Journal of Mathematical Problems, Equations and Statistics 2022; 3(2): 120-125

Journal of Mathematical Problems, Equations and Statistics

E-ISSN: 2709-9407
P-ISSN: 2709-9393
JMPES 2022; 3(2): 120-125

www.mathematicaljournal.com
Received: 14-04-2022
Accepted: 19-05-2022

Phong Luu
Department of Mathematics, University of North Georgia, Georgia, USA

Noah Yoon
Columbus High School, Columbus, Georgia, USA

Corresponding Author:
Phong Luu
Department of Mathematics, University of North Georgia, Georgia, USA

Classifying the time series of stocks into different types of stochastic models

Phong Luu and Noah Yoon

DOI: https://doi.org/10.22271/math.2022.v3.i2b.68

Abstract
Different trading strategies have been developed for different types of models. So being able to identify the asset's models is crucial in the success of the strategies. In this paper, the behavior and the identification of the time series of stocks into Geometric Brownian Motion, Mean Reversion, or Trend Following processes will be studied. In particular, the Hurst Exponent method will be implemented to classify the stochastic processes, and numerical examples are reported to demonstrate the technique.

Keywords: Geometric Brownian motion, mean reversion, trend following, Hurst exponent

Introduction
In connection with trading strategies, Hurst [3] studied the long-term storage capacity of reservoirs, but the results can also be applied for a number of other natural systems. While looking for a way to model the levels of the river Nile so that architects could construct an appropriately sized reservoir system, he gave life to a statistical methodology for distinguishing random from non-random systems and to identify the persistence of trends. Many years later, while investigating the fractal nature of financial markets – specifically, the tendency of a time series to regress strongly to its mean or to cluster in a direction – Benoit Mandelbrot [1] happened to come across Hurst's work and introduced to fractal geometry, in Hurst's honor, the term Generalized Hurst Exponent which is used as a measure of the long-term memory of a time series [4].

In this paper, we will study different types of stochastic models and implement a statistical test called the Hurst Exponent to identify them. This paper is organized as follows. In §2, we describe the Hurst Exponent method and simulation results. In §3, we demonstrate numerical examples for identifying the underlying stocks in financial market. Finally, in §4, we conclude the paper by making some remarks.

2. Classification of stochastic processes
2.1 The Hurst Exponent
The Hurst exponent is used as a measure of long-term memory of time series. It relates to the autocorrelations of the time series and the rate at which these decrease as the lag between pairs of values increases.

The goal of the Hurst Exponent is to provide us with a scalar value that will help us to identify (within the limits of statistical estimation) whether a series is mean reverting, random walking or trending.

The idea behind the Hurst Exponent calculation is that we can use the variance of a price series to assess the rate of diffusive behavior [2]. The volatility, sampled at intervals of τ, is

$$\text{Volatility}(\tau) = (\text{Variance}(X_t - X_{t-\tau}))^{1/2}$$

Denote $V(\tau) = \text{Variance}(X_t - X_{t-\tau})$, the variance over many sample times.

If Xt follows a geometric random walk, $V(\tau) = \tau$. If Xt is mean reverting, it does not wander away from their initial value as fast as a random walk. If Xt is trending, it wanders away faster. In general, we can write

$$V(\tau) = \tau^{2H}$$

where H is called the Hurst exponent, and it is equal to 0.5 for a true geometric random walk, but will be less than 0.5 for mean reverting asset, and greater than 0.5 for trending asset.
Taking the logarithm on both sides of equation previous gives

$$\text{Log}\,(V(\tau)) = 2H \log(\tau)$$

Now we can estimate 2H by taking multiple values for τ, and perform simple linear regression of log (V (τ)) on log (τ).

2.2 Simulation Results and Discussion

In this section, we consider a numerical example with the following specifications:

1. The values of lags τ ranging from 2 through 99 are used in the estimation of the Hurst Exponent H.
2. Simulating the Geometric Brownian Motion (GBM), Mean-Reverting (MR), and Trend-Following (TR) processes using sample sizes n = 200, 300, 400, 500, 1000, 2000, 5000, 10000, 20000.
3. For each sample size, generating 1000 sample paths of GBM, 1000 sample paths of MR, and 1000 sample paths of TR, and estimating the Hurst Exponents for each sample path respectively.
4. Calculating the 95% confidence intervals for the Hurst Exponents of GBM, MR, and TR.

We obtain the following result of the 95% confidence intervals for the Hurst Exponents.

Table 1: 95% Confidence intervals for the Hurst Exponents of GBM, MR, and TR with varying sample sizes.

Sample Size	Process Type	Lower Bound	Upper Bound
200	GMB	0.280925	0.297578
200	MR	-0.002278	-0.001473
200	TR	0.307565	0.325563
300	GMB	0.372954	0.388044
300	MR	-0.000809	-8.00E-06
300	TR	0.421566	0.437729
400	GMB	0.41422	0.427578
400	MR	-0.000614	1.60E-05
400	TR	0.494553	0.508376
500	GMB	0.432579	0.44469
500	MR	-0.000346	0.000175
500	TR	0.540601	0.552845
1000	GMB	0.467533	0.475431
1000	MR	-0.00028	3.20E-05
1000	TR	0.6941	0.700008
2000	GMB	0.483536	0.489045
2000	MR	-8.30E-05	0.00011
2000	TR	0.80481	0.807551
5000	GMB	0.495232	0.498492
5000	MR	-4.20E-05	6.80E-05
5000	TR	0.887544	0.8886
10000	GMB	0.496677	0.49902
10000	MR	-6.80E-05	8.00E-06
10000	TR	0.91912	0.919791
20000	GMB	0.49763	0.499346
20000	MR	-2.00E-05	3.20E-05
20000	TR	0.93712	0.937601

It can be seen from the previous table that as the sample size increases, the confidence interval for the Hurst Exponent gets closer to 0.5 for GBM, to 0 for MR, and to 1 for TR. Moreover, the classification of the processes will be clearer if the sample size is at least 1000. The following example uses the sample size of 1000. The Hurst Exponents for GBM, MR, and TR are 0.493981, 0.002859, and 0.746335 respectively.

Geometric Brownian motion

There is no correlation between the observations and a future observation; being higher or lower than the current observation is equally likely. Series of this kind are hard to predict. The Hurst Exponent for the data plotted below is estimated to be 0.493981.

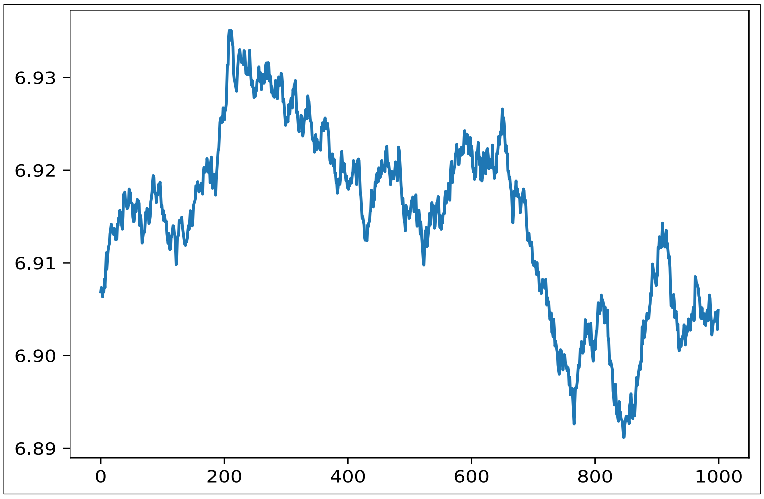

Fig 1: GBM series (H = 0.493981).

Mean Reversion: An increase will most likely be followed by a decrease or vice-versa (i.e., values will tend to revert to a mean). This means that future values have tendency to return to a long-term mean. The Hurst Exponent for the data plotted below is estimated to be 0.002859.

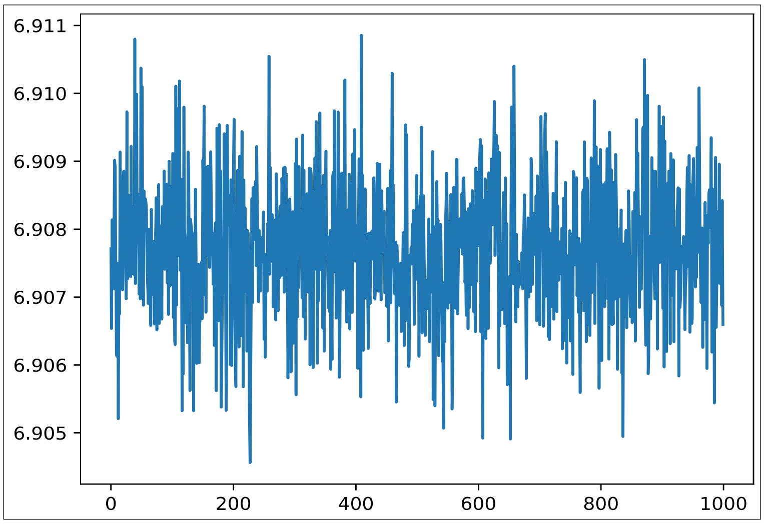

Fig 2: MR series (H = 0.002859).

Trend Following: An increase in values will most likely be followed by an increase in the short term and a decrease in values will most likely be followed by another decrease in the short term. The Hurst Exponent for the data plotted below is estimated to be 0.746335.

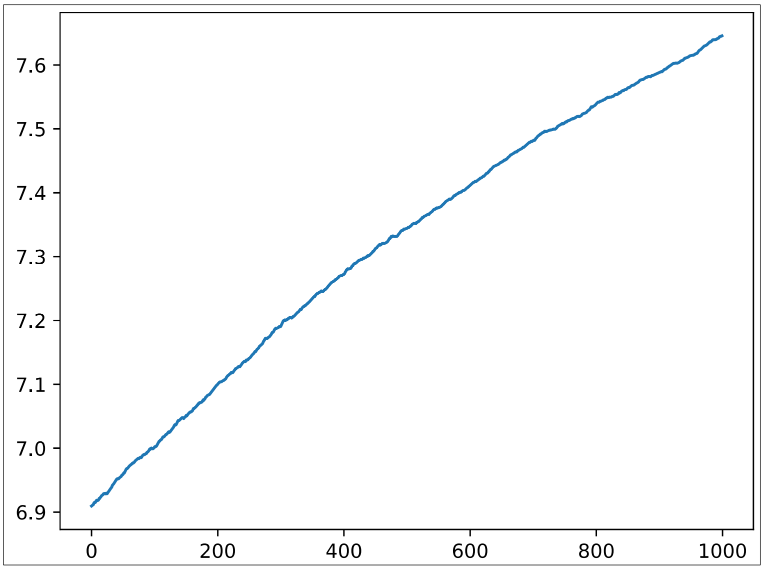

Fig 3: TR series (H = 0.746335).

3 Examples of real stocks
In this section, we will classify the time series of technology stocks. The data is collected from Yahoo Finance from 10-01-2016 to 10-01-2021 (1258 data points). We will compute the Hurst Exponents for 305 technology companies.

The following output is arranged in ascending order of the Hurst Exponents:

Table 2: Hurst Exponents for Technology Stocks

Symbol	Hurst Exponent	Symbol	Hurst Exponent	Symbol	Hurst Exponent	Symbol	Hurst Exponent
JNPR	0.334706118	BOX	0.394370916	ABST	0.417690871	ECOM	0.437873721
AMSWA	0.336183201	PRGS	0.394400664	LSCC	0.420516338	OSGX	0.438301922
MLAB	0.336191802	MNDT	0.394477959	TXN	0.421486409	PLXS	0.438802646
FIS	0.339895957	IBM	0.394838879	CTSH	0.422374495	ADBE	0.439231335
ZIXI	0.340448394	NPTN	0.395746382	SQ	0.423112846	PANW	0.439775044
CACI	0.341716154	DIOD	0.395977948	G	0.42339921	VIAV	0.439968679
HCKT	0.342936389	NOW	0.396435133	RDWR	0.424161965	AMKR	0.440301968
QLYS	0.343534373	AKAM	0.397532001	MODN	0.424245589	FLT	0.440361187
SAIC	0.347420315	ALLT	0.397838646	SSYS	0.424295456	SATS	0.440427225
AUDC	0.350904394	NLOK	0.39809933	AMD	0.424337018	UEIC	0.440921836
LDOS	0.351830257	MINDP	0.399118955	AZPN	0.424646243	DOX	0.441161488
SPI	0.35559532	CNXN	0.400303737	UI	0.424804385	FTV	0.441216823
MANT	0.364533954	SCWX	0.400498148	LPSN	0.424861532	MG	0.441610359
GDDY	0.365119774	ANY	0.400614651	TER	0.425205245	TDC	0.442447639
CCMP	0.366203514	MCHP	0.401693311	SSNC	0.425536033	ACLS	0.443743839
MSFT	0.366580535	INSG	0.402252058	DDD	0.425679347	NTGR	0.443886256
VRSN	0.366987076	INTC	0.402652288	SANM	0.426339304	EXLS	0.444229472
WDAY	0.367223202	CHKP	0.403090838	AVNW	0.42639507	EVTC	0.445367524
ADSK	0.367469804	ERIC	0.40325119	GPRO	0.426926067	DLB	0.445444738
NATI	0.368913658	SHOP	0.403629541	AVGO	0.427572433	CEVA	0.445490724
PLAB	0.369035019	LIEN	0.404391104	SWKS	0.427606869	SPNS	0.446975789
FICO	0.369745379	SCSC	0.40497972	ANET	0.427927866	VMW	0.447267466
GSIT	0.370071675	GILT	0.406413418	PAYC	0.428002027	MANN	0.447406502
FORTY	0.370300155	GLOB	0.406441933	HUBS	0.428183802	PLUS	0.448309167
NTCT	0.370361042	SLAB	0.407472031	STM	0.428854858	CDW	0.448365218
FISV	0.372419986	ORCL	0.407768903	AMBA	0.430472743	IMOS	0.448368681
ADI	0.372627842	SMCI	0.408881065	CSGS	0.430951058	BELFA	0.448656413
FN	0.374598771	KLIC	0.40905968	WIX	0.431510905	SONY	0.449969175
OM	0.376558741	ITRI	0.409943971	CTXS	0.432163664	AXTI	0.450459182
CRM	0.379468165	LYTS	0.410118506	RNG	0.43300251	VRNS	0.450659689
TYL	0.380198478	GWRE	0.411944347	CALX	0.433678638	INTU	0.451013435
INS	0.383264972	FARO	0.41366699	LRCX	0.434280461	MPWR	0.45110989
MGIC	0.384434963	SNPS	0.414938454	CDNS	0.434692591	KVHI	0.451932779
SEDG	0.387297783	ESE	0.415443528	PDFS	0.435942132	TEAM	0.45240956
SMTC	0.387808668	NICE	0.415635593	NOK	0.436841877	LOGI	0.452656462
BRKS	0.388421672	MIME	0.416277163	AGYS	0.437176168	PCP/	0.452701927
ENTG	0.388498714	H LIT	0.416630603	MRVL	0.437526271	PCT]	0.452808889
ATEN	0.38851321	KLAC	0.416723081	QRVO	0.437629757	CLS	0.454260292
PEGA	0.392509902	ENV	0.417636196	SPLK	0.437692632	DGII	0.454657405

Table 2: Hurst Exponents for Technology Stocks (continued)

Symbol	Hurst Exponent	Symbol	Hurst Exponent	Symbol	Hurst Exponent	Symbol	Hurst Exponent
MSI	0.455050542	MEI	0.472941285	SPSC	0.490849614	CTG	0.520059247
CRUS	0.455099457	AAPL	0.473355508	VSAT	0.491845025	SILC	0.52014837
BKI	0.455389537	ITRN	0.473537169	TEL	0.492085298	POLY	0.520657859
ACN	0.456082741	SWIR	0.47418943	LPL	0.492565947	WIT	0.520980798
MSTR	0.456198541	SANS	0.474291095	SPWR	0.492769504	SPAY	0.521524574
SIMO	0.457944092	QTWO	0.474305137	GRMN	0.493407338	VICR	0.521525158
CSIQ	0.457945299	MXL	0.475132126	VPG	0.49401456	INFY	0.525014043
ZEN	0.458022592	MRS	0.475285949	NUAN	0.494375041	EGHT	0.526752126
NTNX	0.458581346	ZBRA	0.476960192	CtCOM	0.494870704	RELL	0.528139887
UEPS	0.459306525	VCRA	0.477076006	ALRM	0.495283955	VNET	0.528946592
RMBS	0.459518313	CAMP	0.477367261	CVLT	0.496268732	BCOV	0.529069309
APH	0.459633099	UM!	0.47782503	GLW	0.497080166	TESS	0.530156183
IMMR	0.460798776	ADTN	0.477886787	XS	0.49793513	TSEM	0.531314481
PAR	0.461392703	WDC	0.47835578	COMM	0.498074283	PRO	0.531348821
NXPI	0.461428007	RUN	0.479647426	ASX	0.498220746	SNX	0.531642532
ST	0.46245838	DO	0.479769239	FLEX	0.502606916	11EC	0.532655602
POW!	0.462667876	MIS!	0.479926713	JBL	0.504133161	NTAP	0.533400466

WNS	0.462874468	ROG	0.480145463	UIS	0.504605858	COHU	0.533598638
NVMI	0.463341053	XPER	0.481740322	DBD	0.504611264	IT	0.534318234
PTC	0.463458119	PI	0.482310107	HPE	0.505137018	EEFT	0.535360736
CGNX	0.463874505	ON	0.482312296	AOSL	0.506868139	TSM	0.535737062
CSCO	0.464390539	FSLR	0.482674876	SOL	0.507278583	XRX	0.536644389
ACIW	0.464741841	MU	0.482780721	DMRC	0.508016632	VRNT	0.537120245
EBIX	0.465315922	ASYS	0.482869816	BELFB	0.508248514	MAXR	0.537567115
MKSI	0.465693301	AV!	0.483061265	OJWI	0.510044304	SABR	0.54019565
INFN	0.465918099	IIVI	0.483144414	KN	0.510052359	CMTL	0.541519669
XLNX	0.466217173	MIXT	0.48315131	NCR	0.510351978	CAI	0.548470108
CYBR	0.466880433	CTS	0.483532725	OSIS	0.510770568	EPAM	0.548854874
MX	0.46697039	FFIV	0.485181802	NVDA	0.510819949	COHR	0.549525923
ARW	0.467071298	SAP	0.486218116	AMAT	0.511536397	DXC	0.551322741
TRMB	0.467396067	BLKB	0.486293541	UMC	0.513139441	BTCM	0.55169825
KEYS	0.467993404	WOLF	0.48635877	NSIT	0.513259958	OHM	0.555610988
VECO	0.468064247	IPGP	0.48698401	BR	0.51609893	FTNT	0.562856566
DELL	0.469100802	GIB	0.487861502	CDK	0.516626348	WK	0.571635394
HIMX	0.470151524	HPQ	0.488405925	DCTR	0.516795915	PRFT	0.591091772
UCTT	0.47039168	WEX	0.489828456	SD(	0.517094576		
LFUS	0.472007901	VSH	0.48995011	VOXX	0.519158111		
FORM	0.472783694	NEWR	0.490026503	CTLP	0.519698792		

It can be seen from the previous table that technology stocks seem to follow GBM (all Hurst Exponents are not far from 0.5). The JNPR (Juniper Networks, Inc.) with H = 0.334706118 is the closest to MR while the PRFT (Perficient, Inc.) with H = 0.591091772 is the closest to TR. The ASX (ASE Technology Holding Co., Ltd.) with moderate H = 0.498220746 is closest to GMB. Followings are the plots of the time series of these stocks.

JNPR has the smallest Hurst Exponent, which is close to Mean Reversion.

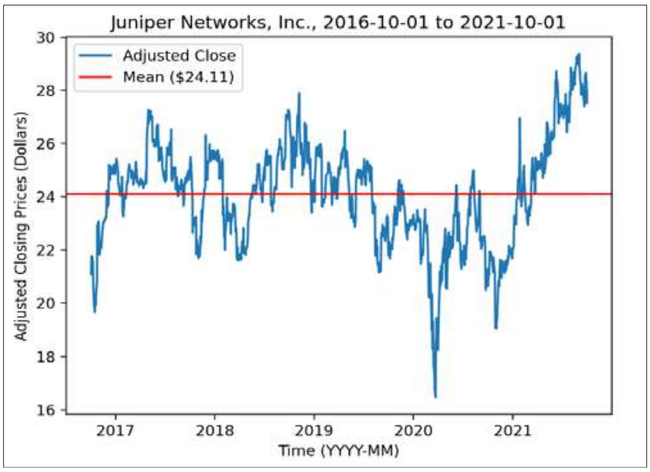

Fig 4: JNPR (H = 0.334706118)

PRFT has the highest Hurst Exponent, which is close to Trend Following.

Fig 5: PRFT (H = 0.591091772)

ASX has a moderate Hurst Exponent, which is close to Geometric Brownian Motion.

Fig 6: ASX (H = 0.498220746)

4. Conclusion

The Hurst Exponent is a useful statistical method for inferring the properties of a time series without making assumptions about stationarity. It is most useful when used in conjunction with other techniques, and has been applied in a wide range of industries. For example, the Hurst Exponent is paired with technical indicators to make decisions about trading securities. However, as can be seen from the examples of technology stocks, mean reverting processes are not common in financial market. So being able to identify such processes is much valuable.

Despite the attractive theory, utilizing the Hurst Exponent estimation requires some deeper thought and analysis than simply plugging some numbers into an algorithm. In the implementation above, we used lags 2-99. This is completely arbitrary. However, if we carefully investigate, we can obtain some potentially very useful insights. The Hurst Exponent will vary depending on the sample sizes. It has been shown in the numerical reports that the Hurst Exponent can be used to clearly classify time series as GBM, MR, or TR when the sample size is sufficiently large. Moreover, the sufficient sample size can vary for different values of lags being used.

5. References

1. Benoit MB, Hudson RL. The (Mis). Behavior of Markets, A Fractal View of Risk, Ruin and Reward, Basic Books, New York; c2004.
2. Chan E. Mean Reversion, Momentum, and Volatility Term Structure. Quantitative Trading; c2016.
3. Hurst HE. Long-term Storage of Reservoirs: an Experimental Study. Transactions of the American Society of Civil Engineers. 1951;116:770-799.
4. Mansukhani S. The Hurst Exponent: Predictability of Time Series. Analytics; c2012.
5. Øksendal B. Stochastic Differential Equations, 6th Ed., Springer-Verlag, New York; c2003.