Hurst Exponent, Efficiency of Moving Averages on NYSE and their Relationship

Martin Zeman

Vedoucí práce: doc. Ing. Jiří Trešl, CSc.

Technical Analysis, Simple Moving Average, Weighted Moving Average

Technical analysis tries to predict future with respect to past data. One of its tools is based on calculation of a simple moving average (SMA) and its confrontation either with current market price or with shorter simple moving average. This confrontation helps us to indicate a new trend at its beginning. The trend is believed to continue for some time. The more values are involved in SMA, the less it is sensitive to sudden changes. Longer SMA are used to detect trends that already last some time, using too short SMA results in a great number of sell/buy signals some of which are considered to be wrong.

Technical analysis suggests:

$$SMA_{t-1}(n_1) \leq SMA_{t-1}(n_2) \wedge SMA_t(n_1) > SMA_t(n_2) \Rightarrow BUY$$

$$SMA_{t-1}(n_1) \geq SMA_{t-1}(n_2) \wedge SMA_t(n_1) < SMA_t(n_2) \Rightarrow SELL$$

where $SMA_t(n_1) = \dfrac{1}{n}\sum_{i=0}^{n-1} X_{t-i}$

n is the length of the period of which data are took into account in SMA(n)

$n_1 < n_2$ and $n_1 \in \{1, 2, 3, ...\}$

The suggestion above also applies to weighted moving average which is given by formula

$$WMA_t(n) = \frac{2}{n \cdot (n+1)}\left[nX_t + (n-1)X_{t-1} + ... + 2X_{t-n+2} + X_{t-n+1} \right]$$

An interesting feature of this recommendation is that buy and sell signals take turns therefore the number of buy signals and the number of sell signals differ at most by 1.

Hurst exponent (H)

Hurst exponent is a characteristic of a data set.

„Interpretation of Hurst exponent is following:

- If H=0.50, then is time series generated by i.i.d. (Independent Identically Distributed) process. R/S analysis is a non-parametric procedure and it does not demand any assumptions about process parameters.

● In the range are corresponding processes called persistent and they are characterized by long memory. There is no characteristic time scale, which is typical for fractal time series. They are suitable for the modelling of time series of stock returns."[1]

„Estimating the Hurst exponent for a data set provides a measure of whether the data is a pure random walk or has underlying trends. … The Hurst exponent is not so much calculated as estimated. A variety of techniques exist for doing this and the accuracy of the estimation can be a complicated issue."[2]

Relationship between Hurst Exponent and Technical Analysis

As it was already mentioned, Hurst exponent tries to resolve the difference between a random process and a process, where some trends exists. This is connected with one of technical analysis' assumption.

Technical analysis is based on the belief that „price movement creates always certain trend, in which the price persists for certain time"[3]

I believe that quantifying this relation could help traders to make better decision on employing technical analysis. It could also help to create some picture about the Hurst exponent.

Certainly it would be better to express the relationship between Hurst exponent and some tool of technical analysis in a mathematical way but I think it would be very difficult to do that and I work only with empirical observations.

There may exist some implications concerning this relationship, which would suggest the advisability of the use of technical analysis. I assume it's not extra-profitable to use technical analysis with respect to the future in a market where the Hurst exponent equals 0,5 even if some tool of technical analysis proves to be profitable at that market. If Hurst exponent equal to 0,5 implies an independent process then what justification has the use of technical analysis when H=0,5? On the other hand, the success of a technical analysis when H>0.5 is more likely not to be a matter of coincidence. Supposing Hurst exponent expresses the difference between randomness and predictability well, it is better to justify the use of a technical analysis tool not only by its success in the past but also by the value of the Hurst exponent because the success in the past could be caused by a very precise selection of the tool and its parameters rather than by a dependent and predictable character of the process.

Proved correlation between Hurst exponent and profitability of a certain tool could suggest how important the value of the Hurst exponent is for our decision to employ that tool for trading on a particular market. Likewise the F-value in analysis of variance influences our decision about suitable regression model and measures the significance of that model.

The key question is how precisely does the Hurst exponent reflect the difference between randomness and predictability. I don´t know the answer but having read some interpretations of the Hurst exponent and performed several simulations I suppose it definitely has the informative value regarding predictability.

---

1 Trešl, Jiří: Statistical Methods and Capital Markets, 2003, p. 75

2 http://www.bearcave.com/misl/misl_tech/wavelets/hurst/ (15.5.2007)

3 Trešl, Jiří: Statistical Methods and Capital Markets, 2003, p. 83

Data processing

I analyse the index Dow Jones Industrial Average (DJIA). The period 1. 10. 1928–29. 6. 1999 contains 18 811 observations (closing values of DJIA). All the data were collected from

http://www.miras.cz/akcie/indexy-djia.php

18810 returns were calculated according to $\mathrm{Ret}_t = \dfrac{DJIA_t}{DJIA_{t-1}}$.

The set of returns is characterised by high kurtosis, it reaches the value 24,3.

The whole period was divided into 19 subperiods of 990 returns each.

season	beginning	end
1	1.10.1928	3.2.1932
2	4.2.1932	6.6.1935
3	7.6.1935	21.9.1938
4	22.9.1938	7.1.1942
5	8.1.1942	26.4.1945
6	27.4.1945	1.11.1948
7	3.11.1948	25.4.1952
8	26.4.1952	27.3.1956
9	28.3.1956	2.3.1960
10	3.3.1960	7.2.1964
11	10.2.1964	12.1.1968
12	15.1.1968	24.1.1972
13	25.1.1972	26.12.1975
14	29.12.1975	27.11.1979
15	28.11.1979	26.10.1983
16	27.10.1983	28.9.1987
17	29.9.1987	27.8.1991
18	28.8.1991	27.7.1995
19	28.7.1995	29.6.1999

For the estimation of the **Hurst exponent** (related to each subperiod) the procedure which is described in *Statistical Methods and Capital Markets (2003), p. 75 , Jiří Trešl,* was used with the parameters n=10, N=990, m=99.

In this study, the standard deviation necessary for the counting up the Hurst exponent is calculated $s_j = \sqrt{\dfrac{1}{n}\sum_{i=1}^{n}\left(x_{ij} - \bar{x}_j\right)^2}$.

Calculating with $s_j = \sqrt{\dfrac{1}{n-1}\sum_{i=1}^{n}\left(x_{ij} - \bar{x}_j\right)^2}$ resp. $s_j = \sqrt{\dfrac{1}{n+1}\sum_{i=1}^{n}\left(x_{ij} - \bar{x}_j\right)^2}$ would increase resp. decrease the value of the Hurst exponent approximately by 0,008. The statistical significance of Hurst exponents here is always high.

s.	H
1	0,557
2	0,542
3	0,592
4	0,501
5	0,607
6	0,523
7	0,593
8	0,565
9	0,605
10	0,591
11	0,585
12	0,618
13	0,566
14	0,533
15	0,554
16	0,548
17	0,492
18	0,528
19	0,487

This procedure overestimates the value of H as I've simulated 3000 trials with 990 random numbers. To achieve a more random character, each of these numbers was calculated as a sum of 30 pseudorandom numbers in Excel. The average value of these 3000 Hurst exponents was 0,5403 , the standard deviation 0,043 and the average absolute deviation 0,035. The computed Hurst exponent seems to be overestimated approximately by 0,0403. This is necessary to take into account when evaluating the randomness of the DJIA estimated in this paper. However the distortion must not imply the inferior quality of using this particular estimation procedure because the estimated Hurst exponent still reflects the degree of randomness. Following tables shows the results achieved for three different processes. Its parameter reflects the influence of the previous values.

$$X_t = a \, X_{t-1} + u_t$$

$$u_t \sim N(0;1)$$

a	1	0,5	0,1	-0,1	-0,5	-1
H	0,988	0,616	0,552	0,527	0,472	0,162

$$X_t = b \, X_{t-1} \, u_{t-1} + u_t$$

$$u_t \sim N(0;1)$$

b	1	0,5	0,1	-0,1	-0,5	-1
H	0,534	0,554	0,54	0,54	0,553	0,533

$$X_t = c\, X_{t-1} + (1-c^2)^{1/2}\, u_t \;\; ; u_t \sim N(0;1)$$

c	0,999	0,8	0,5	0,3	0	-0,3	-0,5	-0,8	-0,999
H	0,987	0,71	0,614	0,581	0,54	0,504	0,472	0,395	0,167

The value H is an average of 500 Hurst exponents that were calculated for a given parameter. Apparently for the random walk process is H=0,54 in this procedure. Higher contribution of previous values affects the Hurst exponent depending on a sign of the constant parameter.

Moving averages $SMA_{t-1}(n)$ and $SMA_t(n)$ were computed for n=1, n=5, n=21 and n=100.

The recommendation
$$SMA_{t-1}(n_1) \le SMA_{t-1}(n_2) \wedge SMA_t(n_1) > SMA_t(n_2) \Rightarrow BUY$$
$$SMA_{t-1}(n_1) \ge SMA_{t-1}(n_2) \wedge SMA_t(n_1) < SMA_t(n_2) \Rightarrow SELL$$

and analogous comprising $WMA_t(n)$ will be analysed for these couples:

$n_1=1$, $n_2=21$ Each couple represents a strategy, which is strictly followed by

$n_1=5$, $n_2=21$ a trader. Mutual intersection of corresponding lines representing

$n_1=5$, $n_2=100$ simple moving average signals BUY or SELL the stock for

$n_1=21$, $n_2=100$ the current price.

The recommendation is tested for each closing price with exception for the first 21 observations of the first subperiod for the strategies where $n_2=21$ and for the first 100 observations of the same subperiod for for those two strategies where $n_2=100$. Otherwise there are 990 moments when the trading signal can arise in every subperiod.

The profit made under a particular strategy is calculated rather simply but quite faithfully. The difference between a sum of selling prices and a sum of buying prices shall be adjusted because the frequency of buys and sells can differ by 1.

$$NP = \frac{f_{sells}}{f_{buys}} \cdot \sum P_{sell} - \sum P_{buy}\,, \qquad NP - \text{nominal profit}$$

f_{sells}, f_{buys} - number of sells, resp. buys

P_{sell}, P_{buy} - the price at which the index was sold, resp. bought

For more objective comparison of the subperiods nominal profits are related to the level of the index.

$$RR = \frac{NP}{SMA_{21}(21)} + 1 \qquad RR - \text{rate of return} \quad (<1 \text{ denotes depreciation; } >1 \text{ appreciation})$$

It is also useful to calculate a profit which would indicate wheter the use of a technical analysis strategy was beneficial or not, let it be the rate of extra return.

$$RER = RR - B\&H \text{ where } B\&H = \frac{SMA_{i+1,21}(21)}{SMA_{i,21}(21)} \text{ denotes the growth of DJIA during}$$

a given subperiod i. It indicates the profit reached with the use of buy-and-hold strategy. It's better to substract B&H from RR than devide the first by the other because the deduction truly reflects the amount earned in addition. Putting them into the fraction would make e.g. RR=1,6 and B&H=1 more favourable for technical analysis than RR=3 and B&H=2.

There is no need to to calculate the profit per transaction even if the strategies based on longer SMA suggests significantly less trade signals because there cannot appear two buying signals subsequently, one after another. Therefore the amount of money invested is not greater then the last buying price at any moment. Only when transaction costs mattered, a trader would have to penalize the strategies for emerging signals.

The obtained profits gained with the use of buy/sell signals do not take into consideration a possible yield that could be earned through investing money during the time between sell signal and following buy signal. For this reason the profitability of SMA and WMA strategies is underrated in this paper. Only high transaction costs could compensate it or even make overrated.

The table shows the frequencies of trading signals.

strategy	SMA(1) and SMA(21)		SMA(5) and SMA(21)		SMA(5) and SMA(100)		SMA(21) and SMA(100)	
subperiod	Buys	Sales	Buys	Sales	Buys	Sales	Buys	Sales
1	52	53	27	28	17	18	5	6
2	46	46	27	27	29	28	8	7
3	48	48	30	30	12	12	4	4
4	48	47	31	30	24	24	6	7
5	37	37	24	24	17	16	6	5
6	57	57	29	30	24	25	8	8
7	46	46	24	25	27	28	6	7
8	41	40	28	27	12	11	6	5
9	41	41	27	27	14	15	5	6
10	51	50	27	27	19	18	8	7
11	43	43	28	28	16	16	2	3
12	39	39	27	27	21	21	8	7
13	55	55	34	34	26	26	7	7
14	53	53	35	35	24	25	6	7
15	56	57	31	32	23	22	8	7
16	50	50	29	29	15	15	5	5
17	60	59	33	32	19	19	7	7
18	60	60	30	30	21	21	5	5
19	56	57	28	29	17	17	5	5

The following table summarizes the results obtained using the above formulas.

s.	H	B&H	SMA(1) and SMA(21)			SMA(5) and SMA(21)			SMA(5) and SMA(100)			SMA(21) and SMA(100)		
			NP1	RR1	RER1	NP2	RR2	RER2	NP3	RR3	RER3	NP4	RR4	RER4
1	0,56	0,32	70	1,28	0,96	103	1,42	1,09	18	1,07	0,75	6	1,02	0,70
2	0,54	1,47	31	1,38	-0,08	20	1,25	-0,22	-8	0,89	-0,57	-6	0,92	-0,54
3	0,59	1,21	51	1,43	0,22	30	1,25	0,04	37	1,31	0,10	12	1,10	-0,11
4	0,50	0,78	27	1,19	0,41	22	1,15	0,38	-7	0,95	0,17	3	1,02	0,24
5	0,61	1,49	11	1,10	-0,39	-5	0,95	-0,54	24	1,22	-0,27	24	1,22	-0,27
6	0,52	1,07	15	1,09	0,02	7	1,04	-0,03	-13	0,92	-0,15	-12	0,93	-0,14
7	0,59	1,47	85	1,48	0,01	64	1,36	-0,11	9	1,05	-0,42	-3	0,98	-0,49
8	0,57	1,96	58	1,22	-0,74	7	1,03	-0,93	66	1,25	-0,71	10	1,04	-0,92
9	0,61	1,20	53	1,10	-0,10	13	1,03	-0,18	154	1,30	0,10	41	1,08	-0,12
10	0,59	1,30	69	1,11	-0,19	37	1,06	-0,24	95	1,15	-0,14	-4	0,99	-0,31
11	0,59	1,09	163	1,20	0,12	33	1,04	-0,05	53	1,07	-0,02	57	1,07	-0,02
12	0,62	1,05	350	1,40	0,36	141	1,16	0,12	101	1,12	0,07	-34	0,96	-0,09
13	0,57	1,00	222	1,24	0,24	-97	0,89	-0,11	105	1,12	0,11	123	1,14	0,14
14	0,53	0,92	98	1,11	0,19	-225	0,75	-0,17	11	1,01	0,09	-146	0,84	-0,08
15	0,55	1,49	7	1,01	-0,48	-129	0,85	-0,64	6	1,01	-0,48	-52	0,94	-0,55
16	0,55	1,88	862	1,69	-0,19	692	1,56	-0,33	-198	0,84	-1,04	-338	0,73	-1,16
17	0,49	1,29	-389	0,83	-0,46	-562	0,76	-0,53	-6	1,00	-0,29	-223	0,90	-0,39
18	0,53	1,54	-757	0,75	-0,80	-957	0,68	-0,86	-462	0,85	-0,70	-622	0,79	-0,75
19	0,49	2,29	1227	1,26	-1,03	1307	1,28	-1,01	-1164	0,75	-1,54	-2141	0,54	-1,75
				Sum	**-1,92**			**-4,31**			**-4,95**			**-6,60**

Sums of RER indicate that none of these strategies would be more successful than buy-and-hold.

It is better to add up the rates of extra returns than multiply the rates of return because the amount invested in any period does not depend on returns during other periods. In the case of multiplying the rate of returns, a strategy with $RR_1=0,9$ and $RR_2=0,9$ would be better than a strategy with $RR_1=40$ and $RR_2=0,01$. That is not desirable.

Correlation matrix for four SMA strategies.

| r_{xy}	H	B&H	RR1	RER1	RR2	RER2	RR3	RER3	RR4	RER4
H	1									
B&H	-0,155	1								
RR1	0,319	0,105	1							
RER1	0,293	-0,887	0,366	1						
RR2	0,148	0,083	0,874	0,328	1					
RER2	0,214	-0,875	0,329	0,971	0,410	1				
RR3	0,755	-0,274	0,032	0,271	-0,106	0,199	1			
RER3	0,368	-0,955	-0,082	0,855	-0,105	0,822	0,548	1		
RR4	0,650	-0,507	-0,030	0,460	-0,143	0,394	0,821	0,695	1	
RER4	0,314	-0,969	-0,095	0,862	-0,110	0,833	0,461	0,986	0,705	1

The values in the first column indicate a positive correlation between a profit made under any of these four strategies and the Hurst exponent. Using (abs(H-0,5403)) instead of H to express „deviation from randomness" results in decrease of the values of r_{xy} in the first column. However is it more relevant to use the H and not the deviation because the SMA strategies are based on a belief in persistency rather than in non-randomness.

Relation between the third strategy and H

Line regression seems to be the most suitable model for RR or RER and H.

RR3 = a + b H

r	0,755
R^2	0,570
St. Error of Estimation	0,107
Observations	19

H and RR3
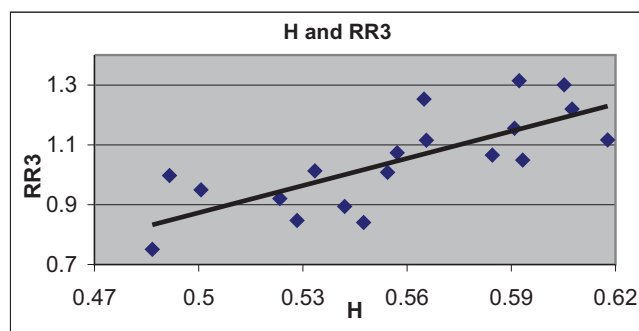
ANOVA	*SS*	*MS*	*F stat*	*P-Value*
Regression	0,259644	0,259644	22,56282	0,00018546
Residual	0,195629	0,011508		
Total	0,455273			
	parameter	*St.Error*	*t stat*	*P-Value*
Intercept	-0,63843	0,355555	-1,79559	0,09035568
Slope	3,023769	0,636578	4,750034	0,00018546

The relationship is statistically significant

RER3 = a + b H

This regression is shows that the benefit of the use of the third strategy on NYSE during periods with high H is not too significant. However bearing in mind[4] the heteroskedasticity and the nature of the stock-exchange sector, the Hurst exponent still provides some clue.

H and RER3

| r	0,368
R^2	0,136
St. Error of Estimation	0,491
Observations	19

---
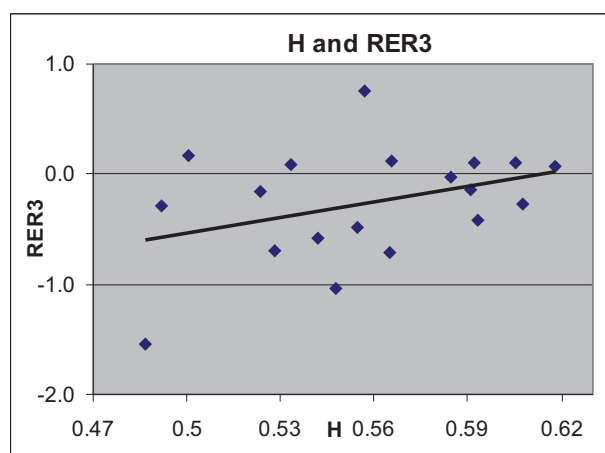
4   If the errors are less variable around the extremes of one of the predictors in the model, Ordinary Least-Squares Estimation tends to overestimate the standard error. (http://www.comm.ohio-state.edu/ahayes/hcse.pdf (19. 5. 2007))

ANOVA	SS	MS	F stat	P-Value
Regression	0,644501	0,644501	2,670975	0,12057412
Residual	4,102066	0,241298		
Total	4,746567			
	parameter	*St.Error*	*t stat*	*P-Value*
Intercept	-2,91488	1,628139	-1,79031	0,09122885
Slope	4,763993	2,914984	1,634312	0,12057412

Relation between the fourth strategy and H

RER4 = a + b H
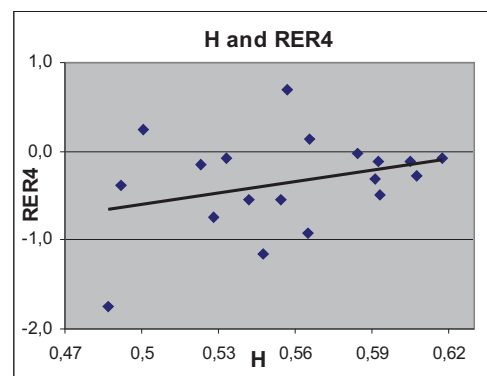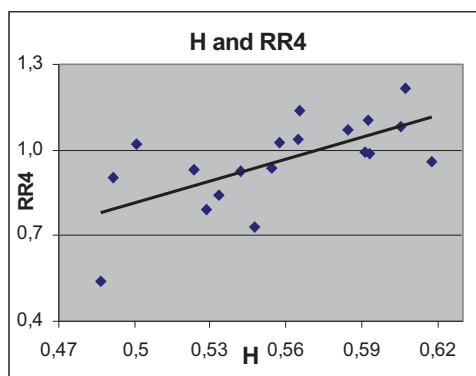
Returns with the use of WMA instead of SMA

			WMA(1) and WMA(21)			WMA(5) and WMA(21)			WMA(5) and WMA(100)			WMA(21) and WMA(100)		
s.	H	B&H	NP5	RR5	RER5	NP6	RR6	RER6	NP7	RR7	RER7	NP8	RR8	RER8
1	0,56	0,32	75	1,30	0,98	22	1,09	0,76	94	1,38	1,06	42	1,17	0,85
2	0,54	1,47	28	1,35	-0,12	16	1,20	-0,27	-24	0,70	-0,77	23	1,29	-0,18
3	0,59	1,21	33	1,28	0,07	-10	0,91	-0,30	26	1,22	0,01	14	1,12	-0,09
4	0,50	0,78	35	1,25	0,47	-1	0,99	0,22	3	1,02	0,25	-13	0,91	0,13
5	0,61	1,49	17	1,15	-0,34	-12	0,89	-0,60	29	1,26	-0,23	22	1,20	-0,29
6	0,52	1,07	24	1,15	0,08	-13	0,92	-0,15	10	1,06	-0,01	11	1,07	0,00
7	0,59	1,47	38	1,21	-0,26	33	1,18	-0,29	13	1,07	-0,40	9	1,05	-0,42
8	0,56	1,96	79	1,31	-0,66	-26	0,90	-1,06	60	1,23	-0,73	19	1,07	-0,89
9	0,61	1,20	85	1,17	-0,03	2	1,00	-0,20	144	1,28	0,08	125	1,25	0,04
10	0,59	1,30	113	1,18	-0,12	-83	0,86	-0,44	113	1,18	-0,12	35	1,06	-0,24
11	0,58	1,09	263	1,33	0,24	62	1,08	-0,01	75	1,09	0,01	54	1,07	-0,02
12	0,62	1,05	393	1,45	0,41	78	1,09	0,04	204	1,24	0,19	66	1,08	0,03
13	0,57	1,00	246	1,27	0,27	-341	0,62	-0,38	176	1,19	0,19	154	1,17	0,17
14	0,53	0,92	115	1,13	0,21	-405	0,55	-0,36	28	1,03	0,11	-15	0,98	0,07
15	0,55	1,49	101	1,12	-0,37	-54	0,93	-0,55	62	1,07	-0,41	-74	0,91	-0,58
16	0,55	1,88	-11	0,99	-0,89	673	1,54	-0,34	-354	0,71	-1,17	-241	0,81	-1,08
17	0,49	1,29	-536	0,77	-0,52	-266	0,89	-0,40	122	1,05	-0,24	335	1,14	-0,15
18	0,53	1,54	-966	0,68	-0,86	999	1,33	-0,21	-502	0,83	-0,71	-353	0,88	-0,66
19	0,49	2,29	1205	1,26	-1,03	4637	2,00	-0,30	-1277	0,73	-1,57	-971	0,79	-1,50
				Sum	-2,48			-4,83			-4,47			-4,82

Sums of RER indicate that none of these strategies would be more successful than the buy-and-hold.

Correlation matrix

r_{xy}	H	B&H	RR5	RER5	RR6	RER6	RR7	RER7	RR8	RER8
H	1									
B&H	-0,155	1								
RR5	0,435	-0,206	1							
RER5	0,292	-0,934	0,541	1						
RR6	-0,303	0,595	-0,070	-0,536	1					
RER6	-0,077	-0,705	0,192	0,676	0,151	1				
RR7	0,607	-0,559	0,334	0,602	-0,635	0,128	1			
RER7	0,328	-0,958	0,273	0,923	-0,674	0,584	0,773	1		
RR8	0,475	-0,418	0,313	0,474	-0,525	0,051	0,538	0,505	1	
RER8	0,261	-0,969	0,261	0,928	-0,651	0,617	0,624	0,957	0,629	1

Most of the values in the first column (besides the first two rows) are positive.

The Hurst exponent and the yield of the fifth and seventh strategy
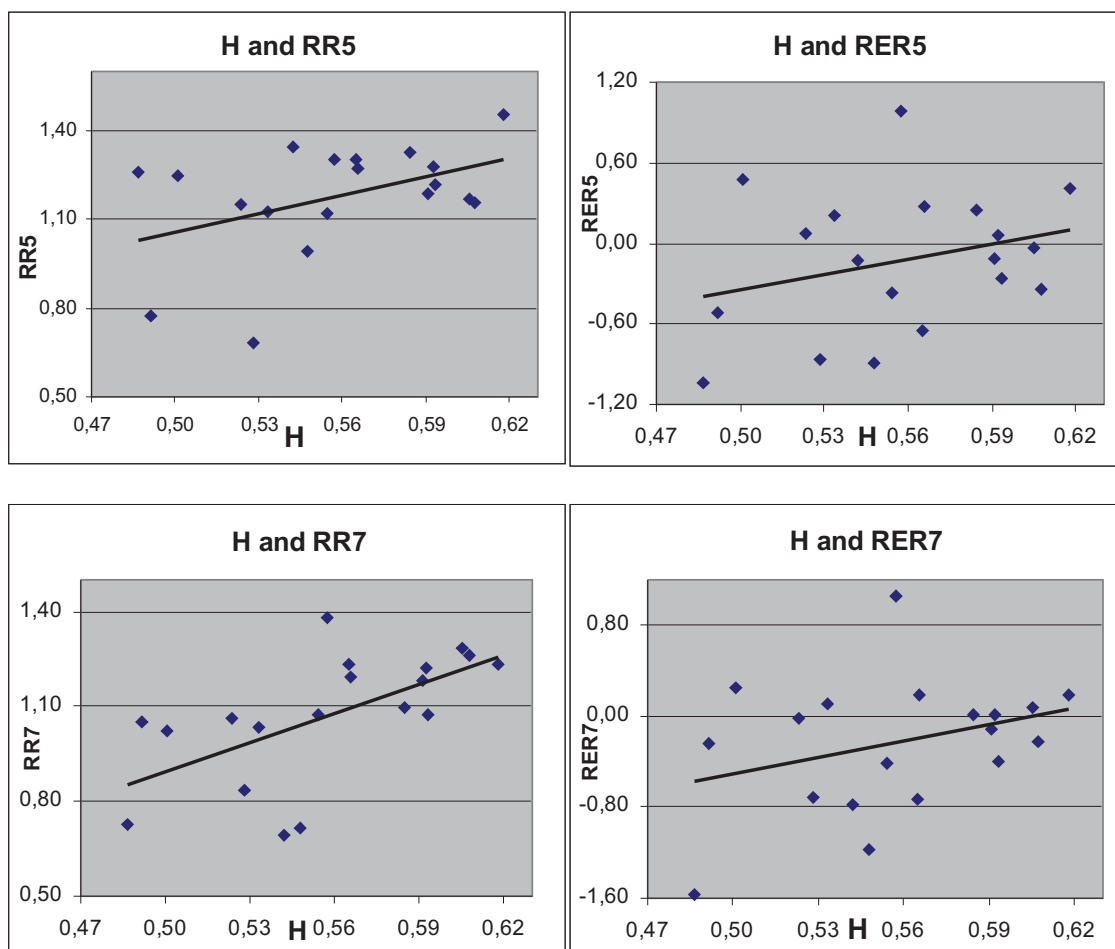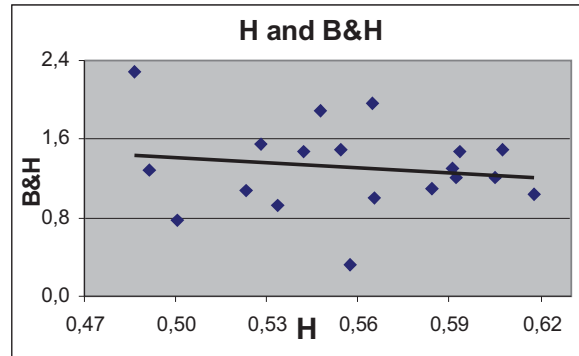
The graphs depicting the rates of extra returns denote a decreasing variance about the regression line. One of its reasons is apparent from the last graph. The 7 observations with the highest H are close to the line. And the RERs are calculated from B&H.

The regression lines of RERs slope up not only because of a positive correlation between H and RRs but also as the B&H-yield is slightly lower during periods characterised by higher H. On the other hand, in average, the RRs grows as the Hurst exponent increases even if the market grows rather slowly (apparent from B&H) during the periods with higher H.

Conclusion

The computed correlation coefficients indicate a positive dependence between the Hurst exponent and the success of the technical analysis tool which is based on a belief in presence of trends. The results are in accordance with prior expectations.

Supposedly there is a correlation between Hurst exponent and the profitability of some tools of technical analysis and these correlations vary with the tool. A significant general correlation between the H and the tool could be helpful in the decision-making about the advisability of that tool in a particular market. The Hurst exponent could be the clue that estimates the profitability of the tool in the future. This reminds the Bayes' theorem which suggests updating our beliefs in light of new information. Let us consider two strategies, the profitability of the first generally correlates positively with Hurst exponent, while the profitability of the second don't correlate . Let us suppose that they were equally profitable in a given market in the past while the Hurst exponent was 0,7 so far. Then it is better to choose the first strategy because it suits more to the nature of that market. To say whether it generally correlates or not, the 19 observations mentioned in this paper were not sufficient.

Taking the Hurst exponent into account could be adventageous because the profitability of a technical analysis tool could be sensitive to its parameters and the quantity of possible parameters is infinitely large. However it is very likely that calculating the past profits on employing the technical analysis for a large nubmer of different parameters and examining its distribution would be more relevant for a high-quality decision-making. For the last reason, Hurst exponent could be helpful to the fast but imprecise verification of a particular strategy or, I believe that would be more useful, to the search for an appropriate technical analysis tool for a given financial instrument in a particular market.

References

[1]   BRADA, Jaroslav: Technická analýza. 2000.

[2]   TREŠL, Jiří: *Statistical Methods and Capital Markets*. 2003.

[3]   http://www.miras.cz/akcie/indexy-djia.php (4. 5. 2007)

[4]   http://www.bearcave.com/misl/misl_tech/wavelets/hurst/ (15. 5. 2007)

[5]   http://www.comm.ohio-state.edu/ahayes/hcse.pdf (19. 5. 2007)