻交易价格 OP 的值。

x_t^{OP} 是初始值为 1，公差为 1 的等差数列。

通过 Matlab 计算得到残差模=4.0023，OP- t 拟合曲线与残差模量如图 3 所示。

同理，股票交易价格 FP 与时间 t 的变化依赖关系为：

$$y_t^{\text{FP}} = 4.6 \times 10^{-7} x_t^{\text{FP}4} - 1.1 \times 10^{-4} x_t^{\text{FP}3} + 8.1 \times 10^{-3} x_t^{\text{FP}2} - 0.2 x_t^{\text{FP}} + 16 \quad (10)$$

通过 Matlab 计算得到残差模=3.4719，FP- t 拟合曲线与残差模量如图 4 所示。

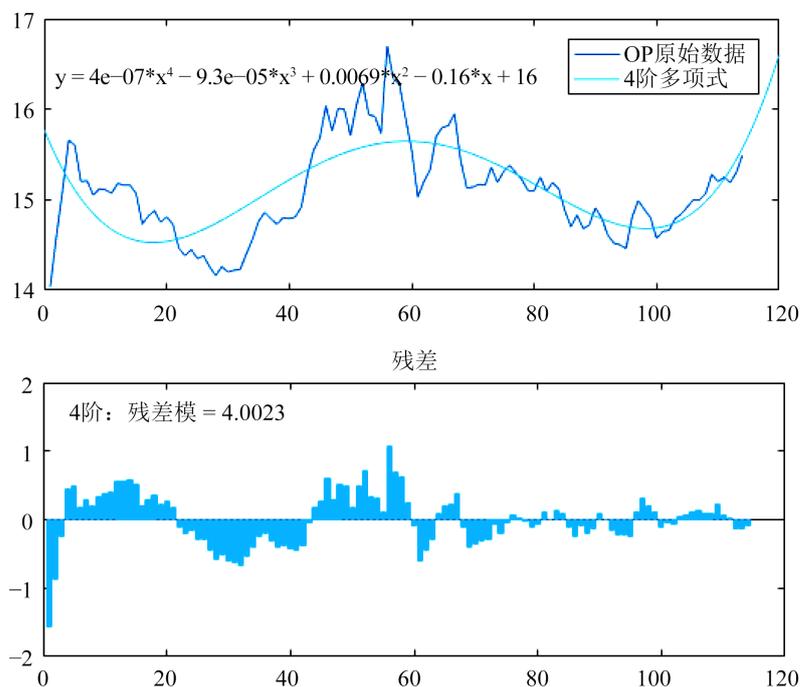


Figure 3. OP- t polynomial fitting curve and residual plot

图 3. OP- t 多项式拟合曲线与残差图

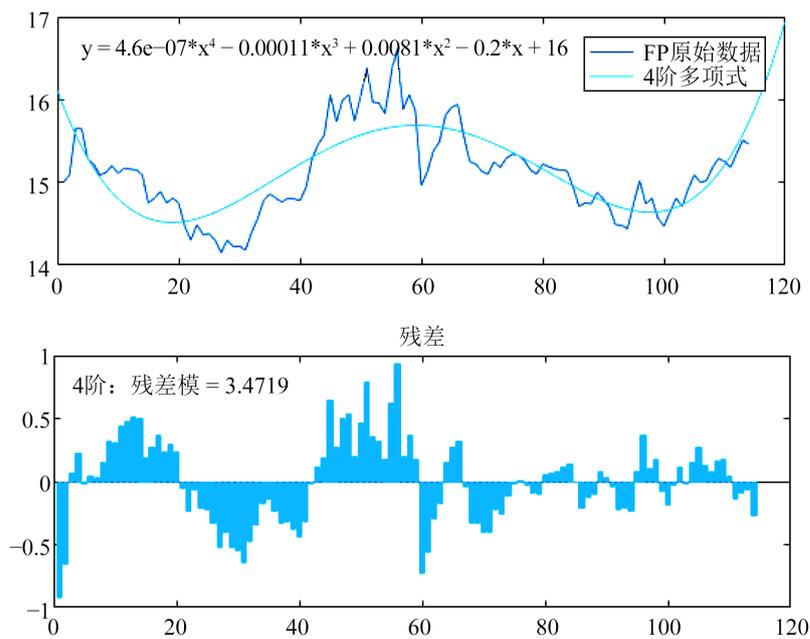


Figure 4. FP- t polynomial fitting curve and residual plot

图 4. FP- t 多项式拟合曲线与残差图

股票交易价格 SP 与时间 t 的变化依赖关系为:

$$y_t^{SP} = 4.9 \times 10^{-7} x_t^{SP4} - 1.2 \times 10^{-4} x_t^{SP3} + 8.6 \times 10^{-3} x_t^{SP2} - 0.21 x_t^{SP} + 16 \quad (11)$$

通过 Matlab 计算得到残差模=3.794, SP- t 拟合曲线与残差模量如图 5 所示。

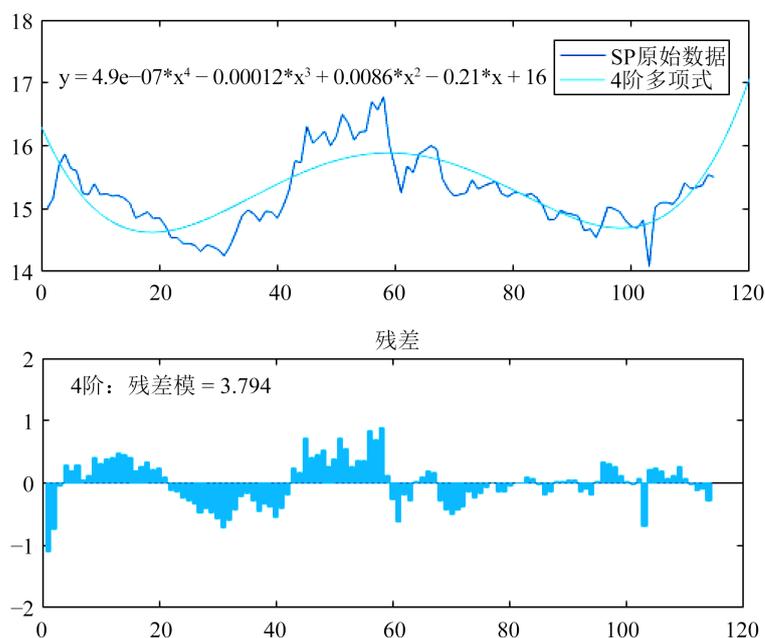


Figure 5. SP- t polynomial fitting curve and residual plot

图 5. SP- t 多项式拟合曲线与残差图

股票交易价格 LP 与时间 t 的变化依赖关系为:

$$y_t^{LP} = 3.7 \times 10^{-7} x_t^{LP4} - 8.6 \times 10^{-4} x_t^{LP3} + 6.4 \times 10^{-3} x_t^{LP2} - 0.15 x_t^{LP} + 16 \quad (12)$$

通过 Matlab 计算得到残差模=3.5708, LP- t 拟合曲线与残差模量如图 6 所示。

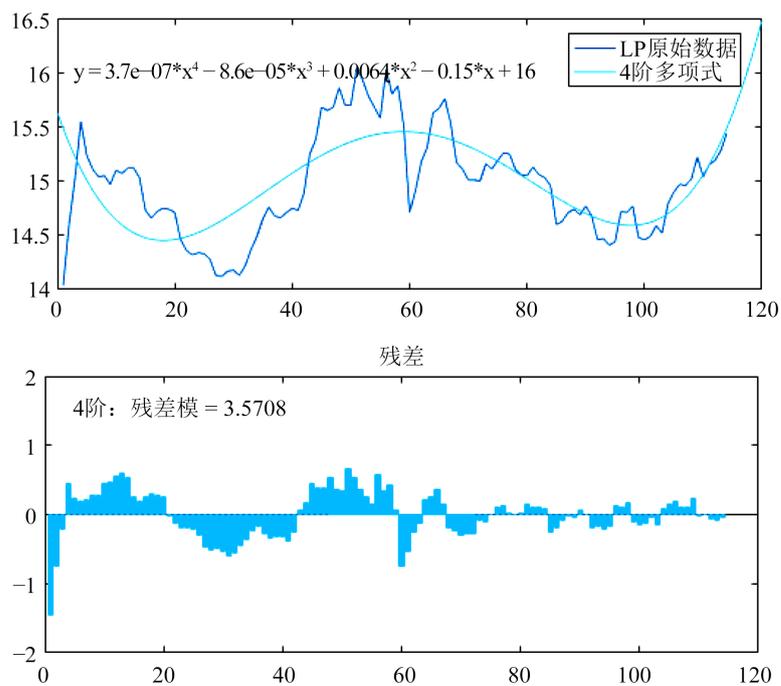


Figure 6. LP- t polynomial fitting curve and residual plot

图 6. LP- t 多项式拟合曲线与残差图

3. 股票价格预测

3.1. 时间序列 ARIMA 模型建立

ARIMA 模型也记作 $ARIMA(p, d, q)$, 是统计模型中最常见的一种用来进行时间序列预测的模型, 三个参数: p : 代表预测模型中采用的时间序列数据本身的滞后数, 也称为 Auto-Regressive (AR)项; d : 代表时序数据需要进行几阶差分才是稳定的, 也称为 Intergrated 项; q : 代表预测模型总采用的预测误差的滞后数, 也称为 Moving Average (MA)项。

ARIMA 模型十分简单, 只需内生变量而不需借助其他外生变量。但模型要求时间序列是稳定的, 或者通过过差分运算后是稳定的。

模型的识别和定阶主要是确定 p, d, q 三个参数, 差分的阶数 d 可以通过差分运算的差分图进行确定, 一般 2 阶差分图就已足够稳定, 假设 y_t 表示 t 时刻 Y 的差分, 那么:

- ① 当 $d = 0$ 时, $y_t = Y_t$;
- ② 当 $d = 1$ 时, $y_t = Y_t - Y_{t-1}$;
- ③ 当 $d = 2$ 时, $y_t = (Y_t - Y_{t-1}) - (Y_{t-1} - Y_{t-2}) = Y_t - 2Y_{t-1} + Y_{t-2}$ 。

p 和 q 阶数的确定与自相关系数 ACF 和偏自相关系数 PACF 以及截尾、拖尾有关, 具体确定依据见表 2。

Table 2. Criteria for determining p and q
表 2. p 和 q 的确定准则

模型(序列)	AR	MA	ARMA
ACF	拖尾	第 q 个后截尾	拖尾
PACF	第 p 个后截尾	拖尾	拖尾

其中, 自相关系数 ACF (Auto Correlation Function)描述的是时间序列观测值与其过去的观测值之间的线性相关性, 计算公式如式下:

$$ACF(k) = \frac{Cov(y_t, y_{t-k})}{Var(y_t)} \tag{14}$$

式中: k 表示滞后期数, 如果 $k = 2$, 那么描述的就是 y_t 和 y_{t-2} 之间的线性相关性。

偏相关系数 PACF (Partial Auto Correlation Function)描述的是在给定中间观测值的条件下, 时间序列预期观测值与过去观测值之间的线性相关性。几个简单的例子, 假设 $k = 3$, 那么我们描述的是 y_t 和 y_{t-3} 之间的相关性, 但是这个相关性还受到 y_{t-1} 和 y_{t-2} 的影响。PACF 剔除了这个影响, 而 ACF 包含这个影响。

表 2 中拖尾是指序列以指数率单调递减或震荡衰减, 而结尾指序列从某个时点变得非常小。

截尾判断依据: ① 在最初的 d 阶明显大于 2 倍标准差范围; ② 之后几乎 95%的(偏)自相关系数都落在 2 倍标准差范围以内; ③ 由非零自相关系数衰减为在零附近小值波动的过程非常突然。

拖尾判断依据: ① 超过 5%的样本(偏)自相关系数都落入 2 倍标准差范围之外; ② 或者由显著非零的(偏)自相关系数衰减为小值波动的过程比较缓慢或者非常连续。

通过上述方法可以确定 ARIMA 模型的参数 p, d, q 。所以当 p, d, q 已知时, ARIMA 的数学表达式为:

$$\hat{y}_t = \mu + \phi_1 \cdot y_{t-1} + \dots + \phi_p \cdot y_{t-p} + \theta_1 \cdot e_{t-1} + \dots + \theta_q \cdot e_{t-q} \tag{15}$$

其中: \hat{y}_t 表示 Y 的预测值, ϕ 表示 AR 的系数, θ 表示 MA 的系数。

3.2. ARIMA 模型的求解

根据上述建模方法,我们以股票交易价格 OP 数据为例进行模型求解。

Step 1: 由图 2 可以看出交易价格 OP 数据并不是平稳时间序列,因此由式 13 进行 1 阶和 2 阶差分处理。差分结果如图 7, 图 8 所示。

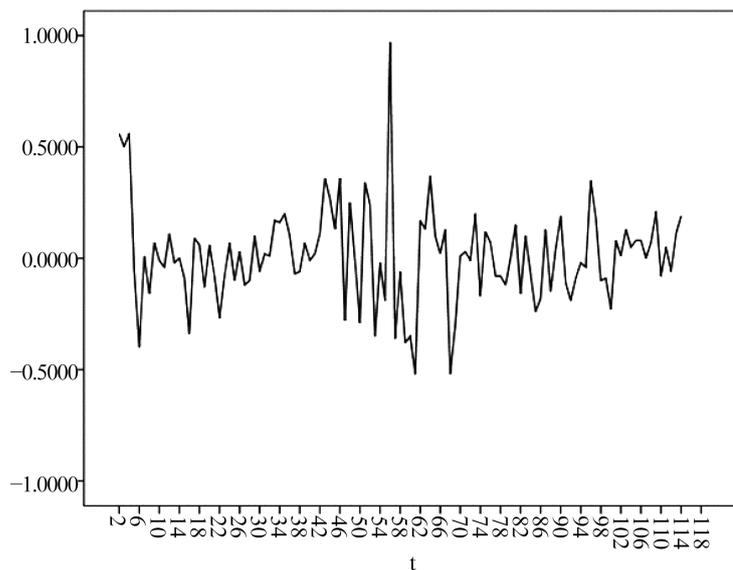


Figure 7. OP first order difference graph

图 7. OP 一阶差分图

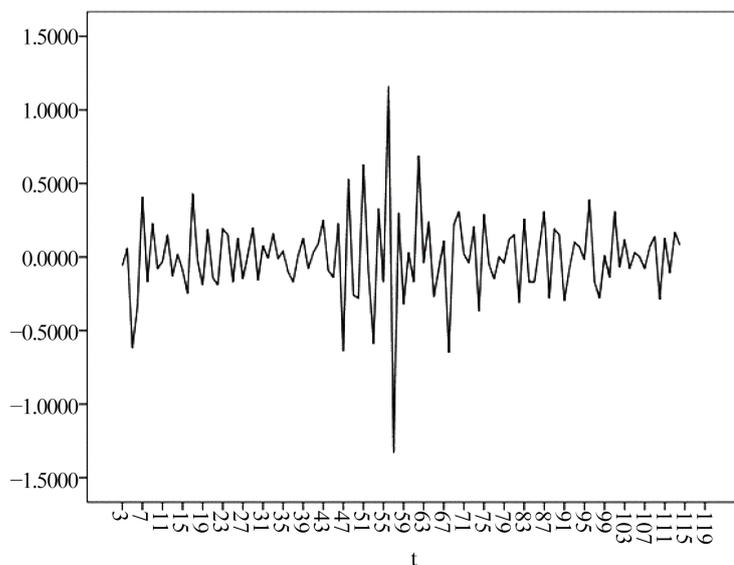


Figure 8. OP second-order difference graph

图 8. OP 二阶差分图

通过图 7 和图 8 得知 OP 一阶差分变化趋势已经基本稳定(在数值 0 上下波动),符合平稳时间序列的特点,所以可初步确定参数 d 为 1。

Step 2: 做出 OP 的自相关图和偏相关图,如图 9 与图 10 所示。

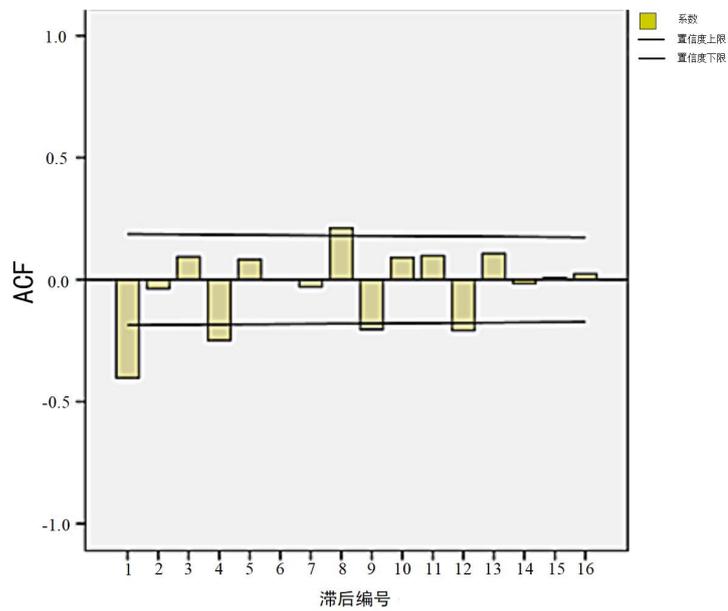


Figure 9. OP autocorrelation coefficient
图 9. OP 自相关系数

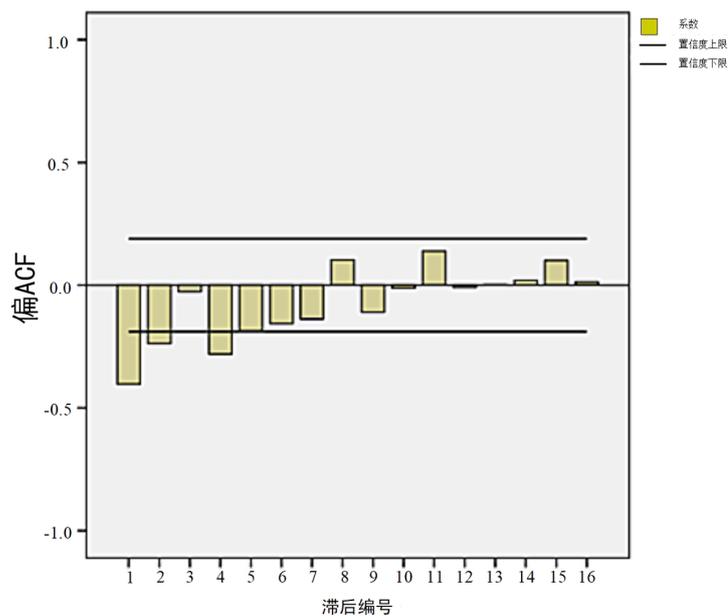


Figure 10. OP partial correlation coefficient
图 10. OP 偏相关系数

图 9 中 ACF 在第 1 阶后呈截尾状，图 10 中 PACF 在第 4 阶后呈拖尾状，因此可初步确定差分后序列适合 ARIMA (1, 1, 4)模型。

Step 3: 利用 SPSS 软件求解该模型，得到如图 11，图 12 的结果。图 11 可得显著性水平 Sig 为 0.298 大于 0.05，且 R 方为 0.826，拟合程度较好，初步判断模型的有效性。由图 12 可得 AR, MA 的系数分别为 0.253 和 0.15，因此对股票交易数据 OP 的预测模型结果为：

$$\hat{y}_t = \mu + 0.253y_{t-1} + 0.15e_{t-1} \tag{16}$$

模型	预测变量数	模型拟合度统计		杨-博克斯Q(18)			离群值数
		平稳R方	R方	统计	DF	显著性	
OP(元)-模型_1	1	.012	.826	18.449	16	.298	0

Figure 11. Statistical results of OP time series model

图 11. OP 时间序列模型统计结果

				估算	标准误差	t	显著性
OP(元)-模型_1	OP(元)开盘价	不转换	常量	.033	.048	.692	.491
			AR 延迟1	.253	.872	.289	.773
		差异	MA 延迟1	.150	.891	.168	.876
			t	不转换	分子	延迟0	.000

Figure 12. OP time series model parameters

图 12. OP 时间序列模型参数

Step 4: 以 9.22~9.26 的交易价格 OP 值作为预测目标, 利用式 16 作为预测方程, 预测结果如图 13 所示, 其中 UCL 表示预测值上限, LCL 表示预测值下限。将拟合结果与预测结果进行可视化处理, 如图 14 所示。

		115	116	117	118	119
CP(元)-模型1	预测	15.5065	15.5064	15.5019	15.4961	15.4897
	UCL	15.9432	16.1566	16.3178	16.4508	16.5658
	LCL	15.0697	14.8562	14.6860	14.5413	14.4135

Figure 13. OP price prediction result

图 13. 交易价格 OP 预测结果

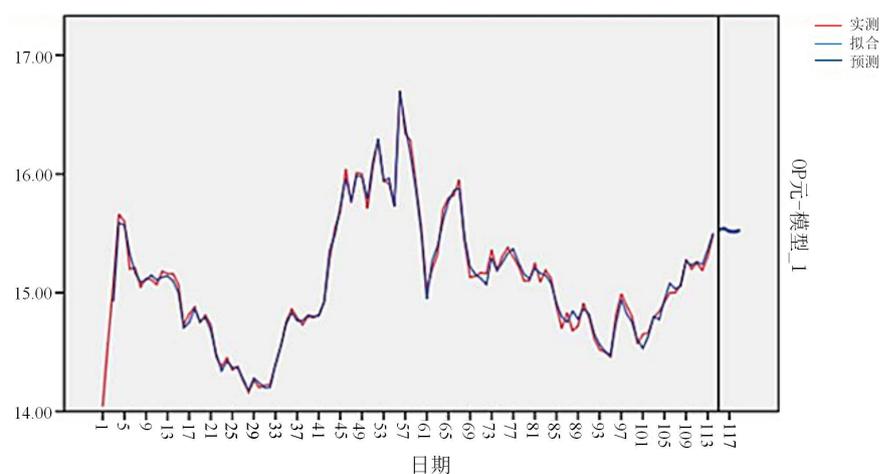


Figure 14. OP prediction model fitting results and prediction results

图 14. OP 预测模型拟合结果与预测结果

Step 5: 对 ARIMA (1, 1, 4)模型的参数进行残差分析检验, 如图 15 所示, 可以看出残差图基本上处于平稳状态, 且模型预测结果均在预测区间内, 因此可以断定 OP 预测模型 ARIMA (1, 1, 4)预测结果是合理的, 且拟合效果较好。

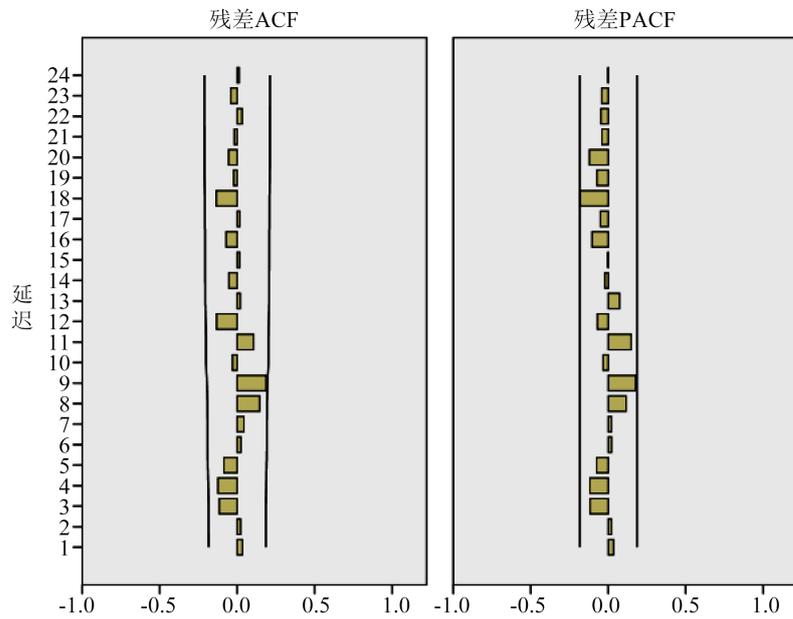


Figure 15. ARIMA (1, 1, 4) model residual plot
图 15. ARIMA (1, 1, 4)模型残差图

3.3. ARIMA 模型的应用

以股票交易价格 OP 为例，进行了模型的求解，初步验证了时间序列模型 ARIMA 的有效性。因此利用同样的方法对剩下的交易价格 FP, SP, LP 求解，求解过程与上述 OP 求解过程一样，不在详细叙述，在这里只展示经过残差分析后 ARIMA 模型的拟合结果与预测结果。

		115	116	117	118	119
FP(元)-模型1	预测	15.4880	15.4861	15.4370	15.4732	15.4903
	UCL	15.8895	16.0747	16.1198	16.2670	16.3391
	LCL	15.0866	14.8975	14.7541	14.6794	14.6416

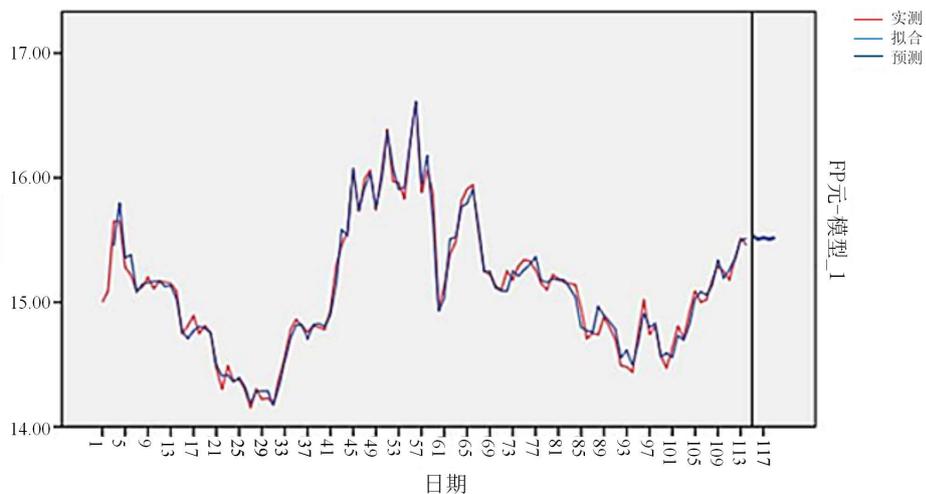


Figure 16. FP prediction model fitting results and prediction results
图 16. FP 预测模型拟合结果与预测结果

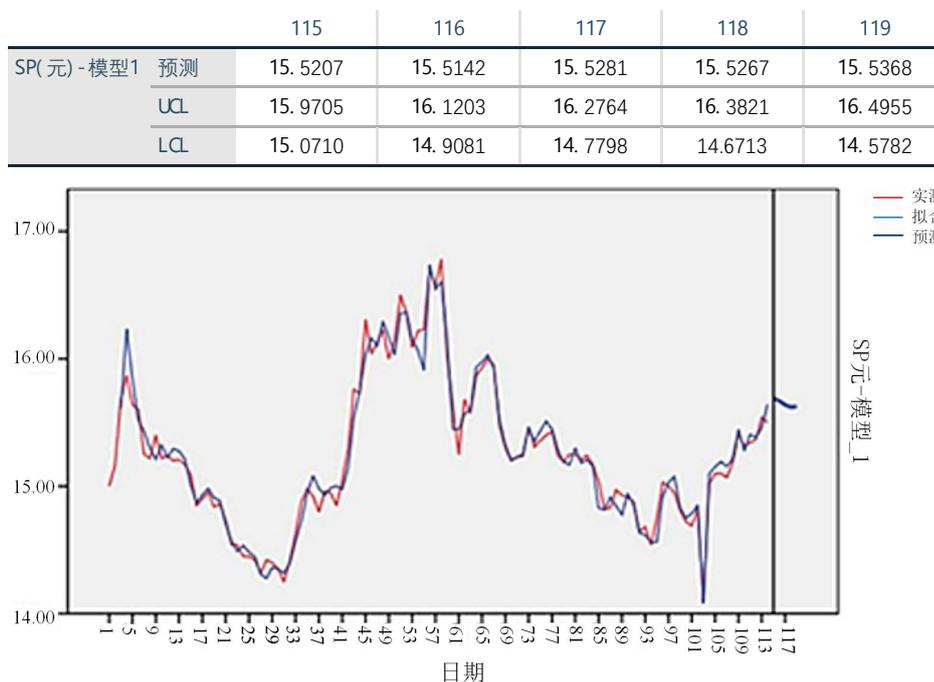


Figure 17. SP prediction model fitting results and prediction results

图 17. SP 预测模型拟合结果与预测结果

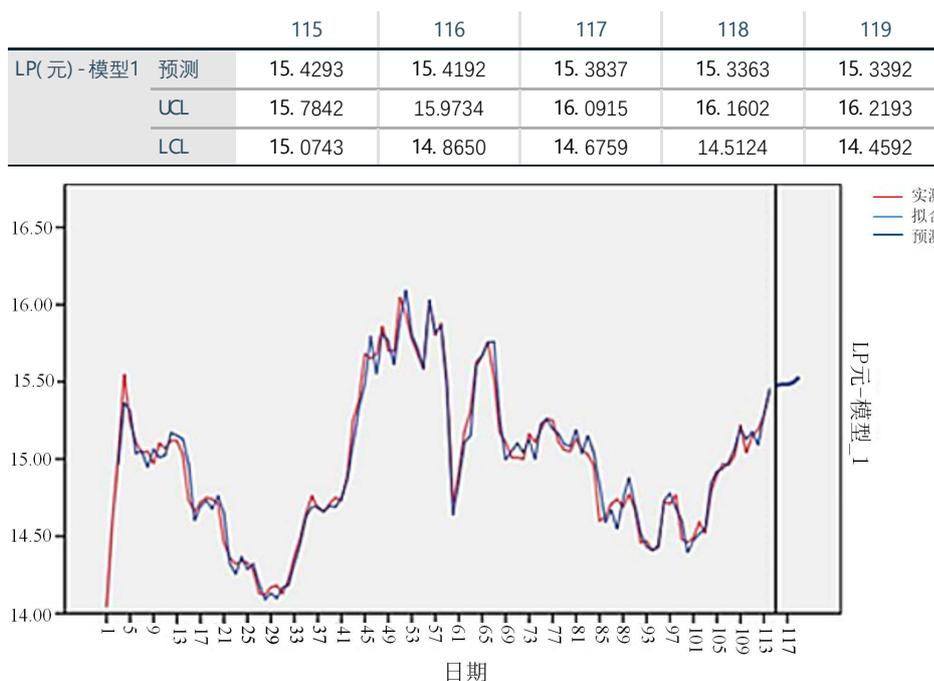


Figure 18. LP prediction model fitting results and prediction results

图 18. LP 预测模型拟合结果与预测结果

通过图 14~18, 可以说明 ARIMA 模型建立较为简单, 求解过程不复杂, 且拟合效果较为精确。该模型不仅可以应用于股票价格预测, 还可以应用于交通主干道年交通量预测、商品价格预测、仓库进出货量预测等等。

4. 结束语

通过提出算法的拟合与预测, ARIMA 模型[8]-[12]可以较为准确刻画股票交易价格波动变化特征,在一定程度上可以预测股票的价格走势。但是该模型仅仅适用于短期预测,波动频率较高或长期预测偏差较大,因为变幻莫测的交易市场,影响其价格的因素不仅与自身成本有关,还与国家宏观经济政策等息息相关。

参考文献

- [1] 苑小康. 基于随机系统理论的股票定价模型的研究[D]: [硕士学位论文]. 北京: 北京交通大学, 2018.
- [2] 荣馨兵. 基于相似搜索的时间序列预测方法及其应用[D]: [硕士学位论文]. 大连: 大连海事大学, 2018.
- [3] 卢昊宽. 基于隐马尔科夫的预测关键技术研究与应用[D]: [硕士学位论文]. 荆州: 长江大学, 2019.
- [4] Zhou, Z.B., Gao, M., Liu, Q., *et al.* (2019) Forecasting Stock Price Movements with Multiple Data Sources: Evidence from Stock Market in China. *Physica A: Statistical Mechanics and Its Applications*, Article ID: 123389. (In Press) <https://doi.org/10.1016/j.physa.2019.123389>
- [5] Senapati, M.R., Das, S. and Mishra, S. (2018) A Novel Model for Stock Price Prediction Using Hybrid Neural Network. *Journal of the Institution of Engineers (India): Series B*, **99**, 555-563. <https://doi.org/10.1007/s40031-018-0343-7>
- [6] Wang, W.N. (2018) A Big Data Framework for Stock Price Forecasting Using Fuzzy Time Series. *Multimedia Tools and Applications*, **77**, 10123-10134. <https://doi.org/10.1007/s11042-017-5144-5>
- [7] 汉王科技股票原始数据. <http://quote.eastmoney.com/sz002362.html>
- [8] 官银, 李新月, 朱家明. 基于 ARIMA 模型对我国农村人口的预测分析[J]. 山东农业工程学院学报, 2019, 36(7): 15-20.
- [9] 朱家明, 胡玲燕. 基于 ARIMA 和 BP 神经网络对人民币汇率预测的比较分析——以美元人民币汇率为例[J]. 重庆理工大学学报(自然科学), 2019, 33(5): 207-212.
- [10] 王珂, 王雅文, 朱家明. 基于 ARIMA 模型的安徽省就业发展趋势分析[J]. 佳木斯大学学报(自然科学版), 2018, 36(4): 621-624+654.
- [11] 张祎, 朱家明. 基于 ARIMA 和二次指数平滑模型的人口组合预测[J]. 牡丹江师范学院学报(自然科学版), 2017(1): 9-12.
- [12] 林杰, 朱家明, 陈富媛. 基于 ARMA 模型对股票“青岛海尔”成交量的分析预测[J]. 哈尔滨师范大学自然科学学报, 2016, 32(6): 6-10.