Front. Phys. China, 2008, 3(4): 489–494
DOI 10.1007/s11467-008-0038-7

RESEARCH ARTICLE

Wei WAN, Jian-Wei ZHANG

Long-term memory of the returns in the Chinese stock indices

Abstract The modified R/S statistic (MRS) and the local Whittle method (LWM) are used to analyze the long-range dependence on various indices of the Chinese stock markets. The MRS accepts the null hypothesis of no long-range dependence while the LWM rejects it. We also find that the long-range dependence phenomena presented in these markets depend on the time in which they are measured.

Keywords long-range dependence, modified R/S statistic, local Whittle estimator

PACS numbers 89.65.Gh, 05.45.Tp, 02.50.-r

1 Introduction

Long-range dependence, also commonly referred to as long-term memory or correlation, has a long history and remains an active topic of research in the study of financial time series. A lot of empirical studies have been done on the long-range dependence of stock markets. In 1970s Mandelbrot [1−3] introduced the rescaled range statistic (R/S) method, which was originally invented by Hurst [4], to evaluate the Hurst exponent H in analyzing the long-term memory of security returns, and found that it was quite larger than 1/2 for stock prices and returns of interest rates. It displays a strong positive long-range correlation, and challenges the efficient market hypothesis and the related random walk model. A clear description of these issues can be found in Ref. [5]. Similar results were also obtained by Greene and Fielitz [6]. As we know, many other financial variables (such as volatility, trade sign, *et al.*) possess long-term memory, which are less debated. However, there is a tremendous discord between affirmation and denial of the long-term memory for security returns. In 1988 Fama and French [7] argued that the long-range dependence of stock returns is weak for weekly and daily data, and returns for a longer time have a negative correlation. Aydogan and Booth [8] believed that no long-term dependence exists in stock returns. In 1991 Lo [9] declared that the classical R/S analysis is sensitive to the presence of short-range dependencies. Evidence of dependence using the classical R/S statistic may come merely from short-range dependencies instead of long-range ones. He proposed a modified R/S (MRS) statistic to determine the long-term memory. Nevertheless, some economists said that an acceptance of the null hypothesis of no long-range dependence based on the MRS statistic should never be viewed as the "final word". Teverovsky, Taqqu and Willinger [10] identified a number of problems associated with Lo's method and its use in practice. They verified that Lo's MRS statistic has a strong preference for accepting the null hypothesis of no long-range dependence.

Another widely used methodology, called the Whittle estimator, evaluates the long-term memory parameter through the log-periodogram regression estimator proposed by Geweke and Porter-Hudak [11]. Unlike the MRS, this method is frequency-domain based and applies maximum likelihood techniques. However, it has been criticized because of its finite-sample bias. Robinson [12] proposed a least biased semi-parametric estima-

Wei WAN[1], Jian-Wei ZHANG[1, 2, 3] (✉)

[1] School of Physics, Peking University, Beijing 100871, China
[2] Econophysics Laboratory, Normal College, Shihezi University, Xinjiang 832003, China
[3] Institut für Theoretische Physik, Freie Berlin Universität, Arnimalle 14, 14195 Berlin, Germany
E-mail: james@pku.edu.cn

Received April 24, 2008; accepted May 20, 2008

tor, H, called the local Whittle method (LWM), which presents robustness to data seasonality and short-range dependence [13].

Although disputes on the presence of long-range dependence of returns of stocks are still on-going, empirical studies have found that some stock markets have long-range dependence [14−17]. Chinese stock markets are emerging markets, and a few papers have studied their distributions and specifications of returns of stocks and indices [18−23]. Specifically, Gu and Zhou [24] found evidence of long-term memory in the ensemble daily price returns within several time periods of 500 stocks traded in the Shanghai Stock Exchange by using the detrended fluctuation analysis, the R/S analysis, and the MRS analysis. Moreover, there are quite a few other papers [25−27] which have applied the generalized spectral derivative method [28] to the Chinese stock markets to demonstrate that Chinese stock markets are far from being weak in terms of market efficiency. The main focus of this paper is to calculate the Hurst exponent and its variance in returns of various indices of the Chinese stock markets by using the LWM, and to analyze long-range dependence by using the MRS statistic.

The remainder of this paper is organized as follows. In Section 2, the Local Whittle method and the modified R/S statistic are reviewed. In Section 3, the calculations are described and the results are discussed. Finally, Section 4 presents some conclusions.

2 Estimators

Here we review two estimators, i.e., the local Whittle estimator and the modified R/S statistic, used in this paper to calculate the Hurst exponent of indices of the Chinese stock markets. The local Whittle method is expected to provide an exact result with its standard error. The modified R/S statistic determines the short-range dependence. Combining with the V-statistic will verify the acceptance or rejection of a null hypothesis of no long-range dependence.

2.1 The local Whittle method

It is essential to know the degree of long-term memory. The local Whittle estimator given by Robinson is used to calculate the Hurst exponent H. The local Whittle method is a semi-parametric estimator, which only requires specifying the parametric form of the spectral density when the frequency λ is close to zero,

$$f(\lambda) \sim G\lambda^{1-2H}, \quad \text{as} \quad \lambda \to 0+ \tag{1}$$

where $G \in (0, \infty)$ and $H \in (0, 1)$. Let N denote the size of the time series. This method involves an additional parameter, m, an integer less than $N/2$, which must satisfy the condition that $N \to \infty$,

$$\frac{1}{m} + \frac{m}{N} \to 0 \tag{2}$$

This means that m must tend towards infinity for consistency, which it must do more slowly than n. For a spectral density of the form Eq. (1) and some assumptions including Eq. (2), the Whittle approximation of the Gaussian likelihood function is obtained by minimizing

$$Q(G,H) = \frac{1}{m}\sum_{j=1}^{m}\left[\lg(G\lambda_j^{1-2H}) + \frac{I(\lambda_j)}{G\lambda_j^{1-2H}}\right] \tag{3}$$

Replacing G by its estimator $\hat{G}$,

$$\hat{G} = \frac{1}{m}\sum_{j=1}^{m} I(\lambda_j)\lambda_j^{2H-1} \tag{4}$$

the estimate

$$(\hat{G}, \hat{H}) = \arg\min Q(G,H) \tag{5}$$

exists, which we can also write as:

$$\hat{H} = \arg\min R(H) \tag{6}$$

where

$$R(H) = \lg \hat{G}(H) - \frac{2H-1}{m}\sum_{j=1}^{m} \lg \ \lambda_j$$

$$\hat{G} = \frac{1}{m}\sum_{j=1}^{m} I(\lambda_j)\lambda_j^{2H-1}$$

converges in probability to the actual value H

$$m^{1/2}(\hat{H} - H) \to \ _d Normal\,(0, 1/4) \tag{7}$$

2.2 The Modified R/S Statistic

The R/S statistic, which was originally developed by Hurst [4] in his studies of river discharges, is usually applied to detect long-range dependence. A global parameter measures the persistence effect that now bears Hurst's name. Specifically, consider a sample of returns X_1, X_2, $\cdots$, X_n, and let $\bar{X}$ denote the sample mean, i.e, $\bar{X} = \frac{1}{n}\sum_j X_j$. The classical R/S statistic, denoted by $\tilde{Q}_n = R/S_n$, is defined as:

$$\tilde{Q}_n \equiv$$

$$\frac{1}{s_n}\left[\max_{1\leqslant k\leqslant n}\sum_{j=1}^{k}(X_j - \bar{X_n}) - \min_{1\leqslant k\leqslant n}\sum_{j=1}^{k}(X_j - \bar{X_n})\right] \tag{8}$$

where s_n is the usual maximum likelihood standard deviation estimator

$$s_n^2 \equiv \frac{1}{n}\sum_{j=1}^{n}(X_j - \bar{X}_n)^2$$

Thence we can make a linear regression to $\lg \tilde{Q}_n$ and $\lg n$ to get the Hurst exponent (the slope) by the equation

$$\lg \tilde{Q}_n = c + H \lg n \tag{9}$$

However, Lo suggested that the R/S would be sensible to short-range dependencies and he proposed the modified R/S statistic. The statistic $Q_n = R/\sigma_n$ defined by Lo is

$$Q_n \equiv \frac{1}{\hat{\sigma}_n(q)}\left[\max_{1\leqslant k\leqslant n}\sum_{j=1}^{k}(X_j - \bar{X}_n) - \min_{1\leqslant k\leqslant n}\sum_{j=1}^{k}(X_j - \bar{X}_n)\right] \tag{10}$$

where

$$\hat{\sigma}_n^2(q) \equiv \frac{1}{n}\sum_{j=1}^{n}(X_j - \bar{X}_n)^2 + \frac{2}{n}\sum_{j=1}^{q}\omega_j(q)\left[\sum_{i=j+1}^{n}(X_i - \bar{X}_n)(X_{i-j} - \bar{X}_n)\right] = \hat{\sigma}_x^2 + 2\sum_{j=1}^{q}\omega_j(q)\hat{\gamma}_j$$

$$\omega_j(q) \equiv 1 - \frac{j}{q+1}, \quad q < n$$

and $\hat{\sigma}_x^2$ and $\hat{\gamma}_j$ are the usual sample variance and the autocovariance estimators of X. The parameter q is called the truncation lag and must be chosen with some consideration of the data on hand.

As n increases without bound, the rescaled range converges in distribution to a well-defined random variable V or V-statistic when properly normalized, i.e.,

$$\frac{1}{\sqrt{n}}Q_n \equiv V_n \tag{11}$$

This distribution function of V is given explicitly by Kennedy [29] and Siddiqui [30] as:

$$F_V(v) = 1 + 2\sum_{k=1}^{\infty}(1 - 4k^2v^2)\mathrm{e}^{-2(kv)^2} \tag{12}$$

Lo then defined the modified R/S statistic $V_q(N)$, by setting $V_q(N) = N^{-1/2}Q_N$. Since

$$\lim_{N\to\infty} \mathcal{P}\{V_q(N) \in [0.809, 1.862]\} = 0.95$$

he used the interval [0.809, 1.862] as the 95 % (asymptotic) acceptance region for testing the null hypothesis

$$H_0 = \{\text{no long-range dependence, i.e., } H = 0.5\}$$

against the composite alternative

$$H_1 = \{\text{there is long-range dependence, i.e., } 0.5 < H < 1\}$$

3 Numerical calculations and results

The returns are generally computed as the difference in the logarithm of the closure prices. Let P_t denote the price of the index at time t, and define the h-period return at time $t+h$ as $r_{t+h,h} = \lg P_{t+h} - \lg P_t$. Then the daily volatilities of the price index are calculated as the difference between successive daily log-index, $r_{t+1} = \lg P_{t+1} - \lg P_t$.

Five indices are used in this study. These are the Composite Index (SH000001), the Commercial Index (SH001005), the Real Estate Index (SH001006) of SHSE, the Composite Index (SZ399106), and the Component Index (SZ399001) of the Shenzhen Stock Exchange (SZSE). For simplicity, we label these indices by code 11, 12, 13, 21, and 22, respectively. Table 1 shows information about the indices. Table 2 demonstrates Hurst exponents for closure-closure returns of various indices calculated by the local Whittle method and the classical R/S statistic (see also Fig. 1). We find that Hurst exponents calculated by the classical R/S statistic are much larger.

Table 1 Information about the indices.

Code	Begin date	End date	Number of data
11	01/02/1991	03/28/2008	4206
12	12/06/1994	03/28/2008	3178
13	12/06/1994	03/28/2008	3178
21	12/23/1991	03/28/2008	3953
22	04/03/1991	03/28/2008	4175

Table 2 The Hurst exponents calculated by the local Whittle method and the classical R/S statistic.

Code	H (LWM)	H (RS)
11	0.5595	0.6515
12	0.5483	0.6074
13	0.5156	0.5735
21	0.5428	0.6333
22	0.5560	0.6331

For the purpose of understanding the intraday and the overnight effects, it is convenient to consider several

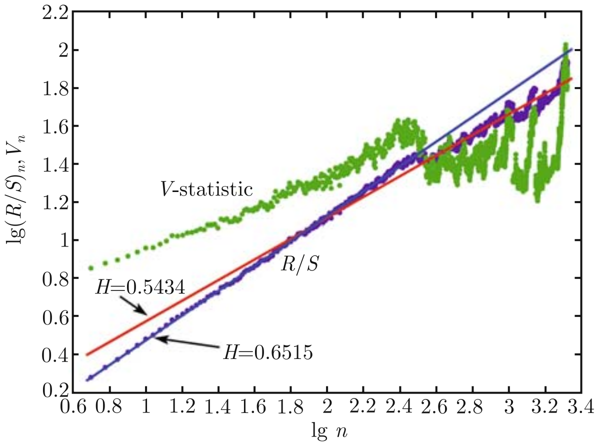

Fig. 1 The R/S fit (*lower*) and the V-statistic (*upper*) for the Hurst exponent of the SHSE Composite Index, by using daily returns in the period of Jan. 1991–Mar. 2008.

distinct returns with unique features. In particular, employing closure and opening prices of an index, we build for closure-closure (Cl-Cl) and opening-opening (Op-Op) returns, and also for opening-closure (Op-Cl) returns (a measure of the intraday return) and closure-opening (Cl-Op) returns (the latter from day t to day $t+1$, which can be seen as an overnight return). We believe that differences in the long-range dependence phenomena intensity of these indices may be interpreted as a segmentation of the dynamics of the market. Table 3 presents Hurst exponents and the standard errors for closure-closure, opening-closure, opening-opening and closure-opening returns. Since the probability of $\hat{H}$ is given by Eq. (7), the standard error is then equal to $1/\sqrt{2N}$, where N is the total number of the time series.

Table 3 Hurst exponents H and standard errors s.

Code		Cl-Cl	Op-Cl	Op-Op	Cl-Op
	s	H	H	H	H
11	0.0109	0.5595	0.4468	0.5101	0.5703
12	0.0125	0.5483	0.4947	0.4981	0.5860
13	0.0125	0.5156	0.5092	0.4823	0.5550
21	0.0112	0.5428	0.5027	0.5121	0.6104
22	0.0109	0.5560	0.4889	0.5118	0.6079

The Wald statistic is applied to check whether the long-range dependence hypothesis is accepted. The statistic is given by $W=(H-0.5)^2/s^2$ and has a χ^2 distribution. It tests the null hypothesis that the estimated Hurst exponent H is equal to 0.5. A p-value under 0.05 suggests that the null hypothesis of absence of long-range dependence can be rejected at the 5 % significance level. This statistic with the associated p-values is presented in Table 4, where s stands for the standard error associated with H and p for the p-value.

From Tables 3 and 4 we can see that for closure-closure and closure-opening returns, only one case accepts the absence of long-range dependence. On the other hand, for the opening-closure and opening-opening returns the rejection of absence of long-range dependence is weak. Therefore, dynamics of these time series seem to be quite different.

Table 4 Wald statistic and p-value.

Code	Cl-Cl		Op-Cl		Op-Op		Cl-Op	
	Wald	p	Wald	p	Wald	p	Wald	p
11	29.798	0.00	23.789	0.00	0.8520	0.36	41.587	0.00
12	14.834	0.00	0.1819	0.67	0.0238	0.88	47.050	0.00
13	1.5545	0.21	0.5353	0.46	1.9877	0.16	19.181	0.00
21	14.484	0.00	0.0558	0.81	1.1543	0.28	96.416	0.00
22	26.182	0.00	1.0270	0.31	1.1597	0.28	97.207	0.00

Table 5 presents the MRS analysis of the five indices, respectively. Four values of q are tried: q_{opt}, $n/10$, $n/5$, $n/2$. Data labeled by a symbol "*" indicates that the result is not in the range [0.809, 1.862], with q chosen by Andrews' data-dependent formula [9] as follows:

$$q=[k_n], \quad k_n \equiv \left(\frac{3n}{2}\right)^{1/3}\left(\frac{2\hat{\rho}}{1-\hat{\rho}^2}\right)^{2/3} \tag{13}$$

where $[k_n]$ denotes the greatest integer less than or equal to k_n, and $\hat{\rho}$ is the estimated first-order autocorrelation coefficient of the data.

In Fig. 1, results of the classical R/S analysis and the V-statistic are depicted for the SHSE Composite Index. If the series is a random walk the V-statistic $V_n=(R/S)_n/\sqrt{n}$ will be quite smooth and show no systematic deviations. The figure displayed shows a clear inflection point at approximately 250 trading days. We have plotted two lines, with slopes (Hurst exponents) to illustrate the effect of a sudden decreasing slope which was first observed by Mandelbrot and Wallis [31] for sunspot numbers and heights of waves, respectively. Mandelbrot and Wallis argued that this phenomenon is a characteristic of a time series with a large cyclic component. This slope variation was also discussed for the ensemble returns of individual stocks in SHSE [24]. Details of the cyclic trend is beyond the scope of the current paper and will be discussed elsewhere.

While we focus only on lag $n=N$, the length of the series, the modified R/S statistic $V_q(N)$ is calculated. Figure 2 sketches a plot of $V_q(N)$ vs. q of the SHSE Composite Index. However, there is no evidence about the existence of long-term dependence.

Table 5 Results of the modified R/S statistic for the SHSE Composite Index (a), the SHSE Commercial Index (b), the SHSE Real Estate Index(c), the SZSE Composite Index (d), and the SZSE Component Index (e).

	n	$\hat{V}_n$	$V_n(q = q_{\text{opt}})$	%Bias	$V_n\left(q = \frac{n}{10}\right)$	%Bias	$V_n\left(q = \frac{n}{5}\right)$	%Bias	$V_n\left(q = \frac{n}{2}\right)$	%Bias
(a)	100	1.328	1.268	4.673	1.286	3.210	1.331	−0.278	1.631	−18.594
	250	1.577	1.392	13.350	1.341	17.662	1.386	13.785	1.633	−3.391
	500	1.389	1.335	4.073	1.313	5.785	1.395	−0.441	1.964*	−29.285
	750	1.509	1.435	5.148	1.381	9.291	1.605	−5.951	1.875*	−19.493
	1000	1.678	1.220	6.718	1.406	19.296	1.460	14.866	1.453	15.435
	1500	1.302	1.220	6.718	1.312	−0.822	1.271	2.419	1.479	−12.013
(b)	100	1.288	1.231	4.625	1.259	2.340	1.320	−2.398	1.692	−23.901
	250	1.327	1.238	7.175	1.298	2.199	1.447	−8.316	1.908*	−30.471
	500	1.242	1.163	6.811	1.285	−3.331	1.475	−15.823	2.324*	−46.559
	750	1.409	1.293	8.967	1.335	5.579	1.542	−8.600	2.093*	−32.684
	1000	1.694	1.475	14.856	1.475	14.856	1.538	10.098	1.592	6.366
	1500	1.699	1.551	9.499	1.443	17.742	1.476	15.105	1.967*	−13.631
(c)	100	1.215	1.200	1.192	1.263	−3.821	1.381	−12.044	1.807	−32.797
	250	1.247	1.212	2.884	1.350	−7.625	1.491	−16.386	2.112*	−40.968
	500	1.181	1.159	1.838	1.392	−15.181	1.595	−25.966	2.381*	−50.403
	750	1.438	1.407	2.252	1.551	−7.260	1.732	−16.966	2.424*	−40.674
	1000	1.669	1.628	2.532	1.673	−0.204	1.704	−2.036	1.717	−2.769
	1500	1.864	1.811	2.968	1.726	7.990	1.735	7.469	2.504*	−25.546
(d)	100	1.281	1.263	1.414	1.278	0.233	1.345	−4.711	1.748	−26.688
	250	1.443	1.400	3.042	1.374	5.001	1.421	1.555	1.710	−15.627
	500	1.459	1.402	4.112	1.306	11.788	1.356	7.656	1.583	−7.831
	750	1.678	1.613	4.050	1.358	23.617	1.408	19.240	1.571	6.827
	1000	1.471	1.409	4.383	1.296	13.451	1.335	10.157	1.379	6.641
	1500	1.419	1.364	4.051	1.307	8.578	1.256	12.961	1.586	−10.518
(e)	100	1.277	1.236	3.287	1.251	2.057	1.328	−3.853	1.685	−24.193
	250	1.433	1.358	5.519	1.348	6.317	1.418	1.090	1.776	−19.291
	500	1.534	1.443	6.285	1.347	13.903	1.404	9.270	1.863*	−17.669
	750	1.708	1.591	7.369	1.409	21.165	1.479	15.468	1.625	5.079
	1000	1.978	1.846	7.136	1.435	37.832	1.369	44.478	1.383	42.994
	1500	1.430	1.315	8.703	1.206	18.556	1.209	18.316	2.153*	−33.601

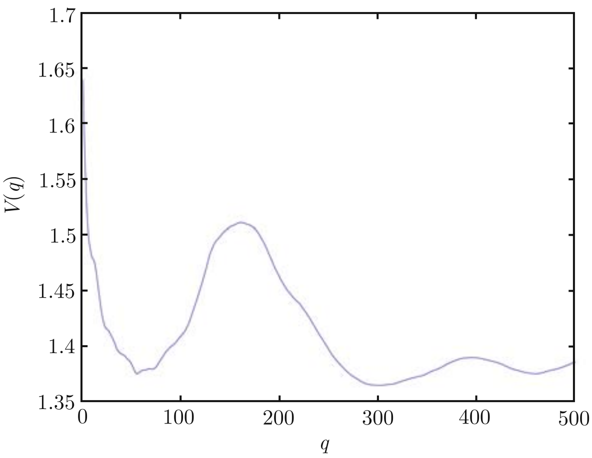

Fig. 2 The relation between $V(q)$ and q.

4 Conclusions

In this paper, we used two methods, i.e., the modified R/S estimator and the local Whittle estimator, to calculate the Hurst exponent to test the presence of the long-range dependence phenomena in daily returns of typical indices of the Chinese stock markets. The V-statistic and the Wald statistic with p-values test were respectively used to validate the null hypothesis of no long-range dependence. The modified R/S statistic accepts this null hypothesis, while the local Whittle method rejects it. Numerical results demonstrate that if there would exist a long-term memory, it will likely be a weakly persistent dependence. Moreover, there are signs showing that long cycles or trends may also have various impacts on the evaluations.

We also note that the long-range dependence phenomena presented in this market depends on the time it is measured. It displays complex dynamics in the involved time series. Based on our findings described in this paper, we believe that one should not rely on using Lo's modified R/S statistic as the sole technique to test for long-term memory in the data set of the given financial indices.

Acknowledgements This project was partially supported by the State Key Development Programme for Basic Research of China (Grant No. 2001CB309308). The authors are grateful to Drs. Y. N. Tang, Y. Zhang, and C. Y. Xu for useful discussions.

References

1. B. B. Mandelbrot, Econometrica, 1970, 38, Suppl.: 122
2. B. B. Mandelbrot, Econometrica, 1971, 39, Suppl.: 68
3. B. B. Mandelbrot, Rev. Econ. Stat., 1971, 53: 225
4. H. Hurst, Trans. Amer. Soc. Civil Eng., 1951, 116: 770
5. W.-X. Zhou, A Guide to Econophysics, Shanghai: Shanghai University of Finance and Economics Press, 2007 (in Chinese)
6. M. T. Greene and B. D. Fielitz, J. Fin. Econ., 1977, 4: 339
7. E. F. Fama and K. R. French, J. Pol. Econ., 1988, 96: 246
8. K. Aydogan and G. G. Booth, South. Econ. J., 1988, 55: 141
9. A. W. Lo, Econometrica, 1991, 59: 1279
10. W. Willinger and M. S. Taqqu, Fin. Sto., 1999, 3: 1
11. J. Geweke and S. Porter-Hudak, J. Time Series Ana., 1983, 4: 221
12. P. M. Robinson, Ann. Stat., 1995, 23: 1630
13. A. Montanari, M. S. Taqqu, and V. Teverovsky, Math. Comput. Model., 1999, 29: 217
14. W. Götzmann, J. Bus., 1993, 66: 249
15. N.-N. Huang and C. W. Yang, Appl. Econ. Lett., 1995, 2: 67
16. S. Sadique and P. Silvapulle, Intern. J. Fin. Econ., 2001, 6: 59
17. R. L. Costa and G. L. Vasconcelos, Physica A, 2003, 329: 231
18. C. Yan, J. W. Zhang, Y. Zhang, and Y. N. Tang, Physica A, 2005, 353: 425
19. J. W. Zhang, Y. Zhang, and H. Kleinert, Physica A, 2007, 377: 166
20. G.-F. Gu, W. Chen, and W.-X. Zhou, Physica A, 2008, 387: 495
21. Z.-Q. Jiang, L. Guo, and W.-X. Zhou, Eur. Phys. J. B, 2007, 57: 347
22. G.-F. Gu, W. Chen, and W.-X. Zhou, Eur. Phys. J. B, 2007, 57: 81
23. J. L. Ma and F. T. Ma, Front. Phys. China, 2007, 2(3): 368
24. G.-F. Gu and W.-X. Zhou, Physica A, 2007, 383: 497
25. Y.-C. Zhang and Y.-G. Zhou, J. Fin. Res., 2001, 249: 34 (in Chinese)
26. M. Chen and Y.-M. Hong, Lecture Notes in Computer Science, 2003, 2658: 90
27. D.-D. Chen and Y.-M. Hong, China Economic Quarterly, 2003, 3: 97 (in Chinese)
28. Y.-M. Hong, J. Amer. Stat. Assoc., 1999, 94: 1201
29. D. Kennedy, J. Appl. Prob., 1976, 13: 371
30. M. Siddiqui, Water Resources Res., 1976, 12: 1271
31. B. B. Mandelbrot and J. R. Wallis, Water Resources Research, 1969, 5: 967