International Journal of Economics and Finance Vol. 2, No. 1
February 2010
www.ccsenet.org/ijef

Nonlinear Noise Estimation in International Stock Markets: Coarse-grained Entropy Method

Yong Fang
Post-doctoral Research Center for Applied Economics
Shanghai University of Finance and Economics, Shanghai 200433, China
Department of Applied Mathematics
Shanghai Finance University, Shanghai 201209, China
E-mail: yongf72@163.com

The research is financed by China Postdoctoral Science Foundation (Grant No. 20090450075), Shanghai Municipal Natural Science Foundation (Grant No. 09ZR1421900) and Innovation Program of Shanghai Municipal Education Commission (Grant No. 10YZ184). (Sponsoring information)

Abstract

With the step-by-step opening of China Stock Market and gradual strengthening of international linkage, how to efficiently measure and manage risk, evaluate and improve market operation efficiency is an important project in present financial research. According to nonlinear dynamics and chaos and fractal theory, we apply phase space reconstruction technique and coarse-grained entropy method to estimate the nonlinear noise levels in stock markets of Chinese Mainland, Hong Kong, US, UK and Japan, and we emphasize on discussing the standard deviation of nonlinear noise σ and noise-to-signal ratio NSR which are two important indexes about risk measurement and efficiency evaluation, and further we make a comprehensive comparison analysis on the risk and operation efficiency of stock markets of above countries or areas.

Keywords: Coarse-grained entropy, International stock markets, Market efficiency, Noise estimation, Nonlinear time series, Phase space reconstruction

1. Introduction

Traditional financial models are linear equilibrium models based on rational expectation. Excluding the influence of psychological factors such as greed and fear, these models depict the negative feedback leading to the stabilization of financial system and indicate that the response to exogenous disturbance of financial system is reverting to the equilibrium in a continuous and linear way.

Nevertheless, what traditional financial models depict is only an ideal state and is widely divergent from the truth. Subjected to noise trading, positive feedback trading, overreaction and herding behavior which are universal in financial market (see De Long, Shleifer, Summers, and Waldmann (1990a, 1990b), De Bondt and Thaler (1985), Banerjee (1992)), the fluctuation of asset price presents strong nonlinearity. Moreover, the evolution of financial market is driven by numerous heterogeneous investors with bounded rationality, so financial market is a complex dynamic system which is essentially characterized by intrinsic randomness of deterministic system, nonrepetitive aperiodic cycle, equilibrium which is far from balance, sensitivity to initial conditions, tendency, self-similarity and mutation. Therefore, nonlinear dynamics and chaos and fractal theory which have developed rapidly in recent years provide a new perspective for us to study complex financial system.

Among numerous financial research projects, estimating the nonlinear noise level of complex financial dynamic system is of great practical significance. It can provide financial supervision and management departments with efficient decision supports such as evaluating the operation efficiency of financial market, guarding against and managing financial risk and exporting so-called policy function developed by Boldrin and Montrucchio (1986). Nonlinear noise of complex financial system has two kinds of sources, and one kind is measurement noise from outside of system while another kind is dynamical noise from inside of system. Some scholars have put forward a series of methods on estimating nonlinear noise level, among which the methods developed by Cawley and Hsu (1992), Schreiber (1993), Diks (1996) and Oltmans and Verheijen (1997) only apply to the estimation of low level measurement noise while the coarse-grained entropy method developed by Urbanowicz and Holyst (2003) can efficiently estimate high level measurement noise as well as dynamical noise.

In recent years, many scholars have been drawn to the nonlinearity of stock markets. Longbing Xu and Rong Lu (1999) studied the nonlinearity of Chinese stock market by use of R/S method, and their empirical results indicated that both

Shanghai and Shenzhen stock market were characterized by nonlinearity, long memory and volatility clustering. By use of R/S analysis and the calculation of correlation dimension, auto correlation function and Lyapunov index, Haihua Wu and Daoye Li (2001) found that Shanghai stock market was distinctly characterized by fractal and chaos, and was a nonlinear system where strange attractors existed. By implementing a series of tests such as normality test, ADF unit root test, BDS test and ARCH test and calculating correlation dimension and Lyapunov index, Xusong Xu and Yanbin Chen (2001) found that both nonlinearity and chaos existed in Chinese stock market.

But existing literatures are in large limited to testing the nonlinearity and chaos in stock markets, and few literatures focus on how to measure nonlinear risk, how to evaluate the efficiency of a stock market, how to form an efficient investment strategy and how to manage systematic risk. These problems are worthy of deeper research. In this paper, we will make an exploration on studying these problems. We apply phase space reconstruction technique and coarse-grained entropy method to estimate the nonlinear noise levels in stock markets of Chinese Mainland, Hong Kong, US, UK and Japan, and further a comprehensive comparison analysis on risk and operation efficiency of above stock markets is made.

2. Sample data and descriptive statistics

We choose the daily return time series of SSE Composite Index (Chinese Mainland), Hang Seng Index (Hong Kong), S&P 500 (US), FTSE 100 (UK) and Nikkei 225 (Japan) as our research objects. The period studied is from December 20, 1990 to June 21, 2007 and the sample data is from RESSET (http://www.resset.cn) and Yahoo Finance (http://finance.yahoo.com).

To be exempt from the influence of outliers on statistical results, we take values which distances from the mean of time series exceed four times standard deviation as outliers, and then get rid of them from the time series. The descriptive statistics of daily return time series of various stock indexes after getting rid of outliers are given in table 1.

As shown from table 1 and line graph and histogram of time series, the above five daily return time series of stock indexes all significantly deviate from normal distribution and are all characterized by leptokurtic, fat tail and volatility clustering.

3. Preliminary tests

3.1 Nonlinearity tests

We apply BDS test to detect the nonlinearity of various time series. At first, linear correlation in the time series is filtered by the use of ARMA model, and then BDS test is conducted for residual series. BDS statistics with various thresholds r and embedded dimensions m are given in table 2 (σ_ε denotes the standard deviation of residual series). As shown from table 2, the null hypothesis that residuals are i.i.d. is significantly rejected (the critical value of normal distribution is 1.96 under the significance level of 5%), and it indicates that the stock markets of above five countries or areas are all of significant nonlinearity.

3.2 Determinism tests

So-called determinism of system is that the future states can be determined by the past states. The system with determinism usually presents some degree of tendency and self-similarity. We apply the method of recurrence plot developed by Eckmann, Kamphorst and Ruelle (1987) to test the determinism of various time series. At first, phase space reconstruction is conducted by the use of nonlinear dynamics, and the time delays τ and embedded dimensions m (the maximum embedded dimension is 15) of various time series which help to reconstruct phase space are given in table 3, and they are determined by average mutual information method and false nearest neighborhood method.

After phase space reconstruction, recurrence plots of various time series can be made (thresholds are the standard deviations of various time series), and recurrence quantitative analysis (RQA) can be further conducted. By observing the recurrence plots of above five time series, we find that there scatter some small bands parallel with the diagonal, which indicates obvious determinism. Results of RQA on original data and on the new data after randomly disturbing the original sequence of various time series are respectively given in table 4. As contrasted to original data, the determinism of the new data after randomly disturbing the original sequence significantly declined and this indicates that our determinism tests are robust.

3.3 Chaos tests

According to G-P algorithm (see Grassberger and Procaccia (1983)), double logarithmic scatter plots of correlated integral $C(r)$ with respect to threshold r with various embedded dimensions (as shown in figure 1) are made and we can find the following things. On one hand, with the gradual increase of embedded dimension m, the slope of linear part in $\ln C(r) - \ln r$ graph (correlated dimension) gradually goes to stabilization, which indicates that there exists strange attractor in the system. On the other hand, before arriving at stabilization correlated dimension presents obvious jump (SSE Composite Index acts significantly). From these findings, we may believe that various time series are all characterized by chaos in some degree.

4. Nonlinear noise estimation

4.1 coarse-grained entropy method

Let $\{x_i\}, i = 1,2,\cdots,N$ be a nonlinear time series. After selecting suitable time delay τ and embedded dimension m we reconstruct phase space, and then get m dimension vector sequence

$$\overrightarrow{y_i} = (x_i, x_{i+\tau}, \cdots, x_{i+(m-1)\tau}), i = 1,2,\cdots,M \tag{1}$$

where $M = N - (m-1)\tau$. The correlated integral of embedded phase space is defined as the follows

$$C_m(r) = \frac{2}{M(M-1)} \sum_{1 \le i < j \le M} H(r - \left\| \overrightarrow{y_i} - \overrightarrow{y_j} \right\|) \tag{2}$$

where $H(\cdot)$ is Heaviside function, and r is threshold and $\left\| \cdot \right\|$ is maximum norm. Therefore, coarse-grained entropy can be estimated as follows

$$K_2(r) \approx -\frac{d \ln[C_m(r)]}{dm} \tag{3}$$

Let σ be the standard deviation of the noise (including measurement noise and dynamical noise) of nonlinear time series $\{x_i\}, i = 1,2,\cdots,N$, and the observation values of coarse-grained entropy $K_{noisy}(r)$ can be fitted by the following formula

$$K_{noisy}(r) = -\frac{d \ln[C_m(r)]}{dm}$$

$$= -cg(\frac{r}{2\sigma}) \ln r + [\kappa + b \ln(1 - ar)] \times (1 + \sqrt{\pi} \frac{\sqrt{r^2/3 + 2\sigma^2} - r/\sqrt{3}}{r}) \tag{4}$$

where κ, a, b, c, σ are parameters under estimation, and the function $g(z)$ is defined as follows

$$g(z) = \frac{2}{\sqrt{\pi}} \frac{ze^{-z^2}}{erf(z)} \tag{5}$$

where $erf(\cdot)$ is error function.

In formula (4), $g(\cdot)$ shows the influence of noise on correlated dimension, and $\kappa + b \ln(1 - ar)$ is coarse-grained entropy of signal uncontaminated by noise while $[\kappa + b \ln(1 - ar)] \times (\sqrt{\pi} \frac{\sqrt{r^2/3 + 2\sigma^2} - r/\sqrt{3}}{r})$ is increment of coarse-grained entropy caused by noise. After estimating the standard deviation of noise σ by the use of fitting formula (4), finally we can calculate noise-to-signal ratio (NSR) of time series as follows

$$NSR = \frac{\sigma}{\sigma_{data}} \times 100\% \tag{6}$$

and we take it as a measurement of noise level of time series.

4.2 Estimation results

Taking the essentiality that correlated integral $C_m(r)$ varies with embedded dimension m and threshold r into account, we choose m from 1 to 20, and r takes 100 different values in interval $[\frac{\sigma_{data}}{2}, 2\sigma_{data}]$. Observation value of coarse-grained entropy $K_{noisy}(r)$ can be estimated by the use of OLS method for the following linear regression

$$\ln C_m(r) = k \cdot m + e \tag{7}$$

where the negative of slope k is just $K_{noisy}(r)$.

Then we apply Levenberg-Marquardt method and general global optimization method to fit $K_{noisy}(r)$ according to (4).

The estimation value of σ, NSR, decision coefficient of fitting R^2 and the root of mean square errors (RMSE) are given in table 5. The statistics in table 5 and figure 2 all sufficiently indicate that the fitting makes great effect.

4.3 Results analysis

Traditional financial models are all based on the hypothesis that stock return is normally distributed (i.e. acts as stationary distribution with characteristic parameter $\alpha = 2$, and its variance is limited and stationary), in which variance of return is applied to measure risk. A large amount of empirical research indicates, however, stock return significantly deviate from normal distribution and is characterized by leptokurtic, fat tail and state continuing. The above practical characteristics of stock return can be well depicted by stationary distribution with characteristic parameter $\alpha \in (1,2)$. As population variance in this case is uncertain or infinity, taking sample variance as the measurement of risk is of no significance.

Nonlinear dynamics and chaos and fractal theory which have rapidly developed in recent years provide a new perspective for us to study modern financial risk management and portfolio selection. Complex financial dynamic system is essentially characterized by intrinsic randomness of deterministic system and the fluctuation of the deterministic part is completely predicted in a short period (so we may believe this part of fluctuation to be of no risk in a short period), so as the essential financial risk measurement, the natural choice is the standard deviation of nonlinear noise of financial asset return time series σ. Not only can we apply σ to guard against and control financial risk, but also we can use it to optimize portfolio selection (see Urbanowicz and Holyst (2004)).

By observing the estimation values of σ, we can find that the risks of eastern stock markets including Chinese Mainland, Hong Kong and Japan are higher than that of western stock markets including US and UK. The main reason is that western stock markets are more mature than eastern markets and the later are influenced to a large extent by the free input or output of large amount of capital and regular changes in trading rules. Moreover, we can find that the risks of Hong Kong and Japan stock markets are higher than that of Chinese Mainland stock market. Strong "policy market" characteristic, limitation of margin of rise or fall, relatively weak international linkage and small size of QFII may account for the above finding. As we know, the trial of direct investment in Hong Kong stock market for individual investors of Chinese Mainland will be on the way, and in spite of the great significance of this policy, how to efficiently guard against risk is

the primary challenge faced by investors.

Besides σ discussed above, NSR is another important index and it indicates the extent of operation efficiency of stock market. Higher NSR is (i.e. the proportion of intrinsic random fluctuation in the whole fluctuation is higher), market is more mature and operates more efficiently and is closer to EMH. By observing the values of NSR, we can find that the other four stock markets are more and more mature than the stock market of Chinese Mainland. The statistical results which respectively indicate nonnormality, leptokurtic, fat tail, nonlinearity, determinism and mutation also significantly support above finding. This arises from two aspects. On one hand, efficient securities legislation system, credit system and multi-dimensional supervision and management system are not completely established, and some illegal behaviors such as inside trading, manipulating price and issuing false information usually occur. On the other hand, a rational investment culture which core idea is "value investment in the long run" and a healthy and harmonious ecological environment in securities market are not formed and the size of institutional investors is relatively small, so there exist strong effects of "positive feedback" and "herding behavior", and with the accumulation of these effects in a long period financial bubble will be prone to arise, and more heavily financial crisis will probably break out.

5. Conclusion

With the step-by-step opening of China Stock Market and gradual strengthening of international linkage, how to efficiently measure and manage risk, evaluate and improve market operation efficiency is an important project in present financial research. According to nonlinear dynamics and chaos and fractal theory, we apply phase space reconstruction technique and coarse-grained entropy method to estimate the nonlinear noise levels in stock markets of Chinese Mainland, Hong Kong, US, UK and Japan, and we emphasize on discussing the standard deviation of nonlinear noise σ and noise-to-signal ratio NSR which are two important indexes about risk measurement and efficiency evaluation, and further we make a comprehensive comparison analysis on the risk and operation efficiency of stock markets of above countries or areas. The following research work will be applying these two indexes to the simulation of stock market based on agent to study the influence of microindividuals on the macromarket.

References

Banerjee, A. (1992). A simple model of herd behavior. *Quarterly Journal of Economics*, 107, 797-817.

Boldrin, M., &Montrucchio, L. (1986). On the indeterminism of capital accumulation paths. *Journal of Economic Theory*, 40, 26-39.

Cawley, R. & Hsu, G. H. (1992). Local-geometric projection method for noise reduction in chaotic maps and flows. *Physical Review A*, 46, 3057-3082.

De Bondt, W. F. M. &Thaler, R. H. (1985). Does the stock market overreact?. *Journal of Finance*, 40,793-805.

De Long, J. B., Shleifer, A., Summers, L. H. &Waldmann, R. J. (1990a). Noise trader risk in financial markets. *The Journal of Political Economy*, 98, 703-738.

De Long, J. B., Shleifer, A., Summers, L. H. &Waldmann, R. J. (1990b). Positive feedback investment strategies and destabilizing rational speculation. *Journal of Finance*, 45, 379-395.

Diks, C. (1996). Estimating invariants of noisy attractors. *Physical Review E*, 53, R4263-R4266.

Eckmann, J. P., Kamphorst, S. O. & Ruelle, D. (1983). Recurrence plots of dynamical systems. *Euro-physics Letters*, 4, 973-977.

Grassberger, P. & Procaccia, I. (1983).Characterization of strange attractors. *Physical Review Letters,* 50, 346-349.

Haihua Wu & Daoye Li. (2001). Nonlinear dynamics analysis on stock market-in case of Shanghai stock market. *Journal of Qingdao University*,16,1-7.

Longbing Xu & Rong Lu. (1999) Exploring the nonlinearity of Chinese stock market based on R/S analysis. *Forecasting*,18 ,59-62.

Oltmans, H. & Verheijen, P. J. T. (1997). The influence of noise on power law scaling functions and an algorithm for dimension estimations. *Physical Review E*, 56, 1160-1170.

Schreiber, T. (1993). Determination of the noise level of a chaotic time series. *Physical Review E*, 48, R13-R16.

Urbanowicz, K. & Holyst, J. A. (2003) .Noise-level estimation of time series using coarse-grained entropy. *Physical Review E*, 67, 046218.

Urbanowicz, K. & Holyst, J. A. (2004). Investment strategy due to the minimization of portfolio noise level by observations of coarse-grained entropy. *Physica A*, 344, 284-288.

Xusong Xu. & Yanbin Chen. (2001). An empirical study on the nonlinearity of Shanghai and Shenzhen stock market. *The Journal of Quantitative & Technical Economics*,18,110-113.

Table 1. Descriptive statistics of daily return time series of various stock indexes

	SSE Composite Index	Hang Seng Index	S&P 500	FTSE 100	Nikkei 225
Number of observations	4027	4065	4142	4148	4062
Mean	0.000489	0.000572	0.000364	0.000294	0.000023
Standard deviation	0.020895	0.014261	0.009335	0.009556	0.014206
Maximum	0.1191	0.062	0.0393	0.0406	0.0796
Minimum	-0.1118	-0.0607	-0.0391	-0.0397	-0.0698
Skewness	0.01118	-0.037123	-0.026864	-0.100655	0.147949
Kurtosis	8.742128	4.892077	4.788616	4.540808	5.249315
Jarque-Bera	5532.517	607.2887	552.6178	417.3261	871.1254
P-value	0.000000	0.000000	0.000000	0.000000	0.000000

Table 2. BDS statistics of daily return time series of various stock indexes

		SSE Composite Index	Hang Seng Index	S&P 500	FTSE 100	Nikkei 225
$r = 0.5\sigma_\varepsilon$	$m = 2$	18.62	8.68	7.73	9.66	5.21
	$m = 3$	25.56	10.79	12.30	13.14	8.43
	$m = 4$	34.07	13.63	15.81	17.25	10.89
	$m = 5$	44.73	16.69	20.15	21.00	14.17
$r = \sigma_\varepsilon$	$m = 2$	20.86	9.55	7.80	11.09	5.84
	$m = 3$	25.92	12.70	12.99	14.93	9.11
	$m = 4$	30.42	15.85	16.60	18.88	11.54
	$m = 5$	34.38	18.47	20.63	22.25	14.13
$r = 1.5\sigma_\varepsilon$	$m = 2$	22.36	10.17	8.49	12.59	6.67
	$m = 3$	26.47	13.95	13.57	16.67	9.99
	$m = 4$	29.40	17.03	16.82	20.27	12.19
	$m = 5$	31.40	19.38	20.13	23.16	14.12

Table 3. Time delays and embedded dimensions of daily return time series of various stock indexes which help to reconstruct phase space

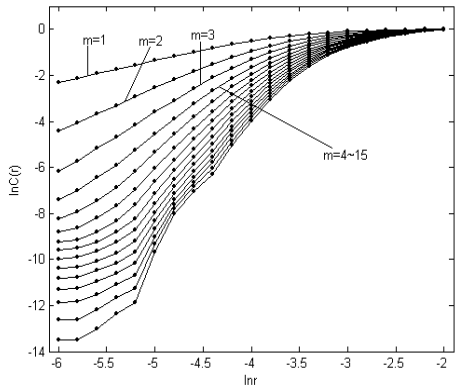
	SSE Composite Index	Hang Seng Index	S&P 500	FTSE 100	Nikkei 225
τ	2	1	1	2	1
m	9	11	9	12	13

Table 4. RQA results of daily return time series of various stock indexes

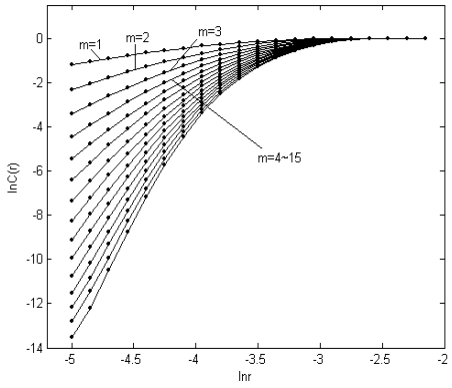
		SSE Composite Index	Hang Seng Index	S&P 500	FTSE 100	Nikkei 225
Original data	%recurrence	8.62%	1.06%	2.37%	0.63%	0.41%
	%determinism	37.27%	92.55%	92.59%	3.83%	93.66%
	Longest diagonal line segment	271	45	55	23	51
New data after randomly disturbing the original sequence	%recurrence	1.55%	0.22%	0.61%	0.11%	0.06%
	%determinism	1.64%	82.58%	81.85%	0.10%	80.95%
	Longest diagonal line segment	13	19	16	4	12

Table 5. Nonlinear noise estimation of daily return time series of various stock indexes

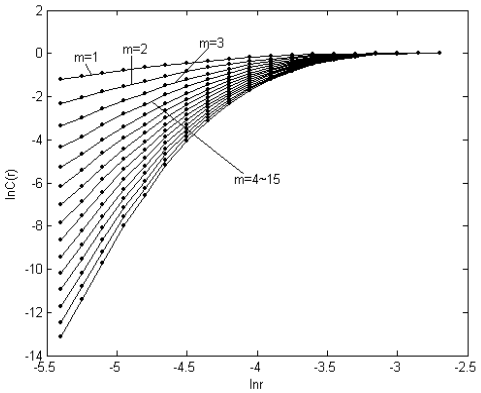
	SSE Composite Index	Hang Seng Index	S&P 500	FTSE 100	Nikkei 225
σ	0.0079	0.0101	0.0069	0.0066	0.0090
NSR	37.8%	70.8%	73.9%	69.1%	63.4%
R^2	0.9931	0.9994	0.9988	0.9994	0.9986
RMSE	0.00699	0.00437	0.00645	0.00555	0.00594



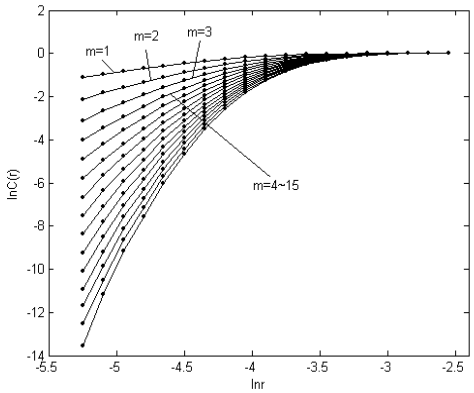
(a) SSE Composite Index



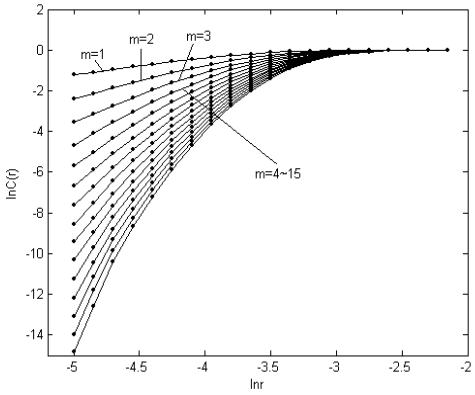
(b) Hang Seng Index



(c) S&P 500

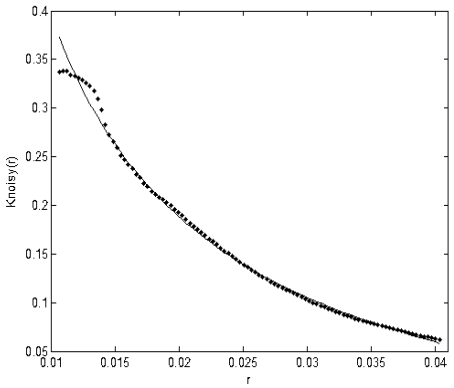


(d) FTSE 100

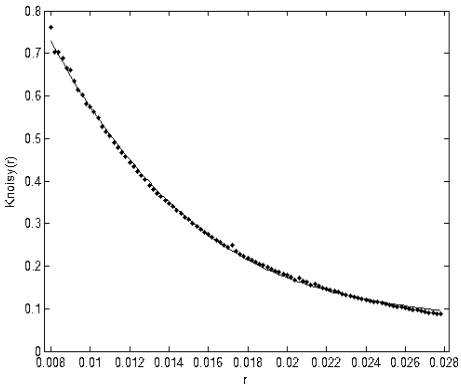


(e) Nikkei 225

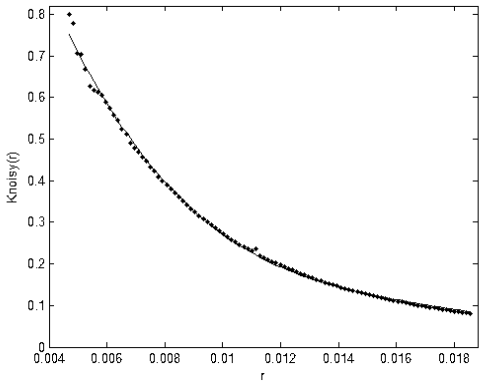
Figure 1. $\ln C(r) - \ln r$ graphs of daily return time series of various stock indexes



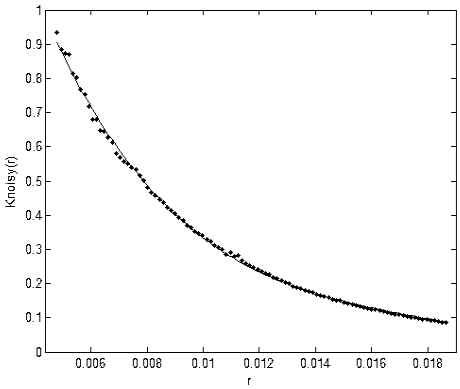
(a)SSE Composite Index



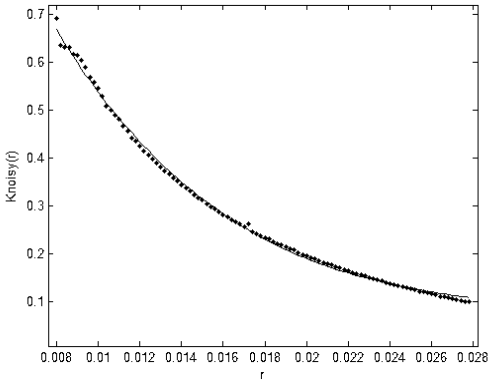
(b) Hang Seng Index



(c) S&P 500



(d) FTSE



(e) Nikkei 225

Figure 2. Fitting plots of coarse-grained entropy of daily time series of various stock indexes