

JADER DA SILVA JALE

*Commodities* agrícolas do agronegócio brasileiro: análise multifractal e  
análise da complexidade diante da crise financeira mundial *subprime*  
2008/2009

Recife-PE – Junho/2015



UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO  
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO  
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOMETRIA E  
ESTATÍSTICA APLICADA

*Commodities* agrícolas do agronegócio brasileiro: análise multifractal e  
análise da complexidade diante da crise financeira mundial *subprime*  
2008/2009

Tese apresentada ao Programa de Pós-Graduação em Biometria e Estatística Aplicada como exigência parcial à obtenção do título de Doutor em Biometria e Estatística Aplicada.

Área de Concentração: Biometria e Estatística Aplicada

Orientador: Prof. Dr. Borko Stosic

Co-Orientadora: Profa. Dra. Tatijana Stosic

Recife-PE – Junho/2015

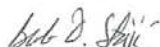
UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO  
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO  
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOMETRIA E ESTATÍSTICA APLICADA

*Commodities* agrícolas do agronegócio brasileiro: análise multifractal e análise da complexidade diante da crise financeira mundial *subprime* 2008/2009

JADER DA SILVA JALE

Tese defendida e aprovada por unanimidade pela banca examinadora em 1 de junho de 2015, Recife-Pernambuco.

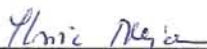
Orientador:



---

Prof. Dr. Borko Stosic  
Universidade Federal Rural de Pernambuco

Banca Examinadora:



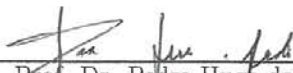
---

Profa. Dra. Tatijana Stosic  
Universidade Federal Rural de Pernambuco



---

Prof. Dr. Moacyr Cunha Filho  
Universidade Federal Rural de Pernambuco



---

Prof. Dr. Pedro Hugo de Figueirêdo  
Universidade Federal Rural de Pernambuco



---

Prof. Dr. Luiz Medeiros de Araújo Lima Filho  
Universidade Federal da Paraíba

*À minha família.*

# Agradecimentos

Agradeço a Deus, autor da vida e *criador* de todas as coisas, pois todas as coisas foram feitas por Ele, e sem Ele nada do que foi feito se fez.

Agradeço aos meus pais e meu irmão, inestimáveis e tão presentes, mesmo a distância. Muito obrigado por tudo!

Obrigado à Fabiana por todo apoio nos bons e maus momentos.

Aos amigos de curso de graduação, em especial aos amigos Darlon, Tiago, Gláuber, Bené, Leandro, Manoel, Kairo, Paulo Airton e Alex. Muito obrigado Darlon por toda ajuda nos momentos mais difíceis. Muito obrigado Bené por sua ajuda. Muito obrigado Manoel pela ajuda ao chegarmos em Recife. Agradeço a todos os professores e funcionários do DEMA-UFC, em especial aos professores Maurício, Aílton, Juvêncio, Vicente, Lassance, André, Ana Maria e Sílvia e às secretárias Marjorie e Luíza. Aos amigos de curso de mestrado pelo convívio, em especial a Cícero, Rosilda, Samuel, Rodrigo, David, Patrícia, Mariese, Rita, Rogério, Aranildo, Dâmocles, Rivelino, Nelson e também aos amigos do Doutorado, em especial a Joseilme e Mácio, muito obrigado pela ajuda nos momentos difíceis. Agradeço aos professores e funcionários do PPGBEA, em especial aos professores Borko, Tatijana, Eufrazio, Moacyr, Gabriel, Cláudio, Tiago e ao secretário Marco.

Ao meu orientador Professor Dr. Borko Stosic, pelos ensinamentos transmitidos e apoio neste trabalho e a Professora Dra. Tatijana Stosic, pelos ensinamentos passados e apoio. Muito obrigado por tudo! Sou eternamente grato!

# Resumo

O crescimento da economia mundial, impulsionado por países emergentes, principalmente a China, gerou mudanças relevantes no mercado de *commodities* a partir de 2002. Observou-se uma mudança nos preços das *commodities*, que mostraram uma elevação expressiva, mostrando condições acirradas entre oferta e demanda desses produtos, impulsionadas pela existência de problemas climáticos que afetaram negativamente a oferta e pelo ritmo de crescimento da demanda. A crise financeira mundial iniciou-se no mercado americano e acabou se tornando a pior crise financeira mundial desde 1929 (quebra da bolsa de Nova York). A falência do banco de investimento Lehman Brothers no dia 15 de setembro de 2008 marca a transformação da crise financeira internacional, e após isso, ocorre uma grande redução do crédito internacional e o dólar dispara no Brasil. Considerando que o setor agrícola é de fundamental importância para a sanidade econômica e por ser um grande investidor em tecnologias ambiental e rural, o Brasil não pode sucumbir à ideia de uma desaceleração neste setor, pois o agronegócio brasileiro representou, em 2008, 36.7% das exportações brasileiras, geração de 37% dos empregos e 28% do Produto Interno Bruto (PIB). Neste trabalho investigou-se a assincronia, a transferência de informação e o comportamento das correlações cruzadas dos retornos de seis *commodities* agrícolas do agronegócio brasileiro, para os períodos anterior (2006-2009) e posterior a crise financeira mundial (2010-2014). Utilizou-se o método *Cross-Sample Entropy* para quantificar a assincronia entre todas as séries de retornos das *commodities*. Utilizou-se os métodos *Multifractal Detrended Cross-Correlation Analysis* (MF-DCCA), *Multifractal Detrended Fluctuation Analysis* (MF-DFA) e *Detrended Cross-Correlation Analysis* (DCCA) para investigar correlações cruzadas e auto correlações. Os resultados da análise multifractal mostram que para todas as séries temporais, a multifractalidade diminuiu após a crise financeira mundial, indicando menor variedade do tamanho das flutuações que apresentam invariância de escala, exceto o algodão, que apresentou comportamento contrário. Com base nos resultados obtidos, pode-se concluir que a análise multifractal e a análise de complexidade podem ser úteis nos estudos da dinâmica do agronegócio brasileiro, dada a sua importância, diante do cenário econômico mundial seja para adoção de políticas monetárias e fiscal dos órgãos responsáveis, agentes econômicos ou pelo governo federal.

**Palavras-chave:** Crise financeira mundial, *Commodities*, Análise Multifractal e Análise de Complexidade.

# Abstract

The growth of the world economy, driven by emerging countries, especially China, has generated significant changes in the commodities market since 2002. The commodity prices have shown a significant increase, reflecting the fierce conditions of supply and demand for these products, driven by the climatic phenomena that have negatively affected the supply, and by the demand growth rate. The global financial crisis began in the US market, and eventually turned out the worst global financial crisis since 1929 (the break of the New York Stock Exchange). The bankruptcy of Lehman Brothers investment bank on September 15, 2008 marks the transformation of the international financial crisis, after which in Brazil there was a great reduction of international credit, accompanied by a sharp increase of the dollar exchange rate. Considering that the agricultural sector is of fundamental importance to the economic health, being a major investor in environmental and rural technologies, Brazil can not succumb to the idea of a slowdown in this sector, as in 2008 the Brazilian agribusiness represented 36.7% of exports, generating 37% of jobs, and 28% of gross domestic product (GDP). This work investigates the returns asynchrony and the behavior of the cross-correlations for six agricultural Brazilian agribusiness commodities, for the period prior to the global financial crisis (2006-2009), and after the crisis (2010-2014). The Cross-Sample Entropy method was used for quantifying the asynchrony among the commodity returns series. In addition, the methods Multifractal Detrended Cross-Correlation Analysis (MF-DCCA), Multifractal Detrended Fluctuation Analysis (MF-DFA) and Detrended Cross-Correlation Analysis (DCCA) were used to investigate cross-correlations and auto correlations in the returns series. The results of multifractal analysis show that for all time series, the multifractality decreased after the global financial crisis, indicating smaller range of the scale invariant fluctuations, except for Cotton, which exhibits precisely the opposite behavior. Based on the obtained results, it can be concluded that the multifractal analysis and the complexity analysis can be useful in the studies of the dynamics of the Brazilian agribusiness, given its importance within the global economic scenario, for adoption of monetary and fiscal policies by the responsible economic agents, or by the federal government.

**keywords:** Global financial crisis, commodities, Multifractal Analysis and Complexity Analysis.





# Sumário

<b>1</b>	<b>Introdução</b>	<b>1</b>
<b>2</b>	<b>Revisão de Literatura</b>	<b>5</b>
2.1	As <i>commodities</i> e a crise financeira . . . . .	5
2.1.1	<i>Commodities</i> . . . . .	5
2.1.2	A Crise <i>Subprime</i> . . . . .	6
2.1.3	Consequências da crise no Brasil . . . . .	7
2.2	Fractais e Multifractais . . . . .	8
2.2.1	Fractais . . . . .	8
2.2.2	Multifractais . . . . .	15
2.2.3	Análise Multifractal . . . . .	15
2.2.4	Processos Fractais . . . . .	16
2.3	Análise de Correlações de Longo Alcance em Séries Temporais Fractais	17
2.3.1	Análise da Função de Autocorrelação . . . . .	17
2.3.2	Teste de Correlação-Cruzada . . . . .	17
2.3.3	Análise Espectral . . . . .	18
2.3.4	Análise de Hurst . . . . .	18
2.3.5	<i>Fluctuation Analysis</i> (FA) . . . . .	20
2.3.6	<i>Detrended Fluctuation Analysis</i> (DFA) . . . . .	21
2.3.7	<i>Detrended Cross-Correlation Analysis</i> (DCCA) . . . . .	23
2.3.8	Coefficiente de Correlação-Cruzada ( $\rho_{DCCA}(n)$ ) . . . . .	24
2.4	Séries Temporais Multifractais . . . . .	25
2.4.1	<i>Multifractal Detrended Fluctuation Analysis</i> (MF-DFA) . . . . .	25
2.5	<i>Multifractal Detrended Cross-Correlation Analysis</i> (MF-DCCA) . . . . .	28
2.6	Análise de Complexidade das Séries Temporais . . . . .	29

2.6.1	<i>Sample Entropy (SampEn)</i> . . . . .	29
2.6.2	<i>Cross-Sample Entropy (Cross – SampEn)</i> . . . . .	31
<b>3</b>	<b>Materiais e Método</b>	<b>33</b>
3.1	Conjunto de Dados . . . . .	33
3.2	Descrição dos dados . . . . .	34
3.3	Metodologia . . . . .	38
3.3.1	Análise Fractal . . . . .	38
3.3.2	Análise Multifractal . . . . .	38
3.3.3	Análise da Complexidade . . . . .	38
3.3.4	Testes Estatísticos . . . . .	39
3.3.5	Coefficiente de Assimetria . . . . .	43
3.3.6	Coefficiente de Curtose . . . . .	43
<b>4</b>	<b>Resultados</b>	<b>45</b>
4.1	Preço de fechamento das <i>commodities</i> . . . . .	45
4.2	Retornos das <i>commodities</i> . . . . .	48
4.3	Teste de Correlação-Cruzada . . . . .	51
4.4	<i>Cross-Sample Entropy</i> entre as <i>commodities</i> . . . . .	54
4.5	DCCA . . . . .	58
4.6	Análise Multifractal . . . . .	67
4.7	Análise em Janelas Móveis de Tempo . . . . .	74
4.7.1	<i>SampEn</i> em Janela Móvel de Tempo (2006-2014) . . . . .	74
4.7.2	Expoente de Hurst em Janela Móvel de Tempo no período 2006-2014 . . . . .	76
4.7.3	Expoente DFA em Janela Móvel de Tempo no período 2006-2014 . . . . .	83
<b>5</b>	<b>Conclusões</b>	<b>89</b>
	<b>Anexo</b>	<b>103</b>

# Lista de Figuras

2.1	Fractal determinístico. Triângulo de Sierpinski. . . . .	9
2.2	Exemplo de fractal estocástico: brócolis romanesco. . . . .	9
2.3	Segmento de reta, quadrado e cubo divididos em partes menores, semelhantes a inicial. Fonte: [1]. . . . .	10
2.4	Conjunto de Cantor. Fonte: [2]. . . . .	11
2.5	Curvas de Koch. Fonte: [2]. . . . .	12
2.6	Esponja de Menger. Fonte: [1] . . . . .	12
2.7	Exemplo do uso do método <i>box-counting</i> . . . . .	13
2.8	Exemplo de série temporal estacionária. . . . .	21
2.9	Exemplo de série temporal não-estacionária. . . . .	21
4.1	Séries temporais dos preços de fechamento diário das <i>commodities</i> : Açúcar 4.1(a), Algodão 4.1(b), Arroz 4.1(c), Boi 4.1(d), Café 4.1(e) e Trigo 4.1(f), respectivamente. . . . .	46
4.2	Séries temporais do retorno diário das <i>commodities</i> : Açúcar 4.2(a), Algodão 4.2(b), Arroz 4.2(c), Boi 4.2(d), Café 4.2(e) e Trigo 4.2(f), respectivamente. . . . .	48
4.3	Distribuições dos retornos das <i>commodities</i> Açúcar 4.3(a), Algodão 4.3(b), Arroz 4.3(c), Boi 4.3(d), Café 4.3(e) e Trigo 4.3(f), respectivamente. . . . .	49
4.4	Gráfico log-log da estatística do teste $Q_{cc}(p)$ versus os graus de liberdade $p$ para o período 2006-2014 ( $N = 1960$ dias), em que $p = 1, \dots, 1000$ . . . . .	52
4.5	Gráfico log-log da estatística do teste $Q_{cc}(p)$ versus os graus de liberdade $p$ para o período pré-crise 2006-2009 ( $N = 915$ dias), onde $p = 1, \dots, 900$ . . . . .	53
4.6	Gráfico log-log da estatística do teste $Q_{cc}(p)$ versus os graus de liberdade $p$ para o período pós-crise 2010-2014 ( $N = 1045$ dias), em que $p = 1, \dots, 1000$ . . . . .	53

4.7	<i>Cross-Sample Entropy</i> entre cada <i>commodity</i> e as demais, para Açúcar 4.7(a), Algodão 4.7(b), Arroz 4.7(c), Boi 4.7(d), Café 4.7(e) e Trigo 4.7(f), respectivamente. . . . .	55
4.8	Coefficiente $\rho_{DCCA}(n)$ entre Açúcar e as demais <i>Commodities</i> : Algodão 4.8(a), Arroz 4.8(b), Boi 4.8(c), Café 4.8(d) e Trigo 4.8(e), respectivamente. . . . .	60
4.9	Coefficiente $\rho_{DCCA}(n)$ entre Algodão e as demais <i>Commodities</i> : Açúcar 4.10(a), Arroz 4.9(b), Boi 4.9(c), Café 4.9(d) e Trigo 4.9(e), respectivamente. . . . .	61
4.10	Coefficiente $\rho_{DCCA}(n)$ entre Arroz e as demais <i>Commodities</i> : Açúcar 4.10(a), Algodão 4.10(b), Boi 4.10(c), Café 4.10(d) e Trigo 4.10(e), respectivamente. . . . .	62
4.11	Coefficiente $\rho_{DCCA}(n)$ entre Boi e as demais <i>Commodities</i> : Algodão 4.11(b), Açúcar 4.11(a), Arroz 4.11(c), Café 4.11(d) e Trigo 4.11(e), respectivamente. . . . .	63
4.12	Coefficiente $\rho_{DCCA}(n)$ entre Café e as demais <i>Commodities</i> : Algodão 4.12(b), Açúcar 4.12(a), Arroz 4.12(c), Boi 4.12(d) e Trigo 4.12(e), respectivamente. . . . .	64
4.13	Coefficiente $\rho_{DCCA}(n)$ entre Trigo e as demais <i>Commodities</i> : Algodão 4.13(b), Açúcar 4.13(a), Arroz 4.13(c), Boi 4.13(d) e Café 4.13(e), respectivamente. . . . .	65
4.14	Expoente de Rényi $\tau(q)$ das <i>commodities</i> analisadas: Açúcar 4.14(a), Algodão 4.14(b), Arroz 4.14(c), Boi 4.14(d), Café 4.14(e) e Trigo 4.14(f), respectivamente. . . . .	67
4.15	Espectro multifractal $f(\alpha)$ dos retornos das <i>commodities</i> Açúcar, Algodão e Arroz, para as séries original e randomizada, nos períodos 2006-2009 e 2010-2014. . . . .	72
4.16	Espectro multifractal $f(\alpha)$ dos retornos das <i>commodities</i> Boi, Café e Trigo, para as séries original e randomizada, nos períodos 2006-2009 e 2010-2014. . . . .	73
4.17	Estatística <i>SampEn</i> em Janela Móvel de Tempo (tamanho 240 dias), com passo de um dia entre cada janela, totalizando 1721 janelas de tamanho 240 para cada <i>commodity</i> no período 2006-2014. . . . .	74
4.18	Expoente de Hurst em Janela Móvel de Tempo (tamanho 240 dias), com passo de um dia entre cada janela, totalizando 1721 janelas de tamanho 240 para cada <i>commodity</i> no período 2006-2014. . . . .	76

4.19	Expoente de Hurst em Janela Móvel de Tempo (tamanho 480 dias), com passo de um dia entre cada janela, totalizando 1481 janelas de tamanho 480 para cada <i>commodity</i> no período 2006-2014. . . . .	78
4.20	Expoente de Hurst em Janela Móvel de Tempo (tamanho 1000 dias), com passo de um dia entre cada janela, totalizando 961 janelas de tamanho 1000 para cada <i>commodity</i> no período 2006-2014. . . . .	79
4.21	Expoente DFA em Janela Móvel de Tempo (tamanho 240 dias), com passo de um dia entre cada janela, totalizando 1721 janelas de tamanho 240 para cada <i>commodity</i> no período 2006-2014. . . . .	83
4.22	Expoente DFA em Janela Móvel de Tempo (tamanho 480 dias), com passo de um dia entre cada janela, totalizando 1481 janelas de tamanho 480 para cada <i>commodity</i> no período 2006-2014. . . . .	84
4.23	Expoente DFA em Janela Móvel de Tempo (tamanho 1000 dias), com passo de um dia entre cada janela, totalizando 961 janelas de tamanho 1000 para cada <i>commodity</i> no período 2006-2014. . . . .	85
1	DCCA entre Açúcar e as demais <i>commodities</i> : Algodão 1(a), Arroz 1(b), Boi 1(c), Café 1(d) e Trigo 1(e), respectivamente. . . . .	103
2	DCCA entre Algodão e as demais <i>commodities</i> : Açúcar 2(a), Arroz 2(b), Boi 2(c), Café 2(d) e Trigo 2(e), respectivamente. . . . .	104
3	DCCA entre Arroz e as demais <i>commodities</i> : Açúcar 3(a), Algodão 3(b), Boi 3(c), Café 3(d) e Trigo 3(e), respectivamente. . . . .	105
4	DCCA entre Boi e as demais <i>commodities</i> : Açúcar 4(a), Algodão 4(b), Arroz 4(c), Café 4(d) e Trigo 4(e), respectivamente. . . . .	106
5	DCCA entre Café e as demais <i>commodities</i> : Açúcar 5(a), Algodão 5(b), Arroz 5(c), Boi 5(d) e Trigo 5(e), respectivamente. . . . .	107
6	DCCA entre Trigo e as demais <i>commodities</i> : Açúcar 6(a), Algodão 6(b), Arroz 6(c), Boi 6(d) e Café 6(e), respectivamente. . . . .	108
7	Expoente de Hurst Generalizado (EHG) entre Açúcar e as demais <i>Commodities</i> (2006-2009): Algodão 7(a), Arroz 7(b), Boi 7(c), Café 7(d) e Trigo 7(e), respectivamente. . . . .	109
8	Expoente de Hurst Generalizado (EHG) entre Algodão e as demais <i>Commodities</i> (2006-2009): Açúcar 8(a), Arroz 8(b), Boi 8(c), Café 8(d) e Trigo 8(e), respectivamente. . . . .	110

9	Expoente de Hurst Generalizado (EHG) entre Arroz e as demais <i>Commodities</i> (2006-2009): Açúcar 9(a), Algodão 9(b), Boi 9(c), Café 9(d) e Trigo 9(e), respectivamente. . . . .	111
10	Expoente de Hurst Generalizado (EHG) entre Boi e as demais <i>Commodities</i> (2006-2009): Açúcar 10(a), Algodão 10(b), Boi 10(c), Café 10(d) e Trigo 10(e), respectivamente. . . . .	112
11	Expoente de Hurst Generalizado (EHG) entre Café e as demais <i>Commodities</i> (2006-2009): Açúcar 11(a), Algodão 11(b), Arroz 11(c), Boi 11(d) e Trigo 11(e), respectivamente. . . . .	113
12	Expoente de Hurst Generalizado (EHG) entre Trigo e as demais <i>Commodities</i> (2006-2009): Açúcar 12(a), Algodão 12(b), Arroz 12(c), Boi12(d) e Café 12(e), respectivamente. . . . .	114
13	Expoente de Hurst Generalizado (EHG) entre Açúcar e as demais <i>Commodities</i> (2010-2014): Algodão 13(a), Arroz 13(b), Boi 13(c), Café 13(d) e Trigo 13(e), respectivamente. . . . .	115
14	Expoente de Hurst Generalizado (EHG) entre Algodão e as demais <i>Commodities</i> (2010-2014): Açúcar 14(a), Arroz 14(b), Boi 14(c), Café 14(d) e Trigo 14(e), respectivamente. . . . .	116
15	Expoente de Hurst Generalizado (EHG) entre Arroz e as demais <i>Commodities</i> (2010-2014): Açúcar 15(a), Algodão 15(b), Boi 15(c), Café 15(d) e Trigo 15(e), respectivamente. . . . .	117
16	Expoente de Hurst Generalizado (EHG) entre Boi e as demais <i>Commodities</i> (2010-2014): Açúcar 16(a), Algodão 16(b), Boi 16(c), Café 16(d) e Trigo 16(e), respectivamente. . . . .	118
17	Expoente de Hurst Generalizado (EHG) entre Café e as demais <i>Commodities</i> (2010-2014): Açúcar 17(a), Algodão 17(b), Arroz 17(c), Boi 17(d) e Trigo 17(e), respectivamente. . . . .	119
18	Expoente de Hurst Generalizado (EHG) entre Trigo e as demais <i>Commodities</i> (2010-2014): Açúcar 18(a), Algodão 18(b), Arroz 18(c), Boi18(d) e Café 18(e), respectivamente. . . . .	120
19	Espectro Multifractal $f(\alpha)$ entre Açúcar e as demais <i>commodities</i> : Arroz 19(b), Algodão 19(a), Boi 19(c), Café 19(d) e Trigo 19(e), respectivamente (2006-2009). . . . .	121
20	Espectro Multifractal $f(\alpha)$ entre Algodão e as demais <i>commodities</i> : Arroz 20(b), Açúcar20(a), Boi 20(c), Café 20(d) e Trigo 20(e), respectivamente (2006-2009). . . . .	122

---

21	Espectro Multifractal $f(\alpha)$ entre Arroz e as demais <i>commodities</i> : Açúcar 21(a), Algodão21(b), Boi 21(c), Café 21(d) e Trigo 21(e), respectivamente (2006-2009). . . . .	123
22	Espectro Multifractal $f(\alpha)$ entre Boi e as demais <i>commodities</i> : Açúcar 22(a), Algodão22(b), Arroz 22(c), Café 22(d) e Trigo 22(e), respectivamente (2006-2009). . . . .	124
23	Espectro Multifractal $f(\alpha)$ entre Café e as demais <i>commodities</i> : Açúcar 23(a), Algodão23(b), Arroz 23(c), Boi 23(d) e Trigo 23(e), respectivamente (2006-2009). . . . .	125
24	Espectro Multifractal $f(\alpha)$ entre Trigo e as demais <i>commodities</i> : Açúcar 24(a), Algodão 24(b), Arroz 24(c), Boi 24(d) e Café 24(e), respectivamente (2006-2009). . . . .	126
25	Espectro Multifractal $f(\alpha)$ entre Açúcar e as demais <i>commodities</i> : Arroz 25(b), Algodão 25(a), Boi 25(c), Café 25(d) e Trigo 25(e), respectivamente (2010-2014). . . . .	127
26	Espectro Multifractal $f(\alpha)$ entre Algodão e as demais <i>commodities</i> : Arroz 26(b), Açúcar26(a), Boi 26(c), Café 26(d) e Trigo 26(e), respectivamente (2010-2014). . . . .	128
27	Espectro Multifractal $f(\alpha)$ entre Arroz e as demais <i>commodities</i> : Açúcar 27(a), Algodão27(b), Boi 27(c), Café 27(d) e Trigo 27(e), respectivamente (2010-2014). . . . .	129
28	Espectro Multifractal $f(\alpha)$ entre Boi e as demais <i>commodities</i> : Açúcar 28(a), Algodão28(b), Arroz 28(c), Café 28(d) e Trigo 28(e), respectivamente (2010-2014). . . . .	130
29	Espectro Multifractal $f(\alpha)$ entre Café e as demais <i>commodities</i> : Açúcar 29(a), Algodão29(b), Arroz 29(c), Boi 29(d) e Trigo 29(e), respectivamente (2010-2014). . . . .	131
30	Espectro Multifractal $f(\alpha)$ entre Trigo e as demais <i>commodities</i> : Açúcar 30(a), Algodão 30(b), Arroz 30(c), Boi 30(d) e Café 30(e), respectivamente (2010-2014). . . . .	132





# Lista de Tabelas

1.1	Balança comercial brasileira entre 2000 e 2013 (em US\$ bilhões). . . . .	2
3.1	Estatísticas descritivas para preços de fechamento diário da <i>commodity</i> Açúcar. . . . .	37
3.2	Estatísticas descritivas para preços de fechamento diário da <i>commodity</i> Algodão. . . . .	37
3.3	Estatísticas descritivas para preços de fechamento diário da <i>commodity</i> Arroz. . . . .	37
3.4	Estatísticas descritivas para preços de fechamento diário da <i>commodity</i> Boi. . . . .	37
3.5	Estatísticas descritivas para preços de fechamento diário da <i>commodity</i> Café. . . . .	37
3.6	Estatísticas descritivas para preços de fechamento diário da <i>commodity</i> Trigo. . . . .	37
4.1	Estatísticas descritivas dos retornos das <i>commodities</i> . . . . .	50
4.2	Cross-Sample Entropy entre todos os retornos das séries temporais das <i>commodities</i> analisadas (2006-2009). . . . .	56
4.3	Cross-Sample Entropy entre todos os retornos das séries temporais das <i>commodities</i> analisadas (2010-2014). . . . .	56
4.4	DCCA entre todos os retornos das séries temporais das <i>commodities</i> analisadas (2006-2009). . . . .	58
4.5	DCCA entre todos os retornos das séries temporais das <i>commodities</i> analisadas (2010-2014). . . . .	58

---

4.6	Teste de hipóteses sobre igualdade dos coeficientes angulares $\beta_1$ e $\beta_2$ da tendência linear nos períodos pré- crise e pós- crise, ao nível de significância de 5%. . . . .	66
4.7	Comportamento dos parâmetros multifractais. . . . .	68
4.8	Teste Wilcoxon-Mann-Whitney para os parâmetros multifractais. . . . .	69
4.9	Comportamento dos parâmetros multifractais (2006-2009). . . . .	71
4.10	Comportamento dos parâmetros multifractais (2010-2014). . . . .	71
4.11	Estatísticas descritivas da estatística <i>SampEn</i> para 1721 janelas móveis de tempo de 240 dias (2006-2014). . . . .	75
4.12	Estatísticas descritivas do Expoente de Hurst ( $H$ ) com janela móvel de tempo de 240 dias (2006-2014). . . . .	80
4.13	Estatísticas descritivas do Expoente de Hurst ( $H$ ) com janela móvel de tempo de 480 dias (2006-2014). . . . .	81
4.14	Estatísticas descritivas para o Expoente de Hurst ( $H$ ) com janela móvel de tempo de 1000 dias (2006-2014). . . . .	82
4.15	Estatísticas descritivas para o Expoente DFA ( $\alpha$ ) com janela móvel de tempo de 240 dias (2006-2014). . . . .	86
4.16	Estatísticas descritivas para o Expoente DFA ( $\alpha$ ) com janela móvel de tempo de 480 dias (2006-2014). . . . .	87
4.17	Estatísticas descritivas para o Expoente DFA ( $\alpha$ ) com janela móvel de tempo de 1000 dias (2006-2014). . . . .	88

# Lista de Abreviaturas e Símbolos

$\alpha_{m\acute{a}x}$	Valor mximo do espectro multifractal.....	27
$\alpha_{m\acute{i}n}$	Valor mnimo do espectro multifractal.....	27
$\beta$	Expoente espectral.....	18
$\gamma$	Expoente de correlao.....	15
$\langle g(t) \rangle$	Mdia aritmtica da srie $g(t)$ .....	34
$\mu_3$	Coefficiente de Assimetria.....	43
$\mu_3$	Coefficiente de Curtose.....	43
$\psi(\cdot)$	Funo indicadora para a integral de correlao.....	14
$\rho_{DCCA}(n)$	<i>Coefficiente de Correlao-Cruzada</i> .....	24
$\sigma_g$	Desvio-padro da srie $g(t)$ .....	34
$\theta$	Expoente de escala para funo de flutuao $F(n)$ .....	20
$\theta_0$	Valor da mediana sob $H_0$ no Teste de Wilcoxon-Mann-Whitney – Uma Amostra 40	
$\varepsilon$	Tamanho da grade retangular no mtodo <i>box-counting</i> .....	13
$b$	Nmero de vezes em que uma escala de uma estrutura  reduzida.....	11
$C(M, \varepsilon)$	Integral de correlao.....	14
$C(t)$	Coefficiente de Correlao-Cruzada.....	17
$C_{6,2}$	Combinao simples de 6, 2 a 2.....	17

$D$	Dimensão euclideana .....	13
$D_0$	Dimensão <i>box-counting</i> .....	13
$D_2$	Dimensão de correlação .....	14
$d_f$	Dimensão fractal .....	11
$F(n)$	Função de flutuação .....	20
$F_{ac}(t)$	Função de autocorrelação .....	17
$F_{RS}(n)$	Função de flutuação de Hurst .....	19
$g(t)$	Retorno logaritmo diário .....	34
$H$	Expoente de Hurst .....	19
$H_0$	Hipótese de nulidade .....	18
$H_1$	Hipótese alternativa .....	18
$L$	Dimensão linear do sistema na análise multifractal .....	15
$m$	Fator de aumento de um objeto geométrico auto-semelhante .....	10
$Md$	Mediana .....	39
$n$	Tamanho do segmento .....	20
$n(\varepsilon)$	Número das caixas que contém pelo menos um ponto do objeto fractal .....	13
$N(l)$	Número de unidades de uma estrutura em escala $l$ .....	11
$N(l/b)$	Número de unidades de uma estrutura em escala reduzida $l/b$ .....	11
$N_n$	Número de segmentos de tamanho $n$ .....	20
$n_p$	Número de partes de um objeto geométrico auto-semelhante .....	10
$p$	Graus de liberdade da distribuição Qui-quadrado .....	17
$P_1$	População 1 no Teste de Wilcoxon-Mann-Whitney – Duas Amostras Independentes .....	39

$P_1$	População no teste Wilcoxon-Mann-Whitney – Uma Amostra.....	40
$P_2$	População 2 no Teste de Wilcoxon-Mann-Whitney – Duas Amostras Independentes.....	39
$q$	$q$ -ésimo momento multifractal.....	15
$Q_{cc}(p)$	Estatística de correlação-cruzada.....	17
$R$	Região de dimensão linear.....	13
$R(t)$	Retorno diário normalizado.....	34
$s$	Índice para a contagem de caixas/segmentos.....	20
$S(f)$	Espectro de potências.....	18
<i>SampEn</i>	<i>Sample Entropy</i> .....	29
$V(R)$	Volume de uma região de dimensão linear $R$ .....	13
$W$	Largura do espectro multifractal.....	27
$W_o$	Largura do espectro da série original.....	70
$W_r$	Largura do espectro da série randomizada.....	70
ACR	Açúcar.....	33
ALG	Algodão.....	33
ARR	Arroz.....	33
BACEN	Banco Central do Brasil.....	33
BM&FBOVESPA	Bolsa de Valores, Mercadorias e Futuros.....	6
BOI	Boi.....	33
CAF	Café.....	33
CONAB	Companhia Nacional de Abastecimento.....	2
CRB	Commodity Research Boureau.....	33
DCCA	<i>Detrended Cross-Correlation Analysis</i> .....	23

---

DFA	<i>Detrended Fluctuation Analysis</i> .....	21
EHG	Expoente de Hurst Generalizado .....	26
FA	<i>Fluctuation Analysis</i> .....	20
IC-Br	Índice de Commodities Brasil .....	33
IRGA	Instituto Rio Grandense do Arroz.....	35
PIB	Produto Interno Bruto .....	2
TRG	Trigo.....	33

## Introdução

O crescimento da economia mundial, impulsionado por países emergentes, principalmente a China, gerou mudanças relevantes no mercado de *commodities* a partir de 2002. Observou-se uma mudança nos preços das *commodities*, que mostraram uma elevação expressiva, mostrando condições acirradas entre oferta e demanda desses produtos, impulsionadas pela existência de problemas climáticos que afetaram negativamente a oferta e pelo ritmo de crescimento da demanda [3].

Diante da crescente demanda da China (e da Índia, em uma certa medida), além de uma economia mundial em expansão, observou-se a elevação dos preços das *commodities*, sobretudo do petróleo, minerais e alimentos a partir do final de 2004 e final de 2007. Essa elevação foi tão grande quanto o primeiro choque da elevação dos preços do petróleo na década de 1970 [4].

Entre 2001 e 2007 o Brasil obteve um aumento na participação do comércio mundial: de 1.2% em 2001 para 1.5%, em 2007. A participação das exportações agrícolas brasileiras, nesse período, cresce de 4.8% para 6.7%, acima da evolução das exportações agrícolas mundiais, impactando a expansão do mercado brasileiro no mercado mundial [5].

O agronegócio brasileiro representou, em 2008, 36.7% das exportações brasileiras, geração de 37% dos empregos e 28% do Produto Interno Bruto (PIB), que é a soma de todos os bens e serviços finais produzidos durante um determinado período de tempo (geralmente um ano) dentro dos limites territoriais de um país [6]. O Brasil exportou US\$ 58.4 bilhões de dólares em produtos agrícolas no ano de 2008. No entanto, em 2009 houve um declínio de 14% nos preços, comparado com 2008. Essa queda nas cotações se deu, principalmente, em função da crise financeira mundial ocorrida em meados de 2008 [7].

A Tabela 1.1 descreve a balança comercial brasileira (em US\$ bilhões) entre os anos

de 2000 e 2013. A linha referente ao ano de 2009 encontra-se em negrito, indicando o período em que ocorre queda na exportação, importação e no saldo do agronegócio, após uma sequência crescente observada nos anos anteriores.

Tabela 1.1: Balança comercial brasileira entre 2000 e 2013 (em US\$ bilhões).

Período	Exportações	Importações	Saldo Geral	Saldo Agronegócio
2000	55.119	55.851	-0.732	14.811
2001	58.287	55.602	2.685	19.016
2002	60.439	47.243	13.196	20.347
2003	73.203	48.326	24.878	25.848
2004	96.677	62.836	33.842	34.134
2005	118.529	73.600	44.929	38.416
2006	137.807	91.351	46.457	42.727
2007	160.649	120.617	40.032	49.696
2008	197.942	172.985	24.958	57.714
<b>2009</b>	<b>152.995</b>	<b>127.722</b>	<b>25.272</b>	<b>54.800</b>
2010	201.915	181.768	20.147	63.000
2011	256.040	226.243	29.796	77.510
2012	242.580	223.142	19.438	79.408
2013	242.178	239.617	2.561	82.907

Fonte: CONAB / Min. da Agricultura / Min. do Desenvolvimento [5; 8; 9].

Considerando que o setor agrícola é de suma importância para a sanidade econômica e por ser um grande investidor em tecnologias ambiental e rural, o Brasil não pode sucumbir à ideia de uma desaceleração neste setor. O Governo Federal brasileiro tomou algumas medidas adequadas para diminuir os efeitos da crise econômica de 2008-2009. No setor agrícola, houve aumento do limite dos compulsórios e liberação de mais recursos para poupança rural e cooperativas [7].

A partir do quarto trimestre de 2008 até esse mesmo período de 2009, a economia brasileira passou por um momento em que os índices de crescimento despencaram para abaixo dos 2% negativos em média por trimestre, um reflexo da forte retração pela qual uma economia dependente da exportação de *commodities* passaria diante de um cenário de intensa crise de demanda [10].

A turbulenta crise financeira mundial iniciada com a crise do mercado *subprime* americano, interrompeu a trajetória do crescimento da economia mundial, causando retração nas exportações.



O *subprime* é um crédito a habitação de alto risco destinado a uma fatia da população com rendimentos mais baixos e com uma situação econômica mais instável [11]. De acordo com Galle *et al.* [12], trata-se de modalidade de empréstimo de segunda linha que apresenta maior risco de inadimplência por se tratar de pessoas de baixa renda, muitas vezes com histórico de pagamento não muito bom.

Os efeitos da crise do *subprime*, associados à valorização do real frente ao dólar surtiram efeitos na balança comercial do agronegócio brasileiro já no ano de 2009, tendo as exportações recuadas em 11.27% em comparação com o ano anterior e as importações queda de 29.71%. Em 2008 o agronegócio apresentou saldo positivo em sua balança comercial, fato decorrente da elevação dos preços das *commodities* no mercado internacional e da expansão agrícola durante o ano [7].

O objetivo geral dessa tese é avaliar a dinâmica da flutuação dos preços das *commodities* no agronegócio brasileiro, por meio de métodos da análise dos sistemas complexos diante da pior crise financeira internacional (2008/2009) desde a quebra da bolsa de Nova York em 1929.

Objetivos específicos foram:

- (i) Aplicar os métodos Expoente de Hurst ( $H$ ), Expoente de Hurst em Janela Móvel de Tempo (*Rolling Window*), Coeficiente de Correlação Cruzada ( $\rho_{DCCA}(n)$ ), *Multifractal Detrended Cross-Correlation Analysis* (MF-DCCA), *Multifractal Detrended Fluctuation Analysis* (MF-DFA) e *Detrended Cross-Correlation Analysis* (DCCA) para comparar as correlações cruzadas e autocorrelações nos retornos das *commodities* analisadas para os períodos pré-crise e pós-crise financeira internacional.
- (ii) Aplicar o método *Sample-Entropy* em Janela Móvel de Tempo (*Rolling Window*) para os retornos das *commodities* e utilizar o método *Cross-Sample Entropy* para análise da complexidade entre as séries temporais dos retornos das *commodities* avaliadas nos períodos anterior e posterior a crise financeira internacional.

A estrutura desta tese está dividida em 5 capítulos. O primeiro capítulo corresponde a esta introdução.

No capítulo 2, apresentam-se os aspectos gerais da crise financeira mundial que ocorreu em meados de 2008 e algumas características relacionadas às *commodities*, tais como: a forma em que são negociadas e a sua importância econômica para a negociabilidade a nível mundial. Ainda no capítulo 2, há uma descrição dos métodos fractais e multifractais, juntamente com análise de correlações de longo alcance, por

meio dos métodos: Análise de Hurst, Análise de Hurst em Janela Móvel de Tempo (*Rolling Window*), *Detrended Fluctuation Analysis* (DFA), *Detrended Fluctuation Analysis* em Janela Móvel de Tempo (*Rolling Window*), *Detrended Cross-Correlation Analysis* (DCCA), Coeficiente de Correlação-Cruzada ( $\rho(\cdot)$ ), *Multifractal Detrended Fluctuation Analysis* (MF-DFA), *Multifractal Detrended Cross-Correlation Analysis* (MF-DCCA). Além disso, a análise de complexidade em séries temporais é apresentada por meio dos métodos *Sample Entropy*, *Sample Entropy* em Janela Móvel de Tempo (*Rolling Window*) e *Cross-Sample Entropy*.

O conjunto de dados, sua descrição e os preços de fechamentos diários das 6 *commodities* agrícolas analisadas e a metodologia empregada são discutidos no capítulo 3.

Os resultados podem ser observados no capítulo 4, onde os métodos descritos do capítulo 2 são aplicados.

Finalmente, o capítulo 5 apresenta as conclusões e sugestões para trabalhos futuros.

## Revisão de Literatura

### 2.1 As *commodities* e a crise financeira

No ano de 1948 houve uma reunião das Nações Unidas em Havana–Cuba, para tentar criar uma organização internacional para o comércio. Nessa reunião, definiu-se o termo *commodities* como todo produto de atividades agrícolas, florestal e pesqueira, e todo mineral, obtidos *in natura* ou que sofressem alguma transformação suficiente para a sua negociabilidade em quantidades significativas nos mercados mundiais. Com o passar do tempo, essa definição sofreu algumas alterações e absorveram novas mercadorias, incorporando basicamente materiais não duráveis e produtos intermediários como: ouro, petróleo, cobre, minério de ferro, etanol, soja, milho, trigo, açúcar, suco de laranja concentrado congelado e até energia elétrica [13].

#### 2.1.1 *Commodities*

*Commodity* é um termo de língua inglesa (plural *commodities*), que significa mercadoria. É utilizado nas transações comerciais de produtos de origem primária nas bolsas de mercadorias. É um termo usado para produtos de base e em estado bruto (matérias-primas) ou com pequeno grau de industrialização, de qualidade quase uniforme, produzidos em larga escala e por diferentes produtores. Estes produtos *in natura*, cultivados ou de extração mineral, podem ser armazenados por determinado período sem perda significativa de qualidade. Possuem cotação e negociabilidade globais, utilizando bolsas de mercadorias [9].

O que torna as *commodities* muito importantes na economia é o fato de que, embora sejam mercadorias primárias, ou minimamente industrializadas possuem negociabilidade global. Isto ocorre em bolsas de mercadorias, portanto seus preços são definidos

em nível global, pelo mercado internacional de acordo com fatores de oferta e demanda. Por este motivo são suscetíveis a oscilações nas cotações de mercado, em virtude de perdas e ganhos nos fluxos financeiros no mundo [14].

Em geral, as *commodities* são produzidas em grandes quantidades por vários produtores/empresas. Não apresentam diferenciação (apresentam baixo valor agregado), marca de referência ou serviço que as diferenciem. Basicamente, são negociadas em duas formas: mercado à vista e futuro (fecha-se um contrato para entrega/pagamento futuro) e nas Bolsas de Mercadorias, onde são negociadas em quantidades padrões: por exemplo, na BM&FBOVESPA o dólar é negociado em contratos de US\$ 10.000 e o café em contratos de 100 sacas de 60 Kg. Os principais tipos de *commodities* são [14]:

- (i) Agrícolas. Exemplos: café, trigo, soja, milho, açúcar, farelo de soja;
- (ii) Minerais. Exemplos: ouro, petróleo, ferro, alumínio;
- (iii) Financeiras. Exemplos: dólar, euro, real, índices futuros;
- (iv) Ambientais. Exemplos: créditos de carbono, condições climáticas médias em regiões do planeta;
- (v) Recursos energéticos. Exemplo: energia elétrica;
- (vi) Químicas. Exemplos: ácido sulfúrico, sulfato de sódio, fertilizantes.

### 2.1.2 A Crise *Subprime*

O mercado *subprime*<sup>1</sup> entrou na economia após a crise das empresas “ponto com” nos Estados Unidos (EUA). O Federal Reserve (Fed), uma espécie de banco central americano, reduziu a taxa de juros para acelerar a economia. Com isso, o setor imobiliário aproveitou o crescente momento de crédito (os juros chegaram a 1% ao ano). Era viável e acessível para a maioria das pessoas realizar compra da casa própria. Assim, muitas pessoas compraram casas financiadas e quem já era proprietário de imóvel, hipotecava usando o dinheiro para quitar dívidas ou gastar mais. A compra de imóveis se transformou em investimento e acreditava-se na valorização. Os *subprime* foram transformados em papéis e vendidos para outras instituições financeiras, que por sua vez, emprestavam novas quantias antes mesmo de receber pelos papéis comprados. Isso tudo se tornou em um ciclo de dívidas [12].

---

<sup>1</sup>Para maiores detalhes, o filme *Inside Job* (dublado em português) descreve a crise financeira de 2008. Este filme foi vencedor de vários prêmios. Assista em: <http://www.filmesonlinegratis.net/assistir-trabalho-interno-dublado-online.html>.

Com a elevação das taxas de juros, o crédito encareceu, o mercado imobiliário desaqueceu, houve desvalorização do preço dos imóveis, pois já não era mais interessante comprar imóveis com juros altos. Conseqüentemente, quem comprou casas não conseguiu vendê-las [12]. Em 2006, iniciou-se um acelerado processo de execução de hipotecas dos EUA, o que resultaria em uma crise de proporções globais [15]. Resultado: a pior crise desde 1929.

O último trimestre de 2008 marcou o agravamento da crise financeira mundial originada no mercado de hipotecas de alto risco (*subprime*) dos Estados Unidos. Esta é considerada a maior crise do capitalismo moderno desde 1929 (quando a bolsa de Nova York quebrou) [12; 16; 17]. A falência do banco de investimento Lehman Brothers<sup>2</sup> no dia 15 de setembro de 2008 marca a transformação da crise financeira internacional [18]. A crise passou a ser global, afetando todas as classes de títulos financeiros, comprometendo o desempenho do comércio internacional e das economias emergentes e avançadas [19]. Ainda em seu início, a crise causou uma perda acumulada de 2.2 trilhões de dólares [20].

### 2.1.3 Consequências da crise no Brasil

Até setembro de 2008, não se poderia prever uma valorização do real frente ao dólar americano, devido a política monetária rigorosa adotada pelo governo brasileiro. Muitas empresas apostaram que o dólar continuaria caindo até o final do ano de 2008. Quando o dólar começou a subir, houve a necessidade de cobrir os prejuízos que somaram mais de 5 bilhões de reais. Em setembro de 2008, com a queda do banco norte-americano Lehman Brothers, ocorre uma grande redução do crédito internacional e o dólar dispara no Brasil [11].

A crise internacional de confiança nos bancos e a falta de crédito externo afetaram os pequenos e médios bancos no Brasil. O Banco Central brasileiro realizou mudanças no recolhimento de depósitos compulsórios, beneficiando bancos menores e instituições que trabalhavam com *leasing*<sup>3</sup> [7].

---

<sup>2</sup>Filme legendado em português: os últimos dias do Lehman Brothers. Assista em <https://www.youtube.com/watch?v=s6QRJT1vBD4>

<sup>3</sup>*Leasing*: é um sistema de arrendamento mercantil, ou seja, de aluguel com opção de compra [21].

## 2.2 Fractais e Multifractais

### 2.2.1 Fractais

Modelos com estrutura fractal são usados para representar ou descrever uma grande variedade de processos naturais em vários campos da ciência, desde a geologia à astrofísica e da física à biologia.

Os fractais são conjuntos cuja forma é extremamente irregular ou fragmentada. Além disso, têm essencialmente a mesma estrutura em todas as escalas [22; 23].

Algumas aplicações dentre muitas incluem: compactação de imagens [24], codificação e decodificação de áudio e vídeo [25], classificação de imagens e estudo de paisagens [26].

As principais propriedades que caracterizam e que permitem definir os conjuntos fractais são as seguintes [27]:

- (i) Autossimilaridade: partes do objeto se assemelham ao todo. A autossimilaridade pode ser exata (determinística) ou estatística (estocástica). Isto significa que o sistema é invariante (mantém a mesma forma e estrutura) sob uma transformação de escala (transformação que reduz ou amplia o objeto ou parte dele);
- (ii) Extrema irregularidade: no sentido de rugosidade (não-suavidade) ou fragmentação;
- (iii) Dimensão fractal: possuem, em geral, uma dimensão fractal não-inteira. Essa dimensão quantifica, de certo modo, o grau de irregularidade ou fragmentação do conjunto considerado. Uma alta dimensão monofractal significa um maior grau de complexidade, uma estrutura de forma mais irregular [23; 28; 29].

Os fractais geométricos podem ser classificados em duas categorias [23; 30–32]:

- (i) Fractais determinísticos: são gerados por meio de iterações e possuem autossimilaridade exata em todas as escalas. Exemplos de fractais determinísticos: conjunto de Cantor, tapete de Sierpinski, triângulo de Sierpinski, curvas de Peano, floco de neve de Koch, etc.
- (ii) Fractais estocásticos: são fractais gerados por processos estocásticos, que possuem autossimilaridade estatística. São conhecidos também por fractais naturais. Exemplos de fractais naturais: árvores, nuvens, linhas costeiras, sistemas fluviais, relâmpagos, redes arteriais, neurônios, etc.

As primeiras quatro iterações do fractal determinístico do triângulo Sierpinski estão apresentadas na Figura 2.1 e autossimilaridade estocástica pode ser observada na Figura 2.2.



Figura 2.1: Fractal determinístico. Triângulo de Sierpinski.



Figura 2.2: Exemplo de fractal estocástico: brócolis romanesco.

A complexidade geométrica de um fractal pode ser descrita, de uma forma geral, por sua dimensão [23; 30]. Dentre as características que definem um fractal, sua dimensão é a mais importante. A dimensão fractal representa o grau de irregularidade e ocupação no espaço. Assim, quanto maior a irregularidade de uma forma, maior é a sua dimensão fractal. Essa característica da dimensão fractal a torna uma ferramenta muito útil para a comparação de diferentes formas fractais [33].

A dimensão fractal representa o grau de ocupação de uma estrutura fractal no espaço que a contém [1]. A seguir, apresenta-se um exemplo do cálculo de dimensão de objetos conhecidos. A Figura 2.3 ilustra três objetos: uma reta, um quadrado e um cubo.



Figura 2.3: Segmento de reta, quadrado e cubo divididos em partes menores, semelhantes a inicial. Fonte: [1].

De acordo com a Figura 2.3, o segmento de reta é dividido em 5 partes iguais, o quadrado em 9 e o cubo em 8. Cada uma dessas partes se apresenta semelhante à figura inicial. Observa-se que ao multiplicar cada parte do segmento de reta por 5 ( $5^1$ ), obtém-se a reta toda. Portanto, o fator de aumento é 5. Em relação ao quadrado, o fator de aumento é 3 ( $3^2 = 9$ ). Por último, o fator de aumento é 2 para o cubo ( $2^3 = 8$ ). As dimensões são respectivamente: 1, 2 e 3.

Portanto, tem-se a seguinte relação:

- (i) No segmento de reta, o número de partes menores é 5, podendo ser escrito como  $5 = 5^1$ , onde 5 é o fator de aumento e 1 é a dimensão do segmento.
- (ii) Para o quadrado, tem-se 9 partes menores, isto é, existem  $9 = 3^2$  quadrados menores. Assim, 3 é o fator de aumento e 2 representa a dimensão do quadrado.
- (iii) Em relação ao cubo, obtém-se 8 partes menores, ou seja,  $8 = 2^3$  cubos menores. Portanto, o fator de aumento é 2 e a dimensão do cubo é 3.

O número  $n_p$  de partes de um objeto geométrico auto-semelhante, em geral, é dado por:

$$n_p = m^D \quad (2.1)$$

onde  $m$  é o fator de aumento e  $D$  é a dimensão do objeto geométrico inicial. Assim, para obter a dimensão de um objeto qualquer, basta encontrar o valor de  $D$ , aplicando o logaritmo em ambos os lados da Equação 2.1. Portanto, segue que

$$D = \frac{\log n_p}{\log m} \quad (2.2)$$

onde  $n_p$  é o número de partes em que o objeto foi dividido e  $m$  é o fator de aumento.



Para fractais determinísticos, a dimensão fractal pode ser calculada da seguinte forma: seja  $N(l)$  o número de unidades da estrutura em escala  $l$ . A retração da escala  $b$  vezes gera um novo número de unidades estruturais.

$$N\left(\frac{l}{b}\right) = N(l)b^{d_f} \tag{2.3}$$

Assim, a dimensão fractal  $d_f$  pode ser calculada por

$$d_f = \frac{\log \left[ \frac{N(l/b)}{N(l)} \right]}{\log b} \tag{2.4}$$

Alguns exemplos de fractais determinísticos e suas dimensões [34]:

- (i) O conjunto de Cantor: para obter o conjunto de Cantor, deve-se remover um terço central do segmento. Na primeira iteração, sobram dois segmentos de tamanho  $1/3$ . Removendo um terço central de cada um deles, sobram  $2^2$  segmentos de tamanho  $1/3^2$ . Na iteração  $n$ , sobram  $2^n$  segmentos de tamanho  $1/3^n$ . A Figura 2.4 exibe algumas iterações do Conjunto de Cantor.

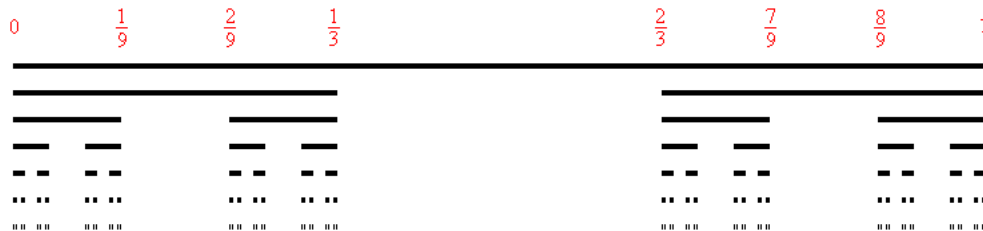


Figura 2.4: Conjunto de Cantor. Fonte: [2].

Como o conjunto é formado pelas duas cópias com escala de  $1/3$ , sua dimensão é  $d_f = \frac{\log 2}{\log 3} = 0.6309$ .

- (ii) Curvas de Koch: um segmento é substituído por um segmento polinomial formado por 4 segmentos de tamanho  $1/3$ . O processo é aplicado indefinidamente em cada um dos segmentos. Na etapa  $n$ , terá uma linha poligonal com  $4^n$  segmentos de tamanho  $1/3^n$ . Esta curva é contínua, mas não tem derivadas em nenhum ponto.

A curva de Koch é formada pelas 4 cópias com redução de  $1/3$ . Logo, sua dimensão fractal é de  $d_f = \frac{\log 4}{\log 3} = 1.2618$ . A Figura 2.5 descreve o Conjunto de Koch.

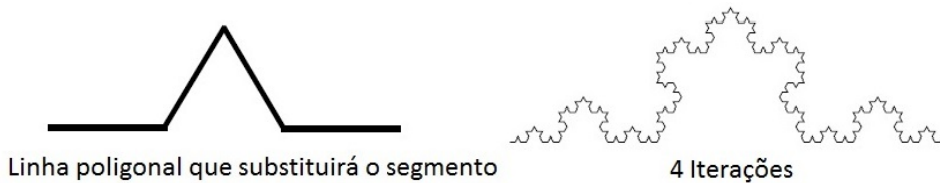


Figura 2.5: Curvas de Koch. Fonte: [2].

(iii) A Esponja de Menger é construída a partir de um cubo, por meio do seguinte processo recursivo [1]:

- (1) Considere um cubo qualquer (Figura 2.6 (a));
- (2) Divida cada face do cubo em 9 quadrados. Desse modo, o cubo inicial fica subdividido em 27 cubos menores;
- (3) Remova o cubo localizado no meio de cada face e o cubo central, deixando apenas 20 cubos restantes (Figura 2.6 (b)). Este é o primeiro nível da Esponja de Menger.
- (4) Repita os passos 2 e 3 para cada um dos 20 pequenos cubos restantes do nível anterior. Assim, obtemos o segundo nível da Esponja (Figura 2.6 (c)). Note que, neste nível, ocorre a divisão de cada um dos 20 cubos do nível anterior em outros 20 cubos menores, obtendo no final  $20^2$  cubos.
- (5) A Esponja de Menger é o limite deste processo depois de um número infinito de iterações.

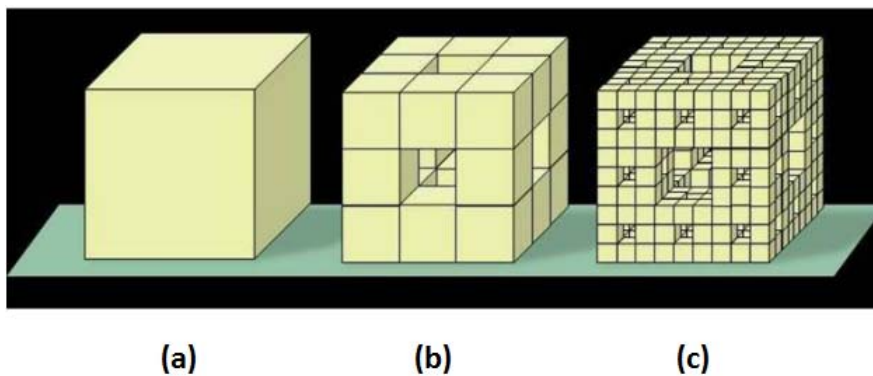


Figura 2.6: Esponja de Menger. Fonte: [1]

A esponja de Menger é formada pelas 20 cópias com redução de  $1/3$ . Logo, sua dimensão fractal é de  $d_f = \frac{\log 20}{\log 3} = 2.727$ .

Para fractais estocásticos, a dimensão fractal pode ser calculada por:

$$V(R) \propto R^{d_f} \quad (2.5)$$

onde  $V(R)$  é o volume da região de dimensão linear  $R$  e  $d_f$  a dimensão fractal, um número não inteiro menor que a dimensão euclidiana  $D$  do espaço em que o fractal se encontra. Existem vários métodos para calcular a dimensão fractal, dentre eles, os mais usados são: método da contagem de caixas (*box-counting*), dimensão de informação e dimensão de correlação [31].

### Método da contagem de caixas

Para calcular a dimensão fractal, utiliza-se uma grade retangular de tamanho  $\varepsilon$ . Cobre-se o objeto com uma grade de tamanho  $\varepsilon$  e contam-se o número das caixas  $n(\varepsilon)$  que contém pelo menos um ponto do objeto, onde  $n(\varepsilon) \sim \varepsilon^{-D_0}$ . Repete-se o procedimento para diferentes tamanhos de caixas e faz-se um gráfico log-log de  $n(\varepsilon)$  versus  $\varepsilon$ . A inclinação negativa desse gráfico é a dimensão *box-counting*, que é definida formalmente por [35]:

$$D_0 \equiv \lim_{\varepsilon \rightarrow 0} \frac{\log \frac{1}{n(\varepsilon)}}{\log \varepsilon} = - \lim_{\varepsilon \rightarrow 0} \frac{\log n(\varepsilon)}{\log \varepsilon} \quad (2.6)$$

A estimativa da dimensão *box-counting* é a inclinação negativa da reta no gráfico log-log de  $n(\varepsilon)$  versus  $\varepsilon$  [35–37]. A Figura 2.7 ilustra um exemplo do método *box-counting* cuja dimensão é 0.5642.

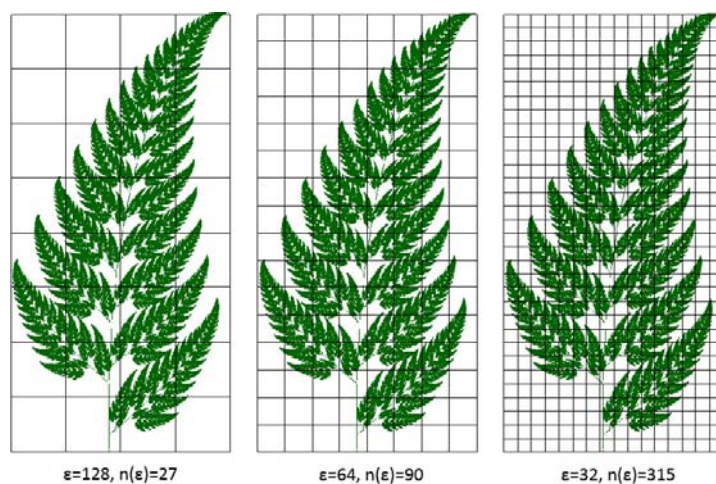


Figura 2.7: Exemplo do uso do método *box-counting*.

### Dimensão da informação

A dimensão da informação está relacionada com a probabilidade de encontrar um ponto da estrutura dentro de uma caixa. Cobre-se o objeto com uma grade de tamanho  $\varepsilon$  e conta-se o número de pontos  $M_i$  dentro da  $i$ -ésima caixa. A probabilidade de encontrar um ponto escolhido ao acaso dentro da  $i$ -ésima caixa é  $P_i = \frac{M_i}{M}$ , onde  $M$  é o número total de pontos da estrutura.

A dimensão da informação é definida por:

$$D_1 \equiv \lim_{\varepsilon \rightarrow 0} \frac{\sum_{i=1}^M P_i \log P_i}{\log \varepsilon} \quad (2.7)$$

A estimativa da dimensão de informação é a inclinação da reta no gráfico log-log de  $\sum_{i=1}^M P_i \log P_i$  versus  $\varepsilon$  [35–37].

### Dimensão de correlação

A dimensão de correlação, proposta por Grassberger & Procaccia [38; 39], está relacionada com o número de pontos cuja distância entre si é inferior a  $\varepsilon$ .

A dimensão de correlação é definida por:

$$D_2 \equiv \lim_{\varepsilon \rightarrow 0, M \rightarrow \infty} \frac{\log C(M, \varepsilon)}{\log \varepsilon} \quad (2.8)$$

onde

$$C(M, \varepsilon) = \frac{1}{M(M-1)} \sum_{i \neq j} \psi(\varepsilon - \|x_i - x_j\|) \quad (2.9)$$

é a integral de correlação em que  $\psi(\cdot)$  representa a função indicadora e  $M$  é o número total de pontos da estrutura.

A estimativa da dimensão de correlação é dada pela inclinação da reta no gráfico log-log de  $C(M, \varepsilon)$  versus  $\varepsilon$ .

## 2.2.2 Multifractais

O termo Multifractal se refere a uma coleção de simples fractais entrelaçados, em que, cada um possui sua própria dimensão fractal. Assim, um objeto multifractal necessita de uma hierarquia de expoentes de escala para caracterizar sua estrutura. Conseqüentemente, pode-se dizer que um objeto multifractal pode ser visto como uma extensão de objetos fractais [30; 40]. Alguns exemplos de objetos multifractais: tempestade com raios, nuvens, partituras musicais [41].

## 2.2.3 Análise Multifractal

A análise multifractal fornece mais informações acerca do espaço de propriedades do que a análise monofractal. A vantagem da análise multifractal é que ela caracteriza as propriedades de escalas locais, além das propriedades globais [28; 30]. Para o cálculo da dimensão generalizada  $D_q$ , pode-se utilizar o método *box-counting* em que analisa-se o número de partículas dentro de uma região da estrutura [42; 43]. Para calcular a dimensão multifractal generalizada, deve-se cobrir a estrutura analisada com caixas de aresta de tamanho  $\varepsilon$ , para diferentes valores de  $\varepsilon$ , registrando os valores de  $M_i$  dentro da  $i$ -ésima caixa, onde  $M_0$  é número total de partículas do sistema e  $L$  é a dimensão linear do sistema. A dimensão generalizada  $D_q$  é definida por:

$$\sum_i \left( \frac{M_i}{M_0} \right)^q \sim \left( \frac{\varepsilon}{L} \right)^{(q-1)D_q} \quad (2.10)$$

onde  $q$  é uma variável contínua que permite observar as propriedades fractais em diferentes escalas. Valores positivos de  $q$  representam regiões com maior densidade e valores negativos de  $q$  regiões com baixa densidade. Para estruturas monofractais, todas as dimensões generalizadas são iguais, resultando em um único valor de dimensão fractal. Para uma abordagem alternativa da multifractalidade, usa-se o espectro  $f(\alpha)$  [30; 43; 44], onde:

$$N(\alpha) = L^{-f(\alpha)} \quad (2.11)$$

representa o número de caixas  $N(\alpha)$  tal que a probabilidade  $P_i$  de encontrar uma partícula dentro da  $i$ -ésima caixa segue uma lei de potência

$$P_i = L^{\alpha_i} \quad (2.12)$$

sendo  $f(\alpha)$  a dimensão fractal da união de regiões com singularidade entre  $\alpha$  e  $\alpha + d\alpha$ , com  $\alpha$  variando em  $(-\infty, \infty)$ . A relação entre a função  $D_q$  e o espectro  $f(\alpha)$  é feita

via transformada de Legendre

$$f(\alpha(q)) = q \alpha(q) - \tau(q) \quad (2.13)$$

em que

$$\alpha(q) = \frac{\partial \tau(q)}{\partial q} \quad (2.14)$$

e

$$\tau(q) \equiv (q - 1)D_q \quad (2.15)$$

No caso de estrutura monofractal, a dimensão fractal não depende de  $q$  ( $D_q \equiv D$ ) e usando as Equações 2.13, 2.14 e 2.15, tem-se que  $f(\alpha) = D$  e o espectro  $f(\alpha)$  consiste em um único ponto, onde  $f(\alpha)$  é igual a dimensão fractal da estrutura. Para multifractais  $F_q()$  é uma função decrescente em relação a  $q$  e  $f(\alpha)$  uma função côncava para baixo [30; 45].

## 2.2.4 Processos Fractais

Flutuações em séries temporais possuem a característica de autossimilaridade estatística de forma análoga a objetos fractais nas várias escalas espaciais. Desta forma, características fractais podem ser estendidas a séries temporais com flutuações em várias escalas de tempo.

Considere uma série temporal  $\{x(t)|t = 1, \dots, N\}$  que possui propriedade de autossimilaridade com o parâmetro  $\alpha$  se:

$$x(t) = a^\alpha x\left(\frac{t}{a}\right) \quad (2.16)$$

em que  $=$  se refere a igualdade de propriedades estatísticas. Essa igualdade surge devido a mudanças de escalas para  $t$  e  $x(t)$  usando diferentes fatores:  $t \rightarrow t/a$  e  $x(t) \rightarrow a^\alpha x(t)$ . O expoente  $\alpha$  é denominado de parâmetro de autossimilaridade [46].

## 2.3 Análise de Correlações de Longo Alcance em Séries Temporais Fractais

### 2.3.1 Análise da Função de Autocorrelação

Sabe-se que uma série temporal é estacionária quando ela flutua em torno de uma mesma média ao longo do tempo, conseqüentemente todas as distribuições unidimensionais são invariantes sob translação do tempo. Para séries temporais estacionárias, com média  $\bar{x}$  e variância  $\sigma^2$ , a função de autocorrelação é definida como:

$$F_{ac}(t) = \frac{E\left(\left[x(i) - \bar{x}\right]\left[x(i+t) - \bar{x}\right]\right)}{\sigma^2} \quad (2.17)$$

Para uma série não correlacionada,  $F_{ac}(t) = 0$  para  $t > 0$ . Para uma série que possui correlação de curto alcance  $F_{ac}(t)$  tem um decaimento exponencial onde  $F_{ac}(t) \sim \exp(-t/t_0)$ , sendo  $t_0$  uma escala característica. Para uma série que possui correlação de longo alcance,  $F_{ac}(t)$  tem decaimento seguindo uma lei de potência  $F_{ac}(t) \sim t^{-\gamma}$ , com o expoente de correlação  $0 < \gamma < 1$ . No caso de séries não-estacionárias, a média não está bem definida e para grandes escalas  $t$ ,  $F_{ac}(t)$  flutua ao redor de zero, o que dificulta a obtenção do expoente de correlação  $\gamma$  [47].

### 2.3.2 Teste de Correlação-Cruzada

Utilizou-se o teste de correlações-cruzadas proposto por Podobnik *et al.* [48]. O objetivo deste teste é qualificar as correlações-cruzadas em existentes ou não existentes. O teste foi realizado para cada par entre as *commodities* analisadas ( $C_{6,2} = 15$  pares). Este teste é análogo ao teste de Ljung-Box [49], sendo amplamente usado para testar correlações-cruzadas [50–55].

A correlação-cruzada entre duas séries  $\{x(t)|t = 1, 2, \dots, N\}$  e  $\{y(t)|t = 1, 2, \dots, N\}$  é definida como

$$Q_{cc}(p) = N^2 \sum_{t=1}^p \frac{C^2(t)}{N-t} \quad (2.18)$$

onde o coeficiente de correlação-cruzada  $C(t)$  é definido por

$$C(t) = \frac{\sum_{k=t+1}^N x(k)y(k-t)}{\sqrt{\sum_{k=1}^N x^2(k) \sum_{k=1}^N y^2(k)}} \quad (2.19)$$

Podobnik *et al.* [48] propuseram que a estatística de correlação-cruzada  $Q_{cc}(p)$  segue aproximadamente uma distribuição Qui-quadrado com  $p$  graus de liberdade, ou seja,  $Q_{cc}(p) \approx \chi^2(p)$ . Ela pode ser usada para testar a hipótese de nulidade ( $H_0$ ) de que nenhum dos  $p$  primeiros coeficientes de correlações-cruzadas é diferente de zero [48].

Estas são as hipóteses a serem testadas:

$$\begin{cases} H_0 : Q_{cc}(p) = 0 & \text{(Não há correlação-cruzada entre as séries temporais } x(t) \text{ e } y(t)) \\ H_1 : Q_{cc}(p) \neq 0 & \text{(Há correlação-cruzada entre as séries temporais } x(t) \text{ e } y(t)) \end{cases}$$

### 2.3.3 Análise Espectral

O espectro de potências é definido como a transformada de Fourier da função de autocorrelação [56; 57]:

$$S(f) = F_{ac}(0) + 2 \sum_{n=1}^{\infty} F_{ac}(n) \cos(2\pi f n) \quad (2.20)$$

Para séries temporais que possuem correlação de longo alcance  $S(f)$  segue uma lei de potência

$$S(f) \sim f^{-\beta} \quad (2.21)$$

com  $\beta = 1 - \gamma$ . O expoente espectral  $\beta$  e o expoente de correlação  $\gamma$  são obtidos através da inclinação da reta no gráfico log-log de  $S(f)$  versus  $f$ . A análise espectral é aplicada apenas a séries temporais estacionárias.

### 2.3.4 Análise de Hurst

Este método é baseado no passeio aleatório e foi desenvolvido pelo engenheiro inglês H. E. Hurst. Ele estudou a hidrologia do Rio Nilo, Egito, investigando os períodos de cheia [58–60]. A seguir, apresentam-se os passos para a realização da análise clássica de Hurst.

- (i) Divide-se a série original  $\{x(t)|t = 1, \dots, N\}$  em  $N_n = \text{int}(N/n)$  segmentos de tamanho  $n$  e, em cada segmento  $s$ , com  $s = 1, \dots, N_n$ , a série original é integrada pela equação:

$$X_{sk} = \sum_{i=1}^k \left( x((s-1)n + i) - \langle x_s \rangle \right) \quad (2.22)$$

onde  $\langle x_s \rangle = \frac{1}{n} \sum_{i=1}^n x((s-1)n + i)$  é a média local da série no  $s$ -ésimo segmento.



(ii) Para cada segmento, calcula-se a diferença:

$$R_s(n) = \max\{X_{sk}\} - \min\{X_{sk}\}; \quad k = (s-1)n + 1, \dots, sn \quad (2.23)$$

e o desvio-padrão é

$$S_s(n) = \sqrt{\frac{1}{n} \sum_{i=1}^n (x(s-1)n + i) - \langle x_s \rangle)^2} \quad (2.24)$$

A função de flutuação de Hurst é dada por:

$$F_{RS}(n) = \sqrt{\frac{1}{N_n} \sum_{s=1}^{N_n} \frac{R_s(n)}{S_s(n)}} \sim n^H \quad \text{para } n > 1 \quad (2.25)$$

O expoente de Hurst  $H$  está relacionado com o expoente de correlação  $\gamma$  e com o coeficiente  $\beta$  da análise espectral pela equação  $2H = 1 + \beta = 2 - \gamma$  [60].

O expoente de Hurst  $H$  é interpretado da seguinte forma [30; 58]:

- (i) Em relação ao ruído branco (séries não correlacionadas)  $H = 0.5$  a função de autocorrelação diminui exponencialmente;
- (ii) O valor  $0.5 < H < 1$  indica que a série original possui correlações de longo alcance persistentes, ou seja, valores grandes (pequenos) têm maior probabilidade de serem seguidos por valores grandes (pequenos) e a função de autocorrelação diminui seguindo uma lei de potência  $C(n) \sim n^{-\gamma}$ , com  $\gamma = 2 - 2H$ ;
- (iii) Para  $0 < H < 0.5$  a série é antipersistente, significando que os valores grandes (pequenos) tem maior probabilidade de serem seguidos por valores pequenos (grandes);
- (iv) O valor  $H = 1$  representa um ruído tipo  $1/f$  (o espectro de potências segue uma lei de potências)  $S(f) \sim f^{-\beta}$ , com  $\beta = 2H - 1 = 1$ ;
- (v) Para  $H > 1$  as correlações existem, porém não diminuem seguindo uma lei de potência. O caso especial  $H = 1.5$ , indica um ruído Browniano, que é a integração do ruído branco.

### Expoente de Hurst em Janela Móvel de Tempo (*Rolling Window*)

Este método avalia o comportamento do Expoente de Hurst  $H$  em janelas móveis de tempo, com tamanhos de janela alternativos. Alguns estudos incluem janelas de

tamanho 4 anos de janela de tempo (aproximadamente 1000 dias) [61–66], enquanto em outros estudos fixou-se o tamanho da janela em 2 anos (aproximadamente 480 dias) [67], 3 anos (aproximadamente 720 dias) [68] e 1 ano (aproximadamente 240 dias)[69; 70].

Por exemplo: considere uma série temporal de tamanho 300 e janela móvel de 240 dias, com passo de um dia entre cada janela. As janelas de tempo variam da seguinte forma:  $[1, \dots, 240]$ ,  $[2, \dots, 241]$ ,  $[3, \dots, 242]$ ,  $\dots$ ,  $[61, \dots, 300]$ .

### 2.3.5 *Fluctuation Analysis* (FA)

*Fluctuation Analysis* (FA) ou Análise de Flutuações é um método baseado no passeio aleatório. Seja  $\{x(t)|t = 1, \dots, N\}$  uma série temporal de comprimento  $N$ . Primeiramente, integramos (soma acumulada) a série original, formando uma nova série, tal que,  $X(k) = \sum_{i=1}^k x_i$ , com  $k = 1, \dots, N$ . Em seguida, a série  $\{X(k), k = 1, \dots, N\}$  é dividida em  $N_n$  segmentos não sobrepostos de tamanho  $n$ , a partir do início da série.

A função de flutuação  $F(n)$  é obtida para cada segmento  $s = \{1, \dots, N_n\}$ , pela equação:

$$F_{FA}^2(s, n) = \left[ X((s-1)n+1) - X(sn) \right]^2 \quad (2.26)$$

A função de flutuação  $F(n)$  para o tamanho de segmento  $n$  é dada pela expressão:

$$F(n) = \left[ \frac{1}{N_n} \sum_{s=1}^{N_n} F_{FA}^2(s, n) \right]^{1/2} \quad (2.27)$$

Nos casos de correlações de longo alcance, a função de flutuação  $F(n)$  aumenta segundo uma lei de potência, ou seja,  $F(n) \sim n^\theta$  [71].

### 2.3.6 *Detrended Fluctuation Analysis (DFA)*

O método *Detrended Fluctuation Analysis* (Análise de Flutuação Sem Tendência) foi introduzido por Peng *et al.* [71] para quantificar as correlações entre séries temporais não-estacionárias, ou seja, séries temporais que flutuam em torno de uma média durante um tempo e posteriormente oscilam em torno de outro nível, ou ainda, que variam em uma direção durante um tempo e mudam para outra direção no decorrer do tempo.

As Figuras 2.8 e 2.9 exibem, respectivamente, exemplos de série temporal estacionária e não estacionária.

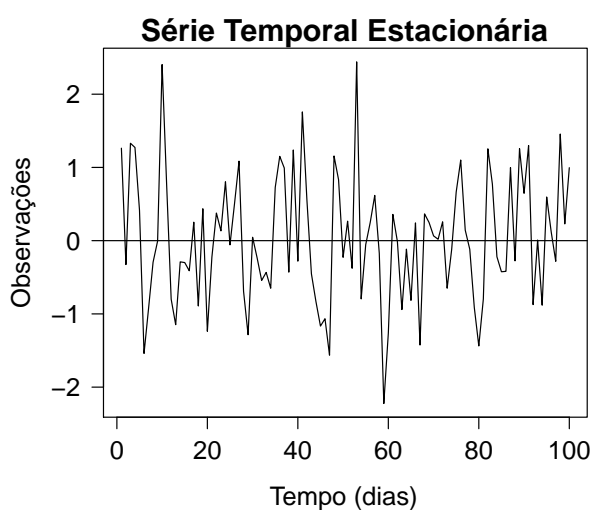


Figura 2.8: Exemplo de série temporal estacionária.

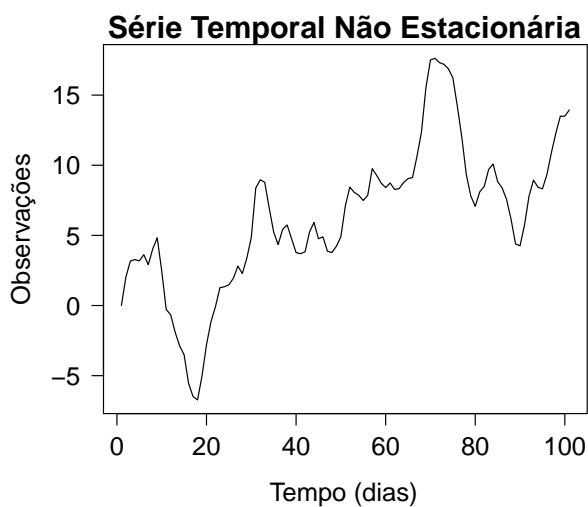


Figura 2.9: Exemplo de série temporal não-estacionária.

O método DFA é realizado com os seguintes passos:

- Inicialmente a série temporal original  $\{x(t)|t = 1, \dots, N\}$ , é integrada, produzindo:

$$X(k) = \sum_{i=1}^k [x(i) - \langle x \rangle] \quad (2.28)$$

onde  $\langle x \rangle = \frac{1}{N} \sum_{i=1}^N x_i$  é a média da série  $\{x(t)|t = 1, \dots, N\}$ ;

- A série integrada  $\{X(k)|k = 1, \dots, N\}$  é dividida em  $N_n = \text{int}(N/n)$  segmentos não sobrepostos de tamanho  $n$  e, em cada segmento,  $s = \{1, \dots, N_n\}$ ,  $X_{n,s}(k)$  representa a tendência local (ajuste linear ou polinomial) que é estimada e subtraída da série integrada  $\{X(k)|k = 1, \dots, N\}$ ;

- A variância sem tendência é calculada como:

$$F_{DFA}^2(n) = \frac{1}{nN_n} \sum_{s=1}^{N_n} \sum_{k=(s-1)n+1}^{n \times s} [X(k) - X_{n,s}(k)]^2 \quad (2.29)$$

- Repetimos o cálculo do passo anterior para diferentes tamanhos dos segmentos, obtendo a relação entre a função de flutuação  $F_{DFA}(n)$  e o tamanho do segmento  $n$ . Se a série original  $\{x(t)|t = 1, \dots, N\}$  apresentar correlações de longo alcance, então  $F_{DFA}(n)$  segue uma lei de potência.

$$F_{DFA}(n) \sim n^\alpha \quad (2.30)$$

Onde o expoente de escala  $\alpha$  pode ser obtido como o coeficiente angular (inclinação) da reta do gráfico log-log entre  $F_{DFA}(n)$  versus  $n$ . Para séries temporais não-correlacionadas, temos  $\alpha = 0.5$ , para séries persistentes  $\alpha > 0.5$  e para séries anti-persistentes  $\alpha < 0.5$  [72]. Algumas das inúmeras aplicações do método DFA incluem: climatologia [73], meteorologia [74], hidrologia [75], geofísica [76].

### Expoente DFA em Janela Móvel de Tempo (*Rolling Window*)

Este método possibilita analisar o comportamento do Expoente DFA em janelas móveis de tempo, onde podem ser utilizados tamanhos de janela alternativos [77]. Alguns estudos incluem janelas de tamanho 4 anos de janela de tempo (aproximadamente 1000 dias) [61–66], enquanto em outros estudos fixou-se o tamanho da janela em 2 anos (aproximadamente 480 dias) [67], 3 anos (aproximadamente 720 dias) [68] e 1 ano (aproximadamente 240 dias)[69; 70].

### 2.3.7 *Detrended Cross-Correlation Analysis (DCCA)*

O método *Detrended Cross-Correlation Analysis* (DCCA) (Análise de Correlação Cruzada Sem Tendência) foi introduzido recentemente por Podobnik & Stanley [78]. Este método representa uma generalização do *Detrended Fluctuation Analysis* (DFA) e serve para quantificar as correlações entre duas séries temporais registradas no mesmo período. Este método pode ser aplicado para detectar e quantificar as correlações de longo alcance entre duas séries temporais não-estacionárias. Trata-se de um método muito utilizado em várias áreas como climatologia [79], finanças [80], engenharia [81], sinais biológicos [82] e fisiologia [83]. A implementação do algoritmo DCCA é semelhante a do algoritmo DFA descrita na seção anterior (2.3.6). A diferença entre DCCA e DFA, é que, em vez da variância  $F_{DFA}^2(n)$ , calcula-se a covariância sem tendências para o método DCCA.

$$F_{DCCA}^2(n) = \frac{1}{nN_n} \sum_{s=1}^{N_n} \sum_{k=(s-1)n+1}^{n \times s} [X(k) - X_{n,s}(k)] \times [Y(k) - Y_{n,s}(k)] \quad (2.31)$$

em que  $X(k) = \sum_{i=1}^k [x_i - \langle x \rangle]$  e  $Y(k) = \sum_{i=1}^k [y_i - \langle y \rangle]$ , para  $k = \{1, \dots, N\}$ , são as séries integradas a partir de duas séries registradas simultaneamente  $\{x(t)|t = 1, \dots, N\}$  e  $\{y(t)|t = 1, \dots, N\}$ , e  $X_{n,s}(k)$  e  $Y_{n,s}(k)$  representam as tendências locais (ajuste linear ou polinomial) nos segmentos  $s = \{1, \dots, N_n\}$ , onde  $\langle x \rangle = \frac{1}{N} \sum_{i=1}^N x_i$  e  $\langle y \rangle = \frac{1}{N} \sum_{i=1}^N y_i$ .

Este cálculo se repete para os diferentes tamanhos de segmentos, proporcionando uma relação entre a função de flutuação  $F_{DCCA}(n)$  e o tamanho  $n$  do segmento. Se existirem correlações de longo alcance entre as duas séries originais  $\{x(t)|t = 1, \dots, N\}$  e  $\{y(t)|t = 1, \dots, N\}$ , a função de flutuação  $F_{DCCA}(n)$  segue uma lei de potência, ou seja,  $F_{DCCA}(n) \sim n^\lambda$ . O expoente de escala  $\lambda$  pode ser obtido como o coeficiente angular (inclinação) da reta, no gráfico log-log de  $F_{DCCA}(n)$  versus  $n$  e tem interpretação similar ao expoente  $\alpha$  do DFA. A existência das correlações entre duas séries simultâneas significa que cada série possui a memória dos seus valores anteriores e também a memória dos valores anteriores da outra série [78].

### 2.3.8 Coeficiente de Correlação-Cruzada ( $\rho_{DCCA}(n)$ )

Proposto por Zebende [84], o coeficiente de correlação-cruzada  $\rho_{DCCA}(n)$  tem como objetivo quantificar a correlação-cruzada entre duas séries temporais não-estacionárias  $\{x(t)|t = 1, 2, \dots, N\}$  e  $\{y(t)|t = 1, 2, \dots, N\}$ , com base nos métodos DFA e DCCA.

Define-se o coeficiente de correlação-cruzada  $\rho_{DCCA}(n)$  como a relação entre a função de covariância sem tendência  $F_{DCCA}^2(n)$  e a função de variância sem tendência  $F_{DFA}(n)$ :

$$\rho_{DCCA}(n) = \frac{F_{DCCA}^2(n)}{F_{DFA_x}(n) \times F_{DFA_y}(n)} \quad (2.32)$$

Portanto,  $\rho_{DCCA}(n)$  é um coeficiente adimensional com intervalo de variação entre  $-1 \leq \rho_{DCCA}(n) \leq 1$ . Para  $\rho_{DCCA}(n) = 0$ , interpreta-se que não há correlação-cruzada entre as séries analisadas. Para valores positivos, ou seja, se  $\rho_{DCCA}(n) > 0$ , significa que há correlação-cruzada, sendo esta perfeita somente quando  $\rho_{DCCA}(n) = 1$ . Para valores negativos, isto é, se  $\rho_{DCCA}(n) < 0$ , significa que existe anticorrelação-cruzada, onde esta é dita anticorrelação-cruzada perfeita somente quando  $\rho_{DCCA}(n) = -1$  [84].

## 2.4 Séries Temporais Multifractais

Existem muitas séries que não podem ser explicadas por um simples expoente de escala. Assim, é necessária uma hierarquia de expoentes de escala para melhor descrever as correlações temporais de uma série. Portanto, a análise multifractal é a metodologia a ser aplicada [45].

### 2.4.1 *Multifractal Detrended Fluctuation Analysis* (MF-DFA)

A Análise de Flutuação Sem Tendência (DFA), proposta por Peng *et al.* [72], foi um método bastante utilizado para a determinação das propriedades de escala de séries temporais monofractais não-estacionárias. Porém, muitas séries não exibem um simples comportamento monofractal, conseqüentemente, isso torna o DFA em um método impreciso na caracterização das propriedades de escala de séries temporais multifractais. Kantelhardt *et al.* [45] generalizaram o método DFA, denominado de Análise Multifractal de Flutuação Sem Tendência ou *Multifractal Detrended Fluctuation Analysis* (MF-DFA), o qual é uma poderosa ferramenta para investigar multifractalidade em séries temporais financeiras [85; 86]. Por esse método é possível detectar propriedades multifractais de inúmeras séries temporais. Algumas aplicações do método MF-DFA: séries temporais financeiras [67], séries temporais de velocidade do vento [87], modelagem de tráfego de redes [88], autômatos celulares [89].

O algoritmo do MF-DFA é dado por:

- (i) Inicialmente, a série temporal original  $\{x(t)|t = 1, \dots, N\}$  é integrada:

$$X(k) = \sum_{i=1}^k [x(i) - \langle x \rangle], \quad k = 1, \dots, N \quad (2.33)$$

em que  $\langle x \rangle = \frac{1}{N} \sum_{i=1}^N x_i$  representa a média da série  $\{x(t)|t = 1, \dots, N\}$ .

- (ii) A série integrada  $X(k)$  é dividida em  $N_n = \text{int}(N/n)$  segmentos não sobrepostos de tamanho  $n$ . Em cada segmento  $s = 1, \dots, N_n$ , a tendência local  $X_{n,s}(k)$  é estimada, através de ajuste linear ou polinomial, e subtraída da série  $X(k)$ .

- (iii) A variância sem tendência

$$F^2(n, s) = \frac{1}{n} \sum_{k=(s-1)n+1}^{s \times n} [X(k) - X_{n,s}(k)]^2 \quad (2.34)$$

é calculada para cada segmento. Finalmente, o valor médio para todos os segmentos é usado para obter a função de flutuação de ordem  $q$ :

$$F_q(n) = \left\{ \frac{1}{N_n} \sum_{s=1}^{N_n} [F^2(n, s)]^{q/2} \right\}^{1/q} \quad (2.35)$$

onde  $q$  pode assumir qualquer número real, exceto zero ( $q \neq 0$ ). Este cálculo se repete para diferentes tamanhos de segmentos  $n$ , fornecendo uma relação entre a função de flutuação  $F_q(n)$  e o tamanho do segmento  $n$ . Se a série original  $\{x(t)|t = 1, \dots, N\}$  possui correlações de longo alcance, a função de flutuação segue uma lei de potência:  $F_q(n) \sim n^{h(q)}$ . O expoente de escala  $h(q)$  pode ser obtido como o coeficiente angular (inclinação) da reta no gráfico log-log de  $F_q(n)$  versus  $n$ .

O valor de  $h(q)$  correspondente ao limite  $h(q)$  quando  $q \rightarrow 0$  não pode ser determinado diretamente usando a Equação 2.35 porque o expoente diverge. Assim, um procedimento de cálculo logarítmico da média deve ser utilizado:

$$F_0(n) = \left\{ \frac{1}{2N_n} \sum_{s=1}^{N_n} \ln [F^2(n, s)]^{1/2} \right\} \sim n^{h(0)} \quad (2.36)$$

Para o caso de séries temporais estacionárias  $h(2)$  é igual ao expoente clássico de Hurst, assim  $h(q)$  é chamado de Expoente de Hurst Generalizado (EHG) [85; 90]. Para valores negativos de  $q$ ,  $h(q)$  descreve comportamento de segmentos com pequenas flutuações, por outro lado, para valores positivos de  $q$ ,  $h(q)$  descreve o comportamento de segmentos com grandes flutuações. Para séries temporais monofractais  $h(q)$  é independente de  $q$ , enquanto que para séries temporais multifractais  $h(q)$  é uma função decrescente de  $q$  [90]. Expoentes generalizados  $h(q)$  estão relacionados com expoentes de Rényi  $\tau(q)$  que são definidos na análise multifractal clássica.

$$\tau(q) = qh(q) - 1 \quad (2.37)$$

No caso de séries temporais monofractais  $\tau(q)$  é uma função linear de  $q$ , enquanto para séries temporais multifractais  $\tau(q)$  é uma função não linear de  $q$  [30]. Adicionalmente, um processo multifractal pode também ser caracterizado de acordo com o espectro de singularidade  $f(\alpha)$ , o qual está relacionado com o expoente de Rényi, via transformada de Legendre, de acordo com as expressões [85]:

$$\alpha(q) = \frac{d\tau(q)}{dq} \quad (2.38)$$

$$f(\alpha) = q\alpha - \tau(q) \quad (2.39)$$



onde  $f(\alpha)$  denota a dimensão fractal do subconjunto da série temporal com valores caracterizados pelo expoente de Hölder  $\alpha$ . O espectro  $f(\alpha)$  de uma série temporal monofractal é caracterizada por um único ponto, enquanto para séries temporais multifractais  $f(\alpha)$  é uma função côncava para baixo que pode ser ajustada por um polinômio quadrático [30; 90; 91]. A amplitude  $W = \alpha_{máx} - \alpha_{mín}$  pode ser usada como medida do grau de multifractalidade da série temporal [90]. Ao ajustar a curva do espectro multifractal por um polinômio, alguns parâmetros podem ser usados para medir a complexidade das séries temporais por meio do seguinte ajuste polinomial [87]

$$f(\alpha) = A + B(\alpha - \alpha_0) + C(\alpha - \alpha_0)^2 + D(\alpha - \alpha_0)^3 + E(\alpha - \alpha_0)^4 \quad (2.40)$$

para calcular os parâmetros do espectro multifractal: posição do máximo  $\alpha_0$ ; largura do espectro  $W = \alpha_{máx} - \alpha_{mín}$ ; parâmetro de assimetria  $r = (\alpha_{máx} - \alpha_0)/(\alpha_0 - \alpha_{mín})$ , onde  $r = 1$  para formas simétricas,  $r > 1$  para assimetria à direita e  $r < 1$  para assimetria à esquerda.

Pequenos valores de  $\alpha_0$  sugerem que o processo é mais regular, aparentemente. A largura do espectro mede o grau de multifractalidade da série: quanto maior é a largura do espectro mais forte é a multifractalidade e mais rica é a estrutura da série.

O parâmetro  $r$  determina quais expoentes fractais são dominantes: expoentes fractais que descrevem invariância de escalas de pequenas flutuações são caracterizados por assimetria à direita ( $r > 1$ ), enquanto que expoentes fractais caracterizados por invariância de escalas de grandes flutuações apresentam assimetria à esquerda ( $r < 1$ ). Estes parâmetros podem ser usados para medir a complexidade das séries: um sinal com grande valor de  $\alpha_0$ , uma grande amplitude  $W$  do espectro multifractal e assimétrico à direita ( $r > 1$ ) pode ser considerado mais complexo que um sinal com características opostas [87; 92].

Há dois tipos de multifractalidade que podem ocorrer em séries temporais:

- (1) Multifractalidade causada pela função densidade de probabilidade dos valores da série;
- (2) Multifractalidade causada pelas diferentes correlações em subconjuntos com pequenas e grandes flutuações [45].

Para determinar qual tipo de multifractalidade está presente na série original, deve-se aplicar o método MF-DFA na série randomizada. Caso a multifractalidade seja do tipo (1), a randomização não afeta a função densidade de probabilidade dos valores da série temporal, resultando na permanência da função original  $h(q)$ . Para a multifractalidade do tipo (2), a randomização da série temporal destrói as correlações de

longo alcance e  $h(q) = 0.5$ . Se ambos os casos de multifractalidade estão presentes, a série temporal randomizada apresenta uma multifractalidade mais fraca que a série original [45].

## 2.5 Multifractal Detrended Cross-Correlation Analysis (MF-DCCA)

O método MF-DCCA é semelhante ao MF-DFA apresentado na Seção 2.4.1, com a diferença que, em vez da variância  $F^2(n, s)$ , calcula-se a covariância sem tendências:

$$F_{xy}(n, s) = \frac{1}{n} \sum_{k=(s-1)n+1}^{n \times s} |X(k) - X_{n,s}(k)| |Y(k) - Y_{n,s}(k)| \quad (2.41)$$

onde  $X(k) = \sum_{i=1}^k [x(i) - \langle x \rangle]$  e  $Y(k) = \sum_{i=1}^k [y(i) - \langle y \rangle]$ , com  $k = 1, \dots, N$ .  $X_{n,s}(k)$  e  $Y_{n,s}(k)$  denotam as tendências locais nos segmentos  $s = 1, \dots, N_n$  [93]. Onde  $\langle x \rangle = \frac{1}{N} \sum_{i=1}^N x_i$  e  $\langle y \rangle = \frac{1}{N} \sum_{i=1}^N y_i$ .

Calcula-se o valor médio para todos os segmentos e determina-se a função de flutuação de ordem  $q$ :

$$F_{xy}(q, n) = \frac{1}{N_n} \sum_{s=1}^{N_n} \left\{ [F(n, s)]^{q/2} \right\}^{1/q} \quad (2.42)$$

onde  $q$  pode assumir qualquer número real, exceto zero. Repete-se esse procedimento para diferentes tamanhos de segmentos  $n$ , determinando-se a relação entre a função de flutuação  $F_{xy}(n)$  e o tamanho do segmento  $n$ . Se existirem correlações de longo alcance entre as duas séries temporais originais, então  $F_{xy}(q, n)$  segue uma lei de potência:  $F_{xy}(q, n) \sim n^{h_{xy}(q)}$ . O expoente de escala  $h_{xy}(q)$  é calculado como coeficiente angular (inclinação) da reta no gráfico log-log de  $F_{xy}(q, n)$  versus  $n$ . Se as séries temporais  $\{x(i)|i = 1, \dots, N\}$  e  $\{y(i)|i = 1, \dots, N\}$  são iguais, o método MF-DCCA é idêntico ao método MF-DFA [45]. Obtendo-se os expoentes generalizados  $h_{xy}(q)$ , o expoente de Rényi  $\tau_{xy}(q)$  e o espectro de singularidade  $f_{xy}(\alpha)$  podem ser calculados usando as Equações 2.37, 2.38 e 2.39 [93]. O modelo MF-DCCA foi utilizado em estudos de: modelagem de processos biológicos [94], séries temporais financeiras [50], fluxo de tráfego [95], hidrologia [96] e sismologia [97].

## 2.6 Análise de Complexidade das Séries Temporais

### 2.6.1 *Sample Entropy* (*SampEn*)

O método *Sample Entropy* introduzido por Richman & Moorman [98] é definido como sendo o logaritmo natural da probabilidade condicional de que duas sequências similares (dentro de um nível de tolerância  $r$ ) para  $m$  pontos, permanecem similares para  $m + 1$  pontos.

Seja  $\{x(t)|t = 1, \dots, N\}$  uma série temporal de tamanho  $N$ . A seguir, o algoritmo *SampEn*( $m, r, n$ ) é descrito:

- (i) Construir  $N - m$  vetores de tamanho  $m$  onde:

$$x(i) = [x(i), x(i + 1), \dots, x(i + m - 1)]; i = 1, \dots, N - m + 1 \quad (2.43)$$

- (ii) Definir a distância  $d_m$  entre os vetores  $x(i)$  e  $x(j)$ , onde:

$$d_m[x(i), x(j)] = \max\{|x(i + k) - x(j + k)| : 0 \leq k \leq m - 1\} \quad (2.44)$$

- (iii) Para cada  $i = 1, \dots, N - m$ , calcula-se:

$$B_i^m(r) = \frac{B_i}{(N - m - 1)} \quad \text{e} \quad A_i^m(r) = \frac{A_i}{(N - m)} \quad (2.45)$$

onde  $B_i$  é o número dos vetores  $x(j)$  de tamanho  $m$  que são similares aos vetores  $x(i)$  dentro da distância  $r$ :  $d_m[x(i), x(j)] \leq r$  e  $A_i$  é o número dos vetores  $x(j)$  de tamanho  $(m + 1)$  que são similares aos vetores  $x(i)$  dentro da distância  $r$ .

Calcula-se:

$$B^m(r) = \frac{1}{N - m} \left( \sum_{i=1}^{N-m} B_i^m(r) \right) \quad (2.46)$$

e

$$A^m(r) = \frac{1}{N - m} \left( \sum_{i=1}^{N-m} A_i^m(r) \right) \quad (2.47)$$

onde  $B^m(r)$  é a probabilidade de dois vetores serem similares para  $m$  pontos e  $A^m(r)$  é a probabilidade de dois vetores serem similares para  $(m + 1)$  pontos.

(iv) Finalmente, calcula-se o índice *Sample Entropy*:

$$SampEn(m, r, N) = -Ln\left(\frac{A^m(r)}{B^m(r)}\right) \quad (2.48)$$

que representa a probabilidade condicional em que duas sequências que são similares para  $m$  pontos permaneçam similares para  $m + 1$  pontos (quando um dado consecutivo é adicionado) [99]. *Sample Entropy* é uma modificação do método *Approximate Entropy* (ApEn) [100]. Ambos os métodos servem para quantificar a complexidade em séries temporais não lineares. Estes métodos foram amplamente usados em análise de processos fisiológicos [101] e séries climáticas [102]. Maiores valores de *SampEn*, significam uma maior complexidade (menor regularidade) da série temporal [98]. O método *Sample Entropy* foi aplicado nos seguintes estudos: fisiologia [98; 99; 103], neurologia [104] e cardiologia [105].

### ***Sample Entropy* em Janela Móvel de Tempo (*Rolling Window*)**

Este método pode ser usado para avaliar o comportamento da estatística *SampEn* em janelas móveis de tempo. Podem ser usados tamanhos alternativos de janelas [77].

### 2.6.2 *Cross-Sample Entropy (Cross – SampEn)*

*Cross-Sample Entropy* é um método proposto para quantificar o grau de sincronia e correlação-cruzada entre duas séries temporais registradas simultaneamente [98; 106; 107]. Algumas aplicações incluem sinais fisiológicos [108] e séries financeiras [109].

Sejam  $u = (u(1), u(2), \dots, u(N))$  e  $v = (v(1), v(2), \dots, v(N))$  duas séries temporais de tamanho  $N$ . Sejam  $m$  e  $r$  os parâmetros, onde  $m$  é o comprimento do vetor a ser comparado e  $r$  o nível de tolerância. A aplicação do método consiste nos seguintes passos:

(i) Constrói-se uma sequência de vetores:

$$\begin{aligned} x_m(i) &= (u(i), u(i+1), \dots, u(i+m-1)), \quad 1 \leq i \leq N-m \\ y_m(j) &= (v(j), v(j+1), \dots, v(j+m-1)), \quad 1 \leq j \leq N-m \end{aligned} \quad (2.49)$$

a partir de  $u$  e  $v$  respectivamente.

(ii) Para cada  $i \leq N-m$ , calcula-se:

$$B_i^m(r)(v||u) = \frac{\#(1 \leq j \leq N-m), \text{ tal que, } d[x_m(i), y_m(j)] \leq r}{N-m} \quad (2.50)$$

onde  $d[x_m(i), y_m(j)] = \max\{|u(i+k) - v(j+k)| : 0 \leq k \leq m-1\}$ , representa a diferença máxima entre seus respectivos escalares.

(iii) Define-se:

$$B^m(r)(v||u) = \frac{\sum_{i=1}^{N-m} B_i^m(r)(v||u)}{N-m} \quad (2.51)$$

que corresponde ao valor médio de  $B_i^m(r)(v||u)$ .

(iv) De forma análoga, define-se:

$$A_i^m(r)(v||u) = \frac{\#(1 \leq j \leq N-m), \text{ tal que, } d[x_{m+1}(i), y_{m+1}(j)] \leq r}{N-m} \quad (2.52)$$

e

$$A^m(r)(v||u) = \frac{\sum_{i=1}^{N-m} A_i^m(r)(v||u)}{N-m} \quad (2.53)$$

que denota o valor médio de  $A_i^m(r)(v||u)$ .

(v) Por fim, calcula-se a estatística *Cross-Sample Entropy*:

$$Cross - SampEn(m, r, N) = -Ln \left\{ \frac{A^m(r)(v||u)}{B^m(r)(v||u)} \right\} \quad (2.54)$$

Aplicou-se o método *Cross-Sample Entropy* com parâmetros  $m = 2$  e  $r = 0.2$ . A seleção dos parâmetros para  $m$  e  $r$  depende do intervalo de confiança da estatística *SampEn* usada pela estatística *Cross - SampEn* [98].

Uma boa estatística *SampEn* deve obedecer o seguinte critério definido por Lake *et al.* [99]:

$$r \geq \max(\sigma_{PC}/PC, \sigma_{PC}/-\log(PC)PC) \quad (2.55)$$

onde  $PC$  é a probabilidade condicional de que duas sequências sejam similares para  $m + 1$  pontos dado que são similares para  $m$  pontos.

Este critério é igual a condição:

$$IC(r, 95\%) = -\log(PC) \mp 1.96(\sigma_{PC}/PC) \quad (2.56)$$

onde  $IC(Cross - SampEn, 95\%)$  significa um intervalo de confiança de 95% para a estatística *Cross - SampleEn*, a qual assume-se ser normalmente distribuída. Utilizou-se esse método bem como a escolha dos parâmetros  $m = 2$  e  $r = 0.2$ , a qual é adotada como aceitável. Tal escolha é similar a trabalhos anteriores [98; 106; 109].

## Materiais e Método

### 3.1 Conjunto de Dados

O conjunto de dados analisado consiste nas séries temporais de preço de fechamento diário de seis *commodities* agrícolas: Açúcar, Algodão, Arroz, Boi, Café e Trigo. Neste trabalho, tais *commodities* agrícolas são representadas simbolicamente por: ACR, ALG, ARR, BOI, CAF e TRG, respectivamente.

De acordo com o Banco Central do Brasil (BACEN), o Índice de Commodities Brasil (IC-Br) é composto por preços de matérias-primas brasileiras. Este equivale ao indicador internacional *Commodity Research Bureau* (CRB). O indicador IC-Br é constituído por três subgrupos de *commodities*: agrícolas, metálicas e energéticas. O IC-Br é divulgado mensalmente pelo Banco Central do Brasil. Trata-se de um indicador que objetiva identificar a parcela das variações de preços das *commodities* nos mercados internacionais relevante para explicar a dinâmica da inflação doméstica [110].

As *commodities* agrícolas que compõem o indicador IC-Br são: carne de boi, algodão, óleo de soja, trigo, açúcar, milho, café, arroz e carne de porco. As *commodities* metálicas são: alumínio, minério de ferro, cobre, estanho, zinco, chumbo e níquel. As *commodities* minerais são constituídas por petróleo *brent*, gás natural e carvão [110].

Neste trabalho analisou-se uma quantidade expressiva de *commodities* agrícolas, totalizando 6 das 9 *commodities* agrícolas do índice IC-Br. Isto representa 6 das 19 *commodities* que compõem o IC-Br.

O conjunto de dados analisados neste trabalho refere-se aos preços de fechamento diários registrados desde 3 de abril de 2006 até 7 de abril de 2014, totalizando 1960 dias. Todas as séries de preços analisadas são não-estacionárias. Os dados estão divididos em dois períodos: de 2006 a 2009 (conjunto de Dados A) e de 2010 a 2014 (conjunto de

Dados B). O conjunto Dados A (915 dias) refere-se ao período anterior a crise financeira mundial (2006-2009), enquanto o conjunto Dados B (1045 dias) faz referência ao período posterior a crise financeira mundial.

Na prática, utilizaram-se séries temporais padronizadas. Denotando o preço de fechamento das *commodities* no tempo  $t$  como  $\{x(t)|t = 1, \dots, N\}$ , o retorno diário logaritmo é definido por:

$$g(t) = \log(x(t+1)) - \log(x(t)), \quad \forall t = 1, \dots, N-1 \quad (3.1)$$

O retorno diário normalizado é definido como

$$R(t) = \frac{g(t) - \langle g(t) \rangle}{\sigma_g}, \quad \forall t = 1, \dots, N-1 \quad (3.2)$$

onde  $\sigma_g$  é o desvio-padrão da série  $g(t)$  e  $\langle g(t) \rangle$  é a média aritmética da série  $g(t)$ , ou seja,  $\langle g(t) \rangle = \frac{\sum_{t=1}^{N-1} g(t)}{N-1}$ , em que  $N-1$  é o comprimento da série  $g(t)$ .

## 3.2 Descrição dos dados

Foram analisadas as séries temporais do preço de fechamento diário de seis *commodities* agrícolas brasileiras: Açúcar, Algodão, Arroz, Boi, Café e Trigo. Os dados foram obtidos do *site* do Centro de Estudos Avançados em Economia Aplicada - ESALQ/USP a partir do endereço eletrônico <http://www.cepea.esalq.usp.br/>.

As *commodities* são negociadas da seguinte forma [111]:

- (i) Açúcar Cristal: negociado a partir do indicador ESALQ/BM&FBOVESPA pelo preço (R\$/sc) de fechamento diário da saca de 50 Kg. Especificação do produto: Açúcar Cristal - com mínimo de polarização de 99.7 graus, máximo de 0.10% de umidade, cor ICUMSA mais frequente 130-180, máximo de 0.07% de cinzas, ensacado em sacas novas. Moeda/Unidade de medida: R\$/saca de 50 kg. Região de referência: Estado de SP - dividido em 6 regiões: Piracicaba, Ribeirão Preto, Jaú, Assis, São José do Rio Preto e Araçatuba. Agentes consultados na pesquisa: unidades produtoras (usinas), compradores (consumidores) e grandes intermediários de vendas.
- (ii) Algodão: negociado a partir do indicador ESALQ/BM&FBOVESPA pelo preço de fechamento diário em centavos de reais ou em centavos de dólar por libra-  
peso (0,453597 Kg). Especificação do produto: algodão tipo 41, folha 4 - cor



estritamente abaixo da média (*strict low middling*) - (antigo tipo 6, fibra 30/32 mm, sem característica). Unidade de medida: libra-peso de pluma (0.453597 Kg); divulgados em centavos de real por libra-peso (ex: 125.95 centavos de R\$/lp ou R\$ 1.2595/lp). O Indicador também é calculado em dólar americano, pelo câmbio comercial, preço de venda cotado diariamente às 16h30. Região de referência: negócios feitos nas principais regiões produtoras e consumidoras de algodão do Brasil, entre elas: Santa Catarina, Rio de Janeiro, São Paulo, Rio Grande do Norte, Pernambuco, Alagoas, Paraíba, Ceará, Goiás, Bahia, Minas Gerais, Paraná, Tocantins, Mato Grosso e Mato Grosso do Sul. Base de Ponderação das regiões: média aritmética dos negócios coletados.

- (iii) Arroz em casca: negociado a partir do indicador ESALQ/BM&FBOVESPA pelo preço (R\$/sc) de fechamento diário da saca de 50 Kg. Especificação do produto: arroz em casca, com rendimento de 58% de grãos inteiros e 10% de grãos quebrados (renda de 68%), para produção de arroz branco Tipo 1, seguindo a classificação da Portaria 269/88 do Ministério da Agricultura. Unidade de medida: saca de 50 Kg (R\$/sc). Para a conversão desse indicador em dólar norte-americano (US\$), utiliza-se a Taxa de Câmbio comercial 16h30 de venda. Região de referência: Estado do Rio Grande do Sul, dividido em seis regiões administrativas segundo o Instituto Rio-Grandense do Arroz (IRGA): Fronteira Oeste, Zona Sul, Planície Costeira Interna, Planície Costeira Externa, Depressão Central e Campanha. Base de Ponderação das regiões: Participação das regiões no beneficiamento total do Estado, segundo os dados levantados pelo IRGA para safra 2008. Tendo todas as informações como preço à vista, líquido dos impostos e posto na indústria, calcula-se a média aritmética para cada uma das seis regiões. Estas médias regionais são ponderadas pela respectiva participação do beneficiamento de cada região no total do estado. Resulta, então, o valor em Reais do Indicador do arroz em casca do Rio Grande do Sul ESALQ/BM&FBOVESPA.
- (iv) Boi: o Indicador de Preços do Boi Gordo ESALQ/BM&FBOVESPA é uma média diária ponderada de preços à vista do boi gordo no estado de São Paulo. Região de referência (onde ocorrem os abates): Estado de SP, dividido em 4 regiões: Presidente Prudente, Araçatuba, Bauru e São José do Rio Preto. Unidade de medida: arroba do boi gordo (R\$/arroba). Moeda: O Indicador é calculado originalmente em Reais, mas é também divulgado em dólar americano, tratando-se de simples conversão do valor em Real. O câmbio considerado é o comercial, preço de venda, às 16h30. Agentes consultados na pesquisa: frigoríficos, pecuaristas,

escritórios de compra e venda de gado e leiloeiras. Especificação do produto: Bovino macho, castrado ou inteiro, comum e rastreado com 16 arrobas ou mais - animais inteiros são considerados desde 1 de agosto/2011.

- (v) Café Arábica: negociado a partir do indicador ESALQ/BM&FBOVESPA pelo preço (R\$/sc) de fechamento diário da saca de 60 Kg. Especificação do produto: café arábica tipo 6. Unidade de medida: saca de 60 Kg (R\$/sc). Região de referência: Cerrado e Sul de Minas Gerais, Mogiana Paulista (SP) e noroeste do Paraná. Agentes consultados na pesquisa: cooperativas, corretores, torrefadores e exportadores. Os preços referem-se a negociações no mercado físico de lotes (entre empresas). Os Indicadores em dólares americanos são calculados pela taxa de câmbio comercial para venda cotado às 16h30.
- (vi) Trigo: negociado a partir do indicador ESALQ/BM&FBOVESPA pelo preço (R\$/t) de fechamento diário em reais por tonelada. Especificação do produto: Trigo tipo 1, PH 78, com umidade máxima de 13%, teor de matérias estranhas impurezas máximo de 1%, máximo de grãos danificados por insetos de 0.5%, máximo de grãos danificados pelo calor, mofados e ardidos de 0.5% e até 1.5% de grãos quebrados, chochos e triguilho. Para o Paraná considera-se o trigo pão ou melhorador e, para o Rio Grande do Sul, o brando. Unidade de medida: reais por tonelada (R\$/t). Entrega: os valores se referem a negócios no mercado disponível (mercado de lotes, entre empresas), a retirar na empresa/armazém de venda na região de cada estado. Regiões de referência do Paraná: Oeste, Norte, Sudoeste e Ponta Grossa. Regiões de referência do Rio Grande do Sul: Passo Fundo, Ijuí, Santa Rosa e Santa Maria. Agentes consultados na pesquisa: vendedores (cooperativas/empresas), compradores (moinhos, indústrias e comerciantes) e intermediários (corretores).

As Tabelas abaixo exibem algumas estatísticas descritivas dos preços de fechamento das *commodities* analisadas. Os valores  $Q1$  e  $Q3$  referem-se aos quartis 1 e 3 respectivamente. Observa-se uma grande variação nos preços de um período para o outro nos casos das *commodities* Algodão e Café.

Tabela 3.1: Estatísticas descritivas para preços de fechamento diário da *commodity* Açúcar.

Período	Mínimo	Q1	Mediana	Média	Q3	Máximo	D. Padrão
2006-2009	23.00	26.80	34.80	36.50	45.50	63.50	10.47
2010-2014	40.00	48.60	54.80	56.60	64.50	77.00	10.35

Tabela 3.2: Estatísticas descritivas para preços de fechamento diário da *commodity* Algodão.

Período	Mínimo	Q1	Mediana	Média	Q3	Máximo	D. Padrão
2006-2009	110.00	117.00	124.00	125.00	131.00	145.00	8.84
2010-2014	135.00	160.00	181.00	199.00	215.00	399.00	57.43

Tabela 3.3: Estatísticas descritivas para preços de fechamento diário da *commodity* Arroz.

Período	Mínimo	Q1	Mediana	Média	Q3	Máximo	D. Padrão
2006-2009	16.10	21.10	25.10	25.70	28.90	36.40	5.22
2010-2014	18.80	25.60	28.20	29.20	34.00	39.40	5.33

Tabela 3.4: Estatísticas descritivas para preços de fechamento diário da *commodity* Boi.

Período	Mínimo	Q1	Mediana	Média	Q3	Máximo	D. Padrão
2006-2009	47.10	54.90	73.20	68.90	78.60	92.20	13.23
2010-2014	73.20	93.10	97.80	97.30	102.00	128.00	10.35

Tabela 3.5: Estatísticas descritivas para preços de fechamento diário da *commodity* Café.

Período	Mínimo	Q1	Mediana	Média	Q3	Máximo	D. Padrão
2006-2009	213.00	244.00	256.00	255.00	266.00	296.00	16.51
2010-2014	241.00	292.00	354.00	371.00	446.00	555.00	86.08

Tabela 3.6: Estatísticas descritivas para preços de fechamento diário da *commodity* Trigo.

Período	Mínimo	Q1	Mediana	Média	Q3	Máximo	D. Padrão
2006-2009	335.00	466.00	499.00	521.00	566.00	781.00	104.30
2010-2014	398.00	453.00	494.00	581.00	739.00	991.00	165.19

## 3.3 Metodologia

Para avaliar a dinâmica das séries temporais dos preços das *commodities* analisadas, dividiu-se as séries temporais em dois períodos distintos, com o objetivo de comparar os períodos pré-crise financeira mundial (2006-2009) e pós-crise financeira mundial (2010-2014), utilizando análise multifractal e análise da complexidade.

### 3.3.1 Análise Fractal

Os métodos *Detrended Cross-Correlation Analysis* (DCCA), Expoente de Hurst, Expoente de Hurst em Janela Móvel de Tempo e o Coeficiente de Correlação-Cruzada ( $\rho_{DCCA}(n)$ ) foram aplicados para analisar as séries temporais dos retornos das *commodities* em dois períodos: período pré-crise (2006-2009) e período pós-crise financeira mundial (2010-2014).

### 3.3.2 Análise Multifractal

Utilizou-se o método *Multifractal Detrended Fluctuation Analysis* (MF-DFA) para analisar as séries temporais dos retornos das *commodities*, tanto das séries originais quanto das séries aleatórias, onde comparou-se a largura dos espectros multifractais nos períodos pré-crise e pós-crise, para avaliar a dinâmica dos preços das *commodities*. A aplicação do método *Multifractal Detrended Cross-Correlation Analysis* (MF-DCCA), descreve a natureza das correlações-cruzadas entre os retornos das *commodities* analisadas, nos períodos anterior e posterior a crise financeira mundial.

### 3.3.3 Análise da Complexidade

O método Sample Entropy em Janela Móvel de Tempo (*Rolling Window*) foi utilizado para analisar o comportamento dos retornos das seis *commodities*. Verificou-se a assincronia entre os retornos das *commodities* aplicando-se o método *Cross-Sample Entropy*, utilizado para quantificar a assincronia de duas séries temporais de retornos.

### 3.3.4 Testes Estatísticos

Aplicaram-se os testes de Wilcoxon-Mann-Whitney, Shapiro-Wilk e t-Student com objetivo de validar os resultados obtidos.

#### Teste de Wilcoxon-Mann-Whitney – Duas Amostras Independentes

O teste Wilcoxon-Mann-Whitney para duas amostras independentes é indicado para avaliar a hipótese de que duas populações têm mesma distribuição. Isto é, a hipótese da nulidade ( $H_0$ ) é a de que as amostras provenham de populações tais que a probabilidade de que uma observação, tomada ao acaso de um grupo, seja maior do que outra observação, tomada ao acaso do outro grupo, é 0.5 (teste bilateral). Esse teste também é conhecido como teste  $U$  de Mann-Whitney, teste dos postos somados de Wilcoxon e Teste de Mann-Whitney.

Aplicou-se o teste Wilcoxon-Mann-Whitney para duas amostras com objetivo de comparar a estatística *Cross-SampEn*, o expoente  $\alpha$  do método DFA, o expoente  $\lambda$  do método DCCA e os parâmetros multifractais  $W$ ,  $r$  e  $\alpha_0$  entre duas amostras de dados. As duas amostras referem-se aos dois conjuntos de dados: Dados A (2006-2009) e Dados B (2010-2014). Estes conjuntos de dados referem-se aos períodos anterior e posterior a crise financeira mundial.

Sejam  $P_1$  e  $P_2$  duas populações onde não há informações sobre suas distribuições de probabilidade, porém, as variáveis envolvidas tenham escala ao menos ordinal. Isso quer dizer que pode-se abordar o caso de variáveis aleatórias qualitativas ordinais ou quantitativas. Considere duas amostras independentes dessas duas populações. O objetivo do teste é verificar se as distribuições são iguais em localização, ou seja, se uma população tende a ter valores maiores do que a outra, ou se elas têm a mesma mediana ( $Md$ ). Este é o teste de Wilcoxon-Mann-Whitney [112; 113].

O teste de Wilcoxon-Mann-Whitney é baseado nos postos dos valores obtidos das duas amostras combinadas. Ordena-se esses valores, em ordem crescente, independentemente de qual população cada valor é proveniente.

Seja  $X_1, X_2, \dots, X_m$  uma amostra aleatória da população  $P_1$  e  $Y_1, Y_2, \dots, Y_n$  uma amostra aleatória da população  $P_2$  de modo que tanto os  $X_i$ 's quanto os  $Y_i$ 's sejam independentes e identicamente distribuídos. Além disso, suponha que os  $X_i$ 's e os  $Y_i$ 's sejam mutuamente independentes. Considere a amostra de  $Y$  aquela com menor tamanho amostral, isto é,  $n \leq m$ .

Para aplicar o teste, supõe-se que  $F$  e  $G$  sejam as funções de distribuição correspondentes às populações  $P_1$  e  $P_2$ , respectivamente. A hipótese de nulidade é dada

por:

$$\begin{cases} H_0 : F(t) = G(t) \quad \forall t \\ H_1 : F(t) \neq G(t) \quad \forall t \end{cases}$$

para o teste bilateral. Para os testes unilaterais à esquerda ou à direita, basta fazer  $H_1 : F(t) < G(t) \quad \forall t$  ou  $H_1 : F(t) > G(t) \quad \forall t$ .

Em seguida, basta ordenar os valores, de ambas as amostras, em ordem crescente e colocar os postos associados. Seja  $S_m$  e  $S_n$  as somas dos postos relacionados às amostras de  $X$  e  $Y$  respectivamente. Após isso, calcula-se os valores de  $U_m$  e  $U_n$  da seguinte forma:

$$U_m = S_m - \frac{1}{2}m(m+1) \quad (3.3)$$

e

$$U_n = S_n - \frac{1}{2}n(n+1) \quad (3.4)$$

Como  $S_m$  e  $S_n$  é igual a soma de todos os postos de ambas as amostras, então:

$$S_m + S_n = \frac{1}{2}(m+n)(m+n+1) \quad (3.5)$$

em que  $U_m$  e  $U_n$  estão relacionados por

$$U_m = mn - U_n \quad (3.6)$$

### Teste de Wilcoxon-Mann-Whitney – Uma Amostra

O teste de Wilcoxon-Mann-Whitney para uma amostra foi utilizado nos métodos de Janelas Móveis de Tempo (*Rolling Window*), para testar se as medianas dos expoentes  $H$  e  $\alpha$  dos métodos de Hurst e DFA, respectivamente, são estatisticamente diferentes de 0.5.

Considere uma população  $P$  em que se retira uma amostra aleatória  $X_1, X_2, \dots, X_n$ . O objetivo do teste é estudar o comportamento da mediana populacional com relação a um valor  $\theta_0$  especificado, sob  $H_0$  verdadeira.

O teste de Wilcoxon-Mann-Whitney (para uma amostra) é baseado nos postos (*ranks*) dos valores obtidos. Postos são as posições, representados por números, que os valores ocupam quando colocados em ordem crescente.

As hipóteses do teste são as seguintes:

$$\begin{cases} H_0 : \theta = \theta_0 \\ H_1 : \theta \neq \theta_0 \end{cases}$$

para o teste bilateral. Para os testes unilaterais à esquerda ou à direita, basta fazer  $H_1 : \theta < \theta_0$  ou  $H_1 : \theta > \theta_0$ , respectivamente.

Inicialmente deve-se subtrair  $\theta_0$  de cada valor  $x_i$ ,  $i = 1, \dots, n$ , da amostra. Obtendo assim um novo conjunto de dados  $Z_1, Z_2, \dots, Z_n$ , onde  $Z_i = X_i - \theta_0$ .

Coloca-se em ordem crescente o novo conjunto de dados  $\{|Z_1|, |Z_2|, \dots, |Z_n|\}$  e associa-se a cada valor  $Z_i$  o posto correspondente. Em seguida, definem-se variáveis indicadoras  $\psi_i$ ,  $i = 1, \dots, n$ , dadas por

$$\psi_i = \begin{cases} 1, & \text{se } Z_i > 0 \\ 0, & \text{se } Z_i < 0 \end{cases}$$

Obtém-se os  $n$  produtos  $R_1\psi_1, \dots, R_n\psi_n$ . Cada produto  $R_i\psi_i$  é chamado posto positivo de  $Z_i$ . Se  $Z_i > 0$ ,  $R_i\psi_i$  é igual ao posto  $R_i$  correspondente. Por outro lado, se  $Z_i < 0$ ,  $R_i\psi_i$  é igual a zero.

A estatística  $T^+$  é definida como a soma dos postos com sinal positivo, ou seja

$$T^+ = \sum_{i=1}^n R_i\psi_i \quad (3.7)$$

### Teste de Shapiro-Wilk

O teste Shapiro-Wilk foi proposto em 1965. É um teste baseado na estatística  $W_{cal}$ , dada por [114–117]:

$$W_{cal} = \frac{b^2}{\sum_{i=1}^n (x_{(i)} - \langle x \rangle)^2} \quad (3.8)$$

em que  $x_{(i)}$ 's são os valores da amostra ordenados, ou seja,  $x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)}$ , onde  $x_{(1)}$  representa o valor mínimo e  $x_{(n)}$  o valor máximo. A média dos valores é representada por  $\langle x \rangle$ . As variáveis aleatórias  $X_{(1)}, X_{(2)}, \dots, X_{(n)}$  são denominados estatísticas de ordem. A constante  $b$  é determinada da seguinte forma:

$$b = \begin{cases} \sum_{i=1}^{n/2} a_{n-i+1} \times (x_{(n-i+1)} - x_{(i)}) & , \text{ se } n \text{ é par} \\ \sum_{i=1}^{(n+1)/2} a_{n-i+1} \times (x_{(n-i+1)} - x_{(i)}) & , \text{ se } n \text{ é ímpar} \end{cases}$$

em que os  $a_{n-i+1}$ 's referem-se a constantes geradas pelas médias, variâncias e covariâncias das estatísticas de ordem de uma amostra de tamanho  $n$  de uma distribuição normal.

A formulação do teste Shapiro-Wilk é dada a seguir:

(1) Enunciar hipóteses:

$$\begin{cases} H_0 : \text{A amostra provém de uma população Normal} \\ H_1 : \text{A amostra não provém de uma população Normal} \end{cases}$$

(2) Fixar o nível de significância do teste. Geralmente  $\alpha = 0.05$ ;

(3) Calcular a estatística do teste: Ordenar as  $n$  observações da amostra:  $x_{(1)}, x_{(2)}, \dots, x_{(n)}$ ;

Calcular  $\sum_{i=1}^n (x_i - \langle x \rangle)$ ,  $b$  e  $W_{cal}$ ;

(4) Tomar a decisão: rejeitar  $H_0$  ao nível de significância  $\alpha$  se  $W_{cal} < W_{tab}$ . Aceitar, caso contrário. Onde  $W_{tab}$  é o valor crítico encontrado na tabela de Shapiro-Wilk.

### Teste t-Student

Uma variável aleatória contínua  $X$  segue distribuição  $t$  Student com  $v$  graus de liberdade se sua função densidade de probabilidade é dada por [118]:

$$f(x) = \frac{\Gamma\left(\frac{v+1}{2}\right)}{\sqrt{v\pi}\Gamma\left(\frac{v}{2}\right)} \left(1 + \frac{x^2}{v}\right)^{-\left(\frac{v+1}{2}\right)}, \quad x \in (-\infty, \infty). \quad (3.9)$$

A formulação das hipóteses para o teste do coeficiente angular  $\beta_1$  da reta de regressão é dada por:

$$\begin{cases} H_0 : \beta_1 = b \\ H_1 : \beta_1 \neq b \end{cases}$$

A estatística do teste é

$$t_{cal} = \frac{\hat{\beta}_1}{s(\hat{\beta}_1)} \sim t_{(n-2)} \quad (\text{sob } H_0)$$

onde

$$s^2(\hat{\beta}_1) = \frac{QME}{S_{xx}} \quad \text{e} \quad QME = \frac{SQ_E}{n-2}$$

O valor  $t_{cal}$  deve ser comparado com a distribuição t-Student com  $n - 2$  graus de liberdade para um nível de significância  $\alpha$  fixado, em que  $n$  é o tamanho da amostra. Utilizou-se o teste t-Student para comparar as tendências lineares dos coeficientes de correlações-cruzadas  $\rho_{DCCA}(n)$  nos períodos anterior (2006-2009) e posterior (2010-2014) a crise financeira mundial.



### 3.3.5 Coeficiente de Assimetria

O coeficiente de assimetria permite identificar as distribuições assimétricas. Valores negativos indicam que a cauda esquerda da função densidade de probabilidade é maior que a da direita. Por outro lado, valores positivos indicam que a cauda direita é maior que a esquerda. Valor próximos de zero indicam distribuição próximo à simetria [119]. O coeficiente de assimetria é dado por:

$$\mu_3 = \frac{1}{n} \sum_{i=1}^n \left[ \frac{x_i - \langle x \rangle}{s} \right]^3 \quad (3.10)$$

onde  $\langle x \rangle$  é a média e  $s$  é o desvio-padrão dos valores.

### 3.3.6 Coeficiente de Curtose

O coeficiente de Curtose é uma medida de dispersão. Ela caracteriza o “achatamento” da função densidade de probabilidade. Valores positivos indicam um pico mais agudo, corpo mais fino e cauda mais pesada que a distribuição normal. Valores negativos indicam pico mais tênue, corpo mais grosso e uma cauda mais fina que a distribuição normal. Assim, curtose positiva indica cauda mais alongada e curtose negativa indica cauda mais curta [120]. O coeficiente de curtose é dado por:

$$\mu_4 = \frac{1}{n} \sum_{i=1}^n \left[ \frac{x_i - \langle x \rangle}{s} \right]^4 - 3 \quad (3.11)$$

onde  $\langle x \rangle$  é a média e  $s$  é o desvio-padrão dos valores.



# Capítulo 4

## Resultados

### 4.1 Preço de fechamento das *commodities*

O conjunto de dados refere-se aos preços de fechamento diários registrados desde 3 de abril de 2006 até 7 de abril de 2014, totalizando 1960 dias. Os dados estão divididos em dois períodos: de 2006 a 2009 (conjunto de Dados A) e de 2010 a 2014 (conjunto de Dados B). O conjunto Dados A (915 dias) refere-se ao período anterior a crise financeira mundial (2006-2009), enquanto o conjunto Dados B (1045 dias) faz referência ao período posterior a crise financeira mundial.

A Figura 4.1 descreve o preço de fechamento das *commodities* (2006-2014). A linha vertical (em cinza) indica a divisão dos períodos pré-crise (2006-2009) e pós-crise (2010-2014) financeira mundial. De acordo com a Figura 4.1, todas as séries de preços analisadas são não-estacionárias, pois não seguem uma única tendência durante todo o período.

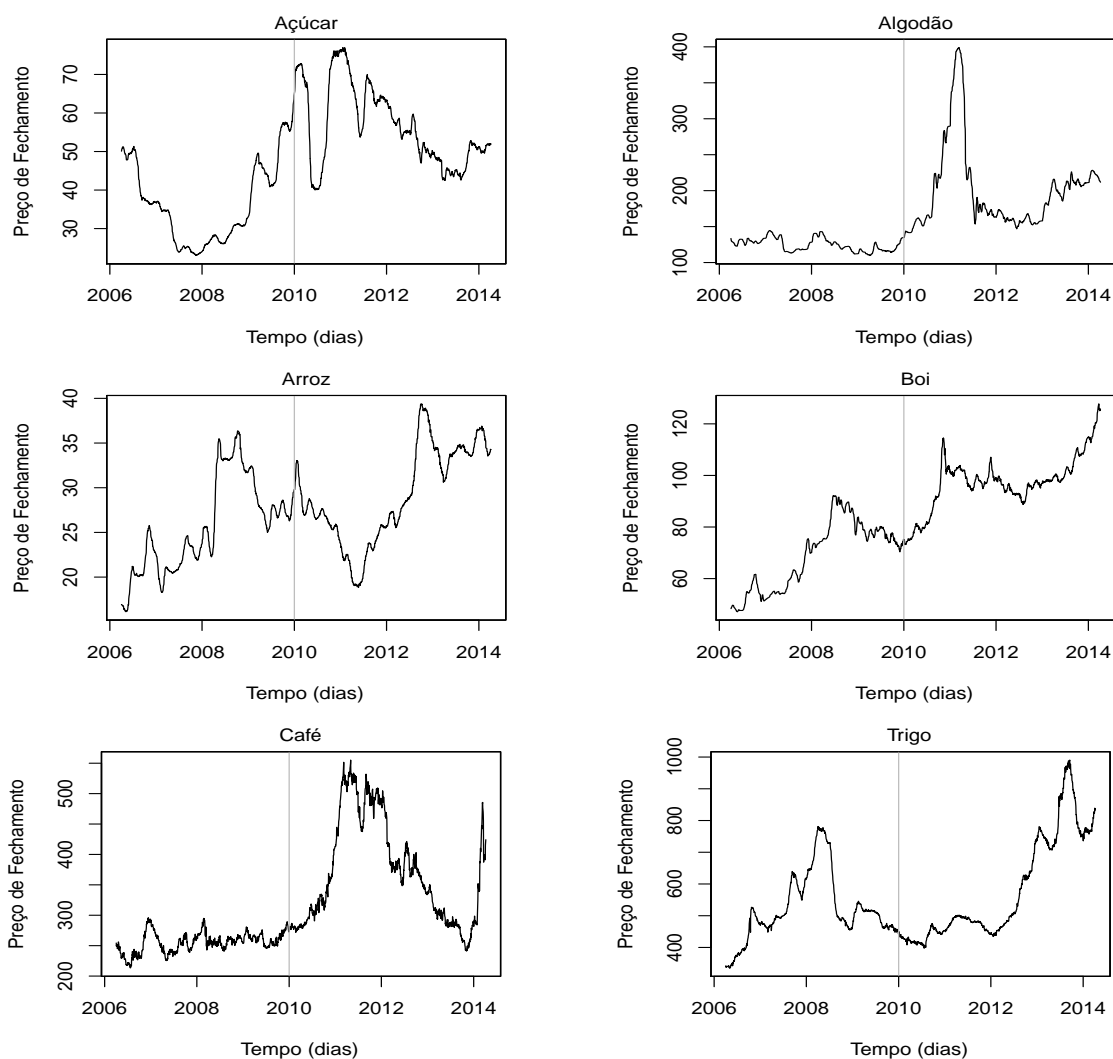


Figura 4.1: Séries temporais dos preços de fechamento diário das *commodities*: Açúcar 4.1(a), Algodão 4.1(b), Arroz 4.1(c), Boi 4.1(d), Café 4.1(e) e Trigo 4.1(f), respectivamente.

A partir da Figura 4.1, observa-se que as *commodities* Café e Algodão apresentam preço mais estável no período compreendido entre 2006 e 2009. Entre 2010 e 2014 os preços sobem até atingir o valor máximo das respectivas séries.

Em relação ao Açúcar, no início da série os preços caem. Entre o início de 2008

e final de 2009 o preço exibe tendência crescente. No início de 2010 há uma queda acentuada e em seguida um novo aumento. Após esse aumento, a série apresenta comportamento de queda para o restante da série.

Arroz e Trigo apresentam tendência crescente nos preços até 2008. Após atingirem os maiores valores no período 2006-2008, há tendência de queda em meados de 2008. Desde o início de 2011 até o final de 2014, ocorre comportamento crescente nos preços de ambas as *commodities*.

Por fim, os preços da *commodity* Boi aumentam até aproximadamente metade de 2008. Após isso, os preços caem. A partir de 2010, há tendência crescente.

## 4.2 Retornos das *commodities*

A Figura 4.2 exibe as séries temporais dos retornos das *commodities* (2006-2014). A linha vertical (em cinza) indica a divisão dos períodos pré-crise (2006-2009) e pós-crise (2010-2014) financeira mundial. As séries de retornos são estacionárias, pois os valores flutuam em torno da média zero.

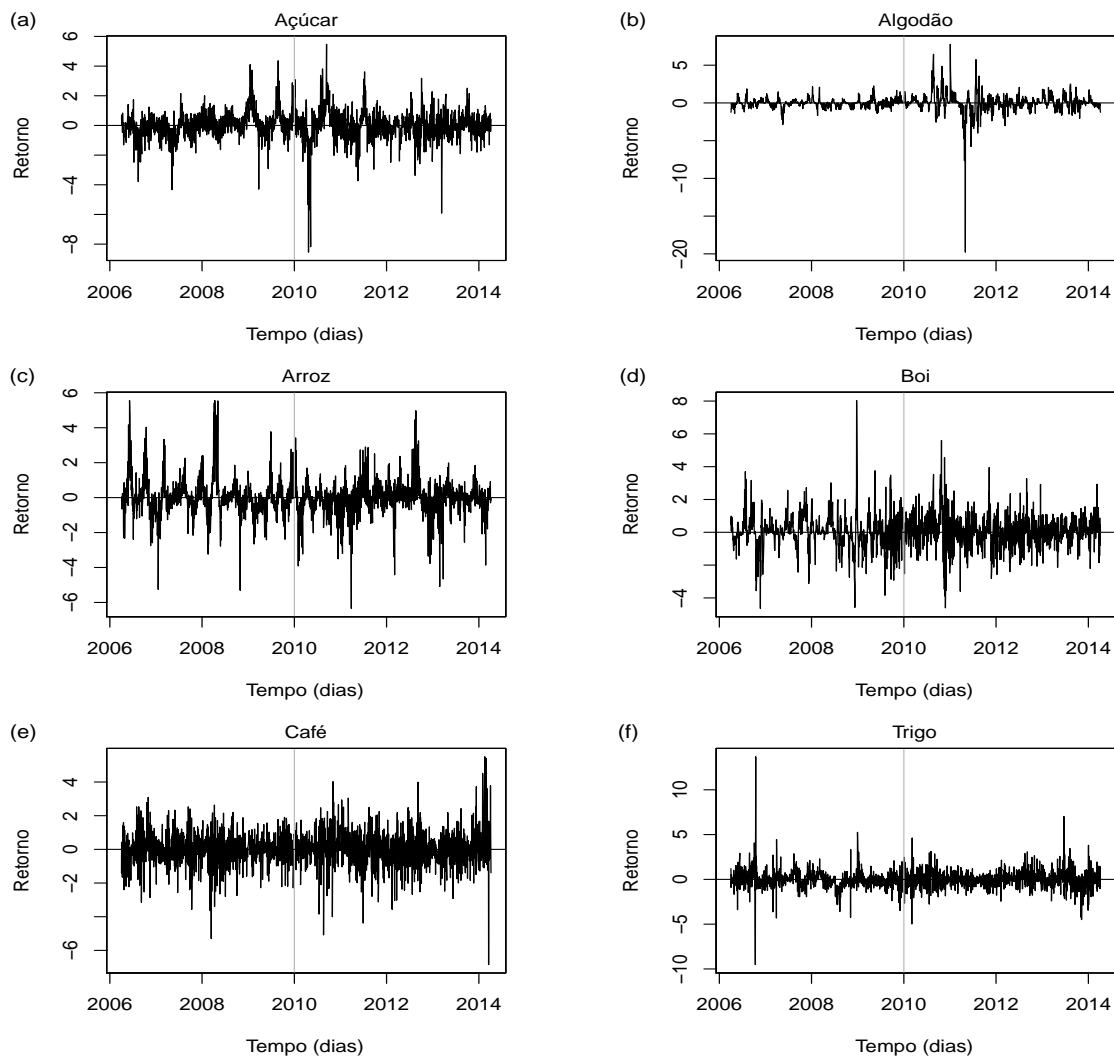


Figura 4.2: Séries temporais do retorno diário das *commodities*: Açúcar 4.2(a), Algodão 4.2(b), Arroz 4.2(c), Boi 4.2(d), Café 4.2(e) e Trigo 4.2(f), respectivamente.

A Figura 4.3 apresenta as densidades dos retornos (linha azul) das *commodities* para todo o período analisado (2006-2004). A curva, em vermelho, representa a densidade normal ajustada para os dados dos retornos. Observa-se que a curva normal não se ajusta bem aos retornos.

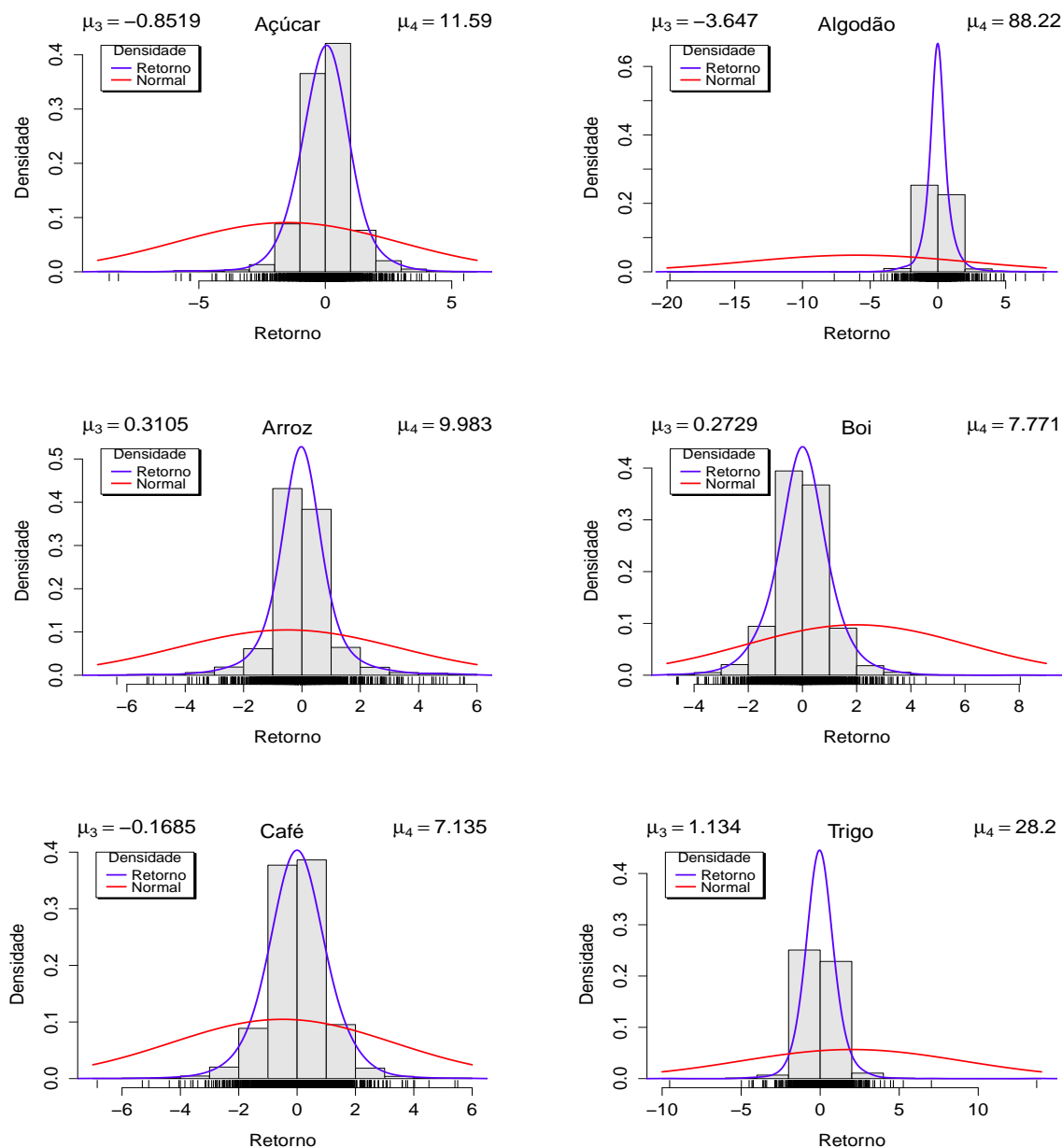


Figura 4.3: Distribuições dos retornos das *commodities* Açúcar 4.3(a), Algodão 4.3(b), Arroz 4.3(c), Boi 4.3(d), Café 4.3(e) e Trigo 4.3(f), respectivamente.

A Tabela 4.1 apresenta as estatísticas descritivas dos retornos das *commodities*, para todo o período analisado (2006-2014):

Tabela 4.1: Estatísticas descritivas dos retornos das *commodities*.

	ACR	ALG	ARR	BOI	CAF	TRG
Média	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
Mediana	0.04415	-0.01504	-0.06040	-0.02146	0.00637	-0.02999
Máximo	5.47469	7.78139	5.56167	8.03859	5.50391	13.69781
Mínimo	-8.54349	-19.78365	-6.34503	-4.64215	-6.84990	-9.52551
Desv. Padrão	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000
Assimetria ( $\mu_3$ )	-0.85191	-3.64735	0.31046	0.27291	-0.16851	1.13399
Curtose ( $\mu_4$ )	11.58965	88.21771	9.98273	7.77099	7.13542	28.19551
P-valor Shapiro	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000

Observa-se que todas as séries de retornos apresentam coeficientes de curtose positivos ( $\mu_4 > 0$ ), indicando um pico mais agudo, corpo mais fino e cauda mais pesada que a distribuição normal. O teste de Shapiro-Wilk afirma que os retornos são seguem distribuição normal, ao nível de significância de 5%.



### 4.3 Teste de Correlação-Cruzada

Utilizou-se o teste de correlações-cruzadas proposto por Podobnik *et al.* [48]. O objetivo deste teste é qualificar as correlações-cruzadas em existentes ou não existentes. O teste foi realizado para cada par ( $C_{6,2} = 15$  pares) entre as *commodities* analisadas. Este teste é análogo ao teste de Ljung-Box [49], sendo amplamente usado para testar correlações-cruzadas [50–55].

Podobnik *et al.* [48] propuseram que a estatística de correlação-cruzada  $Q_{cc}(p)$  segue aproximadamente uma distribuição Qui-quadrado com  $p$  graus de liberdade, ou seja,  $Q_{cc}(p) \approx \chi^2(p)$ . Ela pode ser usada para testar a hipótese de nulidade ( $H_0$ ) de que nenhum dos  $p$  primeiros coeficientes de correlações-cruzadas é diferente de zero [48].

As hipóteses a serem testadas são:

$$\begin{cases} H_0 : Q_{cc}(p) = 0 & \text{(Não há correlação-cruzada entre as séries temporais } x(t) \text{ e } y(t)) \\ H_1 : Q_{cc}(p) \neq 0 & \text{(Há correlação-cruzada entre as séries temporais } x(t) \text{ e } y(t)) \end{cases}$$

Podem-se observar nas Figuras 4.4, 4.5 e 4.6, os gráficos log-log das estatísticas de correlações-cruzadas  $Q_{cc}(p)$  entre as séries temporais dos preços de fechamento das *commodities* versus os graus de liberdade  $p$ . Os graus de liberdade variam entre 1 e 1000, para os períodos 2006-2014 e 2010-2014. Para a série de preços no período 2006-2009 os graus de liberdade variam de 1 até 900. Note que a variação dos graus de liberdade é limitada pelo tamanho da série, ou seja,  $p \leq N$ . Para uma melhor comparação, apresentam-se os valores críticos para a distribuição Qui-quadrado (pontos em vermelho) ao nível de significância de 5%. Para uma ampla variedade dos valores de  $p$ , todas as estatísticas do teste pertencem a região de rejeição de  $H_0$ , ou seja,  $Q_{cc}(p) > \chi_{0.95}^2(p)$ . Portanto, pode-se rejeitar a hipótese de nulidade ( $H_0$ ) de que não existe correlação-cruzada entre as séries temporais do preço de fechamento das *commodities* analisadas, para os três períodos considerados.

A Figura 4.4 descreve o gráfico log-log entre a estatística do teste  $Q_{cc}(p)$  e os graus de liberdade  $p$  no período 2006-2014 ( $N = 1960$  dias), em que  $p = 1, \dots, 1000$ .

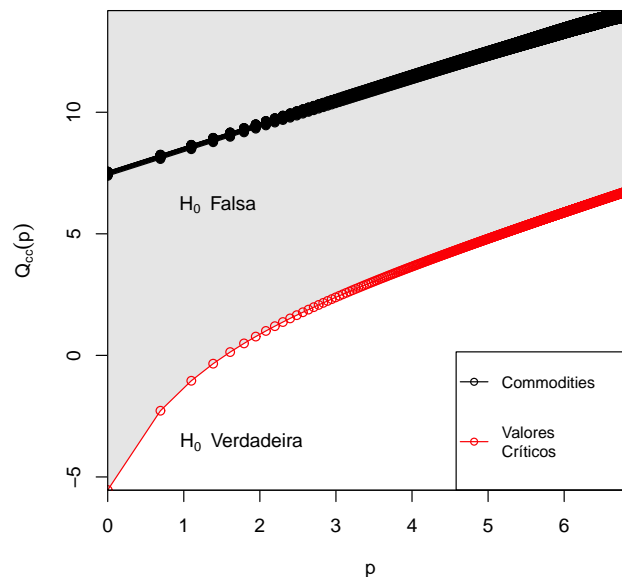


Figura 4.4: Gráfico log-log da estatística do teste  $Q_{cc}(p)$  versus os graus de liberdade  $p$  para o período 2006-2014 ( $N = 1960$  dias), em que  $p = 1, \dots, 1000$ .

Na Figura 4.5 pode-se observar o gráfico log-log entre a estatística do teste  $Q_{cc}(p)$  e os graus de liberdade  $p$  durante o período pré-crise 2006-2009 ( $N = 915$  dias), onde  $p = 1, \dots, 900$ .

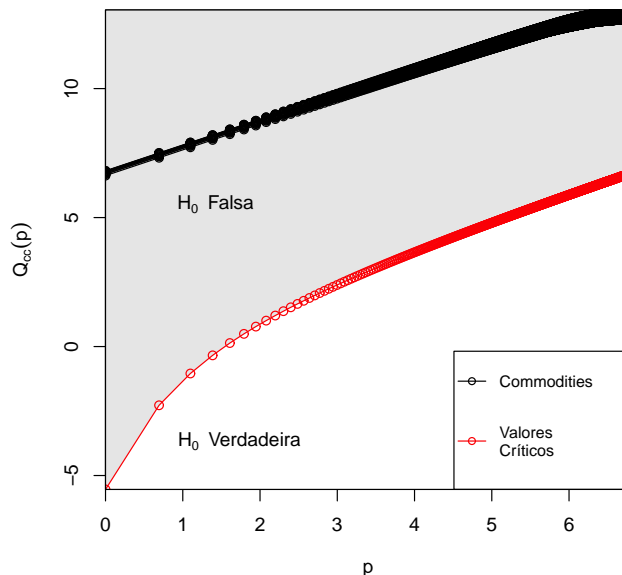


Figura 4.5: Gráfico log-log da estatística do teste  $Q_{cc}(p)$  versus os graus de liberdade  $p$  para o período pré-crise 2006-2009 ( $N = 915$  dias), onde  $p = 1, \dots, 900$ .

Pode-se observar na Figura 4.6 o gráfico log-log entre a estatística do teste  $Q_{cc}(p)$  e os graus de liberdade  $p$  durante o período pós-crise 2010-2014 ( $N = 1045$  dias), em que  $p = 1, \dots, 1000$ .

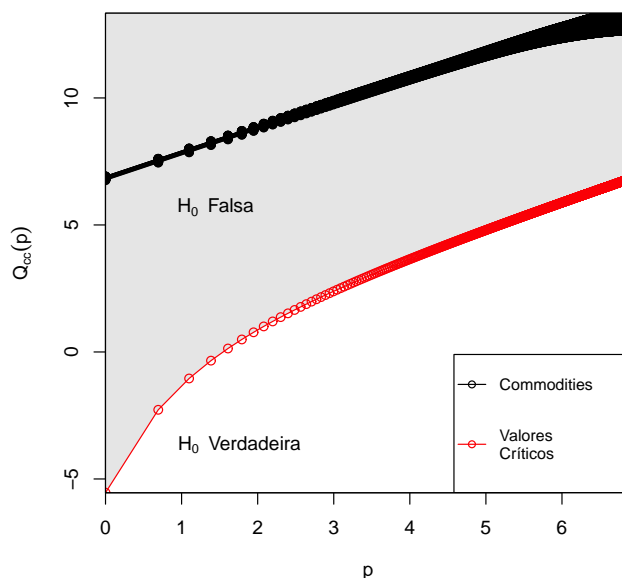


Figura 4.6: Gráfico log-log da estatística do teste  $Q_{cc}(p)$  versus os graus de liberdade  $p$  para o período pós-crise 2010-2014 ( $N = 1045$  dias), em que  $p = 1, \dots, 1000$ .

## 4.4 *Cross-Sample Entropy* entre as *commodities*

Nesta seção apresenta-se a análise da estatística *Cross-SampEn*. Aplicou-se o método *Cross-Sample Entropy* entre as séries temporais dos retornos padronizados das *commodities* analisadas.

Para cada par de séries temporais dos retornos padronizados nos conjuntos de Dados A (2006-2009) e Dados B (2010-2014), aplicou-se o método *Cross-Sample Entropy*, com os parâmetros  $m = 2$  e  $r = 0.2$ . A Figura 4.7 descreve o comportamento da estatística *Cross-SampEn* entre as *commodities*.

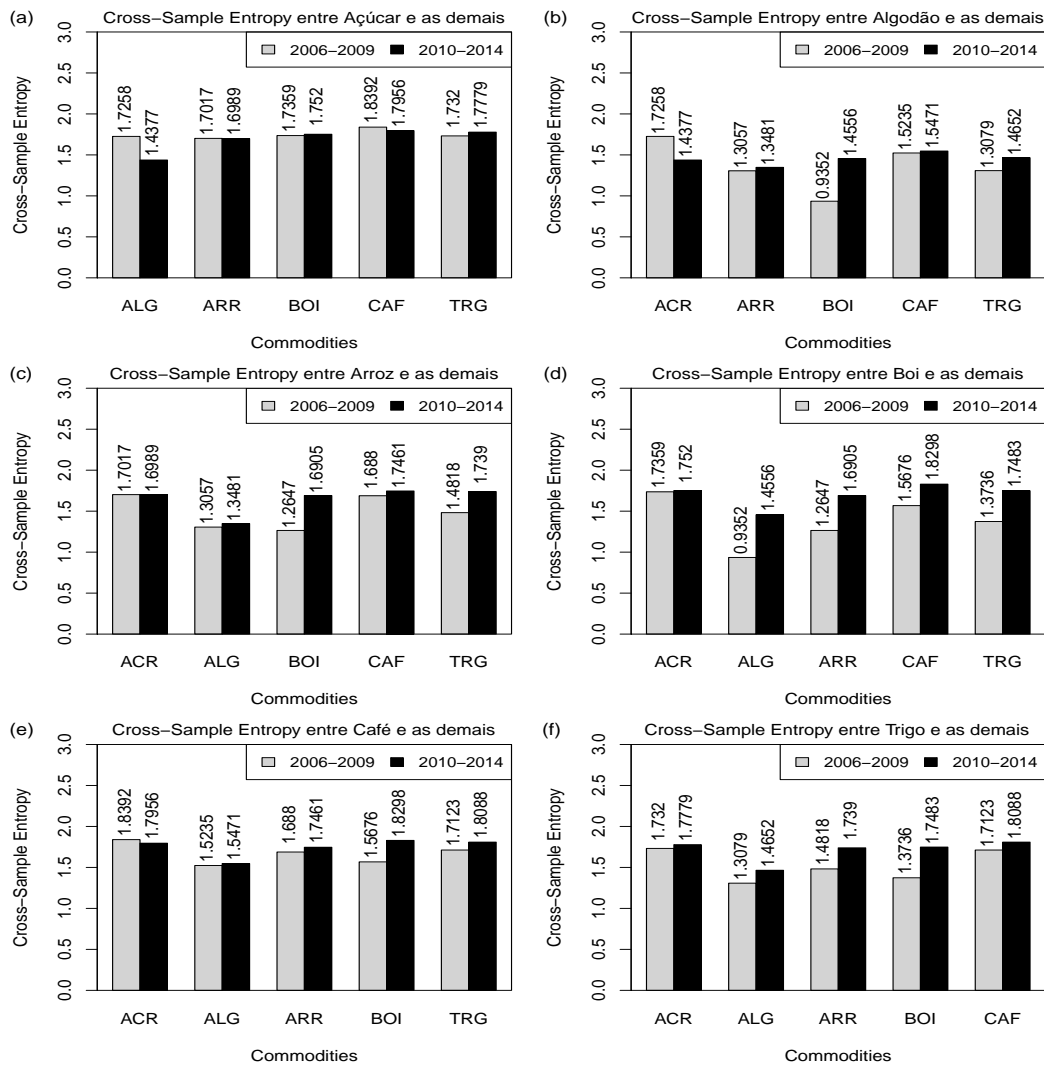


Figura 4.7: *Cross-Sample Entropy* entre cada *commodity* e as demais, para Açúcar 4.7(a), Algodão 4.7(b), Arroz 4.7(c), Boi 4.7(d), Café 4.7(e) e Trigo 4.7(f), respectivamente.

As Tabelas 4.2 e 4.3 mostram a estatística *Cross-SampEn*, nos períodos (2006-2009) e (2010-2014), como medida de assincronia e correlação cruzada entre as séries temporais das *commodities* Açúcar, Algodão, Arroz, Boi, Café e Trigo, respectivamente. Os valores em negrito referem-se a estatística *SampEn*.

Tabela 4.2: Cross-Sample Entropy entre todos os retornos das séries temporais das *commodities* analisadas (2006-2009).

Cross-SampEn	Açúcar	Algodão	Arroz	Boi	Café	Trigo
Açúcar	<b>1.6050</b>	1.7258	1.7017	1.7359	1.8392	1.7320
Algodão	1.7258	<b>0.8717</b>	1.3057	0.9352	1.5235	1.3079
Arroz	1.7017	1.3057	<b>1.2960</b>	1.2647	1.6880	1.4818
Boi	1.7359	0.9352	1.2647	<b>0.9532</b>	1.5676	1.3736
Café	1.8392	1.5235	1.6880	1.5676	<b>1.6240</b>	1.7123
Trigo	1.7320	1.3079	1.4818	1.3736	1.7123	<b>1.4550</b>

Tabela 4.3: Cross-Sample Entropy entre todos os retornos das séries temporais das *commodities* analisadas (2010-2014).

Cross-SampEn	Açúcar	Algodão	Arroz	Boi	Café	Trigo
Açúcar	<b>1.5610</b>	1.4377	1.6989	1.7520	1.7956	1.7779
Algodão	1.4377	<b>0.8355</b>	1.3481	1.4556	1.5471	1.4652
Arroz	1.6989	1.3481	<b>1.5240</b>	1.6905	1.7461	1.7390
Boi	1.7520	1.4556	1.6905	<b>1.5880</b>	1.8298	1.7483
Café	1.7956	1.5471	1.7461	1.8298	<b>1.6450</b>	1.8088
Trigo	1.7779	1.4652	1.7390	1.7483	1.8088	<b>1.5750</b>

Aplicou-se o teste de Wilcoxon-Mann-Whitney para duas amostras independentes: Dados A (2006-2009) e Dados B (2010-2014), respectivamente. Esse teste é indicado para avaliar a hipótese de que duas populações têm mesma distribuição. Isto é, a hipótese da nulidade é a de que as amostras provenham de populações tais que a probabilidade de que uma observação, tomada ao acaso de um grupo (Dados A), seja maior do que outra observação, tomada ao acaso do outro grupo (Dados B), é 0.5 (teste bilateral). Esse teste também é conhecido como teste  $U$  de Mann-Whitney, teste dos postos somados de Wilcoxon e Teste de Mann-Whitney. O teste Wilcoxon-Mann-Whitney resultou em um p-valor de 0.06592. Isto significa que não há diferença significativa ao nível de significância de 5%, entre os dois grupos avaliados, para a estatística *Cross-SampEn*.

O menor valor da estatística *Cross-SampEn* entre todos ocorre entre Algodão e Boi ( $Cross - SampEn(ALG, BOI) = 0.9352$ ), no conjunto de Dados A, ou seja, a maior sincronia ocorre entre Algodão e Boi.

Por outro lado, os maiores valores da estatística *Cross-SampEn* dentre todos são encontrados entre Açúcar e Café ( $Cross - SampEn(ACR, CAF) = 1.8392$ ) no conjunto de Dados A e, entre Boi e Café ( $Cross - SampEn(BOI, CAF) = 1.83$ ) no conjunto de Dados B, isto é, ocorre uma maior assincronia entre Açúcar e Café e entre Boi e Café.

Entre Açúcar e as demais *commodities* houveram pequenas diferenças para os dois períodos considerados, exceto entre Açúcar e Algodão, com uma diferença mais significativa:  $Cross - SampEn(ACR, ALG) = 1.7258$  para os Dados A, enquanto que para os Dados B, obteve-se  $Cross - SampEn(ACR, ALG) = 1.4377$ . Dessa forma, houve uma maior sincronia (menor *Cross-SampEn*) entre Açúcar e Algodão após a crise.

Em relação ao Algodão e as demais *commodities*, observa-se uma maior assincronia (maior *Cross-SampEn*) após a crise, exceto entre o par ALG-ACR, o qual teve comportamento contrário, ou seja, após a crise, Algodão e Açúcar aumentaram a sincronia em relação ao período anterior a crise. Observam-se pequenas diferenças entre os pares ALG-ARR, ALG-CAF.

Verifica-se uma maior assincronia (maior *Cross-SampEn*) entre Arroz e as demais *commodities* após a crise, entre os pares ARR-BOI e ARR-TRG, com pequenas diferenças entre os pares restantes.

Já entre Boi e as demais *commodities* observa-se uma maior assincronia (maior *Cross-SampEn*) após a crise, sendo que entre Boi e Açúcar não houve diferença significativa entre os dois períodos.

Observa-se também entre Café e as demais *commodities* uma maior assincronia (maior *Cross-SampEn*) após a crise, exceto entre Café e Açúcar que descreve comportamento contrário aos demais pares de *commodities*, ou seja, o par CAF-ACR apresenta maior sincronia após a crise, com uma pequena diferença.

Para o relacionamento entre Trigo e as demais *commodities*, observa-se uma maior assincronia entre todos os pares (TRG-ACR, TRG-ALG, TRG-ARR, TRG-BOI e TRG-CAF) após a crise.

Ao todo, em apenas 3 pares (ACR-ALG, ACR-ARR e ACR-CAF) de *commodities* (20%), a estatística *Cross-SampEn* foi maior a partir dos Dados A (antes da crise) do que a partir dos Dados B (após a crise). Assim, os 12 pares restantes de *commodities* (80%), resultaram em valores de *Cross-SampEn* maiores a partir dos Dados B (depois da crise) do que a partir dos Dados A (antes da crise). Isso quer dizer que houve uma maior assincronia entre as *commodities* (80%) no período após a crise (2010-2014).

## 4.5 DCCA

As Tabelas 4.4 e 4.5 apresentam os valores dos coeficientes de correlação-cruzada ( $\lambda$ ) do método DCCA para os dois períodos, antes e após a crise financeira mundial. Observa-se que a maioria das *commodities* (15 casos = 71.43%) apresentaram coeficiente de correlação-cruzada comportamento mais forte no período pré-crise (2006-2009). Apenas os seguintes casos (6 casos = 28.57%) apresentaram um coeficiente de correlação cruzada mais forte após a crise: ALG-ARR, ALG-CAF, ARR, ARR-CAF, CAF e CAF-TRG. Os valores em negrito correspondem ao expoente de auto-correlação  $\alpha$  no método DFA.

Tabela 4.4: DCCA entre todos os retornos das séries temporais das *commodities* analisadas (2006-2009).

DCCA	Açúcar	Algodão	Arroz	Boi	Café	Trigo
Açúcar	<b>1.0040</b>	0.9370	0.9495	0.8839	0.7714	1.0013
Algodão	0.9370	<b>0.7759</b>	0.8660	0.7545	0.6647	0.8570
Arroz	0.9495	0.8660	<b>0.8644</b>	0.8412	0.7245	0.9436
Boi	0.8839	0.7545	0.8412	<b>0.6995</b>	0.6624	0.8531
Café	0.7714	0.6647	0.7245	0.6624	<b>0.5375</b>	0.7216
Trigo	1.0013	0.8570	0.9436	0.8531	0.7216	<b>0.9215</b>

Tabela 4.5: DCCA entre todos os retornos das séries temporais das *commodities* analisadas (2010-2014).

DCCA	Açúcar	Algodão	Arroz	Boi	Café	Trigo
Açúcar	<b>0.9725</b>	0.8947	0.9075	0.8602	0.7696	0.8697
Algodão	0.8947	<b>0.7659</b>	0.8796	0.7077	0.7313	0.7782
Arroz	0.9075	0.8796	<b>0.9113</b>	0.7538	0.7781	0.8793
Boi	0.8602	0.7077	0.7538	<b>0.6171</b>	0.6084	0.7088
Café	0.7696	0.7313	0.7781	0.6084	<b>0.6107</b>	0.7415
Trigo	0.8697	0.7782	0.8793	0.7088	0.7415	<b>0.8439</b>

Aplicou-se o teste de Wilcoxon-Mann-Whitney para duas amostras independentes: Dados A (2006-2009) e Dados B (2010-2014), respectivamente. Esse teste é indicado para avaliar a hipótese de que duas populações têm mesma distribuição. Isto é, a hipótese da nulidade é a de que as amostras provenham de populações tais que a probabilidade de que uma observação, tomada ao acaso de um grupo (Dados A), seja



---

maior do que outra observação, tomada ao acaso do outro grupo (Dados B), é 0.5 (teste bilateral). Esse teste também é conhecido como teste  $U$  de Mann-Whitney, teste dos postos somados de Wilcoxon e Teste de Mann-Whitney. O teste Wilcoxon-Mann-Whitney resultou em um p-valor de 0.4395. Isto significa que não há diferença significativa ao nível de significância de 5%, entre os dois grupos avaliados, para o expoente de correlação-cruzada  $\lambda$  do método DCCA.

A Figura 4.8 exibe o coeficiente  $\rho_{DCCA}(n)$  entre Açúcar e as demais *commodities*.

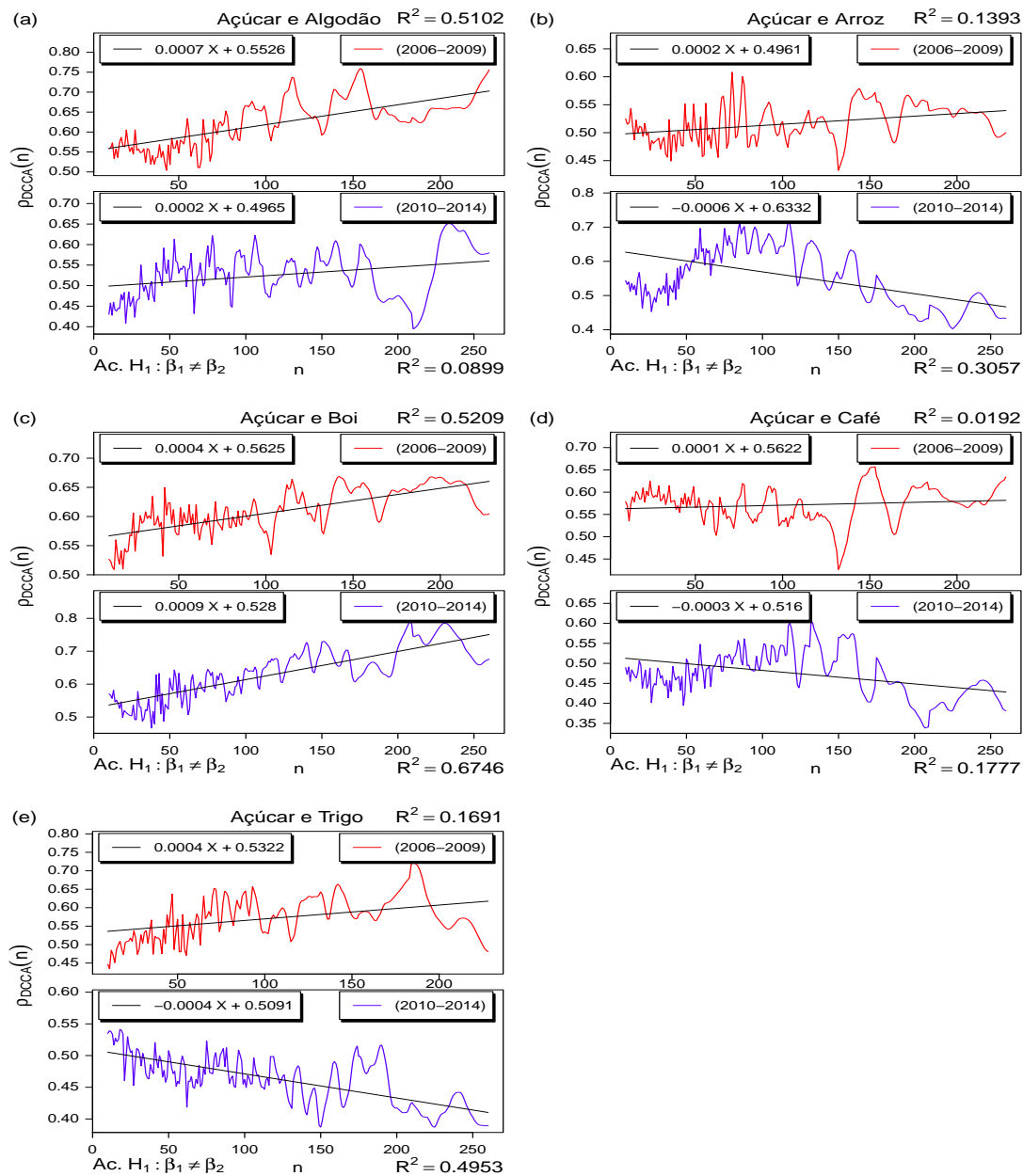


Figura 4.8: Coeficiente  $\rho_{DCCA}(n)$  entre Açúcar e as demais *Commodities*: Algodão 4.8(a), Arroz 4.8(b), Boi 4.8(c), Café 4.8(d) e Trigo 4.8(e), respectivamente.

Pode-se observar na Figura 4.8 que o comportamento do coeficiente de correlação  $\rho_{DCCA}(n)$  segue, em geral, uma tendência (reta em vermelho) menor ou negativa para o período após a crise (2010-2014) em comparação com o período anterior a crise (2006-2009), com algumas pequenas diferenças na inclinação da tendência linear entre os pares de *commodities* ACR-BOI e ACR-CAF.

A Figura 4.9 descreve o coeficiente  $\rho_{DCCA}(n)$  entre Algodão e as demais *commodities*.

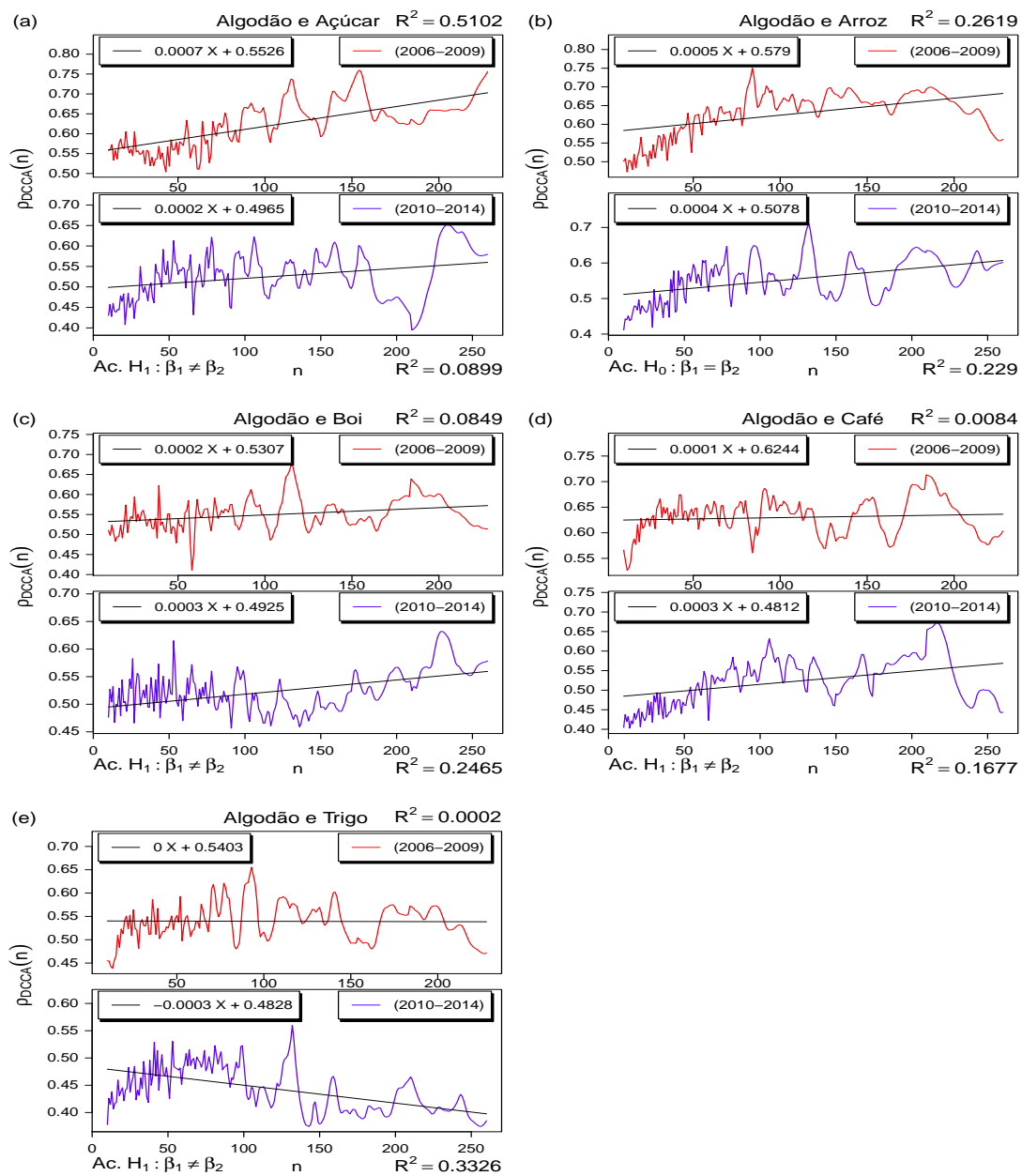


Figura 4.9: Coeficiente  $\rho_{DCCA}(n)$  entre Algodão e as demais *Commodities*: Açúcar 4.10(a), Arroz 4.9(b), Boi 4.9(c), Café 4.9(d) e Trigo 4.9(e), respectivamente.

Na Figura 4.9 pode-se observar que o comportamento do coeficiente de correlação  $\rho_{DCCA}(n)$  segue, em geral, uma tendência (reta em vermelho) menor ou negativa para o período após a crise (2010-2014) em comparação com o período anterior a crise (2006-2009), com exceção do par de *commodities* ALG-CAF que descreve comportamento contrário, com tendência crescente após a crise financeira mundial.

A Figura 4.10 descreve o coeficiente  $\rho_{DCCA}(n)$  entre Arroz e as demais *commodities*.

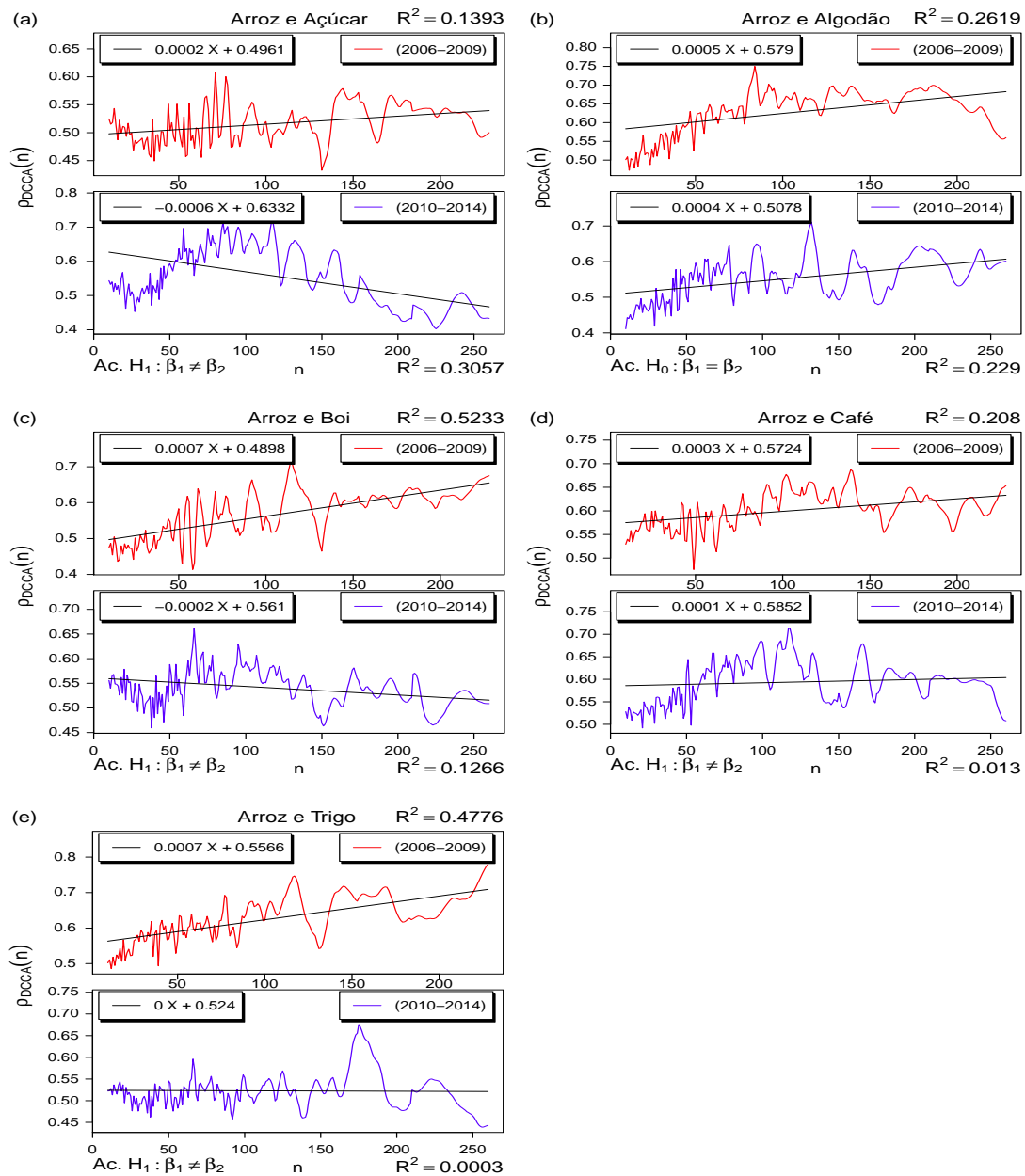


Figura 4.10: Coeficiente  $\rho_{DCCA}(n)$  entre Arroz e as demais *Commodities*: Açúcar 4.10(a), Algodão 4.10(b), Boi 4.10(c), Café 4.10(d) e Trigo 4.10(e), respectivamente.

A Figura 4.10 descreve o comportamento do coeficiente de correlação  $\rho_{DCCA}(n)$  que segue tendência (reta em vermelho) menor ou negativa para o período após a crise (2010-2014) em comparação com o período anterior a crise (2006-2009).

A Figura 4.11 descreve o coeficiente  $\rho_{DCCA}(n)$  entre Boi e as demais *commodities*.

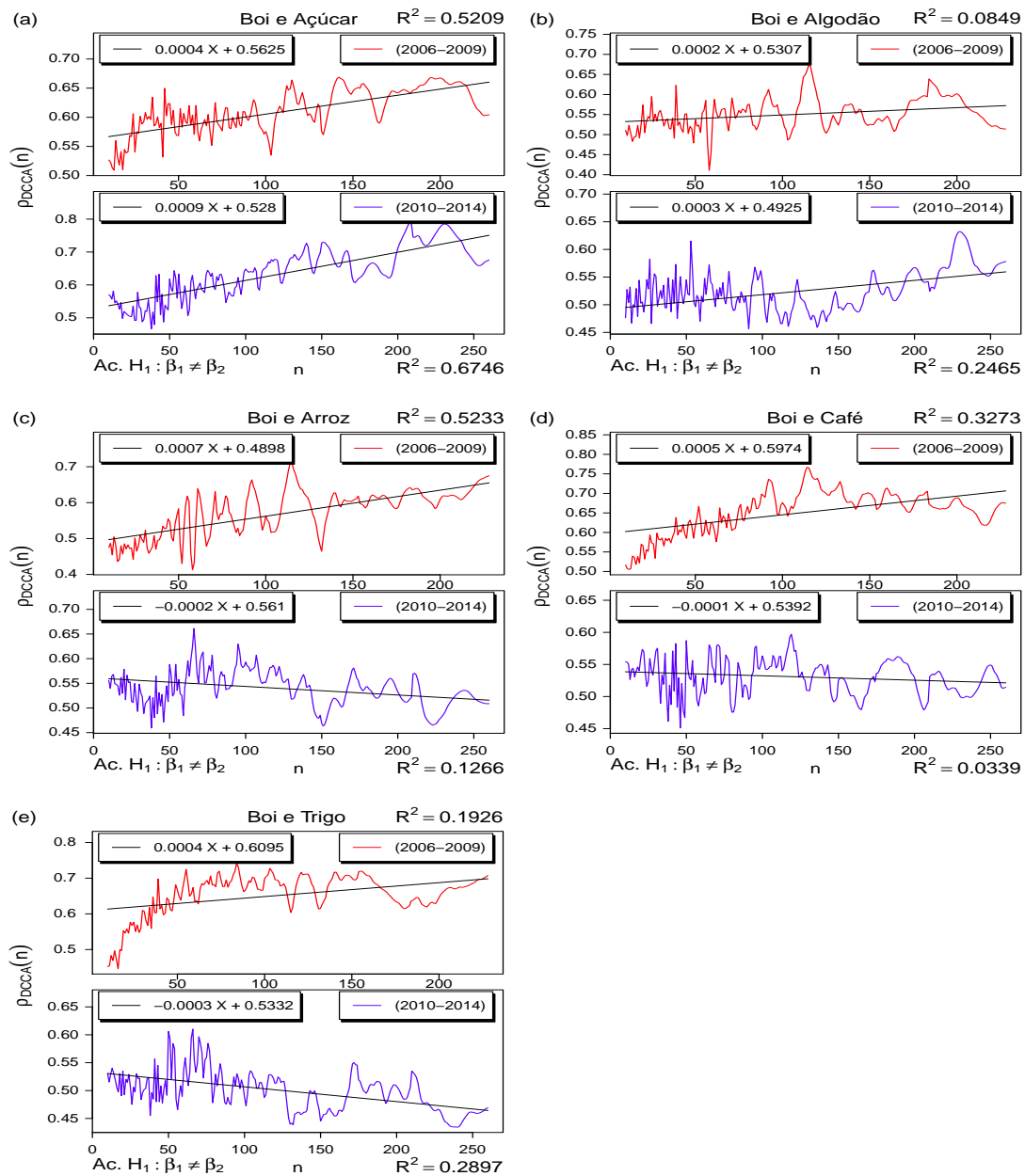


Figura 4.11: Coeficiente  $\rho_{DCCA}(n)$  entre Boi e as demais *Commodities*: Algodão 4.11(b), Açúcar 4.11(a), Arroz 4.11(c), Café 4.11(d) e Trigo 4.11(e), respectivamente.

A Figura 4.11 apresenta o comportamento do coeficiente de correlação  $\rho_{DCCA}(n)$  que segue tendência (reta em vermelho) menor ou negativa para o período após a crise (2010-2014) em comparação com o período anterior a crise (2006-2009), com exceção do par de *commodities* BOI-ACR que apresenta comportamento contrário, já apresentado na Figura 4.8.

A Figura 4.12 descreve o coeficiente  $\rho_{DCCA}(n)$  entre Café e as demais *commodities*.

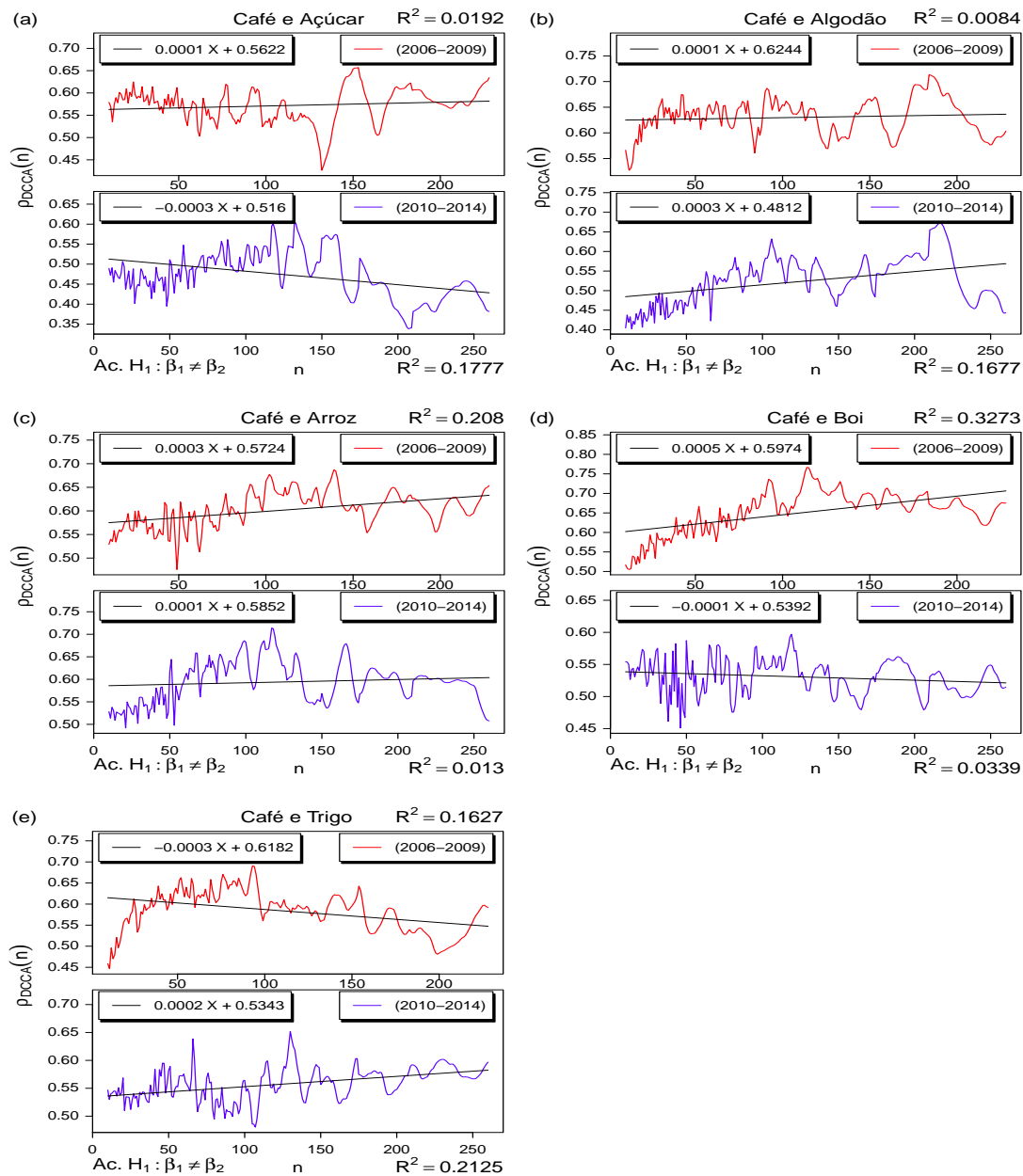


Figura 4.12: Coeficiente  $\rho_{DCCA}(n)$  entre Café e as demais *Commodities*: Algodão 4.12(b), Açúcar 4.12(a), Arroz 4.12(c), Boi 4.12(d) e Trigo 4.12(e), respectivamente.

A Figura 4.11 exibe o comportamento do coeficiente de correlação  $\rho_{DCCA}(n)$  que segue tendência (reta em vermelho) menor ou negativa para o período após a crise (2010-2014) em comparação com o período anterior a crise (2006-2009), com exceção dos pares de *commodities* CAF-ALG e CAF-TRG que apresentam comportamento contrário, com tendência crescente no período pós-crise.

A Figura 4.13 descreve o coeficiente  $\rho_{DCCA}(n)$  entre Trigo e as demais *commodities*.

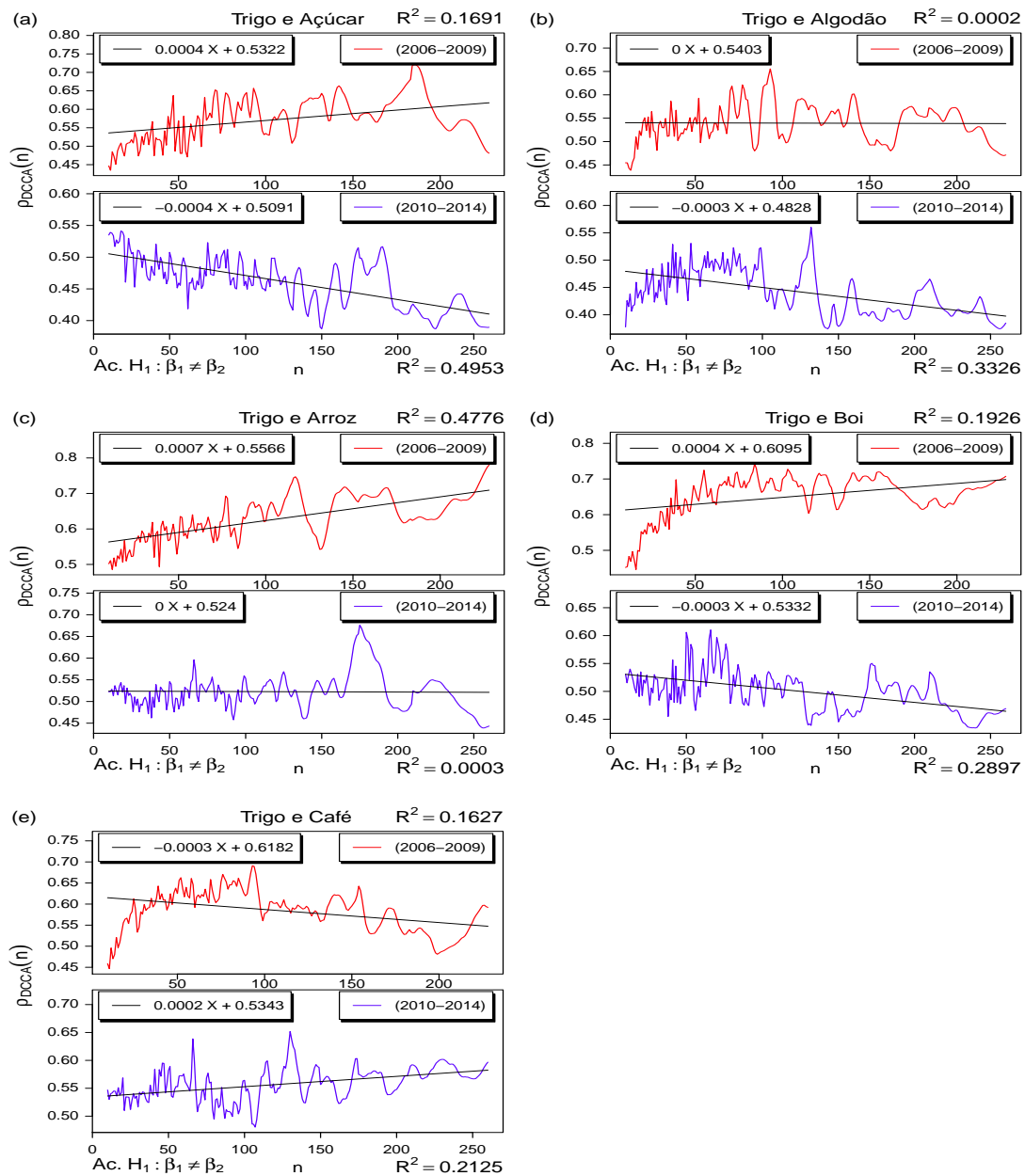


Figura 4.13: Coeficiente  $\rho_{DCCA}(n)$  entre Trigo e as demais *Commodities*: Algodão 4.13(b), Açúcar 4.13(a), Arroz 4.13(c), Boi 4.13(d) e Café 4.13(e), respectivamente.

Na Figura 4.13 pode-se observar comportamento do coeficiente de correlação  $\rho_{DCCA}(n)$  entre Trigo e as demais *commodities* o qual segue tendência (reta em vermelho) menor ou negativa para o período pós-crise (2010-2014) em comparação com o período pré-crise (2006-2009), com exceção do par de *commodities* TRG-CAF, com tendência crescente no período pós-crise.

A Tabela 4.6 indica as hipóteses aceitas na comparação entre os coeficientes angulares  $\beta_1$  e  $\beta_2$  das retas de tendência linear e seus respectivos coeficientes de determinação  $R^2$  para os dois períodos analisados. Em negrito, destaca-se o único par (ALG-ARR) em que não há diferença significativa na inclinação da reta de tendência linear para os dois períodos comparados. O teste t-Student foi utilizado para testar  $H_0$  versus  $H_1$ , ao nível de significância de 5%. Os modelos utilizados para análise de regressão foram:  $Y = \beta_0 + \beta_1 X$  e  $Y = \beta_0 + \beta_2 X$ , para os períodos 2006-2009 e 2010-2014, respectivamente.

As hipóteses testadas são:

$$\begin{cases} H_0 : \beta_1 = \beta_2 & (\text{coeficientes angulares são iguais}) \\ H_1 : \beta_1 \neq \beta_2 & (\text{coeficientes angulares não são iguais}) \end{cases}$$

Tabela 4.6: Teste de hipóteses sobre igualdade dos coeficientes angulares  $\beta_1$  e  $\beta_2$  da tendência linear nos períodos pré-crise e pós-crise, ao nível de significância de 5%.

<i>Commodities</i>	Hipótese Aceita	P-valor	2006-2009		2010-2014	
			$\beta_1$	$R^2$	$\beta_2$	$R^2$
ACR-ALG	$H_1 : \beta_1 \neq \beta_2$	0	0.0007	0.5102	0.0002	0.0899
ACR-ARR	$H_1 : \beta_1 \neq \beta_2$	0	0.0002	0.1393	-0.0006	0.3057
ACR-BOI	$H_1 : \beta_1 \neq \beta_2$	0	0.0004	0.5209	0.0009	0.6746
ACR-CAF	$H_1 : \beta_1 \neq \beta_2$	0	0.0001	0.0192	-0.0003	0.1777
ACR-TRG	$H_1 : \beta_1 \neq \beta_2$	0	0.0004	0.1691	-0.0004	0.4953
<b>ALG-ARR</b>	<b><math>H_0 : \beta_1 = \beta_2</math></b>	<b>0.2959</b>	<b>0.0005</b>	<b>0.2619</b>	<b>0.0004</b>	<b>0.229</b>
ALG-BOI	$H_1 : \beta_1 \neq \beta_2$	0.0042	0.0002	0.0849	0.0003	0.2465
ALG-CAF	$H_1 : \beta_1 \neq \beta_2$	0	0.0001	0.0084	0.0003	0.1677
ALG-TRG	$H_1 : \beta_1 \neq \beta_2$	0	0	0.0002	-0.0003	0.3326
ARR-BOI	$H_1 : \beta_1 \neq \beta_2$	0	0.0007	0.5233	-0.0002	0.1266
ARR-CAF	$H_1 : \beta_1 \neq \beta_2$	0	0.0003	0.208	0.0001	0.013
ARR-TRG	$H_1 : \beta_1 \neq \beta_2$	0	0.0007	0.4776	0	0.0003
BOI-CAF	$H_1 : \beta_1 \neq \beta_2$	0	0.0005	0.3273	-0.0001	0.0339
BOI-TRG	$H_1 : \beta_1 \neq \beta_2$	0	0.0004	0.1926	-0.0003	0.2897
CAF-TRG	$H_1 : \beta_1 \neq \beta_2$	0	-0.0003	0.1627	0.0002	0.2125



## 4.6 Análise Multifractal

A Figura 4.14 descreve o Expoente de Rényi  $\tau(q)$  para as *commodities* analisadas. Observa-se que o expoente de Rényi exibe comportamento não linear. Este comportamento é típico de processos multifractais.

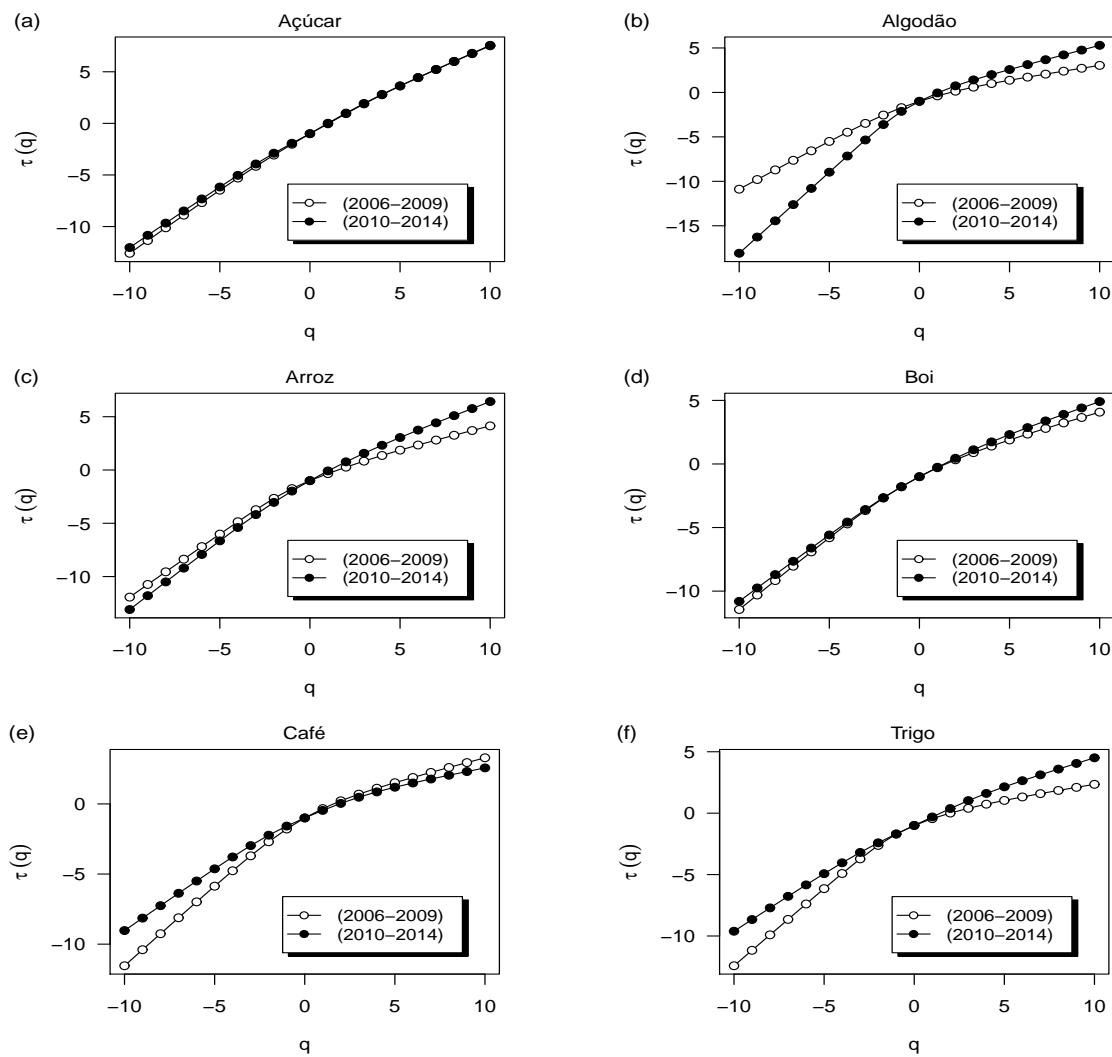


Figura 4.14: Expoente de Rényi  $\tau(q)$  das *commodities* analisadas: Açúcar 4.14(a), Algodão 4.14(b), Arroz 4.14(c), Boi 4.14(d), Café 4.14(e) e Trigo 4.14(f), respectivamente.

A Tabela 4.7 exibe o comportamento dos parâmetros multifractais. Destacam-se os valores com sublinhado simples, representando uma maior presença de multifractalidade na série temporal no período pós-crise, em detrimento do período pré-crise, cuja ocorrência observa-se apenas para a *commodity* Algodão.

Tabela 4.7: Comportamento dos parâmetros multifractais.

<i>Commodities</i>	2006-2009			2010-2014		
	$W$	$r$	$\alpha_0$	$W$	$r$	$\alpha_0$
ACR	0.4773	0.8598	1.0166	0.4132	1.0947	0.9718
ACR-ALG	0.6245	1.5190	1.0529	0.4549	1.0729	0.9733
ACR-ARR	0.4243	1.2614	1.0404	0.3753	0.9900	0.9495
ACR-BOI	0.6592	1.3412	1.0115	0.2174	1.0101	0.8645
ACR-CAF	0.4355	1.2614	0.7888	0.2914	1.8429	0.7611
ACR-TRG	0.3454	1.3976	0.9729	0.2860	1.1630	0.8186
ALG	<u>0.7659</u>	1.3140	0.6529	<u>1.2872</u>	1.7639	1.0028
ALG-ARR	0.6474	1.4875	1.0827	0.5378	1.4568	0.9746
ALG-BOI	1.0624	1.5190	1.0069	0.6738	1.6533	0.8624
ALG-CAF	0.7568	1.7639	0.8098	0.6325	1.7260	0.8039
ALG-TRG	0.6641	1.2614	0.9829	0.5732	1.4268	0.8484
ARR	0.7512	1.9701	0.6890	0.6411	1.2614	0.9396
ARR-BOI	0.6641	1.2614	0.9829	0.4806	1.0306	0.8437
ARR-CAF	0.4354	1.2111	0.8188	0.3700	1.3140	0.7650
ARR-TRG	0.5738	0.4113	0.9972	0.3949	0.6864	0.8206
BOI	0.7219	1.2614	0.7373	0.5550	0.9900	0.7800
BOI-CAF	0.5795	1.7639	0.7700	0.4005	0.7611	0.6440
BOI-TRG	0.7409	0.8426	0.9560	0.3640	1.2111	0.7297
CAF	0.8038	1.1630	0.7127	0.6328	1.0306	0.5692
CAF-TRG	0.4645	0.8598	0.7322	0.3446	1.2614	0.6259
TRG	1.0137	1.9265	0.5912	0.4975	1.0101	0.7043
Média	0.6482	1.3170	0.8765	0.4964	1.2265	0.8216
Desvio-padrão	0.1853	0.3761	0.1531	0.2225	0.3193	0.1228

De acordo com a Tabela 4.7 pode-se observar que para todas as séries temporais das *commodities*, a multifractalidade diminuiu após a crise financeira mundial, indicando menor variedade do tamanho das flutuações que apresentam invariância de escala,

exceto o Algodão, que antes da crise apresenta um valor de  $W = 0.7659$ . Após a crise observa-se  $W = 1.2872$ .

Os valores dos parâmetros de assimetria  $r$  indicam uma multifractalidade mais influenciada por escalas de grande flutuação para valores assimétricos à esquerda ( $r < 1$ ). Por outro lado, para multifractalidade mais influenciada por escalas de pequenas flutuações, tem-se valores assimétricos à direita ( $r > 1$ ).

A Tabela 4.8 apresenta o resultado do teste de Wilcoxon-Mann-Whitney para duas amostras independentes: Dados A (2006-2009) e Dados B (2010-2014), respectivamente. Esse teste é indicado para avaliar a hipótese de que duas populações têm mesma distribuição. Isto é, a hipótese da nulidade é a de que as amostras provenham de populações tais que a probabilidade de que uma observação, tomada ao acaso de um grupo (Dados A), seja maior do que outra observação, tomada ao acaso do outro grupo (Dados B), é 0.5 (teste bilateral). Esse teste também é conhecido como teste  $U$  de Mann-Whitney, teste dos postos somados de Wilcoxon e Teste de Mann-Whitney.

Tabela 4.8: Teste Wilcoxon-Mann-Whitney para os parâmetros multifractais.

Parâmetro	P-valor
$W$	0.002757
$r$	0.2212
$\alpha_0$	0.1665

De acordo com o teste, não há diferença significativa para os parâmetros  $r$  e  $\alpha$ . Por outro lado, há diferença significativa, a 5%, para o parâmetro  $W$ , entre os dois períodos estudados. Como houve diferença, repetiu-se o teste para avaliar o sentido dessa diferença. Assim, realizou-se um teste unilateral à direita e obteve-se um  $p$  –  $valor = 0.001378$ , ao nível de significância de 5%. Portanto, conclui-se que as amostras provenham de populações tais que a probabilidade de que uma observação, tomada ao acaso de um grupo (Dados A), seja maior do que outra observação, tomada ao acaso do outro grupo (Dados B), é maior que 0.5. Isso quer dizer que  $W$  é estatisticamente maior para o período 2006-2009 em relação ao período 2010-2014.

A amplitude do expoente de Hurst generalizado, a convexidade do expoente de Rényi e a largura do espectro multifractal podem ser utilizadas para avaliar o grau de multifractalidade do processo. Para verificar se a multifractalidade é causada pela função densidade de probabilidade dos valores da série ou pelas correlações diferentes em subconjuntos com grandes e pequenas flutuações, o método MF-DFA é aplicado nas séries temporais randomizadas. Para uma série de tamanho  $n$ , uma randomização consiste no embaralhamento de  $10000 \times n$  pares de valores escolhidos aleatoriamente, no qual o espectro  $f(\alpha)$  é recalculado. Repetiu-se esse processo 100 vezes para avaliar o espectro randomizado médio e o desvio-padrão, para cada uma das seis séries dos retornos das *commodities*. A largura do espectro multifractal denotada por  $W = \alpha_{máx} - \alpha_{mín}$ , é o melhor indicador do grau de multifractalidade de um processo estocástico. Maiores valores de  $W$  indicam maior multifractalidade. As Tabelas 4.9 e 4.10 apresentam os parâmetros multifractais para as séries temporais originais e randomizadas para todas as *commodities*, antes (2006-2009) e após (2010-2014) a crise financeira mundial.

De acordo com as Tabelas 4.9 e 4.10, observa-se que para o período pré-crise (2006-2009), a *commodity* Café apresenta maior multifractalidade da série original em relação a série randomizada, ou seja,  $W_o = 0.8038 > W_r = 0.7648$ . Isto quer dizer que houve multifractalidade de ambos os tipos: multifractalidade causada pela função densidade de probabilidade dos valores da série (tipo 1) e multifractalidade causada pelas correlações em subconjuntos com grandes e pequenas flutuações (tipo 2). Por outro lado, a *commodity* Trigo não apresentou diferença significativa na largura do espectro multifractal para os dois períodos analisados. Isso indica que houve multifractalidade do tipo 1, ou seja, a randomização não afetou a função densidade de probabilidade dos valores da série. Para o período pós-crise, apenas a *commodity* Algodão apresenta maior multifractalidade da série original em detrimento da série randomizada, apresentando ambos os tipos de multifractalidade, para o período após a crise (2010-2014).

Tabela 4.9: Comportamento dos parâmetros multifractais (2006-2009).

<i>Commodities</i>	2006-2009					
	Original			Aleatória		
	$W$	$r$	$\alpha_0$	$W$	$r$	$\alpha_0$
ACR	0.4773	0.8598	1.0166	0.6325	1.5513	0.5594
ALG	0.7659	1.3140	0.6529	0.9160	1.9701	0.5981
ARR	0.7512	1.9701	0.6890	0.8594	1.6892	0.6069
BOI	0.7219	1.2614	0.7373	0.9326	1.1398	0.6186
CAF	0.8038	1.1630	0.7127	0.7648	1.4268	0.5751
TRG	1.0137	1.9265	0.5912	0.9990	0.6179	0.6703
Média	0.7556	1.4158	0.7333	0.8507	1.3992	0.6047
Desvio-padrão	0.1719	0.4417	0.1479	0.1327	0.4715	0.0386

Tabela 4.10: Comportamento dos parâmetros multifractais (2010-2014).

<i>Commodities</i>	2010-2014					
	Original			Aleatória		
	$W$	$r$	$\alpha_0$	$W$	$r$	$\alpha_0$
ACR	0.4132	1.0947	0.9718	0.8474	0.8774	0.6341
ALG	1.2872	1.7639	1.0028	1.0856	0.7304	0.6835
ARR	0.6411	1.2614	0.9396	0.9574	1.5513	0.6235
BOI	0.5550	0.9900	0.7800	0.7399	1.4568	0.5827
CAF	0.6328	1.0306	0.5692	0.7769	1.1868	0.5909
TRG	0.4975	1.0101	0.7043	0.8005	1.0306	0.5971
Média	0.6711	1.1918	0.8279	0.8679	1.1389	0.6186
Desvio-padrão	0.3137	0.2971	0.1722	0.1304	0.3226	0.0374

A Figura 4.15 ilustra o espectro multifractal  $f(\alpha)$  dos retornos das *commodities* Açúcar, Algodão e Arroz, para as séries original e randomizada, nos períodos 2006-2009 e 2010-2014.

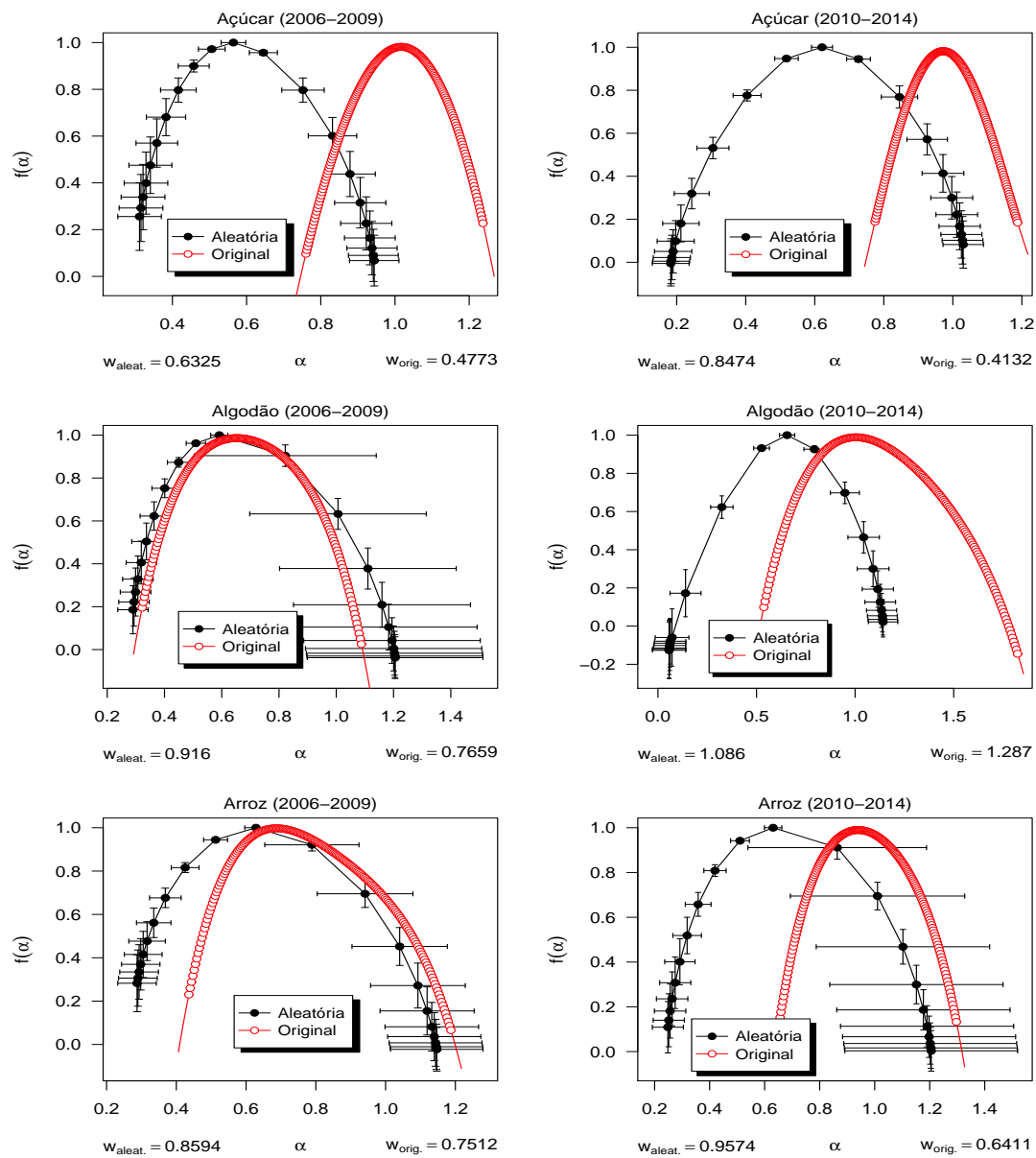


Figura 4.15: Espectro multifractal  $f(\alpha)$  dos retornos das *commodities* Açúcar, Algodão e Arroz, para as séries original e randomizada, nos períodos 2006-2009 e 2010-2014.

Pode-se observar na Figura 4.16 o comportamento do espectro multifractal  $f(\alpha)$  dos retornos das *commodities* Boi, Café e Trigo, para as séries original e randomizada, nos períodos 2006-2009 e 2010-2014.

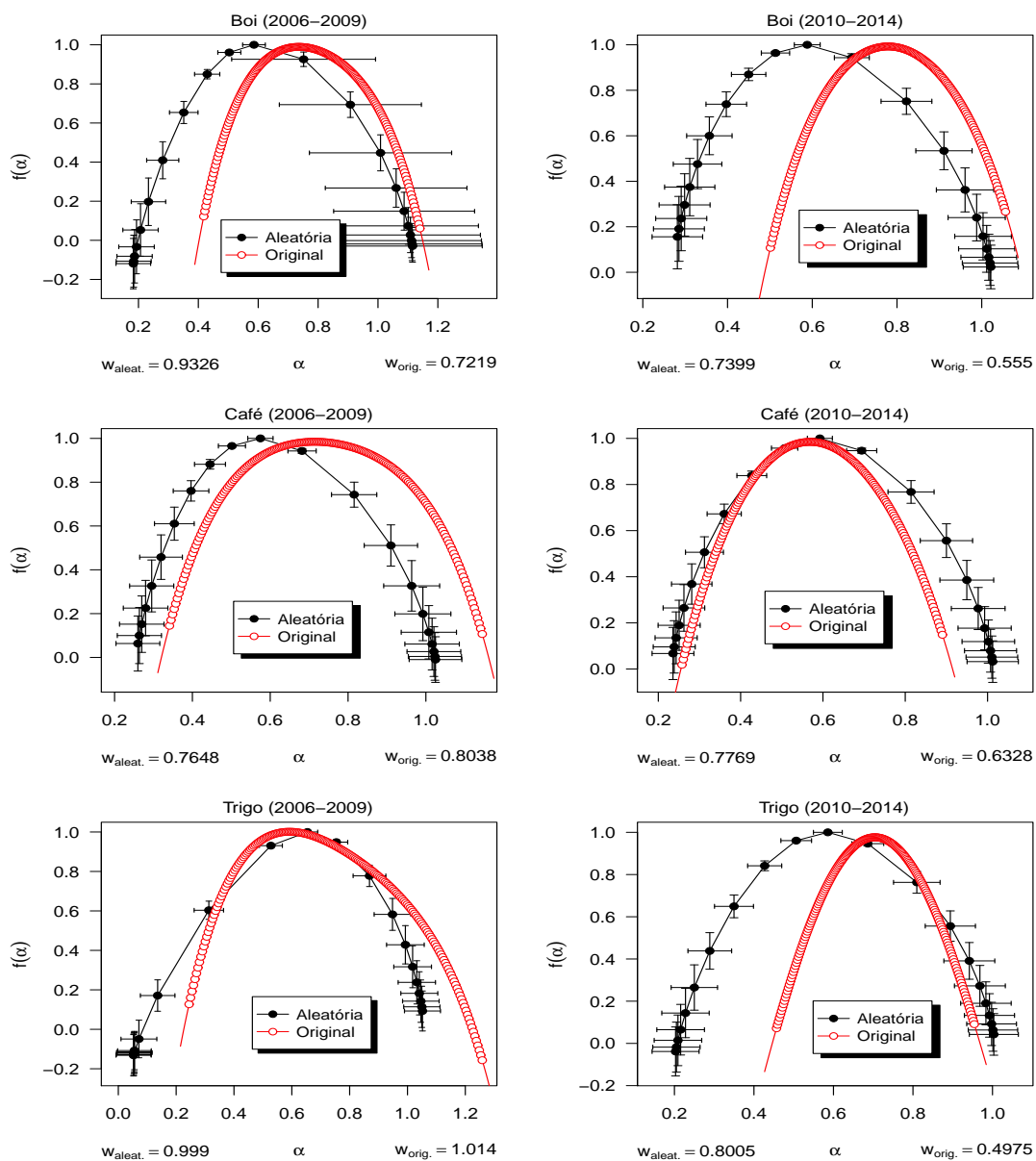


Figura 4.16: Espectro multifractal  $f(\alpha)$  dos retornos das *commodities* Boi, Café e Trigo, para as séries original e randomizada, nos períodos 2006-2009 e 2010-2014.

## 4.7 Análise em Janelas Móveis de Tempo

### 4.7.1 *SampEn* em Janela Móvel de Tempo (2006-2014)

A Figura 4.17 exibe a estatística *SampEn*, para janela de tempo de 240 dias (aproximadamente 1 ano) com passo de um dia entre cada janela. Analisaram-se 1721 janelas para cada *commodity*. A reta em vermelho exibe a tendência linear.

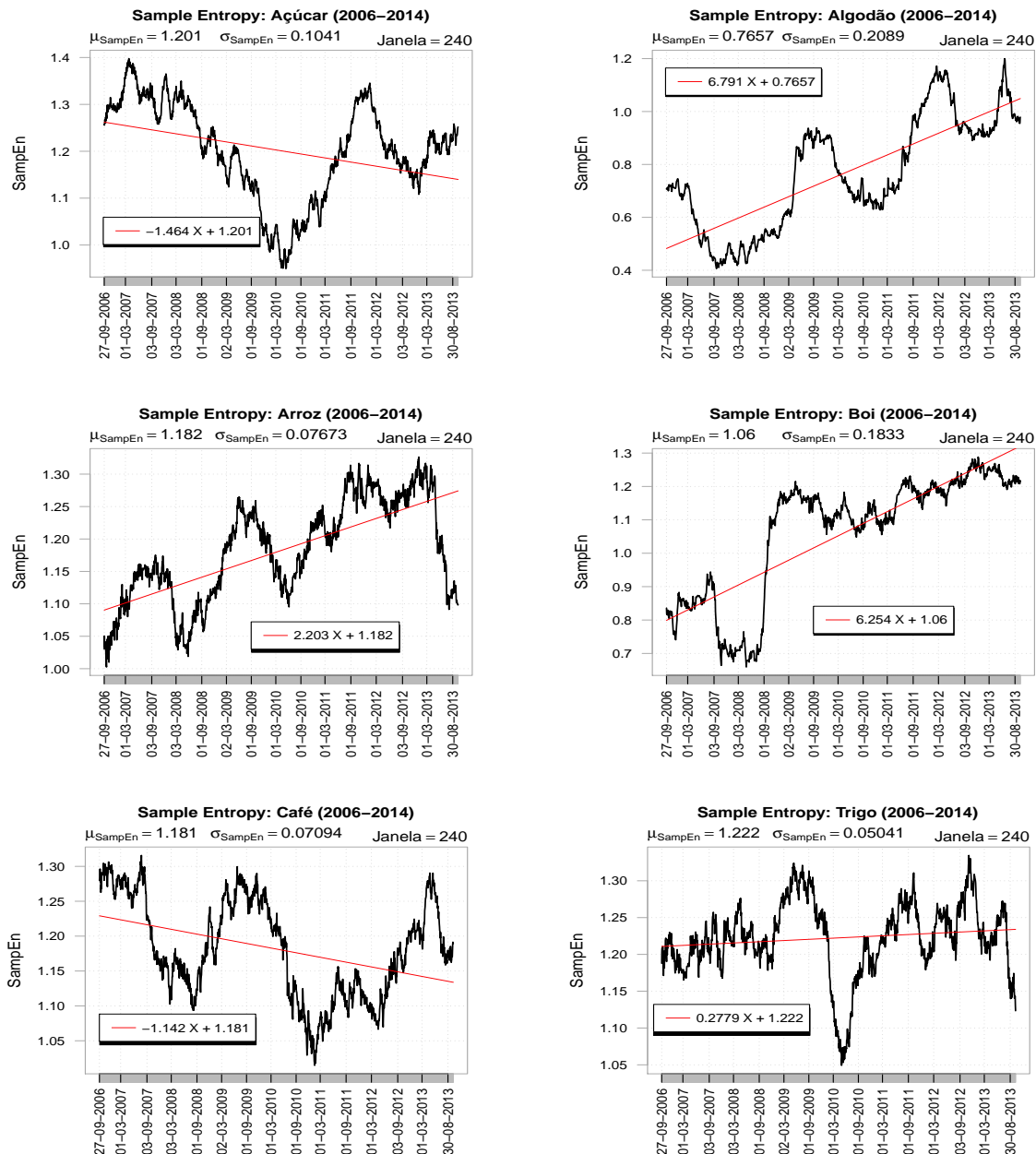


Figura 4.17: Estatística *SampEn* em Janela Móvel de Tempo (tamanho 240 dias), com passo de um dia entre cada janela, totalizando 1721 janelas de tamanho 240 para cada *commodity* no período 2006-2014.



Observa-se, para a *commodity* Açúcar, que há uma tendência de queda para a complexidade durante o período pré-crise (2006-2009). Ao contrário, o período pós-crise (2010-2014) apresenta uma complexidade com tendência crescente.

Para o *Algodão*, a série apresenta uma queda (menor complexidade) até o final de 2007 e após isso há uma tendência crescente (aumento de complexidade) até o início de 2009. Após isso, a série apresenta alternância entre maior e menor complexidade. Durante a série inteira, há alternância entre maior e menor complexidade, para a série da estatística *SampEn*, com tendência crescente para o período considerado.

Até meados de 2007, observa-se uma maior complexidade da série, para a *commodity* Boi. Após isso, há uma forte queda da estatística *SampEn*, e no começo de setembro de 2008 há um forte aumento da complexidade com ápice em março de 2009. Após isso, observa-se que a complexidade apresenta uma média em torno de 1.2.

O Café apresenta complexidade em queda a partir de setembro de 2007 e, após setembro de 2008 há um aumento da complexidade com ápice por volta de junho de 2009. Em seguida, há uma tendência de queda na complexidade. A partir de março de 2011 a estatística *SampEn* exibe complexidade crescente.

Para a *commodity* Trigo, a complexidade permanece constante, em média, até meados de 2008. Há uma maior complexidade a partir do início de 2009 até julho de 2009. Após isso, há uma queda acentuada da complexidade até o mês de maio de 2010. A partir disso, há uma tendência crescente para o período restante.

A Tabela 4.11 exibe as estatísticas descritivas da estatística *SampEn* para 1721 janelas móveis de 240 dias para cada *commodity* entre 2006-2014.

Tabela 4.11: Estatísticas descritivas da estatística *SampEn* para 1721 janelas móveis de tempo de 240 dias (2006-2014).

	ACR	ALG	ARR	BOI	CAF	TRG
Média	1.20	0.77	1.18	1.06	1.18	1.22
Mediana	1.21	0.73	1.18	1.13	1.18	1.22
Máximo	1.40	1.20	1.33	1.29	1.32	1.33
Mínimo	0.95	0.41	1.00	0.66	1.01	1.05
Desv.Padrão	0.10	0.21	0.08	0.18	0.07	0.05

### 4.7.2 Expoente de Hurst em Janela Móvel de Tempo no período 2006-2014

A Figura 4.18 descreve o Expoente de Hurst em 1721 janelas móveis de 240 dias (aproximadamente 1 ano), com passo de um dia entre cada janela. A reta em vermelho representa a tendência linear.

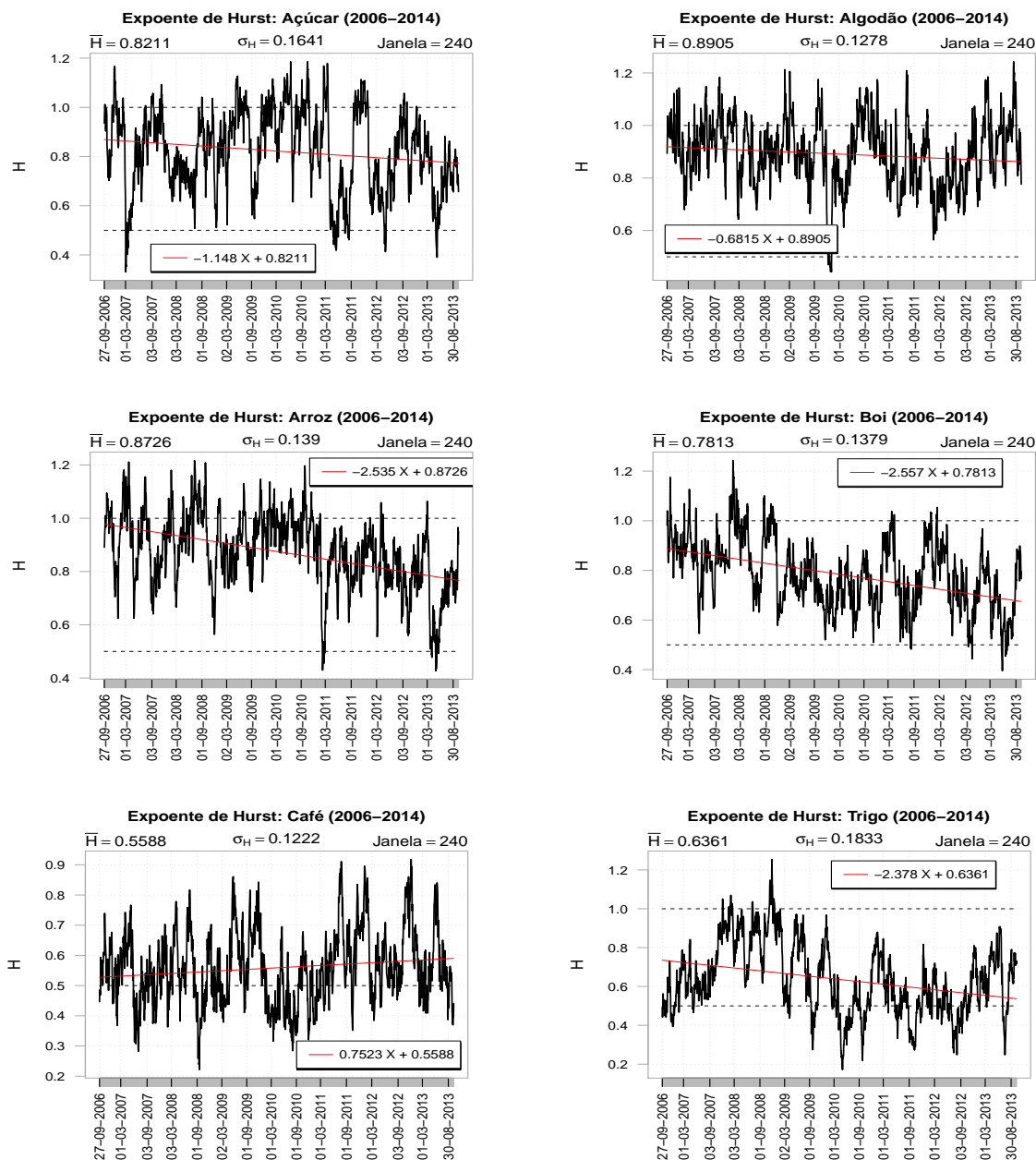


Figura 4.18: Expoente de Hurst em Janela Móvel de Tempo (tamanho 240 dias), com passo de um dia entre cada janela, totalizando 1721 janelas de tamanho 240 para cada *commodity* no período 2006-2014.

Para a *commodity* Açúcar, pode-se observar uma forte queda do expoente  $H$  em 01/03/2007, exibindo comportamento antipersistente ( $H < 0.5$ ). Após essa data, observa-se um aumento acentuado e, a seguir, entre 03/09/2007 e 01/09/2008 uma nova queda do expoente  $H$ , exibindo comportamento fracamente persistente. No geral, há uma tendência decrescente do expoente  $H$  ao longo de todo período considerado.

Em relação a *commodity* Algodão, observa-se o expoente  $H$  flutuando em torno de uma média de aproximadamente 0.9 até setembro/outubro de 2009, com comportamento persistente. Após esse período, no fim de 2009, observa-se uma forte queda do expoente  $H$ , passando a ter comportamento antipersistente, próximo a 0.5.

Para as *commodities* Arroz e Boi observa-se uma tendência geral de queda no comportamento do expoente de Hurst  $H$ . Para a *commodity* Trigo observa-se um comportamento persistente crescente até meados de 2008. Após esse período observa-se uma queda do expoente  $H$  até atingir valor próximo a 0.2, revelando comportamento antipersistente. Após isso o expoente de Hurst flutua em torno de uma média de aproximadamente  $H = 0.6$ .

Por outro lado, para *commodity* Café, o expoente de Hurst apresenta uma tendência geral de comportamento levemente persistente, com uma forte queda em meados de 2008 e início de 2009.

A Figura 4.19 exibe o Expoente de Hurst em 1481 janelas móveis de 480 dias (aproximadamente 2 anos), com passo de um dia entre cada janela. A reta em vermelho representa a tendência linear.

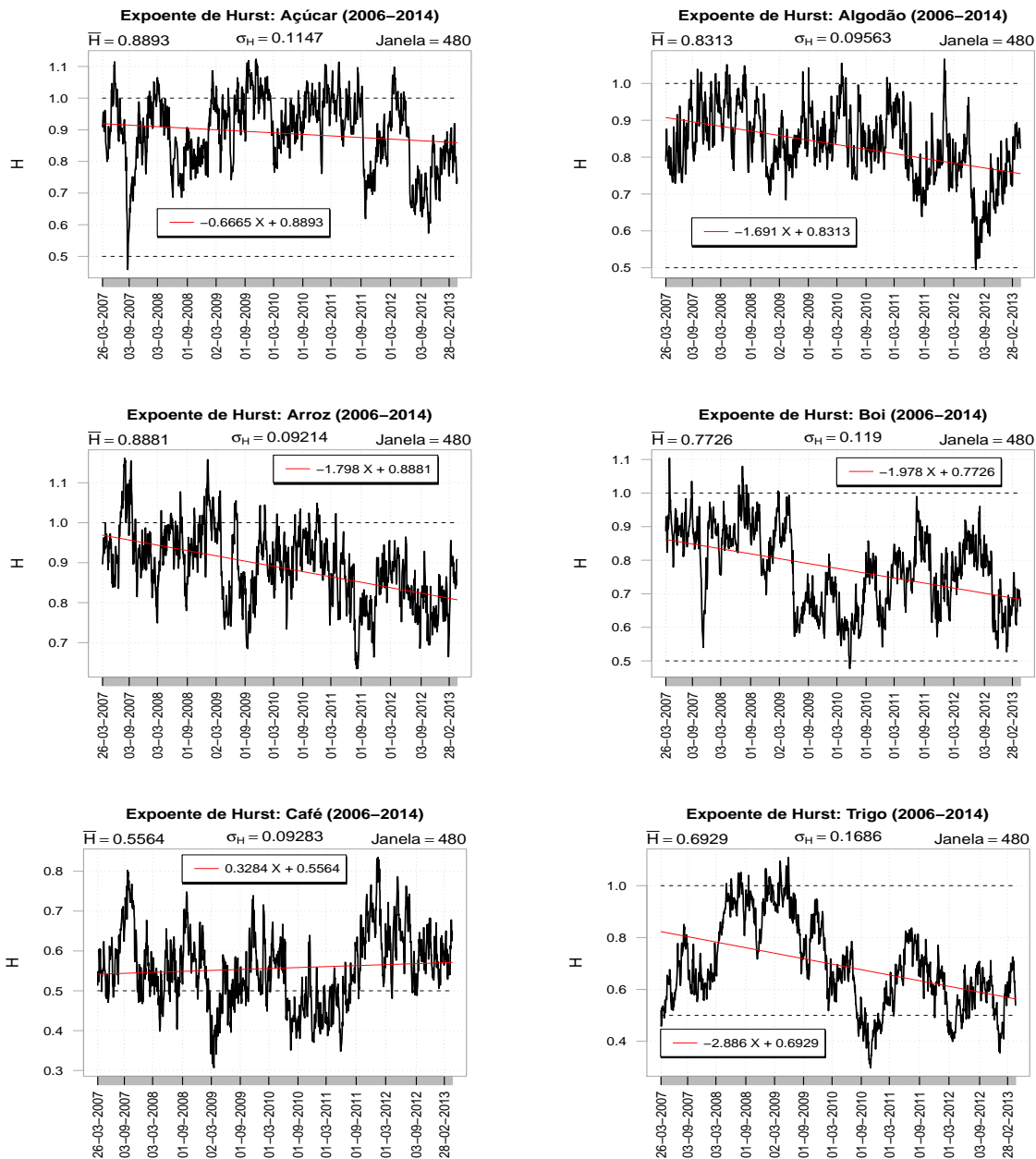


Figura 4.19: Expoente de Hurst em Janela Móvel de Tempo (tamanho 480 dias), com passo de um dia entre cada janela, totalizando 1481 janelas de tamanho 480 para cada *commodity* no período 2006-2014.

A Figura 4.20 ilustra o Expoente de Hurst em 961 janelas móveis de 1000 dias (aproximadamente 4 anos), com passo de um dia entre cada janela. A reta em vermelho representa a tendência linear.

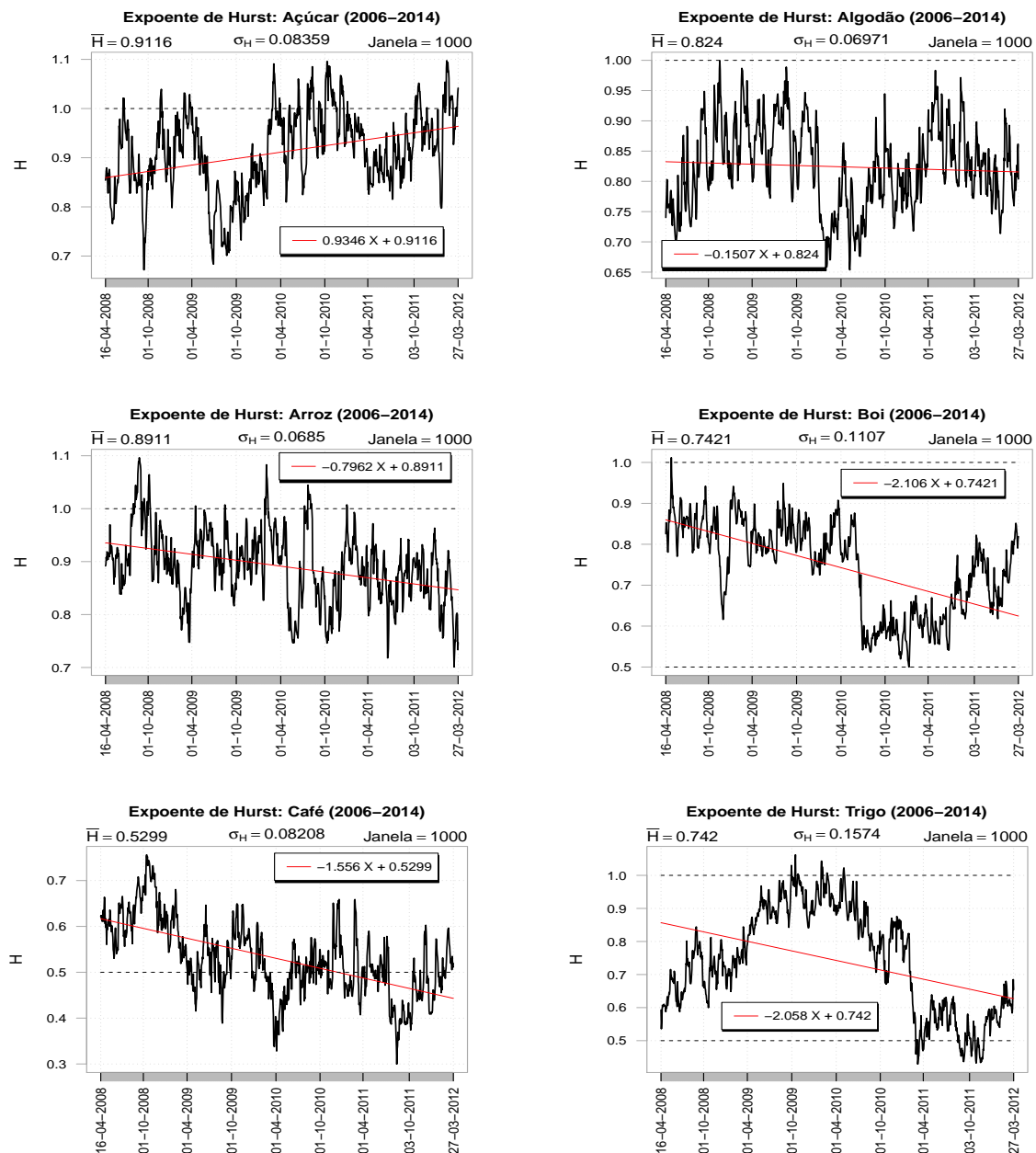


Figura 4.20: Expoente de Hurst em Janela Móvel de Tempo (tamanho 1000 dias), com passo de um dia entre cada janela, totalizando 961 janelas de tamanho 1000 para cada commodity no período 2006-2014.

A Tabela 4.12 exibe as estatísticas descritivas do Expoente de Hurst ( $H$ ) para 1721 janelas móveis de 240 dias para cada *commodity* entre 2006-2014.

Tabela 4.12: Estatísticas descritivas do Expoente de Hurst ( $H$ ) com janela móvel de tempo de 240 dias (2006-2014).

	ACR	ALG	ARR	BOI	CAF	TRG
Média	0.82	0.89	0.87	0.78	0.56	0.64
Mediana	0.83	0.89	0.89	0.78	0.55	0.62
Máximo	1.19	1.24	1.22	1.24	0.92	1.25
Mínimo	0.33	0.44	0.43	0.40	0.22	0.17
Desv.Padrão	0.16	0.13	0.14	0.14	0.12	0.18
$(0.5 < H < 1)$ (%)	82.45	79.78	81.52	92.97	66.94	72.81
$(0 < H < 0.5)$ (%)	3.02	0.70	0.99	1.34	33.06	25.10
$(H > 1)$ (%)	14.53	19.52	17.49	5.69	0.00	2.09
$(H = 0.5)$ (%)	0.00	0.00	0.00	0.00	0.00	0.00
P-valor Shapiro	0.00	0.00	0.00	0.00	0.00	0.00
Mediana-0.5	0.33	0.39	0.39	0.28	0.05	0.12
P-valor Wilcoxon	0.00	0.00	0.00	0.00	0.00	0.00

Observa-se que as *commodities* exibem um alto percentual de persistência ( $0.5 < H < 1$ ). O teste de normalidade de Shapiro-Wilk mostra que o Expoente de Hurst não é normalmente distribuído. Utilizou-se o Teste Wilcoxon-Mann-Whitney para uma amostra, como objetivo de verificar se a mediana do índice de Hurst é estatisticamente igual a 0.5. O teste rejeitou essa hipótese para todas as *commodities* analisadas. Valores  $0.5 < H < 1$  indicam que a série possui correlações de longo alcance persistentes, ou seja, maiores valores (menores) têm maior probabilidade de serem seguidos por maiores valores (menores) e a função de autocorrelação diminui seguindo uma lei de potência.

A série exibe comportamento antipersistente, se  $0 < H < 0.5$ , significando que os maiores valores (menores) tem maior probabilidade de serem seguidos por menores (maiores) valores. Se  $H = 0.5$ , a série não apresenta autocorrelação. Para  $H > 1$  as correlações existem, porém não diminuem seguindo uma lei de potência. Assim, todas as séries de preços das *commodities* avaliadas apresentaram comportamento persistente, na maior parte do período observado, para 1721 janelas móveis com tamanho de 240 dias (1 ano).