Price analysis of Paris polyphylla Smith based on ARIMA time series mode

Yan Guo*, Ziwei Hui1, Wenchuan Cheng2, and Yu Zou3

*1,2,3Sichuan Agricultural University, College of Information Engineering, Ya'an, Sichuan, China
*1Sichuan Agricultural University, Key Laboratory of Agricultural Information Engineering of Sichuan Province, Ya'an, Sichuan, China

Abstract. Paris polyphylla Smith, whose place of introduction is located in Yunnan, Sichuan and parts of Guizhou, has important medicinal value. For many years, scholars from home and abroad have done a lot of research on prediction of agricultural products and medicinal materials. However, there is little prediction about Paris polyphylla Smith. This paper establishes ARIMA time series model based on the market price of Paris polyphylla Smith from July 2013 to August 2018 and forecasts the price from September 2018 to July 2019. Compared with the price from September 2018 to July 2019, R^2 fitting coefficient is 0.954 and the relative error is within 0.072. The prediction of price strengthens the research about price fluctuation of Paris polyphylla Smith and improves the ability of drug farmers to resist risks of the price fluctuation of Paris polyphylla Smith, which will enable drug farmers to improve revenue. It has extremely significant meaning in the stability of the pharmaceutical market as well as guarantee the income of farmers.

1 Introduction

Paris polyphylla Smith is a type of precious traditional Chinese medicine with bitter taste and cold nature.Paris polyphylla Smith is the species that is most in demand and is widely used in traditional medicine, including treatment of back pain and bleeding. In traditional Chinese medicine, it is used as a component of topical medicaments indicated for treatment of boils, carbuncles, sore throat, venomous snake bite, and traumatic pain[1]. At the same time, it is one of the medicinal materials that are in block trading in China. But in recent years, with the development of society, due to the increasing demand for Paris polyphylla Smith, wild Paris polyphylla Smith has been over-exploited and the number of heavy buildings has decreased rapidly. Relevant technicians have optimized the planting and development of Paris polyphylla Smith through research, making it easier to plant Paris polyphylla Smith artificially and more drug farmers can plant it. Forecasting the future price of Paris polyphylla Smith can help drug farmers make more rational decisions and avoid market fluctuations and large losses.

Use At present, there are two main methods to predict the price of medicinal materials: factor analysis and price-time series analysis. Factor analysis method needs to comprehensively analyze the factors that cause price fluctuation, such as production cycle, natural disasters, supply-demand relationship, logistics transportation situation, to establish prediction models. The disadvantage of this method is that it can not collect complete and comprehensive information and can not determine the extent of the impact of various factors on price. Price-time series analysis is a widely used forecasting method at present such as ANFIS [2], grey neural network model, Markov chain [3, 4] and BP neural network [5], etc. All of them are based on historical data to establish forecasting models and forecast the future price trend according to the change law of historical data. On the one hand, this model only needs to collect historical price data to establish models. On the other hand, their drawback is that the sensitivity to data mutation is not very high and the requirement of integrity and continuity of historical data is relatively high. Although BP neural network has high self-learning ability and can approximate the non-linear function with arbitrary precision. It has ability to associative memory and prediction, which can be used to identify the price forecasting model. However, the randomness of its initial weights and thresholds can easily lead to slow convergence or fall into local optimal solution. ARIMA time series model [6] also belongs to the price-time series analysis method. The advantage of ARIMA is that it is more suitable for agricultural products with seasonal changes. Some scholars, such as Darekar, Ashwini, Reddy, etc, use ARIMA model to predict the market price of soybeans and obtain very reliable results [7]. That is to say ARIMA time series is very suitable for forecasting the price of Paris polyphylla Smith.

2 Modeling

The common stochastic time series analysis methods can be divided into stationary time series analysis and non-stationary time series analysis. Stationary time series analysis contains AR model, MA model and ARMA
model. Three kinds of non-stationary time series models mainly include ARIMA model and seasonal model. Among the five models mentioned above, AR (p) model, MA (q) model and ARMA (p, q) model are only suitable for stationary time series modeling and prediction. They can be regarded as special cases of ARIMA model. However, SARIMA model is suitable for time series analysis and modeling with seasonal periodic characteristics.

2.1 AR(p)p-order autoregressive model

In this model, the present value xt of time series is represented by a linear combination of the past values of the series and a white noise perturbation term ϵt.

\[ x_t = \varnothing_0 + \varnothing_1 x_{t-1} + \varnothing_2 x_{t-2} + \cdots + \varnothing_p x_{t-p} + \epsilon_t \]  

The above formula includes three restrictions:

\[ \varnothing_p \neq 0 \]  

• Random interference sequence is a zero-mean white noise sequence, namely \( \epsilon_t \sim WN(0, \delta^2) \)  

• Current random interference is independent of past sequence values, namely \( Ex_t \epsilon_s = 0, s < t \)  

After centralization the formula is:

\[ \Phi(B) x_t = \epsilon_t \]  

Among them:

\[ \Phi(B) = 1 - \varnothing_1 B - \varnothing_2 B^2 - \cdots - \varnothing_p B^p \]  

2.2 MA(q)q-order moving average model

The present value xt of time series is expressed by a linear combination of white noise perturbation terms. The formula for the calculation is as follows:

\[ x_t = \theta_0 - \theta_1 \epsilon_{t-1} - \theta_2 \epsilon_{t-2} - \cdots - \theta_q \epsilon_{t-q} \]  

The formula includes three restrictions:

\[ \theta_q \neq 0 \]  

• Random interference sequence is a zero-mean white noise sequence, namely \( \epsilon_t \sim WN(0, \delta^2) \)  

After centralization the formula is:

\[ x_t = \Theta(B) \epsilon_t \]  

Among them:

\[ \Theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q \]  

2.3 ARMA (p, q) Autoregressive Moving Average Model

The relationship between the present value xt of time series and its own value at the previous time and the disturbance term at the past time as shown in formula.

\[ x_t = u + \varnothing_1 x_{t-1} + \cdots + \varnothing_p x_{t-p} + \epsilon_t - \theta_1 \epsilon_{t-1} - \theta_2 \epsilon_{t-2} - \cdots - \theta_q \epsilon_{t-q} \]  

The constraints of the model are the same as those of AR(p) model and MA(q) model.

After centralization the formula is:

\[ \Phi(B) x_t = \Theta(B) \epsilon_t \]  

All unknown parameters, such as \( \varnothing, \theta, \delta \), need to be estimated using observation data.

2.4 Differential operation

Suppose \( \{x, t \in T\} \) is a time series and the subtraction operation between the two sequence values of one period apart is called the first-order difference operation. \( \nabla x_t \) is denoted as the first-order difference of \( x_t \), that is:

\[ \nabla x_t = x_t - x_{t-1} \]  

After processing the sequence of the first-order difference, we can obtain the second-order difference. \( \nabla^2 x_t \) is denoted as the second-order, that is:

\[ \nabla^2 x_t = \nabla x_t - \nabla x_{t-1} \]  

By analogy, the p-order difference can be obtained after the sequence of p-1 order difference processing the first-order difference operation. \( \nabla^p x_t \) is denoted as the x's second-order difference, that is:

\[ \nabla^p x_t = \nabla^{p-1} x_t - \nabla^{p-1} x_{t-1} \]  

The subtraction operation between two sequence values of the interval k is called k-step difference operation, which is denoted \( \nabla_k \) and represents the subtraction operation between \( x_t \) and \( x_{t+k} \), that is:

\[ \nabla_k = x_t - x_{t+k} \]  

2.5 ARIMA (p, d, q) Sum and Autoregressive Moving Average Model

If \( \{x, t \in T\} \) is stationary after d-order difference, the model should satisfy the following formula, which is in the form of de-centralization:

\[ \varnothing(B) \nabla^d x_t = \theta(B) \epsilon_t \]  

\[ E(\epsilon_t) = 0, Var(\epsilon_t) = \delta^2, E(\epsilon_t \epsilon_s) = 0, s \neq t \]  

\[ Ex_t \epsilon_s = 0, \forall s < t \]  

Above the formula, \( |B| \leq 1 \), \( \varnothing(B) \) and \( \theta(B) \) are relatively prime numbers and \( \varnothing_0 \theta_0 \neq 0 \). d is a differential order, p is an autoregressive order, q is a moving average order, and \( \{\epsilon_t\} \) is a zero-mean white noise sequence.

From the above points, ARIMA model is, in essence, the combination of differential operation and ARMA model. Therefore, if any non-stationary time series can become stationary after proper order difference and the difference sequence can process modeling fitted. At present, the analysis method of ARIMA model is very mature, so the analysis of differential stationary series will be very simple and reliable.

3 Experiments

3.1 Data collection

The price of Chinese herbal medicines is influenced by many factors, such as weather, supply-demand relationship, economic basis, etc. However, there are complicated relationships among these factors, so it is difficult to use factor analysis method. Here, we take the price of the Paris polyphylla S smith over the years as a time series, draw its change rule according to the past data, and establish a prediction model for forecasting the future change and development.
The data in this paper are from Tiandi Pharmaceutical Material Network. The monthly price data of the Paris polyphylla Smith market in Lotus Pool Pharmaceutical Market from July 2013 to June 2019 are collected. The price curve is shown in Table I in yuan.

Table 1 Market Prices of the Paris polyphylla Smith in Southwest China from 2013.7 to 2019.8

| date  | Price | date  | Price | date  | Price |
|-------|-------|-------|-------|-------|-------|
| 2013.7 | 360   | 2015.7 | 750   | 2017.7 | 950   |
| 2013.8 | 430   | 2015.8 | 750   | 2017.8 | 950   |
| 2013.9 | 450   | 2015.9 | 750   | 2017.9 | 950   |
| 2013.10 | 450    | 2015.10 | 750   | 2017.10 | 1200   |
| 2013.11 | 450   | 2015.11 | 750   | 2017.11 | 1200   |
| 2013.12 | 450   | 2015.12 | 750   | 2017.12 | 1150   |
| 2014.1 | 480   | 2016.1 | 750   | 2018.1 | 1100   |
| 2014.2 | 480   | 2016.2 | 750   | 2018.2 | 1100   |
| 2014.3 | 480   | 2016.3 | 750   | 2018.3 | 1100   |
| 2014.4 | 580   | 2016.4 | 750   | 2018.4 | 1100   |
| 2014.5 | 580   | 2016.5 | 780   | 2018.5 | 1150   |
| 2014.6 | 650   | 2016.6 | 780   | 2018.6 | 1150   |
| 2014.7 | 700   | 2016.7 | 800   | 2018.7 | 1150   |
| 2014.8 | 700   | 2016.8 | 800   | 2018.8 | 1120   |
| 2014.9 | 700   | 2016.9 | 900   | 2018.9 | 1120   |
| 2014.10 | 740   | 2016.10 | 900   | 2018.10 | 1100   |
| 2014.11 | 740   | 2016.11 | 900   | 2018.11 | 1100   |
| 2014.12 | 740   | 2016.12 | 900   | 2018.12 | 1050   |
| 2015.1 | 740   | 2017.1 | 900   | 2019.1 | 1050   |
| 2015.2 | 740   | 2017.2 | 900   | 2019.2 | 1050   |
| 2015.3 | 740   | 2017.3 | 900   | 2019.3 | 1050   |
| 2015.4 | 750   | 2017.4 | 900   | 2019.4 | 1050   |
| 2015.5 | 750   | 2017.5 | 900   | 2019.5 | 980    |
| 2015.6 | 750   | 2017.6 | 900   | 2019.6 | 980    |

This paper takes 62 data from July 2013 to August 2018 as an example, uses time series analysis method to analyze and model the data, and uses the model to forecast the price from September 2018 to June 2019, comparing with the price data from September 2018 to June 2019, to see whether it is consistent with the reality and whether it can be used in actual forecast. Finally, the price data for the next six months are predicted.

3.2 Stationary Judgment and Processing of Price Data of the Paris polyphylla Smith

The data of the Paris polyphylla Smith prices from July 2013 to June 2018 are analyzed. Figure 1 is a time series diagram of the Paris polyphylla Smith prices from July 2013 to June 2018.

As can be seen from Figure 2, the price of the Paris polyphylla Smith has shown an upward trend as a whole, followed by a small decline, but on the whole it is still an upward trend. Therefore, it is considered that the sequence is a non-stationary time series, so differential operation should be performed on the sequence to eliminate the linear trend of the sequence and make it finally stable. Therefore, the second-order difference time series diagram of the Paris polyphylla Smith data is shown in Figure 2.

From the change trend of the sequence, the natural logarithm of the original time series and the second-order difference make the distribution of the sequence basically uniform, and fluctuate smoothly above and below zero, basically reaching a stable level. Therefore, the second-order difference data are identified and ranked. After repeated experiments, the final time series model is found to be ARIMA (9, 2, 0). The square of R is 0.954, and the
fitting degree is very good. The AR coefficient is not significantly 0. Therefore, the model is reasonable.

Table 2 Market Prices of the Paris polyphylla Smith from 2018.7 to 2019.6

| predicted value | 1106.173 | 1093.203 | 1083.006 |
|-----------------|----------|----------|----------|
| actual value    | 1120     | 1100     | 1079.737 |
| predicted value | 1084.098 | 1083.019 | 0.01235 |
| actual value    | 1050     | 1050     | 1079.373 |
| relative error  | 0.00618  | 0.01545  | 0.032474 |
| predicted value | 1073.901 | 1062.545 | 1049.94 |
| actual value    | 1050     | 1050     | 1079.737 |
| relative error  | 0.00618  | 0.01545  | 0.032474 |
| predicted value | 1084.098 | 1083.019 | 0.01235 |
| actual value    | 1050     | 1050     | 1079.737 |
| relative error  | 0.00618  | 0.01545  | 0.032474 |

Table 2 shows that the relative error between the actual value and the predicted value in each year is within 7.2%. Therefore, it can be considered that the prediction effect of this model is relatively good. The model can be used to forecast the market price of the Paris polyphylla Smith from July 2019 to December 2019. The results are shown in Table 3.

Table 3 Forecasted market prices for the Paris polyphylla Smith from 2019.7 to 2019.12

| year   | 2019.7 | 2019.8 | 2019.9 |
|--------|--------|--------|--------|
| predicted value | 962.84  | 941    | 923    |
| year   | 2019.1 | 2019.11 | 2019.12 |
| predicted value | 905    | 890    | 866    |

4 Conclusion and discussion

This paper makes time series analysis of the prices of the Paris polyphylla Smith market price in southwest China from July 2013 to June 2018, establishment a model of ARIMA(9, 2, 0). It is found that the fitting result of the model is feasible. Finally, the model is used to forecast the prices of the Paris polyphylla Smith market price from July 2018 to June 2019, the relative error is basically controlled within 7.2%, the market price of the Paris polyphylla Smith in the next few years is predicted accurately, the predicted results are satisfactory. Looking at the forecast results for the next six months, the market price of the Paris polyphylla Smith will continue to show a small downward trend. This is in line with the current situation of economic development in the pharmaceutical market, because in recent years, the global economic development has slowed down, and the economies of various countries have also shown a trend of low-speed development.

Because of the influence of many random factors, we only considered its own factors but not other factors, so the forecast value should only be used as a reference. Practical application should also be combined with other factors affecting the price of the Paris polyphylla Smith, such as the environment, supply-demand relationship and so on. From the perspective of price change, this paper tries to achieve a more accurate prediction of price change. However, the predicted results from the model are only a predicted value, and the change of the price of the Paris polyphylla Smith is a complex and changeable dynamic system. When the country's macro-policy is adjusted, the development environment is changed, or science and technology progress, its price will change accordingly. Therefore, medicinal herb grower should always pay attention to the risk of price changes in the Paris polyphylla Smith market price, and change their planting strategies accordingly, in order to prevent the loss caused by price fluctuations.

Acknowledgement

The authors are very grateful to the journal editorial team and reviewers who have provided valuable suggestions for improving the quality of this paper. This work was supported by Key Laboratory of Agricultural information engineering of Sichuan Province, Research Topic of Government Affairs in Sichuan Province in 2019 (Research on the Development of Intelligent Agriculture in Sichuan Province).

References

1. A. B. Cunningham, J. A. Brinckmann, Y. -F. Bi, S. -J. P  e i,  U.  S c h i p p m a n n,  P.  L u o,  J. Ethnopharmacol,  222, 208-216(2018).
2. Barak, Sasan, and S. S. Sadegh, International Journal of Electrical Power & Energy Systems, 82, 92-104(2016).
3. C. B. Li, K. C. Wang, J CENT SOUTH UNIV T, 14, 713-718(2007).
4. Y. C. Hu, P. Jiang, P. C. Lee, J OPER RES SOC, 1-9(2018).
5. J. Zong, Q. Zhu, IEEE International Conference on Software Engineering & Service Science, (2012).
6. J. Contreras, R. Espinola, F. J. Nogales, MPER, 22, 57-57(2002).
7. A. Darekar, A. A. A. Reddy, SSRN Electronic Journal, 30, 73-76(2017).
8. Z. Gong, F. Gong, J. Peng, Chinese Agricultural Science Bulletin, (2015).