

A New Method of Grain Output Prediction Based on R/S Analysis

Teijun Yang, Na Yang and Chunhua Zhu*

School of Information Science and Engineering, Henan University of Technology, Zhengzhou 450001, China

*Corresponding author

Abstract—The grain production data series present some regularity and nonlinear characteristics in time. To overcome the shortcomings of the linear prediction low precision in traditional grain yield, the application of R/S analysis which based on the nonlinear fractal theory in grain output prediction is studied in this paper. Firstly, the hurst index of grain production series is calculated, thereby its average cycle is analyzed, finally, the prediction of grain production is obtained by the power value of grain production series. Simulation results have shown that the proposed method in this paper can realize the data prediction with nonlinear variation law, and has higher prediction accuracy, compared with the traditional grain output prediction method based on the arima model.

Keywords-grain output, R/S analysis, prediction

I. INTRODUCTION

To predict the grain output accurately is one of significant work in china, which plays an important role in decision-making among the planting plan, the adjustment of consumption structure and macroeconomic regulation and control in agricultural sector. Due to grain output affected by many factors, such as production factors, price factors, policy factors and meteorological factors, the grain output time series present complexly because the combined action of these factors which also influence the long-term trend of grain output and annual fluctuations. At present, there are many prevailing methods in grain output prediction, namely the meteorological yield prediction method, remote sensing method, artificial neural network prediction method, regression prediction method, and gray system forecast method, etc[1-3]. The first three kinds of forecasting methods are the causal prediction for considering the influence factors, but the technical implementation is complex because these prediction methods need a large amount of data; Last two prediction methods are the trend extrapolation prediction for not considering the influence factors of time series, the prediction model is simple but low accuracy. The identification and prediction of under-forecast data with nonlinear change law can be realized by R/S analysis method which based on the theory of nonlinear fractal, and it has been widely applied in many time series prediction fields such as hydrology, meteorology, the stock market because of its simple operation and high precision [4-6], but rarely in grain output prediction. So it has positive significance to explore the application in grain output prediction according to the basic principle of R/S analysis in this paper.

II. THE PRINCIPLE OF R/S ANALYSIS PREDICTION

R/S analysis method was first proposed as a new statistical method by British scientists hurst (H. E. Hurst) studying observation data of the Nile years hydrological, then was improved by Mandelbrot and Wallis in theory, and finally was developed into the fractal theory [7,8]. The principle of time series prediction based on R/S analysis is shown in figure 1.

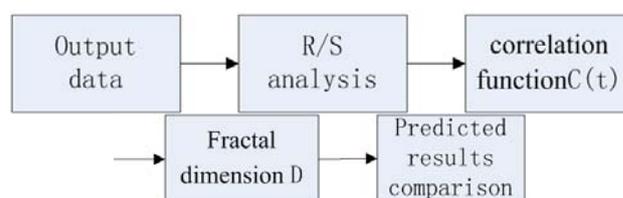


FIGURE I. THE PREDICTION PRINCIPLE DIAGRAM OF TIME SERIES BASED ON R/S ANALYSIS

A. The Fundamental Principle of R/S Analysis

For a given time series at time $t_1, t_2, t_3, \dots, t_n$, the corresponding time series are $x_1, x_2, x_3, \dots, x_n$, and the time span of time series is

$$\tau = t_n - t_1 \quad (1)$$

During τ , the average value of the time series is

$$\bar{x}_n = \frac{1}{n} \sum_{i=1}^n x_i \quad (2)$$

Computing the accumulated deviation at the time of t_j

$$X(t_j, n) = \sum_{i=1}^j (x_i - \bar{x}_n) \quad (3)$$

The difference between maximum and minimum value of $X(t)$ corresponding to the same n is called the range, hence $R(\tau)$,

$$R(t_n - t_1) = R(\tau) = \max X(t, n) - \min X(t, n) \quad t_1 \leq t \leq t_n \quad (4)$$

Hurst, using the standard error:

$$S(\tau) = \sqrt{\frac{1}{\tau} \sum_{i=1}^n (x_i - \bar{x}_n)^2} \quad t_1 \leq t \leq t_n \quad (5)$$

proposed a non-dimensional ratio R/S to rescale R

$$\frac{R(\tau)}{S(\tau)} = (a\tau)^H \quad (6)$$

Where H is called the Hurst exponent. It also can be written as

$$\frac{\max X(t, n) - \min X(t, n)}{\sqrt{\frac{1}{\tau} \sum_{i=1}^n (x_i - \bar{x}_n)^2}} = (a\tau)^H \quad (7)$$

Fitting a model with the method of least squares in double logarithmic coordinate system

$$\ln(R(\tau)/S(\tau)) = H \ln \tau + H \ln c \quad (8)$$

The slope is the power of the Hurst value.

B. Hurst Exponent and the Fractal Dimension

The correlation function $C(t)$ is adopted to express the influence of the past to the future trend more intuitional:

$$C(t) = 2^{2H-1} - 1 \quad (9)$$

The experiments shows when $H=0.5$ and $C=0$, the sequence accords with independent random process, and it has no time correlation; When $0.5 < H < 1$ and $C > 0$, it shows time correlation has positive effect; When $0 < H < 0.5$ and $C < 0$, the time correlation has negative effect.

Besides, R/S analysis can produce fractal dimension D ($D=2-H$) of time series. The fractal dimension measured the uneven of time series. The scope of D is $1 < D < 2$, and if the D value is bigger, it represents more jagged line, more irregular and more complex; On the contrary, if the D value is smaller, it represents less jagged line, more laws and more simple.

C. Determine the Average Cycle of Time Sequence

Introducing the statistic $V(\tau)$ to quantitatively characterize the average cycle system[4],

$$V(\tau) = \frac{R(\tau)/S(\tau)}{\sqrt{\tau}} \quad (11)$$

$V(\tau)$ can not only be used to test the R/S stability, but also be used to determine the sequence whether there is a cycle or not and estimate the cycle length[11-12]. When time series showing continuous ($H > 0.5$), the ratio will increase and the curve of statistic $V(\tau)$ will rise; If the sequence present a random walk ($H = 0.5$) or persistent, statistic $V(\tau)$ will

generally remain unchanged or drab. When $V(\tau)$ change along with the variation trend of $\ln \tau$, namely the curve appears obvious turning point, the future influence of historical state disappeared, the corresponding time span τ is the average cycle of the system.

D. The Prediction Process of R/S Analysis

In order to predict the change trend of grain output x_{n+1} [12], it can be seen from equation (2):

$$x_{n+1} = (n+1)\bar{x}_{n+1} - (x_1 + x_2 + \dots + x_n) \quad (12)$$

Setting $R(n+1)/S(n+1) = [a(n+1)]^H = K$, then combining with the equation(4) and (5), we can get

$$\frac{x_{n+1} - x_1}{\sqrt{\frac{1}{n+1}[(x_1 - \bar{x}_n)^2 + (x_2 - \bar{x}_n)^2 + \dots + (x_n - \bar{x}_n)^2 + (x_{n+1} - \bar{x}_n)^2]}} = K \quad (13)$$

Then the equation (13) can be rewritten by equation (12) as

$$\bar{x}_{n+1} = \frac{-B + \sqrt{B^2 - 4AC}}{2A} \quad (14)$$

$$\begin{aligned} \text{Here,} \quad A &= (n+1)[(n+1)^2 - nK^2] \quad , \\ B &= -2(n+1)[(n+1)(x_1 + u) - K^2u] \quad , \\ C &= (n+1)(x_1 + u)^2 - Q^2(u^2 + v) \quad , \\ u &= x_1 + x_2 + \dots + x_n, \quad v = x_1^2 + x_2^2 + \dots + x_n^2. \end{aligned}$$

Finally, the change trend of time series can be obtained as

$$x_{n+1} = (n+1)\bar{x}_{n+1} - u \quad (15)$$

III. SIMULATION REALIZATION

The overall trend of grain production in China can be studied by fractal R/S analysis method and correlation function $C(t)$ of fractional Brownian motion, thus whether the development of China's food production to meet hurst rule can be verified and it will provide a new way for the prediction of grain output.

A. Data Collection

China's grain output data is from the *China statistical yearbook* [13](1978-2014), the staple food in china is cereals, beans and potato, among them cereals include paddy, wheat and corn.

B. The Simulation Results

The output data from 2010 to 2014 is predicted by using the method described above, then comparing with the results of the ARIMA model and based on this the yield data after 2015 is predicted, while the production data in 1978-2014 which higher than 520 million tons as a dividing line in the calculate

process, and the data in 1978 as zero point. The concrete implementation steps to forecast production data by using R/S analysis method are as following:

(1)According to the principle of R/S analysis method, the corresponding $R(\tau)/S(\tau)$ with different time span τ is calculated, the data points curve about $\ln(R(\tau)/S(\tau)) - \ln \tau$ is drawn by using *MATLAB*, and then carry on the linear fitting, the fitting slope is the H value, the result is shown in figure 1.

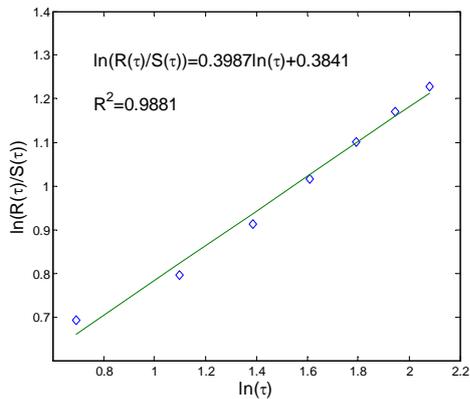


FIGURE II. LINEAR FITTING FIGURE WITH $\ln(R(\tau)/S(\tau)) - \ln \tau$

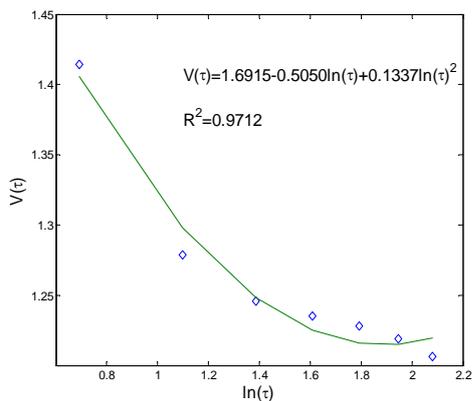


FIGURE III. R/S ANALYSIS OF GRAIN OUTPUT AND V STATISTIC FIGURE

TABLE I. GRAIN OUTPUT HURST INDEX AND CORRELATION FUNCTION $C(t)$, FRACTAL DIMENSION D

Parameter	H	$C(t)$	D
Grain Output	0.3987	-0.1310	1.6013

It can be seen from figure 1 and table 1 that the Hurst exponent of grain output which calculated by R/S analysis method is 0.3987, the correlation function $C(t)$ is -0.1310. So it can be concluded that the overall trend of China's grain output is steady. The growth rate is dropped, which can be verified in the latter results. The fractal dimension D is 1.6013, it represents the grain production sequence fluctuated remarkably,

and also fit with our country's reform and open policy before and after the change of agricultural policy [2].

(2) Calculating statistic of $V(\tau)$ and mapping data points of $\ln(\tau) \sim V(\tau)$ according to the equation (11). Then fitting the curve with quadratic (as shown in figure 2) in this paper so as to determine the average cycle of system. The location of the significant turning point is determined by calculating the inflection point of the curve, and the time point which decided by the abscissa of the turning point is the average cycle system.

As the figure 2 showing that China's grain output V statistic curve presents a falling trend, it also reflects the fact that China's grain production system is complex. The corresponding abscissa of inflection point which calculated by the method in the literature [11] are 0.7, 1.4 and 1.95 respectively, so the average cycle are 2, 4 and 7 years. It means that the correlation of grain output in china change quickly, and the output data has the characteristic of nonlinear dynamic system.

(3) According to the introduction of 2.4, the values of x_{n+1} and u are typed into formula (12) and (14), then the grain output prediction in 2010-2014 are calculated and shown in table 2. The forecast data of grain output based on ARIMA model is also shown in table 2 [14].

TABLE II. THE ERROR COMPARISON ABOUT GRAIN OUTPUT PREDICTION USING R/S ANALYSIS AND ARIMA MODEL (UNIT: 10000 TONS)

Yea	Actual	R/S	error	ARIMA	error
201	54648	54210	0.8%	53571	1.9%
201	57121	56371	1.3%	55179	3.3%
201	58958	59542	0.9%	58159	1.3%
201	60194	60756	0.9%	59929	0.4%
201	60710	60819	0.2%	60874	0.3%

Table 2 shows that the results of grain output prediction overall error by using R/S analysis method is less than using ARIMA model. So it can be concluded that the prediction accuracy of grain output can be improved effectively by using R/S analysis method.

(4) Verifying the accuracy of the results through predicting China's grain output from 2015 to 2019 by using the calculation method which described above. The result is shown in table 3.

TABLE III. USING R/S ANALYSIS METHOD TO FORECAST CHINA'S GRAIN PRODUCTION FROM 2015-2019 (UNIT: 10000 TONS)

number	Year	Cumulative output	R/S analysis output
1	2015	413080	61227
2	2016	474670	61586
3	2017	536570	61901
4	2018	598750	62182
5	2019	661190	62436

From table 3 and the calculation method, it can be calculated that x_{n+1} is 32.9605, and x_{n+1} is 37.6443, the result rounded to 38, which means the year of 2015 in the time series by making the year of 1978 as the calculation zero point. It also indicates that the grain output in 2015 will exceed 520 million tons, which accord with the predicted results.

Based on this, the increase of grain output since 1960 can be calculated with ten years outputs as a stage in order to analysis the China's long-term growth trend of grain production, of which using predictive value as 2015-2019 output number and negative number represent the decline in value. The result is shown in table 4.

TABLE IV. INCREASE OF GRAIN OUTPUT IN CHINA FROM 1960-2019(UNIT:10000T)

Year	Output range	Output difference	Output increase or decrease value
1960-1969	14386-21097	6711	• • •
1970-1979	23996-33211	9215	2504
1980-1989	32055-40754	8699	-516
1990-1999	44624-50838	6214	-2485
2000-2009	46217-53082	6865	651
2010-2019	57637-62436	4799	-2066

It can be seen from table 4 that China's grain production presents a stage of decreasing trend since 1960. The growth rate dropped significantly than before though there are some growth years, it also verified the production trend which has been reflected by H index.

IV. CONCLUSION

In this paper, the complex factors in the analysis process of traditional prediction algorithm is avoided by using R/S analysis method, and the fractal dimension of time sequence of grain output is determined through analyzing trend factors of time series. Meanwhile output trend is estimated accurately after analyzing its long-term trend and the average cycle. The prediction accuracy of this method is obviously higher than that of ARIMA model to predict the results, then infer the grain output estimate value in the next few years by the use of this method, at the same time achieve qualitative and quantitative forecast of grain production. But this method has the requirement of data quantity, therefore the next step of work is to explore its composite applications with other prediction methods.

ACKNOWLEDGEMENT

This research was financially supported by the National food industry commonweal special scientific research projects.

REFERENCES

[1] Lili Peng, Chongyu Yang, Fang Long. Food policy factors influencing yield fluctuation in China, J. Journal of Hunan Agricultural University (Social Sciences). 2010,11(6):16-21.

[2] Xigang Zhu. China's grain supply and demand balance analysis, J. Agricultural economy (monthly). 2004 (12):12-19.

[3] Yongbing Zheng. Single fractal test and R/S prediction of grain output time series, J. Journal of Liaoning Technical University(Social Science Edition). 2008,10(4):357-359.

[4] Kun Huang, Ning Luo, Xiaohua Zhu,etc. Prediction of the Variety of Nature Gas Consumption With R/S Analysis Method, J. Journal of Southwest Petroleum University. 2007,29(1):73-76.

[5] Xiaoli Wang. R/S analysis method of trend and aberrance point on hydrological time series, J. Engineering Journal of Wuhan University. 2002,35(2):10-12.

[6] Ying Fan, Yiming Wei, Xi Zhang,etc. R/S Analysis Based Study on Fractals in China's Stock Markets, J. Systems Engineering. 2004,22(11):46-51.

[7] Zhenghua Chen, Zihua Chen, Xi Zhang,etc. Prediction of the Variety of Nature Gas Consumption With R/S Analysis Method, J. Mining Safety & Environment Protection. 2010,37(1):36-44.

[8] Opong K K.The behavior of some UK equity indices: An application of Hurst and BDS tests, J.Journal of Empirical Finance,1999(6):267-282.

[9] Yong Huang, Zhifang Zhou, Jinguo Wang,etc. Application of R/S Method to Dynamic Groundwater Analysis, J. Journal of Hehai University. 2002,30(1):83-87.

[10] Yongzheng Wu. Standard process and its application, J. Theory Explore. 2004,(8):23-24.

[11] Qian Liu,JiuZhen Liang. Study of average loop cycle for stock market based on R/S method, J. Computer Engineering and Design. 2009,30(21): 4942-4944.

[12] Yan Jin,Yongbing Zheng,Wenzhang Chen. The fractal method for prediction of grain yield, J. Liaoning Agricultural Sciences. 2005(1):45-46.

[13] National Bureau of Statistics of China. China Statistical Yearbook2014.Beijing: China Statistics Press,2015.

[14] Teijun Yang,Na Yang,Chunhua Zhu.Preformance Investigation of Grain output Prediction Based on ARIMA Model, J. Journal of Henan University of Technology(Natural Science Edition). 2015, 36(5):24-27.