

# ポートフォリオ再構成アルゴリズム OLMAR が 実際の株価データにおいて超過収益が得られる要因

## Why portfolio reconstruction algorithm OLMAR gets excessive revenue?

海野 一則<sup>1</sup> 菊地 剛正<sup>1</sup> 國上 真章<sup>1</sup> 山田 隆志<sup>2</sup> 寺野隆雄<sup>1</sup>

Kazunori Umino<sup>1</sup>, Takamasa Kikuchi<sup>1</sup>, Masaaki Kunigami<sup>1</sup>, Takashi Yanada<sup>2</sup> and Takao Terano<sup>1</sup>

<sup>1</sup> 東京工業大学

<sup>1</sup>Tokyo Institute of Technology

<sup>2</sup> 山口大学

<sup>2</sup> Yamaguchi University

**Abstract:** 本研究は、高いパフォーマンスを示すとされている OLMAR 法を対象に、超過収益を得られる要因を分析し、なぜ長期にわたって優位性を保つことができているのかを検証した。我々は、ポートフォリオ再構築アルゴリズムである OLMAR 法が、広範囲な株式市場で長期にわたって機能するためには、株価の変動において、長期に持続し、変動の上昇・下落の方向性が利用可能な、特定の変動特性が存在すると考えた。そして、ポートフォリオ再構成アルゴリズムは、特定の変動特性を積極的に利用する必要があるはずである。本研究は、そのような仮定をもとに、データセットに注目した分析を行った。本論文では、新たに用意した日米の長期株式データにおいて、1) OLMAR 法で超過収益が得られること。2) OLMAR 法が基本とする移動平均乖離率を用いて日米の長期株式データにおいて特定の変動特性が存在すること。3) 特定の変動特性を無効化したデータでは、OLMAR 法によって超過収益が得られないことを確認した。

## 1. はじめに

本研究の目的は、数十年以上の実際の株価データセットにおいて、取引コストを考慮してもインデックスをアウトパフォーマンスしている OLMAR (On-Line Moving Average Reversion)法に注目し、なぜ超過収益を得られるのかを分析することである[1]。OLMAR 法は、移動平均から一時的に大きく下落(乖離)した株式が、移動平均に近づく現象により超過収益を得るポートフォリオ再構成手法である[2][3]。

本稿の分析内容は、1) OLMAR 法が用いている手法を価格時系列データに適用することにより、実際の株価データセットにおいて何らかの株価の変動特性が検出可能か?、2) 何らかの株価の変動特性により超過収益が得られるのであれば、その特性をできる限り小さくすることで OLMAR のパフォーマンスはどう変化するのか? の 2 点である。

多くの研究者により検証が行われ、高いパフォーマンスが得られることが報告されている[3][4][5][6][7]。

しかしながら、決められたデータセットに対して、最適化を行いパフォーマンスを上げることに注力しており、データセットそのものに対する分析は行わ

れていない[3][6][7][8]。株価変動がランダムに近ければ、超過収益を長期にわたりアウトパフォームすることは困難であり、OLMAR 法が何らかの株価変動の特性を利用しているとすれば、今までに明らかにされてこなかった株価変動の特性と、それを利用するテクニックが存在するはずである。

OLMAR は優秀なアルゴリズムでありながら、“なぜ、超過収益が得られるのか?”については研究されていない。本研究では、株価変動の特性を分析し、OLMAR 法が安定的に超過収益を上げるメカニズムの解明を試みる。

本論文の主要な貢献は次のとおりである。

- 新たなデータセット用意し、OLMAR のアルゴリズムが有効であることを示し、データセットが長期に持続する株価変動特性を持つことを明らかにする。
- OLMAR 法が利用する特定の変動特性を無効化する方法を示す。
- OLMAR による超過収益が、本研究で明確にした株価変動特性に依存していることを示す。

本論文では、OLMAR 法の基本となっている移動平均乖離率を適用した場合の株価変動特性について

の分析を行い、超過収益が発生するメカニズムを分析する。

2章では本論文で用いる指標およびOLMAR法に関する説明を行い、3章で本論文の分析手法を説明し、4章で新たに用意した複数のデータセットの検証結果を示し、5章においてまとめとする。

## 2. 関連研究

### 2・1 本論文で使用する指標

時刻  $t$  における資産の価格を  $P_t$  とすると、この資産を時刻  $t-1$  から時刻  $t$  まで保持した時のリターン  $R_t$  は次のようになる。

$$R_t = \frac{P_t}{P_{t-1}},$$

以下、本論文で用いる指標について説明する。

(MAD: Moving Average Deviation rate 移動平均乖離率)

$$\text{simpleMA}_t^n = \frac{1}{n} \sum_{k=1}^n P_{t-k+1},$$

$$\text{MAD}_t^n = 100(P_t - \text{simpleMA}_t^n) / P_t,$$

(AR: Annualized Return)

$$R^{day} = \frac{1}{(n-1)} \sum_{k=2}^n \left( \frac{P_k}{P_{k-1}} \right),$$

$$AR = (R^{day})^{252},$$

(MDD: Maximum Draw Down)

$$M_t = \max_{u \in [0, t]} P_u,$$

$$D_t = M_t - P_t,$$

$$\text{MDD}_t = \max_{u \in [0, t]} D_u,$$

### 2・2 ポートフォリオ再構成手法

今まで平均と分散をコントロールすることから、より積極的に価格変動の特性を生かしたアルゴリズムが開発された。その中で最も高いパフォーマンスを示した手法が **Follow-the-Loser** というカテゴリーに属する手法であり、その代表的手法が **OLMAR** 法であると報告されている[3][9]。

**Follow-the-Loser** とは、負け組(価格が下落している)株式群を常に選択して、価格が下落から上昇に転じて平均回帰することを期待する手法である。

**OLMAR** 法は、移動平均乖離率(MAD)の高い株式群を常に選択する手法であり、全てのデータセットで最もリターンの高い手法であると言われている。ただし、欠点としては、

1) MDD(最大ドロウダウン)が大きく、一時的な資産

価格の下落幅が大きい、2) 頻りに資産入れ替えが起こり、取引コストの僅かな増加でリターンが大幅に減少することである[2]。

## 3. 検証方法

本検証で使用する記号と意味を以下に定義する。

$N$ : ポートフォリオを構成するアセット数

$i$ : 株式  $i(i=1, \dots, N)$

$t$ : 時点  $t(t=1, \dots, S)$  ただし、全期間のデータが存在しない場合は、データが存在する時点のみ値を返す。

$R_{i,t}$ : 時点  $t$  における株式  $i$  のリターン

$\text{MAD}_{i,t}^{40}$ : 期間 40 日を用いた、時点  $t$  における株式  $i$  の移動平均乖離率

$\text{MDD}_{i,t}^{40}(t1:t2)$ : 時点  $t1$  から  $t2$  までの株式  $i$  の最大ドロウダウン

### 3・1 株価変動特性の分析

本研究では、OLMAR 法で用いられている移動平均乖離率に基づく株式のフィルタリングの有効性と、そこから超過収益が得られる株価変動特性を分析する。以下の手順となる。

<1> 株価データセットの全株式の MAD (移動平均乖離率) とリターンを求める。

<2> MAD の小さな順にソートした上で、近傍 1000~4000 点の平均を求め、その値をリターンの平均とし、リターン曲線を得る。

<3> データ全体の下位  $L=[0.25, 0.5, 1, 3, 5]$  パーセントの平均値をリターン曲線より求める。

### 3・2 株価変動特性へのデータ補正操作

本研究では、MAD により超過リターンの得られる特性が検出できれば、その特性を減少または、非常に小さくすることで **OLMAR** 法により超過収益が減少または得られないことを検証する。本研究では、移動平均乖離率によりリターン曲線に不均衡がある場合に、リターン曲線に従いデータを補正する。

<1> 期間  $k$  の  $\text{MAD}_{i,t}^k$  を求めた上で、リターン曲線を求める。上昇株式群と下落株式群の歪みを以下の式に基づき、データの値を書き換える。不均衡のある  $L=1.0(\%)$  以下の値を  $1.0$  以上のリターンのデータ  $dU$  と、それ以外の  $dD$  に分け、それぞれの標準偏差を求め、 $dU$  と  $dD$  の標準偏差が等しくなるようにデータを操作し、 $\text{modify\_dU}$  を  $dU$  に置き換える(1,2)。データを操作により、リターンの総和に差が生まれた場合は、その差である  $\text{add\_mean}$  を  $\text{MAD} < 0$  の全データ

に均等に配分する(3).

$$\text{ratio} = \frac{\text{stddev}(dD)}{\text{stddev}(dU)}, \quad (1)$$

$$\text{modify\_dU} = 1.0 + \text{ratio}(dU - 1.0), \quad (2)$$

$$\text{add\_mean} = \frac{\text{sum}(dU) - \text{sum}(\text{modify\_dU})}{(\text{MAD} < 0 \text{ の全データ数})}, \quad (3)$$

〈2〉 上記に示した補正(データの操作)を, 期間  $k=[20,40,80]$  の順で繰り返し操作を行い, 複数の移動平均に対する下位 1%の歪み(上昇・下落データの標準偏差の差)を縮小させる(全修正).

〈3〉 必要に応じて, 歪みを修正比率  $a$  に応じて修正することも可能である.(一部修正). 全修正の場合は  $a=1.0$  となり, 一部修正の場合は  $0 < a < 1.0$  となる(4).

$$\text{modify\_dU} = 1.0 + (1 - a(1 - \text{ratio}))(dU - 1.0), \quad (4)$$

### 3・3 OLMAR\_random 法

ランダムに株式を選択し, ポートフォリオを構成する平均株式数および, 株式の平均保持期間が OLMAR 法に等しい手法であり, 取引回数も同じである. OLMAR 法によるパフォーマンスの比較対象とする.

### 3・3 検証方法

本研究では, 以下の検証を行う.

- (1) 移動平均乖離率(MAD)を用い, MAD がマイナスの下落局面に注目し, 特性(リターンの不均衡: リターン曲線の歪みや傾向)を分析する. OLMAR 法は, 下落している株式のみを扱うアルゴリズムだからである[3].
- (2) リターンの不均衡があれば, それに対して補正操作を行い上昇・下落の標準偏差の違いを修正する. これを, 複数の代表的な移動平均に対して行う.
- (3) OLMAR 法および OLMAR\_random 法により, オリジナル株価データおよび補正操作を行ったデータに対してシミュレーションを行い, その違いを分析する.

## 4. 検証結果

### 4・1 データセット

本検証で用いるデータセットは, オリジナル株価データセット 2種類と, 4.2節で示すデータの補正操作を行ったものの計 4種類である. 各データセットの特徴を以下に示す(表 1). Hurst 指数の平均の違いを見ると, SP データに対する補正操作では, 平均回帰からトレンド傾向に変化しているが, NK データではほとんど変化が起こっていない[10][11].

表 1 利用データセット

補正操作により, オリジナルから特性が大きく変化している. SP\_mod では, 資産倍率が大きく上昇した株式によりリターンの平均が上がっており, NK\_mod では異なる変化が起こり, 大きく下落する銘柄が生成されている. 今後, そのようなデータの発生機序に関する分析が必要である.

データ名	データサイズ	その他の特徴
SP:S&P500	1999.1~2016.4 4359 日,391 銘柄  Hurst 指数(平均)= 0.46	リターン (平均)= 8.86 (中央値)=4.60 資産倍率= 0.06~434.4 平均 MDD= 71.6
NK:Nikkei225	1999.1~2016.4 4254 日,183 銘柄  Hurst 指数(平均)= 0.48	リターン (平均)= 2.62 (中央値)=1.62 資産倍率= 0.14~56.6 平均 MDD= 78.7
SP_mod: SP に補正操作を 行い特定の変動 特性を無効化し たデータ	1999.1~2016.4 4359 日,391 銘柄 *4.2 節参照  Hurst 指数(平均)= 0.54	リターン (平均)= 44.4 (中央値)= 11.0 資産倍率= 0.024~6098.9 平均 MDD= 70.7
NK_mod: NK に補正操作 を行い特定の 変動特性を無効化 したデータ	1999.1~2016.4 4254 日,183 銘柄 *4.2 節参照  Hurst 指数(平均)= 0.49	リターン (平均)= 1.95 (中央値)=1.21 資産倍率= 5.6×10 <sup>-6</sup> ~27.9 平均 MDD= 79.4

### 4・2 価格変動特性

SP 全株式データ(N=1688338)を, MAD とその翌日のリターンをプロットし, 拡大したものを図 1 に示し, SP および NK の MAD とリターンを(表 2,3)に示す. SP データでは, MAD が下位 1%以下であれば, 統計的に有意な高いリターンが得られることがわかった. 過去 18年かつ 391 銘柄のデータであり, 当初に仮定した長期にわたる株価変動特性が確認できた. NK データでも, 適切に移動平均を選択すれば, MAD が下位 1%以下で高いリターンを得られることが示された.

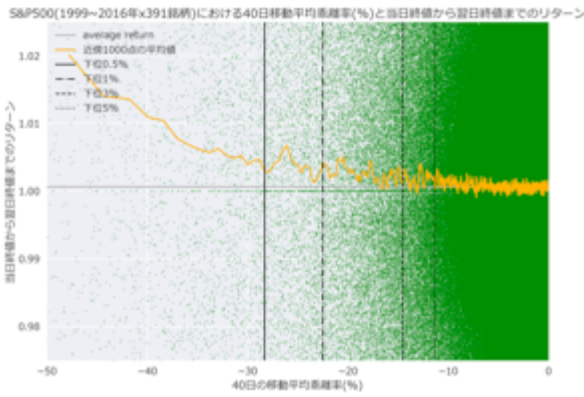


図1 SPの全データを用いた移動平均乖離率(40日のMOD)と翌日のリターンの分布。下位1.0%以下では、上昇・下落の不均衡が起り、MADの値が小さいほど、翌日のリターンは高くなっている。なお、各縦線は左より移動平均乖離率下位0.5,1,3,5%を示している。

表2 SPデータにおける40日MADと翌日のリターン:リターンがMADのマイナス幅の拡大により翌日の平均リターンが増加している。

下位%	移動平均乖離率(MAD)	当日終値から翌日終値までのリターン
0.25	-32.0	1.0055
0.5	-27.3	1.0044
1.0	-22.0	1.0033
3.0	-14.3	1.0025
5.0	-11.2	1.0014

表3 NKデータにおける40日MADと翌日のリターン:リターンがMADのマイナス幅の拡大により翌日の平均リターンが増加している。

下位%	移動平均乖離率(MAD)	当日終値から翌日終値までのリターン
0.25	-26.7	1.0058
0.5	-23.6	1.0037
1.0	-20.3	1.0040
3.0	-15.1	1.0014
5.0	-12.6	1.0009

MADにより価格変動特性を抽出可能であり、下位1%以内の外れ値と見なされる領域で持続的に高いリターンが発生しており、OLMAR法の超過収益に關係している可能性が示された。稀な状況(MADのマイナス幅の大きさ)のみに重点を置き、積極的にリスクを取ることで超過収益を得られる可能性があることが示された。

#### 4・3 価格変動のデータ補正操作

全株式データ(N=1688338)を、MADとその翌日のリターンをプロットの歪みを3.2節で示した手順により下位1%の変動の特性を減少させた結果を図2に示す。MADに対するリターンの上昇傾向が是正されている。本ケースの40日間のMAD=[-40,-45]は全体の0.1および0.025%に相当する。

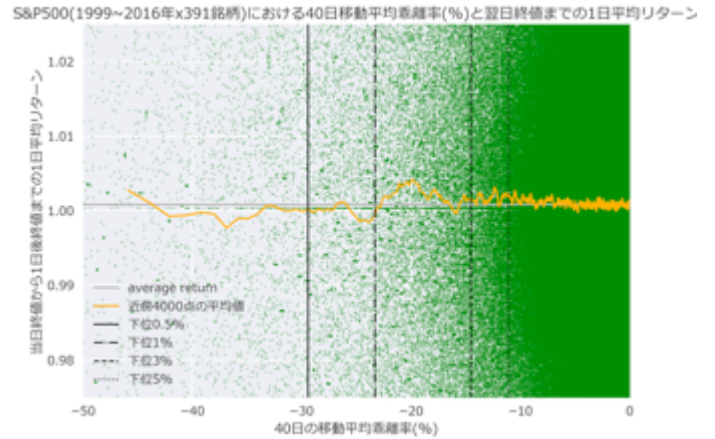


図2 SPデータにおける操作期間k=[20,40,80]の補正結果。図1と同じ条件であるが、MADに応じてリターンが上昇する傾向が補正されている。

なお、MADは、3.2<2>において、3つの期間における操作をおこなっているが、代表的な期間の補正を行うことで、操作以外の期間(例えば60,100日)などの移動平均についても有効であった(図3)。

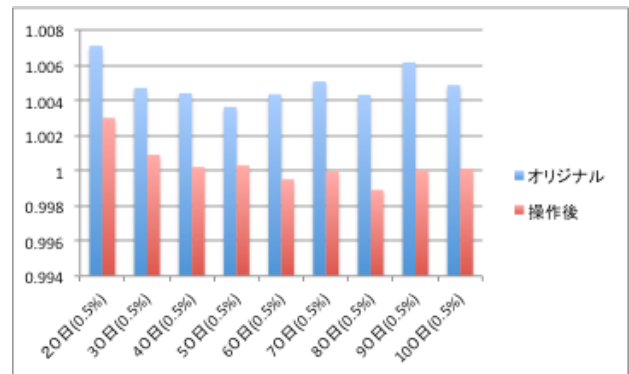


図3 SPデータにおける操作期間k=[20,40,80]以外のケースも含めて0.5%のリターンを比較する。20日以外は、下のデータのリターンの高さが補正されている。また、補正の対象とした[20, 40, 80]以外のデータも補正されている。

主要な移動平均に基づき、上昇・下落の分布における不均衡が補正されたことで、他の期間における不均衡も補正されている。ただし、短い期間(20日)の補正は精度が悪く、今後改良する必要がある。

#### 4・4 OLMAR法およびOLMAR\_random法によるシミュレーション

### < 検証1 > パフォーマンスの検証

データセットにおいてOLMAR 法およびOLMAR\_random 法により超過リターンが得られるかを検証した。取引コストを0.15%としたケースでのSPおよびSP\_modを用いたシミュレーション結果を図4,5に示す。

データ SP: インデックスおよびOLMAR,OLMAR\_randomの資産価格の変化



図4 SP データにおいて、取引コスト0.3%で40日の移動平均用いたケースの結果である。OLMAR法は、インデックスをアウトパフォーマンスし、資産倍率は170倍程度となっているが、OLMAR\_randomは取引コストもかかるため、インデックスをアンダーパフォーマンスしている。

図4の結果より、OLMAR法は18年間のデータにおいてインデックスの年平均リターンは13.4%であるのに対し、OLMAR法は34.5%となっている。ただし、最大ドローダウン(MDD)は49.5%から98.28%に拡大しており、一時的に資産価格は1/50以下になっている。これにより、論文[2]で記載されている高いドローダウンの問題が再確認された。

データ SP\_mod: インデックスおよびOLMAR,OLMAR\_randomの資産価格の)



図5 SP\_mod データにおいて、取引コスト0.3%で40日の移動平均用いたケースの結果である。OLMAR法は、インデックスが上昇しているにもかかわらず下落している。OLMAR法のように、株価の変動特性の一部を利用した手法では、インデックスの変化と関係なく優位性のみがパフォーマンスに影響することが明らかになった。よって、OLMAR法に適した変動特性を持つ株式を選択する必要がある。

次に、データ SP に補正操作を行ったデータである SP\_mod を用いたシミュレーション結果(図5)では、インデックスは40倍以上に資産が上昇しているのに、OLMAR法の資産は大きく下落している。4.1節でも述べた Hurst 指数の平均が、平均回帰傾向( $H < 0.5$ )から、トレンド傾向( $H > 0.5$ )に変化したことで、下落する株式はさらに下落する傾向が影響した可能性がある。また、リターンの特性(大きさ)のわずかな違いが、資産価格に対して指数関数的な影響を与えるため、パフォーマンスに大きな変化を与える[10][11]。

同様の検証を、データ NK および NK\_mod による結果を図6,7に示す。やはり、OLMAR法は利用していた特性がなくなるため、インデックスはより大きなリターンを示しているにもかかわらず、資産価格は1/10程度にまで下落している。SP\_mod および NK\_mod データを用いた場合は、OLMAR\_random 法では、取引コストによるパフォーマンスの低下が見られるが、インデックスの動きに追随していることが確認できた。

データ NK: インデックスおよびOLMAR,OLMAR\_randomの資産価格の変化



図6 NK\_mod データにおいて、取引コスト0.3%で110日の移動平均用いたケースの結果である。NKデータにおいても、SPデータと同様に、OLMAR法はインデックスをアウトパフォーマンスしているが、その差は小さい。

データ NK\_mod: インデックスおよびOLMAR,OLMAR\_randomの資産価格の変化



図7 NK\_mod データにおいて、取引コスト0.3%で110日の移動平均用いたケースの結果である。NK\_mod データ

においても、SP\_mod データと同様に、OLMAR 法は最も下落しており、インデックスを大きくアンダーパフォームしている。

これらの結果から、以下のことがわかった。

- 1) 新たな日米のデータセットで、OLMAR 法は取引コストを考慮しても、インデックスを上回ることが確認でき、変動特性の無効化によりその優位性がなくなった。
- 2) OLMAR 法は、移動平均乖離率を用いて、乖離率の下位 1 パーセント以下で起こるリターンが高くなる傾向を利用した手法である。
- 3) OLMAR\_random 法による資産価格は、インデックスに追従するが、OLMAR 法は非常に限られた特性を用いる手法のため、その性質を操作することで、全く異なる価格変動を起こす。図 5 が表すように、OLMAR 法のパフォーマンスが、MAD の下位 1% の変動特性に依存していることがわかった。
- 4) OLMAR 法は、高いリターンを示す場合であっても、最大ドローダウンリスクが大きく、1/50 まです資産価格が下がる可能性がありリスクの高い手法であることが再確認された。

## 5. まとめ

本研究では、OLMAR 法が何らかのロバストな価格変動特性に基づき超過収益を得られていることを示した。

- 1) 日米の長期の株価データセットにおいて、移動平均乖離率が大きくマイナスの状況(下位 1%以下)では、翌日のリターンが大きくなる傾向が日米の株式データセットで見られた。
- 2) OLMAR 法は、高い頻度で移動平均乖離率が大きなマイナス値を示す株式を選択することで、超過収益を得ている。
- 3) 超過収益が得られる特性を、補正操作(無効化)を行うことで、OLMAR 法の優位性がなくなることを確認した。これは、移動平均乖離率の下位 1%以下のリターンが高くなる特性に依存していることがわかった。

今後の課題としては、補正操作の精度を上げること。また、移動平均乖離とリターンの関係をより詳細に分析することが挙げられる。今後 OLMAR 法による選択の特性と最大ドローダウンが大きくなる問題について分析したいと考えている。

## 参考文献

- [1] Li, Bin, and Steven CH Hoi. "On-line portfolio selection with moving average reversion." *arXiv preprint arXiv:1206.4626* (2012).
- [2] Li, Bin, et al. "Moving average reversion strategy for on-line portfolio selection." *Artificial Intelligence* 222 (2015): 104-123.
- [3] Li, Bin, and Steven CH Hoi. "Online portfolio selection: A survey." *ACM Computing Surveys (CSUR)* 46.3 (2014): 35.
- [4] Li, Bin, Doyen Sahoo, and Steven CH Hoi. "OLPS: a toolbox for on-line portfolio selection." *Journal of Machine Learning Research* 17.35 (2016): 1-5.
- [5] Nyikosa, Favour M., Michael A. Osborne, and Stephen J. Roberts. "Adaptive Bayesian Optimisation for Online Portfolio Selection." *Workshop on Bayesian Optimization at NIPS*. Vol. 2015. 2015.
- [6] Ha, Youngmin. "Online portfolio selection with transaction costs including market impact costs." *Browser Download This Paper* (2016).
- [7] Gao, Li, and Weiguo Zhang. "Weighted Moving Average Passive Aggressive Algorithm for Online Portfolio Selection." *Intelligent Human-Machine Systems and Cybernetics (IHMSC), 2013 5th International Conference on*. Vol. 1. IEEE, 2013.
- [8] Lin, Xiao, et al. "Boosting Moving Average Reversion Strategy for Online Portfolio Selection: A Meta-learning Approach." *International Conference on Database Systems for Advanced Applications*. Springer, Cham, 2017.
- [9] Kroll, Yoram, Haim Levy, and Harry M. Markowitz. "Mean - variance versus direct utility maximization." *The Journal of Finance* 39.1 (1984): 47-61.
- [10] Rasheed, Bo Qian Khaled, and B. Qian. "Hurst exponent and financial market predictability." *Iasted conference on Financial Engineering and Applications (FEA 2004)*. 2004.
- [11] Carbone, Anna, Giuliano Castelli, and H. Eugene Stanley. "Time-dependent Hurst exponent in financial time series." *Physica A: Statistical Mechanics and its Applications* 344.1 (2004): 267-271.