UNIVERSIDADE DE SÃO PAULO - USP
ESCOLA DE ARTES, CIÊNCIAS E HUMANIDADES - EACH
PROGRAMA DE MESTRADO EM MODELAGEM DE SISTEMAS COMPLEXOS

EDER LUCIO DA FONSECA

O estudo das propriedades multifractais de séries temporais financeiras

Dissertação apresentada como requisito para a obtenção do Título de Mestre em Ciências.

Orientador: Fernando Fagundes Ferreira

SÃO PAULO
2012

EDER LUCIO DA FONSECA

O estudo das propriedades multifractais de séries temporais financeiras

Comissão julgadora

Disertação para obtenção do grau de mestre

1. Professor Doutor Fernando Fagundes Ferreira (Presidente da Banca)
   Escola de Artes, Ciências e Humanidades - USP
   Assinatura: _____

2. Professor Doutor Camilo Rodrigues Neto
   Escola de Artes, Ciências e Humanidades - USP
   Assinatura: _____

3. Professor Doutor Antônio Fernando Crepaldi
   Faculdade de Engenharia - UNESP Bauru
   Assinatura: _____

[Versão Corrigida]

SÃO PAULO, 01 DE MARÇO

2012

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

CATALOGAÇÃO-NA-PUBLICAÇÃO
Biblioteca
Escola de Artes, Ciências e Humanidades da Universidade de São Paulo

Fonseca, Eder Lucio da

O estudo das propriedades multifractais de séries temporais financeiras / Eder Lucio da Fonseca ; orientador, Fernando Fagundes Ferreira. - São Paulo, 2012.

86 f.: il.

Dissertação (Mestrado em Ciências) - Programa de Mestrado em Modelagem de Sistemas Complexos, Escola de Artes, Ciências e Humanidades da Universidade de São Paulo.

1. Finanças: Modelos matemáticos. 2. Análise de séries temporais. 3. Sistemas complexos. I. Ferreira, Fernando Fagundes, orient. II. Título.

CDD 22.ed. - 332.015118

Dedico este trabalho à minha esposa Andréa Almeida P. da Silva, amor da minha vida, que sempre esteve ao meu lado em todas as situações, incondicionalmente. Que eu possa um dia retribuir à altura toda a sua dedicação...

Agradecimentos

Agradeço primeiramente a Deus, por me trazer a sensação de amparo e conforto nos momentos em que mais precisei. Por sempre ouvir minhas preces e por ser um importante guia que impulsiona a minha vida. Agradeço principalmente por me trazer até aqui!

Agradeço com muita ternura aos meus pais Osvaldir Lucio da Fonseca e Sônia Aparecida Borges Fonseca, por apoiarem incondicionalmente todos os meus projetos de vida. Agradeço pelo amor, dedicação e esforço na minha criação e na criação dos meus dois irmãos. Agradeço principalmente por serem exemplos na minha vida de caráter e integridade. A eles todo o meu amor, respeito e gratidão.

Aos meus dois amados irmãos Eber Lucio e Franciele Borges, por serem presenças constantes e queridas na minha vida. Agradecimento aos meus cunhados Marcelo Pereira e Cristina Silva. Agradecimento à Yasmin Fonseca, minha primeira sobrinha.

Aos meus queridos sogros Plínio Teixeira e Graça Silva, por serem pessoas maravilhosas e por me receberem e apoiarem como a um filho. Não poderia deixar de agradecer às minhas duas cunhadas/irmãs Adriana Almeida e Ângela Silva (Man) pelo apoio e pelas risadas despendidas gratuitamente.

Especial agradecimento à Andréa Almeida P. da Silva, amor da minha vida, por me apoiar em todos os momentos, incondicionalmente. Razão do meu viver, luz da minha vida! Agradeço também ao nosso cão Malik, por estar sempre ao meu lado, vigilante, principalmente nas solitárias madrugadas de estudo.

Aos amigos de longa data, Rafael Rosa, Fernando Morais, Bruno Nunes e Danilo Adrian. Aos novos amigos conquistados durante a realização de meus estudos: Leandro Ferreira, Ayrton Junior, André Machado, Daniel Savoy, Victor Corrêa, Paula Trottmann, Tiago Colliri e Valéria Reis.

Ao Professor Fernando Fagundes Ferreira, meu orientador, por sempre me apoiar e contribuir para o meu crescimento científico, intelectual e profissional.

Aos Professores Carlos de Brito, Flávia Sarti, Camilo Neto e Ana Amélia, por serem exemplos de conhecimento e excelência, fazendo jus à grande universidade que representam.

Agradecimento especial aos professores Fernando Ferreira, Camilo Neto, Esteban Tuesta, Chang Chiann e Rogério Mugnaini pelo gentil apoio à continuação de meus estudos.

Aos funcionários da secretaria de Pós-Graduação da EACH, por toda a assistência

prestada. Especial agradecimento à Jussara, por ser sempre prestativa a toda e qualquer solicitação.

À Universidade de São Paulo e à EACH, pela oportunidade e insumos fornecidos durante a realização do meu curso de mestrado.

Por fim, agradeço a CAPES pelo apoio financeiro.

Nuvens não são esferas, montanhas não são cones, litorais não são círculos, cascas não são lisas e raios não viajam em linha reta.

Bernoit Mandelbrot - The Fractal Geometry of Nature.

Resumo

Séries temporais financeiras, como índices de mercado e preços de ativos, são produzidas por interações complexas dos agentes que participam do mercado. As propriedades fractais e multifractais destas séries fornecem evidências para detectar com antecedência a ocorrência de movimentos bruscos de mercado (*crashes*). Tais evidências são obtidas ao aplicar o conceito de Calor Específico Análogo $C(q)$, proveniente da equivalência entre a Multifractalidade e Termodinâmica. Na proximidade de um *crash*, $C(q)$ apresenta um "ombro anômalo" à direita de sua curva, enquanto que na ausência de um *crash*, possui o formato parecido com uma distribuição gaussiana. Com base neste comportamento, o presente trabalho propõe um novo indicador temporal $I_A(i)$, definido como a taxa de variação da área sob a curva de $C(q)$. O indicador foi construído por intermédio de uma janela temporal de tamanho s que se movimenta ao longo da série, simulando a entrada de dados na série ao longo do tempo. A análise de $I_A(i)$ permite detectar com antecedência a ocorrência de grandes movimentos, como os famosos *crashes* de 1929 e 1987 para os índices Dow Jones, S&P500 e Nasdaq. Além disso, a análise simultânea de medidas como a Energia Livre, a Dimensão Multifractal e o Espectro Multifractal, sugerem que um *crash* de mercado se assemelha a uma transição de fase. A robustez do método para diferentes ativos e diferentes períodos de tempo, demonstra a importância dos resultados. Além disso, modelos estatísticos não lineares para a volatilidade foram empregados no trabalho para estudar grandes flutuações causadas por *crashes* e crises financeiras ao longo do tempo.

Palavras-chaves: Séries Temporais Financeiras. Análise R/S de Hurst. Detrended Fluctuation Analysis. Multifractalidade. Termodinâmica. Indicador de Crash. Sliding Observation Box.

Abstract

Financial time series such as market index and asset prices, are produced by complex interactions of agents that trade in the market. The fractal and multifractal properties of these series provides evidence for early detection of the occurrence of sudden market movements (crashes). This evidence is obtained by applying the concept of Analog Specific Heat $C(q)$, from the equivalence between the Multifractal Analysis and Thermodynamics. In the vicinity of a crash, $C(q)$ exhibits a shoulder at the right side of its curve, while in the absence of a crash, $C(q)$ presents a form similar to a Gaussian distribution curve. Based on this behavior, it is proposed in this work a new temporal indicator $I_A(i)$ defined here as the area variation rate over the Specific Heat function. We have constructed the mentioned indicator from a window of data with the first points (size s), that moves throughout the series, simulating the actual input of data over time. The indicator $I_A(i)$ allows one detecting in advance the occurrence of large financial market movements, such as those occurred in 1929 and 1987 for the marked indexes Dow Jones, Nasdaq and S&P500. Moreover, the simultaneous analysis of measures such as the Free Energy, Multifractal Dimension and Multifractal Spectrum suggest that a market crash resembles a phase transition. The robustness of the method for others assets and different periods of time demonstrates the importance of the results. Moreover, nonlinear statistical models for volatility have been employed in the work to study large fluctuations caused by crashes and financial crises over time.

Key-words: Financial Time Series. R/S Hurst Analysis. Detrendend Fluctuation Analysis. Multifractality. Thermodynamics. Crash Index. Sliding Observation Box.

Sumário

Lista de Figuras **ix**

Lista de Tabelas **xi**

1 Introdução **1**

2 Revisão da Literatura **4**

3 Análise Estatística **11**

3.1 Introdução 11
3.2 Fonte de dados consultada 12
3.3 Incrementos, Retornos e Volatilidade 12
3.4 Modelos ARCH e GARCH 16
3.5 Resultados 18
3.5.1 Índice Dow Jones 18
3.5.2 Índice S&P500 21
3.5.3 Índice Nasdaq 23
3.5.4 Índice Ibovespa 26

4 Análise Fractal e Multifractal **30**

4.1 Introdução 30
4.2 Fractais, Auto-Afinidade e Leis de Escala 31
4.3 Análise R/S de Hurst Clássica 34
4.4 Detrended Fluctuation Analysis - DFA 36
4.5 Formalismo Multifractal e Termodinâmica 38
4.6 Algoritmo Online 41
4.7 Indicadores de *Crash* 43
4.8 Resultados 47
4.8.1 Análise R/S de Hurst Clássica 47
4.8.2 Detrended Fluctuation Analysis - DFA 49
4.8.3 Análise Multifractal e Termodinâmica 50

5 Conclusão **58**
5.1 Considerações Finais 58
5.2 Sugestões para pesquisas futuras 59

A Modelos ARIMA Simulados - Identificação **60**

B Tabelas com os Ajustes dos Modelos ARMA-GARCH **65**
B.1 Índice Dow Jones: Jan/1980 a Ago/2011 65
B.2 Índice S&P500: Jan/1980 a Ago/2011 66
B.3 Índice Nasdaq: Jan/1980 a Ago/2011 67
B.4 Índice Ibovespa: Jan/1998 a Ago/2011 68

Referências Bibliográficas **69**

Lista de Figuras

3.1 Índice S&P500 - (a) Série de Fechamentos, (b) Série de Retornos, (c) Histograma, (d) Gráfico Normal Q x Q plot. 14
3.2 Dow Jones - ACF e PACF para a série de retornos e retornos ao quadrado. 19
3.3 Dow Jones - Comparação entre a volatilidade estimada $\sqrt{h(t)}$ e os retornos $r(t)$ 20
3.4 Dow Jones - ACF e PACF para a série de resíduos padronizados $\widetilde{a}(t)$ e resíduos padronizados ao quadrado $[\widetilde{a}(t)]^2$. 20
3.5 S&P500 - ACF e PACF para a série de retornos e retornos ao quadrado. . 21
3.6 S&P500 - Comparação entre a volatilidade estimada $\sqrt{h(t)}$ e os retornos $r(t)$ 22
3.7 S&P500 - ACF e PACF para a série de resíduos padronizados $\widetilde{a}(t)$ e resíduos padronizados ao quadrado $[\widetilde{a}(t)]^2$. 23
3.8 Nasdaq - ACF e PACF para a série de retornos e retornos ao quadrado. . . 24
3.9 Nasdaq - Comparação entre a volatilidade estimada $\sqrt{h(t)}$ e os retornos $r(t)$ 25
3.10 Nasdaq - ACF e PACF para a série de resíduos padronizados $\widetilde{a}(t)$ e resíduos padronizados ao quadrado $[\widetilde{a}(t)]^2$. 25
3.11 Ibovespa - ACF e PACF para a série de retornos e retornos ao quadrado. . 26
3.12 Ibovespa - Comparação entre a volatilidade estimada $\sqrt{h(t)}$ e os retornos $r(t)$ 27
3.13 Ibovespa - ACF e PACF para a série de resíduos padronizados $\widetilde{a}(t)$ e resíduos padronizados ao quadrado $[\widetilde{a}(t)]^2$. 28

4.1 A Figura da esquerda apresenta o Conjunto de Cantor (conjunto auto-similar), enquanto a Figura da direita apresenta o Brócolis (conjunto auto-afim) 32
4.2 Característica fractal presente em séries reais e simuladas 33
4.3 Índice Dow Jones - Curvas do Calor Específico Análogo $C(q)$. 40
4.4 Índice S&P500 - Curvas do Calor Específico Análogo $C(q)$. 40
4.5 Esquema para a evolução temporal de $C(q)$ do Algoritmo Online. 42
4.6 Evolução temporal da Área sob $C(q)$ - Índice Dow Jones, Índice S&P500 e Índice Nasdaq. 43

4.7 Evolução temporal do indicador $I_A(i)$: Índice Dow Jones, Índice S&P500 e Índice Nasdaq. 45
4.8 Evolução temporal do indicador ${I_A}^n(i)$: Dow Jones, S&P500 e Nasdaq. . . 46
4.9 Dow Jones (1929): (a) Energia Livre Análoga $\tau(q)$, (b) Calor Específico análogo $C(q)$, (c) Dimensão Multifractal $D(q)$ e (d) Espectro Multifractal $f(\alpha)$ versus α. 51
4.10 Dow Jones (1929): (a) Incrementos Absolutos $|\Delta x(t,1)|$, (b) Indicador de *Crash* $I_A(i)$ 51
4.11 Dow Jones (1987): (a) energia livre análoga $\tau(q)$, (b) Calor Específico Análogo $C(q)$, (c) Dimensão Multifractal $D(q)$ e (d) Espectro Multifractal $f(\alpha)$ versus α. 52
4.12 Dow Jones (1987): (a) Incrementos Absolutos $|\Delta x(t,1)|$, (b) Indicador de *Crash* $I_A(i)$ 53
4.13 S&P500 (1987): (a) Energia Livre Análoga $\tau(q)$, (b) Calor Específico Análogo $C(q)$, (c) Dimensão Multifractal $D(q)$ e (d) Espectro Multifractal $f(\alpha)$ versus α. 53
4.14 S&P500 (1987): (a) Incrementos Absolutos $|\Delta x(t,1)|$, (b) Indicador de *Crash* $I_A(i)$ 54
4.15 Nasdaq (1987): (a) Energia Livre Análoga $\tau(q)$, (b) Calor Específico Análogo $C(q)$, (c) Dimensão Multifractal $D(q)$ e (d) Espectro Multifractal $f(\alpha)$ versus α. 55
4.16 Nasdaq (1987): (a) Incrementos Absolutos $|\Delta x(t,1)|$, (b) Indicador de *Crash* $I_A(i)$ 55
4.17 Ruído Branco Gaussiano: (a) Energia Livre Análoga $\tau(q)$, (b) Calor Específico Análogo $C(q)$, (c) Dimensão Multifractal $D(q)$ e (d) Espectro Multifractal $f(\alpha)$ versus α. 56
4.18 Ruído Branco Gaussiano: (a) Incrementos Absolutos $|\Delta x(t,1)|$, (b) Indicador de *Crash* $I_A(i)$ 56

A.1 (a) Processo $AR(1)$ simulado: $r(t) = 0.8 \cdot r(t-1) + w(t)$, com $\phi = 0.8$ e n =100. (b) Função de Autocorrelação (ACF) do processo simulado. (c) Função de Autocorrelação Parcial (PACF) do processo simulado. 61
A.2 (a) Processo $MA(1)$ simulado: $r(t) = w(t) + 0.9 \cdot w(t-1)$, com $\theta = 0.9$ e n =100. (b) Função de Autocorrelação (ACF) do processo simulado. (c) Função de Autocorrelação Parcial (PACF) do processo simulado. 63
A.3 (a) Processo $AR(1)$ simulado: $r(t) = 0.8 \cdot r(t-1) + w(t) + 0.9w(t-1)$, com $\phi = 0.8$,$\theta = 0.9$ e $n = 100$. (b) Função de Autocorrelação (ACF) do processo simulado. (c) Função de Autocorrelação Parcial (PACF) do processo simulado. 64

Lista de Tabelas

4.1 Expoente de Hurst H, Dimensão Fractal D_f , Coeficiente Espectral β e o Coeficiente de Correlação γ dos Índices Dow Jones, S&P500 e Nasdaq; além do Movimento Browniano Fracionário Simulado (MBFS), com $H = 0.2$, $H = 0.5$ e $H = 0.8$. Os valores de γ marcados com *(*)* não fazem sentido, uma vez que $0 < \gamma < 1$. Isto só ocorrerá para os casos em que $0 < H < 0.5$. 48

4.2 Expoente de Escala α, Dimensão Fractal D_f , Coeficiente Espectral β e o Coeficiente de Correlação γ dos Índices Dow Jones, S&P500 e Nasdaq; além do Movimento Browniano Fracionário Simulado (MBFS), com $H = 0.2$, $H = 0.5$ e $H = 0.8$. Os valores de γ marcados com *(*)* não fazem sentido, uma vez que $0 < \gamma < 1$. Isto só ocorrerá para os casos em que $0 < \alpha < 0.5$. 49

B1 Ajuste do modelo AR-GARCH ao índice Dow Jones 65

B2 Ajuste do modelo AR-GARCH ao índice S&P500 66

B3 Ajuste do modelo AR-GARCH ao índice Nasdaq 67

B4 Ajuste do modelo AR-GARCH ao índice Ibovespa 68

Capítulo 1

Introdução

Vivemos em uma nova e empolgante era onde modelos computacionais traçam novas fronteiras para o conhecimento humano, forçando pesquisadores a repensar a forma de se fazer ciência. Tudo começa com a idéia de que para vários sistemas (sociais, biológicos, econômicos, entre outros), o conceito de Reducionismo simplesmente não se aplica. Segundo Anderson [1], entender as características dos agentes que compõe determinado sistema, não garante o entendimento deste como um todo. Esta afirmação resume a idéia de que em tais sistemas, o "todo é diferente da soma das partes".

O comportamento cooperativo dos agentes que compõe um sistema pode ser classificado como emergente, dado que este resulte da ausência de um controle central e da interação local de seus membros. As propriedades emergentes de um sistema podem ser entendidas como efeitos de larga escala. Tais propriedades são geralmente difíceis de prever, mesmo no caso de interações simples. Pode-se dizer então que um sistema formado por uma grande população de agentes conectados é dito complexo, se existe uma dinâmica global emergente resultante da ação de suas partes, ao invés de ser imposta por um controle central [2].

De acordo com Mitchell [3], os sistemas complexos são grandes redes de componentes sem controle central, em que simples regras de operação geram comportamento coletivo complexo, avançado processamento de informação e adaptação via aprendizagem ou evolução. Em geral, a ação coletiva desses agentes altera o meio, e com o decorrer do tempo, os agentes reavaliam e possivelmente trocam suas estratégias de forma a se tornarem cada vez mais competitivos. Ecologia, organismos vivos e estrutura política dos países são alguns exemplos desses sistemas.

A Economia pode ser vista também como um sistema complexo, cujos agentes (consumidores, produtores, empresas, governos, etc.) otimizam constantemente os próprios objetivos individuais. As ações individuais afetam os demais agentes, resultando em um comportamento coletivo adaptativo e complexo. Em uma economia de mercado, empresas individuais estão inevitavelmente ligadas entre si por meio de flutuações nas condições econômicas. Isso faz com que os eventos de *default* das empresas, por exemplo, sejam

muitas vezes correlacionados. Desta forma, é possível pensar na teoria da complexidade, como a chave para a compreensão da emergência dos padrões coletivos. O interesse reside em estabelecer uma abordagem alternativa, em que se busca compreender inúmeros fenômenos por meio de ingredientes microscópicos, como a heterogeneidade e as regras de interação dos agentes constituintes desses sistemas, com o objetivo de compreender padrões ou estruturas macroscópicas. [4, 5, 6].

Assim como a Economia, os mercados financeiros também podem ser tratados como sistemas complexos adaptativos. Do ponto de vista microscópico, a interação dos agentes que vão ao mercado negociar ativos é traduzida em um conjunto de séries de preços, taxas de juros, câmbio, índices, volumes negociados, etc. O preço por sua vez, é influenciado pela expectativa positiva ou negativa dos compradores, saúde financeira das empresas, conjuntura política local e global, entre outros fatores. Neste cenário, Mantegna e Stanley [7], classificam os mercados financeiros como sistemas complexos cujos agentes (pessoas, traders, instituições, etc.) interagem de maneira não linear. Além disso, os autores argumentam que os mercados possuem regras subjacentes estáveis e sua evolução temporal é continuamente monitorada. De forma complementar, Bocarra [2] classifica os mercados financeiros como sistemas complexos, cujos agentes podem ser divididos em duas grandes categorias: comerciantes e ativos. Bancos, corretoras, fundos mútuos e investidores comerciais são classificados como comerciantes. Ações, títulos, futuros e opções são classificados como ativos.

De acordo com Sornette [8], é essencial perceber que o comportamento temporal de sistemas complexos econômicos é muitas vezes controlado por eventos raros. Os temidos *crashes* de mercado são exemplos destes eventos, que podem destruir em instantes trilhões de dólares, moldando o estado psicológico de investidores, empresas e governos. Tais eventos são interessantes e dão identidade à classe de eventos conhecida como "eventos extremos". Preis e Stanley [9], afirmam que a valorização insustentável de um mercado de ações pode gerar uma "bolha financeira" que irá eventualmente "estourar". Os autores classificam *crashes* como violentas flutuações no preço dos ativos, causando grande impacto na Economia. É possível citar como exemplos de *crashes*, as famosas *Black Monday* de 28/10/1929 e a *Black Thursday* de 19/10/1987, onde os principais índices de ações nos EUA caíram em torno de 30%. Mantegna e Stanley [10] ainda argumentam que muitos pesquisadores, por razões teóricas e práticas, têm se dedicado à distribuição da variação dos preços das ações, além de sua evolução dinâmica. Desta forma, a teoria de *sistemas complexos* pode ser útil em explicar e prevenir estas falhas financeiras, que são comportamentos emergentes repentinos, inesperados e que ocorrem durante curtos intervalos de tempo.

Este trabalho se insere portanto, no contexto de sistemas complexos adaptativos. Sua principal motivação está buscar indícios para a ocorrência de grandes flutuações no preço

de ativos financeiros, por meio de ferramentas estatísticas tradicionais, juntamente com aquelas típicas da mecânica estatística. Muitas críticas têm sido feitas à inexistência de um método ou modelo para evitar crises. Tais críticas não são críticas gratuitas, uma vez que as crises prejudicam financeiramente os governos, as empresas e as famílias.

Inicialmente, para modelar as características não lineares de séries temporais financeiras, foram utilizadas ferramentas estatísticas tradicionais como os famosos modelos ARCH/GARCH. Por intermédio destes modelos, o objetivo aqui foi entender a contribuição da estrutura linear e não linear presente nos dados.

Com o objetivo de caracterizar as correlações de longo alcance presente em séries temporais financeiras, foram utilizadas ferramentas como a Análise R/S de Hurst e o Detrended Fluctuation Analysis (DFA). Estas por sua vez, estão relacionadas ao caráter escalar e fractal presente nas séries. Especial interesse estará no Multifractal Detrended Fluctuation Analysis (MFDFA), generalização do tradicional DFA, utilizado na caracterização multifractal das séries. As técnicas mencionadas levam a relações do tipo leis de potência, característica predominante em sistemas complexos.

Por fim, o objetivo principal do trabalho foi o estudo das propriedades multifractais de séries temporais financeiras na vizinhança de um *crash*. Para isso, foram utilizadas ferramentas advindas da equivalência entre o Formalismo Multifractal e a Termodinâmica. O interesse esteve em determinar a dinâmica temporal de medidas como o Expoente de Escala Multifractal, a Dimensão Multifractal Generalizada, o Espectro Multifractal e o Calor Específico Análogo, calculado por meio do Expoente de Escala. A palavra "Análogo" neste caso, refere-se ao fato da medida ser calculada para uma série temporal financeira e não para um sistema físico. Por meio da caracterização multifractal de séries temporais, foram construídos indicadores de *crash* baseados no comportamento do Calor Específico Análogo. Tais indicadores fornecem indícios para detectar com antecedência a ocorrência de *crashes* presentes nos dados.

O trabalho está dividido da seguinte maneira: O Capítulo 2 apresenta a revisão da literatura pertinente ao trabalho. O Capítulo 3 apresenta a fonte consultada para aquisição dos dados reais. Além disso, o capítulo apresenta os conceitos de incrementos, retornos e volatilidade. O capítulo ainda apresenta os tradicionais modelos ARCH/GARCH, apropriados para explicar o comportamento não linear presente séries temporais financeiras. O Capítulo 4 é dedicado ao ferramental fractal e multifractal. A Análise R/S de Hurst e o DFA são utilizadas para detectar correlações de longo alcance e leis de escala. Também são apresentadas ferramentas para detectar propriedades multifractais como o MFDFA e o Formalismo Multifractal, via equivalência com a Termodinâmica. Este capítulo ainda apresenta a definição de dois indicadores de *crash*. O Capítulo 5 é dedicado às considerações finais sobre os resultados obtidos no trabalho, além de indicações para pesquisas futuras. São apresentadas informações adicionais no Apêndice e por fim, são listadas todas as referências bibliográficas utilizadas.

Capítulo 2

Revisão da Literatura

A abordagem de sistemas complexos envolve mapear interconexões e as relações entre os agentes que formam o sistema. Nestes sistemas, existe geralmente um grande número de agentes interconectados, que em muitos casos são influenciados pelo seu ambiente externo. Por meio do contato com o ambiente externo, os agentes auto-organizam sua estrutura interna, proporcionando novas e surpreendentes propriedades macroscópicas [8].

O resultado da dinâmica destes sistemas pode ser traduzido em uma grande quantidade de informações a serem processadas. Em especial, os sistemas complexos econômicos são traduzidos em inúmeras séries temporais, como índices de mercado, preços de ativos, volumes de negociação, etc. Conforme argumenta Bouchaud [16], a possibilidade de acessar e processar facilmente grandes quantidades de informação de mercados financeiros em todo o mundo, abriu caminho para novas metodologias, onde a comparação sistemática entre a teoria e dados reais não só se torna apenas possível, mas obrigatória. Essa perspectiva tem estimulado a aplicação de ferramentas advindas da física estatística a problemas econômicos, com a esperança de se desenvolver novas metodologias para lidar com sistemas complexos.

A aplicação de conceitos físicos a sistemas econômicos tornou-se uma área de pesquisa interdisciplinar conhecida como Econofísica. Conceitos como leis de escala, universalidade, sistemas auto-organizados, etc., têm sido úteis na análise e modelagem de sistemas econômicos e financeiros. Outra importante contribuição desta área de pesquisa é a ênfase à análise empírica de dados [7].

A análise estatística de dados financeiros não é recente, remontando à publicação da tese de Doutorado de Louis Bachelier [17], em 1900. Se $X(t)$ é o preço de uma ação ou commoditie no tempo t (independente do período de tempo utilizado), o modelo de Bachelier assume que os incrementos $X(t+T)-X(t)$ são variáveis aleatórias independentes e normalmente distribuídas, com média $\mu = 0$ e variância σ^2 proporcional ao intervalo de tempo T. Isto implica que a probabilidade de ocorrerem grandes flutuações é bem pequena.

A grande quantidade de dados acumulados desde o trabalho de pioneiro de Bachelier,

levou Bernoit Mandelbrot [18] a propor uma nova abordagem, em que dados financeiros não podem ser modelados utilizando-se uma distribuição gaussiana. Em seu modelo, Mandelbrot sugeriu a modelagem dos log-retornos $logX(t+T) - logX(t)$ como variáveis aleatórias independentes e identicamente distribuídas e com distribuição de Lévy (processo estocástico que obedece ao teorema do limite central generalizado). A sua descoberta foi ratificada pouco tempo depois, por meio dos resultados apresentados por Fama [19], ao analisar o índice de ações NYSE (New York Stock Exchange).

Um recente impulso às pesquisas em Econofísica é o também pioneiro trabalho de Mantegna e Stanley [10], sobre o comportamento escalar do índice S&P500 (Standard and Poors). Os autores confirmaram que a distribuição de probabilidade dos retornos do índice S&P500 não deve ser ajustada por meio de uma distribuição Normal (Gaussiana) e sim por um processo de Lévy truncado. Eles observaram ainda que as caudas da distribuição de probabilidade da série de retornos seguem uma lei de potência (universalidade). Os autores que argumentam que este fenômeno é semelhante a uma transição de fase em um ponto crítico. Os resultados, assim como os apresentados por Mandelbrot [18], são uma clara oposição ao modelo de Bachelier [17], indicando que um processo aleatório simples não é capaz de descrever completamente a dinâmica de sistemas financeiros.

Nos últimos anos, no entanto, inúmeras pesquisas têm levado em consideração uma quantidade enorme de dados financeiros reais, incluindo volume, preço, data e hora de cada transação. No final dos anos 1990, Gopikrishnan et al. [20] [21] e Plerou et al [22] decidiram analisar cada transação de todas as ações individuais nos principais mercados dos EUA. Eles encontraram que grandes flutuações são mais comuns do que se pensava, com a cauda de seu histograma seguindo uma "lei de potência inversa de quarta ordem", ao invés de ser gaussiana. Os resultados reforçam ainda mais o fato de que o modelo proposto por Bachelier não é capaz de explicar as grandes variações no preço das ações. No entanto, ainda não se conhece um modelo (processo estocástico) que explique satisfatoriamente os dados, por isso, há um grande interesse científico na investigação de possíveis modelos.

Muitas classes de modelos estão sendo propostas para avançar na compreensão do mecanismo que governa a dinâmica do mercado. De forma geral, pode-se dizer que se destacam quatro classes:

1. **Modelos de difusão** (processos estocásticos em tempo contínuo), baseados na teoria do movimento Browniano. Estes modelos são clássicos na teoria financeira, por fornecer uma elegante solução matemática graças ao desenvolvimento de ferramentas tais como o cálculo de Itô e a teoria de Martingais. [17] [23].

2. **Modelos para a Volatilidade Condicional**, os chamados modelos ARCH/GARCH e suas extensões. Este modelos são extensões dos modelos ARIMA e foram os primeiros a considerar que a variância evolui com o tempo; fenômeno conhecido como

heterocedasticidade condicional [24, 25].

3. Modelos estocásticos aditivos como o **Movimento Browniano fracionário**, **Vôos de Lévy** e **Vôos de Lévy truncados**. Estes modelos contemplam muitos fatos estilizados e ainda são processos que capturam a estrutura monofractal presente nos dados [10] [18] [20, 21, 22].

4. **Modelos baseados em agentes**: dentre os mais promissores está o Jogo da Minoria. Este modelo exibe transição de fase, de tal forma que, os fatos estilizados aparecem na vizinhança do ponto crítico. Este modelo ainda não capturou o mecanismo que rege o mercado, contudo, modelos inspirados nele representam alternativas promissoras [26, 27, 28, 29, 30, 31, 32, 33].

Os modelos citados são construídos com o objetivo de reproduzir as propriedades estatísticas observadas empiricamente em séries temporais financeiras. Estas propriedades são muitas vezes comuns a vários mercados, que por sua vez são regulamentados por regras rígidas para a atuação dos agentes. Norouzadeh e Jafari [34] afirmam que os mercados financeiros possuem uma grande variedade de regras de negociação e mercadorias comercializadas. Desta forma, eles são caracterizados por algumas características genéricas de suas séries financeiras.

Cont [35] sinaliza que este conjunto de propriedades (que são comuns a muitos instrumentos, mercados e períodos de tempo), tem sido observado por estudos independentes em quase todos os mercados e foram classificados como fatos estilizados (regularidades estatísticas). Assim, pode-se listar alguns destes fatos estilizados visualizados empiricamente em muitas séries, a saber: ausência de autocorrelações, caudas pesadas, ganho/perda de assimetria, normalidade agregada, intermitência, clusterização de volatilidade, caudas pesadas condicionais, efeito alavancagem, entre outros. Portanto, de maneira geral, os fatos estilizados nos dão uma idéia da riqueza informacional presente nas séries temporais financeiras.

Outro fato estilizado de grande importância para o estudo de séries temporais financeiras é a detecção de correlações de longo alcance. Esta característica é revelada por meio de estudos de leis de potência presente nos dados. Conforme aponta Mantegna e Stanley [7], apesar das leis de potência serem contra-intuitivas por não apresentarem uma escala característica, elas tem grande aplicação na descrição de sistemas abertos, como os econômicos e financeiros. Desta forma, os autores ressaltam que as leis de potência são cruciais para a compreensão de fenômenos críticos, mesmo quando os sistemas finitos são considerados.

Os modelos ARCH e GARCH são muito importantes no contexto financeiro, por serem os primeiros modelos capazes de modelar as correlações de curto alcance e a heterocedasticidade presente nos dados. Podobnik et al. [37] propuseram uma variação destes

modelos para descrever a série de incrementos índice S&P500, considerando o caráter de leis de potência desta série. Eles mostraram que o processo combinado AR-GARCH com inovações seguindo uma distribuição de Lévy Truncada, é capaz de gerar leis de potência. Os autores ainda demonstraram a estabilidade do processo, ao definirem uma nova variável que é a soma de n variáveis do processo AR-GARCH descrito anteriormente. Desta forma, os autores puderam verificar o surgimento de vários fatos estilizados como: correlações de longo alcance para a variância, estabilidade para um longo intervalo de escalas de tempo e comportamento do tipo lei de potência para a função distribuição de probabilidade do processo.

Motivado pelo objetivo de encontrar uma descrição mais precisa para a dinâmica empírica de incrementos financeiros, Podobnik, Grosse e Stanley [38] propuseram um processo estocástico que produz três propriedades estatísticas destas séries: autocorrelações de curto alcance, correlações de longo alcance no valor absoluto dos incrementos (com cruzamento entre regimes do tipo lei de potência de aproximadamente uma semana) e estabilidade do tipo lei de potência nas caudas da distribuição dos incrementos. Para isso os autores propõe um modelo do tipo AR-GARCH para a série de retornos do índice S&P500, supondo inovações independentes, com distribuição do tipo Lévy truncada. Os autores argumentam que este processo estocástico é capaz de reproduzir muito bem as propriedades estatísticas observadas em dados de alta frequência do índice S&P500.

Uma forma de estudar correlações de longo alcance em séries temporais financeiras é por meio da Análise R/S de Hurst Clássica. Esta técnica foi proposta por Henry Hurst em 1951 [39] [40], com o intuito de testar a presença de correlações de longo alcance em séries temporais empíricas. A idéia por trás da técnica é olhar para o comportamento escalar dos desvios cumulativos reescalados pela média (intervalo em que o sistema varia em função do tempo) e comparar com a hipótese nula de que a série é aleatória. Isto é feito por meio do expoente H de Hurst estimado [34] [41].

Norouzadeh e Jafari [34] fazem uma excelente revisão do método. Além disso, eles aplicam a Análise R/S de Hurst para a série de retornos do TEPIX (Tehran Price Index), indicando que a série de retornos do TEPIX é persistente, caracterizada por correlações de longo alcance em todas as escalas temporais. No entanto os autores alertam que o método de Hurst tende a superestimar o expoente H Hurst, para séries temporais cujos valores são pequenos. Esta limitação é eliminada por intermédio da técnica do Detrended Fluctuation Analysis (DFA).

O DFA é uma alternativa à Análise R/S de Hurst. Esta técnica foi proposta inicialmente com o objetivo de detectar correlações de longo alcance em sequências de DNA [42] [43]. Além disso, esta técnica tem sido utilizada para determinar propriedades de escala monofractal. A técnica ainda traz consigo a vantagem de evitar a detecção espúria de correlações de longo alcance, quando esta é um artefato da não-estacionariedade. [12]

Norouzadeh e Jafari [34] também aplicaram o DFA para a série de retornos do TEPIX

(Tehran Price Index). O resultado corrobora as conclusões obtidas pela Análise R/S de Hurst, mostrando que a série de retornos do TEPIX é realmente persistente, caracterizada por correlações de longo alcance em todas as escalas temporais.

Grech e Mazur [44] utilizaram o DFA para estimar o expoente Hurst do índice de ações DJIA (Dow Jones Industrial Average). Os autores investigaram se esta medida é capaz de gerar com antecedência, indícios para a ocorrência de um *crash* de mercado. Para isso, eles analisaram o comportamento do expoente de Hurst local, antes de grandes choques na série temporal do índice. Os autores analisaram os famosos *crashes* de 1929 e 1987 além de alguns períodos de crises menos graves. Eles concluíram que uma análise cuidadosa do expoente local de Hurst pode ajudar significativamente a determinar a maneira em que o mercado se comportará em cada instante, uma vez que tendências muito claras puderam ser distinguidas, não apenas nas séries financeiras, mas também na evolução temporal do expoente de Hurst. Isto confirma que expoente de Hurst local pode ser utilizado como uma medida real do estado de excitação do mercado.

Bernoit Mandelbrot foi o primeiro a observar evidências de estrutural fractal em dados financeiros [45]. Inúmeros trabalhos foram reportados na literatura ratificando este resultado ([14, 15], [46, 47, 48, 49, 50, 51]). De acordo com Kantelhardt [11], se um sistema complexo é caracterizado por uma dinâmica fractal (ou multifractal), com particular(es) expoente(s) de escala, este resultado irá ajudar na obtenção de previsões sobre o comportamento futuro do sistema e sua reação a perturbações externas. Transições de fase em um sistema complexo são freqüentemente associadas com mudanças em sua dinâmica fractal, permitindo a detecção de tais transições (ou os estados correspondentes) por análise fractal. A dinâmica fractal (ou multifractal) de uma série temporal deve ser estabelecida e não ser assumida a priori. Portanto, há a necessidade de técnicas de análise refinadas, que ajudam a diferenciar a verdadeira dinâmica fractal (ou multifractal) de detecções espúrias, causadas por não estacionariedades presentes nos dados.

Neste contexto, Kantelhardt et al. [12], desenvolveram uma nova metodologia para caracterizar o comportamento multifractal de séries temporais não estacionárias, baseado na generalização do DFA. Os autores denominaram a nova técnica por MFDFA (Multifractal Detrended Fluctuation Analysis). Por meio de vários exemplos eles argumentam que o novo método é capaz de determinar o comportamento escalar multifractal em séries temporais. O método permite ainda detectar a origem do comportamento multifractal (correlações de longo alcance, distribuição de probabilidade dispersa). Os autores ainda mostram que há equivalência entre a técnica apresentada e a técnica de Wavelets.

Outros trabalhos, no entanto, exploram o caráter multifractal de séries temporais baseado na equivalência com a termodinâmica. Baseado na abordagem de enumeração exata, Lee e Stanley [54] encontraram evidências sugerindo a existência de transição de fase em espectro multifractal de modelos DLA (Diffusion-Limited Aggregation). Acima de um ponto crítico, a expansão de momentos mostra uma infinita hierarquia de fases, enquanto

que abaixo de um ponto crítico, uma única fase. Sob o ponto crítico, são encontradas flutuações em todas as escalas de energia, comportamentos singulares da energia e do calor específico. Motivados pela analogia entre multifractalidade e termodinâmica, os autores definem a seguinte função de partição:

$$Z(\beta, L) = \sum_a C_a \sum_i p_{i,a}^{\beta} = \sum_{\epsilon} D(\epsilon, L) L^{-\beta\epsilon},$$

onde C_a é o peso da configuração a e p_i, a é uma medida de probabilidade e L é o "comprimento do sistema". A energia livre $F(\beta, L)$ é obtida por meio da seguinte Equação:

$$\partial F(\beta, L) = -\frac{\ln Z(\beta, L)}{\ln L},$$

com $F(\beta) = \lim_{L \to \infty} F(\beta, L)$. Especial atenção é dada à transformada de Legendre:

$$\begin{cases} S(E, L) \equiv \beta E - F(\beta, L) \\ E = \dfrac{\partial F(\beta, L)}{\partial \beta} \end{cases}.$$

Convenciona-se chamar E de energia, $S(E, L)$ de entropia e $F(\beta, L)$ de energia livre. Para comparar a notação utilizada por Lee e Stanley [54] com outras referências, os autores sugerem a seguinte conversão:

Lee e Stanley [54]	Outras Referências
β	q
E	α
$F(\beta)$	$\tau(q)$
$S(E)$	$f(\alpha)$

O trabalho de Lee e Stanley [54] ainda mostrou que o formato da curva do Calor Específico se altera sensivelmente, quando este é calculado nas proximidades de um valor crítico.

Baseado nestes resultados, Canessa [14] [15] apresentou uma nova abordagem para caracterizar o início de *crashes* de mercado. Isto pode ser feito por intermédio do cálculo do Calor específico análogo $C(q)$ para uma série temporal financeira, usando a definição termodinâmica de multifractalidade. Na ausência de grandes flutuações no mercado, $C(q)$ apresenta um formato parecido com uma distribuição gaussiana. Em períodos de grandes flutuações, observa-se a formação de ombro à direita da curva de $C(q)$. A aplicação da técnica para os dados históricos do índice S&P500 demonstrou estes fatos, ressaltando o fato de que isto acontece devido a violentos agrupamentos de volatilidade. Para grandes valores de atrasos (*lags*), $C(q)$ se assemelha a uma transição de fase clássica em um ponto crítico. Essa transição de fase dinâmica pode ser explicada pelo mapeamento com fenômeno multifractal em processos multiplicativos aleatórios.

Vale mencionar que as evidências empíricas de que o processo é multifractal, não ga-

rantem que de fato o seja. Bouchaud, Potters e Meyer [52] apresentaram um estudo em que séries temporais finitas, geradas artificialmente por processos fractais apresentam espectro multifractal. Os resultados apresentados pelos autores sugerem que em séries temporais financeiras, não se pode distinguir entre comportamento multifractal real e aparente. Jiang e Zhou [53] descobriram evidências de que não se pode distinguir o comportamento escalar de séries temporais intradiárias reais, do comportamento escalar de séries aleatórias. Os autores classificam a multifractalidade em índices de mercado como uma "ilusão". Para o presente trabalho, entendemos que este ainda é um assunto em aberto e que as pesquisas deverão evoluir no sentido de promover melhores técnicas para a detecção da estrutura multifractal em séries temporais financeiras, evitando assim detecções espúrias de multifractalidade. No entanto, estamos interessados apenas em investigar se o espectro multifractal empírico contém informações relevantes, capazes de permitir a detecção com antecedência de *crashes* de mercado. Não estamos interessados, por exemplo em definir se a natureza do processo é fractal ou multifractal.

Capítulo 3

Análise Estatística

3.1 Introdução

Este capítulo tem como foco a caracterização estatística de séries temporais financeiras reais. Inicialmente apresenta-se uma descrição das mesmas, além da fonte consultada. Estas séries serão utilizadas tanto na análise estatística, quanto na análise multifractal.

Outro aspecto importante para o trabalho são as transformações das séries originais em incrementos e retornos. Estas transformações são necessárias para eliminar tendências presentes nos dados. Além do mais, estas medidas têm propriedades mais interessantes que as séries originais [36]. O capítulo trata ainda do importante conceito de volatilidade, utilizado para mapear grandes flutuações nas séries financeiras, além de ser importante para definir medidas de risco.

Por fim, serão apresentados os modelos estatísticos ARCH e GARCH. Estes modelos foram os primeiros a considerar a evolução temporal da variância de ativos financeiros, chamada costumeiramente de volatilidade. Tais modelos são úteis para modelar séries financeiras, por apresentarem estacionariedade local para grandes intervalos de tempo, mas ao mesmo tempo apresentarem não estacionariedade local para pequenos intervalos temporais.

3.2 Fonte de dados consultada

Para as análises feitas no presente trabalho, foram utilizados quatro índices financeiros: Dow Jones, S&P500, Nasdaq e Ibovespa. Foram coletados dados em vários períodos de tempo, observando as datas onde ocorreram *crashes* de mercado. As séries utilizadas no presente trabalho possuem frequência diária.

À medida que as séries foram citadas no texto, foram citados também os períodos de tempo considerados. Estas séries foram extraídas do site público Yahoo Finance [55].

Segue a seguir uma pequena descrição sobre cada um dos Índices considerados:

- **Índice Dow Jones**: Indica o desempenho das 30 maiores empresas que negociam na bolsa de Valores de Nova York.
- **Índice S&P500**: Indica o desempenho das 500 empresas mais representativas na Economia dos Estados Unidos.
- **Índice Nasdaq**: Indica o desempenho da bolsa eletrônica americana Nasdaq para empresas de tecnologia.
- **Índice Ibovespa**: Mais importante indicador do desempenho médio das cotações do Mercado de ações brasileiro.

3.3 Incrementos, Retornos e Volatilidade

Ao trabalharmos com ativos financeiros, estaremos necessariamente preocupados com a avaliação do risco, que é freqüentemente medido em termos das variações de preços dos ativos [36]. Denotemos $x(t)$ como o preço de um ativo no instante t, supondo que não há pagamento de dividendos. A variação de preços ou incrementos $\Delta x(t,1)$ entre os instantes $t+1$ e t é dada por:

$$\Delta x(t,1) = x(t+1) - x(t). \tag{3.1}$$

Por sua vez, o retorno simples $r(t,1)$ entre os instantes $t+1$ e t, é dado pela Equação 3.2:

$$r(t,1) = \frac{x(t+1) - x(t)}{x(t)} = \frac{\Delta x(t,1)}{x(t)}. \tag{3.2}$$

Rearranjando os termos da Equação 3.2 e aplicando o logaritmo natural em ambos os lados da igualdade, temos:

$$\ln\left(1 + r(t,1)\right) = \ln\left(\frac{x(t+1)}{x(t)}\right).$$

Temos que $ln(1 + r(t,1)) = r(t,1) - \frac{r^2(t,1)}{2} + \frac{r^3(t,1)}{3} - \cdots$, para $-1 < r(t,1) \leq 1$. Se $r(t,1)$ é uma pequena variação percentual, então os termos de ordem maiores ou iguais a dois serão desprezíveis, ou seja, $\ln(1 + r(t,1)) \approx r(t,1)$. Assim, temos:

$$r(t,1) \approx \ln\left(\frac{x(t+1)}{x(t)}\right). \tag{3.3}$$

A Equação 3.3 é geralmente chamada de retorno continuamente composto ou log-retorno [36].

É possível ainda definir incrementos e retornos multiperíodo [36] [56]. Desta forma, os incrementos de período T e o retorno simples de período T, entre os instantes $t+T$ e t, serão dados pelas Equações 3.4 e 3.5 respectivamente:

$$\Delta x(t,T) = x(t+T) - x(t), \tag{3.4}$$

$$r(t,T) = \frac{x(t+T) - x(t)}{x(t)}. \tag{3.5}$$

Por simplicidade de notação, quando $T = 1$, utilizaremos $r(t)$ no lugar de $r(t,1)$. Na prática ao invés de trabalhar com preços propriamente ditos, é preferível trabalhar com os retornos, que são livres de escala e geralmente têm propriedades estatísticas mais interessantes (fatos estilizados). Estes fatos são comuns a muitos instrumentos financeiros, mercados e períodos de tempo. Os principais fatos estilizados sobre a série de retornos são [36]:

1. Os retornos são em geral não autocorrelacionados
2. Os quadrados dos retornos são autocorrelacionados
3. As séries de retornos apresentam agrupamentos de volatilidades ao longo do tempo
4. A distribuição (incondicional) dos retornos apresenta caudas mais pesadas do que uma distribuição normal.

A título de exemplo, a Figura 3.1 apresenta o gráfico do Índice S&P500, de Jan/1980 a Dez/1992 (a). Na letra (b) da figura apresenta a série de retornos, onde podem ser verificados os fatos estilizados citados anteriormente, como aparente estacionariedade com média ao redor de zero e agrupamento de volatilidades. Na letras (c) e (d) do gráfico, podemos ver que existem valores de retornos muito afastados da média (valores extremos) e também que as caudas das distribuições dos retornos são mais pesadas que a da distribuição normal. Este resultados são verificados empiricamente para diferentes tipos de índices e ativos financeiros, diferentes mercados e até para diferentes períodos de tempo [35].

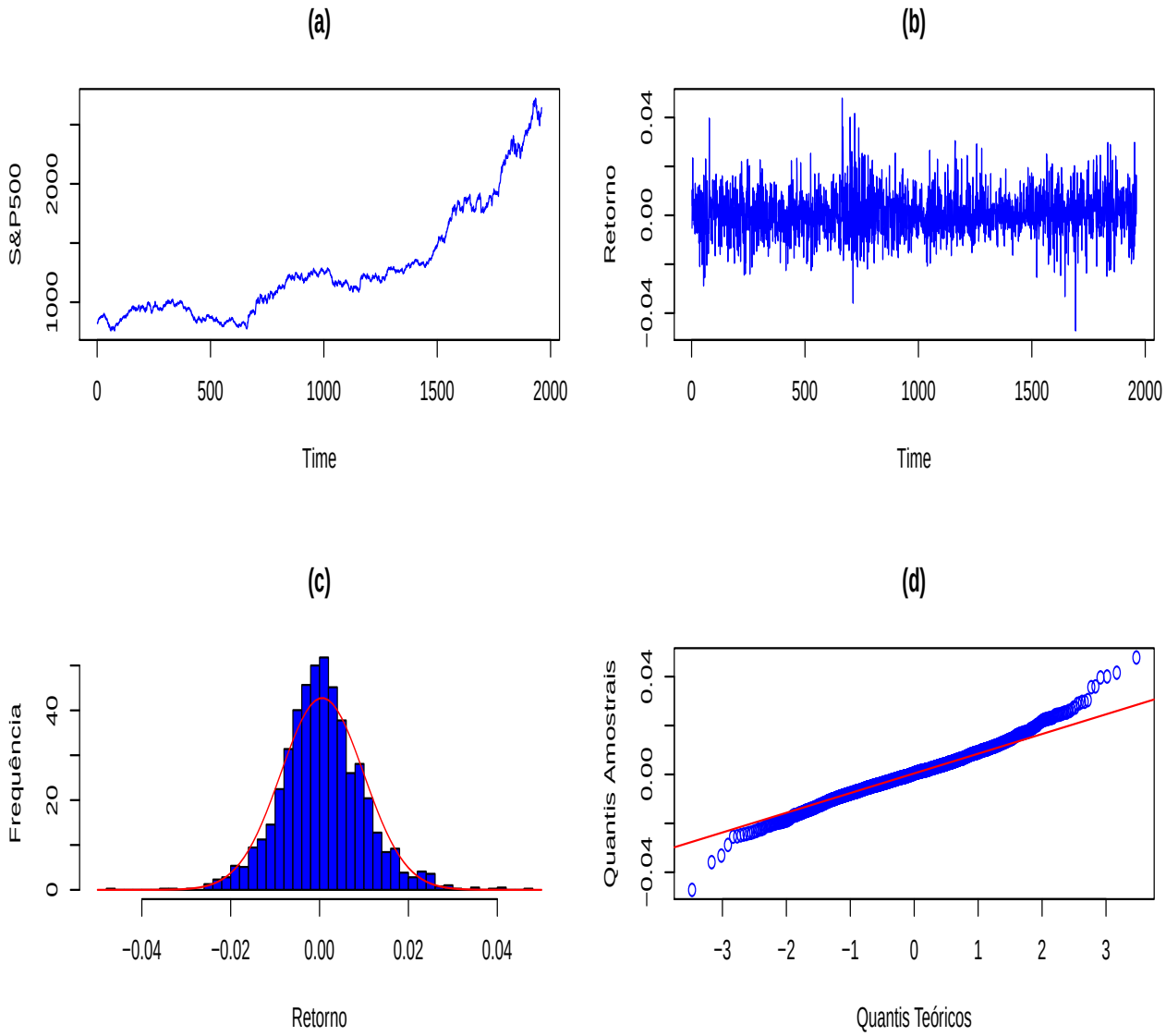

Figura 3.1: Índice S&P500 - (a) Série de Fechamentos, (b) Série de Retornos, (c) Histograma, (d) Gráfico Normal Q x Q plot.

A volatilidade é chamada de variável latente, uma vez que não pode ser observada diretamente, como os preços das ações ou as taxas de juros. O melhor que se pode fazer é uma medida da volatilidade, geralmente por meio de alguma fórmula estatística. Para fazer essa medida é necessário olhar para os dados passados da série de retornos. Por esta razão são definidos diferentes tipos de volatilidades, como *proxies* para a volatilidade real. Podemos citar então alguns tipos de volatilidade: Histórica (Realizada), Implícita e Futura [23].

A volatilidade histórica ou realizada é uma medida da quantidade de aleatoriedade durante algum período no passado. O prazo é sempre especificado assim como o método matemático para o seu cálculo. Às vezes, esta medida é usada como uma estimativa para

a volatilidade futura. Já a volatilidade implícita entra como *input* nas fórmulas de precificação de Black-Scholes; que dá o preço de mercado de uma opção. Ela é freqüentemente descrita como a visão de mercado da volatilidade futura para a vida de uma opção particular. No entanto, ela é também influenciada por outros efeitos, como oferta e demanda. Por fim, o adjetivo "Futuro" pode ser aplicado a muitas formas de volatilidade, e refere-se à volatilidade (real ou implícita) durante algum período no futuro. Existe ainda outra forma de se calcular a volatilidade histórica, onde o quadrado dos retornos é utilizado como *proxie* para a volatilidade [23, 36].

Tal como os retornos, a volatilidade apresenta um conjunto de características que são comuns a vários índices, ativos, mercados e períodos de tempo.

O agrupamento (clustering) de grandes movimentos e pequenos movimentos nos preços foi uma das principais características documentadas do processo de volatilidade de preços de ativos: grandes mudanças no preço de um ativo são geralmente seguidas por grandes mudanças, enquanto que pequenas mudanças são geralmente seguidas de pequenas mudanças.

3.4 Modelos ARCH e GARCH

Os modelos ARCH (Auto-Regressivos com Heterocedasticidade Condicional) foram introduzidos por Engle em 1982 [24], com o objetivo de estimar a variância do índice de inflação do Reino Unido. Estes modelos foram os primeiros a considerar o fato de que a variância evolui temporalmente. Para propor estes modelos, Engle levou em conta os seguintes fatos estilizados: geralmente não há correlação serial nas séries de retornos $r(t)$, mas há correlação serial para a volatilidade $h(t)$ (variância condicional), que por sua vez, depende dos retornos defasados por meio de uma função quadrática. Seja $r(t)$ uma série de retornos.

Definição 3.4.1 *Um modelo geral $ARCH(w)$ é da seguinte maneira [36] [57]:*

$$\begin{aligned} r(t) &= \sqrt{h(t)} \cdot \varepsilon_t \\ h(t) &= \alpha_0 + \alpha_1 \cdot r^2(t-1) + \ldots + \alpha_r \cdot r^2(t-w) \end{aligned}, \tag{3.6}$$

onde w é a ordem do modelo, ε_t é uma sequência de variáveis aleatórias independentes e identicamente distribuídas (i.i.d.) com média zero e variância um, $\alpha_0 > 0, \alpha_i >= 0$ com $i > 0$. Supõe-se também que as inovações $\varepsilon_t \sim N(0,1)$ ou $\varepsilon_t \sim t_v$ (distribuição t de Student com v graus de liberdade).

As condições listadas na definição 3.4.1, garantem que a variância do processo seja finita e positiva. Os modelos ARCH todavia, têm vantagens e desvantagens [36]. Por exemplo, os modelos ARCH apresentam curtose (K) do processo $r(t)$, tal que:

$$K = 3 \cdot \frac{1 - \alpha_1^2}{1 - 3 \cdot \alpha_1^2} > 3. \tag{3.7}$$

Logo, ao admitir que $r(t)$ segue um modelo $ARCH(w)$, serão observadas caudas da distribuição de retornos mais pesadas que as caudas da distribuição normal, o que é uma característica favorável, dadas as propriedades empíricas dos retornos. Por outro lado, o modelo trata retornos positivos e negativos da mesma forma, já que o quadrado dos retornos entra na fórmula da volatilidade. Esta é uma desvantagem do modelo, pois na prática, sabe-se que a volatilidade reage de modo diferente a retornos positivos e negativos (efeito alavancagem). Para a maioria dos ativos, a volatilidade está relacionada de maneira não linear a série de retornos [35] [36].

Uma importante extensão do modelo ARCH foi proposta por Bollerslev [25], que substitui a representação $AR(p)$ para a volatilidade da série modelada por um processo $ARMA(p,q)$, uma vez que o modelo ARMA pode ser mais parcimonioso, no sentido de apresentar menos parâmetros do que um modelo $AR(p)$ ou $MA(q)$ puros [36].

Definição 3.4.2 *Um modelo geral $GARCH(w, z)$ é definido da seguinte maneira [36] [57]:*

$$\begin{aligned} r(t) &= \sqrt{h(t)} \cdot \varepsilon_t \\ h(t) &= \alpha_0 + \sum_{i=1}^{w} \alpha_i \cdot r^2(t-i) + \sum_{j=1}^{z} \beta_j \cdot h(t-j) \end{aligned} \quad , \tag{3.8}$$

onde w e z definem a ordem do modelo, ε_t é uma sequência de variáveis aleatórias independentes e identicamente distribuídas (i.i.d.) com média zero e variância um, $\alpha_0 > 0, \alpha_i >= 0$ com $i > 0$, $\beta_j >= 0$, $\sum_{i=1}^{q} \alpha_i + \beta_i$ com $q = max(w, z)$.

As condições listadas na definição 3.4.2, garantem que a variância do processo seja finita e positiva. Uma variação deste modelo foi proposta por Podobnik et al. [37, 38], que propõem uma nova abordagem em que as inovações ε_t seguem uma distribuição de Lévy Truncada.

Antes de aplicar um modelo $GARCH(w, z)$, deve-se verificar se existe dependência serial na série de retornos. Caso exista, deve-se especificar um modelo do tipo $r(t) = \mu + a(t)$, com $\mu \sim ARMA(p, q)$ para removê-la. Desta forma, teremos o modelo do tipo $ARMA(p, q) - GARCH(w, z)$, conforme descrito na Equação 3.9:

$$\begin{aligned} r(t) &= \phi_1 \cdot r(t-1) + \ldots + \phi_p \cdot r(t-p) + \theta_1 \cdot r(t-1) + ... + \theta_q \cdot r(t-q) + a(t) \\ a(t) &= \sqrt{h(t)} \\ h(t) &= \alpha_0 + \sum_{i=1}^{w} \alpha_i \cdot r^2(t-i) + \sum_{j=1}^{z} \beta_j \cdot h(t-j) \end{aligned} \quad . \tag{3.9}$$

No modelo definido na Equação 3.9, o termo $a(t)$ é o resíduo do modelo $ARMA(p, q)$ aplicado à série de retornos $r(t)$. A série $a(t)$ por sua vez, é explicada pelo modelo $GARCH(w, z)$, caso existam efeitos ARCH presentes.

Para verificar se o ajuste do modelo foi adequado, toma-se os resíduos padronizados definidos na Equação 3.10:

$$\widetilde{a}(t) = \frac{a(t)}{\sqrt{h(t)}}, \tag{3.10}$$

que devem ser variáveis aleatórias i.i.d. com distribuição Normal Padrão ou t-Student. Para verificar a condição, geralmente utiliza-se a Estatística Q de Box-Pierce-Ljung para a sequência $\widetilde{a}(t)$. É possível ainda aplicar à sequência $\widetilde{a}(t)$ o Teste de Multiplicadores de Lagrange de Engle para verificar se existe heterocedasticidade condicional nos resíduos do modelo [36].

3.5 Resultados

A seguir, são apresentados os resultados dos ajustes dos modelos GARCH para a série de retornos dos índices Dow Jones, S&P500 e Nasdaq, de Jan/1980 a Ago/2011. Para o índice Ibovespa, foi utilizado o período de Jan/1998 a AGO/2011.

Para cada ajuste, foram apresentadas as Funções de Autocorrelação (ACF) e Autocorrelação Parcial (PACF) empíricas. A análise destas funções foi importante para definir a ordem dos modelos $ARMA(p, q)$ ajustados, onde q refere-se à parte MA (Médias Móveis) do modelo, enquanto p refere-se à ordem de modelos AR (Auto-Regressiva) do modelo. Em todas as análises, tomou-se o nível de significância $\alpha = 0.05$, representado pelas linhas tracejadas azuis nos gráficos das figuras subsequentes. Isso quer dizer que somente os "lags" (ou defasagens) que estejam fora dos intervalos definidos pelas linhas tracejadas são estatisticamente significantes. O último lag fora do intervalo de confiança, define a ordem do modelo (ACF para q e PACF para p). Na sessão de Anexos, são apresentadas as tabelas com o ajuste dos modelos.

3.5.1 Índice Dow Jones

Com o objetivo de checar a aplicabilidade do modelo GARCH aos dados, na Figura 3.2, gráficos (a) e (b), são apresentadas a ACF e a PACF para a série de retornos do Índice Dow Jones. A análise destas funções indica a existência de uma dependência linear entre as observações, principalmente nos *lags* 1, 2, 11, 12, 15, 16 e 18. Nos gráficos (c) e (d), são apresentadas respectivamente as ACF e as PACF para a série de retornos ao quadrado, onde é possível notar a presença de uma forte dependência. Tanto a ACF quanto na PACF tendem a diminuir vagarosamente, indicando que a variância do processo não é estacionária. Desta forma, como existe dependência tanto na série de retornos quanto na série de retornos ao quadrado (volatilidade), é possível concluir que o uso de um modelo $ARMA(p, q) - GARCH(w, z)$ é adequado ao conjunto de dados.

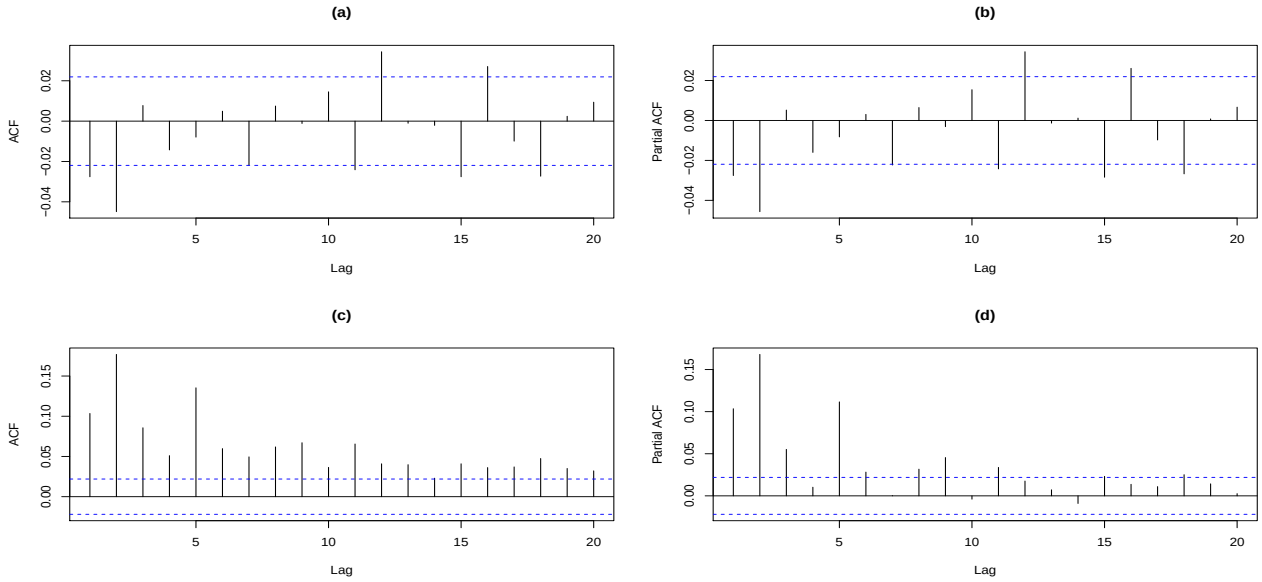

Figura 3.2: Dow Jones - ACF e PACF para a série de retornos e retornos ao quadrado.

Dada a aplicabilidade do modelo aos dados, um modelo $ARMA(18, 0)$ (ou simplesmente $AR(18)$), foi ajustado para retirar a correlação na série de retornos. A ordem $p = 18$ utilizada é referente ao maior *lag* fora dos intervalos de confiança na ACF. Em seguida, após retirar os coeficientes não significativos do modelo, a um nível de 5%, um modelo $GARCH(1, 1)$ foi ajustado na série de resíduos. Por fim, após retirar novamente alguns coeficientes não significativos, o modelo final estimado é descrito Equação 3.11:

$$\begin{aligned} r(t) &= 0.0006 + 0.0292 \cdot r(t-12) - 0.0240 \cdot r(t-15) + a(t) \\ a(t) &= \sqrt{h_t} \cdot \varepsilon_t \\ h(t) &= 0.000002 + 0.0815 \cdot r^2(t-1) + 0.9067 \cdot h(t-1) \end{aligned} \quad . \tag{3.11}$$

A Figura 3.3 traz a comparação entre a volatilidade condicional $\sqrt{h(t)}$ estimada pelo modelo (curvas azuis) e a série de retornos (curva preta). Conforme é possível notar, a volatilidade estimada captura as principais flutuações na série de retornos. A volatilidade foi multiplicada respectivamente por 1.96 e -1.96, simulando um intervalo com 95% de confiança.

Por fim, resta ainda checar se o modelo foi capaz de capturar a heterocedasticidade condicional presente nos dados. Na Figura 3.4, (a) e (b), são apresentadas a ACF e a PACF da série de resíduos padronizados ($\widetilde{a}(t)$). Nas letras (c) e (d), são apresentadas as ACF e PACF da série de resíduos padronizados ao quadrado ($[\widetilde{a}(t)]^2$). Conforme podemos ver, os resíduos comportam-se como um ruído branco gaussiano, o que nos diz que o modelo foi capaz de capturar a heterocedasticidade presente nos dados, sendo portanto adequado.

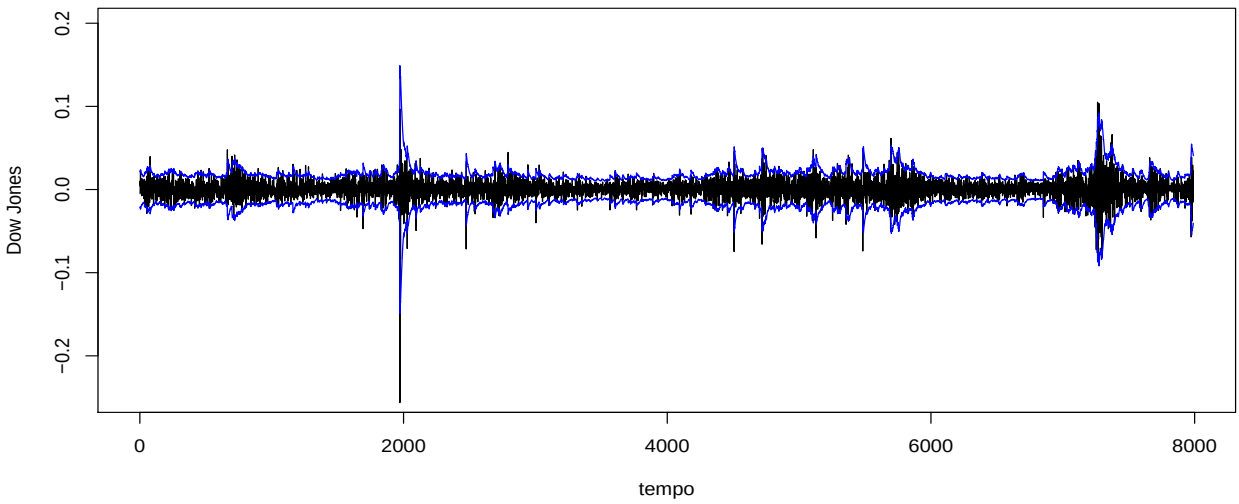

Figura 3.3: Dow Jones - Comparação entre a volatilidade estimada $\sqrt{h(t)}$ e os retornos $r(t)$

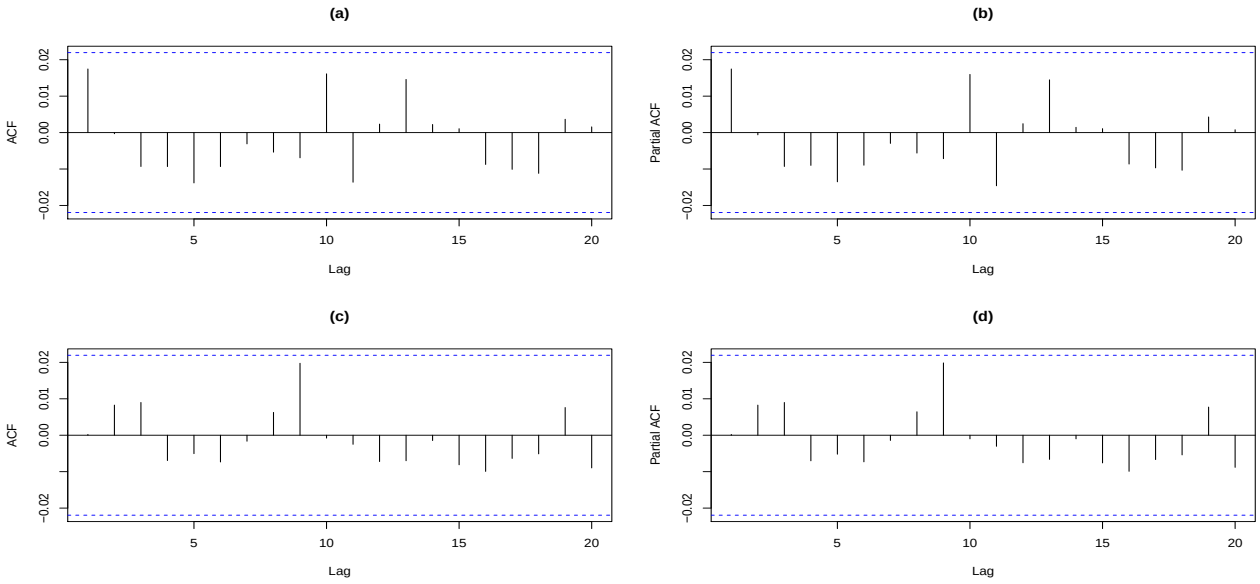

Figura 3.4: Dow Jones - ACF e PACF para a série de resíduos padronizados $\widetilde{a}(t)$ e resíduos padronizados ao quadrado $[\widetilde{a}(t)]^2$.

Para checar se o ajuste do modelo foi satisfatório, foi aplicado também o teste Box-Pierce-Ljung com p-valor < 0.05, para os *lags* 1 a 10, não havendo evidências estatísticas de que exista correlação serial para os resíduos padronizados. Em seguida foi aplicado o Teste de Engle para checar a presença de efeito ARCH. Neste caso com p-valor < 0.05 e para os *lags* de 1 a 10, aceita-se a hipótese nula de que os dados não possuem efeito ARCH presente, ou seja, o modelo foi capaz de capturar a heterocedasticidade presente nos dados.

3.5.2 Índice S&P500

Com o objetivo de checar a aplicabilidade do modelo GARCH aos dados, na Figura 3.5, gráficos (a) e (b), são apresentadas a ACF e a PACF para a série de retornos do Índice S&P500. A análise destas funções indica a existência de uma dependência linear entre as observações, principalmente nos *lags* 1, 2, 7, 12, 15, 16 e 18. Nos gráficos (c) e (d), são apresentadas respectivamente as ACF e as PACF para a série de retornos ao quadrado, onde é possível notar a presença de uma forte dependência. Tanto a ACF quanto na PACF tendem a diminuir vagarosamente, indicando que a variância do processo não é estacionária. Desta forma, como existe dependência tanto na série de retornos quanto na série de retornos ao quadrado (volatilidade), é possível concluir que o uso de um modelo $ARMA(p,q) - GARCH(w,z)$ é adequado ao conjunto de dados.

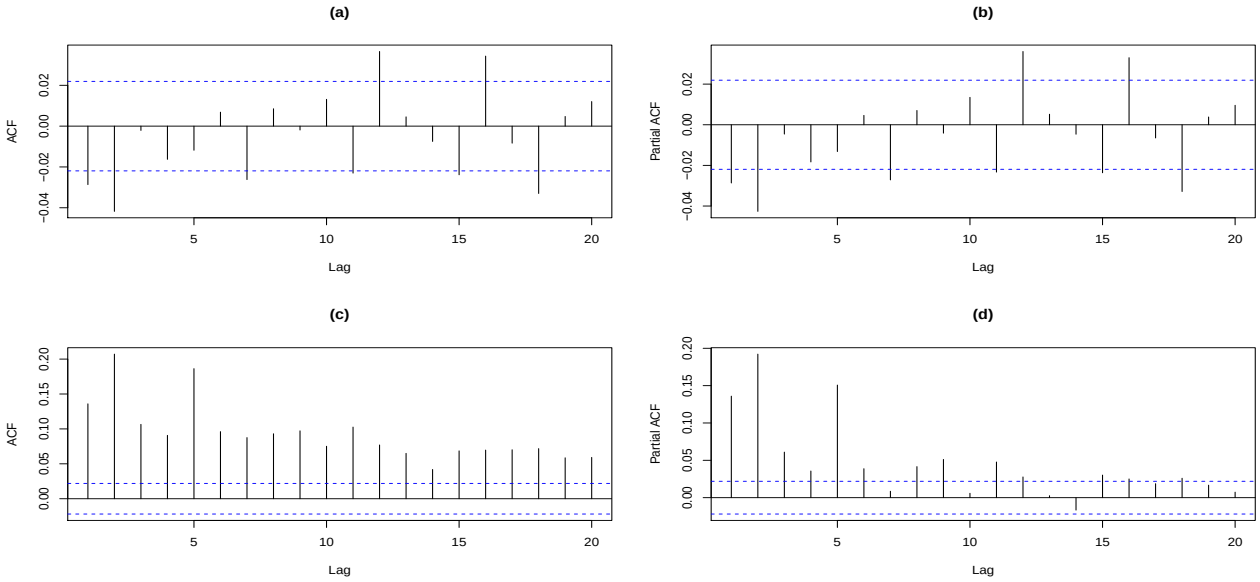

Figura 3.5: S&P500 - ACF e PACF para a série de retornos e retornos ao quadrado.

Dada a aplicabilidade do modelo aos dados, um modelo $AR(18)$ foi ajustado para retirar a correlação na série de retornos. A ordem $p = 18$ utilizada é referente ao maior *lag* fora dos intervalos de confiança na ACF. Em seguida, após retirar os coeficientes não significativos do modelo, a um nível de 5%, um modelo $GARCH(1,1)$ foi ajustado na série de resíduos. Por fim, após retirar novamente alguns coeficientes não significativos, o modelo final estimado é descrito Equação 3.12:

$$\begin{aligned} r(t) &= 0.0005 - 0.0299 \cdot r(t-4) + 0.0303 \cdot r(t-12) + a(t) \\ a(t) &= \sqrt{h_t} \cdot \varepsilon_t \\ h(t) &= 0.000001 + 0.0786 \cdot r^2(t-1) + 0.9124 \cdot h(t-1) \end{aligned} \quad . \tag{3.12}$$

A Figura 3.6 traz a comparação entre a volatilidade condicional $\sqrt{h(t)}$ estimada pelo modelo (curvas azuis) e a série de retornos (curva preta). Conforme é possível notar, a volatilidade estimada captura as principais flutuações na série de retornos. A volatilidade foi multiplicada respectivamente por 1.96 e -1.96, simulando um intervalo com 95% de confiança.

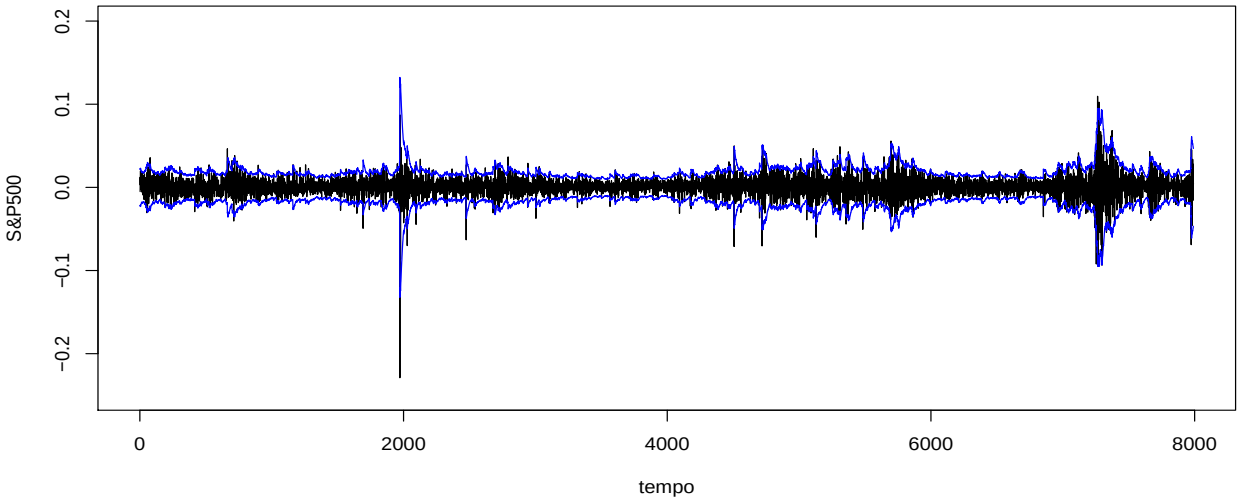

Figura 3.6: S&P500 - Comparação entre a volatilidade estimada $\sqrt{h(t)}$ e os retornos $r(t)$

Por fim, resta ainda checar se o modelo foi capaz de capturar a heterocedasticidade condicional presente nos dados. Na Figura 3.7, (a) e (b), são apresentadas a ACF e a PACF da série de resíduos padronizados ($\widetilde{a}(t)$). Nas letras (c) e (d), são apresentadas a ACF e a PACF da série de resíduos padronizados ao quadrado ($[\widetilde{a}(t)]^2$). Conforme podemos ver, os resíduos comportam-se como um ruído branco gaussiano, o que nos diz que o modelo foi capaz de capturar a heterocedasticidade presente nos dados, sendo portanto adequado.

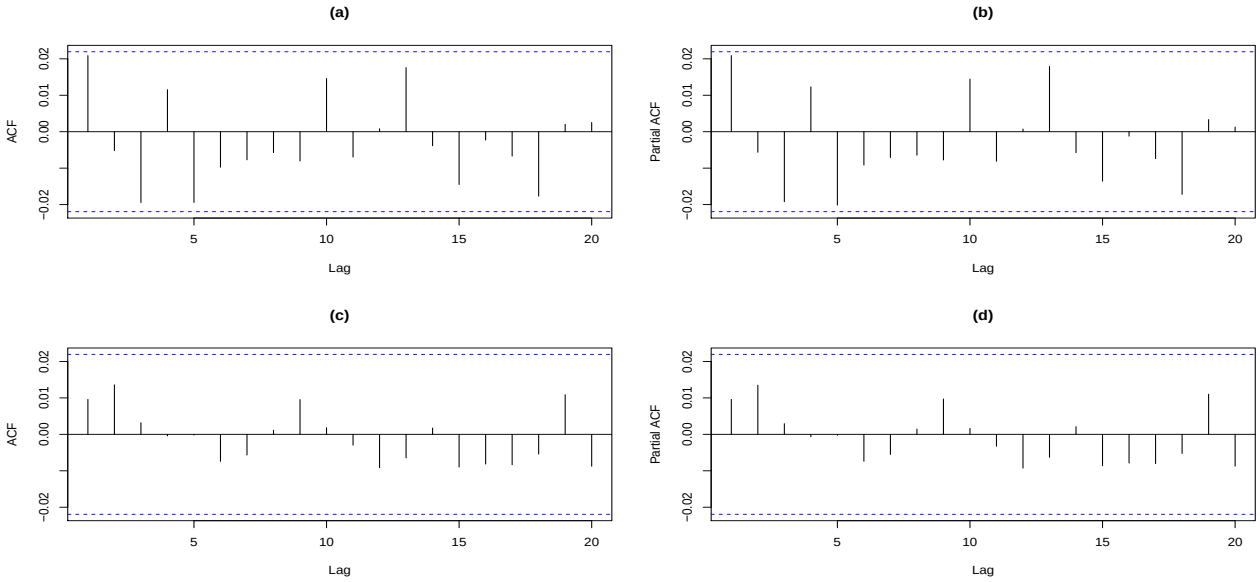

Figura 3.7: S&P500 - ACF e PACF para a série de resíduos padronizados $\widetilde{a}(t)$ e resíduos padronizados ao quadrado $[\widetilde{a}(t)]^2$.

Para checar se o ajuste do modelo foi satisfatório, foi aplicado também o teste Box-Pierce-Ljung com p-valor < 0.05, para os lags 1 a 10, não havendo evidências estatísticas de que exista correlação serial para os resíduos padronizados. Em seguida foi aplicado o Teste de Engle para checar a presença de efeito ARCH. Neste caso com p-valor < 0.05 e para os *lags* de 1 a 10, aceita-se a hipótese nula de que os dados não possuem efeito ARCH presente, ou seja, o modelo foi capaz de capturar a heterocedasticidade presente nos dados.

3.5.3 Índice Nasdaq

Com o objetivo de checar a aplicabilidade do modelo GARCH aos dados, na Figura 3.8, gráficos (a) e (b), são apresentadas a ACF e a PACF para a série de retornos do Índice Nasdaq. A análise destas funções indica a existência de uma dependência linear entre as observações, principalmente nos *lags* 1, 2, 12, 13, 16 e 18. Nos gráficos (c) e (d), são apresentadas respectivamente a ACF e a PACF para a série de retornos ao quadrado, onde é possível notar a presença de uma forte dependência. Tanto a ACF quanto na PACF tendem a diminuir vagarosamente, indicando que a variância do processo não é estacionária. Desta forma, como existe dependência tanto na série de retornos quanto na série de retornos ao quadrado (volatilidade), é possível concluir que o uso de um modelo $ARMA(p,q) - GARCH(w,z)$ é adequado ao conjunto de dados.

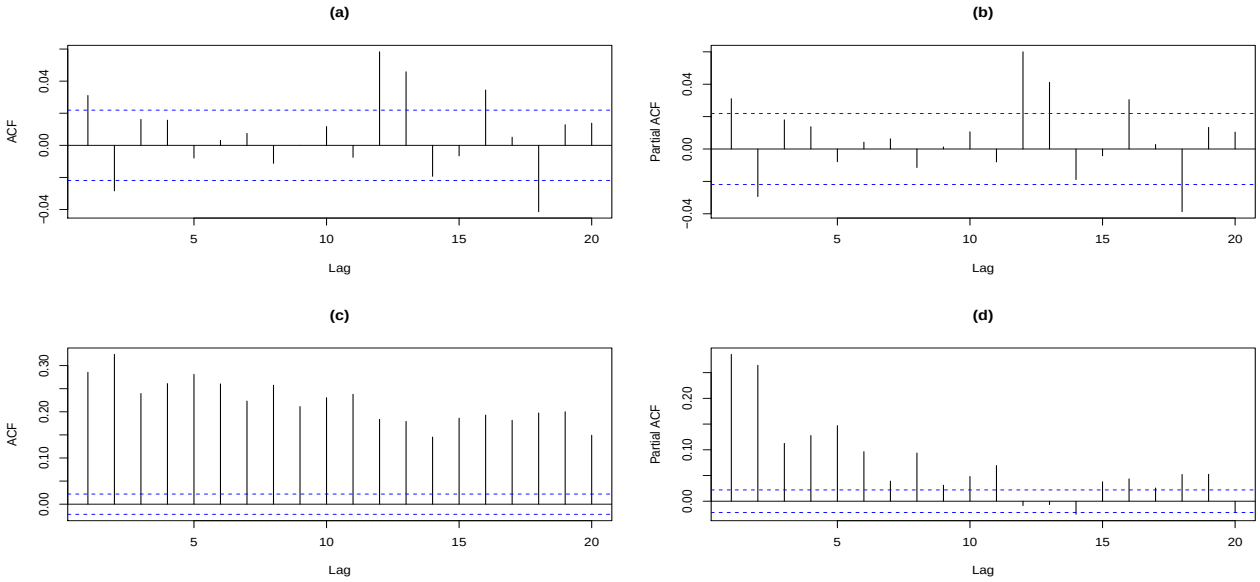

Figura 3.8: Nasdaq - ACF e PACF para a série de retornos e retornos ao quadrado.

Dada a aplicabilidade do modelo aos dados, um modelo $AR(18)$ foi ajustado para retirar a correlação na série de retornos. A ordem $p = 18$ utilizada é referente ao maior *lag* fora dos intervalos de confiança na ACF. Em seguida, após retirar os coeficientes não significativos do modelo, a um nível de 5%, um modelo $GARCH(1, 1)$ foi ajustado na série de resíduos. Por fim, após retirar novamente alguns coeficientes não significativos, o modelo final estimado é descrito Equação 3.13:

$$\begin{aligned} r(t) &= 0.0006 + 0.1422 \cdot r(t-1) + 0.0325 \cdot r(t-12) + a(t) \\ a(t) &= \sqrt{h_t} \cdot \varepsilon_t \\ h(t) &= 0.000001 + 0.1166 \cdot r^2(t-1) + 0.8769 \cdot h(t-1) \end{aligned} \quad . \tag{3.13}$$

A Figura 3.9 traz a comparação entre a volatilidade condicional $\sqrt{h(t)}$ estimada pelo modelo (curvas azuis) e a série de retornos (curva preta). Conforme é possível notar, a volatilidade estimada captura as principais flutuações na série de retornos. A volatilidade foi multiplicada respectivamente por 1.96 e -1.96, simulando um intervalo com 95% de confiança.

Por fim, resta ainda checar se o modelo foi capaz de capturar a heterocedasticidade condicional presente nos dados. Na Figura 3.10, (a) e (b), são apresentadas as ACF e PACF da série de resíduos padronizados ($\widetilde{a}(t)$). Nas letras (c) e (d), são apresentadas a ACF e a PACF da série de resíduos padronizados ao quadrado ($[\widetilde{a}(t)]^2$). Conforme podemos ver, os resíduos comportam-se como um ruído branco gaussiano, o que nos diz que o modelo foi capaz de capturar a heterocedasticidade presente nos dados, sendo portanto adequado.

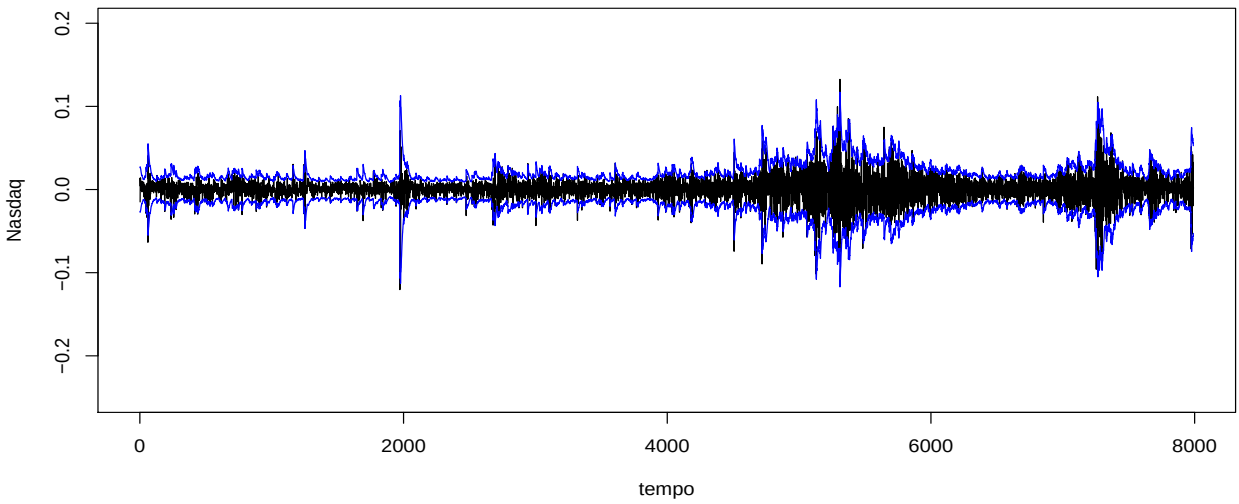

Figura 3.9: Nasdaq - Comparação entre a volatilidade estimada $\sqrt{h(t)}$ e os retornos $r(t)$

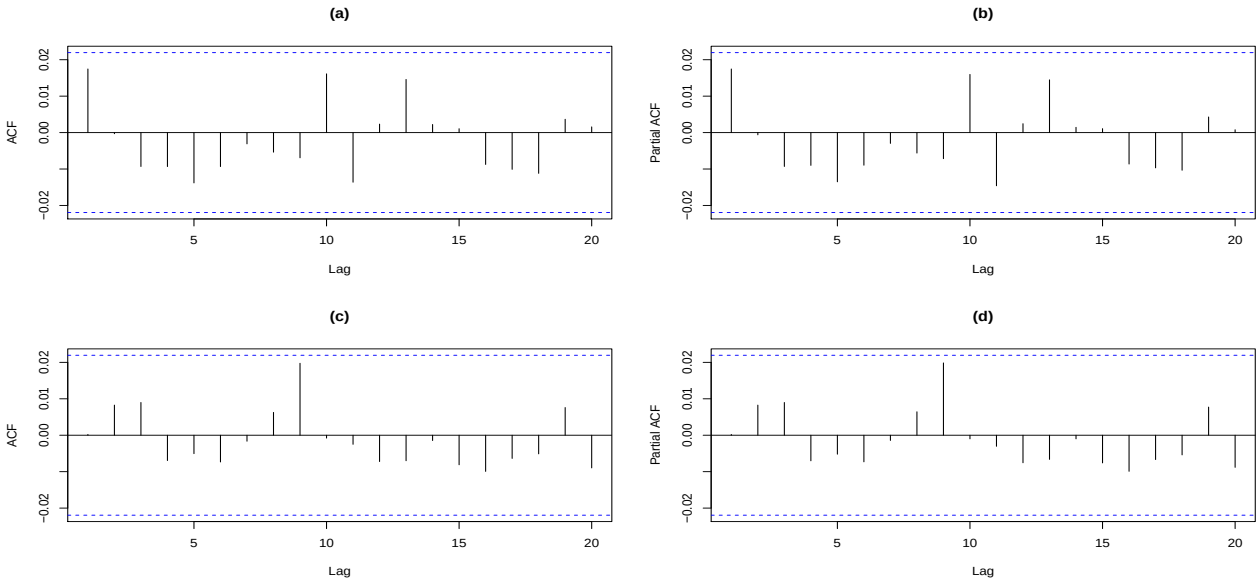

Figura 3.10: Nasdaq - ACF e PACF para a série de resíduos padronizados $\widetilde{a}(t)$ e resíduos padronizados ao quadrado $[\widetilde{a}(t)]^2$.

Para checar se o ajuste do modelo foi satisfatório, foi aplicado também o teste Box-Pierce-Ljung com p-valor < 0.05, para os *lags* 1 a 10, não havendo evidências estatísticas de que exista correlação serial para os resíduos padronizados. Em seguida foi aplicado o Teste de Engle para checar a presença de efeito ARCH. Neste caso com p-valor < 0.05 e para os *lags* de 1 a 10, aceita-se a hipótese nula de que os dados não possuem efeito ARCH presente, ou seja, o modelo foi capaz de capturar a heterocedasticidade presente nos dados.

3.5.4 Índice Ibovespa

Com o objetivo de checar a aplicabilidade do modelo GARCH aos dados, na Figura 3.11, gráficos (a) e (b), são apresentadas a ACF a PACF para a série de retornos do Índice Ibovespa. A análise destas funções indica a existência de uma dependência linear entre as observações, principalmente nos *lags* 3, 4, 10 e 15. Nos gráficos (c) e (d), são apresentadas respectivamente as ACF e as PACF para a série de retornos ao quadrado, onde é possível notar a presença de uma forte dependência. Tanto a ACF quanto na PACF tendem a diminuir vagarosamente, indicando que a variância do processo não é estacionária. Desta forma, como existe dependência tanto na série de retornos quanto na série de retornos ao quadrado (volatilidade), é possível concluir que o uso de um modelo $ARMA(p,q) - GARCH(w,z)$ é adequado ao conjunto de dados.

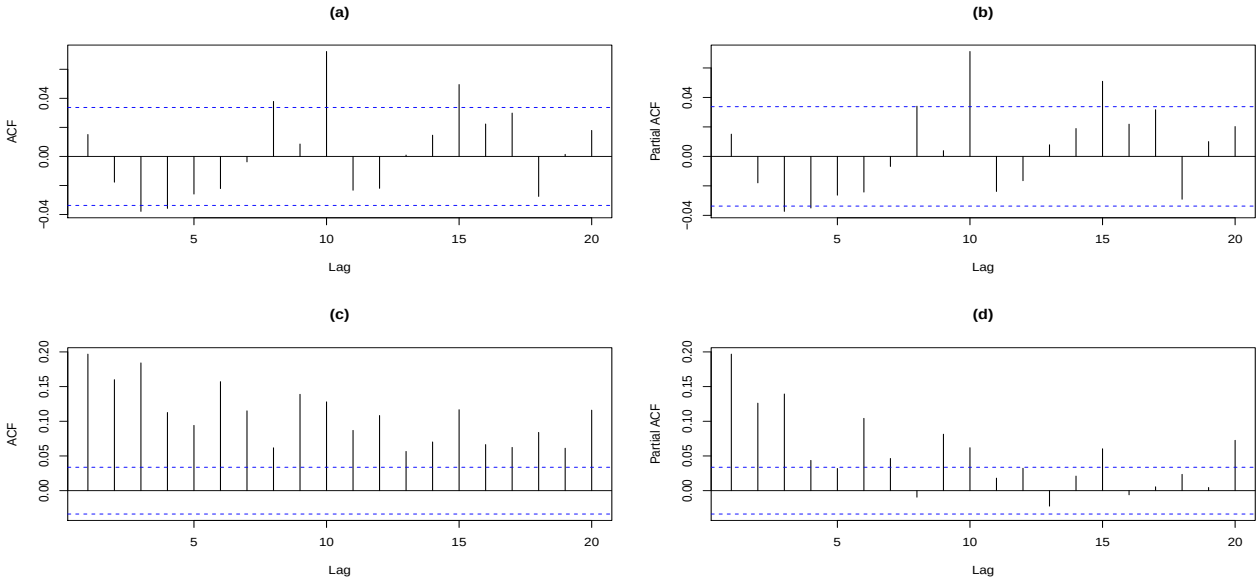

Figura 3.11: Ibovespa - ACF e PACF para a série de retornos e retornos ao quadrado.

Dada a aplicabilidade do modelo aos dados, um modelo $AR(15)$ foi ajustado para retirar a correlação na série de retornos. A ordem $p = 15$ utilizada é referente ao maior *lag* fora dos intervalos de confiança na ACF. Em seguida, após retirar os coeficientes não significativos do modelo, a um nível de 5%, um modelo $GARCH(2,1)$ foi ajustado na série de resíduos. Por fim, após retirar novamente alguns coeficientes não significativos, o modelo final estimado é descrito Equação 3.14:

$$\begin{aligned} r(t) &= 0.0009 + 0.0522 \cdot r(t-10) + a(t) \\ a(t) &= \sqrt{h_t} \cdot \varepsilon_t \\ h(t) &= 0,00001 + 0.0734 \cdot r^2(t-1) + 0.0385 \cdot r^2(t-2) + 0.8635 \cdot h(t-1) \end{aligned} \quad . \qquad (3.14)$$

A Figura 3.12 traz a comparação entre a volatilidade condicional $\sqrt{h(t)}$ estimada pelo modelo (curvas azuis) e a série de retornos (curva preta). Conforme é possível notar, a volatilidade estimada captura as principais flutuações na série de retornos. A volatilidade foi multiplicada respectivamente por 1.96 e -1.96, simulando um intervalo com 95% de confiança.

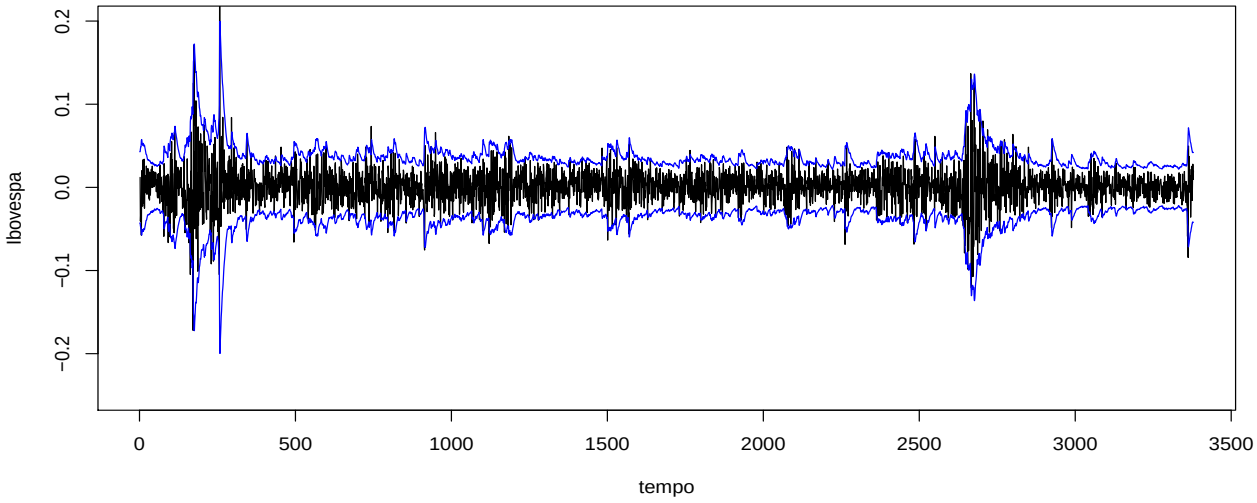

Figura 3.12: Ibovespa - Comparação entre a volatilidade estimada $\sqrt{h(t)}$ e os retornos $r(t)$

Por fim, resta ainda checar se o modelo foi capaz de capturar a heterocedasticidade condicional presente nos dados. Na Figura 3.13, (a) e (b), são apresentadas a ACF e a PACF da série de resíduos padronizados ($\widetilde{a}(t)$). Nas letras (c) e (d), são apresentadas as ACF e PACF da série de resíduos padronizados ao quadrado ($[\widetilde{a}(t)]^2$). Conforme podemos ver, os resíduos comportam-se como um ruído branco gaussiano, o que nos diz que o modelo foi capaz de capturar a heterocedasticidade presente nos dados, sendo portanto adequado.

Para checar se o ajuste do modelo foi satisfatório, foi aplicado também o teste Box-Pierce-Ljung com p-valor < 0.05, para os *lags* 1 a 10, não havendo evidências estatísticas de que exista correlação serial para os resíduos padronizados. Em seguida foi aplicado o Teste de Engle para checar a presença de efeito ARCH. Neste caso com p-valor < 0.05 e para os *lags* de 1 a 10, aceita-se a hipótese nula de que os dados não possuem efeito ARCH presente, ou seja, o modelo foi capaz de capturar a heterocedasticidade presente nos dados.

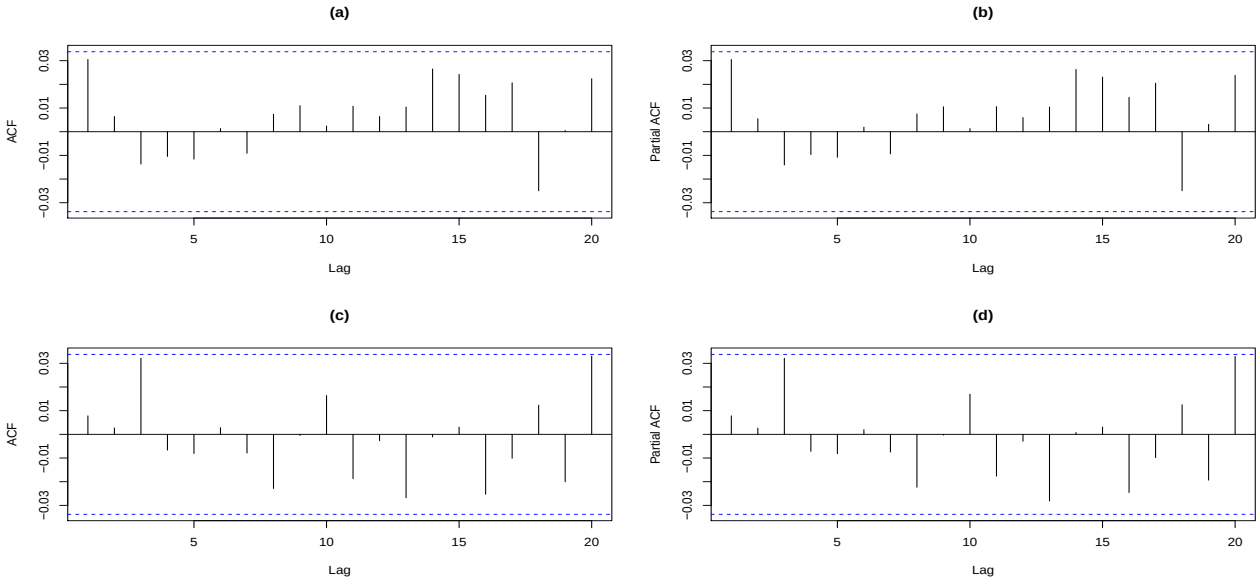

Figura 3.13: Ibovespa - ACF e PACF para a série de resíduos padronizados $\widetilde{a}(t)$ e resíduos padronizados ao quadrado $[\widetilde{a}(t)]^2$.

De uma maneira geral, os modelos ajustados para os índices Dow Jones, Nasdaq, S&P500 e Ibovespa, foram satisfatórios em estimar a volatilidade de grandes períodos de tempo. Para os três primeiros índices, foram estimadas as volatilidades para mais de 30 anos. Para o índice Ibovespa, foram considerados mais de 13 anos.

A comparação da volatilidade estimada com a série de retornos, permite verificar que os modelos ajustados capturaram os agrupamentos de volatilidade, levando em consideração as não-estacionariedades locais dos retornos. Além disso, a volatilidade estimada para os quatro índices mostrou-se estacionária para todo o período considerado.

Na literatura, os retornos são citados costumeiramente como sendo não-correlacionados serialmente [35]. No entanto, seguindo a abordagens de outros autores [36] [56] [57], notamos a necessidade de ajustarmos modelos do tipo ARMA nos quatro modelos, com o intuito de retirar a correlação serial na série de retornos e tornar a aplicação dos modelos GARCH mais confiável. A idéia aqui foi retirar qualquer resquício de correlação serial presente nas séries de retornos. Este procedimento está pautado na literatura, pois conforme aponta Tsay [56], os retornos diários apresentam geralmente correlações seriais inexistentes ou pouco expressivas. O autor afirma ainda que para retornos mensais, quase não já correlação serial presente nos dados.

Apesar de serem os primeiros modelos a considerar a dinâmica temporal da volatilidade, os modelos ARCH/GARCH apresentam alguns pontos fracos [36] [56] [57]. Estes modelos assumem que choques positivos e negativos na serie de retornos afetam a volatilidade da mesma maneira, uma vez que esta é modelada por meio do quadrado de retornos na iteração prévia. Os modelos ARCH/GARCH também não fornecem qualquer informação sobre a fonte de variação presente nos dados, sendo apenas uma fonte mecânica

de obtê-la. Por fim, estes modelos podem estimar grandes volatilidades espúrias (super previsões), uma vez que eles respondem lentamente a grandes choques isoladas para a série de retorno.

No próximo capítulo, serão apresentadas ferramentas pautadas no contexto multifractal. Estas ferramentas tem sido aplicadas no contexto financeiro e gerado importantes *insights* para o entendimento da dinâmica que governa os mercados financeiros.

Capítulo 4

Análise Fractal e Multifractal

4.1 Introdução

Compreender e caracterizar sistemas complexos não é uma tarefa fácil, em função de suas características intrínsecas. Conforme aponta Kantelhardt [11], não é possível dividir tais sistemas em subsistemas menores sem corromper suas propriedades dinâmicas. A maneira de caracterizar a dinâmica destes sistemas é através da identificação de leis de escala, que são válidas para um grande intervalo de tempo.

Se a dinâmica é identificada por meio de um único expoente de escala, dizemos que o processo é fractal. Neste caso, o capítulo abordará a Análise R/S de Hurst Clássica, utilizada para tratamento fractal de séries temporais estacionárias. Ainda no contexto fractal, o capítulo contemplará a técnica do DFA, alternativa à Análise R/S de Hurst. Esta técnica é mais robusta quanto à detecção espúria de correlações de longo alcance, além da vantagem de poder ser aplicada a séries temporais não estacionárias.

Por outro lado, se a dinâmica do processo for identificada por um conjunto de expoentes de escala, dizemos que o processo é multifractal. Para este caso, ferramentas como o MFDFA (generalização do DFA) e o Formalismo Multifractal e sua equivalência com a Termodinâmica, são úteis para caracterizar o comportamento multifractal presente em séries temporais. O Formalismo Multifractal terá importância ímpar no trabalho, uma vez que será utilizada a equivalência com a Termodinâmica para extrair medidas que auxiliarão na construção de indicadores de crashes de mercado.

4.2 Fractais, Auto-Afinidade e Leis de Escala

As leis de escala ganharam notoriedade ao longo do século XX, depois que a Mecânica Estatística voltou sua atenção para uma profunda descoberta por trás destas leis. Os físicos estatísticos perceberam que sistemas aparentemente diferentes podem apresentar comportamentos comuns. Várias grandezas físicas como magnetização, calor específico e susceptibilidade, quando medidas em condições próximas de uma transição de fase, apresentam divergência caracterizada por expoentes com valores muito específicos. Curiosamente, sistemas aparentemente distintos podem apresentar o mesmo expoente. Este fato deu origem ao conceito de classe de universalidade, onde sistemas com os mesmos expoentes são ditos pertencentes à mesma classe. Embora estes sistemas sejam diferentes em vários aspectos, suas propriedades básicas são independentes dos detalhes microscópicos. Este é um fato surpreendente e abriu as portas para se pensar no conceito de Ciência da Complexidade [58].

As leis de escala são regras matemáticas muito simples, estabelecendo que determinada variável, dependerá de outra por intermédio de uma função muito simples:

$$f(x) = A \cdot x^{\alpha}, \tag{4.1}$$

onde A é chamada de constante de normalização e α é o expoente de escala. Uma propriedade que torna esta função muito interessante é a homogeneidade, descrita na equação a seguir:

$$f(\rho x) = \rho^{\alpha} \cdot f(x). \tag{4.2}$$

.

Pode-se notar que a mudança de escala é absorvida na constante de normalização. Por isso, fenômenos descritos por $f(x)$ são ditos livres de escala. Esta invariância de escala ressalta o caráter universal da lei expressa por $f(x)$, chamada de lei de potência.

Existem diferentes e distintos mecanismos que levam a formação de leis de potência. O objetivo será estudar aqueles cuja natureza geométrica é caracterizada pelo conceito de conjuntos auto-similares e auto-afins [58].

Um conjunto de pontos S é dito auto-similar com respeito ao raio de escala ρ, se S for a união de n subconjuntos não sobrepostos $\{S_1, \ldots, S_n\}$, onde cada subconjunto de S é obtido por meio da transformação $S_i = \rho \cdot S$, com $i = \{1, \ldots, n\}$. Por outro lado, um conjunto de pontos S é dito auto-afim com respeito ao vetor raio de escala $\rho = \{\rho_1, \ldots, \rho_n\}$, se S for a união de N subconjuntos não sobrepostos $\{S_1, \ldots, S_n\}$, onde cada $S_i = \rho_i \cdot S$, com $i = \{1, \ldots, n\}$, é obtido de S, por meio de uma transformação auto-afim [59].

As definições de auto-afinidade e auto-similaridade permitem formalizar a idéia de

fractais e multifractais, introduzida por Benoit Mandelbrot nos anos 1960. Um fractal pode ser entendido como um objeto cujas subunidades se assemelham ao todo. Por trás dessa definição está a condição de invariância geométrica do objeto quando observado em escalas diferentes. Ao se fazer um zoom em alguma parte do objeto com padrão fractal, ocorre a contração (ou dilatação) de forma isotrópica. Rigorosamente, pode-se dizer que a auto-similaridade é uma propriedade de simetria que torna o fractal invariante sob uma transformação geométrica isotrópica. Existem fractais, contudo, que são igualmente formados por cópias, mas estas não mantêm fixas as proporções originais, ou seja, eles são invariantes sob transformações anisotrópicas. Neste caso, eles são chamados de fractais auto-afins.

A diferença básica entre as auto-similaridade e auto-afinidade é que a primeira exige apenas um fator de escala, enquanto a segunda exige pelo menos dois fatores distintos, o que evoca a idéia de anisotropia (tendência ou viés em pelo menos uma direção). A Figura 4.1 exibe um exemplo para cada um destes casos.

Figura 4.1: A Figura da esquerda apresenta o Conjunto de Cantor (conjunto auto-similar), enquanto a Figura da direita apresenta o Brócolis (conjunto auto-afim)

A geometria fractal mostrou-se útil para descrever contornos irregulares ou superfícies rugosas. Estas irregularidades ou rugosidades são descritas pela dimensão fractal que mede o grau de ocupação ou cobertura espacial de um conjunto de pontos.

Tem-se claramente a noção que um ponto tem dimensão zero, enquanto uma superfície tem dimensão dois e um cubo tem dimensão três. Isto é o que se chama de dimensão topológica. A dimensão fractal pode ser entendida como algo que envolve a dimensão espacial no sentido de tentar capturar o espaço preenchido ou ocupado por uma figura e admite valores fracionários para esta medida.

As séries temporais também apresentam características fractais [11, 14, 18, 34, 41]. A função $x(t)$ dada na Equação 4.3, geralmente estudada em análise matemática, é contínua e possui derivadas contínuas. Para pequenas regiões ao redor de t_i, são necessários poucos termos da expansão para aproximar a função $x(t)$:

$$x(t) = \alpha_0 + \alpha_1(t - t_i) + \alpha_2(t - t_i)^2 + \ldots + \alpha_h(t - t_i)^{h_i}. \tag{4.3}$$

onde t pertence a uma pequena vizinhança de t_i e h_i é um número não inteiro quantificando a singularidade local de $x(t)$, em $t = t_i$.

Como em outros contextos (biológicos, físicos, etc.), determinadas séries temporais também apresentam comportamento fractal. Nestas séries, alguns pontos não podem ser aproximados pela expansão de Taylor na Equação 4.3, utilizando-se apenas poucos termos na expansão. Além disso, muitas séries temporais empíricas possuem características fractais, ou seja, para dado t_i, x(t) apresenta comportamento singular. Isto significa que para determinado período t_i, a série apresenta componentes com potências não inteiras, que se caracterizam por degraus ou cúspides (angulações ou pontas) na série, chamadas de singularidades. A Figura 4.2 apresenta alguns exemplos deste comportamento. A série (**a**) é um ruído com correlações de longo alcance e regime do tipo lei de potência, gerado aleatoriamente. A série (**b**) apresenta os batimentos cardíacos de um paciente saudável. A série (**c**) apresenta outro ruído gerado aleatoriamente, com correlações de longo alcance e espectro de potência que decresce com a potência. Nas séries, o formato de cúspide presente nos dados é destacado em verde, enquanto que os degraus são destacados em azul [60].

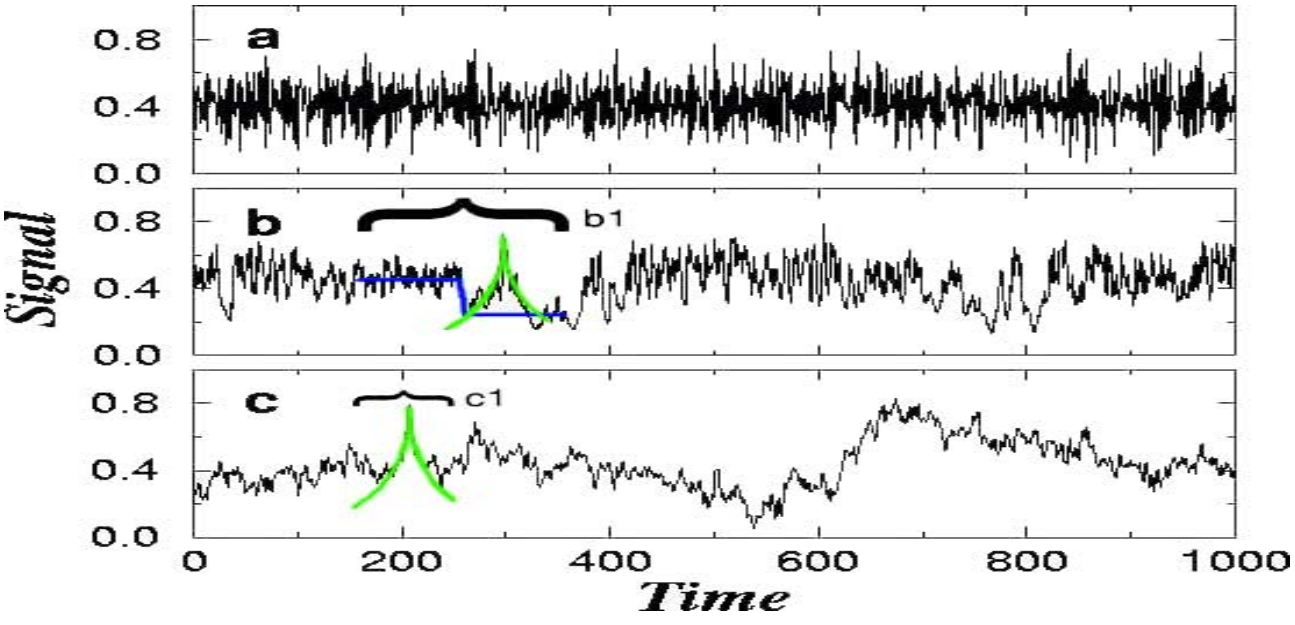

Figura 4.2: Característica fractal presente em séries reais e simuladas

4.3 Análise R/S de Hurst Clássica

A Análise R/S de Hurst foi proposta por Henry Hurst em 1951 [39] [40], com o intuito de testar a presença de correlações de longo alcance em séries temporais empíricas. A idéia por trás da técnica é olhar para o comportamento escalar dos desvios cumulativos reescalados pela média (intervalo em que o sistema varia em função do tempo) e comparar com a hipótese nula de que a série é aleatória. Isto é feito por meio do expoente H de Hurst estimado [34] [41].

Seja $x(t) = \{x(1), x(2), \ldots, x(S)\}$ uma série temporal de comprimento S. O primeiro passo consiste em aplicar a Equação 3.3 para transformar $x(t)$ numa série de log-retornos de comprimento $L = S - 1$: $r(t) = \{r(1), ..., r(L)\}$.

O segundo passo consiste em dividir $r(t)$ em N segmentos não sobrepostos de tamanho s, onde $N = int(L/s)$. Cada segmento é chamado de v, com $v = \{1, 2, \ldots, N\}$. Para cada subperíodo v, calcula-se a média $\overline{r}_v$:

$$\overline{r}_v = \frac{1}{s}\sum_{i=1}^{s}\{r[(v-1)s+i]\}. \tag{4.4}$$

O terceiro passo consiste em calcular as diferenças acumuladas $X_v(s)$, para cada intervalo v:

$$X_v(s) = \sum_{i=1}^{s}\{r[(v-1)s+i] - \overline{r}_v\}. \tag{4.5}$$

É possível notar que a série de valores $X_v(s)$ na Equação 4.5, sempre terminará com zero para cada intervalo v de comprimento s [34].

O quarto passo será computar na Equação 4.6, a diferença entre o máximo e mínimo valor de $X_v(s)$, representando o intervalo coberto pela série temporal em relação à média, em cada intervalo v:

$$R_v(s) = max[X_v(s)] - min[X_v(s)]. \tag{4.6}$$

Em seguida, no quinto passo, calcula-se o desvio padrão em cada intervalo v:

$$S_v(s) = \sqrt{\frac{1}{s}\sum_{i=1}^{s}\{r[(v-1)s+i] - \overline{r}_v\}^2}. \tag{4.7}$$

O sexto passo consiste em calcular a estatística $(R/S)_s$, para os intervalos de comprimento s:

$$(R/S)_s = \frac{1}{N}\sum_{v=1}^{N}\left[\frac{R_v(s)}{S_v(s)}\right]. \tag{4.8}$$

O sétimo e último passo consiste em determinar o comportamento do tipo lei de potência, ao analisar o gráfico log-log de $(R/S)_s$ versus s. Esta relação é obtida ao repetir os seis passos anteriores para diferentes tamanhos de janela s (diferentes escalas). Desta forma, o interesse estará em determinar o expoente H na relação:

$$(R/S)_s \sim s^{\,H}. \tag{4.9}$$

O parâmetro H, chamado de Expoente de Hurst, representa as propriedades de correlação da série temporal, da seguinte maneira:

- Se $H = 0.5$, não existem correlações de longo alcance presentes na série.
- Se $0 < H < 0.5$, a série é antipersistente e possui autocorrelações negativas.
- Se $0.5 < H < 1$, a série é persistente e possui autocorrelações positivas.

O Expoente de Hurst está diretamente relacionado à Dimensão Fractal D_f, que é a medida da rugosidade do sistema considerado. A relação é dada pela Equação 4.10:

$$D_f = 2 - H. \tag{4.10}$$

Quanto menor o expoente H, maior o valor de D_f, indicando maior rugosidade. Consequentemente, a relação contrária também é válida.

O Coeficiente Espectral β e o Coeficiente de Correlação γ, também estão relacionado ao expoente H, por meio das Equações 4.11 e 4.12 respectivamente [11]:

$$\beta = 2H - 1, \tag{4.11}$$

$$\gamma = 2 - 2H. \tag{4.12}$$

Note que $0 < \gamma < 1$, portanto a relação apresentada na Equação 4.12 só vale para o caso em que $0.5 < H < 1$ [11].

4.4 Detrended Fluctuation Analysis - DFA

O DFA é uma alternativa à Análise R/S de Hurst. Esta técnica foi proposta inicialmente com o objetivo de detectar correlações de longo alcance em sequências de DNA [42] [43]. Além disso, esta técnica tem sido utilizada para determinar propriedades de escala monofractal. A técnica ainda traz consigo a vantagem de evitar a detecção espúria de correlações de longo alcance, quando esta é um artefato da não-estacionariedade. [12]

Seja $x(t) = \{x(1), x(2), \ldots, x(S)\}$ uma série temporal de comprimento S. No primeiro passo, aplica-se a Equação 3.3 para transformar $x(t)$ em uma série de log-retornos de comprimento $L = S - 1$: $r(t) = \{r(1), ..., r(L)\}$.

No segundo passo, deve-se calcular a série integrada $Y(i)$:

$$Y(i) = \sum_{k=1}^{i} \{r(k) - \overline{r}\}, \tag{4.13}$$

onde $\overline{r}$ é a média da série de retornos e $i = \{1, \ldots, L\}$. Não é um requisito obrigatório subtrair a média $\overline{r}$ na Equação 4.13, desde que esta seja eliminada no tratamento das tendências locais, no terceiro passo [11].

O terceiro passo consiste em dividir $Y(i)$ em N segmentos não sobrepostos de tamanho s, onde $N = int(L/s)$. Cada segmento é chamado de v, com $v = \{1, 2, \ldots, N\}$.

O quarto passo consiste em calcular, para cada segmento v, a *Função de Flutuação* ou Variância:

$$F^2(s, v) = \frac{1}{s} \sum_{i=1}^{s} \{Y[(v-1)s + i] - y_v^m(i)\}^2, \tag{4.14}$$

com $y_v^m(i)$ representando o ajuste polinomial de ordem m (tendência local) em cada segmento v. Pode-se utilizar polinômios de ordem $m = 1, 2, 3, \ldots$, obtendo respectivamente os métodos $DFA1, DFA2, DFA3, \ldots$.

O quinto passo consiste em calcular a Função de Flutuação Média, para a janela v de tamanho s:

$$F_2(s) = \left(\frac{1}{N} \sum_{v=1}^{N} [F^2(s, v)] \right)^{\frac{1}{2}}. \tag{4.15}$$

O sexto passo consiste em determinar o comportamento escalar da Função de Flutuação Média, ao analisar o gráfico log-log de $F_2(s)$ versus s. Esta relação é obtida ao repetir passos 3 a 5, para diferentes tamanhos de janela s (diferentes escalas). Se a série temporal $x(t)$ possui correlações de longo alcance, o valor de $F_2(s)$ aumenta à medida que o tamanho da janela s aumenta, de acordo com uma lei de potência:

$$F_2(s) \sim s^{\alpha}. \tag{4.16}$$

O parâmetro α, chamado de Expoente de Escala ou de Correlação, representa as propriedades de correlação da série temporal, da seguinte maneira:

- Se $\alpha = 0.5$, não existem correlações de longo alcance presentes na série
- Se $0 < \alpha < 0.5$, a série é antipersistente e possui correlação negativa
- Se $0.5 < \alpha < 1$, a série é persistente e possui correlação positiva

A obtenção do parâmetro α utilizando as Equações 4.15 e 4.16, representa o método de cálculo padrão do DFA [42]. A generalização da Equação 4.15, chamada de Função de Flutuação Média Generalizada, para momentos de ordem q é dada pela Equação 4.17. Este método é chamado de Multifractal Detrended Fluctuation Analysis ou simplesmente de MFDFA [12] é definido como:

$$F_q(s) = \left(\frac{1}{N} \sum_{v=1}^{N} [\{F^2(s,v)\}^{q/2}] \right)^{\frac{1}{q}}, \tag{4.17}$$

onde a variável q pode assumir qualquer valor real exceto zero. O valor de $F_q(s)$ não pode ser determinado diretamente quando $q = 0$, por causa do expoente divergente da Equação 4.17. Este problema é contornado ao utilizar a Equação 4.18:

$$F_0(s) = exp\left(\frac{1}{N} \sum_{v=1}^{N} ln[F(s,v)] \right). \tag{4.18}$$

O interesse está em determinar o quanto a Função de Flutuação Média Generalizada $F_q(s)$ depende da escala temporal s, para diferentes valores de q. Assim como observado na Equação 4.16, $F_q(s)$ deve seguir uma lei de potência:

$$F_q(s) \sim s^{h(q)}. \tag{4.19}$$

Para séries temporais estacionárias, $h(2)$ é idêntico ao Expoente de Hurst H. Então, $h(q)$ é chamada de Expoente de Hurst Generalizado [11].

É possível mostrar que o Expoente de Escala Multifractal $h(q)$, definido na Equação 4.19 relaciona-se com o expoente de escala $\tau(q)$, definido pela Função de Partição (Equação 4.25), no formalismo multifractal [54]. Esta relação é computada pela Equação 4.20:

$$\tau(q) = qh(q) - 1. \tag{4.20}$$

É possível ainda relacionar $h(q)$ com a Dimensão Multifractal Generalizada $D(q)$ [15]:

$$D(q) \equiv \frac{\tau(q)}{q-1} = \frac{qh(q)-1}{q-1}. \tag{4.21}$$

Outra maneira de caracterizar o comportamento multifractal de séries temporais é calcular o Espectro Multifractal Singular $f(\alpha)$, relacionado com $\tau(q)$ via transformada de Legendre:

$$\begin{cases} \alpha(q) = d\tau(q)/dq \\ f(\alpha) = q\alpha - \tau(q) \end{cases} . \tag{4.22}$$

4.5 Formalismo Multifractal e Termodinâmica

O objetivo deste tópico será apresentar a interpretação termodinâmica de multifractalidade, à luz dos resultados obtidos por Enrique Canessa [14, 15]. Baseado no conceito de Calor Específico Análogo $C(q)$, calculado para o índice S&P500, o autor argumenta que é possível determinar com antecedência a ocorrência de um *crash* de mercado. A seguir, são descritos os passos para implementar o método. Seja $x(t) = \{x(1), x(2), \ldots, x(S)\}$ uma série temporal financeira de comprimento S.

O primeiro passo consiste em aplicar a Equação 3.1 em módulo, para transformar $x(t)$ na série de incrementos absolutos de comprimento $L = (S - 1)$. O período utilizado aqui é $T = 1$, pois de acordo com a notação econômica, somente os incrementos $\Delta x(t, T = 1) = \Delta x(t)$ são usualmente relevantes [15]. Esta transformação é representada na Equação 4.23:

$$|\Delta x(t)| = |x(t + 1) - x(t)|, \tag{4.23}$$

onde $t = 1, ..., L$.

O segundo passo consiste em construir a série de medidas normalizadas $\mu(t) > 0$:

$$\mu(t, L) = \frac{|\Delta x(t)|}{|\sum_{t=1}^{L} \Delta x(t)|}, \tag{4.24}$$

onde $t = 1, ..., L$.

O terceiro passo tem como objetivo a construção da Função de Partição $Z(q, L)$, baseada nas medidas normalizadas $\mu(t)$ e nos momentos de ordem q, conforme descrito na Equação 4.25:

$$Z(q, L) = \sum_{t=1}^{L} \mu(t)^q \sim L^{-\tau(q)}. \tag{4.25}$$

Na Equação 4.25, $\tau(q)$ representa a Energia Livre Análoga. Uma vez computado o valor de $Z(q, L)$, o quarto passo consiste em extrair o vetor $\tau(q)$ por meio da relação:

$$\tau(q, L) = \frac{lnZ(q, L)}{lnL}. \tag{4.26}$$

Em seguida, para obter a interpretação termodinâmica de multifractalidade, considera-

se a Equação 4.21, de onde é possível extrair a Dimensão Multifractal $D(q)$. É possível ainda calcular o Espectro multifractal $f(\alpha)$, por meio da Transformada de Legendre definida na Equação 4.22.

O quinto passo é referente ao cálculo do Calor Específico Análogo $C(q)$, dado pela derivada de segunda ordem de $\tau(q)$, conforme descrito na Equação 4.27:

$$- C(q, L) \equiv \frac{\partial^2 \tau(q)}{\partial q^2}. \tag{4.27}$$

As Figuras 4.3 e 4.4 apresentam o Calor Específico Análogo $C(q)$ para os Índices Dow Jones e S&P500, de Jan/1980 a Dez/1992; período que compreende o *crash* Black Monday (19 de outubro de 1987). Para ambas as figuras, com o objetivo de reproduzir os resultados apresentados por Canessa [15], o cálculo de $C(q)$ foi feito sob as seguintes condições: (a) Utilizando os primeiros 499 dias da série de incrementos (Jan/1980 a Dez/1981). Neste caso, $C(q)$ apresenta um formato de sino, uma vez que esta curva é calculada para um período em que os dados não são influenciados pelo *crash* de 1987. (b) Utilizando os primeiros 1971 dias de dados (Jan/1980 a 18/10/1987), onde o conjunto de dados vai até um dia antes do *crash*. Neste caso, é possível observar o início da formação de um "ombro" anômalo à direita da curva $C(q)$. (c) Utilizando 3286 dias de dados (Jan/1980 a Dez/1992), onde apenas o dia 19/10/1987 foi excluído da análise. Neste caso, nota-se com mais clareza a formação de um "ombro anômalo" à direita de $C(q)$. A medida $C(q)$ é robusta, pois mesmo sem a presença do *crash* nos dados, ela é capaz de capturar a influência deste nos dados, em função do aparecimento do "ombro anômalo" à direita da curva. A medida é portanto capaz de capturar a influência do *crash* nos dados, antes e depois deste ocorrer. (d) Utilizando todos os 3287 dias de dados (Jan/1987 a Dez/1992), onde nenhum dia foi excluído da análise.

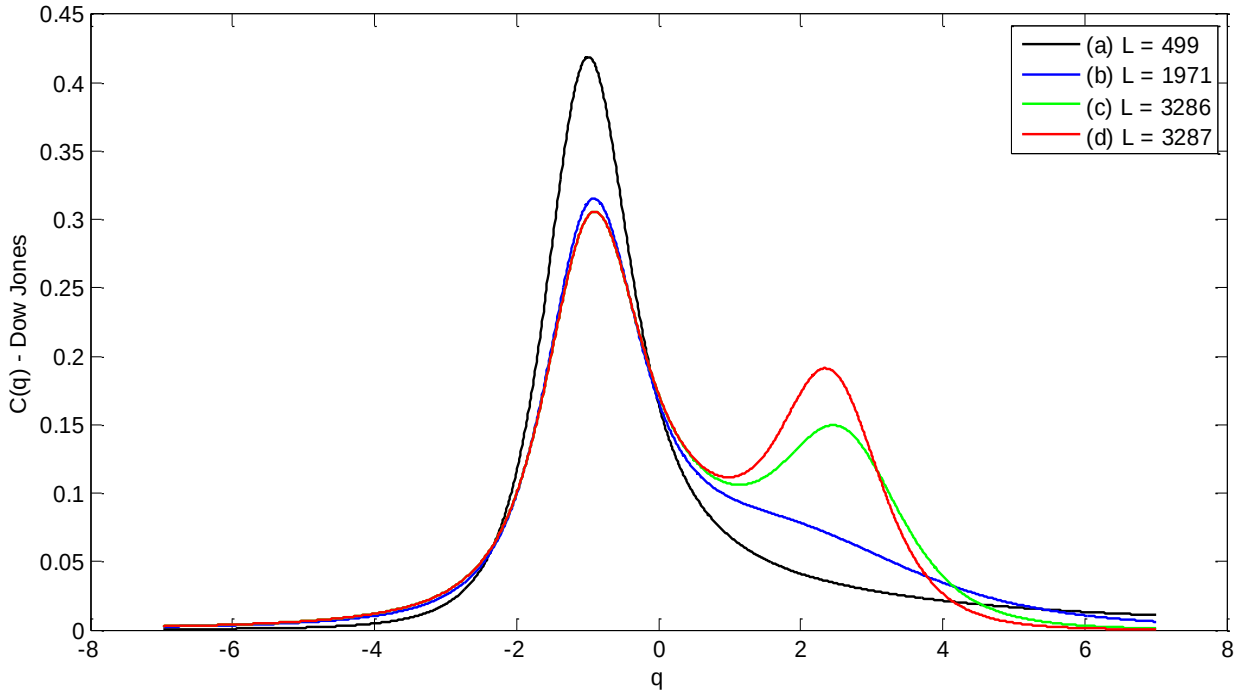

Figura 4.3: Índice Dow Jones - Curvas do Calor Específico Análogo $C(q)$.

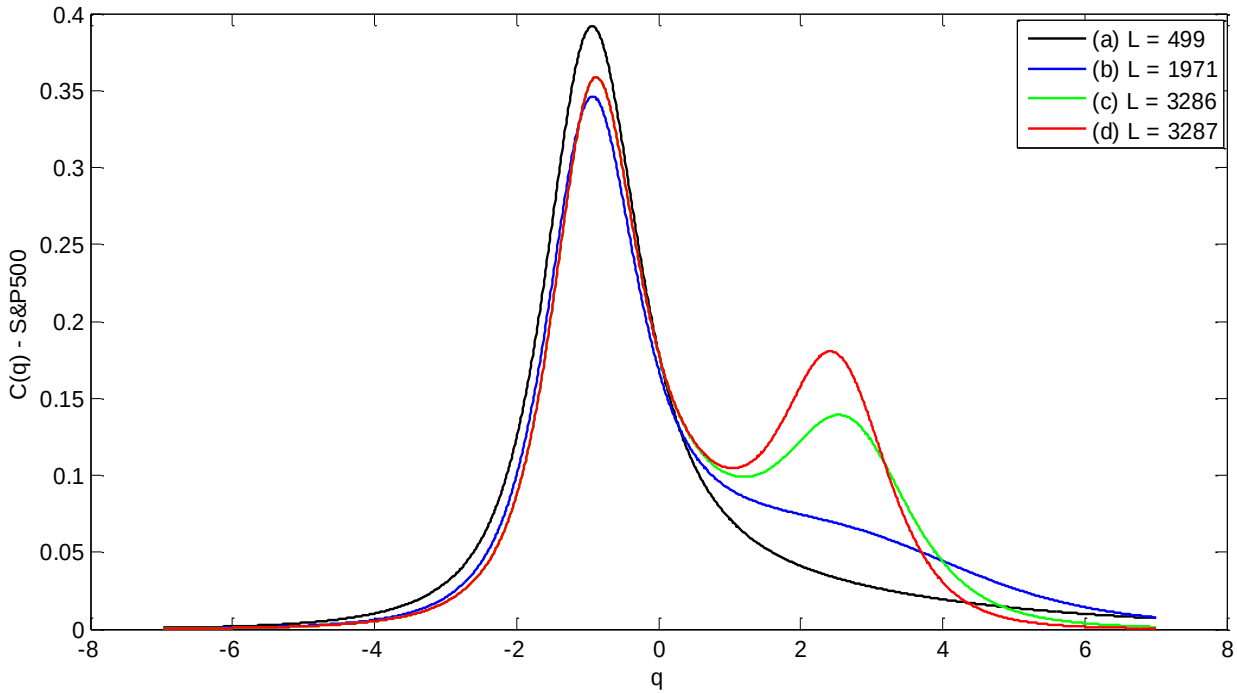

Figura 4.4: Índice S&P500 - Curvas do Calor Específico Análogo $C(q)$.

4.6 Algoritmo Online

Conforme visto anteriormente, $C(q)$ é capaz de gerar informações interessantes sobre a iminência de um *crash* de mercado. Desta forma, o interesse estará na dinâmica do comportamento de $C(q)$, calculado para uma janela temporal de tamanho s, que se desloca a uma taxa constante sobre a série de incrementos absolutos $|\Delta x(t,1)|$. A idéia é que a janela sempre agregue informações mais recentes à medida que descarta informações antigas. Este método, chamado de "Sliding Observation Box", foi empregado por Peng et al. [42] no cálculo do *DFA* para sequências de DNA e traz consigo a vantagem de gerar melhores estatísticas que o método original, em que se divide o conjunto de dados em $N = int(L/s)$ segmentos não sobrepostos [43].

A técnica do "Sliding Observation Box" é portanto muito interessante para o presente trabalho, uma vez que fornece um método de cálculo dinâmico para computar $C(q)$ para uma janela de tamanho s que se desloca temporalmente. Daqui por diante, o método será chamado de Algoritmo Online.

Existirão ao todo $v = (L-s+1)$ janelas de tamanho s, supondo deslocamento unitário. É importante ressaltar que no Algoritmo Online, as medidas μ, Z, τ e C são sempre calculadas para uma janela i de tamanho s, subconjunto da série de Incrementos Absolutos de tamanho L. Outra questão é que a cada iteração, as medidas são recalculadas, pois a janela i é dinâmica (sempre agregando e descartando informações). Para diferenciar as medidas em cada iteração, foi empregada uma nova notação: $\mu_i(t,s)$, $Z_i(q,s)$, $\tau_i(q,s)$ e $C_i(q,s)$, com $i = 1,\ldots,v$. A fórmula para o cálculo de $\mu_i(t,s)$, $Z_i(q,s)$ e $\tau_i(q,s)$ dentro da janela v de tamanho s é ligeiramente diferente do apresentado nas Equações 4.24 e 4.25 e 4.26, conforme descrito nas Equações 4.28, 4.29 e 4.30 respectivamente. Não há alteração para o cálculo de $C_i(q,s)$ (Equação 4.27):

$$\mu_i(t,s) = \frac{|\Delta x(t,1)|}{|\sum_{t=1}^{s} \Delta x(t,1)|}, \tag{4.28}$$

$$Z_i(q,s) = \sum_{i=1}^{s} \mu(t)^q \sim s^{-\tau(q)}, \tag{4.29}$$

$$\tau_i(q,s) = \frac{lnZ_i(q,s)}{lns}, \tag{4.30}$$

onde, $t = 1,2,\ldots,s$ e $i = 1,2,\ldots,v$.

A Figura 4.5 apresenta um esquema para a dinâmica do Algoritmo Online, executado para uma série de incrementos absolutos.

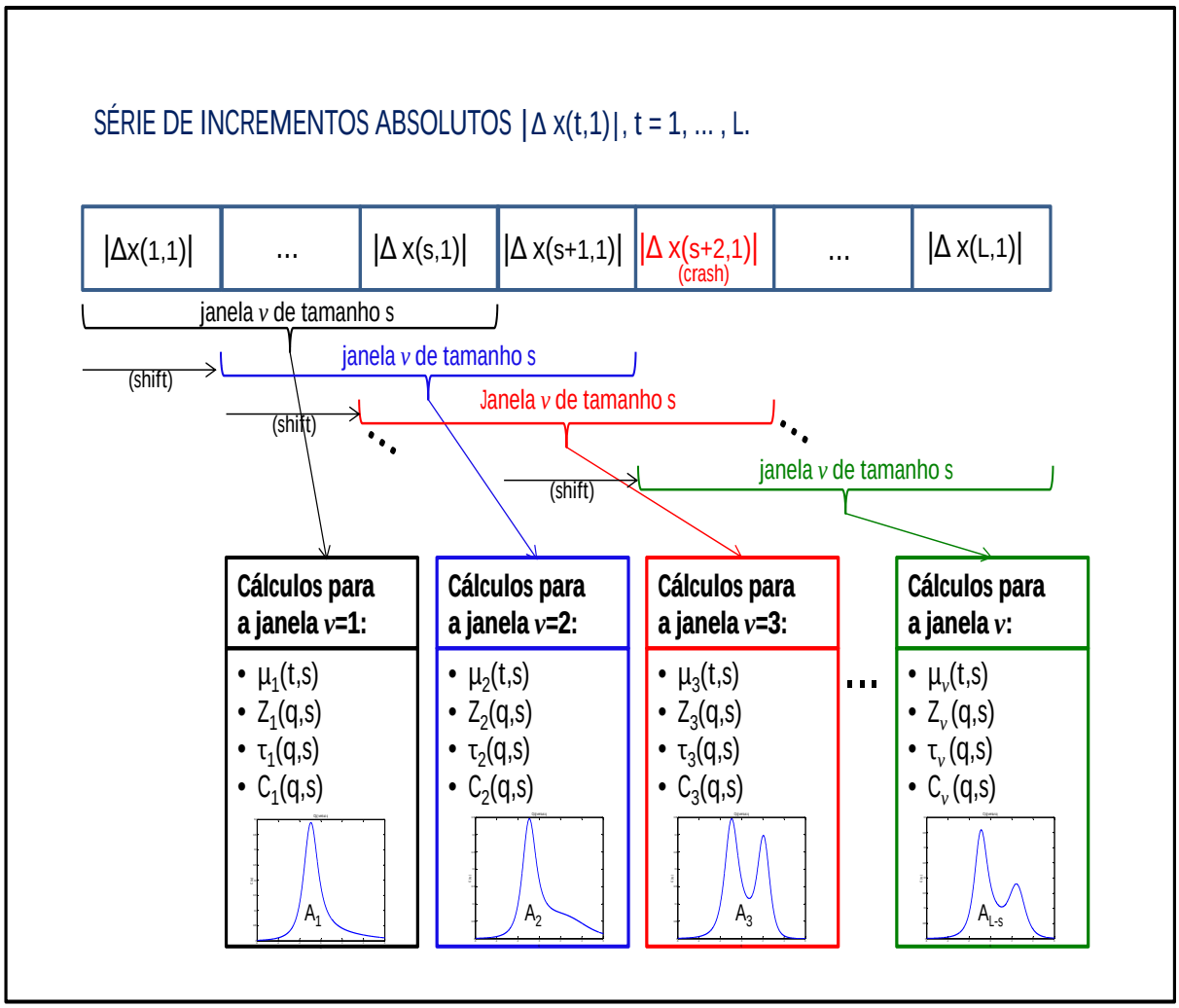

Figura 4.5: Esquema para a evolução temporal de $C(q)$ do Algoritmo Online.

4.7 Indicadores de *Crash*

Como visto anteriormente, o comportamento de $C(q)$ é capaz de fornecer com certa antecedência, indícios sobre a ocorrência de um *crash* de mercado. O interesse reside portanto em traduzir em algum número ou medida, o comportamento empírico de $C(q)$. Conforme visto nas Figuras 4.3 e 4.4, a área sob as curvas de $C(q)$ parece aumentar à medida que são utilizadas informações próximas ou que contemplem o dia do *crash* (em função do aparecimento do ombro à direita da curva $C(q)$). Assim, uma forma intuitiva de traduzir este comportamento seria calcular a Área *A(i)* sob a curva $C_i(q, s)$, onde $i = 1, 2, \ldots, v$ no Algoritmo Online. O objetivo então será aliar a dinâmica do Algoritmo Online com o comportamento de $C_i(q, s)$ (resumido por *A(i)*), na tentativa de obter com antecedência informações sobre a iminência de um *crash* de mercado.

A Figura 4.6 apresenta a evolução temporal da Área $A(i)$ sob a curva $C_i(q, s)$, dos índices Dow Jones, S&P500 e Nasdaq, à medida que a janela v de tamanho $s = 1922$ se desloca sob a série de incrementos no Algoritmo Online. Os dados utilizados referem-se ao período de Jan/1980 a Dez/1992, contemplando o *crash* Black Monday, de 19/10/1987 (ponto 50 do gráfico). Analisar os dados desta forma permite acompanhar a evolução temporal de $A(i)$ num período de 100 dias: 50 dias antes e 50 dias depois do *crash*.

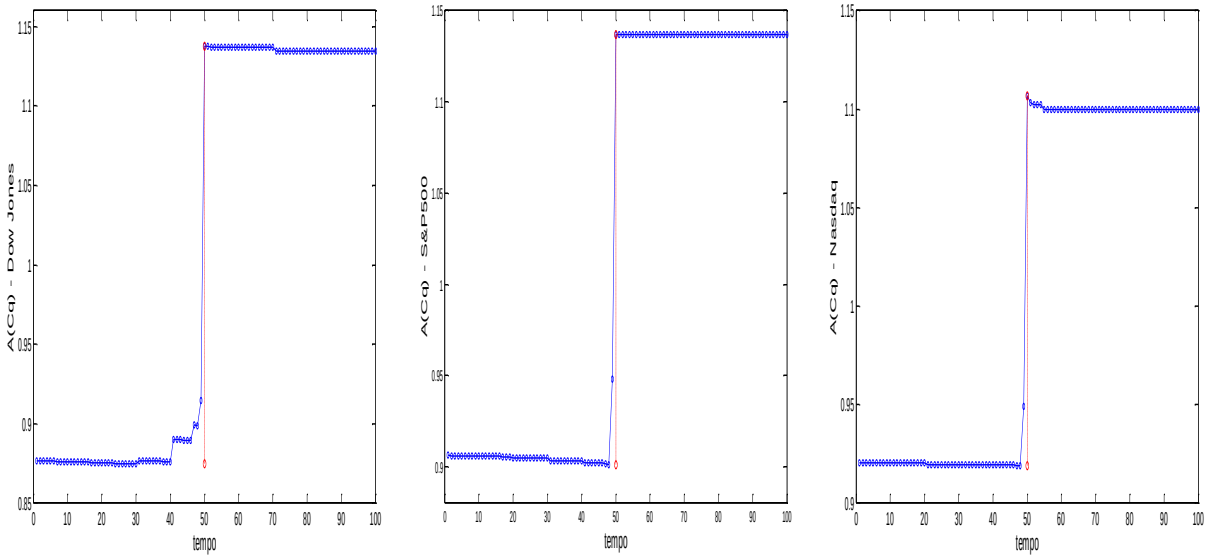

Figura 4.6: Evolução temporal da Área sob $C(q)$ - Índice Dow Jones, Índice S&P500 e Índice Nasdaq.

Seria conveniente que $A(i)$ apresentasse um comportamento característico único. Um comportamento desejado seria, por exemplo, que os valores de $A(i)$ aumentassem com bastante antecedência à medida que a janela móvel se aproxima do *crash*. Como é possível ver na Figura 4.6, este comportamento é evidente para o índice Dow Jones, onde $A(i)$ começa a aumentar à pelo menos 20 dias do *crash*. Este comportamento no entanto, não é observado para os índice S&P500 e Nasdaq, cujas Áreas sob $C_i(q, s)$ diminuem até a véspera do *crash*, dia em que $A(i)$ aumenta consideravelmente. Outra desvantagem de tomar $A(i)$ como o indicador de *crash*, é que esta possui uma memória muito longa, ou seja, mesmo quando o Algoritmo Online agrega informações que estão bem distantes da data do *crash*, a Área $A(i)$ permanece alta em relação ao período antes do *crash*.

Na tentativa de obter um comportamento uniforme nas proximidades de um *crash* de mercado, foram cogitados outros indicadores que são função de $A(i)$. O objetivo é uniformizar a informação gerada por $A(i)$, de maneira a obter um indicador robusto que detecte a iminência de um *crash* em diferentes índices financeiros, ativos, períodos de tempo e mercados.

O primeiro indicador refere-se à variação percentual entre o atual valor de $A(i)$ e a média dos valores passados de $A(i)$, calculados anteriormente no Algoritmo Online. Este indicador é definido na Equação 4.31:

$$I_A(i) = \left| \frac{A(i)}{\overline{A}(1:i)} - 1 \right| = \left| \frac{A(i) - \overline{A}(1:i)}{\overline{A}(1:i)} \right|, \tag{4.31}$$

onde o passo i é definido à medida que a janela de tamanho s se desloca no Algoritmo Online. Na construção do Índice $I_A(i)$, o valor absoluto foi tomado, pois para o cálculo do $C(q, s)$, foram utilizados incrementos absolutos. Porém, esta condição deverá ser testada e relaxada posteriormente, caso seja necessário.

A Figura 4.7 apresenta a evolução temporal do indicador $I_A(i)$ para os índices econômicos Dow Jones, S&P500 e Nasdaq, para o período de Jan/1980 a Dez/1987, com janela móvel de tamanho $s = 1922$. É possível notar que o indicador $I_A(i)$ normaliza a medida da Área $A(i)$, tornando a evolução temporal sempre crescente, à medida que a janela móvel agrega dados nas proximidades do *crash*. Desta forma, verifica-se que o comportamento de $I_A(i)$ é mais consistente que $A(i)$, apresentado na Figura 4.6.

O indicador $I_A(i)$ é interessante, pois apresenta uma tendência crescente durante a evolução temporal no Algoritmo Online, até atingir seu valor máximo no dia onde ocorre o *crash*. A tendência de crescimento de $I_A(i)$ é a informação prévia de que um *crash* está próximo. O indicador possui ainda a vantagem de tender a voltar para o nível anterior ao *crash* mais rapidamente que $A(i)$, passado o período de turbulência provocado pelo *crash*. A desvantagem deste método é tomar a média sobre todos os valores de área $A(i)$ calculados previamente no Algoritmo Online. Isso pode trazer problemas computacionais para uma série de dados muito longa, além da desvantagem de valores muito distantes

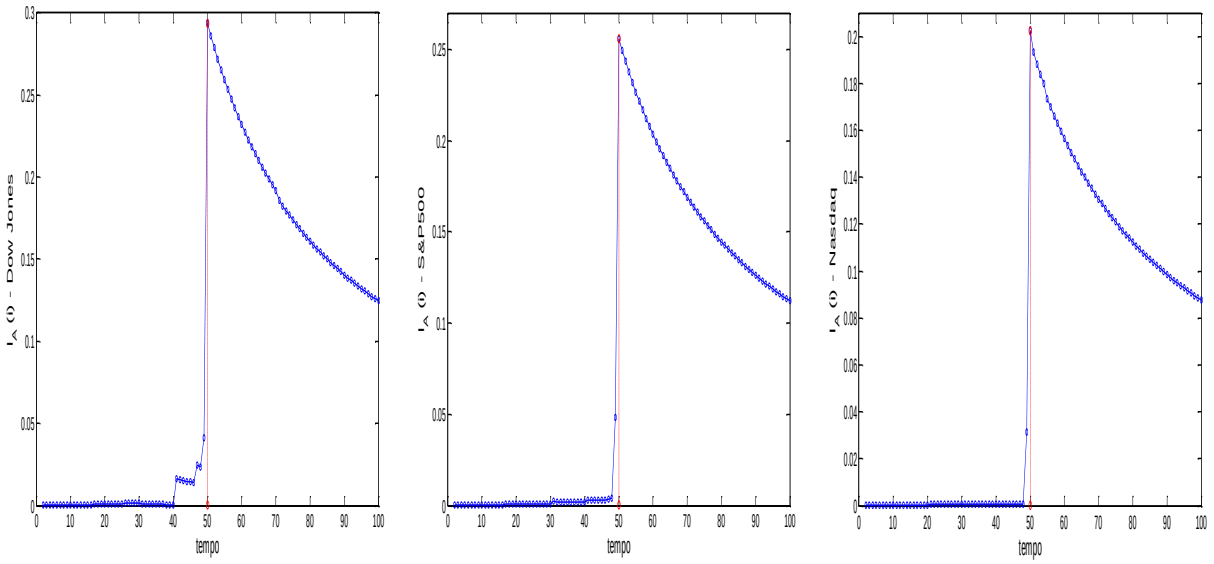

Figura 4.7: Evolução temporal do indicador $I_A(i)$: Índice Dow Jones, Índice S&P500 e Índice Nasdaq.

no passado influenciarem o indicador no presente. Pensando nisso, o segundo indicador proposto é uma pequena variação do primeiro. Neste, ao invés de tomar a média sobre todos os valores de $A(i)$ calculados previamente, calcula-se uma média móvel para os últimos n valores de $A(i)$. O segundo indicador é definido então na Equação 4.32:

$$I_A{}^n(i) = \left| \frac{A(i)}{\overline{A}(i-n:i)} - 1 \right| = \left| \frac{A(i) - \overline{A}(i-n:i)}{\overline{A}(i-n:i)} \right|. \tag{4.32}$$

A média móvel $\overline{A}(i-n:i)$ é calculada para $i \geq n$. Para $i < m$, calcula-se a média $\overline{A}(1:i)$, assim como no cálculo do indicador $I_A(i)$.

A Figura 4.8 apresenta a evolução temporal do índice $I_A{}^n(i)$ para os índices Dow Jones, S&P500 e Nasdaq, para o período de Jan/1980 a Dez/1987, com janela de tamanho $s = 1922$ e parâmetro de média móvel $n = 15$. Assim como o indicador $I_A(i)$, o indicador $I_A{}^n(i)$ normaliza a medida de Área $A(i)$, tornando a evolução temporal sempre crescente, à medida que a janela móvel agrega dados nas proximidades do *crash*. O indicador também atinge o valor máximo no ponto onde ocorre o *crash* e decai rapidamente para um nível próximo de zero. Este comportamento indica que o indicador $I_A{}^n(i)$ não possui tanta memória quanto os indicadores $A(i)$ e $I_A(i)$.

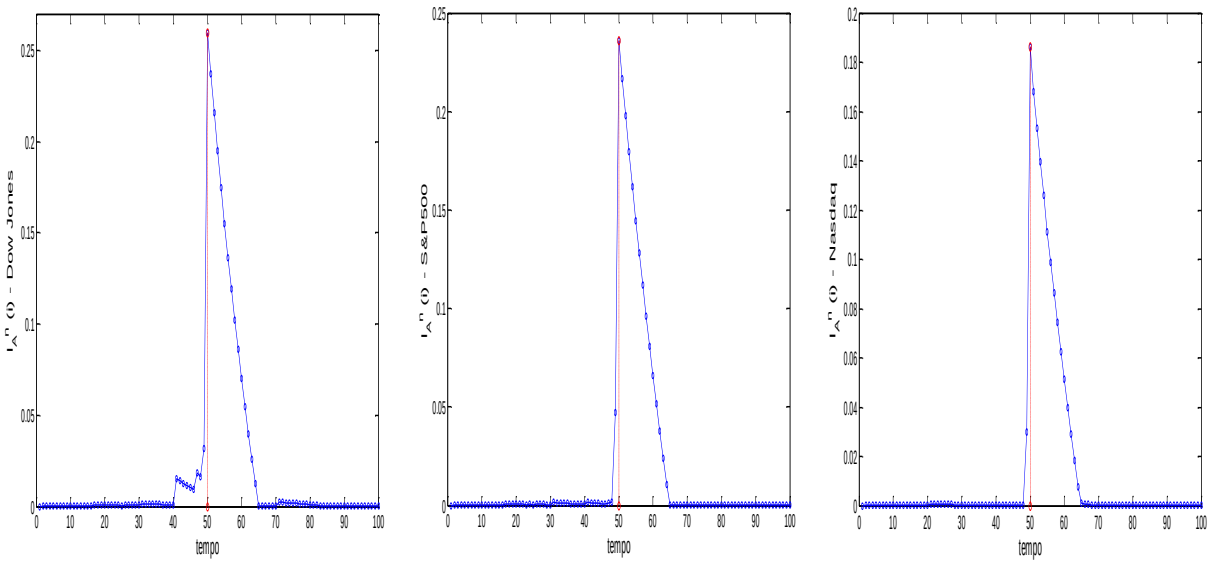

Figura 4.8: Evolução temporal do indicador $I_A{}^n(i)$: Dow Jones, S&P500 e Nasdaq.

4.8 Resultados

4.8.1 Análise R/S de Hurst Clássica

Na Tabela 4.1, são apresentados os resultados da Análise R/S de Hurst (Expoente H de Hurst, Dimensão Fractal D_f, Coeficiente Espectral β e o Coeficiente de Correlação γ) para os Índices Dow Jones, S&P500, Nasdaq e Ibovespa em vários períodos de tempo:

- Out/1928 a Dez/1929: Período que contempla o *crash* de 1929. Apenas o Índice Dow Jones possui dados com esta data [55].
- Jan/1980 a Dez/1992: Período que contempla o *crash* de 1987. Aqui foram testados os índices Dow Jones, S&P500 e Nasdaq.
- Jan/1993 a Dez/1993: Período que não contempla nenhum *crash* . Análise para os Índices Dow Jones, S&P500 e Nasdaq neste período tem caráter comparativo.
- Jan/1998 a Ago/2011: Período que contempla vários eventos de *crash* e crises no mercado, como a crise do Subprime, desencadeada em 2008. Aqui foi feita a análise para o Índice Ibovespa, além de análise para os Índices Dow Jones, S&P500 e Nasdaq,
- Dados Simulados: Foram simuladas três séries do tipo Movimento Browniano Fracionário, com coeficientes $H = 0.2$, $H = 0.5$ e $H = 0.8$ respectivamente, com o intuito de validar a técnica utilizada.

ATIVO	PERÍODO	H	D_f	β	γ
Dow Jones (Crash 1929)	Out/1928 a Dez/1929	0,5715	1,4285	0,1429	0,8571
Dow Jones (Crash 1987)	Jan/1980 a Dez/1992	0,5422	1,4578	0,0845	0,9155
SP500 (Crash 1987)	Jan/1980 a Dez/1992	0,5672	1,4328	0,1345	0,8655
Nasdaq (Crash 1987)	Jan/1980 a Dez/1992	0,6398	1,3602	0,2796	0,7204
Dow Jones (Sem Crise)	Jan/1993 a Dez/1993	0,5491	1,4509	0,0983	0,9017
SP500 (Sem Crise)	Jan/1993 a Dez/1993	0,4812	1,5188	-0,0377	1,0377*
Nasdaq (Sem Crise)	Jan/1993 a Dez/1993	0,5548	1,4452	0,1096	0,8904
Dow Jones (Crise 2010)	Jan/1998 a Ago/2011	0,5479	1,4521	0,0958	0,9042
SP500 (Crise 2010)	Jan/1998 a Ago/2011	0,5609	1,4391	0,1218	0,8782
Nasdaq (Crise 2010)	Jan/1998 a Ago/2011	0,6068	1,3932	0,2136	0,7864
Ibovespa (Crise 2010)	Jan/1998 a Ago/2011	0,5821	1,4179	0,1642	0,8358
MBF (H = 0.2)	3288 pontos simulados	0,2988	1,7012	-0,4025	1,4025*
MBF (H = 0.5)	3288 pontos simulados	0,5312	1,4688	0,0623	0,9377
MBF (H = 0.8)	3288 pontos simulados	0,7712	1,2288	0,5424	0,4576

Tabela 4.1: Expoente de Hurst H, Dimensão Fractal D_f , Coeficiente Espectral β e o Coeficiente de Correlação γ dos Índices Dow Jones, S&P500 e Nasdaq; além do Movimento Browniano Fracionário Simulado (MBFS), com $H = 0.2$, $H = 0.5$ e $H = 0.8$. Os valores de γ marcados com *(*)* não fazem sentido, uma vez que $0 < \gamma < 1$. Isto só ocorrerá para os casos em que $0 < H < 0.5$.

Por meio da Tabela 4.1, é possível notar que apenas Índice S&P500 no período de Jan/1993 a Dez/1993 (período sem nenhuma crise ou crash de mercado), e logicamente a série MBF com $H = 0.2$ apresentaram comportamento antipersistente, ou seja $0 < H < 0.5$. As demais séries apresentaram comportamento persistente, com efeito de memória longa em todas os intervalos de tempo ($0.5 < H < 1$). A análise indica a necessidade de utilizarmos técnicas de memória longa para as séries nos períodos considerados. No entanto, estes valores deverão ser validados na próxima sessão, devido aos conhecidos problemas de detecção espúria de correlações de longo alcance, na Análise R/S de Hurst [11] [12] [42] [43].

4.8.2 Detrended Fluctuation Analysis - DFA

Na Tabela 4.2, são apresentados os resultados para o Expoente de Escala α do DFA, Dimensão Fractal $D(q)$, Coeficiente Espectral β e o Coeficiente de Correlação γ) para os Índices Dow Jones, S&P500, Nasdaq e Ibovespa para os mesmos períodos de tempo utilizado na Análise R/S de Hurst Clássica.

ATIVO	PERÍODO	α	D_f	β	γ
Dow Jones (Crash 1929)	Out/1928 a Dez/1929	0,4646	1,5354	-0,0707	1,0707*
Dow Jones (Crash 1987)	Jan/1980 a Dez/1992	0,5337	1,4663	0,0674	0,9326
SP500 (Crash 1987)	Jan/1980 a Dez/1992	0,5327	1,4673	0,0654	0,9346
Nasdaq (Crash 1987)	Jan/1980 a Dez/1992	0,6370	1,3630	0,2740	0,7260
Dow Jones (Sem Crise)	Jan/1993 a Dez/1993	0,4912	1,5088	-0.0176	1,0176*
SP500 (Sem Crise)	Jan/1993 a Dez/1993	0,4135	1,5865	-0,1730	1,1730*
Nasdaq (Sem Crise)	Jan/1993 a Dez/1993	0,5447	1,4553	0,0895	0,9105
Dow Jones (Crise 2010)	Jan/1998 a Ago/2011	0,5022	1,4978	0,0045	0,9955
SP500 (Crise 2010)	Jan/1998 a Ago/2011	0,5104	1,4896	0,0208	0,9792
Nasdaq (Crise 2010)	Jan/1998 a Ago/2011	0,5659	1,4341	0,1319	0,8681
Ibovespa (Crise 2010)	Jan/1998 a Ago/2011	0,5622	1,4378	0,1243	0,8757
MBF (H = 0.2)	3288 pontos simulados	0,2158	1,7842	-0,5684	1,5684*
MBF (H = 0.5)	3288 pontos simulados	0,5071	1,4929	0,0142	0,9858
MBF (H = 0.8)	3288 pontos simulados	0,8284	1,1716	0,6568	0,3432

Tabela 4.2: Expoente de Escala α, Dimensão Fractal D_f , Coeficiente Espectral β e o Coeficiente de Correlação γ dos Índices Dow Jones, S&P500 e Nasdaq; além do Movimento Browniano Fracionário Simulado (MBFS), com $H = 0.2$, $H = 0.5$ e $H = 0.8$. Os valores de γ marcados com *(*)* não fazem sentido, uma vez que $0 < \gamma < 1$. Isto só ocorrerá para os casos em que $0 < \alpha < 0.5$.

A análise da Tabela 4.2 permite notar que o Índice Dow Jones (Out/1928 a Dez/1929, Jan/1993 a Dez/1993), o Índice S&P500 (Jan/1993 a Dez/1993) e a Série MBF ($H = 0.2$), apresentaram comportamento antipersistente, ou seja $0 < \alpha < 0.5$. As demais séries/períodos considerados apresentaram movimento persistente, ou seja, em período que contém eventos como crashes e/ou crises, espera-se que a série apresente persistência ou antipersistência bem definida, conforme pôde ser notado pelos valores de H e α nas Tabelas 4.1 e 4.2. É possível notar também que a análise R/S claramente superestima o valor do expoente H, favorecendo a detecção espúria de correlações de longo alcance [11] [12] [42] [43].

4.8.3 Análise Multifractal e Termodinâmica

Conforme visto anteriormente, o Calor Específico Análogo $C(q)$ é capaz de gerar com antecedência, informações sobre a iminência de um *crash* de mercado [14] [15]. Como a medida $C(q)$ está diretamente ligada ao conceito de multifractalidade, nada impede a visualização de conceitos como a Energia Livre Análoga $\tau_i(q,s)$, a Dimensão Fractal $D_i(q,s)$ e o Espectro Multifractal $f_i(\alpha)$ durante a execução do Algoritmo Online.

O Algoritmo Online por sua vez, com janela móvel v de tamanho s, foi preparado de maneira que o início de sua execução está a 50 iterações do *crash*. O algoritmo continua a execução por mais 50 iterações resultando num total de 100 iterações (dias). Na tentativa de apresentar de forma didática a evolução das medidas no Algoritmo Online, a cada iteração i, são plotadas as medidas $\tau_i(q,s)$, $C_i(q,s)$, $D_i(q,s)$ e $f_i(\alpha)$. Porém, para diferenciar as curvas antes, durante e depois do *crash*, foram tomadas cores diferentes para períodos diferentes, conforme descrito a seguir:

- Curvas plotadas em **Preto**: Representam as 39 primeiras iterações no Algoritmo Online. A idéia é mapear o comportamento de $\tau_i(q,s)$, $C_i(q,s)$, $D_i(q,s)$ e $f_i(\alpha)$ à medida que a janela v de tamanho s começa a agregar informações cada vez mais próximas do *crash*.

- Curvas plotadas em **Azul**: Representam as iterações de número 40 a 49. Na iteração número 49, a *janela* v está a um dia do *crash*.

- Curva plotada em **Vermelho**: Representa a iteração única de número 50. Nesta data ocorre o *crash*.

- Curvas plotadas em **Verde**: Representam as iterações de número 51 a 100, à medida que a janela v agrega informações posteriores ao *crash*.

Índice Dow Jones - *Crash* de 1929 (Out/1928 a Dez/1929)

A Figura 4.9 apresenta a evolução temporal das medidas multifractais $\tau(q)$, $C(q)$, $D(q)$ e $f(\alpha)$ versus α (espectro) para o índice Dow Jones, no *crash* de 1929. É possível notar que poucos dias antes do *crash*, um "ombro anômalo" [14] [15] começa a se formar à direita do gráfico de $C(q)$ (curvas em azul), ficando bastante evidente na ocorrência do crash (curva em vermelho). As demais medidas também apresentam um deslocamento, indicando a passagem para um período turbulento de mercado. Este deslocamento se assemelha a uma transição de fase [54].

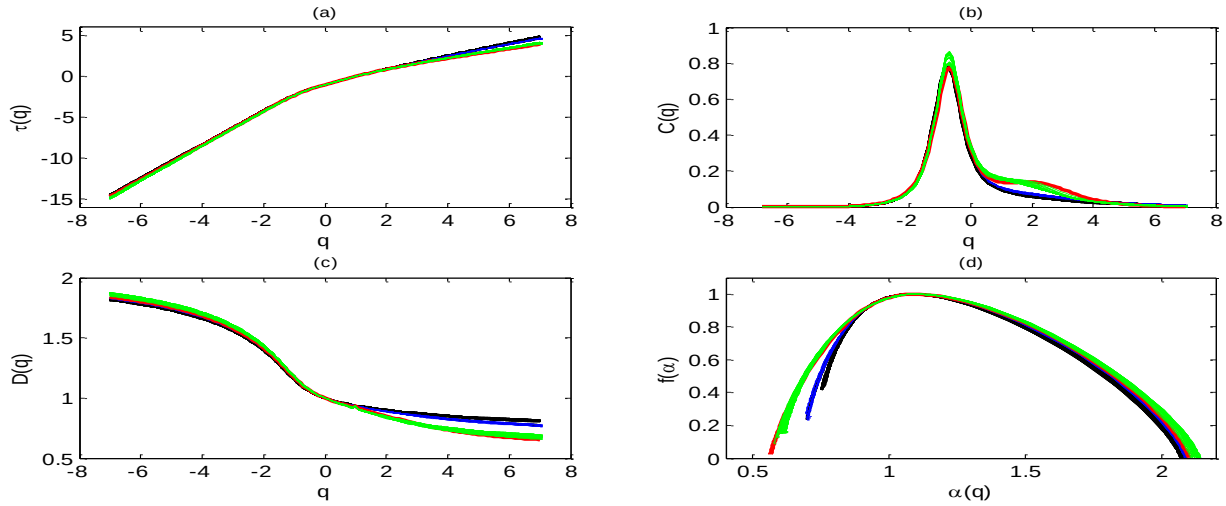

Figura 4.9: Dow Jones (1929): (a) Energia Livre Análoga $\tau(q)$, (b) Calor Específico análogo $C(q)$, (c) Dimensão Multifractal $D(q)$ e (d) Espectro Multifractal $f(\alpha)$ versus α.

Simultaneamente ao cálculo das medidas multifractais, a Figura 4.10 apresenta a evolução do indicador $I_A(i)$, que tende a crescer rapidamente e muda seu nível médio pouco depois da iteração 30. A mudança de nível do indicador, juntamente com início da mudança de regime das medidas multifractais, são os indícios prévios para a ocorrência do *crash*. Após atingir o pico máximo na iteração 50, o Índice $I_A(i)$ começa a decrescer vagarosamente, indicando que ainda existem efeitos do *crash* sobre os dados.

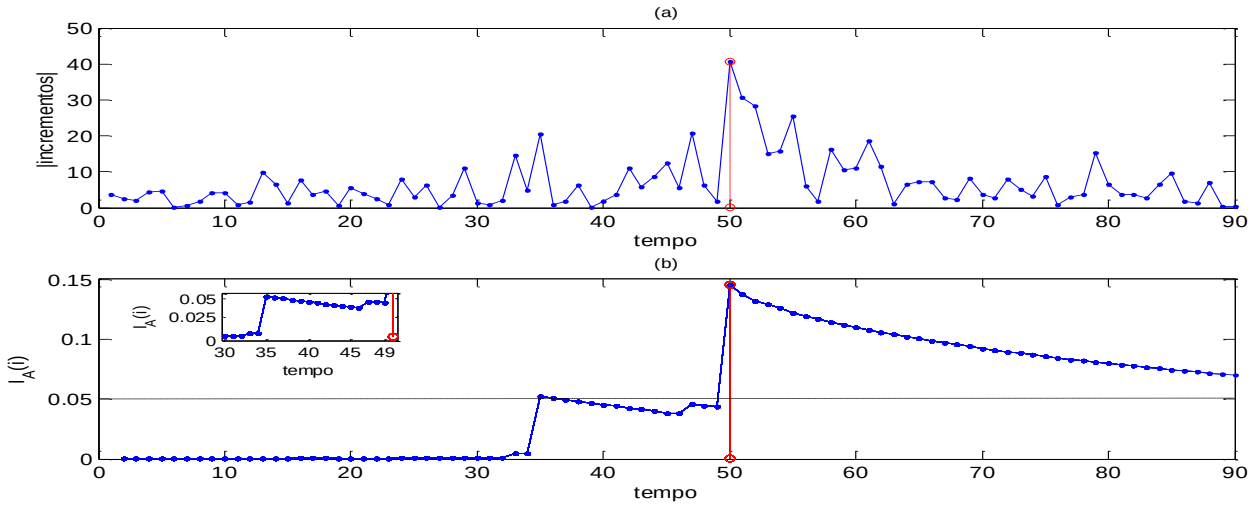

Figura 4.10: Dow Jones (1929): (a) Incrementos Absolutos $|\Delta x(t,1)|$, (b) Indicador de *Crash* $I_A(i)$

Índice Dow Jones - *Crash* de 1987 (Jan/1980 a Dez/1987)

A Figura 4.11 apresenta a evolução temporal das medidas multifractais $\tau(q)$, $C(q)$, $D(q)$ e $f(\alpha)$ versus α (espectro) para o índice Dow Jones, no *crash* de 1987. Assim como no *crash* de 1929, à poucos dias antes do *crash* de 1987, um "ombro anômalo" [14] [15] começa a se formar à direita do gráfico de $C(q)$ (curvas em azul), ficando bastante evidente na ocorrência do *crash* (curva em vermelho). As demais medidas também apresentam um deslocamento nas curvas indicando a passagem para um período turbulento de mercado. Este deslocamento também se assemelha à uma transição de fase [54].

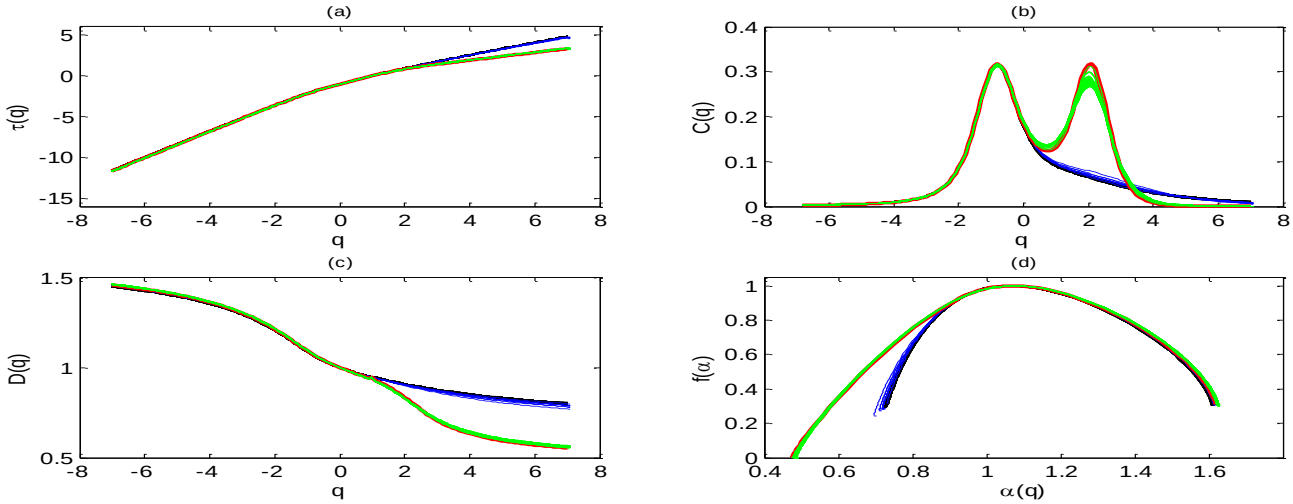

Figura 4.11: Dow Jones (1987): (a) energia livre análoga $\tau(q)$, (b) Calor Específico Análogo $C(q)$, (c) Dimensão Multifractal $D(q)$ e (d) Espectro Multifractal $f(\alpha)$ versus α.

Simultaneamente ao cálculo das medidas multifractais, a Figura 4.12 apresenta a evolução do indicador $I_A(i)$, que tende a crescer rapidamente e muda seu nível médio pouco depois da iteração 40. A mudança de nível do indicador, juntamente com início da mudança de regime das medidas multifractais, são os indícios prévios para a ocorrência do *crash*. Após atingir o pico máximo na iteração 50, o Índice $I_A(i)$ começa a decrescer vagarosamente, indicando que ainda existem efeitos do *crash* sobre os dados.

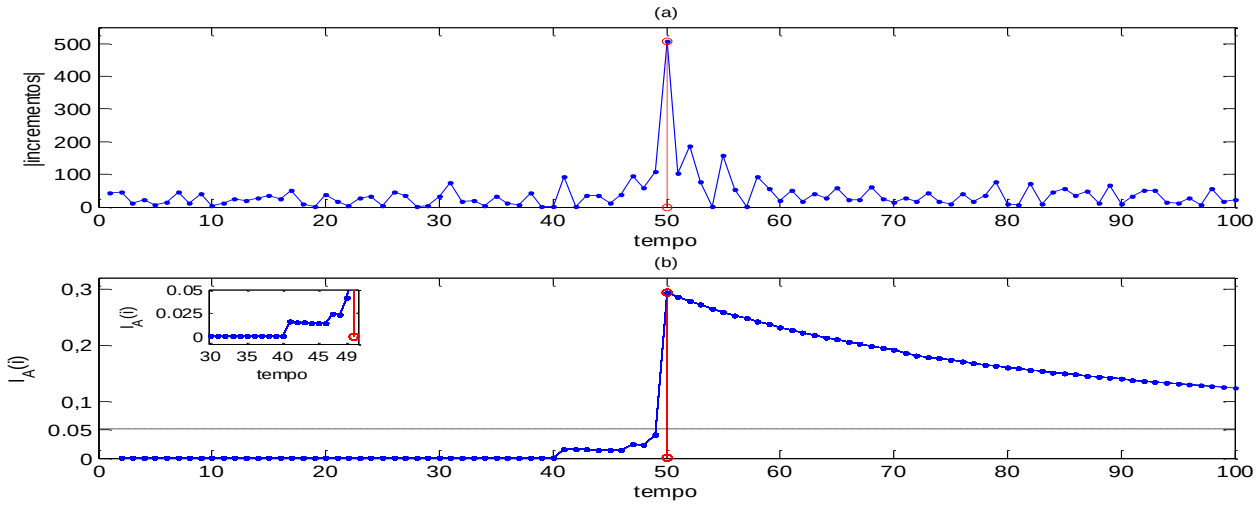

Figura 4.12: Dow Jones (1987): (a) Incrementos Absolutos $|\Delta x(t,1)|$, (b) Indicador de *Crash* $I_A(i)$

Índice S&P500 - *Crash* de 1987 (Jan/1980 a Dez/1987)

A Figura 4.13 apresenta a evolução temporal das medidas multifractais $\tau(q)$, $C(q)$, $D(q)$ e $f(\alpha)$ versus α (espectro) para o índice S&P500, no *crash* de 1987. É possível notar que à poucos dias antes do *crash*, um "ombro anômalo" [14] [15] começa a se formar à direita do gráfico de $C(q)$ (curvas em azul), ficando bastante evidente na ocorrência do *crash* (curva em vermelho). As demais medidas também apresentam um deslocamento, indicando a passagem para um período turbulento de mercado. Este deslocamento se assemelha a uma transição de fase [54].

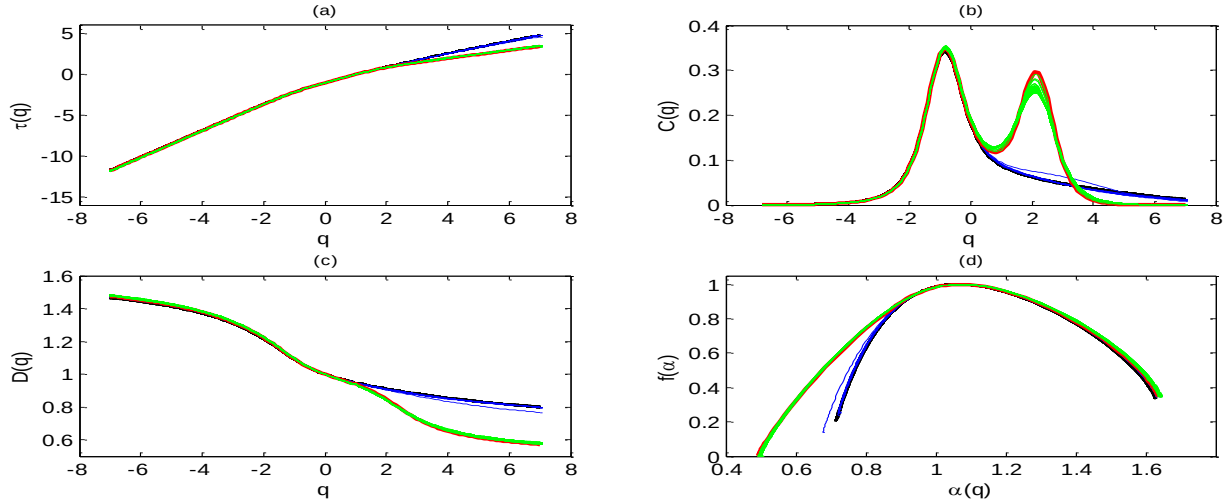

Figura 4.13: S&P500 (1987): (a) Energia Livre Análoga $\tau(q)$, (b) Calor Específico Análogo $C(q)$, (c) Dimensão Multifractal $D(q)$ e (d) Espectro Multifractal $f(\alpha)$ versus α.

Simultaneamente ao cálculo das medidas multifractais, a Figura 4.14 apresenta a evolução do indicador $I_A(i)$, dá um salto crítico na iteração 49, um dia antes do *crash*. A

mudança de nível do indicador, juntamente com início da mudança de regime das medidas multifractais, são os indícios prévios para a ocorrência do *crash*. Após atingir o pico máximo na iteração 50, o Índice $I_A(i)$ começa a decrescer vagarosamente, indicando que ainda existem efeitos do *crash* sobre os dados.

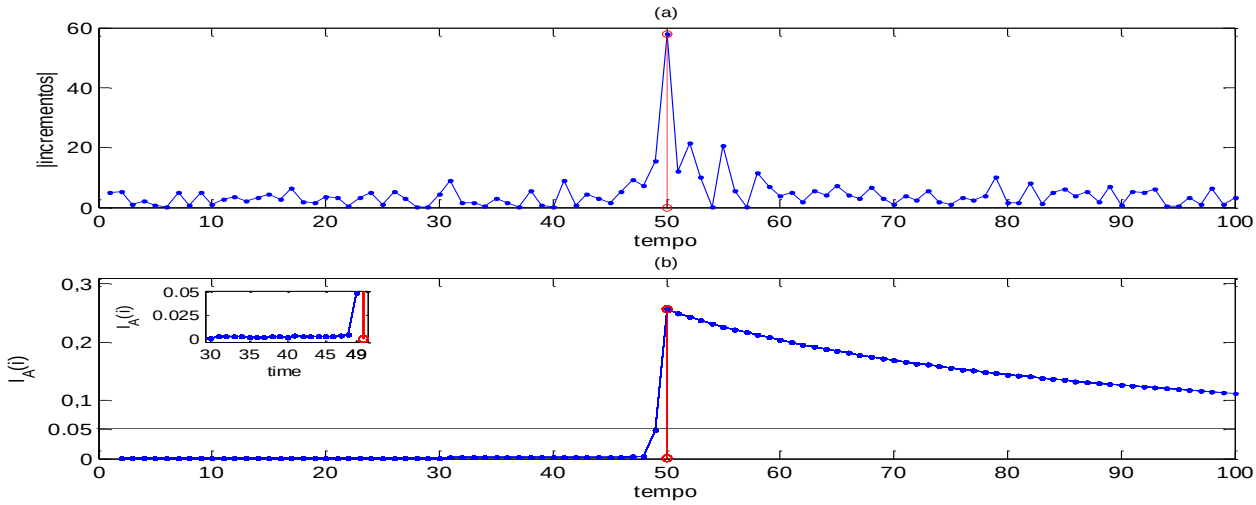

Figura 4.14: S&P500 (1987): (a) Incrementos Absolutos $|\Delta x(t,1)|$, (b) Indicador de *Crash* $I_A(i)$

Índice Nasdaq - *Crash* de 1987 (Jan/1980 a Dez/1987)

A Figura 4.15 apresenta a evolução temporal das medidas multifractais $\tau(q)$, $C(q)$, $D(q)$ e $f(\alpha)$ versus α (espectro) para o índice S&P500, no *crash* de 1987. É possível notar que à poucos dias antes do *crash*, um "ombro anômalo" [14] [15] começa a se formar à direita do gráfico de $C(q)$ (curvas em azul), ficando bastante evidente na ocorrência do *crash* (curva em vermelho). As demais medidas também apresentam um deslocamento, indicando a passagem para um período turbulento de mercado. Este deslocamento se assemelha a uma transição de fase [54].

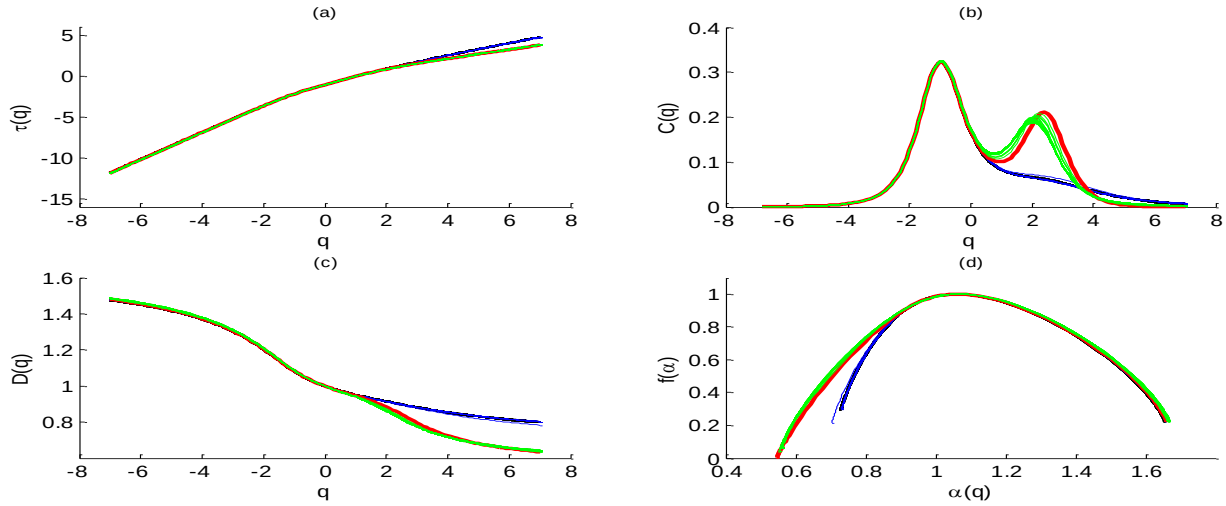

Figura 4.15: Nasdaq (1987): (a) Energia Livre Análoga $\tau(q)$, (b) Calor Específico Análogo $C(q)$, (c) Dimensão Multifractal $D(q)$ e (d) Espectro Multifractal $f(\alpha)$ versus α.

Simultaneamente ao cálculo das medidas multifractais, a Figura 4.16 apresenta a evolução do indicador $I_A(i)$, dá um salto crítico na iteração 49, um dia antes do *crash*. A mudança de nível do indicador, juntamente com início da mudança de regime das medidas multifractais, são os indícios prévios para a ocorrência do *crash*. Após atingir o pico máximo na iteração 50, o Índice $I_A(i)$ começa a decrescer vagarosamente, indicando que ainda existem efeitos do *crash* sobre os dados.

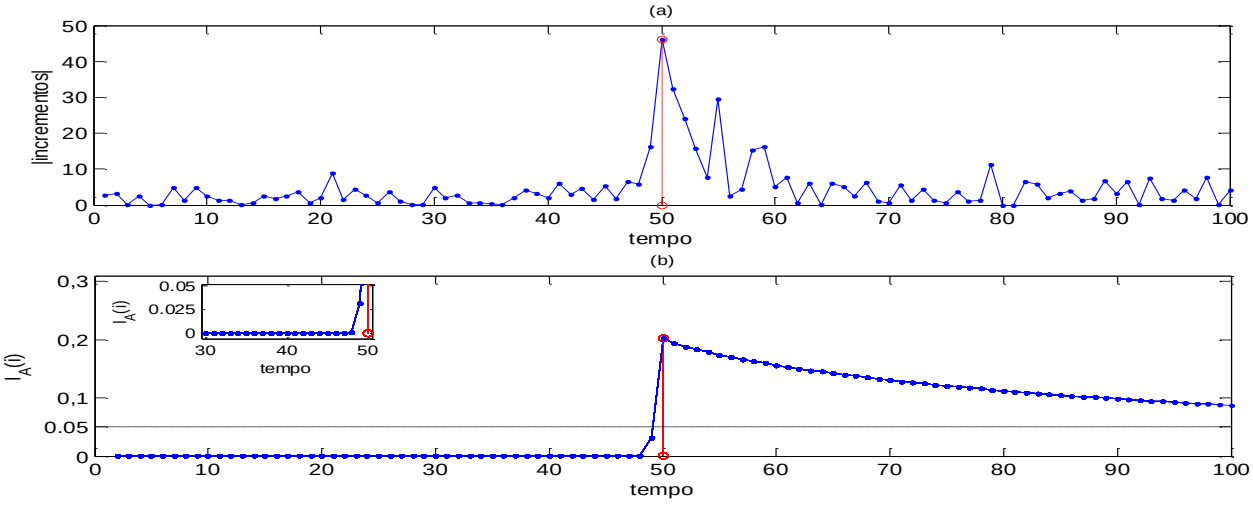

Figura 4.16: Nasdaq (1987): (a) Incrementos Absolutos $|\Delta x(t,1)|$, (b) Indicador de *Crash* $I_A(i)$

Ruído Branco Gaussiano

A Figura 4.17 apresenta a evolução temporal das medidas multifractais $\tau(q)$, $C(q)$, $D(q)$ e $f(\alpha)$ versus α (espectro) para 2022 valores simulados de um Ruído Branco Gaus-

siano. A série simulada possui o mesmo comprimento das séries Dow Jones, S&P500 e Nasdaq tomadas nas análises do *crash* de 1987. Conforme é possível notar, não há qualquer movimento acentuado nas medidas, assim como verificado para as análises anteriores. Também não há qualquer semelhança com transição de fase [54].

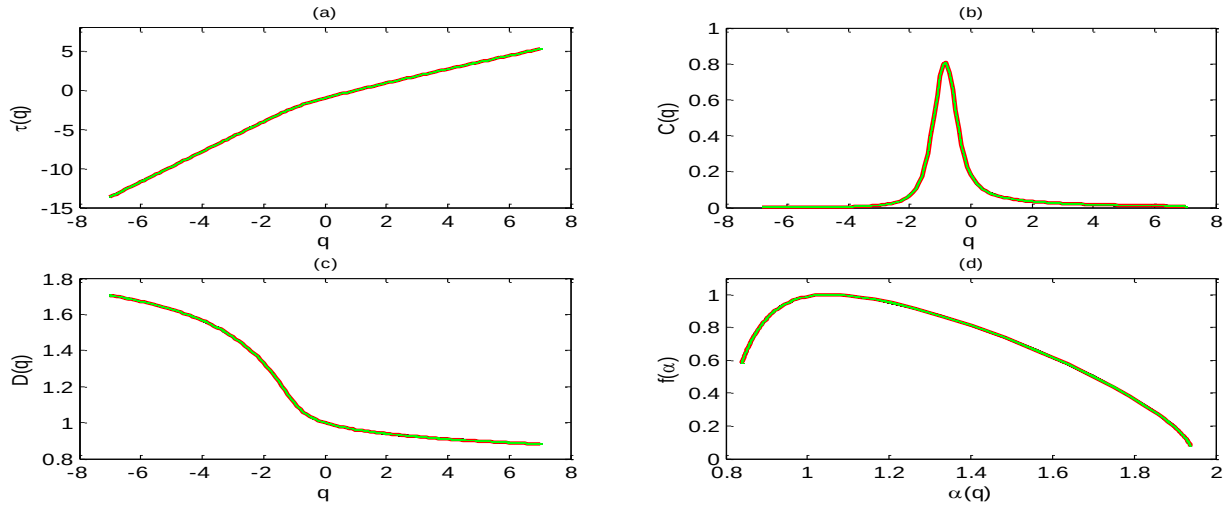

Figura 4.17: Ruído Branco Gaussiano: (a) Energia Livre Análoga $\tau(q)$, (b) Calor Específico Análogo $C(q)$, (c) Dimensão Multifractal $D(q)$ e (d) Espectro Multifractal $f(\alpha)$ versus α.

Simultaneamente ao cálculo das medidas multifractais, a Figura 4.18 apresenta a evolução do indicador $I_A(i)$, que permanece numa escala bem menor do que o verificado nas análises anteriores. Logo, não há qualquer indício da ocorrências de um *crash* nos dados simulados.

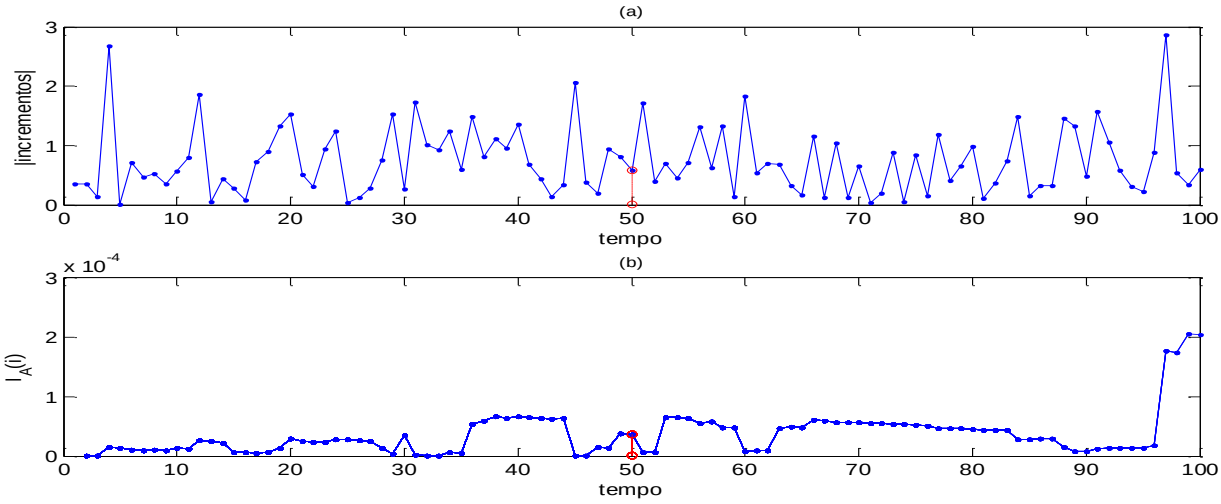

Figura 4.18: Ruído Branco Gaussiano: (a) Incrementos Absolutos $|\Delta x(t,1)|$, (b) Indicador de *Crash* $I_A(i)$

Este capítulo é interessante por trazer a caracterização fractal e multifractal para quatro índices financeiros. De acordo com os resultados obtidos na Análise R/S de Hurst, a maioria das séries, na maioria dos períodos analisados, apresentaram comportamento persistente, com efeito de memória longa em todas as escalas de tempo ($0.5 < H < 1$). O índice S&P500 no período de Jan/1993 a Dez/1993 e logicamente a série MBF com $H = 0.2$, apresentaram comportamento antipersistente, ou seja $0 < H < 0.5$. Estes resultados são interessantes e confirmam que a série de retornos não pode ser modelada por um ruído branco. Esta é uma justificativa adicional para a utilização de modelos ARMA para a média da série de retornos, no ajuste dos modelos Garch do Capítulo 3. Os modelos ARMA para a média foram detectados na ocasião por meio das funções de autocorrelação e autocorrelação parcial.

Para confirmar os resultados obtidos na Análise R/S de Hurst, foram aplicados às mesmas séries e períodos de tempo, a técnica do DFA. Os resultados demonstram que para o Índice Dow Jones (Out/1928 a Dez/1929, Jan/1993 a Dez/1993), o Índice S&P500 (Jan/1993 a Dez/1993) e a logicamente a série MBF com $H = 0.2$ apresentaram comportamento antipersistente, ou seja $0 < \alpha < 0.5$. Os resultados divergem dos apresentados na Análise R/S de Hurst. No entanto, mesmo divergindo entre comportamento persistente e antipersistente, os resultados mostram que os retornos não parecem ser ruídos brancos gaussianos, ou aleatórios. Este é mais um ponto a favor da utilização de modelos ARMA para a modelagem da média da série de retornos, na modelagem GARCH do Capítulo 3.

Por fim, basedos no conceito de Análise Multifractal, testamos a capacidade preditiva do indicador $I_A(i)$, construído por meio do cálculo da área sob $C(q)$, que evolui temporalmente no algoritmo online. Conforme visto, os resultados são interessantes e demonstram que o indicador ou outros indicares basedos nele podem trazer novos avanços para a compreensão de movimentos bruscos de mercado. Com ele, foi possível detectar, com pelo menos um dia de antecedência, a ocorrência de um *crash*. Os resultados são robustos, uma vez que as medidas multifractais também evoluíram temporalmente, se assemelhando a uma transição de fase [54]. A robustez do método foi verificada ao aplicar o algoritmo para uma série aleatória, que não apresentou o comportamento característico observado para séries temporais reais, nas proximidades de um *crash*.

Capítulo 5

Conclusão

5.1 Considerações Finais

No presente trabalho procurou-se fazer uma revisão dos modelos econométricos não lineares para séries temporais financeiras para mapear (estimar) os períodos mais voláteis do mercado e verificar se estes são capazes de gerar alguma informação relevante sobre movimentos bruscos de mercado. Aqui, os modelos não foram aplicados no contexto preditivo, pois o objetivo era utilizar a capacidade explicativa dos modelos para caracterizar períodos com grande volatilidade.

Para testar as propriedades de correlação de longo alcance dos mesmos índices, foram aplicadas as análises R/S de Hurst e o DFA. O objetivo foi caracterizar empiricamente as propriedades de séries temporais financeiras e futuramente, discutir sobre qual família de processos estocásticos seria candidata para modelar este sistema. Esta família de processos estocásticos deverá replicar os expoentes observados empiricamente.

Por último, buscou-se reproduzir os resultados obtidos por Enrique Canessa em [14] [15], com o objetivo de caracterizar o início de *crashes* de mercado, por meio do cálculo do Calor Específico Análogo $C(q)$ para os índices Dow Jones, S&P500 e Nasdaq. O interesse estará nos indicadores $I_A(i)$ e ${I_A}^n(i)$, construídos por meio da área $A(i)$ sob $C(q)$. Tais índices foram construídos utilizando o "Sliding Observation Box", chamado no trabalho de Algoritmo Online. O mérito destes indicadores está em traçar um denominador comum para o comportamento aparentemente difuso de $A(i)$ na véspera de *crashes*. Conforme pôde ser visto, os indicadores são interessantes por fornecerem com certa antecedência, indícios consistentes sobre a ocorrência de *crashes* de mercado.

5.2 Sugestões para pesquisas futuras

Futuras pesquisas podem dar continuidade ao presente trabalho. Seria interessante verificar o quanto os índices propostos são robustos a falsos positivos e/ou negativos. Conforme verificado, alguma função da área é um bom palpite para a construção de indicadores de *crash*. Neste contexto, outros indicadores poderiam ser construídos. Outro aspecto interessante é verificar se os indicadores propostos são capazes de antecipar o impacto do *crash*.

Os métodos utilizados até agora para estudar a estrutura multifractal das séries financeiras, como a Análise R/S de Hurst, o DFA e a Função de Partição no Formalismo Multifractal, são apenas aproximações [11]. A utilização de um método mais preciso como as Wavelets, poderia trazer resultados mais robustos que a metodologia empregada neste trabalho.

Por fim, seria interessante aplicar as técnicas utilizadas neste trabalho em outros contextos, como biológicos, sociais e saúde.

Apêndice A

Modelos ARIMA Simulados - Identificação

A metodologia utilizada para ajustar modelos Autorregressivos Integrados de Médias Móveis $ARIMA(p, d, q)$ a um conjunto de dados, é baseado na conhecida metodologia de Box e Jenkings. A estratégia para a construção desta classe de modelos é baseada em um ciclo iterativo, cujos estágios são [36]:

1. Uma classe geral de modelos é considerada para a análise.
2. Há a identificação de um modelo, com base na análise das funções de autocorrelação (ACF) e autocorrelação parcial (PACF) amostrais.
3. Estimam-se os parâmetros do modelo identificado.
4. Faz-se um diagnóstico do modelo ajustado por meio de uma análise de resíduos, para saber se este é adequado.
5. Caso o modelo não seja adequado, o ciclo é repetido, voltando-se à fase de identificação.

Os modelos Auto-Regressivos são baseados na idéia de que o valor atual da série $r(t)$, pode ser explicado como uma função dos p valores passados, $r(t-1)$, $r(t-2)$, ... , $r(t-p)$, onde p determina o número de passos no passado necessários para gerar uma previsão do valor atual [57]. Assim, podemos definir:

Definição A.1 *Um Modelo Auto-Regressivo de ordem p, abreviadamente AR(p), tem a seguinte forma:*

$$r(t) = \phi_1 \cdot r(t-1) + \phi_2 \cdot r(t-2) + \ldots + \phi_p \cdot r(t-p) + w(t), \tag{A.1}$$

onde $r(t)$ *é estacionário,* $\phi_1, \phi_2, \ldots, \phi_p$ *são constantes e* $\phi_p \neq 0$. *Embora não seja necessário, assume-se que* $w(t)$ *seja um Ruído Branco Gaussiano com média zero e variância* σ_w^2.

É possível ainda reescrever o processo definido na Equação A.2, utilizando o operador auto-regressivo $\phi(B) = (1 - \phi_1 \cdot B - \phi_2 \cdot B^2 - \ldots - \phi_p \cdot B^p)$ da seguinte maneira:

$$\phi(B) \cdot r(t) = w(t). \tag{A.2}$$

A Figura A.1 mostra um processo $AR(1)$ simulado, com $\phi = 0.8$ e $n = 100$ iterações. Ao lado, são inseridas as Funções de Autocorrelação (ACF) e Autocorelação Parcial do processo (PACF).

Uma característica da ACF de processos $AR(p)$ é que esta decai de acordo com exponenciais e/ou senóides amortecidas, infinitas em extensão. Já a PACF de processos $AR(p)$ apresentam um corte após o *lag* p. As características descritas podem ser confirmadas na Figura A.1.

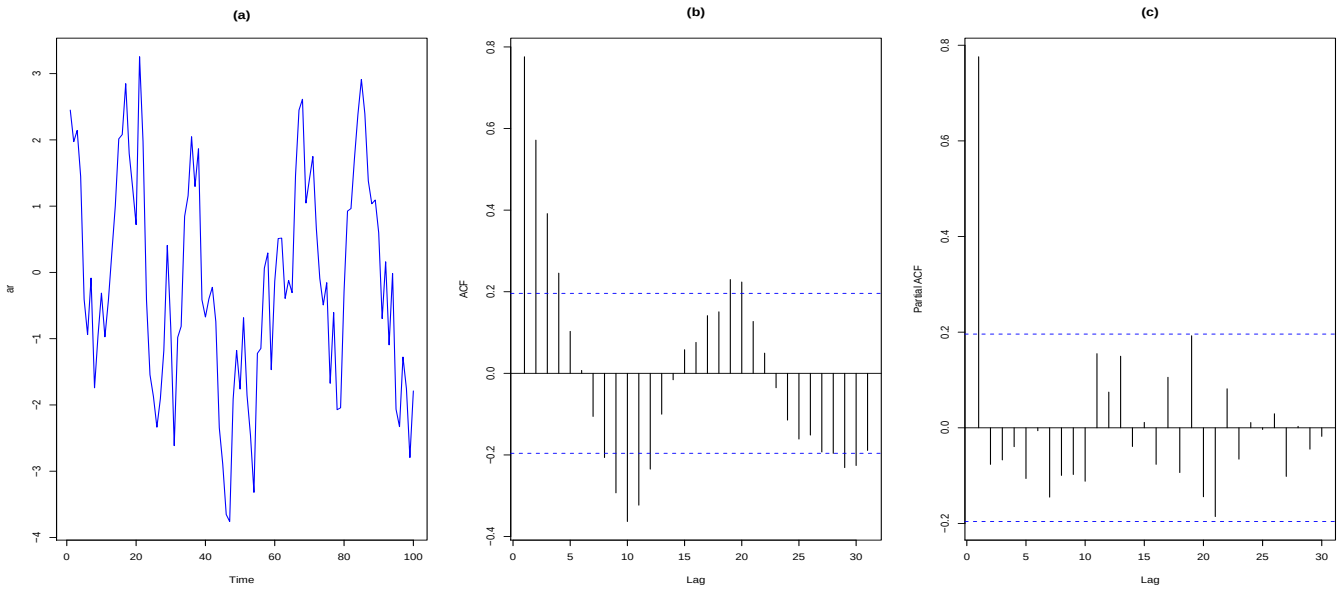

Figura A.1: (a) Processo $AR(1)$ simulado: $r(t) = 0.8 \cdot r(t-1) + w(t)$, com $\phi = 0.8$ e n =100. (b) Função de Autocorrelação (ACF) do processo simulado. (c) Função de Autocorrelação Parcial (PACF) do processo simulado.

Como alternativa à representação Auto-regressiva, o modelo de Médias Móveis de ordem q assume que os dados observados são combinações lineares do Ruído Branco $w(t)$. Logo, podemos definir:

Definição A.2 *Um Modelo de Médias Móveis de ordem q, abreviadamente $MA(q)$, tem a seguinte forma:*

$$r(t) = w(t) + \theta_1 \cdot w(t-1) + \theta_2 \cdot w(t-2) + \ldots + \theta_q \cdot w(t-q), \tag{A.3}$$

*onde existem q atrasos (*lags*) na média móvel e $\theta_1, \theta_2, \ldots, \theta_q$ ($\theta_q \neq 0$) são os parâmetros. Embora não seja necessário, assume-se que $w(t)$ seja um Ruído Branco Gaussiano com média zero e variância σ_w^2.*

É possível ainda reescrever o processo definido na Equação A.2, utilizando o operador de médias móveis $\theta(B) = (1 + \theta_1 \cdot B + \theta_2 \cdot B^2 + \ldots + \theta_p \cdot B^p)$ da seguinte maneira:

$$r(t) = \theta(B) \cdot w(t). \tag{A.4}$$

A Figura A.2 mostra um processo $MA(1)$ simulado, com $\theta = 0.9$ e $n = 100$ iterações. Ao lado, são inseridas as Funções de Autocorrelação (ACF) e Autocorelação Parciais do processo (PACF).

Uma característica da ACF de modelos $MA(q)$ é que esta apresenta um corte após o *lag* q. Já a PACF de um processo $MA(q)$ decai de acordo com exponenciais e/ou senóides amortecidas, infinitas em extensão. Estas características podem ser confirmadas na Figura A.2.

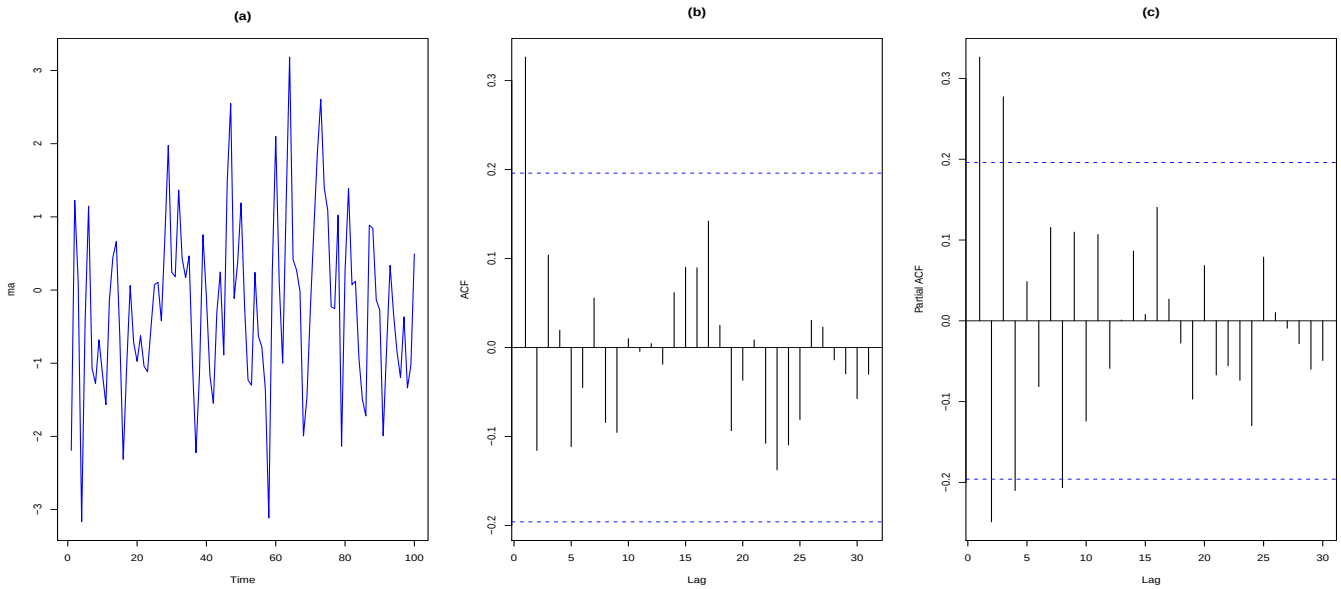

Figura A.2: (a) Processo $MA(1)$ simulado: $r(t) = w(t) + 0.9 \cdot w(t-1)$, com $\theta = 0.9$ e n =100. (b) Função de Autocorrelação (ACF) do processo simulado. (c) Função de Autocorrelação Parcial (PACF) do processo simulado.

Em muitas aplicações, a utilização de modelos $AR(p)$ ou $MA(q)$ puros para descrever adequadamente os dados torna-se complicada, principalmente quando é necessário utilizar estes modelos com ordem muito alta. Para contornar esta dificuldade, foram introduzidos os modelos $ARMA(p, q)$, que são modelo gerais mistos Auto-Regressivo e de Médias Móveis para séries temporais estacionárias:

Definição A.3 *Um Modelo Autorregressivo e de Médias Móveis de ordens p e q, abreviadamente $ARMA(p, q)$, tem a seguinte forma:*

$$r(t) = \phi_1 \cdot r(t-1) + \ldots + \phi_p \cdot r(t-p) + w(t) + \theta_1 \cdot w(t-1) + ... + \theta_q \cdot w(t-q), \quad \text{(A.5)}$$

com $\phi_p \neq 0$, $\theta_q \neq 0$ e $\sigma_w^2 > 0$. Os parâmetros p e q são chamados de ordem auto-regressiva e de médias móveis respectivamente.

É possível ainda reescrever o processo definido na Equação A.5, utilizando o operador auto-regressivo $\phi(B) = (1 - \phi_1 \cdot B - \phi_2 \cdot B^2 - \ldots - \phi_p \cdot B^p)$ e o operador de médias móveis $\theta(B) = (1 + \theta_1 \cdot B + \theta_2 \cdot B^2 + \ldots + \theta_p \cdot B^p)$ da seguinte maneira:

$$\phi(B) \cdot r(t) = \theta(B) \cdot w(t). \quad \text{(A.6)}$$

Podemos notar no modelo $ARMA(p, q)$ descrito nas Equações A.2 e A.3 pode ser reduzido a um processo $AR(p)$ bastando utilizar $q = 0$. Da mesma forma, para reduzir o modelo $ARMA(p, q)$ a um processo $MA(q)$ puro, basta utilizar $p = 0$.

A Figura A.3 mostra um processo $ARMA(1, 1)$ simulado, com $\theta = 0.8$, $\theta = 0.9$

e $n = 100$ iterações. Ao lado, são inseridas as Funções de Autocorrelação (ACF) e Autocorelação Parciais do processo.

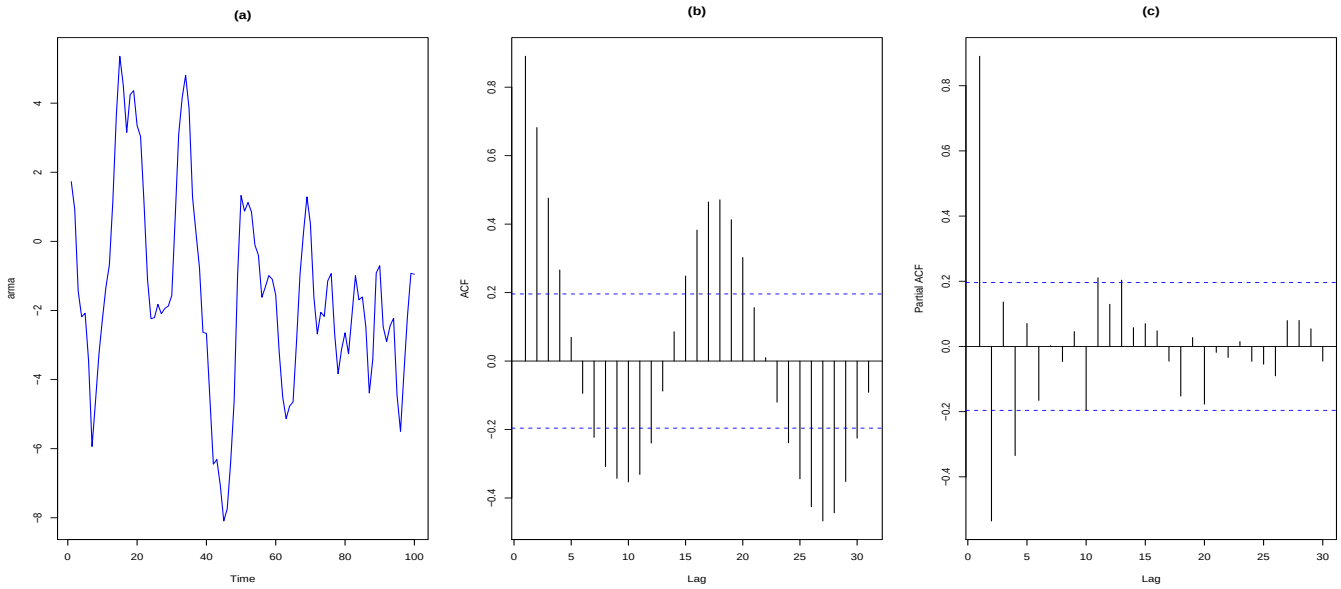

Figura A.3: (a) Processo $AR(1)$ simulado: $r(t) = 0.8 \cdot r(t-1) + w(t) + 0.9w(t-1)$, com $\phi = 0.8$,$\theta = 0.9$ e $n = 100$. (b) Função de Autocorrelação (ACF) do processo simulado. (c) Função de Autocorrelação Parcial (PACF) do processo simulado.

A ACF de modelos $ARMA(p,q)$ é infinita em extensão , decaindo de acordo com exponenciais e/ou senóides amortecidas após o *lag* $q-p$. Já a $PACF$ de um processo $ARMA(p,q)$ decai de acordo com exponenciais e/ou senóides amortecidas, infinitas em extensão. Estas características podem ser confirmadas na Figura A.3.

Os modelos definidos até agora são apropriados para séries temporais estacionárias. Porém, muitas séries reais não são estacionárias, exibindo média e variância não constantes ao longo do tempo. Várias séries econômicas e financeiras não são estacionárias, mas quando diferenciadas tornam-se estacionárias [36]. Assim, podemos definir os modelos $ARIMA(p,d,q)$ cujo objetivo é modelar séries não estacionárias:

Definição A.4 *Um Modelo é dito Autorregressivo, Integrado e de Médias Móveis de ordens p, d e q, abreviadamente $ARIMA(p,d,q)$, se: $\nabla^d r(t) = (1-B)^d \cdot r(t)$ é um processo $ARMA(p,q)$. Em geral, para $\phi(B) = (1 - \phi_1 \cdot B - \phi_2 \cdot B^2 - \ldots - \phi_p \cdot B^p)$ e $\theta(B) = (1 + \theta_1 \cdot B + \theta_2 \cdot B^2 + \ldots + \theta_p \cdot B^p)$, o modelo é escrito da seguinte maneira:*

$$\phi(B) \cdot (1-B)^d \cdot r(t) = \theta(B) \cdot w(t). \tag{A.7}$$

Apêndice B

Tabelas com os Ajustes dos Modelos ARMA-GARCH

B.1 Índice Dow Jones: Jan/1980 a Ago/2011

	Coefficient	Std. Error	z-Statistic	Prob.
C	0.000578	9.69E-05	5.963533	0.0000
AR(12)	0.029238	0.011943	2.448005	0.0144
AR(15)	-0.024031	0.011510	-2.087750	0.0368
Variance Equation				
C	1.66E-06	1.43E-07	11.62431	0.0000
ARCH(1)	0.081527	0.001861	43.81766	0.0000
GARCH(1)	0.906747	0.003264	277.8300	0.0000

R-squared	0.001431	Mean dependent var	0.000324
Adjusted R-squared	0.000804	S.D. dependent var	0.011326
S.E. of regression	0.011321	Akaike info criterion	-6.473273
Sum squared resid	1.021355	Schwarz criterion	-6.468018
Log likelihood	25818.18	F-statistic	2.283941
Durbin-Watson stat	2.051360	Prob(F-statistic)	0.043768

Inverted AR Roots	.74+.13i	.74-.13i	.66+.45i	.66-.45i
	.40+.70i	.40-.70i	.06-.79i	.06+.79i
	-.25-.71i	-.25+.71i	-.49+.58i	-.49-.58i
	-.71-.35i	-.71+.35i	-.81	

B.2 Índice S&P500: Jan/1980 a Ago/2011

	Coefficient	Std. Error	z-Statistic	Prob.
C	0.000533	9.55E-05	5.578255	0.0000
AR(4)	-0.029948	0.011550	-2.592822	0.0095
AR(12)	0.030333	0.011864	2.556632	0.0106
Variance Equation				
C	1.33E-06	1.22E-07	10.85199	0.0000
ARCH(1)	0.078617	0.001890	41.59432	0.0000
GARCH(1)	0.912470	0.002972	307.0084	0.0000

R-squared	0.001001	Mean dependent var	0.000300
Adjusted R-squared	0.000375	S.D. dependent var	0.011525
S.E. of regression	0.011523	Akaike info criterion	-6.473977
Sum squared resid	1.058499	Schwarz criterion	-6.468724
Log likelihood	25830.69	F-statistic	1.597860
Durbin-Watson stat	2.056408	Prob(F-statistic)	0.156965

Inverted AR Roots	.74	.65-.38i	.65+.38i	.38-.65i
	.38+.65i	-.00-.74i	-.00+.74i	-.38+.65i
	-.38-.65i	-.65-.38i	-.65+.38i	-.74

B.3 Índice Nasdaq: Jan/1980 a Ago/2011

	Coefficient	Std. Error	z-Statistic	Prob.
C	0.000693	0.000115	6.034277	0.0000
AR(1)	0.142204	0.012335	11.52840	0.0000
AR(12)	0.032521	0.011071	2.937559	0.0033
Variance Equation				
C	1.62E-06	1.31E-07	12.41062	0.0000
ARCH(1)	0.116604	0.004076	28.60594	0.0000
GARCH(1)	0.876989	0.004278	205.0040	0.0000
R-squared	-0.009004	Mean dependent var		0.000352
Adjusted R-squared	-0.009636	S.D. dependent var		0.013765
S.E. of regression	0.013831	Akaike info criterion		-6.333779
Sum squared resid	1.525090	Schwarz criterion		-6.328526
Log likelihood	25271.44	Durbin-Watson stat		2.211967
Inverted AR Roots	.76	.66-.38i	.66+.38i	.39-.65i
	.39+.65i	.01+.75i	.01-.75i	-.36+.65i
	-.36-.65i	-.64+.38i	-.64-.38i	-.74

B.4 Índice Ibovespa: Jan/1998 a Ago/2011

	Coefficient	Std. Error	z-Statistic	Prob.
C	0.000994	0.000315	3.155000	0.0016
AR(10)	0.052269	0.016758	3.119054	0.0018
Variance Equation				
C	1.10E-05	1.92E-06	5.726872	0.0000
ARCH(1)	0.073446	0.012312	5.965610	0.0000
ARCH(2)	0.038512	0.017572	2.191677	0.0284
GARCH(1)	0.863567	0.012840	67.25350	0.0000
R-squared	0.004458	Mean dependent var		0.000531
Adjusted R-squared	0.002977	S.D. dependent var		0.021958
S.E. of regression	0.021925	Akaike info criterion		-5.120823
Sum squared resid	1.616191	Schwarz criterion		-5.109917
Log likelihood	8629.466	F-statistic		3.010911
Durbin-Watson stat	1.969643	Prob(F-statistic)		0.010245
Inverted AR Roots	.74	.60+.44i	.60-.44i	.23-.71i
	.23+.71i	-.23-.71i	-.23+.71i	-.60-.44i
	-.60+.44i	-.74		

Referências Bibliográficas

[1] Anderson, P. W. More is Different. Science New Series, v. 177 (4047), p. 393-396, 1972.

[2] Boccara, N. Modeling Complex Systems. Springer-Verlag New York, 2008.

[3] Mitchell, M. Complexity - A Guided Tour. Oxford University Press, 2009.

[4] Bak, P. How Nature Works. Springer-Verlag New York, 1996.

[5] Oxley, L.; George, D. A. R. Economics on the edge of chaos: Some pitfalls of linearizing complex systems. Environmental Modelling & Software, v. 22, p. 580-589, 2007.

[6] Tsviliuk, O.; Zhang, D.; Melnik, R. Complex systems in finance: Monte Carlo evaluation of first passage time density functions. Procedia Computer Science I, p. 2381-2389, 2010.

[7] Mantegna, R. N.; Stanley. H. S. An Introduction to Econophysics: Correlations and Complexity in Finance, Cambridge University Press, Cambridge, 2000.

[8] Sornete, D. Why Stock Markets Crash? Critical Events in Complex Financial Systems. Princeton University Press, 2003.

[9] Preis, T.; Stanley, H. E. Bubble Trouble: Can a Law Describe Bubbles and Crashes in Financial Markets? Physics World v. 24(5), p. 29-32, 2011.

[10] Mantegna, R. N.; Stanley, H. E. Scaling Behaviour in the Dynamic of a Economic Index. Letters to Nature, v. 376, p. 46-49, 1995.

[11] Kantelhardt, J. W. Fractal and Multifractal Time Series. arXiv:0804.0747v1 [physics.data-an], 2008.

[12] Kantelhardt, J.W.; Zschiegner, S.A.; Koscielny-Bunde, E.; Bunde,A.; Havlin, S.; Stanley, E. Multifractal Detrended Fluctuation Analysis of Nonstationary Time Series. Physica A. v. 316, p. 87, 2002.

[13] Yu, C. X.; Gilmore, M.; Peebles, W. A.; Rhodes, T. L. Structure function analysis of long-range correlations in plasma turbulence. Phys. Plasmas. v. 10(7), p. 2772-2779, 2003.

[14] Canessa, E. Multifractality in time Series. Physica A. v. 33. p. 3637-3651, 2000.

[15] Canessa, E. Langevin dynamics of financial systems : A second-order analysis. Physical Journal B. October, 127, 123-127.

[16] Bouchaud, J. P.; POTTERS, M. Theory of financial risks, Cambridge University Press, Cambridge, UK, 2000.

[17] Bachelier L. Théorie de la spéculation (Paris: Gauthier-Villars 1900). English translation by A. James Boness, in The Random Character of Stock Market Prices, pp. 17-78, Paul H. Cootner (editor), The MIT Press, 1964.

[18] Mandelbrot B. B., The Variation of Certain Speculative Prices, Journal of Business, v. 35, p. 394-419, 1963.

[19] Fama, E.F. Mandelbrot and the Stable Paretian Hypothesis, J. Business, v. 36, p. 420-429, 1963.

[20] Gopikrishnan, P.; Meyer M.; Amaral, L. A. N.; Stanley, H. E. Inverse Cubic Law for the Distribution of Stock Price Variations, The European Physical Journal B, v. 3, p. 139-140, 1998.

[21] Gopikrishnan, P.; Plerou, V., Amaral L. A. N.; Meyer, M.; Stanley, H. E. Scaling of the Distribution of Fluctuations of Financial Market Indices, Physical Review E, v. 60, p. 5305-5316, 1999.

[22] Plerou, V.; Gopikrishnan, P.; Amaral, L. A. N.; Meyer, M.; Stanley, H. G. Scaling of the Distribution of Price Fluctuations of Individual Companies, Physical Review E, v.60, p. 6519-6529, 1999.

[23] Wilmott, P. Paul Wilmott on Quantitative Finance, John Wiley & Sons LTD, Second Edition, 2006.

[24] Engle, R.F. Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation, Econométrica, Vol. 50, N. 4, 1982.

[25] Bollerslev, T. Generalized Autoregressive Conditional Heteroskedasticity, Journal of Econometrics, v. 31, 1986.

[26] Challet, D.; Zhang, Y. C. Emergence of cooperation and organization in an evolutionary game. Physica A. v. 246, p. 407-418, 1997.

[27] Lamper, S.; Howison, S.; Johnson, N. F. Predictability of Large Future Changes in a Competitive Evolving Population. Phys. Rev. Lett. 88, 137902, 2002.

[28] Johnson, N. F.; Jefferies, P.; Hui, P. M. Financial Market Complexity, Oxford University Press, 2003.

[29] Marsili, M.; Challet, D.; Zhang, Y. C. Minority Game and Beyond: theory and applications of an interacting agent, Oxford University Press, 2004.

[30] Ferreira, F. F.; Marsili, M. Real payoffs and virtual trading in agent based market models. Physica A. v. 345, p. 657-675, 2005.

[31] Ferreira, F. F.; Oliveira, V. C.; Crepaldi, A. F.; Campos, P. R. A. Agent-based model with heterogeneous fundamental prices. Physica A 357, p. 534-542, 2005.

[32] Bianconi, G.; De Martino, A.; Ferreira, F. F.; Marsili, M. Multi-asset Minority Games. Quantitative Finance, v8, p. 255, 2008.

[33] Crepaldi, A. F. ; Rodrigues, C.N. ; Ferreira, F. F.; Francisco, G. . Multifractal Regime Transition in a Modified Minority Game Model. Chaos, Solitons and Fractals, v. 42, p.1364-1371, 2009.

[34] Norouzzadeh, P. Jafari, G. R. Application of Multifractal Measures to Teheran Price. Quantitative Finance Papers, 2004.

[35] Cont, R. Empirical Properties of Asset Returns: Stylized Facts and Statistical Issues. Quantitative Finance. v. 1, p. 223-236, 2001.

[36] Morettin, P.A; Toloi, C.M.C. Análise de Séries Temporais, Editora Elsevier, 2ª Edição. 2006.

[37] Podobnik, B.; Matia, K.; Chessa, A.; Ivanov, P. Ch.; Lee, Y; Stanley, H. E. Time Evolution of Stochastic Processes with Correlations in the Variance: Stability in Power-Law Tails of Distributions, Physica A, v. 300, p. 300-309, 2001.

[38] Podobnik, B.; Grosse, I.; Stanley, H. E. Stochastic process with power-law stability and a crossover in power-law correlations, Physica A, v. 316, p. 153-159, 2002.

[39] Hurst, H. E. Trans. Am. Soc. Civ. Eng, v. 116, p. 770, 1951.

[40] Hurst, H. E.; Black, R. P.; Simaika, Y. M. Long-Term Storage: An Experimental Study, Constable, London, 1965.

[41] Feder, J. Fractals. New York: Plenum Publishing Corporation, 1988.

[42] Peng, C. K.; Buldyrev, S. V.; Havlin, S.; Simons, M.; Stanley, H. E. Golderger, A. L. Mosaic organization of DNA nucleotides. Phys. Rev. E , v.49, p.1685-1689, 1994.

[43] Buldyrev, S. V.; Goldberger, A. L.; Havlin, S.; Mantegna, R. N.; Matsa, M. E.; Peng,C-K.; Simons, M.; Stanley, H. E. Long-Range Correlation Properties of Coding and Noncoding DNA sequences: GenBank Analysis, Physical Review E, v. 51, p.5084-5088, 1995.

[44] Grech, D.; Mazur, Z. Can one make any crash prediction in finance using the local Hurst exponent idea? Physica A, v. 336, p. 133-145, 2004.

[45] Mandelbrot, B. B. A multifractal walk down Wall Street. Scientific American, v. 298 (2), p. 70-73, 1999.

[46] Arneodo, A.; Bacry, E.; Muzy, J. The thermodynamics of fractals revisited with wavelets. Physica A . p. 213, 1995.

[47] Bianchi, S. A cautionary note on the detection of multifractal scaling in finance and economics. Appl. Econ. Lett. v. 12, p. 775-780, 2005.

[48] Borland, L. Statistical Signatures in Times of Panic: Markets as a Self-Organizing System. arXiv:0908.0111v2. [q-fin.ST].

[49] Borland, L.; Bouchaud, J. P.; Muzy, J. F.; Zumbach, G. The dynamics of financial markets - Mandelbrot's multifractal cascades and beyond, p. 1-24, 2008. arXiv:cond-mat/0501292v1.

[50] Los, C. A.; Yalamova, R. M. Multi-fractal spectral analysis of the 1987 stock market crash, SSRN. Int. Res. J. Finance. Econ.. Issue 4, 2006.

[51] Muzy, J.; Bacry, E.; Arneodo, A. Wavelets and multifractal formalism for singular signals: application to turbulent data. Phys. Rev. Lett . v.67(25), p. 3515-3518, 1991.

[52] Bouchaud, J. P.; Potters, M; Meyer, M. Apparent multifractality in financial time series, Eur. Phys. J. B. v. 12, p. 595, 2000.

[53] Jianga, Z-Q.; Zhoua, W-X. Multifractality in stock indexes: Fact or Fiction? Physica A, v. 387, p. 3605-3614, 2008.

[54] Lee, Jysoo; Stanley, H.E. Phase Transition in the Multifractal Spectrum of Diffusion-Limited Agregation. Physical Review Letters. v.61, 1998.

[55] Yahoo Finance. Disponível em: http://br.finance.yahoo.com. Acesso em Junho de 2011.

[56] Tsay, R.S. Analysis of Financial Time Series. A JOHN WILEY & SONS LTD, Second Edition, 2005.

[57] Shumway, R.H.; Stoffer, D.S. Time Series Analysis and its Applications with R exemples. Springer Texts in Statistics, Third Edition, 2011.

[58] Barbosa, R. M. Descobrindo a Geometria Fractal para a Sala de Aula. Editora Autêntica, 2002.

[59] Xua, Z.; Gençay, R. Scaling, Self-Similarity and Multifractality in FX Markets. Physica A , v. 323, p. 578 - 590, 2003.

[60] Physionet: Disponível em: http://www.physionet.org/tutorials/multifractal/behavior.htm, Acesso em Janeiro de 2012.