motivated by e-mail from Isil T.

Once upon a time I was playing with Hurst exponents, then ...

>Then you forgot about it, right?
Uh ... yes. In particular, I forgot to generate a spreadsheet, so we'll correct that now. Here's the scenario:

Figure 1A
log(R/s) *vs* log(n) for various values of n

1. Let the daily returns for some asset be
   $r(1), r(2), ... r(n)$
2. Let m be the **M**ean of these returns:
   $M = (1/n) [r(1)+r(2)+...+r(n)]$
3. Calculate the deviations from the mean:
   $x(1) = r(1) - M$
   $x(2) = r(2) - M$
   ...
   $x(n) = r(n) - M$
   Note that the set of **x**s have zero mean.
   A positive **x** indicates that the return was above the average.
4. Now calculate the Partial Sums:
   $Y(1) = x(1)$
   $Y(2) = x(1) + x(2)$
   ...
   $Y(n) = x(1) + x(2) + ...+ x(n)$
   Note that the set of partial sums, the **Y**'s, are sums of zero-mean variables.
   They will be positive if there's a preponderance of positive **x**'s.
   Note, too, that $Y(k) = Y(k-1) + x(k)$.
5. Let $R(n) = MAX[Y(k)] - MIN[Y(k)]$ This difference between the maximum and minimum of the n **Y**-values is called the **Range**
6. Let $s(n)$ be the standard deviation of the set of n **r**-values.
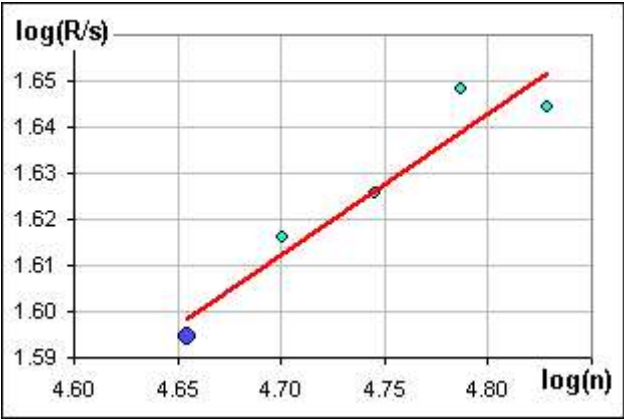7. Calculate **R/s** and plot the point $(\log(n), \log(R/s))$ ... like the **blue** dot in Figure 1A.
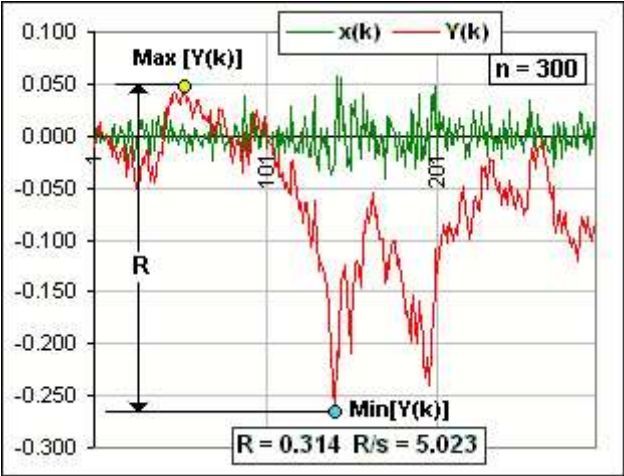

Figure 1B
x(k) and Y(k) for k = 1, 2, 3 ... n=300

>What's Figure 1B?
Them's the values of x(k) and Y(k) for n = 300 ... so k goes from 1 to 300.

>Yeah ... and the slope in Figure 1A is the Hurst Exponent, right?
Well, you do the above steps for increasing values of n, plotting points as you go.
Figure 1B is just one run, with 300 daily returns. Figure 1A is the result of several runs, with n increasing.

For large values of n, one should expect **R/s** to increase like $n^{1/2}$.
That means that log(**R/s**) should increase like (1/2) log(**n**) and that means that ...

>And the slope is the Hurst Exponent, right?
If you plot the points $(\log(n), \log(R/s))$ they should lie close to a line with slope 1/2 ... if the returns are truly random and independent.

>And 1/2 = 0.5, right?
Good for you.

>But we're talkin' large n, right?
Yes. Our calculated Hurst Exponent is just an "estimate".

Note: If the returns **r** are random and independent, a random walk with no memory, no "trends", no correlation – then the slope should be close to 0.5 and ...

>And if the slope isn't close to 0.5, then we can assume there's some correlation between returns, right?
That's what they say.

>And if the slope is significantly greater than 0.5, then we can assume that a positive return is likely to be followed by a positive return, right?
That's what they say.

>And the spreadsheet?
It looks like this:
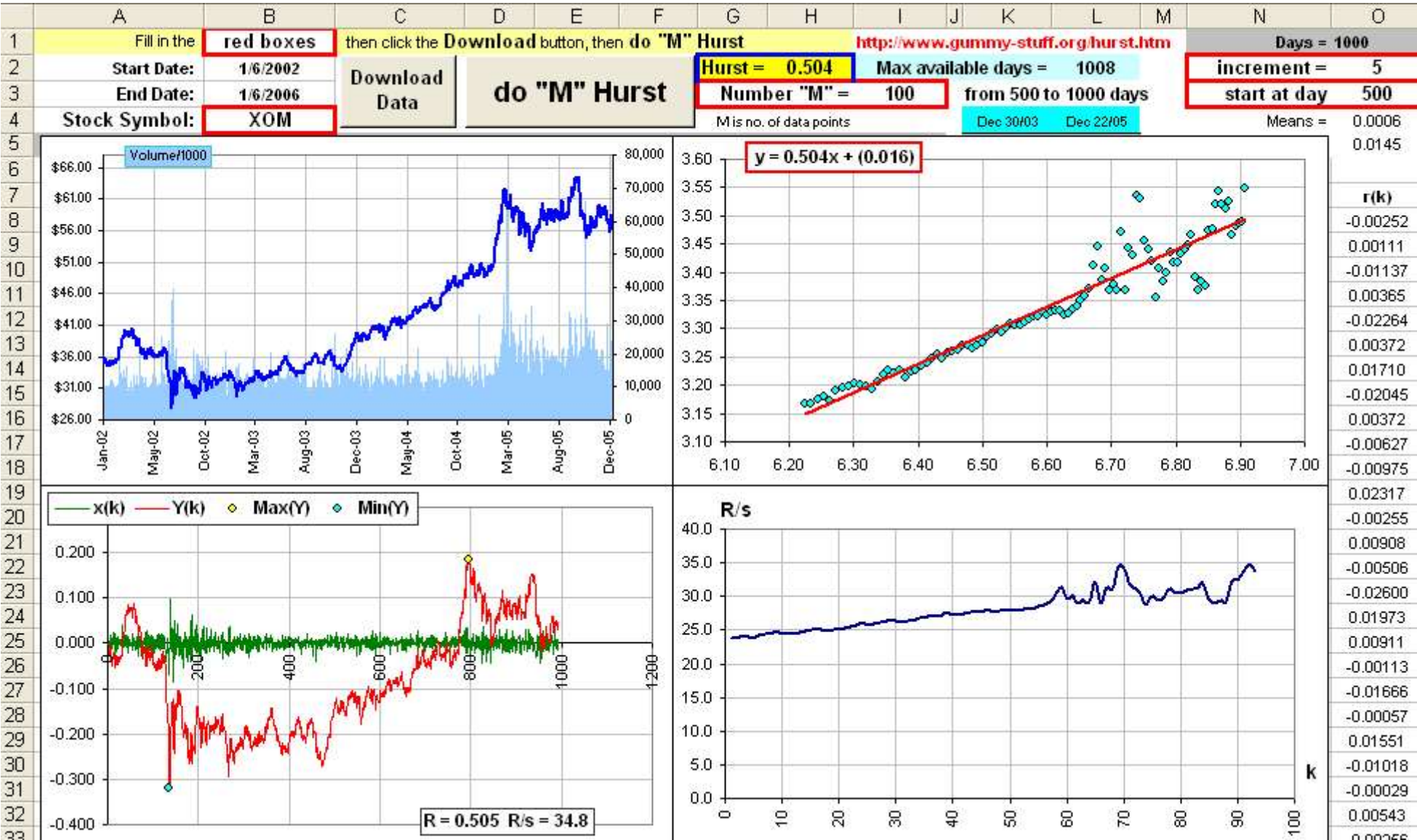
You type in some stock symbol and click the **Download Data** button to get (about) a thousand daily returns.
Then you pick a number **"M"** (example: M = **100**) and click the **do "M" Hurst** button and watch the M points get plotted.

>Huh?
The spreadsheet starts with some n (cell O3) and increments this by whatever you stuck in cell O2 (example: **5**) and, for each n, calculates a point.

>And I click the picture above to download the spreadsheet?
Well, yes, but I should warn you : it's my understanding of Hurst stuff and I can't guarantee it's validity.

In fact, it'll probably change dramatically over the next week or three 😳

>Hey! I notice that, in the picture of the spreadsheet, the Y(k) look like the stock price!
Yeah, I noticed that, too ...

the **Hurst Chart**

Note that, if the daily returns are **r**(1), **r**(2), ... **r**(n), then the stock prices change according to:
$P(1) = Po(1+\mathbf{r}(1))$
$P(2) = Po(1+\mathbf{r}(1))(1+\mathbf{r}(2))$
$P(3) = Po(1+\mathbf{r}(1))(1+\mathbf{r}(2))(1+\mathbf{r}(3))$
etc. etc.
$P(k) = Po(1+\mathbf{r}(1))(1+\mathbf{r}(2))(1+\mathbf{r}(3))... (1+\mathbf{r}(k))$

If we multiply, we'd get:
$P(k) = Po(1+\mathbf{r}(1)+\mathbf{r}(2)+...+\mathbf{r}(k) + \mathbf{r}(1)\mathbf{r}(2)+\mathbf{r}(1)\mathbf{r}(3)+...)$

>Huh?
We get all the product terms, 1-at-a-time, 2-at-a-time, then 3-at-a-time and so on.
For example, $(1+x)(1+y)(1+z) = 1 + (x+y+z) + (xy+xz+yz) + xyz$.
See? 1-at-a-time, then 2-at-a-time, then 3-at-a-time.
Anyway, note that
$\mathbf{Y}(k) = \mathbf{r}(1)+\mathbf{r}(2)+...+\mathbf{r}(k) - k\mathbf{M}$

Hence we can write
$P(k) = Po(1 + \mathbf{Y}(k) + k\mathbf{M} + \textit{other small stuff})$

I say *small stuff* because the daily returns are very small, say r = 0.01, so the products 2-at-a-time, might be $(0.01)(0.01) = 0.0001$ and so on.
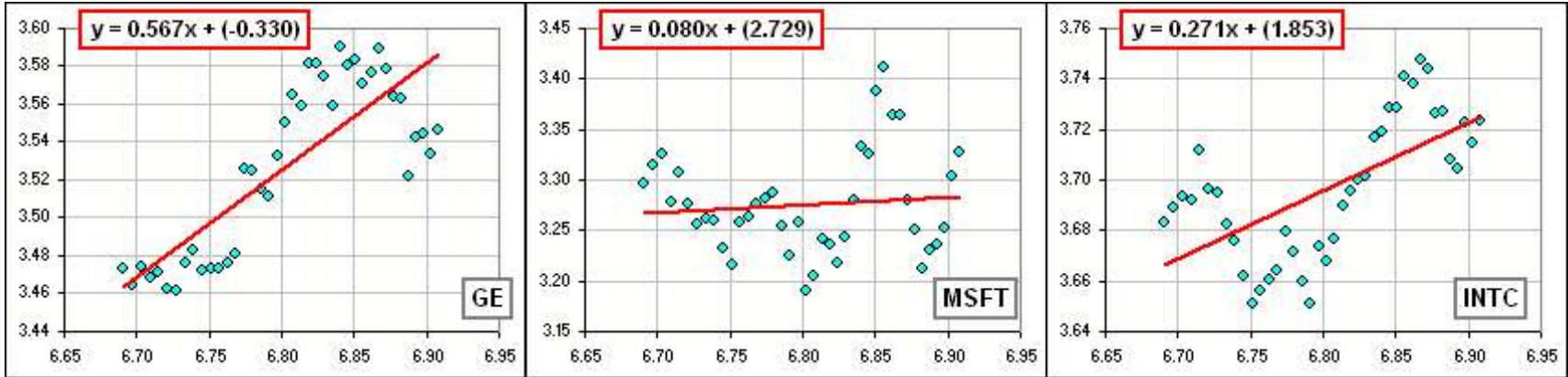
In any case, we see that P(k)/Po **-** 1 and $\mathbf{Y}(k) + k\mathbf{M}$ behave similarly.
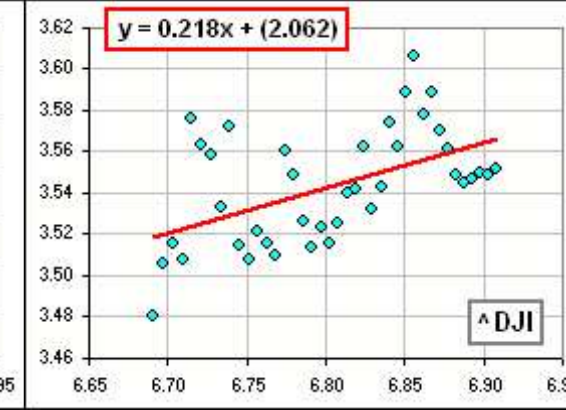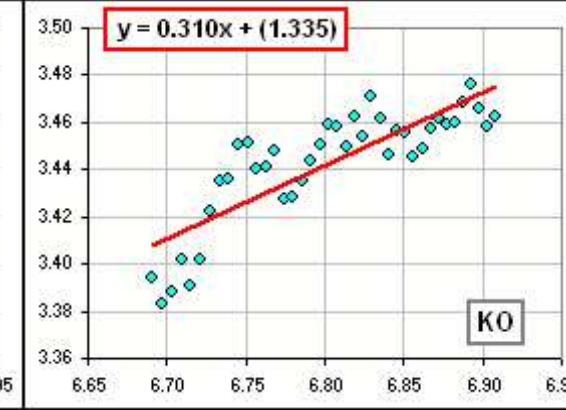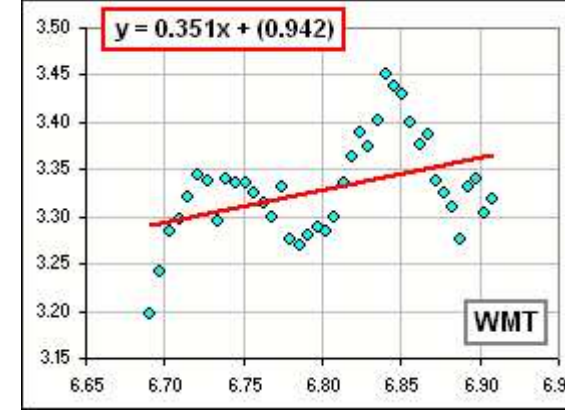
>Except for that *other small stuff*.
Yes.

Here are some examples, each calculated in late 2005, calculating 40 data points using the last n = 800 to n = 1000 days (increment = 5 days).
Actually, I used the daily returns from March 10, 2005 to Dec 22, 2005.
The first 5 returns, then 10, then 15, *etc.* until I had used all 200 returns, generating 40 data points for the scatter plot.

>The chart where the slope is the Hurst Exponent, right?
The chart where the slope is <u>an estimate</u> of the Hurst Exponent 😊

Anyway, they look like this:

for Part III

gummy stuff