Improved risk estimation in multifractal records: Application to the value at risk in finance

Mikhail I. Bogachev and Armin Bunde
Institut für Theoretische Physik III, Justus-Liebig-Universität Giessen, 35392 Giessen, Germany
(Received 16 July 2008; revised manuscript received 3 June 2009; published 31 August 2009)

We suggest a risk estimation method for financial records that is based on the statistics of return intervals between events above/below a certain threshold Q and is particularly suited for multifractal records. The method is based on the knowledge of the probability $W_Q(t;\Delta t)$ that within the next Δt units of time at least one event above Q occurs, if the last event occurred t time units ago. We propose an analytical estimate of W_Q and show explicitly that the proposed method is superior to the conventional precursory pattern recognition technique widely used in signal analysis, which requires considerable fine tuning and is difficult to implement. We also show that the estimation of the Value at Risk, which is a standard tool in finances, can be improved considerably by the method.

DOI: 10.1103/PhysRevE.80.026131 PACS number(s): 89.75.Da, 05.40.−a, 02.50.−r, 05.45.Df

I. INTRODUCTION

Risk estimation is a central issue in finance, seismology, weather forecasting, medicine, and many other fields [1]. By estimating the risk of an extreme event, one aims to learn in advance about the periods of high risk, this way minimizing, for example, the impact of market crashes [2–4], critical physiological conditions [5–8], natural hazards such as large earthquakes, floods, or periods of extreme temperatures [9,10], or destructive effects of mass human behavior [11,12]. The main strategy to estimate the risk of extreme events is to find their characteristic precursory patterns by understanding the laws governing the extreme events occurrence [13].

In conventional signal analysis, the standard strategies are based on finding precursory patterns $y_{n,k}:y_{n-k},y_{n-k+1},\ldots,y_{n-1}$ of k events that typically precede an extreme event $y_n>Q$. The strategies are mainly based on two approaches. In the first approach, one concentrates on the extreme events and their precursory patterns and determines the frequency of these patterns. In the second approach, one considers *all* patterns of k events $y_{n,k}:y_{n-k},y_{n-k+1},\ldots,y_{n-1}$ that precedes any event in the record and determines the probability that a given pattern is a precursor for an extreme event $y_n>Q$ [13,14]. The second approach appears more profound, since it considers information about precursors of *all* events, thus providing additional information on the time series studied, as has been confirmed recently for short- and long-term correlated data [15,16].

Recently, we have suggested a third approach, which is based on the statistics of the return intervals between events above some threshold Q [17]. This return interval approach (RIA) is particularly useful in multifractal records where linear correlations are absent and, hence, simple linear prediction tools are not useful. The most prominent data sets that fall in this category are the price returns $r_i=(P_i-P_{i-1})/P_{i-1}$, where P_i is the daily closing price of a financial asset [18–31].

Here, we compare the prediction efficiency of the conventional precursory pattern recognition technique (PRT) (which has not been applied before to financial data sets) and the RIA. Since the RIA requires less information than the PRT and no fine tuning, it is easier to implement. Using a decision algorithm and the receiver operator characteristics (ROC) analysis, we show that the RIA performs better than the PRT. We also show how the RIA can be used to considerably improve the value at risk, which is one of the central quantities in finances [32].

II. RETURN INTERVAL APPROACH AND PATTERN RECOGNITION TECHNIQUE

In the RIA, the central quantity for risk evaluation is the probability $W_Q(t;\Delta t)$ that within the next Δt units of time at least one extreme event (above Q) occurs, if the last extreme event occurred t time units ago [17]. This quantity is related to the probability density function (pdf) $P_Q(r)$ of the return intervals by

$$W_Q(t;\Delta t)=\int_t^{t+\Delta t}P_Q(r)dr\Big/\int_t^{\infty}P_Q(r)dr. \tag{1}$$

We have shown earlier that in records with nonlinear memory, the return intervals are long-term correlated, and their pdf exhibits pronounced power-law behavior

$$P_Q(r)\sim(r/R_Q)^{-\delta(Q)}, \tag{2}$$

where R_Q is the mean return time which can be obtained directly from the distribution $H(x)$ of the data by $R_Q=1/\int_Q^{\infty}H(x)dx$. The exponent $\delta>1$ depends explicitly on R_Q [17]. The power-law dependence (1) can be clearly seen in financial records [33].

As a consequence of the algebraical decay of $P_Q(r)$, $W_Q(t;\Delta t)=[\delta(Q)-1]\Delta t/t$ for $\Delta t\ll t$. To obtain by numerical simulation a more general expression for $W_Q(t;\Delta t)$ valid for all arguments, we employ a variant of the multiplicative random cascade process (MRC). In the MRC, the data are obtained in an iterative way, where the length of the record doubles in each iteration. In the zero-th iteration ($n=0$) the data set (y_i) consists of one value, $y_1^{(n=0)}=1$. In the n-th iteration, the data $y_i^{(n)}$, $i=1,2,\ldots,2^n$, is obtained from $y_{2l-1}^{(n)}=y_l^{(n-1)}m_{2l-1}^{(n)}$ and $y_{2l}^{(n)}=y_l^{(n-1)}m_{2l}^{(n)}$, where the multipliers $m_{2l-1}^{(n)}$ and $m_{2l}^{(n)}$ are independent and identically distributed (i.i.d.) random numbers taken from a Gaussian distribution with

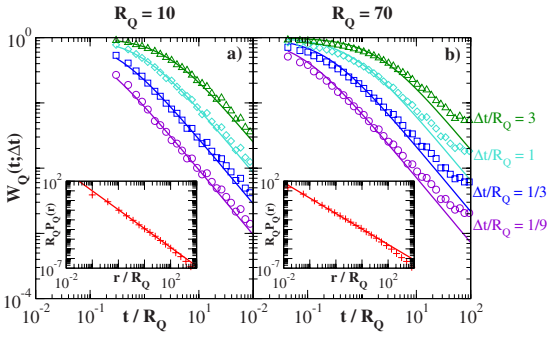

FIG. 1. (Color online) The risk functions $W_Q(t;\Delta t)$ for the MRC record for (a) $R_Q=10$ and (b) $R_Q=70$. The symbols show the numerical estimates for $\Delta t/R_Q=1/9(\bigcirc)$, $1/3(\square)$, $1(\diamond)$, and $3(\triangle)$ for an average over 150 records of length $L=2^{21}$. The corresponding analytical approximations according to (3) are shown by full lines. The pdfs of the return intervals $P_Q(r)$, for the same records, are shown in the insets.

zero mean. The resulting multifractal record has no linear correlations and can serve as a very useful model for the price returns in financial records [17]. Contrary to models for volatility, this model is specific in predicting market losses, since it allows to concentrate solely on predicting large negative returns.

Figures 1(a) and 1(b) show, for the MRC record, the risk function $W(Q;\Delta t)$, for $R_Q=10$ and 70, respectively. In the inset, the related pdf of the return intervals $P_Q(r)$ is shown. Since $W(t;\Delta t)$ is bounded by one for $t/R_Q\to 0$, the power law behavior can only be valid for $t/R_Q>[\delta(Q)-1]\Delta t/R_Q$. For large t/R_Q, strong finite size effects occur in $P_Q(r)$ which become more pronounced for large R_Q values. Since these finite size effects decrease with decreasing R_Q and increasing data length L, they underestimate the denominator in (1) and thus lead to an artificial overestimation of W_Q. To account for the proper short and large time behavior, we are thus led to the ansatz

$$W_Q(t;\Delta t)=\frac{[\delta(Q)-1]\Delta t/R_Q}{(t/R_Q)+[\delta(Q)-1]\Delta t/R_Q}, \tag{3}$$

which for the MRC yields the correct behavior for both small and large arguments t/R_Q (shown in Fig. 1 with full lines). We will use (3) for the risk estimation.

Figure 2 shows $W_Q(t;1)$ for the price returns of three representative financial records (Dow Jones index, IBM stock, British Pound vs U.S. Dollar exchange rate) for (a) $R_Q=10$ and (b) $R_Q=70$. The corresponding pdfs of the return intervals are shown, as in Fig. 1, in the insets. Since the data is short, finite size effects are more pronounced than in the simulated data of Figs. 1(a) and 1(b). For comparison, the results for the simulated data ($\langle m\rangle=0$ and $\sigma_m=1$) of comparable system size $L=2^{14}$ are also shown (dashed lines). The model pdfs represent slightly better the observational data for $R_Q=70$ than for $R_Q=10$, in agreement with the conclusions from [33]. The risk functions from the model and from the observational data agree remarkably well. Due to the short record length and no averaging, the finite size effects

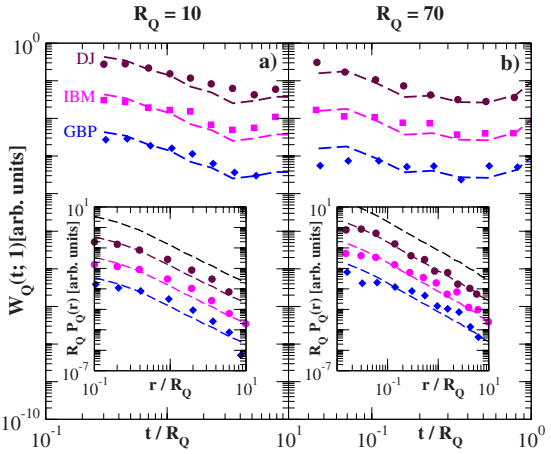

FIG. 2. (Color online) Risk functions $W_Q(t;1)$ for the price returns of three representative financial records: Dow Jones index $(*)$, IBM stock $(\bigcirc)$, British Pound vs U.S. Dollar exchange rate $(\square)$ for (a) $R_Q=10$ and (b) $R_Q=70$. The dashed lines shows the corresponding risk functions obtained from the MRC model obtained for a single representative configuration of size $L=2^{14}$. In the insets, the corresponding pdfs of the return intervals are shown; the dashed lines show the numerical estimate for an average over 150 MRC records of length $L=2^{21}$.

are comparable in the simulated and in the observational data.

The PRT approach is illustrated in Fig. 3. To estimate the risk probability in the pattern recognition approach, one can either use the "learning" observational record itself or use a model representing the record (here the MRC model, where we based our estimations on 150 MRC records of length $L=2^{21}$). First, we choose a pattern length k and create a digital database of all possible patterns $y_{n,k}$ of length k. To this end, we divide the total range of the possible data y_i into l windows such that there is the same number of values in each

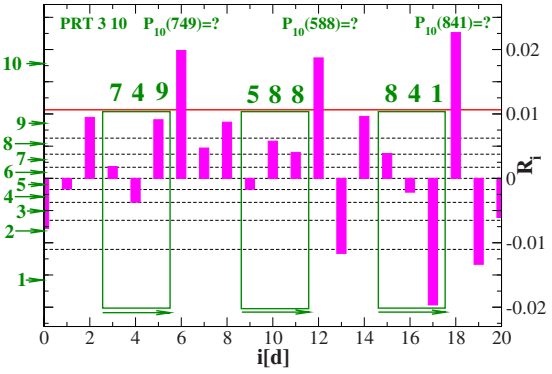

FIG. 3. (Color online) Illustration of the precursory pattern recognition technique for a fragment of the daily returns of the Dow Jones index. For illustration, we have chosen for the pattern length $k=3$ and for the magnitude resolution $l=10$. The dashed lines divide the returns into 10 quantiles. The horizontal solid line shows the threshold Q corresponding to $R_Q=10$. All patterns of three consecutive events are considered in a sliding window. The identificator (ID) of each pattern is a k dimensional consisting of the quantile numbers. For each pattern, we determine the frequency of being a precursor of an event above Q from the learning cohort.

window. Accordingly, there exist l^k different patterns. Next, we determine how often each pattern is followed by an event above Q which after normalization yields the desired probability $P(y_n > Q | y_{n,k})$ that the following event y_n exceeds Q.

The major disadvantage of the PRT (compared with the RIA) is that it needs a considerable amount of fine tuning for finding the optimum parameters l and k that yield the highest prediction efficiency. For transparency, we have kept the total number of patterns $l^k = const$ and concentrated on five pattern lengths $k=2$, 3, 4, 5, and 6. For the predictions in the MRC record, we have chosen $l^k = 10^6$ and obtained the best result for $k=2$. Smaller and larger values of l^k did not improve the prediction efficiency. For predictions based on the observational records where the statistics is limited, it is usually not possible to exceed $l^k = 10^2$ for obtaining ROC curves (see below) that cover the whole area between zero and unit sensitivity. We obtained the best performance for $k=1$ and 2.

III. DECISON-MAKING ALGORITHM AND THE ROC-ANALYSIS

The common strategy for a decision-making algorithm is to seek for that pattern which has the highest probability to be followed by an extreme event and give an alarm when this pattern appears. In nonlinear complex records (e.g., in finance, geophysics, climate, and physiology), this pattern may not be representative, since many other patterns may have comparable probabilities to be followed by an extreme event. In this case, a better approach is to give an alarm when the estimated probability for an extreme event to occur exceeds a certain threshold Q_P. The (arbitrary) selection of Q_P is usually optimized according to the minimal total cost of false predictions made, including false alarms and missed events, after a certain cost of a single false alarm and of a single missed event has been specified, see, e.g. [14].

In order to illustrate the decision-making algorithm, we show in Fig. 4(a) a representative fragment of the Dow Jones returns record [multiplied by (−1) such that large positive values now represent large losses], where we have indicated a threshold Q corresponding to the mean return time $R_Q = 70$ between large negative returns. Figure 4(b) illustrates the estimated probabilities $W_Q(t;1)$ from (3) for the above record. Figure 4(c) illustrates the same probabilities estimated by the PRT with $k=2$ and $l=10$.

The decision threshold Q_P is shown as dashed lines in Figs. 4(b) and 4(c). When the estimated occurrence probabilities exceed Q_P, an alarm is activated. For a certain Q_P value, the efficiency of the algorithm is generally quantified by the sensitivity S_e, which denotes the fraction of correctly predicted events, and the specificity S_p, which denotes the fraction of correctly predicted nonevents. The larger S_e and S_p are, the better is the prediction provided by the algorithm. The overall quantification of the prediction efficiency is usually obtained from the ROC analysis, where S_p is plotted versus S_e for all possible Q_P values. By definition, for $Q_P = 0$, $S_e = 1$ and $S_p = 0$, while for $Q_P = 1$, $S_e = 0$, and $S_p = 1$. For $0 < Q_P < 1$, the ROC curve connects the upper left corner of the panel with the lower right one. If there is no memory in the data, $S_p + S_e = 1$, and the ROC curve is a straight line

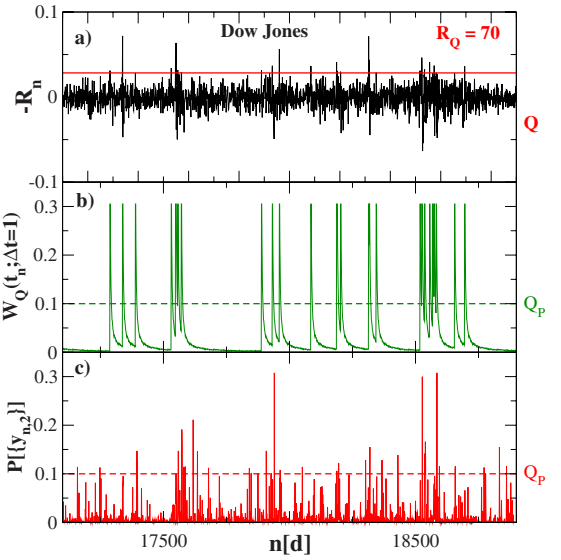

FIG. 4. (Color online) (a) A representative fragment of the Dow Jones index returns record. (b) Risk function $W_Q(t;1)$ estimated by (3). (c) The same risk probabilities obtained by the precursory pattern recognition with $k=2$ and $l=10$. All quantities are given for $R_Q = 70$. An alarm is given, when the probabilities exceed a certain decision threshold Q_P.

between both corners (dashed lines in Fig. 5). The total measure of the predictive power PP, $0 < PP < 1$, is the integral over the ROC curve, which equals one for perfect prediction and equals one half for the random guess.

Figures 5(a) and 5(b) show the ROC curves for a single MRC record of length $L = 2^{21}$ for $R_Q = 10$ and 70. The figure shows that in this case, where the statistics is excellent (compared with observational records), the prediction efficiency of both the RIA and the PRT approach, is quite high and comparable with each other. Figures 5(c)–5(h) show the equivalent curves for the Dow Jones Index (c, d), the IBM Stock (e, f) and the exchange rate between the British Pound and the U.S. Dollar (g, h). In these figures, we have also added the corresponding PRT results obtained from the observational records, where we "learned" from the precursors of large positive returns (gains) to predict large negative returns (losses). This is reasonable since the behavior of large positive returns is in quantitative agreement with that for negative returns [33], see also [34]. For both $R_Q = 10$ and 70, learning on the MRC model generally yields a higher prediction efficiency than learning on the observational records. The figure shows that for the 3 financial records, the ROC curves for the RIA are systematically above the curves from the PRT, especially close to $S_e = 1$. Accordingly, for the same high sensitivities the RIA yields considerably less false alarms than the PRT. The superiority of the RIA approach increases with an increasing return period R_Q when the finite size effects in the observational records become increasingly important. We like to note that we have obtained similar conclusions for all representative records analyzed, consisting of four indices (DJ, FTSE, NASDAQ, and S&P 500), four stocks (BOEING, GE, GM, and IBM), four currency exchange rates (DKK, GBP, GM, and SWF vs USD) and four oil prices (Brent, WTI, Rotterdam, and Singapore).

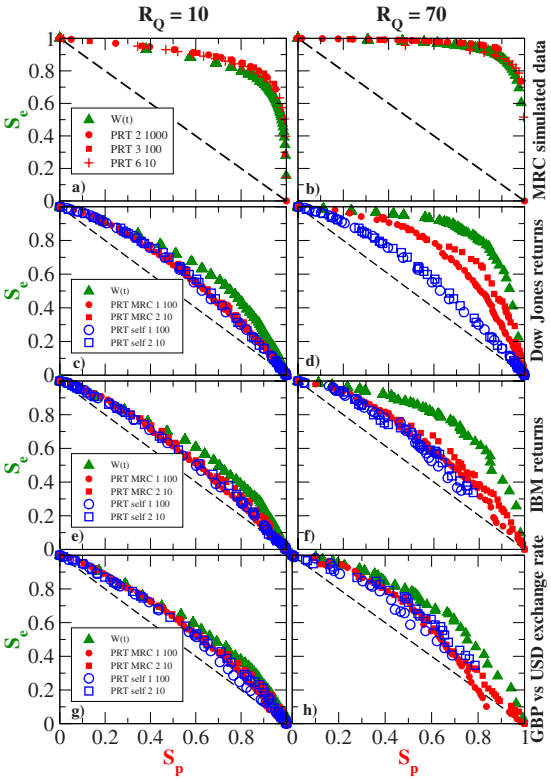

FIG. 5. (Color online) ROC curves quantifying the prediction efficiency obtained in the linearly uncorrelated MRC record, (a) for $R_Q=10$ and (b) for $R_Q=70$, based on the PRT for $k=2$ (●), 3 (■), and 6(+) as well as on RIA, Eq. (3) (▲). Similar curves are presented in (c,d) for the Dow Jones returns, in (e,f) for the IBM stock and in (g,h) for the British Pound vs U.S. Dollar exchange rate. In (c–h) ROC-curves are shown for $k=1$ (●) and 2 (■) for pattern database obtained from the MRC model and for $k=1$ (○) and 2 (□) for the pattern database obtained directly from the observational record.

IV. VALUE AT RISK

Finally, we use the RIA to estimate the Value at Risk (VaR), which is probably the best-known risk estimation technique in finances. The VaR is defined as the loss that, in a given time interval, can only be exceeded with a certain small probability q. In the following, we consider as time interval 1 day. In a first order approximation, when memory effects are being neglected, the VaR can be simply determined from the (global) distribution $P(r)$ of the daily returns via

$$\int_{-Q}^{-\infty} P(r)dr = q. \tag{4}$$

By solving this equation, which is identical to $1/R_{-Q}=q$, one can obtain the return $-Q$ and the corresponding VaR. In order to take into account the fact that the fluctuations in the returns vary in time, one often does not consider the global distribution to estimate $-Q$, but rather a local distribution of the returns based on, e.g., the last 100 days. This technique, which allows to distinguish between volatile and less volatile times, is usually a better estimator of the VaR. For a further

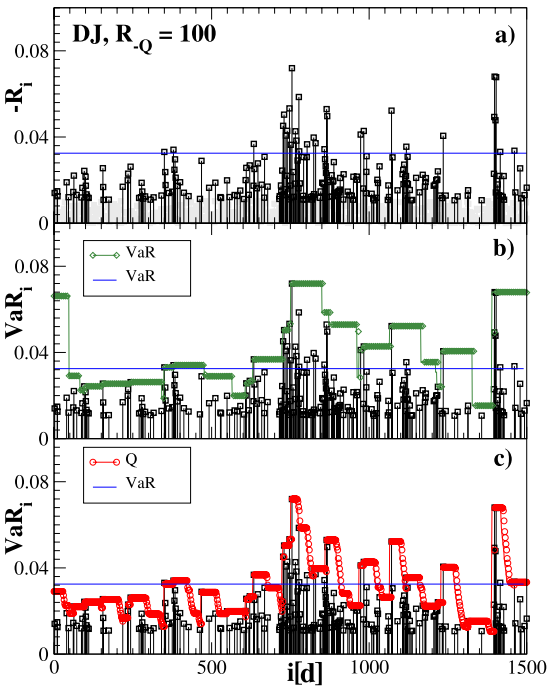

FIG. 6. (Color online) Value-at-Risk estimates for the Dow Jones index. (a) A fragment of the DJ daily returns sequence from 09/10/1934 until 04/10/1940. (b) VaR estimates for the exceedance probability $q=1/100$ for each day, obtained from Eq. (4) by using the global distribution (straight line) or the local distribution of the last 100 days (diamonds). (c) VaR estimates based on the return interval approach (circles).

improvement of the VaR, recently a method based on the return interval between the last two events below $-Q$ has been suggested [35].

Here we show how the RIA can be used to further improve the estimations of the VaR by using explicitly Eq. (3) for the risk function W_Q, which is a symmetric function of Q.

To obtain the VaR within the RIA, we proceed iteratively:

(i) In the first step, we choose q and determine from $1/R_{-Q}=q$ the corresponding loss $-Q$ in zero-th order approximation. Next, we determine the time $t=t_{-Q}$ that has been elapsed after the last return below $-Q$ and use Eq. (3) to determine the new probability W_{-Q}.

(ii) (a) If W_{-Q} is within a certain confidence interval Δq around q, the algorithm is stopped. (b) If W_{-Q} is above the confidence interval we multiply $-Q$ by $1+\Delta Q$ and determine the new elapsed time $t_{-Q(1+\Delta Q)}$ after the last event below $-Q(1+\Delta Q)$. If W is still above the confidence interval, we repeat this step until, in the n^* step, $W_{-Qn^*(1+\Delta Q)}$ is either within or below the confidence interval. Then we stop the algorithm and choose, as the estimate of the VaR, $-Q^*=-Qn^*(1+\Delta Q)$. (c) If W_{-Q} is below the confidence interval, we proceed as in (b), but with a negative increment $\Delta Q<0$ until we are within or above the confidence interval.

Figures 6–8 show the VaR for the three assets discussed in the preceding figures. Each figure consists of three panels. In panel (a), we show the negative returns $-R_i$. For transparency, we only highlight those negative returns with a return period above 10. We consider as extreme events returns with return period above 100 which is shown as straight line in

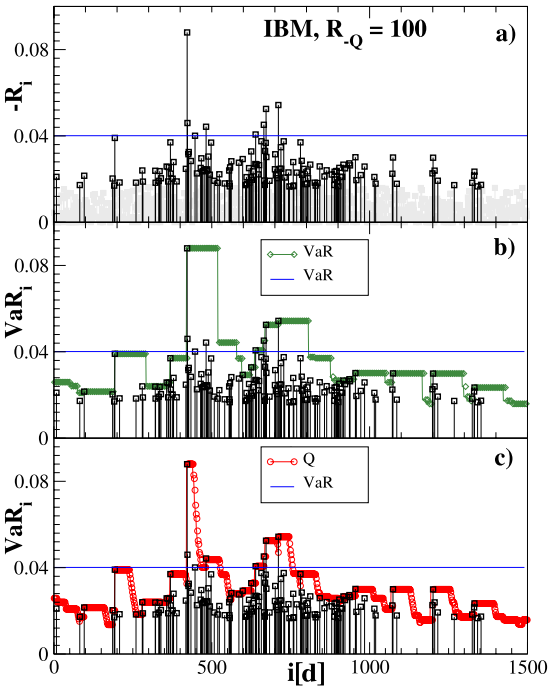

FIG. 7. (Color online) Same as in Fig. 6 but for the IBM stock from 11/01/1972 until 19/12/1977.

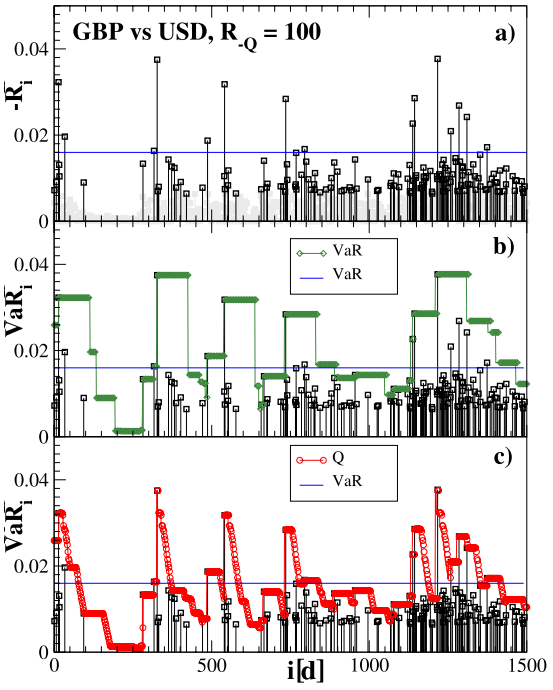

FIG. 8. (Color online) Same as in Fig. 6 but for the GBP vs USD exchange rate from 09/12/1975 until 13/11/1981.

the figures. In panel (b) we show an estimate of the VaR for each day i. We employ Eq. (4) with $q=1/100$ and choose the (local) distribution of the returns from the last 100 days. In this case, the VaR for day i is identical to the maximum negative return between day $i-100$ and $i-1$. In panel (c), we show the estimate of the VaR by the RIA, obtained by the iterative procedure described above. The confidence interval was chosen between 0.0099 and 0.0101, and the size of the Q increments was chosen as $|\Delta Q|=0.025$.

V. CONCLUSION

In summary, we have used two techniques, a PRT known in signal analysis and a RIA to estimate the risk in multifractal records in the absence of linear correlations. We have applied both methods to financial records and have shown that the RIA which is able to exploit long-term memory is superior to the conventional precursory pattern recognition approach that focuses solely on short-term memory. A major additional advantage of the return intervals approach is that it does not require extensive learning or tuning procedures which are needed in the pattern recognition technique, and thus it is considerably easier to implement. Finally, we employed the RIA to estimate the VaR, which is the standard tool of risk estimation in finances. We showed explicitly that the RIA gives significantly better estimates than conventional methods, which are either based on the global or the local distribution of the returns. We like to note that the formalism can also be applied to multifractal data with concomitant linear correlations that can be used to model physiological rhythms, turbulent flows, precipitation, river flows, and network traffic where it also leads to an improved risk prediction.

ACKNOWLEDGMENT

We like to thank our colleagues Shlomo Havlin and Sabine Lennartz for valuable discussions. Part of this work has been supported by DAPHNET.

[1] A. Bunde, J. Kropp, and H. J. Schellnhuber, *The Science of Disasters* (Springer, Berlin, Heidelberg, New York, 2002).

[2] T. Lux and M. Marchesi, Int. J. Theor. Appl. Finance **3**, 675 (2000).

[3] P. Hartmann, S. Straetmans, and C. G. de Vries, in *Asset Price Bubbles: The Implications for Monetary, Regulatory and International Policies*, edited by W. Hunter, G. Kaufman, and M. Pomerleano (MIT Press, Cambridge, MA, 2003).

[4] P. Hartmann, S. Straetmans, and C. G. de Vries, Rev. Econ. Stat. **86**, 313 (2004).

[5] H. E. Stanley, L. A. N. Amaral, A. L. Goldberger, S. Havlin, P. Ch. Ivanov, and C.-K. Peng, Physica A **270**, 309 (1999).

[6] P. Ch. Ivanov, L. A. N. Amaral, Ary L. Goldberger, S. Havlin, M. G. Rosenblum, H. E. Stanley, and Z. R. Struzik, Chaos **11**, 641 (2001).

[7] *Fractals in Biology and Medicine*, edited by G. A. Losa, D.

Merlini, T. F. Nonnenmacher, and E. R. Weibel (Birkhäuser, Basel, 2005).
[8] M. I. Bogachev and A. Bunde, New J. Phys. **11**, 063036 (2009).
[9] M. Mudelsee, M. Börngen, G. Tetzlaff, and U. Grünewald, Nature (London) **425**, 166 (2003).
[10] A. Bunde, J. F. Eichner, J. W. Kantelhardt, and S. Havlin, Phys. Rev. Lett. **94**, 048701 (2005).
[11] D. Helbing, I. Farkas, and T. Viscek, in [1], p. 331.
[12] M. I. Bogachev and A. Bunde, Europhys. Lett. **86**, 66002 (2009).
[13] C. M. Bishop, *Neural Networks for Pattern Recognition* (Oxford University Press, Oxford, 1995).
[14] J. M. Bernado and A. F. M. Smith, *Bayessian Theory* (Wiley, New York, 1994).
[15] S. Hallerberg, E. G. Altmann, D. Holstein, and H. Kantz, Phys. Rev. E **75**, 016706 (2007).
[16] S. Hallerberg and H. Kantz, Phys. Rev. E **77**, 011108 (2008).
[17] M. I. Bogachev, J. F. Eichner, and A. Bunde, Phys. Rev. Lett. **99**, 240601 (2007).
[18] S. Ghashghaie, W. Breymann, J. Peinke, P. Talkner, and Y. Dodge, Nature (London) **381**, 767 (1996).
[19] T. Lux, Appl. Econ. Lett. **3**, 701 (1996).
[20] N. Vandewalle and M. Ausloos, Eur. Phys. J. B **4**, 257 (1998).
[21] K. Ivanova and M. Ausloos, Eur. Phys. J. B **8**, 665 (1999).
[22] N. Vandewalle and M. Ausloos, Int. J. Mod. Phys. C **9**, 711 (1998).
[23] B. B. Mandelbrot, A. Fisher, and L. Calvet, *A Multifractal Model for Asset Returns* (Yale University, New Haven, CT, 1997).
[24] A. Fisher, L. Calvet, and B. B. Mandelbrot, *Multifractality of Deutschemark/U.S. Dollar Exchange Rates* (Yale University, New Haven, CT 1997).
[25] A. Arneodo, J.-F. Muzy, and D. Sornette, Eur. Phys. J. B **2**, 277 (1998).
[26] J.-P. Bouchaud, M. Potters, and M. Meyer, Eur. Phys. J. B **13**, 595 (2000).
[27] B. B. Mandelbrot, *Gaussian Self-Affinity and Fractals* (Springer, New York, Berlin, Heidelberg, 2001).
[28] K. Matia, Y. Ashkenazy, and H. E. Stanley, Europhys. Lett. **61**, 422 (2003).
[29] J. Fillol, Econ. Bull. **3**, 1 (2003).
[30] T. Lux, Int. J. Mod. Phys. C **15**, 481 (2004).
[31] F. Wang, K. Yamasaki, S. Havlin, and H. E. Stanley, Phys. Rev. E **77**, 016109 (2008).
[32] P. Jorion, *Value-at-Risk: The New Benchmark for Managing Financial Risk* (McGraw-Hill, CA, 2001).
[33] M. I. Bogachev and A. Bunde, Phys. Rev. E **78**, 036114 (2008).
[34] Z. Eisler, J. Perelló, and J. Masoliver, Phys. Rev. E **76**, 056105 (2007).
[35] K. Yamasaki, L. Muchnik, S. Havlin, A. Bunde, and H. E. Stanley, *Practical Fruits of Econophysics* (Springer, Tokyo, 2006), p. 43.