

Article

Entropy and Recurrence Measures of a Financial Dynamic System by an Interacting Voter System

Hong-Li Niu and Jun Wang *

Department of Mathematics, Beijing Jiaotong University, Beijing 100044, China; E-Mail: 11118384@bjtu.edu.cn

* Author to whom correspondence should be addressed; E-Mail: wangjun@bjtu.edu.cn; Tel./Fax: +86-10-5168-2062.

Academic Editor: Kevin H. Knuth

Received: 25 January 2015 / Accepted: 20 April 2015 / Published: 23 April 2015

Abstract: A financial time series agent-based model is reproduced and investigated by the statistical physics system, the finite-range interacting voter system. The voter system originally describes the collective behavior of voters who constantly update their positions on a particular topic, which is a continuous-time Markov process. In the proposed model, the fluctuations of stock price changes are attributed to the market information interaction amongst the traders and certain similarities of investors' behaviors. Further, the complexity of return series of the financial model is studied in comparison with two real stock indexes, the Shanghai Stock Exchange Composite Index and the Hang Seng Index, by composite multiscale entropy analysis and recurrence analysis. The empirical research shows that the simulation data for the proposed model could grasp some natural features of actual markets to some extent.

Keywords: financial time series agent-based model; interacting voter system; complexity analysis; multiscaled entropy; recurrence analysis; computer simulation

1. Introduction

In recent decades, the modeling of the dynamics of price fluctuation behaviors in the financial market has sparked considerable interest in both the finance and physics community, which is becoming a key problem in risk management, physical asset valuation and derivatives pricing. It is also becoming increasingly important for understanding the mechanisms of dynamical financial price variations that have exhibited some interesting statistical properties, such as the fat tails phenomenon, power law of logarithmic return and volume, volatility clustering, multifractality of volatility, etc. [1-9]. Various agent-based models have been introduced to make an empirical study of the fluctuations in the market; for instance, see [10–27]. Some of these models are created by applying the interacting particle, system theories and methods, such as the percolation network [11,17,20,23,24,28], the Ising dynamic system [12,13], the contact model [10,25,26] and the voter system [19]. For example, Lux and Marchesi [14] introduced an agent-based model in which chartist agents compete with fundamentalists agents, leading to power law distributed returns as observed in real markets, which contradicts the popular efficient market hypothesis. Fang and Wang [12] developed an interacting-agent model of speculative activity explaining price formation in financial market that is based on the stochastic Ising dynamic system. Through computer simulation and empirical study, it shows that the established financial model can reproduce the main factors and reveal certain statistical characteristics of asset returns. Actually, it is widely accepted that the financial market is an evolving dynamic system that reacts to external investment information to determine the best price for a given asset. It consists of a great number of agents interacting with one another in complicated ways. In modeling of the financial fluctuations, one of the most important thing is to find or define a proper mechanism for interacting information that the market investors hold. This is also what many economists are dedicated to. The establishment of a modeling process enriches the theoretical study of financial stock pricing. The practical relevance of a thorough understanding of the mechanism governing market fluctuation lies in the benefits that this induces in the processes of asset allocation and risk management. The dynamic voter interacting system is the famous statistical physics model, which can be also viewed as a model for non-equilibrium statistical mechanics [29–32]. In the voter process, the voters constantly update their attitudes at independent exponential random variables. At times of reconsideration, a voter chooses one neighbor uniformly from amongst all neighbors and takes that neighbor's opinion. In this case, the voter theory could be taken to describe the decision making mechanism amongst economic agents in the market. In this paper, a financial price model is introduced by the finite-range voter system, in which we also assume that the investors' attitudes towards markets lend to the fluctuations of stock prices and suppose that the interacting particles in the voter system represent the investment opinions. Applying the voter system to the financial modeling may provide a good potential link between economics and physics and give a beneficial way to depict the mechanism of market agents.

Afterwards, we make an empirical study of the statistical behaviors of logarithmic returns for the proposed price model in comparison with two important Chinese stock indexes, the Shanghai Stock Exchange (SSE) Composite Index and the Hang Seng Index (HSI). We mainly focus on the exploration of the complexity of the financial time series. The multiscale entropy (MSE) method was one such method developed to quantify the relative complexity of normalized time series across multiple scales [33], which consists of two steps: (1) a coarse-graining procedure is used to derive the representations of a system's dynamics at different time scales; and (2) the sample entropy (SampEn) algorithm is used to quantify the regularity of a coarse-grained time series at each time scale factor. Various literature has seen the applications of this method to many research areas [33–36]. However, the reliability of SampEn is reduced as a time scale factor is increased. More specifically, the variance of the

entropy estimator grows very quickly as the number of data points is reduced, and in the real application, the data length is not often long enough, which will lead to less reliability in distinguishing time series. A modified algorithm in [37], called composite multiscale entropy (CMSE), was introduced to overcome this weakness, in which the experimental results showed that the CMSE presented better performance on short time series than the MSE and could provide a more reliable entropy estimation by the analysis of both white and 1/f noise series. In the present paper, the CMSE analysis is adopted as a measure of the complexity of simulated data and real ones. Furthermore, nonlinear determinism can potentially explain large movements in financial data that linear stochastic models cannot account for [38]. Thus, we here apply the recurrence plots (which provide visual insight into the complex nonlinear deterministic patterns in time series data [39,40]) and the recurrence quantification analysis (RQA) method (that is able to quantify structure in the recurrence plots through different recurrence measures [41,42]) to the nonlinear analysis of the real market data and the simulated data derived from the financial price model.

In sum, there are two main contributions in this paper. One is to provide a financial price modeling process, in which the information interaction mechanism of agents in the stock market is described by the long-range voter system. The other is to study the complexity behaviors of simulation time series by the CMSE method (a recent proposed method in engineering) and the famous RQA technique, respectively. We hope that our study could enrich the modeling and statistical analysis of the financial market.

2. Price Process Modeling by a Finite-Range Voter System

In this section, we intend to construct a price simulation process according to the mechanisms of a finite-range-biased voter interacting system. First, we give a mathematical description of the finite-range-biased voter model. The voter model is an interacting particle system [29–32], describing the collective behavior of voters who constantly update their political positions. Individuals placed at the points of \mathbb{Z}^d might have one of two possible opinions on a political issue (in favor or against), at independent exponential times; an individual reassesses his view by choosing a neighbor at random with certain probabilities and then adopting his position. Let $\eta_s(x)$ be the position of voter x at time s ($s \ge 0$); the political position of the voters can be denoted by $\eta_s = \{\eta_s(x) : x \in \mathbb{Z}^d\}$, which is a mapping function from \mathbb{Z}^d to $\{0,1\}$, *i.e.*, $\eta_s(x) = 1$ if $x \in \eta_s$ and $\eta_s(x) = 0$ if $x \notin \eta_s$. Then, the dynamics of the voter model can be formulated as follows: (i) if $x \in \eta_s$, then x becomes vacant at a rate equal to the number of vacant neighbors; (ii) if $x \notin \eta_s$, then x becomes occupied at a rate equal to λ times the number of occupied neighbors, where $\lambda(>1)$ is an intensity called the "carcinogenic advantage". When $\lambda = 1$, the model is called the basic voter model, and it is called the biased voter model when $\lambda > 1$. Let η_s^A denote the state at time s with the initial state set $\eta_0^A = A$ and $\eta_s^{\{0\}}(x)$ be the state of $x \in \mathbb{Z}^d$ at time s for $\eta_0^{\{0\}} = \{0\}$. More generally, we consider the initial distribution as v_{θ} , the product measure with density θ (each site is independently occupied with probability θ) and let η_s^{θ} denote the voter model with initial distribution v_{θ} . More formally, the stochastic dynamics of voter model η_s is a Markov process on a configuration space $\{0,1\}^{\mathbb{Z}^d}$, whose generator has the form:

$$\Omega g(\eta) = \sum_{x \in \mathbb{Z}^d} c(x, \eta) \left[g(\eta^x) - g(\eta) \right]$$
(1)

where the functions g is on $\{0,1\}^{\mathbb{Z}^d}$ that depend on finitely many coordinates, and $\eta^x(z) = \eta(z)$ if $z \neq x$, $\eta^x(z) = 1 - \eta(x)$ if z = x, for $x, z \in \mathbb{Z}^d$. $c(x, \eta)$ is the transition rate function for the process, which is given by the following (see [30–32]). For any $x \in \mathbb{Z}^d$, the state of $x \in \mathbb{Z}^d$ flips according to the transition rates:

$$0 \rightarrow 1$$
 at rate $\lambda \sum_{y \in \mathbb{Z}^d} p(x, y) \mathbb{I}_{\{\eta(y)=1\}}$ (2)

$$1 \rightarrow 0$$
 at rate $\sum_{y \in \mathbb{Z}^d} p(x, y) \mathbb{I}_{\{\eta(y)=0\}}$ (3)

where \mathbb{I} is the indicative function, $p(x, y) \geq 0$ for $x, y \in \mathbb{Z}^d$ and $\sum_{y \in \mathbb{Z}^d} p(x, y) = 1$ for all $x \in \mathbb{Z}^d$. Here, we suppose that the transition probability p(x, y) is translation invariant and symmetric and such that the Markov chain with those transition probabilities is irreducible [15,43]. If a site $x \in \mathbb{Z}^d$ is occupied by a one (resp. zero), then at rate one (resp. λ), it picks a site $y \in \mathbb{Z}^d$ with probability p(x, y) and adopts the state of the individual at y. For the biased voter model ($\lambda > 1$), there exists a "critical value" for the process, which is defined as $\lambda_c = \inf\{\lambda : P(|\eta_s^{\{0\}}| > 0, \text{ for all } s \geq 0) > 0\}$, where $|\eta_s^{\{0\}}|$ is the cardinality of $\eta_s^{\{0\}}$. Assume $\lambda > \lambda_c$, then there is a convex set C, so that on $\Omega_{\infty} = \{\eta_s^{\{0\}} \neq \emptyset$, for all $s\}$, we have for any $\epsilon > 0$ and for all s sufficiently large, $(1 - \epsilon)sC \cap \mathbb{Z}^d \subset \eta_s^{\{0\}} \subset (1 + \epsilon)sC \cap \mathbb{Z}^d$. If $\lambda < \lambda_c$, for some positive $\rho(\lambda)$, we have $P(\eta_s^{\{0\}} \neq \emptyset) \leq e^{-\rho s}$. The above results imply that, on a d-dimensional lattice, the process becomes vacant exponentially for $\lambda < \lambda_c$; the process survives with the positive probability for $\lambda > \lambda_c$.

We introduce the graphical representation of a one-dimensional-biased voter model on the configuration space $\{0,1\}^{\mathbb{Z}}$ [31,32], since the graphical representation is very useful to illustrate and simulate the model. We start by constructing the process η_s from a collection of Poisson processes in the case $\lambda \geq 1$. For each pair $x, y \in \mathbb{Z}$ with $|x - y| \leq R$ (R is the finite-range), let $\{T_n^{(x,y)} : n \geq 1\}$ and $\{U_n^{(x,y)} : n \geq 1\}$ be independent Poisson processes with rate one and $\lambda - 1$, respectively. At times $T_n^{(x,y)}$, we draw an arrow from y to x and put a δ at x. At times $U_n^{(x,y)}$, we just draw an arrow from y to x. Then, the process is obtained from the graphical representation as follows: At time $T_n^{(x,y)}$, the state of site x imitates the state of site y, *i.e.*, becomes occupied by a one (resp. zero) if site y is occupied by a one, and the state of site x is not affected if y is occupied by a zero. A figure illustration of the construction of the graphical representation for a one-dimensional-biased voter model with neighbor range R = 3 is presented in Figure 1. We imagine fluid entering the bottom and flowing up the structure. The δ 's are like dams, and the arrows are like pipes, which allow the fluid to flow in the indicated direction.

In the following, a financial price simulation process model is developed by applying the finite-range-biased voter dynamic system with neighbor range R. In this model, we assume that the investors' investment attitudes towards the financial market lead to fluctuations of stock prices and suppose that the investment attitudes are represented by the interacting particles in the biased voter model. We may classify the investment attitudes into buying, selling and neural ones, which correspondingly sort the investors' into three groups according to the attitudes that they hold. We assume that each trader can trade the stock several times at each day $t \in \{1, 2, \dots, N\}$, but at most a unit number of the stock at each time. Let l be the time length of trading day; we denote the stock price at time s in the t-th trading day by $P_t(s)$, where $s \in [0, l]$. Suppose that the stock market consists

of 2M + 1 (*M* is large enough) traders, who are located in a line $\{-M, \dots, -1, 0, 1, \dots, M\} \subset \mathbb{Z}$ (similarly for a *d*-dimensional lattice \mathbb{Z}^d). At the very beginning of each trading day, we select a certain proportion of traders (with the initial distribution v_{θ}) randomly in the system and consider them as those who receive some market news. We define a random variable ζ_t with the values +1, -1, 0 to represent that these investors hold a buying opinion, selling opinion or neutral opinion with probabilities p_{+1} , p_{-1} or $1 - (p_{+1} + p_{-1})$, respectively. Then, these investors send a bullish, bearish or neutral signal to their finite-range neighbors. According to the *d*-dimensional voter process system, investors can affect each other or the news can be spread, which is supposed as the main factor of price fluctuations for the market. The aggregate excess demand for the asset at time *t* is defined by:

$$B_t(s) = \zeta_t \cdot \left| \eta_s^{\theta} \right| / (2M+1) \tag{4}$$

where $|\eta_s^{\theta}| = \sum_{z=-M}^{M} \eta_s^{\theta}(z)$, and *M* may depend on the trading days *N*. From the above description and [15,43,44], we define the simulation formula of a discrete time stock price as follows:

$$P_t(s) = P_{t-1}(s) \exp\{\beta B_t(s)\}, \qquad P_t(s) = P_0 \exp\{\beta \sum_{i=1}^t B_i(s)\}$$
(5)

where $\beta(> 0)$ represents the depth parameter of the market, which measures the sensitivity of price fluctuation in response to the change in excess demand, and P_0 is the stock price at Time 0. The corresponding stock logarithmic return and absolute return from t - 1 to t are defined by:

$$r(t) = \ln P_t(s) - \ln P_{t-1}(s), \quad |r(t)| = |\ln P_t(s) - \ln P_{t-1}(s)|.$$
(6)



Figure 1. Graphical representation of the biased voter model with finite-range R = 3.

According to the above definition and description of the model, we perform the simulation of stock price series and return series with different values of the parameters [45,46], finite-range R and intensity λ in the voter dynamic system. We set the number of traders M = 500 and the initial density of the model $\theta = 0.01$. Three different simulation data are considered for fluctuation behavior analysis with parameter settings { $\lambda = 1.3, R = 1$ }, { $\lambda = 1.5, R = 2$ } and { $\lambda = 1.7, R = 3$ }, respectively, which

cover 6000 data points. In selecting the parameters, we first set the finite range of the voter model R = 1, 2, 3, respectively. For λ , we consider the case of $\lambda > \lambda_c$ and gradually change the value to make the simulated returns be as close as possible to the real market indices. For each setting, 20 repetitions are performed, and the averaged time series is the simulation data for analysis. Meanwhile, we select the daily prices of two important actual market stock indexes in comparison with the simulation ones, the Hang Seng Index (HSI) from 31 December 1990 to 18 March 2013, with 5530 data points, and the Shanghai Stock Exchange (SSE) Composite Index from the period 3 March 1993 to 22 February 2013, with 5142 data points. The normalized plots of these prices and the logarithmic return for simulation data with parameter setting { $\lambda = 1.7, R = 3$ } are shown in Figure 2.



Figure 2. (a) Normalized price series for simulation data and actual data; (b) return and absolute return series for simulation data with parameter $\{\lambda = 1.7, R = 3\}$.

3. Comparison Empirical Analysis of Complexity Behaviors

3.1. Composite Multiscale Entropy Analysis

In this section, we adopt a modified new multiscale entropy analysis algorithm, the composite multiscale entropy (CMSE) method, for complexity analysis of the above-mentioned three simulation data and two actual financial market indexes, which was proposed to overcome the problem that the reliability of the sample entropy of a coarse-grained series is reduced as a time scale factor is increased [37]. The difference between the CMSE and MSE is in the coarse-graining procedure, and the CMSE can be carried out on a time series in the following two steps:

(1) For an one-dimensional time series $\mathbf{x} = \{x_1, x_2, \cdots, x_N\}$, consecutive coarse-grained time series are constructed by averaging a successively increasing number of points within non-overlapping windows. Unlike the MSE algorithm in which each of the coarse-grained time series $\{\mathbf{y}^{(\tau)}\}$ is computed as $y_j^{(\tau)} = \frac{1}{\tau} \sum_{i=(j-1)\tau+1}^{j\tau} x_i$, the k-th coarse-grained time series in the CMSE method for a scale factor τ , $\mathbf{y}_k^{(\tau)} = \{y_{k,1}^{(\tau)}, y_{k,2}^{(\tau)}, \cdots, y_{k,p}^{(\tau)}\}$ is defined as:

$$y_{k,j}^{(\tau)} = \frac{1}{\tau} \sum_{i=(j-1)\tau+k}^{j\tau+k-1} x_i, \qquad 1 \le j \le N/\tau, \ 1 \le k \le \tau.$$
(7)

Note that for $\tau = 1$, the coarse-grained time series is simply the original time series. Figure 3 shows a schematic illustration of the coarse-graining procedure for both MSE (a) and CMSE (b) with $\tau = 2$ and $\tau = 3$, respectively, from which a clear difference between these two methods can be seen.



Figure 3. (a) Schematic illustrations of the MSE and (b) composite multiscale entropy (CMSE) procedures, respectively.

(2) The entropy measure, the sample entropy (SampEn), is calculated for each coarse-grained time series and then plotted as a function of the scale factor. SampEn quantifies the regularity or predictability of a time series, which is defined as the negative logarithm of the conditional probability that a point that repeats itself within a tolerance of ϵ in an *m*-dimensional phase space will repeat itself in an *m* + 1-dimensional phase space:

$$SampEn = -\log[C(m+1,\epsilon)/C(m,\epsilon)]$$
(8)

where $C(m, \epsilon)$ is the number of repeating points in the *m*-dimensional phase space, repeating is defined as points closer than ϵ in a Euclidean sense to the examined points; for details, see [35–37,47]. Finally, the CMSE value is defined as the means of the sample entropies of all coarse-grained time series, that is:

$$CMSE(\mathbf{x},\tau,m,\epsilon) = \frac{1}{\tau} \sum_{k=1}^{\tau} SampEn(\mathbf{y}_k^{(\tau)},m,\epsilon)$$
(9)

while the MSE is computed by only using the first coarse-grained time series $\mathbf{y}_1^{(\tau)}$, *i.e.*, MSE($\mathbf{x}, \tau, m, \epsilon$) = SampEn($\mathbf{y}_1^{(\tau)}, m, \epsilon$).



Figure 4. Composite multiscale entropy of return series r and absolute return form $|r|^q$ with q = 0.25, 0.5, 0.75, 1, 1.5, 2 for simulation data and actual indexes, respectively.

In Figure 4, we calculate the CMSE values of returns and absolute returns with different power exponents, labeled as $|r(t)|^q$, for simulation data and actual data from Scale 1 to Scale 30 ($\tau = 1$ to 30), where q is taken as 0.25, 0.5, 0.75, 1, 1.5 and 2, respectively. It is known that the absolute return is a proxy of the volatility of time series, and $|r|^q$ exhibit obvious different volatility behaviors for different q [19]. The entropy value of each coarse-grained time series is calculated with phase space embedding dimension m = 6 for minimizing the fraction of false neighbors [38] and $\epsilon = 0.15\sigma$, where σ denotes the standard deviation of the original time series. From the figure, we find that values of the composite multiscale entropy of return series monotonically decrease as the scale factor increases for either actual

indexes or simulation data from the financial price model. For the absolute return with different power exponents $|r(t)|^q$, it is observable that the entropy values decrease gradually and, finally, remain almost constant as the scale factor becomes larger, which indicate that each of these time series, unlike the return series, contains correlations and complex structures across multiple time scales. The behavior for each of these volatility time series is similar to that of 1/f series, which has the correlated fluctuations (but the degrees of correlations are different). Meanwhile, it is observed that the entropy values for all given scales become smaller with the power exponents q becoming larger, suggesting decreasing complex structures. The simulative data for the financial price model show similar fluctuation behaviors to the actual SSE and HSI data.

To find some relationships between the correlations and the complexity of time series, we adopt the well-known detrended fluctuation analysis (DFA) method, which was proposed to explore the long-range correlations of nonstationary time series [19], to calculate the Hurst exponents of return series and $|r(t)|^q$ series for both the financial price model and actual market indexes. The results can be found in Table 1. It is seen that the Hurst exponents of return series for both the simulation data and the actual ones are around 0.5, indicating weak correlations. The Hurst exponents for $|r(t)|^q$ are all much larger than 0.5, which means that the time series are long-range auto-correlated. However, for each time series, the Hurst exponents do not show the obvious monotonic relationship with power exponent q. Therefore, we cannot intuitively conclude that the stronger correlations of time series correspond to higher complex structures or entropy values.

data	return	q = 0.25	q = 0.5	q = 0.75	q = 1	q = 1.5	q=2
$\{\lambda=1.3, R=1\}$	0.52026	0.63220	0.74669	0.81593	0.85684	0.88665	0.87644
$\{\lambda=1.5, R=2\}$	0.51066	0.65271	0.68973	0.76012	0.80596	0.84566	0.84251
$\{\lambda=1.7, R=3\}$	0.44587	0.63266	0.74921	0.82767	0.87982	0.92975	0.93350
HSI	0.52570	0.85341	0.88780	0.89769	0.89280	0.85217	0.79179
SSE	0.52271	0.72742	0.79722	0.81659	0.81297	0.76844	0.70621

Table 1. Hurst exponent of $|r|^q$ for simulation data and actual data. HSI, Hang Seng Index; SSE, Shanghai Stock Exchange.

In Figure 5, we present the CMSE results for return series, absolute returns and their corresponding shuffled time series. From Figure 5a, it is seen that the CMSE values of shuffled returns for the price model and the actual indexes are decreasing as the time scale increases, similar to the trends of original returns. In Figure 5b, it is shown that the CMSE values of shuffled absolute return series are also reducing with the time scale enlarging, but the trends are quite different from the original absolute returns. These statistical behaviors may be understand as follows. In the shuffling procedure, the values of time series are put into random order, and thus, all correlations are destroyed; therefore, the composite multiscale sample entropy of shuffled time series presents behaviors similar to the uncorrelated or weak correlated series.

Figure 5. (a) CMSE of return series and the corresponding shuffled time series; (b) CMSE of absolute returns and the corresponding shuffled time series.

3.2. Recurrence Plot and Recurrence Quantification Analysis

In this section, we utilize the recurrence plot and recurrence quantification analysis to explore the complex determinism of return time series for the price model, SSE and HSI. The recurrence plot (RP) is a qualitative tool for visualizing nonlinear dynamics in time series data. RP shows when a point in the phase space is near (at a distance lower than a certain threshold) to another point [39,40]. To perform the recurrence plot analysis, a phase space from a single observation is constructed. The state of the system can be represented by the discrete time delay vector $\mathbf{x}_t = \{x_t, x_{t-\Delta t}, x_{t-2\Delta t}, \dots, x_{t-(m-1)\Delta t}\}$, where *m* denotes the embedding dimension and Δt is the time delay. Then, the Euclidean distance matrix \mathbf{R} using the independence time-delayed coordinates is calculated:

$$\mathbf{R}_{ij} = \Theta(h - ||\mathbf{x}_i - \mathbf{x}_j||), \quad i, j = 1, 2, \cdots, N_R = N - (m - 1)\Delta t,$$
(10)

where $|| \cdot ||$ is a norm, $\Theta(\cdot)$ is the Heaviside step function and h is a threshold value, which has the meaning of the tolerance of recurrence. The distance matrix consists of zeros and ones and corresponds to the state of the system (one, recurrence and zero, no recurrence).

In this paper, the embedding dimension is selected through the false nearest neighbors method [38], and a value m = 6 seems appropriate for the time series under study. The time delay is fixed to $\Delta t = 1$ by the average mutual information method [48]. Concerning the recurrence threshold h, we adopt the 5% ($h \approx 0.02$) and 10% ($h \approx 0.04$) of the maximal phase space diameter of the discussed time series. In Figure 6 and Figure 7, the recurrence plots of returns for the simulated data and two actual market data are depicted in the case of h = 0.02 and h = 0.04, respectively. In Figure 6, more of the recurrence points are observed in the main diagonal lines for these time series, while the recurrence structure consisting of roughly vertical (horizontal) patterns can be found in the plots of Figure 7.

Figure 6. Recurrence plots of returns for simulation data and actual data in the case of h = 0.02.

Figure 7. Recurrence plots of returns for simulation data and actual data in the case of h = 0.04.

Since the graphical representation may be difficult to evaluate, recurrence quantification analysis (RQA) [41,42,49] was developed to provide numerical measures that allow for the quantification of the structure and complexity of RPs. These quantities are based on recurrence points densities, diagonal and vertical line structure. The recurrence rate:

$$\mathbf{RR} = \frac{1}{N_R} \sum_{i,j=1}^{N_R} \mathbf{R}_{i,j} \tag{11}$$

which measures the density of recurrence points in an RP and can be interpreted as the probability that any state may recur. One of the RQA measures based on the line parallel to the main diagonal is determinism (DET), which is defined by:

$$DET = \frac{\sum_{l=l_{\min}}^{N_R} lP(l)}{\sum_{l=1}^{N_R} lP(l)}$$
(12)

where P(l) is a histogram of diagonal lines of the length l, and l_{\min} is the minimal length of a diagonal line that is defined by $l_{\min} = 2$. DET provides an indication of determinism and predictability in the system; thus the larger the value of DET, the more predictable the system with diagonal lines in RP. Another measure on the diagonal line is the Shannon information entropy (L_{ENT}) defined for diagonal line collections:

$$L_{ENT} = -\sum_{l=l_{\min}}^{N_R} p(l) \ln p(l)$$
(13)

where the probability of the line distribution is $p(l) = P(l) / \sum_{l \ge l_{\min}} P(l)$. The rise in L_{ENT} refers to the increase of complexity of the time series. Moreover, the mean length of the diagonal lines $L_{mean} = \sum_{l=l_{\min}}^{N_R} lp(l)$ is a parameter indicating the system stability. Instead of considering diagonal lines, we measure vertical recurrence lines. In analogy to determinism, the laminarity (LAM) is defined for vertical line patterns:

$$LAM = \frac{\sum_{v=v_{\min}}^{N_R} vP(v)}{\sum_{v=1}^{N_R} vP(v)}$$
(14)

where P(v) denotes a histogram of vertical lines of the length v with the minimum line length $v_{\min} = 2$. The larger the laminarity parameter, the more stable the behavior of the system. Finally, the average vertical line length (TT) is given as $TT = \sum_{v=v_{\min}}^{N_R} vp(v)$, and it estimates the mean time that the system remains at a specific state.

The estimated results of RQA measures are presented in Table 2. It is found that for each of these three simulated data and two actual stock indexes, the recurrence rate (RR) for recurrence threshold h = 0.04 is larger than that for h = 0.02; especially, the value for HSI is observed as a higher increase than those for other series. The DET and L_{mean} values decrease when the h changes from 0.02 to 0.04, except the DET value for HSI, which becomes 0.7793 for h = 0.04 from 0.7685 for h = 0.02. Since the higher DET and L_{mean} correspond to a more predictable and stable system, thus the results imply the reduction of determinism of these return series in the case of h = 0.04. It is also noticed that the L_{mean} for HSI under h = 0.02 is much smaller than those for other series, indicating the weakest predictability among them. Furthermore, it is observable that all of the values of L_{ENT} have increased for h = 0.04, and its value for the return series with R = 1 has a relatively considerable rise, which indicate the increasing of

the complexity of these time series. In terms of measures of vertical recurrence lines, the values of LAM and TT enlarge when h becomes 0.04, which means a rise in the fraction of recurrence points forming vertical lines, and it can also obviously be seen in Figure 7.

	RR		DET		L_{mean}			
Data	(h = 0.02, 0.04)		(h = 0.	(h = 0.02, 0.04)		(h = 0.02, 0.04)		
R = 1	0.0005	0.0027	0.9044	0.6649	46.7234	3.6071		
R=2	0.0007	0.0046	0.8373	0.6817	15.6688	3.2662		
R = 3	0.0005	0.0018	0.9336	0.6779	51.0233	4.5073		
HSI	0.0010	0.0163	0.7685	0.7793	7.2597	3.5210		
SSE	0.0006	0.0021	0.8503	0.5994	20.4435	3.9045		
	L_{ENT}		LA	LAM		TT		
Data	(h = 0.02, 0.04)		(h = 0.	(h = 0.02, 0.04)		(h = 0.02, 0.04)		
R = 1	0.3921	1.1084	0.0049	0.0671	2.0000	2.3217		
R=2	0.7699	1.2140	0.0095	0.0989	2.3333	2.4383		
R = 3	0.7629	1.1562	0.0000	0.0449	0.0000	2.6015		
HSI	1.1184	1.6163	0.0072	0.2814	2.0000	2.8932		
SSE	0.4499	0.8834	0.1154	0.1334	2.3116	2.6532		

Table 2. Recurrence quantification analysis (RQA) measures of simulative data and actual data.

4. Conclusions

In the present paper, we have developed a financial price model by using the mechanisms of the interacting dynamic finite-range voter system, which is a well-known statistical physics model. The comparative research of statistical properties of two Chinese stock indexes and the simulated data with different parameters (intensity and neighbor range parameters) is performed. The composite multiscale entropy analysis is applied to investigate the complexity of returns and absolute returns with different power exponents. The findings show that the CMSE for return series decreases as the time scale becomes larger, while the CMSE values for volatility series first reduce gradually, but finally remain almost stable, which is consistent with the auto-correlations of the series. The obviously different complexity degrees of $|r(t)|^q$ for different values of q are also observable. Finally, the recurrence plots and recurrence thresholds h = 0.02 and h = 0.04, the RQA measures present apparently different values, indicating different complex determinism behaviors. Based on the statistical research, we can observe that there is some evidence of similar complex behaviors of the returns of the financial model derived from the finite-range voter system to the real stock markets; this shows that the proposed model can grasp some nature of the real stock market in certain respects and is reasonable for real stock price modeling.

Acknowledgments

The authors were supported in part by the National Natural Science Foundation of China, Grant No. 71271026 and Grant No. 10971010.

Author Contributions

All authors contributed to modelling, analyzing the results, and making the preparation of this paper. All authors have read and approved the final manuscript.

Conflicts of Interest

The authors declare no conflict of interest.

References

- 1. Calvet, L.E.; Fisher, A.J. *Multifractal Volatility: Theory, Forecasting, and Pricing*; Academic Press: Waltham, MA, USA, 2008.
- 2. Chakraborti, A.; Toke, I.M.; Patriarca, M.; Aberget, F. Econophysics review: I. Empirical facts. *Quant. Financ.* **2011**, *11*, 991–1012.
- 3. Gabaix, X.; Gopicrishnan, P.; Plerou, V.; Stanley, H.E. Understanding the cubic and half-cubic laws of financial fluctuations. *Physica A* **2003**, *324*, 1–5.
- 4. Guo, Y.L.; Wang, J. Simulation and statistical analysis of market return fluctuation by Zipf method. *Math. Probl. Eng.* **2011**, *2011*, 253523.
- Kantz, H.; Schreiber, T. *Nonlinear Time Series Analysis*; Cambridge University Press: Cambridge, MA, USA, 1997.
- 6. Liao, Z.; Wang, J. Forecasting model of global stock index by stochastic time effective neural network. *Expert Syst. Appl.* **2010**, *37*, 834–841.
- 7. Lux, T. *Financial Power Laws: Empirical Evidence, Models and Mechanisms*; Cambridge University Press: Cambridge, UK, 2008.
- 8. Mandelbrot, B.B. *Fractals and Scaling in Finance: Discontinuity, Concentration, Risks*; Springer: Berlin/Heidelberg, Germany and New York, NY, USA, 1997.
- 9. Stanley, H.E.; Gabaix, X.; Gopikrishnan, P.; Plerou, V. Economic fluctuations and statistical physics: The puzzle of large fluctuations. *Nonlinear Dyn.* **2006**, *44*, 329–340.
- Cheng, W.Y.; Wang, J. Dependence phenomenon analysis of the stock market. *Europhys. Lett.* 2013, 102, doi:10.1209/0295-5075/102/18004.
- Cont, R.; Bouchaud, J.P.; Herd behavior and aggregate fluctuation in financial markets. *Macroecon. Dyn.* 2000, *4*, 170–196.
- 12. Fang, W.; Wang, J. Statistical properties and multifractal behaviors of market returns by Ising dynamic systems. *Int. J. Mod. Phys. C* **2012**, *23*, doi:10.1142/S0129183112500234.
- 13. Krause, S.M.; Bottcher, P.; Bornholdt, S. Mean-field-like behavior of the generalized voter-model-class kinetic Ising model. *Phys. Rev. E* **2012**, *85*, doi:10.1103/PhysRevE.85.031126.
- 14. Lux, T.; Marchesi, M. Scaling and criticality in a stochastic multi-agent model of a financial market. *Nature* **1999**, *397*, 498–500.
- 15. Ross, S.M. *An Introduction to Mathematical Finance*; Cambridge University Press: Cambridge, UK, 1999.
- 16. Samanidou, E.; Zschischang, E.; Stauffer, D.; Lux, T. Agent-based models of financial markets. *Rep. Prog. Phys.* **2007**, *70*, 409–450.

- fluctuation. *Physica A* 1998, 256, 284–290.
 18. Mills, T.C. *The Econometric Modeling of Financial Time Series*; Cambridge University Press: Cambridge, UK, 1999.
- 19. Niu, H.L.; Wang, J. Volatility clustering and long memory of financial time series and financial price model. *Digit. Signal Process.* **2013**, *23*, 489–498.
- 20. Wang, F.; Wang, J. Statistical analysis and forecasting of return interval for SSE and model by lattice percolation system and neural network. *Comput. Ind. Eng.* **2012**, *62*, 198–205.
- 21. Wang, J.; Wang, Q.Y.; Shao, J.G. Fluctuations of stock price model by statistical physics systems. *Math. Comput. Model.* **2010**, *51*, 431–440.
- 22. Wang, T.S.; Wang, J.; Zhang, J.H.; Fang, W. Voter interacting systems applied to Chinese stock markets. *Math. Comput. Simul.* **2011**, *81*, 2492–2506.
- 23. Xiao, D.; Wang, J. Modeling stock price dynamics by continuum percolation system and relevant complex systems analysis. *Physica A* **2012**, *291*, 4827–4838.
- 24. Yu, Y.; Wang, J. Lattice-oriented percolation system applied to volatility behavior of stock market. *J. Appl. Stat.* **2012**, *39*, 785–797.
- 25. Zhang, J.H.; Wang, J. Modeling and simulation of the market fluctuations by the finite range contact systems. *Simul. Model. Pract. Theory* **2010**, *18*, 910–925.
- 26. Zhang, J.H.; Wang, J.; Shao, J.G. Finite-range contact process on the market return intervals distributions. *Adv. Complex Syst.* **2010**, *13*, 643–657.
- 27. Zivot, E.; Wang, J.H. *Modeling Financial Time Series with S-PLUS*; Springer: New York, NY, USA, 2006.
- 28. Staufffer, D.; Sornette, D. Self-organized percolation models for stock market fluctuations. *Physica A* **1999**, *271*, 496–506.
- 29. Chen, M.F. From Markov Chains to Non-Equilibrium Particle Systems; World Scientific: Singapore, Singapore, 1992.
- Durrett, D. Lecture Notes on Particle Systems and Percolation; Wadsworth & Brooks: California, CA, USA, 1998.
- 31. Liggett, T.M. Interacting Particle Systems; Springer: New York, NY, USA, 1985.
- 32. Liggett, T.M. *Stochastic Interacting Systems: Contact, Voter and Exclusion Processes*; Springer: Heidelberg, Germany, 1999.
- Costa, M.; Goldberger, A.L.; Peng, C.K. Mutltiscale entropy analysis of biological signals. *Phys. Rev. E* 2005, 71, doi:10.1103/PhysRevE.71.021906.
- 34. Cirugeda-Roldan, E.; Cuesta-Frau, D.; Miro-Martinez, P.; Oltra-Crespo, S. Comparative Study of Entropy Sensitivity to Missing Biosignal Data. *Entropy* **2014**, *16*, 5901–5918.
- 35. Martina, E.; Rodriguez, E.; Perez, R.E.; Ramirez, J.A. Multiscale entropy analysis of crude oil price dynamics. *Energy Econ.* **2011**, *33*, 936–947.
- 36. Vargas, L.G.; Rojas, A.R.; Brown, F.A. Multiscale entropy analysis of electroseismic time series. *Nat. Hazards Earth Syst. Sci.* **1008**, *8*, 855–860.
- 37. Wu, S.D.; Wu, C.W.; Liu, S.G.; Wang, C.C.; Lee, K.Y. Time series analysis using composite multiscale entropy. *Entropy* **2013**, *15*, 1069–1084.

- 38. Hsieh, D.A. Chaos and nonlinear dynamics: Application to financial markets. *J. Financ.* **1991**, *46*, 1839–1877.
- 39. Frauch, J.B.; Contreras, D.; Lleso, L.T. Assessing nonlinear structures in real exchange rates using recurrence plot strategies. *Physica D* **2002**, *171*, 249–264.
- 40. Guhathakurta, K.; Bhattacharya, B.; Chowdhury, A.R. Using recurrence plot analysis to distinguish between endogenous and exogenous stock market crashes. *Physica A* **2010**, *389*, 1984–1882.
- 41. Aparicio, T.; Pozo, E.F.; Saura, D. Detecting determinism using recurrence quantification analysis: Three test procedures. *J. Econ. Behav. Organ.* **2008**, *65*, 768–787.
- 42. Marwan, N.; Romano, M.C.; Thiel, M.; Kurths, J. Recurrence Plots for the Analysis of Complex Systems. *Phys. Rep.* **2007**, *438*, 237–329.
- 43. Rolski, T.; Schmidt, V.; Schmidli, H.; Teugels, J. *Stochastic Processes for Insurance and Finance*; Wiley: New York, NY, USA, 1999.
- 44. Lamberton, D.; Lapeyre, B. *Introduction to Stochastic Calculus Applied to Finance*; Chapman and Hall/CRC: London, UK, 2000.
- 45. Gaylord, R.; Wellin, P. Computer Simulations with Mathematica: Explorations in the Physical, Biological and Social Science; Springer: New York, NY, USA, 1995.
- 46. Zheng, Z. *Matlab Programming and the Applications*; China Railway Publishing House: Beijing, China, 2003.
- Dionisio, A.; Menezes, R.; Mendes, D.A. Entropy-Based Independence Test. *Nonlinear Dyn.* 2006, 44, 351–357.
- 48. Frazer, A.M.; Swinney, H.L. Independent coordinates for strange attractors from mutual information. *Phys. Rev. A* **1986**, *33*, 1134–1140.
- 49. Ilinski, K. *Physics of Finance: Gauge Modeling in Non-equilibrium Pricing*; Wiley: New York, NY, USA, 2001.

© 2015 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/4.0/).