Entropy **2014**, *16*, 6415-6433; doi:10.3390/e16126415

OPEN ACCESS

entropy

ISSN 1099-4300

www.mdpi.com/journal/entropy

Article

Detecting Chaos from Agricultural Product Price Time Series

Xin Su [1], Yi Wang [1,*], Shengsen Duan [2] and Junhai Ma [3]

[1] Shandong University of Finance and Economics, Jinan 250014, China; E-Mail: lgj_sx@126.com

[2] Dongfang College, Shandong University of Finance and Economics, Tai'an 271000, China; E-Mail: lzcgh2007@aliyun.com

[3] Management School of Tianjin University, Tianjin 300072, China; E-Mail: lzqsly@126.com

* Author to whom correspondence should be addressed; E-Mail: cghnju@aliyun.com; Tel./Fax: +86-0531-2789-2371.

External Editor: J. A. Tenreiro Machado

Received: 13 July 2014; in revised form: 28 September 2014 / Accepted: 20 November 2014 / Published: 5 December 2014

Abstract: Analysis of the characteristics of agricultural product price volatility and trend forecasting are necessary to formulate and implement agricultural price control policies. Taking wholesale cabbage prices as an example, a multiple test methodology has been adopted to identify the nonlinearity, fractality, and chaos of the data. The approaches used include the R/S analysis, the BDS test, the power spectra, the recurrence plot, the largest Lyapunov exponent, the Kolmogorov entropy, and the correlation dimension. The results show that there is chaos in agricultural wholesale price data, which provides a good theoretical basis for selecting reasonable forecasting models as prediction techniques based on chaos theory can be applied to forecasting agricultural prices.

Keywords: agricultural product wholesale price; time series; chaos; multiple test methodology; selection of forecasting model

PACS Codes: 05.45.Tp time series analysis; 05.45.Gg control of chaos, applications of chaos; 05.45.Pq numerical simulations of chaotic systems

---

1. Introduction

Recently, the prices of China's agricultural products have been growing rapidly with frequent fluctuations. Although cyclical fluctuations in the prices of agricultural products are an inevitable phenomenon, violent price fluctuations will have a bad effect on farmers' production decision-making and people's consumption, and may even cause uneven development of the national economy. To explain the causes of agricultural price volatility, some researchers have discussed the impact of the development of bio-energy by using of the co-integration and error correction model (ECM). Raneses *etc.* [1] analyzed some possible impacts of the potential bio-energy market and its economic effects. Tokgoz [2] showed by running different scenarios that with the increase of the EU crude oil price, the impact of energy prices on the EU agricultural sector was increasing with the emergence of the bio-fuels sector. Benavides [3], using time series models, showed that exchange rates and inventories had a great impact on the fluctuations of corn and wheat prices. Mitra and Boussard [4], applying the nonlinear Cobweb models, thought that inventories deeply affected the food price volatility. Will agricultural prices lead to inflation? This problem has long been a controversial topic in academic circles. Garner [5] thought that commodity prices are a leading indicator of inflation, while Pindyck and Rotemberg [6] thought that macroeconomic or monetary factors will lead to changes in commodity prices. Cheng *etc.* [7] used an econometrics model to perform an analysis and showed that the rise in pork prices obviously spurred a rise in the CPI. The fluctuation of prices of different kinds of agricultural products has the characteristics of conductivity and synchronization. In [8], the author divided the fluctuations of China's agricultural products prices from 1978 to 2006 into five cycles using the HP Filter method and the ECM. He proposed that increasing agricultural product prices would not cause inflation; however, the inflation contributed to the rising of agricultural products prices.

In brief, according to the existing literature, research on price fluctuations of agricultural products broadly includes two aspects. One is the causes of the agricultural product price fluctuation, the other is the effects. Essentially, the research belongs to the analysis of the phenomenon. Namely, causes and effects are analyzed and related suggestions are provided. In this paper, we try to provide effective tools to explain and predict the price fluctuations, through a deep econometric analysis of the characteristics of price time series.

Research on agricultural price data and prediction has a long history. However, some typical analytical methods, such as the linear analysis methods, have encountered challenges when used to predict agricultural prices, so researchers have developed various methods to predict agricultural prices besides the traditional regression analysis method. Recently, with the rapid development of intelligent analysis and information technology, some new technologies including neural networks, genetic algorithms, artificial intelligence and expert systems, and other new intelligent prediction methods have been gradually introduced to forecast prices. Therefore, the selection of an optimal analysis or prediction model according to the different data characteristics is a very important and practical problem. However, there is little literature on the analysis of the dynamics of agricultural market price time series. If the price system is chaotic, then not only the chaos and nonlinear prediction methods should be used as the main prediction approach (for example, see [9–11]), but the price fluctuation can also be understood and interpreted by using chaos theory. Different from [12–14], where the authors used chaotic cobweb models to discuss price fluctuations, in this paper, we aim to find evidence

whether chaos exists in price data, give the chaos criterion, and if so, prove that agricultural product prices can be predicted by chaotic methods.

Mandelbert [15] first described the "high peak and fat tails" in cotton prices. For the nonlinear, fractal, and chaotic features in agricultural markets, we refer to the articles [16–22]. Corazza *et al.* [16] discussed the fractal structure in agricultural futures markets. Barkoulas *et al.* [17] researched the long memory effect in futures prices. Wei and Raymond [18] applied R/S analysis, modified R/S analysis, and the AFIMA model to discuss the long memory in six agricultural futures markets. Yang and Brorsen [19,20] researched the nonlinearity of futures markets. Jacobs and Onochie [21] discussed the relationship between return variability and trading volume in international futures markets. Bhar [22] used an EGARCH model to research return and volatility dynamics in the spot and futures markets in Australia. Although they researched agricultural markets, most of them used futures prices. Few articles discuss the nonlinear, fractal, and chaotic features in agricultural product wholesale markets prices. Factually, besides the futures prices, many factors (mentioned in the Introduction) can affect the wholesale market prices. Furthermore, not all agricultural products can be used as agricultural futures varieties, for example, fresh vegetables. Sometimes, the wholesale markets prices seem to be more indeterminate, and sometimes, the wholesale prices seem to be more stable. In any case, the wholesale prices are different from the futures prices. Since the wholesale market prices can affect the daily lives of consumers, we are interested in the wholesale market prices, but we must point out that generally, the sample capacity of the wholesale prices is small. If we can collect daily data for ten years, the total number of data points is only $365 \times 10 = 3650$, making detecting chaos in the wholesale prices is a challenge. However, chaos detection is very necessary for chaos prediction. If the time series is chaotic, then the chaotic prediction methods can be applied to forecast the future prices, so we can use chaotic detection methods to choose the appropriate forecasting method.

It is very difficult to detect chaos from an agricultural market price time series. On one hand, the main obstacle in empirical analysis arises from the limited data sources. On the other hand, since there are many artificial factors that emerge frequently during the data acquisition process, the level of noise in agricultural market prices data is high, so we must select the appropriate method very cautiously. Most of the techniques used in nonlinear dynamics come from disciplines like mathematics, physics, and engineering, so data quantity and quality are crucial when applying these methods, and not all the techniques are directly applicable to economic time series.

Since a single and reliable test for chaos is not available at present, in this study a multiple test methodology (see [23]) has been adopted to safeguard the reliability of our empirical findings. We start our analysis with the nonlinear and fractal judgment, and then we detect chaos in the time series by employing a variety of different available approaches in order to avoid misleading results. Then, we apply the above ideas to analyze a cabbage price time series, which is derived from the China Price Information Network using the weekly data of Shandong Province from January 2003 to December 2012 (number of observations = 525) as an example.

In this paper, chaos detection is used to analyze agricultural market prices. We select a series of applicable methods to identify chaos in agricultural price time series. Moreover, the normalization method is used to detrend the original data. In our empirical analysis, the log linear detrending approach is compared with the normalization approach. The results show that the latter is more appropriate for detecting chaos in agricultural prices. In addition, we improve the algorithm for determining the optimal

linear scaling region due to the use of a small data set. The weighted averages algorithm is adopted to obtain the mean period.

The paper is organized as follows: in Section 2, we introduce the normalization method to pretreat the original data. In Section 3, we present the methods of testing nonlinearity and fractality, including the BDS method and the R/S analysis. In Section 4, we introduce a series of tests to examine the chaos in agricultural price time series, including the power spectra, the recurrence plot, the largest Lyapunov exponent, the Kolmogorov entropy, and the correlation dimension. Moreover, we improve some algorithms. Applying the above methods, in Section 5, we analyze the cabbage price time series. The results show there is chaos in the series. In Section 6, we present a discussion of the results and some concluding remarks.

2. Data Preprocessing

The data preprocessing in this paper aims at eliminating the tendency of the original data. In chaos and fractal analysis of economic time series, the log linear detrending approach (see [24]) is reasonable and common. In this paper, we present a normalization approach which has not been widely used. There are many different ways to normalize a time series. We adopt the following one:

Assume:

$$X=\{X_t, t=1,2,\cdots,N\} \tag{1}$$

is an original time series, then the transformation $P_t=\dfrac{X_t-\min X}{\max X-\min X}$, $t=1,2,\cdots,N$ changes X to the new time series $\{P_t, t = 1, 2, \ldots, N\}$ in the interval [0, 1]. In addition, we suppose that $\{Q_t, t = 1, 2, \ldots, N\}$ is the resulting log linear detrended time series, where $Q_t = \ln X_t - (a + bt)$; a and b are the intersection and the constant growth rate, respectively.

3. Tests for Nonlinearity and Fractality

3.1. The BDS Test

The BDS test is a powerful test for independence and randomness and is based on the correlation integral concept (see [25]). Given a time series (1), we can form m-dimensional vectors:

$$X_t^m=\{X_t, X_{t+\tau},\cdots,X_{t+(m-1)\tau}\}, t=1,2,\cdots,T \tag{2}$$

where m is the embedding dimension and τ is the delay time, and calculate the correlation integral defined as:

$$C(m,N,r,\tau)=\frac{2}{T(T-1)}\sum_{t<s}H(r-\left\|X_t^m-X_s^m\right\|_\infty), \tag{3}$$

$$H=\begin{cases}0, z\le 0\\ 1, z>0\end{cases}, \tag{4}$$

where $T=N-(m-1)\tau$, $\|\cdot\|_\infty$ is the ∞-norm, and $r > 0$ is the tolerance distance. The ∞-norm means that $\|X\|_\infty=\max\limits_{1\le i\le T}|X_i|$ for any $X = \{X_1, X_2, \ldots, X_T\}$. If X is a random series of independence and identically

distributed (i. i. d.) observations, then as $m \geq 2$ and $N \to \infty$, the BDS statistic converges to the standard normal distribution.

3.2. The R/S Analysis

The R/S test or Rescale Range analysis was established by Hurst and improved later (see [25]). According to the R/S test, long-term dependence for a time series (1) is detected through the Hurst exponent which can be estimated from the relationship:

$$\left(R/S\right)_n = cn^H, \tag{5}$$

where c is a constant, n is the time index for periods of different length, H is the Hurst exponent, and $(R/S)_n$ is the rescaled range statistic. The relation between the Hurst exponent and the fractal dimension is simply $D = 2 - H$. The autocorrelation function is $C = 2^{2H-1} - 1$. Values of $H = 0.5$, $D = 1.5$ and $C = 0$ indicate a true random walk (a Brownian time series). If $0 \leq H < 0.5$, $D > 1.5$, and $-0.5 \leq C < 0$, the values correspond to a profile like curve showing persistent behavior, while for $0.5 < H < 1$, $D < 1.5$, $0 \leq C < 1$ we have an antipersistent behavior. In addition, we can test the significance on the index $t = \dfrac{H - E(H)}{\sqrt{1/N}}$.

4. Detecting Chaos

4.1. Power Spectrum

Chaotic signals are wideband signals, so they can be easily distinguished from periodic signals by looking at their frequency spectra. If the behavior is chaotic, then the power spectrum of the system is expressed in terms of oscillations with a wide continuum of frequencies.

4.2. The Phase Space Reconstruction

It is well known that an attractor, which is topologically equivalent to a scalar time series, can be reconstructed from a dynamic system of m variables by using the time delay coordinates. It is crucial and difficult to select an appropriate time delay and an embedding dimension for reconstruction. There are many methods to determine the reconstructed parameters. For the determination of optimal time delay of nonlinear systems, the average mutual information [26] is better than the autocorrelation function, as the former contains the nonlinear characteristic in a time series. Cao's method [27] improved the false nearest neighbors' method for estimating the embedding dimension. In this paper, we use Cao's method to estimate the embedding dimension and the mutual information method to determine optimal time delay.

4.2.1. The Mutual Information Method

Assume $s_1, s_2, \ldots, s_n$ and $q_1, q_2, \ldots, q_m$ are two nonlinear time series. The entropy of $S = \{s_1, s_2, \ldots, s_n\}$ and $Q = \{q_1, q_2, \ldots, q_m\}$ are as follows:

$$H(S) = -\sum_{i=1}^{n} P_s(s_i)\log_2(P_s(s_i)),$$
$$H(Q) = -\sum_{i=1}^{m} P_q(q_i)\log_2(P_q(q_i)), \tag{6}$$

where $P_s(s_i)$ and $P_q(q_i)$ are the probability distributions of s_i and q_i respectively. For given S, the mutual entropy between S and Q is:

$$I(Q,S) = H(Q) - H(Q \mid S) = \sum_i \sum_j P_{s,q}(s_i, q_j)\log_2\left[\frac{P_{s,q}(s_i, q_j)}{P_s(s_i)P_q(q_j)}\right], \tag{7}$$

where $P_{s,q}(s_i, q_j)$ is the joint probability distribution of s_i and q_i. For the time series (1), we suppose that $S = \{X_t\}$ and $Q = \{X_{t+\tau}\}$, then the mutual entropy between S and Q is:

$$I(\tau) = \sum_{X_t, X_{t+\tau}} P(X_t, X_{t+\tau})\log_2\left[\frac{P(X_t, X_{t+\tau})}{P(X_t)P(X_{t+\tau})}\right]. \tag{8}$$

The first local minimum of $I(\tau)$ can be selected as the optimal time delay.

4.2.2. Cao's Method

Assume vectors of the reconstructed phase space are Equation (2). The value $a(i,m) = \frac{\left\|X_j^m - X_i^m\right\|_\infty^{m+1}}{\left\|X_j^m - X_i^m\right\|_\infty^m}$ can be calculated, where $i = 1, 2, \cdots, N - m\tau$, $\|\cdot\|_\infty$ is the ∞-norm, j is a positive integer on [1, $N-m\tau$]; X_j^m is the nearest neighbor of *Xmi* in the m-dimensional space. The minimum embedding dimension can be found by the plots of $E_1(m)$ and $E_2(m)$, where $E_1(m) = \frac{E(m+1)}{E(m)}$, $E_2(m) = \frac{E^*(m+1)}{E^*(m)}$, $E(m) = \frac{1}{N-m\tau}\sum_{i=1}^{N-m\tau} a(i,m)$ and $E^*(m) = \frac{1}{N-m\tau}\sum_{i=1}^{N-m\tau} |X_{i+m\tau} - X_{j+m\tau}|$.

4.3. Recurrence Plot and Recurrence Quantification Analysis

We detect the determinism by Recurrence Plot and Recurrence quantification analysis. Recurrence Plot (RP) is a graphical tool that evaluates the temporal and phase space distance [28]. It is a method designed to locate hidden recurring patterns, non-stationarity and structural changes. RPs make it instantly apparent whether a system is periodic or chaotic and are particularly suitable to investigate the economic time series that are characterized by noise and short data sets. Assume Equation (1) is the original time series and is reconstructed in an m-dimensional phase space with vectors Equation (2). A recurrence plot is represented by the binary function $R_{i,j} = H(r_0 - \left\|X_i^m - X_j^m\right\|_\infty)$, where r_0 is the threshold radius and H is defined as Equation (4).When $R_{i,j} = 1$ a black dot is plotted on an otherwise white graph. A plot of random time series shows recurrent points distributed in homogenous random patterns. If the analyzed time series is deterministic, then the recurrence plot shows short line segments parallel to the main diagonal; on the other hand, if the series is white noise, then the recurrence plot does not show any structure. Chaotic behavior causes very short diagonals, whereas periodic behavior causes longer diagonals.

However, researchers can only obtain some rough information from the RPs. The recurrence quantification analysis (RQA), proposed by Webber and Zbilut (for example, see [29,30]), can measure the complexity quantitatively. To measure determinism, we introduced the measure variable: determinism(%DET). The ratio of recurrence points on the diagonal structures to all recurrence points is called %DET, defined as $\%DET = \frac{\sum_{l=l_{\min}}^{T} lP(l)}{\sum_{i,j}^{T} R_{i,j}}$, where T is the number of the vectors; $P(l)$ is the frequency distribution of the lengths of the diagonal structures in the RP; l_{min} is the threshold, which excludes the diagonal lines formed by the tangential motion of a phase space trajectory. In this paper we fix $l_{min} = 2$. %DET is a determinism (or predictability) measure of a system.

4.4. The Correlation Dimension and Kolmogorov Entropy

Correlation dimension is a spatial correlation measure of the points on an attractor and can be interpreted as a lower bound to the degrees of freedom of a dynamical system. The Kolmogorov entropy is an important characteristic to distinguish a chaos system and measure the chaotic degree of a chaotic system. The Kolmogorov entropy is defined as the sum of all the positive Lyapunov exponents. If a system is chaotic, its correlation is not an integer and its K-entropy is positive and finite.

The G-P algorithm [31] is often used to estimate the correlation dimension and the K-entropy from a time signal. The outline of calculating the correlation dimension is as follows. First, one should reconstruct the phase space with vectors (Equation (2)). Second, correlation integrals (Equation(3)) are calculated for different r. Finally, the correlation dimension is defined as:

$$D_2(m) = \lim_{r \to 0} \frac{\ln C(m, N, r, \tau)}{\ln r} \tag{9}$$

which can be estimated through $\ln r \sim \ln C(m, N, r, \tau)$ curve fitting by the least square method. When the embedding dimension m is large enough, $D_2(m)$ tends to a non-integer positive number for a chaotic series, while $D_2(m)$ tends to infinity for a stochastic system. The estimation of the Kolmogorov entropy [32] is the Renyi's quadratic entropy defined as:

$$K_2 = \lim_{r \to 0} \lim_{m \to \infty} \frac{1}{\tau} \left(\frac{\ln C(m, N, r, \tau)}{C(m+1, N, r, \tau)} \right) . \tag{10}$$

However, since the G-P algorithm was too subjective, many researchers improved it. In [33], the authors proposed an optimal algorithm for computing simultaneously the correlation dimension and the K_2 entropy from time series. Provided enough data, this algorithm is concise, reliable, and has been widely used. But for small data sets, the results will change a lot when different r are given. Thus, an algorithm, in [25], was proposed to determine the r value by observing the image of the derivative ($D_2(m)$), where:

$$d(D_2(m)) = \lim_{r \to 0} \frac{D_2(m)}{\ln r} . \tag{11}$$

However, when data sets are small, it is difficult to determine r since the computing results of derivatives are very rough. Therefore, based on the above methods, a new algorithm is proposed in this paper for computing simultaneously the correlation dimension and the K_2-entropy from time series with small samples and high noise levels. Our algorithm includes the following steps:

Step 1. Calculate the correlation dimension.

After reconstructing the phase space with vectors (2), one should divide the interval $[\ln r_{\min}, \ln r_{\max}]$ into K equal subinterval assuming $r \in [r_{\min}, r_{\max}]$, where K is selected according to actual situation. Then the sequence $\{\ln r_0, \ln r_1, \ldots, \ln r_K\}$ is obtained, where $\ln r_i = \ln r_{\min} + i\Delta\ln r$, $i = 0, 1, \ldots, K$. Suppose r corresponding to $\ln r_{\min} + i\Delta\ln r$, $i = 0, 1, \ldots, K$ is r_j for $i = 0, 1, \ldots, K$. For a fixed m and different r_j, the corresponding correlation integrals $\ln C(m, N, r_j, \tau)$defined as Equation (3) can be calculated. Then, one can draw the curve of $\ln r \sim \ln C(m, N, r, \tau)$. One also can get a series of curves corresponding to different embedding dimensions. Keep increasing the value of m until the slope in the plot of $\ln r \sim \ln C(m, N, r, \tau)$ reaches a plateau. Here, we omit m, N, and τ of $C(m, N, r, \tau)$for convenience.

Step 2. Determine the linear scaling region (or the non-scale range).

Assume the derivative of a $\ln r \sim \ln C(r)$ curve is:

$$\ln' C(r_j) = \frac{d[\ln C(r_j)]}{d[\ln r_j]} = \frac{\ln C(r_{j+1}) - \ln C(r_j)}{\ln r_{j+1} - \ln r_j}, j = 0,1,\cdots,K-1, \tag{12}$$

and the second-order derivative is:

$$\ln'' C(r_j) = \frac{d[\ln' C(r_j)]}{d[\ln r_j]} = \frac{\ln' C(r_{j+1}) - \ln' C(r_j)}{\ln r_{j+1} - \ln r_j} \tag{13}$$

$$= \frac{\dfrac{\ln C(r_{j+2}) - \ln C(r_{j+1})}{\ln r_{j+2} - \ln r_{j+1}} - \dfrac{\ln C(r_{j+1}) - \ln C(r_j)}{\ln r_{j+1} - \ln r_j}}{\ln r_{j+1} - \ln r_j}, j = 0,1,\cdots,K-2. \tag{14}$$

Since $\ln r$ are chosen equidistantly, $\ln r_{j+2} - \ln r_{j+1} = \ln r_{j+1} - \ln r_j = \Delta\ln r$, we have:

$$\ln'' C(r_j) = \frac{\ln C(r_{j+2}) + \ln C(r_{j+1}) - 2\ln C(r_j)}{(\Delta r)^2} \tag{15}$$

Draw the plots of $\ln r \sim \ln C(r)$, $\ln r \sim \ln C'(r)$ and $\ln r \sim \ln C''(r)$. The optimal scaling region can be selected in the region where the plot of $\ln r \sim \ln C(r)$ appears oblique line, the plot of $\ln r \sim \ln C'(r)$ appears horizontal line, and simultaneously the values of $\ln C''(r)$ are zero.

Step 3. Compute the correlation dimension and the K_2-entropy.

In the linear scaling region determined according to Step 2, assume the embedding dimension is m, and assume x_{mj} and y_{mj} are the abscissas and ordinates of the points on the $\ln r \sim \ln C(r)$ curves, respectively. Then, the correlation dimension D_2 and the K_2-entropy K_2 are $D_2 = \dfrac{\sum_m \sum_j (x_{mj} - \overline{x_m})(y_{mj} - \overline{y_m})}{\sum_m \sum_j (x_{mj} - \overline{x_m})^2}$ and $K_2(m) = \overline{y_m} - D_2 \overline{x_m}$, respectively, where $\overline{x_m}$ and $\overline{y_m}$ is the mean of all the x_{mj} and y_{mj} for fixed m. Draw the plot of $m \sim K_2(m)$. When $K_2(m)$ tends to a stable value, this value can be taken as the estimate of K_2.

4.5. The Largest Lyapunov Exponent

Lyapunov exponents determine the rate of divergence or convergence of initially nearby trajectories in phase space. In general, an *m*-dimensional system has *m* different Lyapunov exponents. For a chaotic system, at least one Lyapunov exponent is positive. Thus, for our purposes it suffices to constrain the analysis solely to the largest Lyapunov exponent. Many algorithms were developed to calculate the largest Lyapunov exponent. In this paper, we select the algorithm of Rosenstein [34], which is fast, easy to implement, and robust to changes in the following quantities: embedding dimension, size of data set, reconstruction delay, and noise level.

4.5.1. Determine the Mean Period.

When we apply the algorithm of Rosenstein, the mean period is an important parameter, but determination of the mean period has been controversial. In this paper, the weighted averages algorithm [35] is adopted to obtain the mean period, which is more reliable. The first step of this algorithm consists of changing the time series by using of the fast Fourier transform (FFT). Next, we assume the amplitudes are $A_1, A_2, \ldots, A_n$ corresponding to the frequencies $f_1, f_2, \ldots, f_n$, respectively. Then, the mean period is:

$$p = \frac{\sum_{i=1}^{n} A_i}{\sum_{i=1}^{n} f_i A_i} \tag{16}$$

4.5.2. Calculate the Largest Lyapunov Exponent.

First, we reconstruct the phase space with vectors Equation (2). Next, the algorithm of Rosenstein locates the nearest neighbor $X_{\hat{j}}^m$ of each point X_j^m on the trajectory. We denote their distance by $d_j(0) = \min \left\| X_j^m - X_{\hat{j}}^m \right\|_2$, $|\hat{j}-j|>p$, where $\|\cdot\|_2$ is the Euclidean norm or 2-norm and p is the mean period of the time series. The Euclidean norm means that $\|X\|_2 = \sqrt{\sum_{i=1}^{T} |X_i|^2}$ for any vector $X = \{X_1, X_2, \ldots, X_T\}$. After i time intervals, the two points separate exponentially. Their distant is changed to $d_j(i) = \min \left\| X_{j+i}^m - X_{\hat{j}+i}^m \right\|_2$. Then, for each i we calculate the average value $S(i)$ of all $\ln d_j(i)$, where $S(i) = \frac{1}{q\Delta t} \sum_{j=1}^{q} \ln d_j(i)$, q is the number of all non-zero $d_j(i)$, and Δt is the time interval. The final step is to perform i~$S(i)$ curve fitting with the least square method. The estimate of the largest Lyapunov exponent is:

$$\hat{\lambda} = \frac{(n+1)\sum_{i=0}^{n} S(i)i - \sum_{i=0}^{n} S(i) \sum_{i=0}^{n} i}{(n+1)\sum_{i=0}^{n} i^2 - (\sum_{i=0}^{n} i)^2}, \tag{17}$$

where n is the number of i.

5. Empirical Analyses

In the China Price Information Network, we obtained the time series of the arithmetic average wholesale cabbage prices of Shandong Province, which yielded 525 weekly data points from January 2003 to December 2014.

5.1. Statistical Description and Stationary Tests

We use X, X_{Nl} and X_{LLD} to represent the original price sequence, the normalized series, and the resulting log linear detrended series, respectively. The basic statistical properties and the results for the ADF and PP tests of X_{Nl} and X_{LLD} are listed in Table 1. We can find out that both of the results for the J-B tests reject the null hypothesis of normal distribution at a 0.001 level of significance; the distribution of XNl has a high peak and a fat tail, but the one of X_{LLD} has not.

Table 1. Statistical description and stationary tests. The ADF and PP tests are in level;trends and Intercepts are not included in the test equations; the lag lengths are selected using the SIC; the results of ADP and PP tests indicate the stationary of the series at a 0.01 level of significance.

	Average	Std.Dev.	Kurtosis	Skewness	Jarque-Bera/Prob.	ADF/Prob.	PP/Prob.
X_{Nl}	0.2777	0.1624	5.2782	1.2559	251.5453/0.0000	−3.3862/0.0007	−2.6754/0.0074
X_{LLD}	7.62×10^{-12}	0.4614	2.9657	−0.4023	14.18629/0.0008	−5.7027/0.0000	−7.452/0.0000

5.2. Tests for Nonlinearity

5.2.1. The BDS Test

Our data were filtered first, by best-fit time series linear models in order to remove any linear dependence. For X_{Nl}, the AIC and SC indicates AR(3) specification as the most suitable. For X_{LLD}, AR(4) is the most suitable model. Then we get the AR(3) residuals and the AR(4) residuals denoting by $X_{Nl\text{-}resid}$ and $X_{LLD\text{-}resid}$, respectively. Tables 2 and 3 show the BDS test statistics and the probabilities for the AR(3) and AR(4) residuals.

Table 2. $X_{Nl\text{-}resid}$BDS test statistics and probabilities. σ_1 is the standard deviation of $X_{Nl\text{-}resid}$.

	$0.5\sigma_1$	$1.0\sigma_1$	$1.5\sigma_1$	$2.0\sigma_1$
$m = 2$	7.8591/0.0000	5.4748/0.0000	4.1701/0.0000	5.968/0.0000
$m = 3$	7.6953/0.0000	2.8132/0.0049	4.9304/0.0000	7.1725/0.0000
$m = 4$	−6.132/0.0000	2.25/0.0045	10.9701/0.0000	10.8574/0.0000
$m = 5$	−3.762/0.0002	−4.7486/0.0000	11.0967/0.0000	13.5797/0.0000

At the 0.05 level of significance, the results of BDS tests reject the null hypothesis of i.i.d. It is well known that the BDS test has good power to detect four types of non- i.i.d. behavior: linear dependence, non-stationarity, chaos and nonlinear stochastic processes. Therefore, combining the stationary test results we have that the time series X_{Nl} and X_{LLD} are nonlinear, but we are not sure that they are chaotic sequences.

Table 3. $X_{LLD\text{-}resid}$ BDS test statistics and probabilities. σ_2 is the standard deviation of $X_{LLD\text{-}resid}$.

	$0.5\sigma_2$	$1.0\sigma_2$	$1.5\sigma_2$	$2.0\sigma_2$
$m = 2$	4.7493/0.0000	3.3066/0.0009	3.0615/0.0000	3.4306/0.0006
$m = 3$	2.9711/0.003	2.9031/0.0037	3.808/0.0000	3.2573/0.0011
$m = 4$	3.7979/0.0001	3.4747/0.0005	4.3133/0.0000	2.9763/0.0029
$m = 5$	18.3401/0.0000	3.4747/0.0005	6.1617/0.0000	2.7693/0.0056

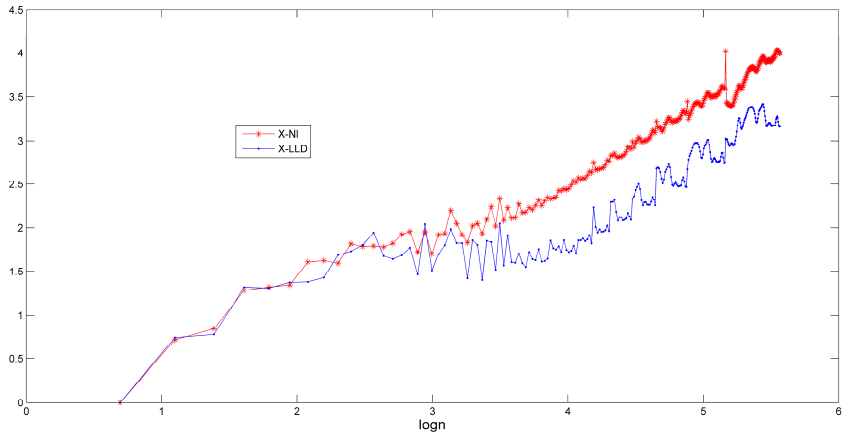
5.2.2. The R/S Analysis

Table 4 contains the Hurst exponents H, the fractal dimensions D, the autocorrelation function values C, and the t values of X_{NI} and X_{LLD}. We can see that the two time series are persistent and fractal with long-term memories. Figure 1 gives the $\log n \sim \log(R/S)_n$ curves of X_{NI} and X_{LLD}, respectively.

Table 4. The R/S analysis. The level of significance is 0.01. The critical value of t is 2.58.

	H	***C***	***D***	***t***
X_{NI}	0.8399	0.6019	1.1601	6.024
X_{LLD}	0.7199	0.3565	1.2801	3.275

Figure 1. The $\log n \sim \log(R/S)_n$ curves of X_{NI} and X_{LLD}. The red curve is for X_{NI} and the blue curve is for X_{LLD}.

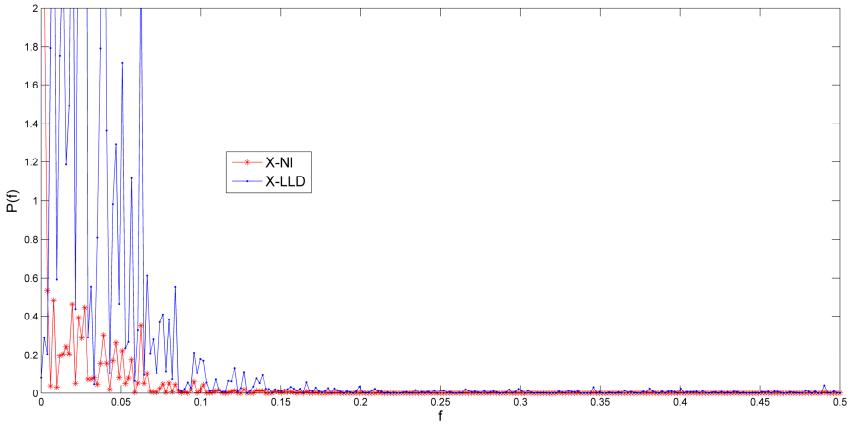


It is worth pointing out that we detrended the original time series by the approach of log first differentiating and the approach of first differentiating. Then we calculated the Hurst exponents of the two resulting detrended series. The two exponents are 0.3242 and 0.3212, which indicate an anti-persistent time series. The reason is that first differentiating detrending is a high-pass filter or noise-amplifier (see [10]). Therefore, we believe that different detrending approaches can affect the results of Hurst exponents. For the agricultural products price time series, the approach of log linear detrending and the normalization approach are more reliable, and the latter is better.

5.3. Detecting Chaos

(1) We plot the power spectrum of X_{NI} and X_{LLD} (see Figure 2), which are decaying, continuous, and not flat.

Figure 2. Power spectrum for X_{NI} and X_{LLD}. The red curve is for X_{NI}. The blue curve is for X_{LLD}. The *x* axis represents the values of frequency, and the *y* axis represents the values of spectrum.



(2) The phase space is reconstructed using delays and dimensions as selected respectively (see Table 5) from mutual information function and false nearest neighbors. From Figures 3 and 4, it can be seen that E_2 increase with m increasing and tend to 1 eventually. Both of the time series exhibit the chaotic behavior instead of the stochastic one.

Figure 3. The m~E_1 and E_2 plot of X_{NI}.

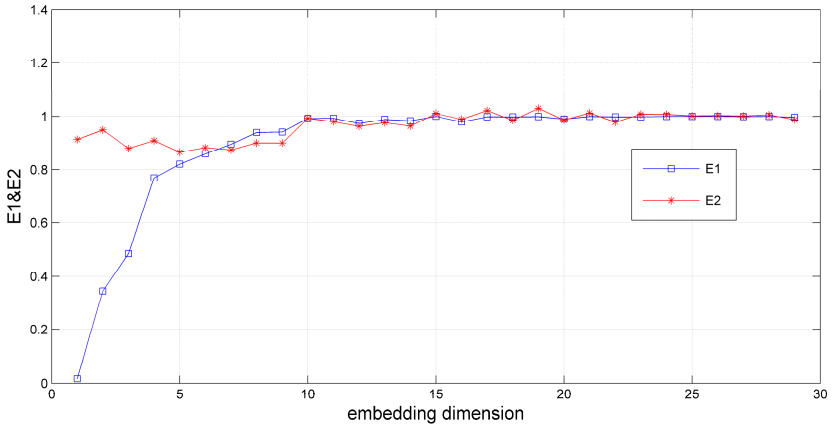


Figure 4. The m~E_1 and E_2 plot of X_{LLD}.

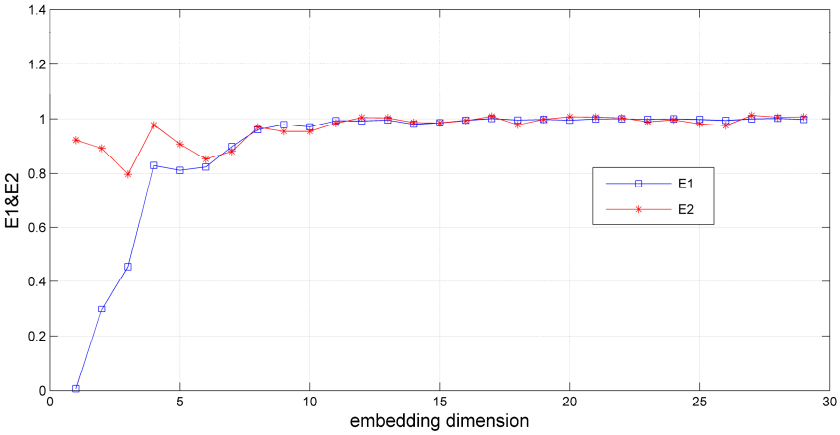


Table 5. The results for τ, m, p, D_2, K_2 and λ.

	τ	m	p	D_2	K_2	λ
X_{NI}	3	10	2.2957	1.2625	0.0045	0.00064
X_{LLD}	3	8	2.0048	4.7185	0.0195	0.00025

(3) The RPs are shown in Figures 5 and 6 after selecting the standard deviations as the threshold radius. Very short diagonals can be seen in these figures, which indicate chaotic behavior.

Figure 5. The RP plot of X_{NI}.

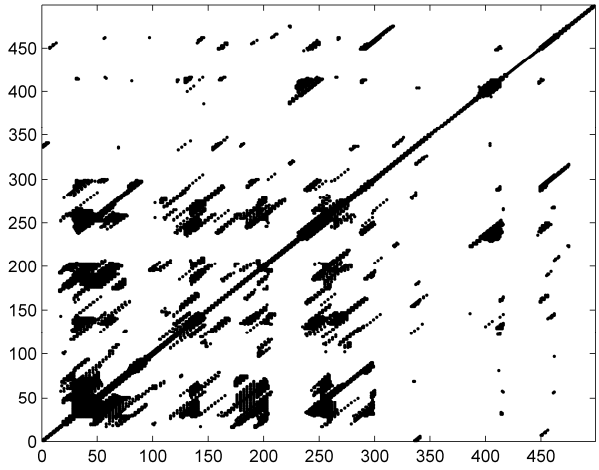
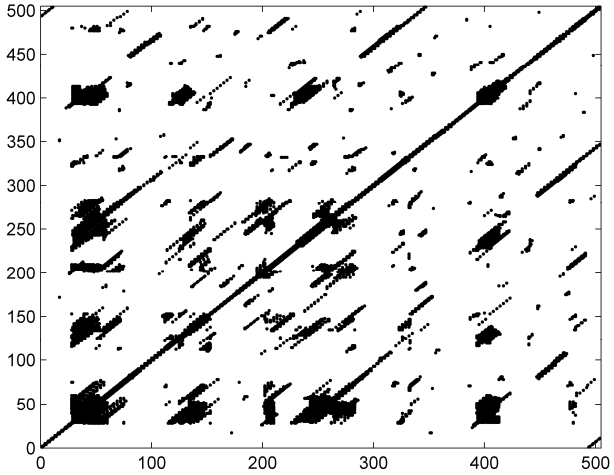
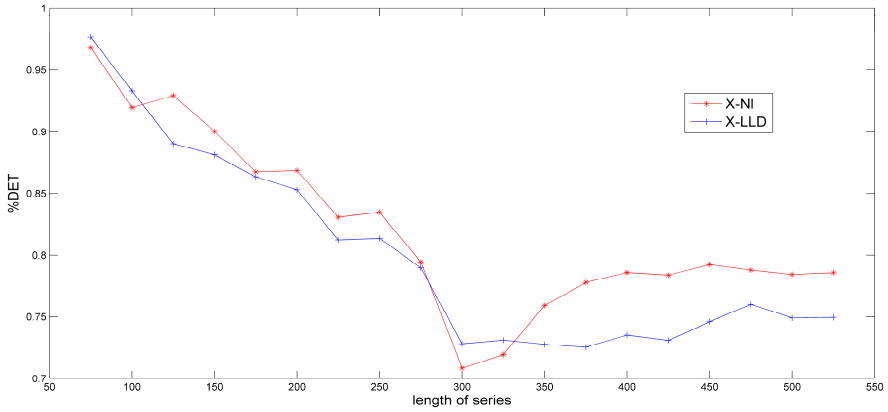


Figure 6. The RP plot of X_{LLD}.



We draw the %DET plots of the two time series which show the %DET changes with increasing time (see Figure 7, the red curve is for X_{NI} and the blue one is for X_{LLD}). We can see that the %DET tend to stable values. The %DET of X_{NI} and X_{LLD} are 0.7842 and 0.7479, respectively. We also calculate the %DET values of a random series, a sine series and a logistic series (μ = 4) with the same sample capacity 525. The results are 0.4229, 0.9969, 0.7627, respectively. Hence, the determinism of our time series are stronger than that of random series, are weaker than that of sine series and are similar to that of logistic series. Since the logistic series is chaotic when (μ = 4), it is possible that our series are chaotic, too. The series X_{NI} shows stronger determinism than X_{LLD}.

Figure 7. %DET curves of X_{NI} and X_{LLD}.



(4) We compute D_2, K_2, and the largest Lyapunov exponents λ of the two time series. All the results are listed in Figure 5. When we observe the plot of ln$r$~ln$C(m, N, r, \tau)$, we can get many linear scaling regions (for example, see Figure 8, the plot of X_{NI}ln$r$~ln$C(m, N, r, \tau)$). Therefore, in order to determine the optimal linear scaling region, we compute the derivatives and the second-order derivatives of ln$r$~ln$C(r)$ curves and draw the plots of ln$r$~ln$C'(r)$ and ln$r$~ln$C''(r)$ (for example, see Figure 9, the X_{NI} plots of ln$r$~ln$C'(r)$ and ln$r$~ln$C''(r)$) according to the new algorithm in this paper. Then we can find out the optimal linear scaling region (for example, see Figure 9, the interval [−0.8,−0.6] is optimal). The values of D_2, K_2 can be obtained in the scaling regions. Following the algorithms of weighted averages and Rosenstein, we can estimate the mean period p and the largest Lyapunov exponent λ (see Figure 5).

Figure 8. The X_{NI} plot of ln$r$~ln$C(m, N, r, \tau)$.

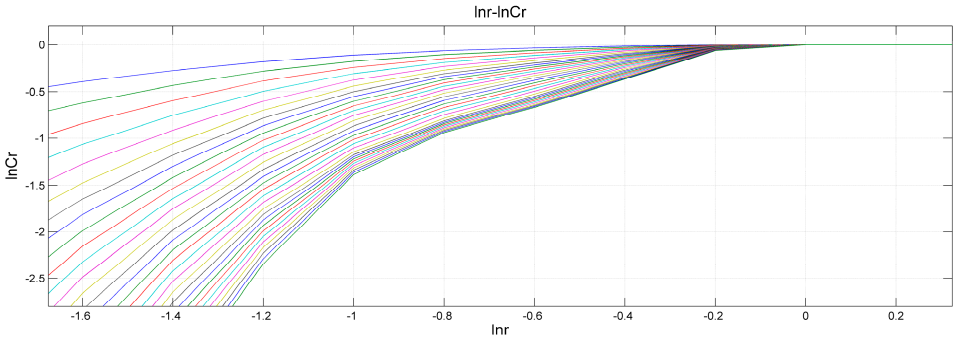
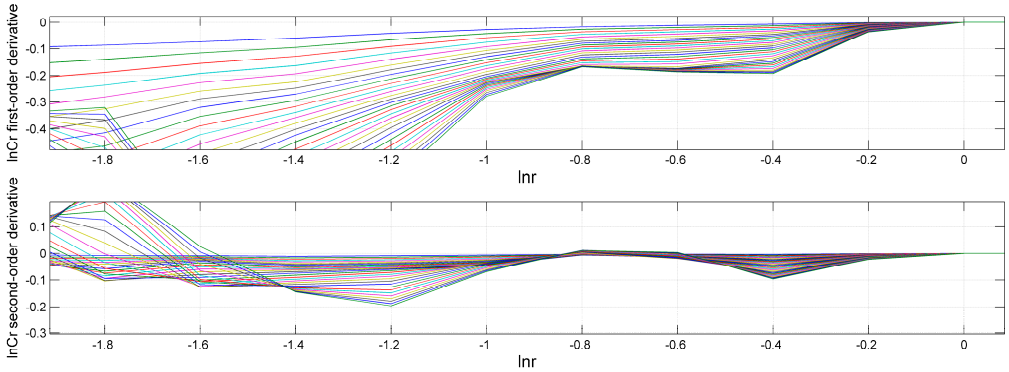


Figure 9. The X_{NI} plots of ln$r$~ln$C'(r)$and ln$r$~ln$C''(r)$).



From Table 5, the two series have the positive Kolmogorov entropy, the positive largest Lyapunov exponents, and non-integer correlation dimensions, which indicate a possible existence of chaos. Since the values of the largest Lyapunov exponents and Kolmogorov entropy are small, the price system exhibits behaviors of weak chaos. Another evidence of weak chaos is the correlation dimension 1.2625 of X_{Nl}. As the reciprocal of the Kolmogorov entropy describes a possible temporal forecasting scale, it is reasonable that the price can be forecasted from 1 to 4 years but cannot be forecasted for a long term.

In addition, the correlation dimension result of XLLD 4.7185 is not reliable. In fact, the minimum data points with a D-dimensional attractor can be estimated (see [24]). Empirically, 500 points are needed for D = 2 and more than 10,000 points for D = 3. So, with only 525 weekly data points, the discovery of economic strange attractors whose dimensionality is higher than 4 is unlikely. This also shows that the normalization method is more suitable for detecting chaos in agricultural price data.

6. Conclusions

In this paper, we consider a time series of cabbage prices, detrend this series by the approaches of LLD detrending and normalization, and analyze the two resulting detrended series. We can make the following conclusions: (1) the results of the BDS test and the stationary tests indicate that our series are nonlinear; (2) both of the two time series have Hurst exponents larger than 0.5, which means they are persistent and fractal with long-term memories; (3) the power spectra of the two data display wide continuum; (4) the two RPs show short line segments parallel to their main diagonals; the %DET of the two series are bigger than that of a random series, are smaller than that of a sine data and are similar to that of a chaotic logistic series, which imply determinism behavior; (5) we show that the largest Lyapunov exponents are positive, the estimates of the Kolmogorov entropy is positive, and the correlation dimensions are non-integer numbers, which are strong indicators for the chaotic behavior of the system; (6) the price can be forecasted for a short term, but cannot be forecasted for a long term; (7) based on the above evidence, we think that there is low-dimension and weak chaos in the price system; (8) the chaotic prediction methods can be used to forecast the agricultural products prices; (9) the multi- test can detect chaos even if the sample capacity is small and the level of noise is big; (10) for a chaotic time series, the chaotic prediction method is evidently more available, hence, when we find out chaos in a price data by using of the multi test, a chaotic or nonlinear prediction method is appropriate, conversely, when we cannot find chaos in a price data, we should use a non-chaotic prediction method.

Empirical analysis also show that: (1) different detrending approaches can affect the detection of chaos, since the first differencing or log first differencing detrending is noise-amplifier and the correlation dimension result of the LLD detrended series is not reliable; (2) the optimal linear scaling region can be quickly obtained by using our new algorithm, and then the correlation dimension and the K_2-entropy can be calculated simultaneously; (3) prediction techniques based on the chaos theory can be applied to forecast agricultural prices.

The results in this paper indicate that there are nonlinear, fractal, and chaotic characteristics in the agricultural price system. Therefore, nonlinear and chaotic models, as well as the prediction techniques based on the chaos theory, can be applied to forecast the agricultural products prices.

It is very important to know whether the agricultural prices are reasonable and stable. This not only affects agricultural development, the circulation and consumption of agricultural products, and the income level of farmers, but also affects the industrial costs and prices and affects the material benefit relationship between the country and the farmers, between the urban and the rural people, and among the farmers. Furthermore, it would have a profound impact on social stability. Nowadays, in China, a new bout of rising agricultural products prices has started: from 2006 to the end of 2010, the wheat, rice, corn and soybean prices rose 47%, 50%, 67% and 68%, respectively; in 2011, Chinese food prices year-on-year rose nearly 15%, where meat and poultry prices rose the most, the growth of pork prices reached 48.7%, mutton prices rose 25.3%, the growth of chicken and beef prices also reached 17.5% and 16.1%, respectively. Therefore, the government not only should follow the rules of the agricultural market, but also should strengthen the macro-control. It is very urgent for the government to develop and implement policies for stabilizing agricultural prices. However, an important prerequisite is to accurately grasp the characteristics of agricultural products prices and trends. Our paper represents exploratory research with this aim.

With the development of Chinese informatization, a lot of provinces, cities, areas and agricultural products wholesale markets regularly publish information about agricultural product prices, including the annual, monthly and daily data, thus how to make full use of these data to predict agricultural products prices is very important that.

In fact, presently, there are a lot of methods used to forecast the agricultural products prices including linear and nonlinear methods, econometrics models, time series models, neural networks, gray prediction methods, *etc.* However, on the one hand, agricultural products have huge number of varieties. The data of every different variety has different characteristics. It is not practical to try all the methods and find the best one. On the other hand, there are many causes which can lead to volatility in agricultural product prices, such as climate, supply and demand, prices of capital goods, energy prices, exchange rates, stocks, import and export trades, emergencies, natural disasters, *etc.* Moreover, the price evolution has greatly uncertainty. Any small changes would be likely to result in unpredictable results. It is usually ineffective to predict the agricultural product prices by establishing models including all the factors that cause the volatility, but if we can determine whether the data is chaotic, then for the chaotic data, we can use the chaotic prediction methods, and for non-chaotic data, we can use the non-chaotic prediction methods. The prediction efficiency will be improved a lot. Furthermore, the prediction methods according to the chaotic characteristics of the data can bypass the dilemma of looking for the volatility causes. Therefore, our method of detecting chaos can be considered as a data classification method for agricultural product prices and provides an effective way to choose the optimal prediction method. In addition, our methodology can also be used in other economic and financial data.

By comparing the present methods of predicting agricultural products prices, we can see that nonlinear methods, especially the neural network methods, have more advantages. However, the neural network method is only one of the chaotic prediction methods. Other chaotic predictions such as the chaotic local adding-weight forecasting algorithm and the Lyapunov exponent forecasting algorithm are rarely used to forecast agricultural product prices. What's more, most of the literature that used the neural network method did not dig into the essential reasons for using this method. We think that the reason why the nonlinear methods, using the neural network method as a representative, have higher

prediction accuracy is that data have nonlinear or chaotic characteristics. Furthermore, we also think the chaotic prediction methods should be widely used as the main method to forecast agricultural product prices.

For the chaotic agricultural product prices, any macro-control policy can have unpredictably serious impacts. Hence, the government should put forward any policy cautiously. Comparing with the research on looking for the volatility reasons of prices, we think that establishing the chaotic prediction models, studying the influence degree of different macro-control policies on chaos through simulations or artificial intelligence methods, and dealing with abnormal fluctuations by using of chaotic control or reverse control are more important and effective. Macroeconomic policies for adjusting agricultural prices should be quantified in order to realize the precise control. In addition, we expect that this study can provide a valuable reference for China and other developing countries for establishing price policies, regulating markets, and inspecting implementation effects in their agricultural product markets.

Acknowledgments

This work is partially supported by the National Natural Science Foundation of China (No. 61340043); this work is also partially supported by Program for Soft Science Research Project of Shandong Province (No. 2013RKB01253). The authors would like to acknowledge the referees for their valuable comments and suggestions which helped to improve the presentation of this paper.

Author Contributions

Xin Su, Yi Wang, Shengsen Duan and Junhai Ma all contributed to this study. Xin Su generated the idea and collected the data. Yi Wang chose the method, analyzed the data and programmed. Xin Su and Yi Wang reviewed the literature and provided research funding for the study. Yi Wang and Xin Su drafted the article. Yi Wang, Xin Su, Shengsen Duan and Junhai Ma reviewed and edited the manuscript. All authors have read and approved the final manuscript.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Raneses, A.R.; Glaser, L.K.; Price, J.M.; Duffield, J.A. Potential biodiesel markets and their economic effects on the agricultural sector of the United States. *Ind. Crops Prod*. **1999**, *9*, 151–162.
2. Tokgoz, S. *The Impact of Energy Markets on the Eu Agricultural Sector*; Working Paper 09-WP485; Center for Agricultural and Rural Development, Iowa State University: Ames, IA, USA, 2009.
3. Benavides, G. *Price Volatility Forecasts for Agricultural Commodities: An Application of Historical Volatility Models, Option Implied and Composite Approaches for Futures Prices of Corn and Wheat*; Central Bank of Mexico: Mexico City, Mexico, 2004.
4. Mitra, S.; Boussard, J. A simple model of endogenous agricultural commodity price fluctuations with storage. *Agric. Econ*. **2012**, *43*, 1–15.

5. Garner, C.A. Commodity prices: Policy target or information variable note. *J. Money Credit Bank.* **1989**, *21*, 508–514.
6. Pindyck, R.S.; Rotemberg, J.J. The excess co-movement of commodity price. *Econ. J.* **1990**, *100*, 1173–1189.
7. Cheng, G.; Hu, B.; Xu, X. An analysis of the impact of the new round of rise in the prices of agricultural produce. *Manag. World* **2008**, *1*, 57–62.
8. Xu, X. The New-round fluctuation cycle of agricultural products prices: Characteristics, mechanism and effects. *J. Financ. Econ*. **2008**, *34*, 110–119.
9. Ma, J.; Liu, L. Multivariate Nonlinear Analysis and Prediction of Shanghai Stock Market. *Discret. Dyn. Nat. Soc.* **2008**, *2008*, 526734.
10. Ma, J.; Cui, Y.; Liu, L. Hopf bifurcation and chaos of financial system on condition of specific combination of parameters. *J. Syst. Sci. Complex.* **2008**, *21*, 250–259.
11. Ma, J.; Wang, Z.; Chen, Y. Prediction techniques of chaotic time series and its applications at low noise level. *Appl. Math Mech.* **2006**, *1*, 7–14.
12. Hommes, C.H. Dynamics of the cobweb model with expectations and nonlinear supply and adaptive demand. *J. Econ. Behav. Organ.* **1994**, *24*, 315–335.
13. Onozaki, T.; Sieg, G; Yokoo, M. Complex dynamics in a cobweb model with adaptive production adjustment. *J. Econ. Behav. Organ.* **2000**, *41*, 101–115.
14. Commendatore, P.; Currie, M. The cobweb, borrowing and financial crises. *J. Econ. Behav. Organ.* **2008**, *66*, 625–640.
15. Mandelbert, B.B. New method in statistical economics. *J. Polit. Econ.* **1963**, *71*, 421–440.
16. Corazza, M.; Malhairs, A.; Nardelli, C. Searching for fractal structure in agricultural futures markets. *J. Futures Mark.* **1997**, *17*, 433–473.
17. Barkoulas, J.T.; Labys, W.C.; Onochie, J.I. Long memory in futures prices. *Finical Rev.* **1999**, *34*, 91–95.
18. Wei, A.; Leuthold, R.M. *Agricultural Futures Prices and Long Memory Processes*; Working Paper 00–04; Office for Futures and Options Research, University of Illinois at Urbana-Champaign: Champaign, IL, USA, 2000.
19. Yang, S.R.; Brorsen, B.W. Nonlinear dynamics of daily cash prices. *Am. J. Agric. Econ*. **1992**, *74*, 706–715.
20. Yang, S.R.; Brorsen, B.W. Nonlinear dynamics of daily futures prices: Conditional heteroscedasticity or chaos? *J. Futures Mark.* **1993**, *13*, 175–191.
21. Michael, J., Jr.; Onochie, J. A bivariate generalized autoregressive conditional heteroscedasticity-in-mean study of the relationship between return variability and trading volume in international futures markets. *J. Futures Mark.* **1998**, *18*, 379–397.
22. Bhar, R. Return and Volatility Dynamics in the Spot and Futures Markets in Australia: An Intervention Analysis in a Bivariate EGARCH-X Framework. *J. Futures Mark.* **2001**, *21*, 833–850.
23. Papaioannou, G.; Karytinos, A. Nonlinear time series analysis of the stock exchange: The case of an emerging market. *Int. J. Bifurc. Chaos* **1995**, *5*, 1557–1584.
24. Chen, P. A random-walk or color-chaos on the stock market?-time-frequency analysis of S&P indexes. *Stud. Nonlinear Dyn. Econom.* **1996**, *1*, 87–103.

25. Rolo-Naranjo, A.; Montesino-Otero, M. A method for the correlation dimension estimation for online condition monitoring of large rotating machinery. *Mech. Syst. Signal Process.* **2005**, *19*, 939–954.
26. Fraser, A.M.; Swinney, H.L. Independent coordinates for strange attractors from mutual information. *Phys. Rev. A* **1986**, *33*, 1134–1140.
27. Cao, L.Y. Practical method for determining the minimum embedding dimension for a scalar time series. *Physica D* **1997**, *110*, 43–50.
28. Faggini, M. Chaos detection in economics. In *Metric versus Topological Tools*; MPRA Paper No. 30928; Munich Personal RePEc Archive: Munich, Germany, 2011.
29. Webber, C.L., Jr. Dynamical assessment of physiological systems and states using recurrence plot strategies. *J. Appl. Physiol.* **1994**, *76*, 965–973.
30. Zbilut, J.P.; Giuliani, A.; Webber, C.L., Jr. Recurrence quantification analysis as an empirical test to distinguish relatively short deterministic *versus* random number series. *Phys. Lett. A* **2000**, *267*, 174–178.
31. Grassberger, P.; Rorcaccia, I. Measuring the strangeness of strange attractors. *Physica D* **1983**, *9*, 189–208.
32. Grassberger, P.; Rorcaccia, I. Estimation of the Kolmogorov entropy from a chaotic signal. *Phys. Rev. Lett.* **1983**, *28*, 2591–2593.
33. Zhao, G.; Shi, Y.; Duan, W.; Yu, H. Computing fractal dimension and the Kolmogorov entropy from chaotic time series. *Chin. J. Comput. Phys*. **1999**, *16*, 309–315.
34. Rosenstein, M.T.; Collins, J.J.; de Luca, C.J. A practical method for calculating largest Lyapunov exponents from small data sets. *Physica D* **1993**, *65*, 117–134.
35. Yang, Y.; Wu, M.; Gao, Z.; Wu, Y.; Ren, X. Parameters selection for calculating largest Lyapunov exponent for small data sets. *J. Vib. Meas. Diagn.* **2012**, *32*, 371–374.