

UNIVERSIDADE DE SÃO PAULO
FACULDADE DE ECONOMIA, ADMINISTRAÇÃO E CONTABILIDADE DE
RIBEIRÃO PRETO
DEPARTAMENTO DE CONTABILIDADE
MESTRADO EM CONTROLADORIA E CONTABILIDADE

NATÁLIA DINIZ

O impacto da janela de Hurst na previsão de séries temporais financeiras.

Orientador: Prof. Dr. Fabiano Guasti Lima

RIBEIRÃO PRETO
2011

Reitor da Universidade de São Paulo
Prof. Dr. João Grandino Rodas

Diretor da Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto
Prof. Dr. Sigismundo Bialoskorski Neto

Chefe do Departamento de Contabilidade de Ribeirão Preto
Profa. Dra. Adriana Maria Procópio de Araújo

Coordenador do Programa de Pós-Graduação em Ciências Contábeis e Controladoria de
Ribeirão Preto
Prof. Dr. André Carlos Brusanelli de Aquino

Versão Corrigida, a original encontra-se disponível no serviço de Pós Graduação da
FEARP/USP

NATÁLIA DINIZ

O impacto da janela de Hurst na previsão de séries temporais financeiras.

Dissertação apresentada ao Programa de Pós-Graduação em Controladoria e Contabilidade da Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto da Universidade de São Paulo como requisito para obtenção do título de Mestre em Ciências.

Orientador: Prof. Dr. Fabiano Guasti Lima

**RIBEIRÃO PRETO
2011**

FICHA CATALOGRÁFICA

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

Diniz, Natália

O impacto da janela de Hurst na previsão de séries temporais financeiras. Ribeirão Preto, 2011.

137p. : il. ; 30 cm

Dissertação de Mestrado, apresentada à Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto/USP.
Área de concentração: Contabilidade.

Orientador: Lima, Fabiano Guasti.

1. Séries Temporais Financeiras. 2. Previsão. 3. Expoente de Hurst. 4. Redes Neurais Recorrentes. 5. Ondaletas

**O IMPACTO DA JANELA DE HURST NA PREVISÃO DE SÉRIES TEMPORAIS
FINANCEIRAS.**

Dissertação apresentada ao Programa de Pós-Graduação em Controladoria e Contabilidade da Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto da Universidade de São Paulo como requisito para obtenção do título de Mestre Ciências.

Área de Concentração: Contabilidade para usuários externos e finanças.

Aprovado em:

Banca Examinadora

Prof. Dr. Fabiano Guasti Lima (Presidente)

Instituição: FEA-RP

Assinatura: _____

Prof. Dr. Henrique Von Dreifus

Instituição: IME

Assinatura: _____

Prof. Dr. Angelo Costa Gurgel

Instituição: FEA-RP

Assinatura: _____

Agradecimento:

Ao iniciar a dissertação pensei que esta parte seria a mais simples de se fazer, mas agora, percebo que não é. Ter que lembrar de cada pessoa que de alguma forma me deu forças, ânimo, me empolgou, me ajudou com teorias, programações, para fazer essa pesquisa e fazê-la direito, não é uma tarefa fácil.

Então para começar gostaria de fazer uma prece de agradecimento a Deus por ter colocado pessoas tão maravilhosas no meu caminho para a realização de mais esta etapa, que no início não passava apenas de um sonho.

Existem pessoas que serão lembradas para sempre...

Agradeço ao melhor professor que já tive, ao meu ver, o mestre que qualquer aluno gostaria de ter, meu “pai acadêmico”, Antônio Carlos, por toda ajuda e pela indispensável contribuição para que este trabalho se concretizasse; É AC... em um “Admirável mundo novo” , podendo escolher, melhor trabalhar para ser um alfa, não é!? Risos...

Deixo um muito obrigada também ao meu querido orientador Fabiano que sempre me ajudou e apoiou ao máximo para o desenvolvimento dessa pesquisa.

Agradeço demais a minha mãe por sempre ser um símbolo de garra, luta e determinação em minha vida. Ao meu amado pai pelo exemplo em respeitar e tratar bem o próximo. Aos meus irmãos Helena e Alexandre pela compreensão, e o meu agradecimento pela companhia da Helena, minha irmã linda, em idas a Ribeirão para reuniões. Não posso deixar de lembrar da amiga e cunhada, Kelly pelas divertidas viagens e conselhos (mesmo que as vezes meio esquisitos neh, kkkkk). Meus sinceros agradecimentos para o meu amigo Guilherme ‘Guigui’ pelos inúmeros concertos no meu computador e pela ajuda com programações e sinceramente... ainda não entendi aquela do C.... risos.... A minha tia Bernadete pelos deliciosos cafés da tarde e pelas sábias palavras de conforto e compreensão. Ao meu amigo Kazuo, pelos conselhos e conversas, a Maria Silvia por me aturar toda semana com paciência e sempre pronta com um conselho para se pensar e ao Alexandre, pelo carinho ao final dessa jornada.

Enfim, a todos que me ajudaram direta e indiretamente com palavras, conselhos, ajudas para a construção dessa pesquisa.

RESUMO

DINIZ, Natália. **O impacto da janela de Hurst na previsão de séries temporais financeiras.** 2011, 137p. Dissertação (Mestrado) – Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto, 2011.

Sabe-se que, na literatura, existem muitos modelos para se fazer previsão para séries temporais financeiras. Sabe-se também que não há um modelo perfeito e que os mais utilizados atualmente são os modelos de redes neurais recorrentes e os da família GARCH. Referências internacionais apontam que existe uma técnica de medição de uma janela temporal para se identificar o tipo de comportamento existente em uma série temporal; tal técnica é conhecida como Expoente de Hurst. É uma medida que qualifica a série como persistente ou anti-persistente. Este trabalho analisou se o Expoente de Hurst, interfere na qualidade das previsões feitas com o modelo de redes neurais recorrentes com e sem o uso do filtro de ondaletas, utilizando os preços diários das principais commodities, ações negociadas no mercado e a taxa de câmbio. no período de janeiro de 1998 a dezembro de 2010. Com a pesquisa observa-se, na maioria dos casos, há uma possível melhora na qualidade das previsões para as séries antipersistentes.

Palavras Chave: Séries Temporais Financeiras, Previsão, Expoente de Hurst, Redes Neurais Recorrentes, ARIMA GARCH, Ondaletas.

ABSTRACT

DINIZ, Natália. **The impact of Hurst's window on the preview of financial time series.** 2011, 137p. Dissertação (Mestrado) – Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto, 2011.

It is known that there are a lot of models to forecast financial time series. It is known, also, that there is not a perfect model and the most used nowadays are the Recurrent Neural Network models and those from the GARCH family. International references point to a technique of measurement using windowing in order to identify the kind of behavior that is present in time series. This technique is known as Hurst Exponent. It is a measure that qualifies the time series as persistent or anti-persistent. This work analyzed if the Hurst Exponent interferes in the quality of the forecasts made with the Neural Network models with and without the wavelet filter, using the main commodities, stock prices, Ibovespa index and the Dollar/Real exchange rate in the period ranging from January 1998 to December 2010. The initial conclusions concerning the models worked out are positives.

Key words: Financial Time Series, Forecasting, Hurst Exponent, Neural Networks, ARIMA GARCH, wavelets.

LISTAS DE ABREVIATURAS E SIGLAS

AIC – Critério de informação akaike

AR – Modelo Autorregressivo

ARCH – heterocedasticidade condicional auto-regressiva

ARIMA – Modelo Autorregressivo Integrado de Médias Móveis

ARMA – Modelo Autorregressivo de Médias Móveis

BIC – Critério de informação Bayesiano

BM&FBOVESPA – Bolsa de mercadorias e futuro da BOVESPA

CSCE – Bolsa de mercado futuro de Nova Iorque

DAX 100 – Índice da bolsa de valores Alemã

ESALQ – Centro de estudos avançados em economia aplicada

EuroStoxx 50 – índice ponderado pela capitalização de 50 ações *blue chips* de empresas pertencentes à zona euro.

EXPN – Bolsa de Valores de Londres

EWMA – Suavização Exponencial

EWV – Volatilidade exponencialmente ponderada

FAC – Função de Autocorrelação

FACP – Função de Autocorrelação Parcial

FTSE 100- índice da bolsa de Londres

GARCH – heterocedasticidade condicional auto-regressiva generalizada

IBRX50 – Índice Brasil, desempenho das 50 ações mais negociadas no mercado à vista da BOVESPA

IC – Critério de informação

ITSA4 – Ação preferencial do Itaú Unibanco

JB – Estatística de Jarque Bera

KOSPI – Bolsa de Valores da Coreia do Sul

LIFFE – Bolsa de mercado futuro de Londres

LM – multiplicadores de Lagrange

LB – teste de Box-LJung

MA – Modelo de Médias Móveis

MAPE – Erro Percentual Absoluto Médio

MVE – Volatilidade Estocástica

NASDAQ – Bolsa de valores Americana

NK225 - Nikkei 225 , índice da bolsa de valores de Tókio
PETR4 – Ação preferencial da Petrobrás
QMLE – Medidor de máxima verossimilhança
SPI – Bolsa Australiana
SHI – bolsa de valores de Shangai
SZBI –Bolsa de valores chinesa
SZA – Bolsa de valores chinesa
S&P - Índice de Mercado Americano
S&P 500- Standard and Poor's , padrão de referência para avaliação de performance do mercado de ações norte americano
TAIEX – Índice da bolsa de Taiwan
TIC – Coeficiente de Desigualdade de Theil
TOPIX – Bolsa de Valores de Tókio
TWSI – Bolsa de valores de Taiwan
 r – Coeficiente de Correlação linear de Pearson.
RN – Redes Neurais
RNA – Redes Neurais Artificiais
RVM – máquina de vetor de relevância
RMSE – Erro quadrático médio absoluto.
VALE5 – Ação preferencial da Vale do Rio Doce
VaR – Perda Máxima de uma Carteira de Ativos

ÍNDICE DE FÓRMULAS

Fórmula 1 – Função de Autocorrelação Parcial	9
Fórmula 2 – Processo Autorregressivo de ordem p.....	12
Fórmula 3 – Média de um Processo Autorregressivo Estacionário	12
Fórmula 4 – Modelo Autorregressivo de ordem 1	12
Fórmula 5 – Média de um Modelo Autorregressivo de ordem 1	12
Fórmula 6 – Processo de Médias Móveis de ordem q.....	13
Fórmula 7 – Processo de Médias Móveis de ordem 1	13
Fórmula 8 – Processo Autorregressivo e de Médias Móveis de ordem (p,q).....	14
Fórmula 9 – Média do Processo Autorregressivo e de Médias Móveis de ordem (p,q) 14	14
Fórmula 10 – Condição para estacionariedade da série	14
Fórmula 11 – Modelo de ARMA (1,1).....	15
Fórmula 12 – Modelo de ARMA (p,q).....	15
Fórmula 13 – Modelo ARIMA	16
Fórmula 14 – Modelo de ARIMA (p,d,q)	16
Fórmula 15 – Operador autorregressivo não estacionário.....	16
Fórmula 16 – Modelo de ARCH de ordem r	18
Fórmula 17 – Modelo de GARCH (r,s)	19
Fórmula 18 – Modelo de GARCH (1,1).....	19
Fórmula 19 – Cálculo Expoente de Hurst	30
Fórmula 20 –Média dos retornos para cada janela	38
Fórmula 21 –Desvio padrão para cada janela.....	38
Fórmula 22 –Estatística R/S para cada janela	38
Fórmula 23 –Média R/S associada ao tamanho da janela	38
Fórmula 24 –Expoente de Hurst para cada janela	38
Fórmula 25 –Regressão pelos mínimos quadrados	39
Fórmula 26 –Série de retornos segundo GARCH (1,1).....	39
Fórmula 27 –Metodologia GARCH (1,1)	39
Fórmula 28 –Processo funcionamento Rede Neural	41
Fórmula 29 –Aplicação de um filtro por uma função de ondaleta	42
Fórmula 30 –Captação de altas e baixas frequências	42
Fórmula 31 – Parte Aproximação do filtro de ondaletas.....	42

Fórmula 32 – Parte Detalhe do filtro de ondaletas	42
Fórmula 33 –MAPE	43
Fórmula 34 –Correlação	43
Fórmula 35 –TIC	43
Fórmula 36 –RMSE.....	44
Fórmula 37 –Estatística teste para correlação	58

ÍNDICE DE FIGURAS

Figura 1 – Decomposição por ondaletas.....	23
Figura 2 – Árvore de decomposição em 3 níveis	23
Figura 3 – Modelo de um neurônio artificial.....	26
Figura 4 – Proposta da pesquisa	37
Figura 5 – Exemplo ilustrado do modelo GARCH	40
Figura 6 – Série temporal do café no período de 02/01/1998 a 30/12/2010 e os respectivos log retornos	46
Figura 7 – Histograma dos log retornos diários do café e retornos diários X distribuição normal.....	47
Figura 8 – Histograma e estatística de Jarque Bera para os log retornos da <i>commoditie</i> de café	48
Figura 9 – Correlograma dos quadrados dos resíduos da <i>commoditie</i> de café.....	51
Figura 10 – Gráfico da <i>commoditie</i> de café real e previsto 4 passos a frente com previsão estatística para Redes Neurais Recorrentes sem o uso de ondaletas	52
Figura 11 – Gráfico da <i>commoditie</i> de café real e previsto 4 passos a frente com previsão estatística para Redes Neurais Recorrentes com o uso de ondaletas	53
Figura 12 – Série do café e os expoentes de Hurst ao longo da série para uma janela de 500 pontos	54
Figura 13 – Correlação em função do expoente de Hurst calculados para uma janela de 500 pontos com o modelo de Redes Neurais Recorrentes com e sem filtro de ondaletas para a <i>commoditie</i> de café	55
Figura 14 – MAPE em função do expoente de Hurst calculados para uma janela de 500 pontos com o modelo de Redes Neurais Recorrentes com e sem filtro de ondaletas para a <i>commoditie</i> de café.....	56
Figura 15 – TIC em função do expoente de Hurst calculados para uma janela de 500 pontos com o modelo de Redes Neurais Recorrentes com e sem filtro de ondaletas para a <i>commoditie</i> de café.....	56
Figura 16 – RMSE em função do expoente de Hurst calculados para uma janela de 500 pontos com o modelo de Redes Neurais Recorrentes com e sem filtro de ondaletas para a <i>commoditie</i> de café	57
Figura 17 – Série da soja e os expoentes de Hurst ao longo da série para uma janela de 500 pontos	65

Figura 18 – Correlação em função do expoente de Hurst calculados para uma janela de 500 pontos com o modelo de Redes Neurais Recorrentes com e sem filtro de ondaletas para a <i>commoditie</i> de soja	66
Figura 19 – MAPE em função do expoente de Hurst calculados para uma janela de 500 pontos com o modelo de Redes Neurais Recorrentes com e sem filtro de ondaletas para a <i>commoditie</i> de soja	66
Figura 20 – TIC em função do expoente de Hurst calculados para uma janela de 500 pontos com o modelo de Redes Neurais Recorrentes com e sem filtro de ondaletas para a <i>commoditie</i> de soja	67
Figura 21 – RMSE em função do expoente de Hurst calculados para uma janela de 500 pontos com o modelo de Redes Neurais Recorrentes com e sem filtro de ondaletas para a <i>commoditie</i> de soja	67
Figura 22 – Série da taxa de câmbio e os expoentes de Hurst ao longo da série para uma janela de 500 pontos	76
Figura 23 – Correlação em função do expoente de Hurst calculados para uma janela de 500 pontos com o modelo de Redes Neurais Recorrentes com e sem filtro de ondaletas para a taxa de câmbio	77
Figura 24 – MAPE em função do expoente de Hurst calculados para uma janela de 500 pontos com o modelo de Redes Neurais Recorrentes com e sem filtro de ondaletas para a taxa de câmbio	77
Figura 25 – TIC em função do expoente de Hurst calculados para uma janela de 500 pontos com o modelo de Redes Neurais Recorrentes com e sem filtro de ondaletas para a taxa de câmbio	78
Figura 26 – RMSE em função do expoente de Hurst calculados para uma janela de 500 pontos com o modelo de Redes Neurais Recorrentes com e sem filtro de ondaletas para a taxa de câmbio	78
Figura 27 – Série do Ibovespa e os expoentes de Hurst ao longo da série para uma janela de 500 pontos	87
Figura 28 – Correlação em função do expoente de Hurst calculados para uma janela de 500 pontos com o modelo de Redes Neurais Recorrentes com e sem filtro de ondaletas para o índice Ibovespa	88
Figura 29 – MAPE em função do expoente de Hurst calculados para uma janela de 500 pontos com o modelo de Redes Neurais Recorrentes com e sem filtro de ondaletas para o índice Ibovespa	88

Figura 30 – TIC em função do expoente de Hurst calculados para uma janela de 500 pontos com o modelo de Redes Neurais Recorrentes com e sem filtro de ondaletas para o índice Ibovespa.....	89
Figura 31 – RMSE em função do expoente de Hurst calculados para uma janela de 500 pontos com o modelo de Redes Neurais Recorrentes com e sem filtro de ondaletas para o índice Ibovespa.....	89
Figura 32 – Série da VALE5 e os expoentes de Hurst ao longo da série para uma janela de 500 pontos	97
Figura 33 – Correlação em função do expoente de Hurst calculados para uma janela de 500 pontos com o modelo de Redes Neurais Recorrentes com e sem filtro de ondaletas para a VALE5	98
Figura 34 – MAPE em função do expoente de Hurst calculados para uma janela de 500 pontos com o modelo de Redes Neurais Recorrentes com e sem filtro de ondaletas para a VALE5	98
Figura 35 – TIC em função do expoente de Hurst calculados para uma janela de 500 pontos com o modelo de Redes Neurais Recorrentes com e sem filtro de ondaletas para a VALE5	99
Figura 36 – RMSE em função do expoente de Hurst calculados para uma janela de 500 pontos com o modelo de Redes Neurais Recorrentes com e sem filtro de ondaletas para a VALE5	99
Figura 37 – Série do ITSA4 e os expoentes de Hurst ao longo da série para uma janela de 500 pontos	107
Figura 38 – Correlação em função do expoente de Hurst calculados para uma janela de 500 pontos com o modelo de Redes Neurais Recorrentes com e sem filtro de ondaletas para o ITSA4	108
Figura 39 – MAPE em função do expoente de Hurst calculados para uma janela de 500 pontos com o modelo de Redes Neurais Recorrentes com e sem filtro de ondaletas para o ITSA4	108
Figura 40 – TIC em função do expoente de Hurst calculados para uma janela de 500 pontos com o modelo de Redes Neurais Recorrentes com e sem filtro de ondaletas para o ITSA4	109
Figura 41 – RMSE em função do expoente de Hurst calculados para uma janela de 500 pontos com o modelo de Redes Neurais Recorrentes com e sem filtro de ondaletas para o ITSA4	109

Figura 42 – Série da PETR4 e os expoentes de Hurst ao longo da série para uma janela de 500 pontos	117
Figura 43 – Correlação em função do expoente de Hurst calculados para uma janela de 500 pontos com o modelo de Redes Neurais Recorrentes com e sem filtro de ondaletas para a PETR4	118
Figura 44 – MAPE em função do expoente de Hurst calculados para uma janela de 500 pontos com o modelo de Redes Neurais Recorrentes com e sem filtro de ondaletas para a PETR4	118
Figura 45 – TIC em função do expoente de Hurst calculados para uma janela de 500 pontos com o modelo de Redes Neurais Recorrentes com e sem filtro de ondaletas para a PETR4	119
Figura 46 – RMSE em função do expoente de Hurst calculados para uma janela de 500 pontos com o modelo de Redes Neurais Recorrentes com e sem filtro de ondaletas para a PETR4	119

ÍNDICE DE TABELAS

Tabela 1 – Teste ADF para os log retornos da <i>commoditie</i> de café	49
Tabela 2 – Valores obtidos pelo teste BDS nas respectivas dimensões para a <i>commoditie</i> de café.....	49
Tabela 3 – Valores obtidos pelo teste de McLeod-Li para a <i>commoditie</i> de café	50
Tabela 4 – Estatística de acurácia das RN recorrentes para a <i>commoditie</i> de café	52
Tabela 5 – Estatística de acurácia das RN recorrentes para a <i>commoditie</i> de café com decomposição de altas e baixas freqüências por ondaletas	53
Tabela 6 – Estatísticas de acurácia das previsões para os Expoentes de Hurst maiores e menores que 0,5 para o café	57
Tabela 7 – Teste de hipóteses de diferença de médias das estatísticas de acurácia com e sem ondaletas para os Expoentes de Hurst maiores e menores que 0,5 para o café	58
Tabela 8 – Varredura das diferentes janelas de Hurst e estatísticas de acurácia da correlação para as previsões com e sem ondaletas e teste de significância para a <i>commoditie</i> de café	60
Tabela 9 – Varredura das diferentes janelas de Hurst e estatísticas de acurácia do MAPE para as previsões com e sem ondaletas e teste de significância para a <i>commoditie</i> de café ..	61
Tabela 10 – Varredura das diferentes janelas de Hurst e estatísticas de acurácia do TIC para as previsões com e sem ondaletas e teste de significância para a <i>commoditie</i> de café.....	62
Tabela 11 – Varredura das diferentes janelas de Hurst e estatísticas de acurácia do RMSE para as previsões com e sem ondaletas e teste de significância para a <i>commoditie</i> de café ..	63
Tabela 12 – Estatísticas de acurácia das previsões para os Expoentes de Hurst maiores e menores que 0,5 para a soja.....	68
Tabela 13 – Teste de hipóteses de diferença de médias das estatísticas com e sem ondaletas para os Expoentes de Hurst maiores e menores que 0,5 para a soja.....	69
Tabela 14 – Varredura das diferentes janelas de Hurst e estatísticas de acurácia da Correlação para as previsões com e sem ondaletas e teste de significância para a <i>commoditie</i> de soja	71
Tabela 15 – Varredura das diferentes janelas de Hurst e estatísticas de acurácia do MAPE para as previsões com e sem ondaletas e teste de significância para a <i>commoditie</i> de soja...72	
Tabela 16 – Varredura das diferentes janelas de Hurst e estatísticas de acurácia do TIC para as previsões com e sem ondaletas e teste de significância para a <i>commoditie</i> de soja.....73	

Tabela 17 – Varredura das diferentes janelas de Hurst e estatísticas de acurácia do RMSE para as previsões com e sem ondaletas e teste de significância para a <i>commoditie</i> de soja...74	
Tabela 18 – Estatísticas de acurácia das previsões para os Expoentes de Hurst maiores e menores que 0,5 para o câmbio79	
Tabela 19 – Teste de hipóteses de diferença de médias das estatísticas de acurácia com e sem ondaletas para os Expoentes de Hurst maiores e menores que 0,5 para o câmbio80	
Tabela 20 – Varredura das diferentes janelas de Hurst e estatísticas de acurácia da Correlação para as previsões com e sem ondaletas e teste de significância para o câmbio82	
Tabela 21 – Varredura das diferentes janelas de Hurst e estatísticas de acurácia da do MAPE para as previsões com e sem ondaletas e teste de significância para o câmbio.....83	
Tabela 22 – Varredura das diferentes janelas de Hurst e estatísticas de acurácia da do TIC para as previsões com e sem ondaletas e teste de significância para o câmbio.....84	
Tabela 23 – Varredura das diferentes janelas de Hurst e estatísticas de acurácia da do RMSE para as previsões com e sem ondaletas e teste de significância para o câmbio.....85	
Tabela 24 – Estatísticas de acurácia das previsões para os Expoentes de Hurst maiores e menores que 0,5 para o Ibovespa90	
Tabela 25 – Teste de hipóteses de diferença de médias das estatísticas de acurácia com e sem ondaletas para os Expoentes de Hurst maiores e menores que 0,5 para o índice Ibovespa91	
Tabela 26 – Varredura das diferentes janelas de Hurst e estatísticas de acurácia da Correlação para as previsões com e sem ondaletas e teste de significância para o Ibovespa92	
Tabela 27 – Varredura das diferentes janelas de Hurst e estatísticas de acurácia do MAPE para as previsões com e sem ondaletas e teste de significância para o Ibovespa93	
Tabela 28 – Varredura das diferentes janelas de Hurst e estatísticas de acurácia do TIC para as previsões com e sem ondaletas e teste de significância para o Ibovespa94	
Tabela 29 – Varredura das diferentes janelas de Hurst e estatísticas de acurácia do RMSE para as previsões com e sem ondaletas e teste de significância para o Ibovespa95	
Tabela 30 – Estatísticas de acurácia das previsões para os Expoentes de Hurst maiores e menores que 0,5 para a VALE5.....100	
Tabela 31 – Teste de hipóteses de diferença de médias das estatísticas de acurácia com e sem ondaletas para os Expoentes de Hurst maiores e menores que 0,5 para as ações da VALE5100	
Tabela 32 – Varredura das diferentes janelas de Hurst e estatísticas de acurácia da Correlação para as previsões com e sem ondaletas e teste de significância para a VALE5102	

Tabela 33 – Varredura das diferentes janelas de Hurst e estatísticas de acurácia do MAPE para as previsões com e sem ondaletas e teste de significância para as ações da VALE5...103	
Tabela 34 – Varredura das diferentes janelas de Hurst e estatísticas de acurácia do TIC para as previsões com e sem ondaletas e teste de significância para a VALE5104	
Tabela 35 – Varredura das diferentes janelas de Hurst e estatísticas de acurácia do RMSE para as previsões com e sem ondaletas e teste de significância para as ações da VALE5...105	
Tabela 36 – Estatísticas de acurácia das previsões para os Expoentes de Hurst maiores e menores que 0,5 para o ITSA4110	
Tabela 37 – Teste de hipóteses de diferença de médias das estatísticas de acurácia com e sem ondaletas para os Expoentes de Hurst maiores e menores que 0,5 para o ITSA4.....110	
Tabela 38 – Varredura das diferentes janelas de Hurst e estatísticas de acurácia da Correlação para as previsões com e sem ondaletas e teste de significância para as ações do ITSA4112	
Tabela 39 – Varredura das diferentes janelas de Hurst e estatísticas de acurácia do MAPE para as previsões com e sem ondaletas e teste de significância para as ações do ITSA4113	
Tabela 40 – Varredura das diferentes janelas de Hurst e estatísticas de acurácia do TIC para as previsões com e sem ondaletas e teste de significância para o ITSA4114	
Tabela 41 – Varredura das diferentes janelas de Hurst e estatísticas de acurácia do RMSE para as previsões com e sem ondaletas e teste de significância para as ações do ITSA4115	
Tabela 42 – Estatísticas de acurácia das previsões para os Expoentes de Hurst maiores e menores que 0,5 para a PETR4120	
Tabela 43 – Teste de hipóteses de diferenças de médias das estatísticas de acurácia com e sem ondaletas para os Expoentes de Hurst maiores e menores que 0,5 para as ações da PETR4120	
Tabela 44 – Varredura das diferentes janelas de Hurst e estatísticas de acurácia da Correlação para as previsões com e sem ondaletas e teste de significância para as ações da PETR4122	
Tabela 45 – Varredura das diferentes janelas de Hurst e estatísticas de acurácia do MAPE para as previsões com e sem ondaletas e teste de significância para as ações do PETR4123	
Tabela 46 – Varredura das diferentes janelas de Hurst e estatísticas de acurácia do TIC para as previsões com e sem ondaletas e teste de significância para a PETR4124	
Tabela 47 – Varredura das diferentes janelas de Hurst e estatísticas de acurácia do RMSE para as previsões com e sem ondaletas e teste de significância para a PETR4.....125	

Sumário

RESUMO.....	VII
ABSTRACT	VIII
LISTAS DE ABREVIATURAS E SIGLAS	IX
ÍNDICE DE FÓRMULAS	XI
ÍNDICE DE FIGURAS	XIII
ÍNDICE DE TABELAS	XVII
1. INTRODUÇÃO	1
1.1 Problema de Pesquisa	3
1.2 Objetivo.....	3
1.3 Metodologia	4
2. REVISÃO TEÓRICA	5
2.1 Breve Histórico sobre Risco	5
2.2 Risco e Séries Temporais.....	6
2.3 Séries Temporais.....	7
2.4 Análise de Séries Temporais.....	8
2.4.1 Estacionariedade de Séries Temporais	8
2.4.2 Verificação de Presença de Heterocedasticidade Condicionada	10
2.5 Modelos de Previsão de Séries Temporais	11
2.5.1 Modelo Autorregressivo (AR).....	12
2.5.2 Modelo de Médias Móveis (MA)	13
2.5.3 Modelo Autorregressivo de Média Móvel (ARMA).....	14
2.5.4 Modelo Autorregressivo Integrado e de Média Móvel (ARIMA)	15
2.5.5 O Modelo de ARCH	17
2.5.6 O Modelo de GARCH	18
2.6 Definição de Filtro	22
2.6.1 Filtro de Ondaletas	22
2.6.2 Previsão de Séries Temporais com Modelos GARCH e <i>Ondaletas</i>	24
2.7 Redes Neurais	25
2.7.1 Previsão de Séries Temporais com Modelos GARCH e Redes Neurais Artificiais	27
2.7.2 Previsão de Séries Temporais com Modelos de Redes Neurais Artificiais e <i>Ondaletas</i>	28
2.7.3 Previsão de Séries Temporais com Modelos GARCH, Ondaletas e Redes neurais Artificiais	29

2.8 Expoente de Hurst.....	30
2.8.1 Expoente de Hurst e Modelos de GARCH em Séries Temporais	32
2.8.2 Expoente de Hurst e <i>Ondaletas</i> em Séries Temporais	32
2.8.3 Expoente de Hurst e Redes Neurais Artificiais em Séries Temporais	34
2.9 Contribuição da Pesquisa.....	34
3. METODOLOGIA.....	35
3.1 Natureza da Pesquisa	35
3.2 Operacionalização da Pesquisa	36
3.2.1 Operacionalização do modelo de Hurst.....	37
3.2.2 Operacionalização do modelo GARCH (1,1).....	39
3.2.3 Operacionalização do modelo de Redes Neurais	41
3.2.4 Operacionalização dos Filtros por Ondaletas	41
3.3 Estatísticas de Erro e Confiança para as previsões	43
3.4 Dados	44
4. RESULTADOS EXPERIMENTAIS	45
4.1 Análise da série da <i>commoditie</i> de café.....	45
4.2 Análises de outras séries observando o comportamento da qualidade da previsão e o Expoente de Hurst.....	64
4.2.1 Análise <i>Commoditie</i> de Soja.....	64
4.2.2 Análise da taxa de câmbio	75
4.2.3 Análise Do Ibovespa.....	86
4.2.4 Análise Da Vale do Rio Doce (VALE5)	96
4.2.5 Análise Itaú (ITSA4)	106
4.2.6 Análise Petrobrás (PETR4)	116
CONCLUSÃO.....	127
REFERÊNCIAS BIBLIOGRÁFIAS	129

1. INTRODUÇÃO

A literatura sobre modelos de previsão de séries temporais financeiras já identificou que, possivelmente, são os modelos não lineares, como os da família ARIMA-GARCH e redes neurais, que melhor apresentam ajustes aos dados e capacidade preditiva, ou seja, menor erro de previsão no que diz respeito às séries temporais financeiras (TSAY, 2005, p. 21).

Porém, este fato já é notório, não existe ainda um circuito fechado e já devidamente testado e validado pelo mercado financeiro como o modelo a ser seguido para se realizar previsão de séries temporais financeiras.

Ressalta-se ainda que nem essa gigantesca presença de tais modelos na literatura foi capaz de explicar crises no mercado financeiro e falhas nos modelos de previsão. Mas um consenso de mercado é que se os modelos ainda não são capazes de prever grandes lucros ou perdas, sem estes modelos seria muito pior. Haja vista que os investidores no mercado precisam conhecer as previsões futuras para que possam analisar suas posições nos respectivos mercados futuros. As previsões se encaixam neste contexto tanto para investidores individuais como empresas e, principalmente, na área de controladoria e finanças.

O mercado também já sinalizou que os modelos precisam de ajustes e até mesmo combinações de técnicas para auxiliar a capacidade dos modelos em lidarem com o imprevisto (GOOIJER E HYNDMAN, 2006, p. 454), principalmente, pelo fato das séries temporais apresentarem comportamentos distintos ao longo do tempo, com períodos que melhor colaboraram para as previsões por serem menos voláteis e períodos onde ocorre o inverso.

O *background* levantado nas referências internacionais aponta que existe uma técnica de medição de uma janela temporal para se identificar um tipo de comportamento existente em uma série temporal. Tal técnica é conhecida como expoente de Hurst. É uma medida que qualifica a série como persistente ou anti-persistente.

A estimação da evolução temporal da memória de longo prazo dos retornos permite a avaliação da evolução da eficiência do mercado de ações e a identificação da reação dessa eficiência a eventos que podem afetar os mecanismos do processo gerador de preço. Além disso permite a melhoria na avaliação do risco associado ao ativo

analisado, pois a volatilidade é uma medida importante de risco (SOUZA, TABAK E CAJUEIRO, 2006).

Assim surge-se a curiosidade: Será que verificando, com a ajuda do expoente de Hurst, numa série temporal financeira que apresenta memória longa, a volatilidade da previsão utilizando os melhores métodos existentes na atualidade para tal fim, que são o modelo de GARCH e Redes Neurais Recorrentes, essa volatilidade diminui? E filtrando-se a série com filtros de ondaletas, melhora ou piora essa volatilidade?

Como Lima (2004) afirma que o modelo de Redes Neurais Recorrentes é superior ao modelo de GARCH para realizar-se pesquisas desse tipo, utilizaremos o modelo de Redes Neurais nesta pesquisa. Este trabalho procura melhorar o erro de previsão de séries temporais financeiras analisando o grau de persistência de tais séries. Inicialmente aplica-se o expoente de Hurst em tal série e depois testa-se o modelo: de Redes Neurais Recorrentes com e sem filtro de ondaletas. Por fim aplica-se os medidores de qualidade de previsão MAPE, TIC, correlação e RMSE. Esses modelos, de acordo com Gooijer e Hyndman (2006), são os mais eficientes para se medir erro nos trabalhos de previsão de série temporal, eles indicam o quanto o modelo de previsão é capaz de reproduzir os dados já conhecidos.

E assim, com esta inspiração e motivação, a pesquisa é desenvolvida, dividida nas seguintes etapas: relata-se o problema de pesquisa, assim como seus objetivos, um breve histórico sobre a existência do risco e séries temporais e uma pesquisa bibliográfica sobre os modelos utilizados no trabalho e mais aceitos na literatura, que são Modelo de GARCH, Redes Neurais, Ondaletas e Expoente de Hurst. A seguir é mostrada a metodologia que foi explicada para se chegar aos resultados, os resultados obtidos, a análise dos dados e, por fim, a conclusão.

A seguir apresenta-se o problema de pesquisa.

1.1 Problema de Pesquisa

O problema de pesquisa pode ser então enunciado: A divisão de séries temporais em janelas de tempo definidas pelo expoente de Hurst melhora ou não a capacidade preditiva dos modelos?

Em outras palavras, serão calculadas janelas de tempo e em cada uma delas feitos os cálculos do expoente de Hurst e das previsões com o modelo de redes neurais. Posteriormente, será analisado o resultado das previsões correlacionando-os ao expoente de Hurst calculado para cada janela. Pretende-se analisar se onde a série é persistente, houve melhora ou piora nas previsões? Idem para a série anti-persistente.

1.2 Objetivo

O objetivo geral é verificar se o expoente de Hurst, aqui usado para uma classificação da qualidade dos dados em cada janela, interfere na qualidade das previsões feitas. Em outras palavras, pretende-se verificar se a influência da janela para determinação da persistência ou não nos dados melhora ou não a capacidade preditiva dos modelos já tradicionalmente testados.

Como objetivo secundário tem-se a identificação do tamanho de uma janela onde se pode reduzir o erro de previsão e consequente garantia de presença de memória nos dados.

Este estudo visa ainda uma contribuição na área contábil, na parte de mercados futuros, à vista de ações e opções estudando: (a) o comportamento dos ativos financeiros, (b) analisando séries temporais financeiras variando o tamanho da janela usada para o cálculo do expoente de Hurst e (c) observando sua volatilidade.

Contribui, dessa forma, para um aperfeiçoamento dos modelos de gestão gerencial de contratos e aplicações futuras.

Vale lembrar ainda que as previsões de séries temporais financeiras sempre estão associadas a uma definição e tentativa de redução do risco envolvido nas operações financeiras. Uma previsão definida dentro de um certo intervalo de confiança trará ao investidor mais tranquilidade ao assumir determinada posição no mercado.

1.3 Metodologia

A metodologia da pesquisa é quantitativa de base econométrica pela abordagem do problema: por realizar previsões de dados com modelos econôméticos e de redes neurais. A pesquisa ainda é descritiva quanto aos objetivos da pesquisa por fazer uma descrição das técnicas de previsão e identificação das janelas pelo expoente de Hurst.

É também uma pesquisa bibliográfica por fazer uma revisão dos principais trabalhos presentes na literatura e em livros que abordam o assunto para ajudar a obtenção da resposta ao problema de pesquisa.

Serão aplicados tais modelos aos principais índices de bolsa de valores e a *commodities* para verificar como se dá o comportamento da detecção da janela e o efeito sobre o cálculo das previsões.

Para realização das previsões pretende-se ainda testar o uso do filtro de ondaletas conforme Lima (2004, p.77) identificou por melhorar a qualidade das previsões com o modelo de redes neurais recorrentes.

As ondaletas aqui são usadas como um filtro redutor de ruídos e facilitador das previsões com redução dos erros finais. Não constituem o foco principal da pesquisa. Serão usados, algoritmos em Matlab para se obterem as devidas previsões.

Detalhes técnicos da metodologia são apresentados no capítulo 3.

A seguir, apresenta-se a revisão teórica necessária para dar suporte a esta pesquisa.

2. REVISÃO TEÓRICA

Conforme explicado no fim do capítulo anterior, a previsão de séries temporais financeiras coaduna-se com a tentativa de se limitarem os riscos assumidos em uma posição no mercado financeiro. Assim, trata-se a seguir de um breve histórico sobre o risco, como ponto inicial das previsões de séries temporais.

2.1 Breve Histórico sobre Risco

O estudo do risco, tem suas raízes mais profundas no século XVI, quando Méré desafiou Pascal a resolver um problema de um jogo de azar. O enigma era como dividir as apostas de um jogo de azar entre dois jogadores, que foi interrompido quando um deles estava ganhando. Como Pascal se confundiu com a pergunta, foi pedir ajuda a seu amigo Pierre Fermat, que, apesar de ser advogado, era brilhante com números (BERNSTEIN,1997).

A partir desta passagem e da resolução de tal problema, Pascal e Fermat iniciaram os estudos e a descoberta da teoria das probabilidades. Esta solução permitiu, pela primeira vez que pessoas tomassem decisões e pudessem tentar antever o futuro com a ajuda dos números. Vários estudos vem sendo feitos e há séculos pesquisadores tentam lidar com uma questão básica: o que é risco? Como mensurá-lo?

No século XVII, Gottfried Von Leibniz conversando com o matemático suíço Jacob Bernoulli afirmou que: “a natureza estabeleceu padrões que dão origem ao retorno dos eventos, mas apenas na maior parte dos casos”. Esta afirmação “maior parte dos casos”, mostrou a essência da existência do risco, pois sem este comentário tudo poderia ser previsto onde não há mudanças de comportamentos, ou seja, onde cada evento acontece igualmente ao anterior.

Assim, estas descobertas levaram Bernoulli a descobrir a Lei dos Grandes Números e os métodos de amostragem estatística, que estão até hoje presentes em várias pesquisas feitas no mundo, como por exemplo: pesquisas de conceitos sobre certos assuntos e ações de novos produtos de mercado.

Ainda no século XVII, Abraham de Moivre criou a curva em sino, mais conhecida como distribuição normal e o conceito de desvio padrão. Esses dois conceitos formaram a Lei das Médias e são importantíssimos para calcular o risco pelas técnicas mais modernas que temos na literatura. Essa descoberta genial foi sintetizada pelo autor

da seguinte forma: “ *Embora o acaso produza irregularidades, as chances serão infinitamente grandes de que, no decorrer do tempo, essas irregularidades não guardem nenhuma proporção com a repetição daquela ordem que é fruto naturalmente do projeto original* ”. (BERNSTEIN,1997)

Gauss se interessou pelo modo que Moivre colocou suas idéias de probabilidade em sino e criou a estrutura que utilizamos hoje em dia. Em verdade o método da curva em sino, não tem por principal objetivo indicar a exatidão, mas sim o erro. Se cada estimativa feita fosse totalmente correta, tudo estaria resolvido. No entanto nenhuma observação individual é um exemplo perfeito da realidade. Ao apresentar uma distribuição normal é como se colocássemos toda amostra em ordem. Francis Galton, louvou tal distribuição da seguinte forma: “ *A “Lei da Frequência dos Erros” ... reina serenamente e com completa discrição em meio a maior das confusões. Quanto maior a turba... mais perfeito é o domínio dela. É a lei suprema da irracionalidade. Sempre que uma grande amostra de elementos caóticos é levada em consideração... uma forma insuspeitada e belíssima de regularidade mostra que esteve latente o tempo todo* ”. (BERNSTEIN, 1997)

Pouco mais tarde Daniel Bernoulli, deu um grande salto nas regras de probabilidades demonstrando como tomar decisões de modo mais perfeito levando em consideração informações novas e antigas.

Após tais avanços, Francis Galton, em 1875, descobriu a regressão à média. Tal teoria afirma que quando lança-se mão de uma decisão baseada na expectativa de que as coisas voltarão ao normal, emprega-se a idéia da regressão à média.

Já no século XIX, Harry Markovitz ganhou um prêmio Nobel ao provar que o risco diminui ao se diversificar uma carteira de ativos, ou seja, para um investidor que deixar de investir em apenas uma empresa para investir em várias a chance de perdas é menor.

2.2 Risco e Séries Temporais

Desde o século XIII já se tinha uma noção de risco. Os mercadores aprenderam desde cedo a usar a diversificação para diminuir o risco de seus negócios. Antônio, o mercador de Veneza na peça de Shakespeare, já mencionava: “ *Todas as minhas cargas não estão confiadas a um só navio, nem as dirijo para um só ponto; nem o total*

de meus bens está a mercê dos contratempos do presente ano. Não são pois, minhas especulações que me fazem ficar triste” (Ato I, Cena 1), (BERNSTEIN, 1997).

Começou-se a se pensar na análise de risco, ou seja, no comportamento dos preços dos retornos ou nas probabilidades de se perder em certo negócio, um pouco mais tarde. Uma série temporal, definida na seção 2.3, é um exemplo de um fenômeno colhido sequencialmente ao longo do tempo, onde se corre algum tipo de riscos dos valores baixarem ou subirem devido a determinados acontecimentos no tempo. Portanto a análise de séries temporais está associada à habilidade de identificar, antecipadamente o comportamento do risco. Assim a próxima seção deste trabalho detalha a definição para séries e as análises das mesmas que serão feitas ao longo da pesquisa.

2.3 Séries Temporais

Box, Jenkins e Reinsel (1994) definem uma série temporal como um conjunto de observações dos resultados de um fenômeno colhido sequencialmente ao longo do tempo.

Qualquer grupo de observações ordenadas no tempo é denominado Série Temporal. São exemplos: os preços diários de uma determinada ação na bolsa de valores, os valores mensais da temperatura em uma determinada região do país, a produção anual de um determinado produto no Brasil e os registros de marés no porto de uma determinada cidade litorânea do mundo.

Com um conjunto de observações colhido ao longo do tempo, pode-se fazer uma abordagem matemática para a análise dos dados. Esta abordagem pode descrever a série por meio de equações que modelam os mecanismos necessários para a geração da série e demonstra como esse comportamento evolui.

Uma série temporal pode ser discreta ou contínua. Diz-se que ela é discreta quando os intervalos de tempos da análise são iguais. Quando menciona-se o exemplo dos preços diários de uma determinada ação na bolsa de valores, percebe-se que esta série temporal é discreta pelo fato do tempo ser diário, ou seja, o intervalo de tempo da série é idêntico. Já uma série temporal contínua, não tem os intervalos de tempo iguais, é o caso da análise dos registros de marés no porto de uma determinada cidade litorânea do mundo.

Além da determinação da série em relação ao seu intervalo de tempo, uma série temporal pode ser medida pelo número de parâmetros que ela contém. Quando se tem

um número finito de parâmetros, ela é paramétrica e se for delimitada por um domínio de freqüências, é não paramétrica.

Muitas vezes não há funções matemáticas que determinam a sequência de uma série temporal perfeitamente, mostrando apenas seu percurso em termos de uma distribuição probabilística; quando isso acontece a série é não determinística. Já quando há uma função matemática que determina exatamente a série ela é determinística.

As séries temporais são amplamente usadas para análise. Existem muitos modelos na literatura que calculam a previsão para os próximos elementos das séries através de funções matemáticas, filtram para um melhor entendimento de seu percurso, ou analisam sua dependência temporal, dentre outros comportamentos.

2.4 Análise de Séries Temporais

Morettin e Toloi (2006) afirmam que dentre os objetivos das análises de séries temporais, destacam-se: a investigação do mecanismo gerador da série temporal, fazer previsões de valores futuros da série, descrever apenas o comportamento da série e procurar periodicidades relevantes nos dados.

Neste trabalho faz-se uma análise para se prever valores futuros da série e entender seu comportamento. Para efetuar esta análise são utilizados modelos probabilísticos ou modelos estocásticos, em um domínio temporal ou de frequências.

A previsão utilizando métodos quantitativos pode ser feita, por exemplo através de vários modelos de regressão. O emprego de cada modelo depende basicamente do comportamento da série temporal que se deseja analisar (PELLEGRINI e FOGLIATTO, 2001).

Muitas análises feitas para séries temporais se baseiam em sua estacionariedade, ou seja, partindo-se do pressuposto de que ela se desenvolve no tempo ao redor de uma média constante. Outras análises são feitas para descrever séries não estacionárias que são modelos autorregressivos integrados à médias móveis.

2.4.1 Estacionariedade de Séries Temporais

Os métodos usados em séries temporais baseiam-se fortemente nos conceitos de estacionariedade, de filtro linear, ergodicidade e independência. Assim, a maioria dos modelos considerados são lineares estacionários fracos, como modelos autoregressivos

(AR), de médias móveis (MA) e mistos (ARMA). Em finanças, existem modelos não estacionários e não lineares. Para se descrever séries não estacionárias com comportamento não explosivo, usam-se modelos autoregressivos integrados de médias móveis – ARIMA (MORETTIN e TOLOI, 2006).

Enders (2003) afirma que um processo estocástico é estacionário se suas média e variância forem constantes ao longo do tempo e o valor da covariância entre dois períodos de tempo depender apenas da distância ou defasagem entre os dois períodos e não do período de tempo efetivo em que a covariância é calculada.

Para identificar a estacionariedade de séries temporais, um instrumento foi proposto por Box, Jenkins e Reinsel (1994), este instrumento é a função de autocorrelação parcial (FACP), obtida pelas estimativas do valor ρ_k como uma função de k

(1)

Sendo que ρ_k é a matriz de autocorrelações para a defasagem k e ρ_0 é a matriz com a última coluna substituída pelo vetor de autocorrelações (MORETTIN, 2002).

A função de autocorrelação parcial (FACP) mede a correlação entre os termos ϵ_t e ϵ_{t-k} depois de eliminada a influência dos elementos intermediários, ao passo que a função de autocorrelação (FAC) calcula apenas a correlação entre ϵ_t e ϵ_{t-1} (MORETTIN e TOLOI, 2004).

Morettin (2002) propõe um resumo do comportamento das FAC e FACP, em relação a características dos modelos AR, MA e ARMA:

Processo	FAC	FACP
AR(p)	Decai, de acordo com exponenciais e/ou senóides amortecidas, infinita em extensão.	Apresenta um corte em p, sendo que o valor para os valores de e sendo para .
MA(q)	É finita, no sentido de que ela apresenta um truncamento (corte) na defasagem q.	Decai, de acordo com exponenciais e/ou senóides amortecidas, infinita em extensão.
ARMA(p,q)	Infinita em extensão, decai, de acordo com exponenciais e/ou senoidais amortecidas após a defasagem q - p.	Decai, de acordo com exponenciais e/ou senóides amortecidas, infinita em extensão.

Quadro 1 Comportamento das funções de FAC e FACP em relação aos modelos de AR, MA e ARMA descritos por Morettin (2002).

2.4.2 Verificação de Presença de Heterocedasticidade Condicionada

Existem dois testes para se verificar se uma série temporal apresenta heterocedasticidade condicionada: o teste de Box-LJung (LB) para a série e o teste de multiplicadores de Lagrange (LM), sendo a hipótese nula para todo na regressão para . A estatística de teste é que segue uma distribuição assintótica com r graus de liberdade em que n é o tamanho da amostra e é o quadrado do coeficiente de correlação múltipla da regressão acima (MORETTIN, 2002).

O teste mostra se há correlação entre os resíduos quadrados. As respostas são dadas da seguinte forma:

: mostra que não há efeitos de ARCH (*autoregressive conditional heterocedasticity*) no modelo analisado.

: mostra que há efeitos de ARCH (*autoregressive conditional heterocedasticity*) no modelo analisado.

2.5 Modelos de Previsão de Séries Temporais

Os modelos probabilísticos de uma série temporal são um método sofisticado de ultrapassar os limites conhecidos dos dados que temos inicialmente na série. Para escolher o modelo ideal para o que se quer, cabe ao usuário dos modelos saber escolher ou criar de acordo com seus objetivos.

Um modelo que descreve uma série não conduz, necessariamente, a um procedimento (ou fórmula) de previsão. Será necessário especificar uma função perda, além do modelo, para se chegar ao procedimento. (MORETTIN E TOLOI, 2006).

Esta função perda, ou seja, o erro do modelo de previsão, mostra o quanto a série normalmente se distorce em relação ao real. Esta estimativa é importante para que o analista se policie a uma possível oscilação da série.

A palavra previsão sugere que tome presciênciade algo, ou seja, saiba de algo que irá acontecer no futuro. DINIZ, et al (1998) afirma que “Observar os dados passados de um fenômeno para antecipar o seu comportamento futuro constitui a essência da previsão.”

Para se prever algo antes que ela exista, ou seja, fazer uma previsão, existem muitos modelos probabilísticos, no entanto é importante dizer que não há nenhuma 100% certa, por isso a palavra previsão utilizada neste trabalho é no sentido de fornecer mais informações para a tomada de decisões e não uma certeza absoluta do que irá acontecer.

Na literatura existem vários modelos para se calcular a volatilidade. O mais simples é o desvio padrão histórico, que considera um peso igual para todas as observações, onde todos os desvios quadráticos dos dados tem o mesmo peso em relação à média. Já a volatilidade exponencialmente ponderada (Suavização exponencial) atribui um peso maior para as observações que são mais recentes. Entretanto esta análise tem a falha de escolher arbitrariamente o peso usado. Os modelos da família GARCH não escolhem arbitrariamente o peso usado e dão ênfase nos passados mais próximos, por isso eles são muito utilizados na literatura para modelos de previsão (MOTA e FERNANDES, 2004).

2.5.1 Modelo Autorregressivo (AR)

No processo autorregressivo de ordem p a observação corrente é gerada por uma média ponderada de observações passadas que recua p períodos, junto com uma perturbação aleatória no período corrente. Denotamos esse processo por AR(p) e escrevemos a equação como (PINDYCK e RUBINFELD, 2004):

(2)

Considerando μ como um termo constante que se relaciona à média do processo estocástico.

Se o processo autorregressivo é estacionário, sua média, que denotamos por μ , deve ser invariável em relação ao tempo; isto é,

A média μ é assim dada por (PINDYCK e RUBINFELD, 2004):

ou

(3)

A condição necessária (não suficiente) para a estacionariedade é:

A versão mais comum de um modelo AR é a que μ depende somente de y_{t-p} , de ϵ_t e de uma constante c . Com estes pré requisitos o modelo se denomina autorregressivo de ordem 1, indicado por AR(1). Segundo Vasconcelos (2000), o modelo algébrico desse modelo é:

(4)

Esse processo tem a média:

(5)

Onde θ é um parâmetro a ser estimado e para $t \neq s$.

Por se tratar de um modelo fracamente estacionário, a variância de (ϵ_t) deve ser constante e as autocovariâncias (ϵ_t) devem ser aproximadamente independentes de t .

2.5.2 Modelo de Médias Móveis (MA)

O método consiste em calcular a média das últimas n observações mais recentes. O valor encontrado, ao seu tempo, é considerado a previsão para o próximo período. (MAKRIDAKIS; et al, 1998).

Pindyck e Rubinfeld (2004) dizem que no processo de médias móveis de ordem q cada observação de y_t é gerada por uma média ponderada de perturbações aleatórias que vão para trás q períodos. Denotamos esse processo como MA(q) e escrevemos a equação como:

(6)

Onde os parâmetros $\alpha_1, \dots, \alpha_q$ podem ser positivos ou negativos. E o erro aleatório é: ϵ_t , é a estimativa de y_t .

Em um modelo de média móvel supõe-se que as oscilações aleatórias sejam distribuídas independentemente ao longo do tempo.

O modelo mais simples de média móvel é o de primeira ordem, denominado MA(1) que é descrito por:

(7)

Como aspectos positivos Tubino (2000) cita que o processo de médias móveis usa dados passados de um número q predeterminado de períodos, normalmente os mais recentes para gerar sua previsão. A cada novo período de previsão o dado mais antigo é substituído pelo mais recente.

Como falhas do modelo, Stengel (1986) diz que as médias móveis são indicadores chamados seguidores de tendências. Esses indicadores possuem uma inércia natural, ou seja, não foram projetados para apontar reversões rapidamente. E Mentzer e

Bienstock (1998) afirmam que o problema com o modelo de médias móveis é que ele se relaciona com a escolha do número de períodos que serão utilizados na previsão.

Na literatura não existem passos exatos para se determinar o período do modelo, para se fazer uma escolha é necessário cautela, pois quanto maior o período, mais suave é o comportamento da média e assim menor o número de ruídos e movimentos curtos a série temporal apresentará. Por outro lado se o período escolhido for pequeno demais a média estará muito exposta à variação dos valores, perdendo uma de suas principais funções que é seguir tendências da série.

2.5.3 Modelo Autorregressivo de Média Móvel (ARMA)

Muitos processos aleatórios não podem ser modelados puramente como médias móveis (MA) ou puramente como autorregressivos (AR), pois possuem qualidades de ambos os tipos de processo. A extensão lógica dos modelos apresentados nas duas últimas seções é o processo misto autorregressivo e de médias móveis de ordem (p,q). Denotamos esse processo como ARMA(p,q) e ele é representado por: (PINDYCK e RUBINFELD, 2004)

(8)

Pressupomos que esse processo é estacionário, e assim sua média é constante ao longo do tempo e é dada pela parcela do processo AR, ou seja:

ou

(9)

Isso nos dá a condição necessária para que o processo seja estacionário, isto é:

(10)

O processo ARMA mais simples é o ARMA(1,1), que é descrito da seguinte forma:

(11)

Box e Jenkins diz que um processo estocástico estacionário, por possuir média, variância e autocorrelação invariantes em relação ao tempo, pode ser otimamente representado por um modelo autorregressivo e/ou médias móveis – ARMA (p,q) – obtido por intermédio da passagem de uma série ruído branco por um filtro linear, o que significa que a série resultante poderá ser vista como uma combinação linear dos termos da série original. O processo resultante dessa passagem, considerando-se este filtro como estável, também será estacionário (RIBEIRO e PAULA, 2000).

2.5.4 Modelo Autorregressivo Integrado e de Média Móvel (ARIMA)

Este modelo é também conhecido como modelo de Box-Jenkins univariados, em homenagem aos dois pesquisadores que mais contribuíram para a construção e aplicação do modelo. Este método consiste em ajustar modelos autorregressivos integrados de médias móveis, ARIMA (p,d,q) a um conjuntos de observações.

O modelo de ARIMA contém: um modelo autorregressivo de ordem p, AR(p), usado quando há autocorrelação entre os dados; um modelo de média móvel de ordem (q), MA(q): utilizado quando existe autocorrelação entre os resíduos da série; um modelo autorregressivo e de média móvel: ARMA(p,q): utilizado quando existe autocorrelação entre os dados e os resíduos; um modelo autorregressivo integrado e de média móvel: ARIMA(p,d,q), utilizado em séries não estacionárias.

De acordo com Morettin e Toloi (2006) se $\Delta^d Y_t$ for estacionária, podemos representar $\Delta^d Y_t$ por um modelo ARMA (p,q), ou seja,

(12)

Se $\Delta^d Y_t$ for uma diferença de d , então $\Delta^d Y_t$ é uma *integral* de Y_t , daí dizemos que Y_t segue um modelo auto-regressivo, *integrado*, de médias móveis, ou modelo ARIMA de ordem (p,d,q),

(13)

Escrevemos ARIMA (p,d,q) , se p e q são as ordens de autoregressão e de diferenças, respectivamente. No modelo da equação (12) todas as raízes de λ estão fora do círculo unitário. Escrever (13) é equivalente a escrever

(14)

Onde ∇^d é um operador auto-regressivo não-estacionário, de ordem $p+d$, com d raízes iguais a um (sobre o círculo unitário) e as restantes p fora do círculo unitário, ou seja,

(15)

Assim, a equação (13) supõe que a d -ésima diferença da série y_t pode ser representada por um modelo ARMA, estacionário e invertível.

O modelo ARIMA (p,d,q) é um caso especial de um processo integrado. Em geral, diz-se que $\nabla^d y_t$ é *integrado de ordem d* se $\nabla^d y_t$ for estacionário, e escrevemos $\nabla^d y_t = I_d y_t$. Se $I_d y_t$ é estacionário.

Xu, Qi e Hua (2009) fazem uma pesquisa sobre a previsão de demanda de *commodities* para períodos após catástrofes naturais. Para isso os autores pegam como base de dados a demanda de produtos agrícolas na China após as tempestades do inverno chinês. Para fazer a previsão eles condensam o método de decomposição empírica (EMD) com o modelo de ARIMA. E constatam que a mistura destes dois modelos melhoram a previsão de demanda de *commodities* após catástrofes naturais.

Usando o método de AR, MA e ARIMA. Wang e Huang (2010) tenta prever o efeito chicote de demanda com os modelos de ARIMA $(1,0,0)$, ARIMA $(1,0,1)$ e ARIMA $(0,0,1)$. Com o controle do efeito por redes varejistas, ou seja, fazendo-se uma previsão deste efeito pelo modelo ARIMA, os varejistas podem traçar melhores parâmetros de funcionamento de suas distribuições de logística. Além disso a aplicação dos parâmetros de previsão investiga os efeitos de flutuação na oferta e demanda de produtos.

Werner e Ribeiro (2003) utilizaram a metodologia de Box-Jenkins para analisar dados históricos de uma empresa de assistência técnica de computadores pessoais e

obter previsões do número de atendimentos. A empresa apresentava 3 tipos de clientes diferentes: contratos, garantia e avulsos. Assim eles utilizaram o modelo AR(1) para prever o número de clientes tipo contrato, ARIMA (2,1,0) para clientes tipo garantia e o modelo sazonal de ARIMA (0,1,0) e (0,1,1) para os clientes tipo avulsos. Com o trabalho os autores concluíram que para os clientes tipo contrato a influência temporal é reduzida, pois a demanda e a concorrência podem crescer na mesma taxa, ou seja, o número de contratos da empresa varia pouco durante o ano. Já os clientes tipo garantia revelam maior dependência temporal, pois as alterações no mercado e os acordos de assistência técnica se refletem gradativamente nos negócios da empresa. Já os clientes tipo avulsos apresentam sazonalidade anual em função dos hábitos de aquisição, os quais estão muito interligados com períodos do ano. Assim o uso destes métodos foi de grande utilidade para a empresa analisada, mostrando uma grande proximidade entre a previsão e a realidade.

Até agora foram mostrados modelos lineares dos tipos: AR, MA, ARMA e ARIMA, que não são adequados para descrever séries temporais que apresentam a variância condicional evoluindo no tempo. Um exemplo de séries deste tipo são as séries financeiras. Para este tipo de série temporal utilizam-se modelos da família de ARCH e suas generalizações que são descritos no próximo tópico.

2.5.5 O Modelo de ARCH

Para a análise de séries temporais financeiras, existem vários modelos não lineares disponíveis na literatura; no entanto, este trabalho se concentrará no modelo de ARCH (*autoregressive conditional heteroscedasticity*). Este modelo (e suas extensões) foi construído por Engle em 1982, com o propósito de estimar a variância da inflação.

O objetivo deste modelo é modelar a variância condicional de uma variável, ou seja, sua volatilidade. Esta variância não pode ser medida diretamente, no entanto ela pode aparecer de várias maneiras como afirma Peña et al (2001):

- 1) A volatilidade aparece em grupos, de maior ou menor volatilidade;
- 2) A volatilidade evolui continuamente no tempo, podendo ser considerada estacionária;
- 3) Ela reage de modo diferente a valores positivos ou negativos da série.

A idéia básica do modelo é que o retorno r_t é não correlacionado serialmente, mas a volatilidade depende de retornos passados por meio de uma função quadrática. Um modelo ARCH(r) é definido por (PINDYCK e RUBINFELD, 2004):

$$\sigma_t^2 = \alpha_0 + \alpha_1 r_{t-1}^2 + \dots + \alpha_p r_{t-p}^2,$$

(16)

Onde $\{r_t\}$ é uma sequência de variáveis aleatórias independentes e identicamente distribuídas (i.i.d) com média zero e variância um, $\alpha_0, \alpha_1, \dots, \alpha_p$. Usualmente supõe-se que $r_t \sim t_v$ (distribuição t de Student com v graus de liberdade). Dependendo do tipo de suposição que colocamos sobre o processo $\{r_t\}$, os coeficientes α_i devem satisfazer certas condições.

Como os quadrados dos retornos entram na fórmula da volatilidade no modelo de ARCH, os retornos positivos e negativos são tratados de forma similar. Outra questão é devido a alguns retornos grandes isolados que podem conduzir a super previsões, o que torna evidente algumas desvantagens deste modelo. No entanto, muitos trabalhos realizados com o modelo ARCH obtiveram sucesso com a melhoria nas previsões.

Um exemplo é Dusan Marcek (2000), que procura demonstrar problemas de previsão e procura avaliar os modelos de ARCH. Com o trabalho o autor procura determinar variância e previsões para a série temporal do índice SAX. Conclui que o processo de previsão ARCH aplicado com uma suavização exponencial quadrática melhora o erro da previsão, mas de uma forma limitada por não poder dizer que melhora para qualquer série temporal, apenas para o índice SAX.

2.5.6 O Modelo de GARCH

Como uma generalização do modelo de ARCH, foi criado o modelo de GARCH (*Generalized Autoregressive Conditional Heterocedasticity*) por Bollerslev em 1986, este é um processo estocástico em que a volatilidade condiciona os rendimentos no

futuro próximo. Este modelo permite a presença de componentes auto-regressivos e de médias móveis na volatilidade dos ativos financeiros.

A volatilidade depende dos quadrados dos erros anteriores e também de sua própria variância em momentos passados. Um modelo de GARCH (r,s) é definido por Morettin e Toloi (2006) da seguinte forma:

$$\sigma_t^2 = \alpha_0 + \alpha_1 e_{t-1}^2 + \alpha_2 e_{t-2}^2 + \dots + \alpha_r e_{t-r}^2 + \beta_1 \sigma_{t-1}^2 + \beta_2 \sigma_{t-2}^2 + \dots + \beta_s \sigma_{t-s}^2, \quad (17)$$

A versão deste modelo mais utilizada para séries financeiras e a utilizada neste trabalho será o GARCH (1,1). Neste modelo é suposto que os erros são normalmente distribuídos e a variância é dada por:

$$\sigma_t^2 = \alpha_0 + \alpha_1 e_{t-1}^2 + \beta_1 \sigma_{t-1}^2, \quad (18)$$

Onde α_1 mede a intensidade em que um impacto no retorno hoje afeta na volatilidade do retorno do próximo dia. A adição de (β_1) mostra a medida de constância na volatilidade, ou seja, a percentagem que reflete como o impacto no retorno hoje se dissemina ao longo do tempo, sobre a volatilidade dos retornos futuros. Assim é importante observar que a alta constância ou persistência do impacto enfraquece-o lentamente.

O modelo de GARCH é o mais avançado que se conhece hoje para medir volatilidade. A razão de ser um dos melhores modelos para se trabalhar com riscos em séries temporais acontece pelo fato de levar em conta o excesso de curtose e a acumulação de volatilidade. Além disso o modelo tem previsões precisas de variâncias e covariâncias dos retornos dos ativos no tempo pelas variâncias condicionais.

Por isso o modelo pode ser aplicado em vários estudos financeiros de séries temporais como: gestão de portfólios e alocação eficiente de capital, mercado de opções, estrutura de taxas de juros, dentre outras.

Portanto, é um mecanismo que inclui as variâncias passadas na explicação das variâncias futuras. Mais especificadamente, GARCH é uma técnica de séries temporais

que permite utilizar o modelo de dependência serial da volatilidade (BOLLERSLEV, 1986).

Sabe-se que o modelo de GARCH não é perfeito e tem suas limitações, como por exemplo: (a) a incapacidade de captar todas as curtos observadas em ativos e (b) a volatilidade pode explicar algumas, mas não todas as razões das irregularidades na distribuição. Para suprir parte dessas limitações, são aplicadas distribuições como t de Student, tentando analisar a série como uma distribuição normal.

Mesmo com suas irregularidades o modelo GARCH é um dos mais usados para fazer previsões de séries temporais pelo fato de ser um dos modelos mais perfeitos para tal função.

Cardoso (2005), propõe um modelo baseado em GARCH para a previsão da demanda da volatilidade de energia elétrica para consumidores livres em regime de curto prazo. O trabalho mostra que o modelo é aplicável para a demanda de energia elétrica.

Galdi e Pereira (2007), estimaram a volatilidade de séries temporais de 3 maneiras: pela suavização exponencial (EWMA), pelo modelo de volatilidade condicional (GARCH) e pelo modelo de volatilidade estocástica (MVE). Após o cálculo da volatilidade eles utilizaram as previsões para o cálculo do VaR, que é uma métrica utilizada para calcular a perda máxima de uma carteira de ativos. Este estudo foi feito para uma janela de 1500 observações. Com o trabalho os pesquisadores concluíram que os testes calculados pelo EWMA sofreram um menor número de violações do que para os outros modelos estudados.

Charles (2008), faz uma pesquisa sobre a previsão de preços nos retornos de 17 ações francesas, com o modelo de GARCH detectando e corrigindo, no entanto, retornos anormais (outliers). Este modelo de detecção foi proposto pela Franses e Ghysels (1999) e estendido para outliers inovadores em 2005 por Charles e Darné. Com o trabalho o autor mostra que os parâmetros da equação que regem a dinâmica da volatilidade são tendenciosos quando não se leva em conta outliers aditivos¹ e inovadores². Mostram que a previsão é melhor quando os dados são limpos de outliers.

¹ *Outliers aditivos: Correspondem a um erro grosso de medição ou gravação afetando uma única observação. Eles são uma componente aleatória adicional, apresentando um comportamento de outliers.*

² *Outliers Inovadores: Ocorre um choque em um determinado período e o efeito se propaga para observações subsequentes. Apresentam ruídos ocasionais e variância maior que os restantes.*

Caporin e McAller (2006) propõe um modelo de GARCH assimétrico e dinâmico. Este modelo acomoda o tamanho e o sinal dos efeitos de choque e introduz algumas dinâmicas na assimetria de variância. Este modelo oferece um modelo mais flexível que o de Engle. Para mostrar a utilidade do modelo, este foi aplicado nos retornos diários do S&P, Nikkei 225, FTSE 100 e EuroStoxx 50. Os pesquisadores concluíram que esta introdução de dinâmica e assimetria foram relevantes para os índices analisados, uma vez que a análise mostra que o efeito multiplicador não está embasado em apenas uma “boa” ou “má” notícia, mas também sobre o tamanho da notícia, ou seja, a repercussão que ela pode causar.

Visser (2010) generaliza o medidor de máxima verossimilhança (QMLE) e por esse modelo tenta fazer previsões de volatilidade dos retornos do S&P 500, ele usa o QMLE para estimar parâmetros de tempo discreto do modelo de GARCH. Com a pesquisa o autor conclui que escolhendo bem o fator de generalização (QMLE), ele reduz a volatilidade do modelo GARCH (1,1).

Feng (2004) propõem uma abordagem semi paramétrica de uma série temporal através da introdução de uma escala harmoniosa em função do modelo de GARCH. Os dados analisados são da S&P 500 e DAX 100. Usando os modelos juntos o resultado foi significativo, ou seja, a volatilidade foi menor do que usando os modelos separadamente.

Adesi, Engle e Mancini (2008) propõe um novo método de precificação de opções utilizando o modelo GARCH com um filtro de resíduos históricos. Esse estudo é feito com as opções da base de dados da S&P 500. Com isso os pesquisadores trazem a precificação de opções para a teoria da probabilidade com volatilidade. O modelo mostra uma flexibilidade maior para a estimativa dos preços das opções com a estimativa de probabilidades.

Karlafits e Vlahogianni (2009) analisam o tráfego de transporte utilizando os modelos de ARIMA, GARCH e um modelo de memória longa, pois afirmam que muitos estudos são feitos com ARIMA GARCH, mas falta-lhes um modelo de memória longa (dual) para melhorar a previsão. Eles descobrem que este modelo de memória dual aplicado juntamente com o de ARIMA GARCH, melhoram a representação do tempo em relação à série original.

O uso do modelo de GARCH para séries temporais foi muito usado na literatura, e com o passar do tempo alguns autores começaram a usá-lo juntamente com outros métodos para tentar diminuir a volatilidade da previsão, aumentando assim o grau de certeza da previsão. Um dos modelos que foi muito utilizado com o modelo de GARCH foram os filtros por ondaletas. Este filtro como o nome já diz, filtra, ou seja, “limpa”, a série temporal transformando-a em várias outras subséries para tentar melhorar a capacidade preditiva. Na próxima seção mostramos como este método funciona detalhadamente.

2.6 Definição de Filtro

Morettin (2002) entende que um filtro funciona na verdade como um processo de transmissão de dados que passaram por um processo de “limpeza”. Esse processo ocorre mediante transformações matemáticas nas séries temporais que possibilitam essa purificação dos elementos da série.

Neste trabalho a definição de um filtro será a mesma adotada por Lima (2004) e terá a finalidade de encontrar uma descrição mais simples do efeito de uma transformação linear de um processo estacionário.

2.6.1 Filtro de Ondaletas

A análise de dados, de acordo com escalas variáveis, no domínio do tempo e da frequência, é a idéia básica da utilização da teoria das ondaletas conhecidas na literatura internacional por *wavelets*. As ondaletas são funções matemáticas que ampliam intervalos de dados, separando-os em diferentes componentes de frequência, permitindo a análise de cada componente em sua escala correspondente. (MISITI et al., 1997).

Misiti et al. (1997), ainda afirma que este processo de decomposição, a partir de uma série x_t , passando pelo filtro de ondaleta, dá origem a duas novas séries, uma chamada aproximação (baixa frequência) e outra detalhe (alta freqüência). E que o processo de decomposição da série pode ser iterativo, com sucessivas decomposições com n caminhos diferentes para a codificação da série. Uma árvore de decomposição é feita da seguinte forma:

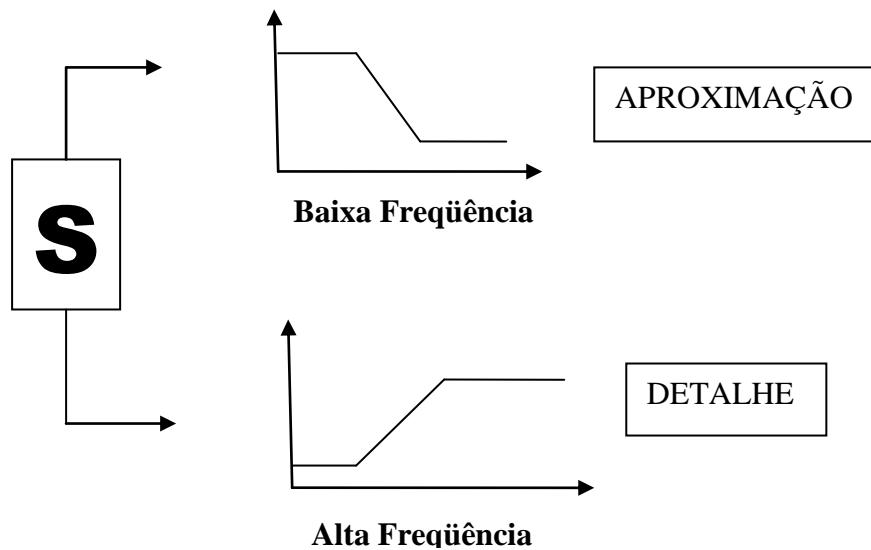


Figura 1: Decomposição por ondaletas.

Fonte: MISITI et al.,1997

O autor em questão afirma também que o processo de decomposição da série pode ser iterativo, com sucessivas decomposições, formando uma árvore de decomposição apresentada da seguinte forma:

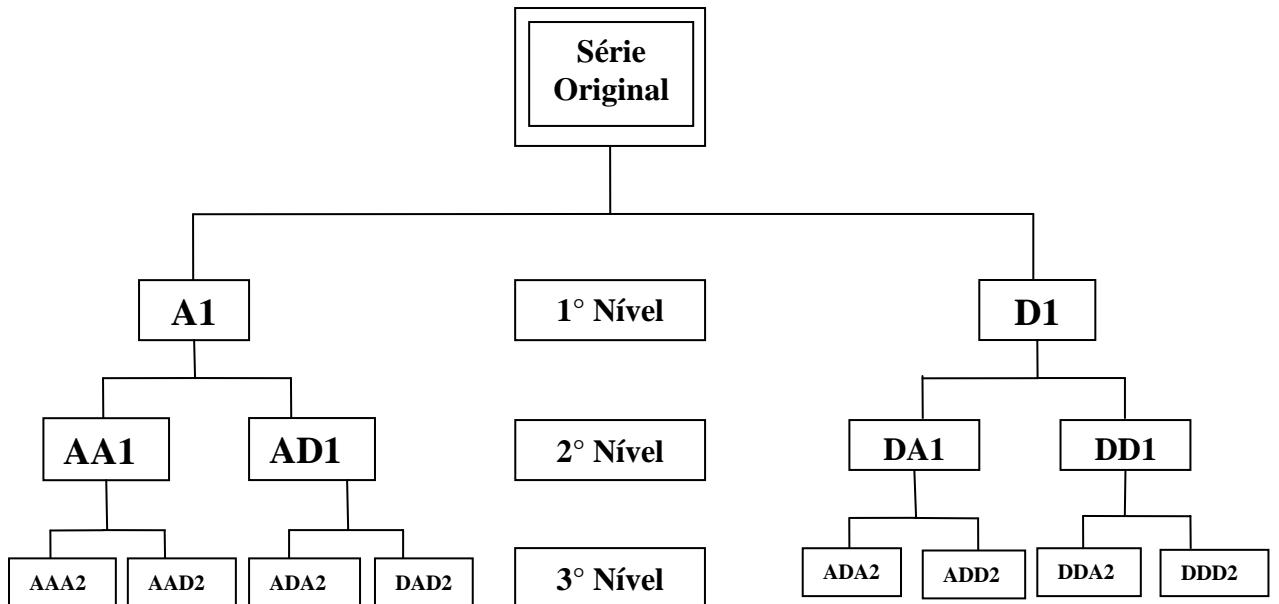


Figura2: Árvore de decomposição em 3 níveis de uma série temporal via decomposição por ondaletas.

Fonte: MISITI et al.,1997

A problemática em questão consiste em obter uma frequência crítica de maneira que os componentes da série original relacionados a frequências menores ou iguais a uma frequência crítica formem subséries relativa a baixas frequências, que também recebem o nome de série “aproximada” e os demais componentes formem outra subsérie relativa a altas frequências, denotada \tilde{x}_t , também chamada de série “detalhe” (HOMSY et al., 2000).

As ondaletas são funções matemáticas que aumentam os intervalos dos dados, e tornam possível que cada componente seja alocado em sua respectiva escala. Este processo tem por finalidade separar os dados da série original em duas outras séries menores por seus constituintes de frequência.

Com esta importante função de transformação de uma série em outras subséries a decomposição por ondaletas vem sendo muito utilizada na literatura para a diminuição da volatilidade da previsão, dentre outros objetivos. Seguem alguns exemplos:

Gosse (2009) mostra uma metodologia geral para se lidar com índices da bolsa durante períodos de alta volatilidade. Ele faz previsões de volatilidade com uma teoria de ondaleta discreta. Este método pode ser aplicado como alternativa ao modelo de Gauss em previsões de volatilidade, pois os resultados são bem próximos.

Grané e Veiga (2010) propõem um modelo de previsão filtrando ao dados das séries temporais e retirando os outliers. Assim ele testa a série completa e outra filtrada de outliers em 7 modelos de previsão baseados nos de Monte Carlo. E conclui que com a filtragem dos outliers o erro de previsão diminui.

2.6.2 Previsão de Séries Temporais com Modelos GARCH e Ondaletas

Tan et al. (2010) combinaram o modelo de GARCH e ondaletas para fazer a previsão com um dia de antecedência no preço de eletricidade no mercado espanhol. Metodologicamente os autores transformam a série em várias subséries; em seguida, para cada subsérie, é aplicado o modelo de previsão GARCH e compõe-se os resultados obtidos em cada subsérie. Com a pesquisa os autores concluíram que o modelo de ondaleta combinado com o método de GARCH melhora bastante o método de previsão.

Hong e Lee (2001) testam o modelo de ARCH e o modelo de ondaletas juntos para tentar melhorar o erro de previsão. E concluem que o filtro por ondaleta ajuda na redução do erro por capturar picos espectrais que o modelo de ARCH não captura.

Nguyen e Nabbney (2010) fazem uma pesquisa para prever a demanda de energia elétrica com um dia de antecedência e a previsão do preço de gás usando o método de ondaleta e GARCH juntos. Os autores concluíram que o uso de um filtro antes de aplicar o modelo de GARCH melhora o erro, ou seja, torna a previsão mais próxima da realidade.

Além dos métodos de ondaletas e GARCH para a previsão ou auxílio na previsão de séries temporais, temos outros métodos usados na literatura. Um método interessante e muito usado também para este fim é o método de redes neurais que tem um objetivo simples e direto: encontrar um método de fazer previsão que realize bem a previsão para dados altamente não lineares. Este modelo será explicado melhor na próxima seção desta pesquisa.

2.7 Redes Neurais

Segundo Haykin (2001) Redes Neurais Artificiais (RNA) são sistemas de processamento de informações distribuídas, compostas por muitos elementos computacionais simples que interagem através de conexões com pesos distintos. Inspiradas na arquitetura do cérebro humano, elas exibem características como a habilidade de “aprender” padrões complexos de informação e generalizar a informação aprendida.

Haykin (2001) afirma ainda que as unidades básicas da rede são neurônios artificiais. Os neurônios se agrupam em camadas. Existem três categorias de camadas: a camada de entrada, a intermediária (que pode também conter mais de uma camada) e a de saída. Os neurônios entre as camadas são conectados por sinapses, os quais refletem a relativa importância de cada entrada com o neurônio. A camada de entrada é responsável pelas variáveis de entrada do modelo; a camada de saída contém um ou mais nós, representando os resultados finais do processamento e as camadas intermediárias, ou camadas ocultas (das quais podem existir uma ou mais), que irão tornar o processamento mais refinado e não linear.

A seguir, segue o modelo de um neurônio, que é a base para o projeto de redes neurais artificiais:

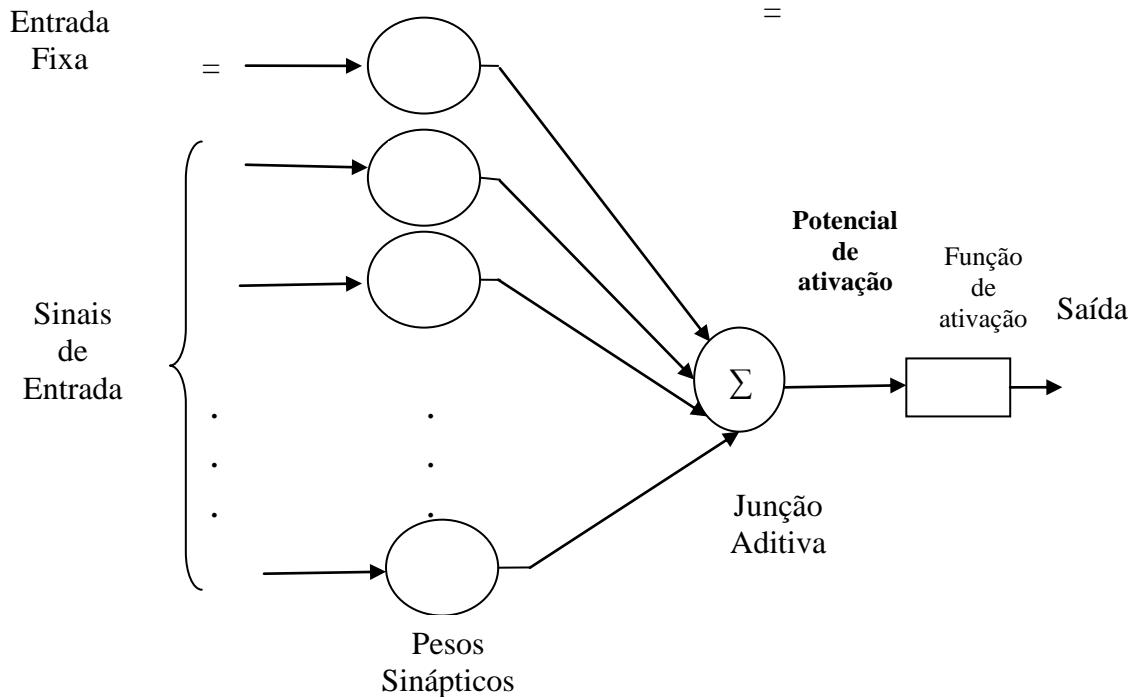


Figura 3: Modelo de um Neurônio Artificial.

Fonte: HAYKIN, 2001.

Na figura 3, os x_1, x_2, \dots, x_n são os sinais de entrada da rede; w_1, w_2, \dots, w_n são os pesos sinápticos do neurônio k ; b é a saída do combinador linear devido aos sinais de entrada; $\varphi(\cdot)$ é o viés; $\varphi(\cdot)$ é a função de ativação e y é o sinal de saída do neurônio. O uso do viés tem o efeito de aplicar uma transformação afim à saída do combinador linear.

Cada elemento computacional não linear é chamado de nó, sendo que os nós são densamente interconectados através de conexões diretas. Os nós operam em passos discretos, de forma análoga a uma função de dois estágios: o primeiro estágio calcula a soma dos sinais de entrada, atribuindo pesos aos sinais; o segundo estágio consiste na aplicação de uma função de saída, chamada de função de ativação.

As RNA(s) são uma alternativa aos modelos lineares e a algumas abordagens não paramétricas para a aproximação de sistemas não lineares. A razão para o uso de uma RNA é simples e direta: o objetivo é encontrar uma abordagem ou método que realize bem previsões para dados gerados para processos que frequentemente são desconhecidos e altamente não lineares, com uma quantidade pequena de parâmetros e que seja mais fácil de estimar que os modelos não lineares paramétricos (DE OLIVEIRA, 2003).

O método de Redes Neurais é amplamente usado e aplicado com outros modelos na literatura atual para a previsão e análise em geral de séries temporais financeiras. Seguem alguns exemplos de trabalhos realizados com este método.

Salcedo-Sanz et al. (2008), inclui redes neurais em um modelo de previsão da velocidade do vento a curto prazo. Esta previsão foi feita com cálculos de velocidades do vento medidos de hora em hora no sudeste da Espanha. Os pesquisadores obtiveram resultados animadores que mostram que a previsão por redes neurais conseguem prever a velocidade do vento em pontos específicos.

Wang e Huang (2010) fizeram um trabalho objetivando fazer um modelo de avaliação de desempenho para aplicar nos sistemas de informações financeiras de empresas. Para isso eles montaram um classificador de rápida adaptação neural; e, para saber se o método estava certo compararam com o método de redes neurais. Eles concluíram que este método é rápido e eficaz para o sistema de informações financeiras das empresas que precisam ser rápidos nas tomadas de decisões.

Aminian et al (2006) investigam o problema de prever fenômenos macroeconômicos com redes neurais, utilizando o Produto Interno Real e a produção industrial americana. Com o trabalho os autores medem a precisão de testar modelos lineares e não lineares com o método. Assim, o trabalho oferece um embasamento de comparação entre a regressão linear e redes neurais.

Titterington (2004) tenta alimentar redes neurais de forma fechada pelo método Bayesiano. O trabalho tenta diminuir ou acabar com os erros de previsão. Assim ele analisa muitos métodos como o de Gauss e de Monte Carlo para comparação, mas chega à conclusão que o método bayesiano não serve para ser tratado juntamente com redes neurais.

2.7.1 Previsão de Séries Temporais com Modelos GARCH e Redes Neurais Artificiais

Yim (2002) analisa se o poder de previsão por redes neurais artificiais pode superar o modelo de GARCH no mercado de ações. Comparando várias estruturas dos dois modelos, a autora conclui que o modelo de redes neurais para a previsão no mercado de ações é superior ao modelo de GARCH. No entanto ela conclui também que o modelo de GARCH é útil como uma complementação da rede neural no quesito de melhor calcular a volatilidade.

Caldeira et al., (2009) têm como objetivo em seu trabalho propor um sistema inteligente capaz de identificar corretamente a ordem dos modelos GARCH, para facilitar a escolha do modelo a ser utilizado para evitar o uso indiscriminado do modelo GARCH (1,1). Assim os autores utilizam séries temporais simuladas para a validação do sistema e os resultados obtidos com o sistema inteligente serão comparados com os dos sistemas AIC e BIC. Como a pesquisa se encontra em andamento, o resultado que eles obtiveram até agora foi que quando se faz uma estrutura de rede e aplica-se uma previsão de GARCH no sistema o índice de acertos é de 46,7% e 41,8% nos métodos AIC e BIC respectivamente. E quando se desconsideram as séries anuais o índice de acertos das previsões é de 39,5% e 35,8% para AIC e BIC respectivamente, mostrando que o aprimoramento do modelo para a previsão tem relevância significativa.

Kamitsuji e Shibata (2005) demonstram em sua pesquisa o poder de previsão do modelo de redes neurais juntamente com um algoritmo que eles desenvolveram para melhorar a probabilidade de previsão do método. A base de dados deles foram os índices da bolsa de Tókio (TOPIX) com os dados de 1994-1996 e previram queda ou aumento para os períodos de 1997-2000. A precisão da previsão foi de 60,28%. O resultado foi muito satisfatório pois o modelo GARCH na mesma condição teve uma precisão de 50,02%; 51,38% e 57,21%.

2.7.2 Previsão de Séries Temporais com Modelos de Redes Neurais Artificiais e Ondaletas

Lu (2010) faz uma análise dos componentes independentes integrados (ICA) com base em redes neurais para a previsão de preços de ações. Primeiro o ICA é utilizado sobre as variáveis de previsão para gerar componentes independentes (ICs); depois os ICs que tem ruído são identificados e removidos e o resto dos ICs são utilizados para fazer a previsão pelo método de redes neurais. Este método foi aplicado no índice de fechamento da TAIEX e abertura do índice Nikkei 225 como ilustração para testar o método proposto. Os resultados mostram que as séries filtradas por ondaletas apresentam uma melhor precisão.

Wang e Vachtsevanos (2001) propõem um modelo de previsão para diagnosticar falhas iminentes no funcionamento de máquinas, assim como o cálculo de sua vida útil. Este modelo mistura o modelo de redes neurais e ondaletas. Para isso eles utilizaram estruturas de rede, algoritmo de aprendizagem, análise de estabilidade, gestão de

incerteza e avaliação de desempenho. A aplicação dos dois modelos juntos para essa função obteve sucesso, de acordo com a pesquisa.

Abiyev (2007) usa em conjunto o método de redes neurais e ondaletas com a integração da lógica fuzzy para a previsão de consumo de eletricidade. O autor usa as redes neurais e ondaletas para criar as regras fuzzy e a funcionalidade do sistema fuzzy é testada pela estrutura de redes neurais. Os resultados demonstram que o uso deste método melhora a previsão de consumos futuros de energia elétrica.

2.7.3 Previsão de Séries Temporais com Modelos GARCH, Ondaletas e Redes neurais Artificiais

Lima (2004), em sua tese de doutorado, explora a possibilidade de utilizar uma metodologia capaz de decompor uma série temporal via ondaletas, conjuntamente com os modelos econométricos e de redes neurais já existentes de previsão e comparar a qualidade de previsões obtidas para sucessões cronológicas lineares e não lineares simuladas. Com o trabalho o autor conclui que é importante fazer um fluxograma implementado para as previsões em separado por ondaletas, ele percebe que a capacidade das redes neurais de lidarem com sucessões cronológicas não lineares é uma grande vantagem, pois as previsões feitas com pré processamento convalidam uma indicação dessa característica.

Huang e Wu (2008) usam o método de *relevance vector machine* (RVM) e de ondaletas para se fazer previsões de séries temporais para os índices de ações. Para fazer a pesquisa eles utilizaram os índices da bolsa de valores americana (NASDAQ), da japonesa NK225 (Nikkei), TWSI (Taiwan) e da KOSPI (Coreia do Sul). As ondaletas foram usadas para se extrair características da escala de tempo da série e o modelo de RVM para se fazer a previsão em si. Assim eles compararam este modelo com o modelo puro de GARCH e depois com o modelo puro de Redes Neurais e concluíram que este modelo de trabalho resulta em um erro menor do que com o de GARCH e Redes Neurais puros.

Esses três métodos analisados anteriormente geralmente são usados para se prever previsões de curto prazo, e não conseguem prever grandes altas ou grandes baixas em previsões pelo fato de procurarem uma normalidade padrão e diminuir o efeito de grandes perdas pela normalização da volatilidade. Um dos métodos para se

prever grandes altas ou baixas é o coeficiente de Hurst que foi criado justamente para esse fim. Por isso falaremos no próximo tópico mais detalhadamente deste processo.

2.8 Exponente de Hurst

Hurst estudou problemas de dimensionamento de reservatórios, ou seja, qual seria sua capacidade ideal, analisando a vazão anual de algumas décadas de uma represa. A idéia básica era determinar os volumes máximo e mínimo no reservatório para prever secagem ou transbordamento. Essa diferença entre o máximo e mínimo de água do reservatório era o *range*, ou seja, a série de volumes do reservatório. Além deste cálculo foram feitos os desvios padrões das vazões de água, sendo que a estatística R/S é o resultado da divisão do *range* pelo desvio padrão, que é um valor adimensional. Hurst percebeu também que havia uma função relacionando o valor da estatística R/S ao número de observações que entram no cálculo (HURST,1951).

Mais tarde este cálculo foi feito para outros fenômenos naturais e percebeu que ele funcionava. Este expoente segue a seguinte fórmula:

(19)

Sendo que:

R/S = diferença da amplitude entre a maior e a menor ocorrência registrada (R),

dividida pelo desvio padrão encontrado na série (S);

N = número de observações

H = fator da lei de potência, que originalmente foi objeto de estudo de Hurst (nesta fórmula chamado de expoente de Hurst, variando entre zero e um).

Em 1969 Mandelbrot e Wallis verificam que há uma relação empírica entre o expoente de Hurst e o movimento browniano fracionário. A descoberta afirma que o coeficiente de Hurst apresenta a mesma forma do movimento browniano, no que se refere ao *rescaled range* (R/S) em função do período utilizado no cálculo (N). Isso indica que a estatística R/S e o expoente de Hurst “ H ” podem ser utilizados para representar as propriedades de memória de longo prazo para séries que apresentam um movimento browniano.

Os valores do expoente de Hurst são interpretados da seguinte forma: $0 < H < 0,5$: indica que a série é anti-persistente, no qual uma tendência positiva no passado é

mais provável de se converter em tendência negativa e vice-versa. Para $H = 0,5$: a série representa *random walk*, ou seja, o movimento é puramente aleatório. E para $0,5 < H < 1$ a série é persistente, no qual uma tendência positiva no passado é mais provável de continuar positiva e vice-versa, e o nível desta persistência é medido quanto maior for o resultado.

O cálculo do Exponente de Hurst por meio de análise R/S clássica é assintótico, o que resulta em erro, o qual será maior quanto menor for o tamanho da amostra submetida à análise (SELVARATNAM E KIRLEY, 2006). A análise clássica não prevê uma correção para a estatística R/S no caso de amostras pequenas. Por este motivo neste trabalho foi usado uma base de dados de pelo menos 5 anos em amostras diárias para que este erro não atrapalhe os resultados.

Tabak e Cajueiro (2004, p.8) desenvolvem também uma pesquisa do comportamento de dependência a longo prazo da estrutura das taxas de juros japonesas. A pesquisa é realizada com o expoente de Hurst para saber se há uma persistência ou antipersistência dos dados. Eles percebem que há uma dependência nesta séries e quanto maior o tempo maior a taxa de juros, concluindo que existe um prêmio pelo prazo.

Matias (2006), realiza seu trabalho de mestrado com o objetivo de analisar a teoria dos fractais sob a ótica de precificação dos ativos, identificando um modelo matemático capaz de testar a eficiência desta teoria. Para isso o autor utiliza 7.922 observações diárias de preços de minério de cobre, compreendendo um período entre 1994 e 2005. Foram realizados testes de normalidade, linearidade, estacionariedade e fractalidade, com o objetivo de identificar se a série apresentou comportamento de uma série temporal durante o período analisado. Com um resultado positivo, identificou-se o parâmetro fractal de Hurst correspondente à série. Dentro dos parâmetros e da janela de tempo utilizada para a comparação da série real da projetada, o trabalho concluiu que apenas esta análise do expoente de Hurst não é satisfatória para previsões, sendo que o expoente de Hurst apenas observa o comportamento da série como persistente e anti-persistente.

Souza, Tabak e Cajueiro (2006), medem a evolução da memória a longo prazo da taxa de câmbio diária, Real contra o dólar dos Estados Unidos, no período de 1995 a 2004. Esta medida é feita através da estatística R/S, com a janela móvel dos dados. Com o trabalho os autores perceberam que a medida de memória de longo prazo obtida para o período de regime cambial de bandas indicou antipersistência. Além disso, percebeu-

se que o regime de bandas manteve as taxas de câmbio do país isoladas das crises que ocorreram no período, como as crises da Ásia e da Rússia.

Eom et al. (2008) investigam a relação entre o grau de eficiência de mercado e a previsibilidade de séries temporais financeiras. Na pesquisa, o Expoente de Hurst foi usado como um medidor de grau de eficiência de mercado, para 60 índices de mercados de vários países. Assim eles descobriram que existe uma forte relação entre o grau de eficiência de mercado (Expoente de Hurst) e a previsibilidade (taxa de sucesso), ou seja, quando o expoente de Hurst é maior a taxa de sucesso tende a ser maior. Esse resultado mostra que o Expoente de Hurst é útil para observar a persistência e não persistência de preços no mercado.

2.8.1 Expoente de Hurst e Modelos de GARCH em Séries Temporais

Cajueiro e Tabak (2003) testam a informação de que os mercados emergentes estão se tornando mais eficientes ao longo do tempo. Para isso eles filtram a série pelo modelo de GARCH e calculam o expoente de Hurst para os resíduos dos retornos. Os dados analisados foram séries temporais de ações dos principais índices dos principais países da América Latina e Ásia. Os resultados empíricos afirmam que isso pode até ser verdade para a maioria dos países, mas para países como o Brasil, Filipinas e Tailândia não condiz com a realidade.

Ainda Cajueiro e Tabak (2005) testam um modelo de previsão filtrando os dados pelo método de GARCH e submetendo os resíduos à análise clássica da estatística R/S. Segundo os autores a filtragem é feita para se extrair dos elementos da série a parcela devida à ocorrência de memória curta e à variação das volatilidades dos retornos. Eles concluem que se o Expoente de Hurst for calculado para retornos não ajustados em relação à volatilidade deles, eles são reconhecidos como persistentes em séries financeiras.

2.8.2 Expoente de Hurst e Ondaletas em Séries Temporais

Xu et al. (2007) estuda sobre o escoamento anual na corrente do rio Tarim usando filtros de ondaletas e o expoente de Hurst, para o período de 1958 a 2005. Os resultados mostram que quando a série é analisada como um todo mostra uma tendência caótica com uma leve tendência decrescente. Filtrando a série, esta é dividida em três

períodos o que mostra para o primeiro deles um coeficiente de Hurst igual a 0,5036 o que mostra que o escoamento anual está em caos. Para os outros dois períodos o coeficiente resulta em um número bem maior que 0,5 mostrando que uma característica perdura no tempo. Com isso os autores concluíram que há uma ligeira tendência do fluxo de água aumentar nos 16 anos à frente de 2005.

Manimaran, Panigrahi e Parikh (2008) realizaram um estudo para medir a caracterização de correlação e dimensionamento de séries temporais não estacionárias do sistema financeiro. Para isso usaram o método de ondaletas e a estatística R/S e aplicaram nos preços diários Nasdaq por um período de 20 anos e BSE Sensex, durante um período de 15 anos. Para o índice BSE os autores encontraram uma forte correlação, já o índice Nasdaq apresentou um comportamento multifractal de memória longa.

Power e Turvey (2009) Analisam a natureza da dependência da memória longa na previsão de preços futuros de 14 commodities agrícolas. Para isso eles usam o coeficiente de Hurst para a análise de previsão e ondaletas para a detecção da auto similaridade e da hipótese H. Os resultados da pesquisa demonstram que para todas as commodities há evidências de memória longa e 9 delas são não estacionárias.

Zunino et al. (2007) discorrem sobre a ineficiência dos índices latino americanos. Eles exploram a ineficiência de previsão da volatilidade em um mercado subdesenvolvido e comparam com a previsão de volatilidade em um país desenvolvido. Para isso os autores aplicam o expoente de Hurst, para a previsão de memória longa e usam ondaletas para filtrar a série e quantificar o grau de memória longa. Com isso eles descobrem que comparado com os resultados de volatilidade de países desenvolvidos, os índices de países subdesenvolvidos é muito irregular e tem um erro muito maior. Eles concluem que isto acontece porque os índices de preços diários de ações em países subdesenvolvidos não apresentam uma distribuição de probabilidade Gaussiana.

Kyaw, Los e Zong (2005) estudam sobre as características de persistência no mercado de ações latino americano. Com o estudo eles descobrem que as séries temporais dos preços das ações do mercado estudado não apresentam traços de normalidade, são não estacionários e apresentam traços de memória longa. Para estudar sobre tal assunto eles utilizaram o expoente de Hurst, já que os dados apresentavam traços de memória longa e o filtro por ondaletas para perceber o grau de dependência da série ao longo do tempo. Com o estudo eles descobriram que deve-se ter muito cuidado ao se fazer previsões de preços em mercados subdesenvolvidos pelo fato de

apresentarem uma alta volatilidade e como a distribuição é não gaussiana há uma maior possibilidade de acontecer uma mudança brusca no mercado.

Los e Yu (2008) fazem o mesmo estudo de Kyaw, Los e Zong, mas para os índices de preços das bolsas chinesas: Shanghai (SHI), Shenzhen a shares (SZI) e B shares (SZBI) antes e após a intervenção do governo chinês no funcionamento e no ditamento de preços para o mercado. Como a pesquisa citada anteriormente os pesquisadores percebem que o preço das ações não apresentam uma distribuição normal de probabilidade e resolvem aplicar o método de ondaletas para calcular a dependência do mercado e o expoente de Hurst para observar os traços de memória longa que a série apresentou. Os autores concluíram que com a regulamentação do governo para com o preço das ações, o mercado está funcionando de forma mais eficiente, ou seja, os dados apresentam persistência de memória longa.

2.8.3 Expoente de Hurst e Redes Neurais Artificiais em Séries Temporais

Selvaratnam e Kirley (2006) fazem a previsão de séries temporais financeiras do índice australiano *index on the Australian Stock Exchange* (SPI). Para isso eles analisam as séries por redes neurais artificiais e para identificar as janelas adequadas para classificar o tempo das séries, usam o expoente de Hurst. Eles concluem que os dois métodos usados para a predição em conjunto melhoraram a previsão.

Lin, Yang e Song (2008) Analisam os dados diários da S&P 500 com o objetivo de filtrar a série para se fazer uma melhor previsão. Para isso os autores inicialmente filtram a série com o cálculo do expoente de Hurst para escolher subséries com maior previsibilidade no tempo. E para a memória a curto prazo aplicam o método de redes neurais. Com o trabalho eles concluem que este método é bastante eficaz para a filtragem dos dados de longo e curto prazo.

2.9 Contribuição da Pesquisa

Assim após estas observações pode-se afirmar que os modelos de Hurst, Redes Neurais Recorrentes e filtro de ondaletas vêm sendo amplamente utilizados na literatura, no entanto, eles nunca foram utilizados ao mesmo tempo para observação de uma melhor previsão para séries temporais financeiras. Neste contexto este trabalho visa acrescentar esta contribuição para a literatura.

3. METODOLOGIA

Nesta parte do trabalho são apresentados: a classificação da pesquisa, os métodos empregados, os procedimentos para a análise, previsão e avaliação dos resultados para o problema proposto inicialmente.

O principal objetivo desta pesquisa é poder contribuir para a área contábil, aprimorando os estudos sobre o comportamento dos ativos financeiros com instrumentos mais modernos e precisos. Para isso pretende-se verificar se o expoente de Hurst, aqui usado para uma classificação da qualidade dos dados em cada janela, interfere na qualidade das previsões feitas.

A literatura indiretamente incentiva o desenvolvimento desta pesquisa à medida que, tem-se mostrado inconclusiva sobre a questão da contribuição do uso combinado de técnicas de previsão. Assim haverá um aperfeiçoamento dos modelos de gestão de contratos e aplicações futuras.

3.1 Natureza da Pesquisa

Segundo o entendimento de Beuren (2006), esta pesquisa pode ser classificada de acordo com seus objetivos de pesquisa, aos procedimentos adotados para se chegar ao constructo e à abordagem do problema.

Quanto aos objetivos, o trabalho apresentado pode ser classificado como exploratório, porque desenvolve uma visão geral dos grandes temas colocados em questão, tornando-os claros para a formulação de suas hipóteses. E também descriptiva por buscar descrever as principais características presentes numa série temporal para se construir um modelo que rege ou pelo menos diminua o erro de previsão para o seu comportamento.

De acordo com os procedimentos adotados para se chegar ao constructo, pode-se dizer que esta pesquisa é bibliográfica, pois para dar suporte teórico e visualização do estado da pesquisa que combina o modelo do expoente de Hurst e outros modelos que permitem melhor adequação à previsão de séries temporais foi necessário realizar uma pesquisa bibliográfica.

A pesquisa foi feita em portais disponibilizados pela faculdade, materiais disponibilizados pelo orientador e livros. A maioria das referências são internacionais, pelo fato de não existirem muitos trabalhos nacionais com grande relevância na área.

Assim selecionaram-se pesquisas que foram realizados a nível mundial que tenham um índice de impacto na academia, que usaram os métodos ou combinaram métodos descritos no trabalho, nas mais diversas áreas do conhecimento, dando mais atenção para os realizados nas séries temporais financeiras.

Não se teve a pretensão de se chegar a uma revisão completa de todos os inúmeros trabalhos realizados no mundo, mas buscou-se chegar o mais próximo possível de onde as técnicas estudadas estão sendo aplicadas e como isto está sendo feito, levando em consideração sua relevância acadêmica.

Quanto à abordagem do problema, a pesquisa realizada é quantitativa de base econométrica, pelo fato dos modelos utilizados terem um desenvolvimento matemático e estatístico. As análises contaram com ferramentas do software Matlab, um dos softwares mais completos que se tem no mundo para esta técnica.

3.2 Operacionalização da Pesquisa

Para a consecução da pesquisa foram construídas programações próprias no software Matlab ® 6.5 para cada modelo. Um para o Expoente de Hurst, outro para o GARCH, outro para Redes Neurais Recorrentes e por fim um para o filtro de ondaletas. Este desenvolvimento tornou-se necessário para que se chegue a todas as etapas descritas na pesquisa. Os modelos de Redes Neurais Recorrentes e Ondaletas em Matlab foram aproveitados do trabalho realizado de Lima (2004) e o modelo de GARCH e Hurst foram construídos para a pesquisa.

Já se sabe que existem vários modelos de previsão para séries temporais financeiras, desde modelos mais simples (pode-se citar o modelo de suavização exponencial) até os mais aceitos hoje na literatura que são os da família GARCH e o de Redes Neurais Recorrentes. Esses modelos fazem suas previsões com base em acontecimentos passados, ou seja, utilizam o efeito de memória das séries.

Sabe-se também que o Expoente de Hurst, calculado ao se delimitar uma certa janela evidencia o efeito de memória de uma série temporal, ou seja, se uma série temporal apresenta ou não memória nos dados.

Assim, têm-se como proposta da pesquisa: delimitando-se uma janela em uma série temporal financeira, calculando-se o Expoente de Hurst para cada janela, verificando o efeito de memória nos dados e, posteriormente, aplicando os métodos de previsão mais aceitos na literatura, a volatilidade, ou seja, o erro de previsão, diminui?

Como complemento, pretendemos saber também o contrário, ou seja, calculando-se o Expoente de Hurst para cada janela, ao se verificar que a série não apresenta um efeito de memória nos dados, o erro de previsão, aumenta?

Salientando que não há uma janela “ótima” para se analisar séries temporais, como objetivo secundário pretende-se também testar vários tamanhos de janelas para observar se este fator influencia na previsão.

As previsões serão feitas com um dos modelos mais aceitos na literatura como já mencionado anteriormente que é o de Redes Neurais Recorrentes. Além desses dois testes faremos também previsões filtrando-se a série temporal financeira com ondaletas antes de se fazer a previsão. A figura 4 ilustra a proposta da pesquisa de forma simplificada.

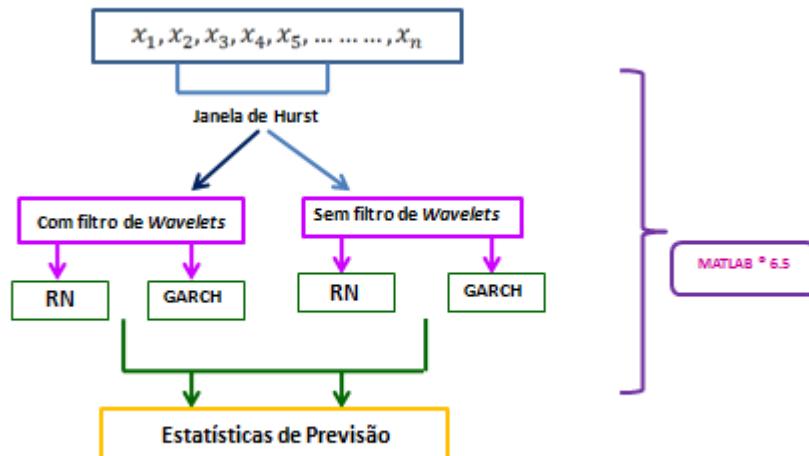


Figura 4 Proposta da pesquisa

Abaixo mostramos a operacionalização de cada modelo utilizado na pesquisa.

3.2.1 Operacionalização do modelo de Hurst

O processo de cálculo desse expoente é descrito a seguir e segue passos descritos em Souza, Tabak e Cajueiro (2006).

Seja R_t a série temporal financeira a ser estudada e R_t o log-retorno desta série e n o tamanho da amostra dos retornos. Para a estimativa da memória longa para a série temporal das volatilidades, utiliza-se como aproximação para a volatilidade instantânea em t o valor absoluto de R_t .

Para se calcular o expoente de Hurst “H” são feitos vários cálculos da estatística R/S para janelas de τ observações, sendo $\tau \leq N$. Para cada valor de τ a estatística R/S é calculada da seguinte forma:

A) A série temporal financeira de N log-retornos é dividida em “n” janelas contíguas de τ elementos, numerados com $1 \leq \tau \leq n$. Em cada janela i , os elementos $R_{t,i}$ são numerados com $1 \leq t \leq \tau$;

B) Calcula-se a estatística R/S associada ao tamanho da janela τ da seguinte forma:

1. Calcula-se a média dos retornos de cada janela:

$$\bar{R}_i = \frac{1}{\tau} \sum_{t=1}^{\tau} R_{t,i} \quad (20)$$

2. Calcula-se o desvio padrão em cada janela:

$$S_i = \left[\frac{1}{\tau} \sum_{t=1}^{\tau} (R_{t,i} - \bar{R}_i)^2 \right]^{\frac{1}{2}} \quad (21)$$

3. Calcula-se, para cada janela i , a estatística R/S_i :

$$R/S_i = \frac{1}{S_i} \left[\max_{1 \leq t \leq \tau} \sum_{k=1}^t R_{k,i} - \bar{R}_i - \min_{1 \leq t \leq \tau} \sum_{k=1}^t R_{k,i} - \bar{R}_i \right] \quad (22)$$

4. Calcula-se a média dos valores R/S_i e se associa ao tamanho da janela τ :

$$R/S_{\tau} = \frac{1}{n} \sum_{i=1}^n R/S_i \quad (23)$$

C) Terminados os cálculos da estatística R/S para diversos valores de comprimento da janela τ , obtém o expoente de Hurst “H” da relação:

$$R/S_{\tau} = \left(\frac{\tau}{2} \right)^H \quad (24)$$

1. Dados os pares R/S_{τ} e τ , obtidos em (B), executa-se uma regressão pelo método dos mínimos quadrados ordinários:

$$\log_{10}(R/S)_\tau = \log_{10} C + H \times \log_{10} \tau + \text{erro} \quad (25)$$

2. O expoente de Hurst “H” é obtido com intervalo de confiança de 95%.

Após fazer o cálculo do expoente de Hurst para os próximos termos da amostra n, faz-se um estudo para verificar se a não linearidade é pela média, variância ou total da série. Se a não linearidade for devido à variância aplica-se na janela o modelo de GARCH (1,1) ou Redes Neurais Recorrentes para previsão.

3.2.2 Operacionalização do modelo GARCH (1,1)

No modelo GARCH padrão, a série de retornos é representada pelo modelo de média constante C com inovações gaussianas (resíduos), segundo um GARCH (1,1), conforme descrito abaixo (BOLLERSLEV,1986):

(26)

A variância da série, ou seja, o modelo GARCH(1,1) propriamente dito é representado, em primeira ordem, pela equação:

(27)

A figura 5 exemplifica modelos ARMA e GARCH que descrevem respectivamente, a média condicional e a variância condicional em uma série:

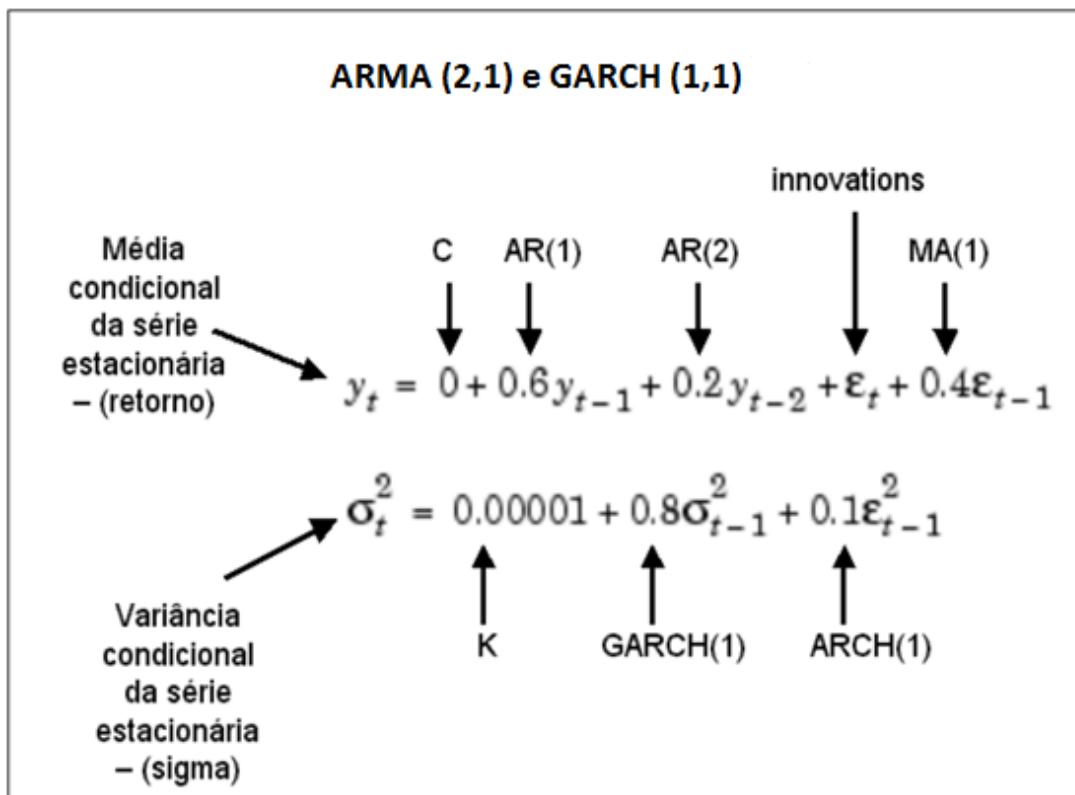


Figura 5 Exemplo do Modelo GARCH para a Média e Variância Condicional.

Fonte: (CARDOSO,2005)

De acordo com BOLLERSLEV (1986), o modelo de média condicional descrito pela primeira equação da figura 5 é, frequentemente, suficiente para descrever a média condicional em uma série de retorno financeira. O modelo de variância condicional pela segunda equação da figura 5, é um modelo que pode ser usado para previsão (além da análise) da variância , que consiste de uma constante K adicionada de uma média ponderada dos últimos períodos de variância e dos últimos períodos do quadrado dos resíduos . Nesse modelo supõe-se a inexistência de efeitos autorregressivos e de média móvel na série de retornos . Trata-se de um modelo simples, mas bastante comum desde que a série seja estacionária.

Já se a não linearidade da série for devido ao total, faz-se a previsão da janela pelo modelo de Redes Neurais Artificiais.

3.2.3 Operacionalização do modelo de Redes Neurais

De acordo com Haykin (2001), matematicamente pode-se descrever o processo de funcionamento da rede neural através das seguintes equações:

(28)

Haykin (2001) afirma que as equações acima descritas a partir da maneira pela qual os neurônios de uma rede neural estão estruturados, ficam diretamente ligadas com o algoritmo de aprendizagem usado para treinar a rede.

A propriedade de aprendizado é crucial para as redes neurais, pois define sua habilidade de aprender e de melhorar o seu desempenho por meio dessa aprendizagem. Essa melhora se deve a uma medida preestabelecida de erro do sistema.

Segundo De Oliveira (2003, p.89), a razão para o uso de uma rede neural é bem simples e direta, já que se quer encontrar uma abordagem para modelagem que aperfeiçoe as previsões para séries temporais financeiras que são dados altamente não lineares com uma quantidade baixa de parâmetros com facilidade de estimação dos mesmos.

Lima (2004) diz que a idéia do uso das redes neurais está baseada em como projetar um modelo de múltiplas camadas e tratar esse modelo como um sistema dinâmico. De acordo com esse modelo, quer-se estimar um valor futuro com base no conhecimento de um conjunto informacional de valores passados que é assumida por conter informação suficiente para se prever a evolução futura do sistema.

3.2.4 Operacionalização dos Filtros por Ondaletas

O objetivo de se usar ondaletas nesta pesquisa é que a partir de uma série temporal financeira possa-se obter suas sub-séries representativas de alta e baixa

frequência numa tentativa de suavizar os efeitos das variações no mercado financeiro através da aplicação de um filtro por uma função de ondaleta.

Segundo Morettin (1999), o objetivo da análise por ondaletas é estender essa função para o campo das funções chamada (IR), isto é, gerar um espaço, a partir de uma função , que é obtida por meio de dilatações através do parâmetro “a” da sentença matemática ou compressões e translações definidos pelo parâmetro “b” de , dada por:

$$\text{---} \quad \text{---} \quad (29)$$

Uma análise de ondaletas é definida pela aplicação da transformada de ondaleta com diversos valores para os parâmetros de dilatação e translação representando a decomposição da série original em seus respectivos componentes localizados no tempo e na frequencia para estes parâmetros.

A captação dos pontos de alta e baixa frequência se dá pela translação e dilatação da forma da ondaleta. Essa representação para uma série temporal denominada é matematicamente denominada por:

(30)

Onde j é o número de componentes e k que varia de 1 ao número de coeficientes do componente específico. Os coeficientes , , ..., são os coeficientes das transformadas de ondaletas dadas pelas projeções:

(31)

Esta parte é denominada parte de aproximação e

(32)

Esta parte é chamada de parte detalhe.

Assim as ondaletas, reduzem os ruídos das séries. De acordo com Donoho e Johnstone (1998) a idéia central é a supressão daqueles coeficientes que provocam ruído

na série de tempo, fazendo com que a série fique mais lisa no ponto de vista técnico e visual.

3.3 Estatísticas de Erro e Confiança para as previsões

Gooijer e Hyndman (2006) levam em conta em seu trabalho as principais medidas de erros nos trabalhos de previsão de série temporal. Eles indicam modelos para verificar o quanto o modelo de previsão é capaz de reproduzir os dados já conhecidos.

Dentre os diversos métodos de se medir a acurácia, os autores descritos acima definem 4 deles como sendo mais eficientes, estes são: MAPE (*Mean Absolute Percentage Error*), r (coeficiente de Correlação Linear de Pearson), TIC (Coeficiente de desigualdade de Theil – *Theil Inequality Coeficient*) e RMSE(*root mean squared error*).

Os modelos são descritos abaixo:

1. MAPE é um erro percentual absoluto médio, para se perceber a margem de acerto em comparação com o valor previsto. É mais adequado para a comparação entre modelos. Quanto menor o seu valor, melhor o modelo.



(33)

2. r: Este coeficiente de correlação linear r mede o grau de relacionamento linear entre os valores emparelhados x e y em uma amostra. O coeficiente de correlação linear varia entre -1 e 1, sendo estes extremos uma indicação de associação linear negativa e positiva perfeita.



(34)

3. TIC , Este coeficiente indica um ajuste perfeito e sempre estará entre zero e um:



(35)

4. RMSE (*root mean squared error*), indica a raiz do erro quadrático médio e quanto menor o resultado obtido, melhor o ajuste:

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$$

(36)

Para todas as previsões feitas são consideradas intervalos de confiança de 95%, centrados na previsão feita para a esperança condicionada dos valores.

3.4 Dados

Os dados a serem utilizados neste trabalho são os preços diários de ações, índice Ibovespa, taxa de câmbio e commodities, divulgadas no período de 02 de janeiro 1998 à 30 de dezembro de 2010. Os dados foram obtidos pelo economática, pelo site da Bolsa de Valores de São Paulo, BM&FBOVESPA e o Centro de estudos avançados em economia aplicada – ESALQ/USP.

4. RESULTADOS EXPERIMENTAIS

Nesta seção são apresentados os resultados das análises comparativas de previsão com diversas janelas do expoente de Hurst e um dos modelos de previsão mais aceitos na literatura, que é o de Redes Neurais Recorrentes, com e sem filtro de ondaletas aplicados em diversas séries temporais financeiras.

4.1 Análise da série da *commoditie* de café

A origem do hábito de tomar café remonta aos costumes das populações árabes no século XV. Este povo apresentou uma relevante contribuição para a difusão do cultivo, a comercialização e o consumo do cafeeiro para os demais países do mundo. Já no século XVIII, o café se consolidou como um dos mais importantes produtos do mercado internacional (Bragança,2003).

Nos dias atuais os países produtores de café estão localizados no hemisfério sul e apresentam clima tropical e subtropical, dentre eles, se destacam: Brasil, Colômbia, Vietnã, Indonésia. De outro lado, a demanda internacional é realizada por países desenvolvidos localizados no hemisfério norte, com exceção do Brasil.

A primeira bolsa de *commodities* agrícolas brasileira foi a bolsa de Mercadorias de São Paulo, fundada em 1917 por empresários paulistas ligados à exportação, ao comércio e à agricultura. O café começou a ser negociado na bolsa de *commodities* agrícolas brasileira juntamente com sua criação em 1917. No início, não se negociava contratos futuros agropecuários. Os contratos futuros foram lançados nas décadas de 70 e 80 para as mercadorias de café, boi gordo e algodão.

Em alguns períodos da década de 1980, o café era a segunda *commoditie* mais negociada no mundo por valor monetário, atrás apenas do petróleo. Este dado estatístico ainda é amplamente citado, mas tem sido impreciso por cerca de duas décadas, devido à queda do preço do café durante a crise do produto na década de 1990, reduzindo o valor total de suas exportações. Em 2003, o café foi o sétimo produto agrícola de exportação mais importante em termos de valor, atrás de culturas como trigo, milho e soja.

De acordo com a divulgação do volume de negociações de contratos agropecuários futuros e de opções pela Bovespa do ano de 2010, o café arábica foi a

segunda *commoditie* mais negociada no ano de 2010, foram: 694.348 negociações simbolizando 25,69% do total dos contratos agropecuários negociados na BM&FBOVESPA.

As razões para o sucesso do contrato futuro de café arábica na BM&F podem ser associadas à alta variabilidade de preço desse produto no mercado à vista (Lamounier,2003). Os contratos futuros do café correspondem a 100 sacas de 60kg cotadas em moeda norte americana e apresentam vencimentos nos meses de março, maio, julho, setembro e novembro (Martins,2005) .

Internacionalmente, a *Coffee, Sugar & Cocoa Exchange* (CSCE) e a *London International Financial Futures Exchange* (LIFFE) são as principais bolsas negociadoras dos contratos futuros de café. A CSCE foi fundada em 1882, em Nova Iorque para a negociação de café. Atualmente é a principal bolsa mundial de contratos futuros e opções de café, açúcar e cacau. Nesta bolsa, são negociados elevados volumes de contratos futuros do café tipo arábica lavado. Ao passo que, na LIFFE, são transacionados café do tipo robusta. (Martins, 2005)

Nesta pesquisa analisaram-se os preços diários da *commoditie* de café arábica, ou seja, o preço à vista da saca de 60kg diária, no período de 02 de janeiro de 1998 a 30 de dezembro de 2010, perfazendo 3231 observações obtidas junto ao CEPEA-ESALQ. A figura 6 mostra a série de preços correspondente aos valores diários nominais deste período e os log retornos da saca de café.

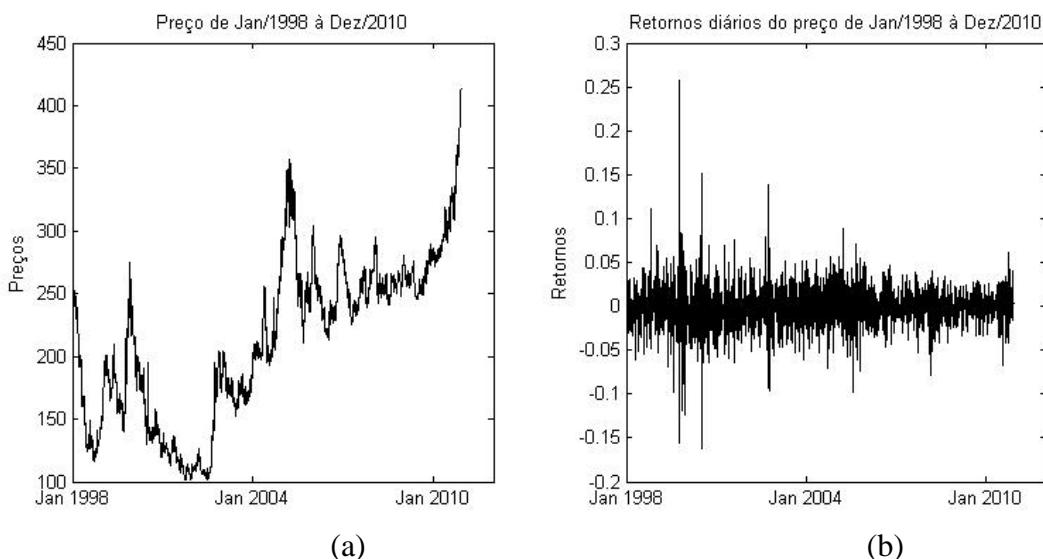


Figura 6: Série temporal do café no período de 02/01/1998 a 30/12/2010 e os respectivos Log Retornos.

Pode-se observar que os log retornos diários oscilam ao redor de zero, com períodos de alta e de baixa variabilidade, sendo essa variabilidade chamada de volatilidade. Nota-se também que existem dias em que o retorno sofre grandes oscilações, tanto positivas quanto negativas, o que não é o normal da série e são comumente chamados de *outliers*. Além disso, observam-se *clusters* de volatilidade que acontecem devido a incertezas e ao mercado. A influência desses fatores é impossível de calcular matematicamente, por isso são chamados de incertezas do mercado.

A figura a seguir, destaca a inspeção visual do gráfico da densidade de probabilidade dos log retornos.

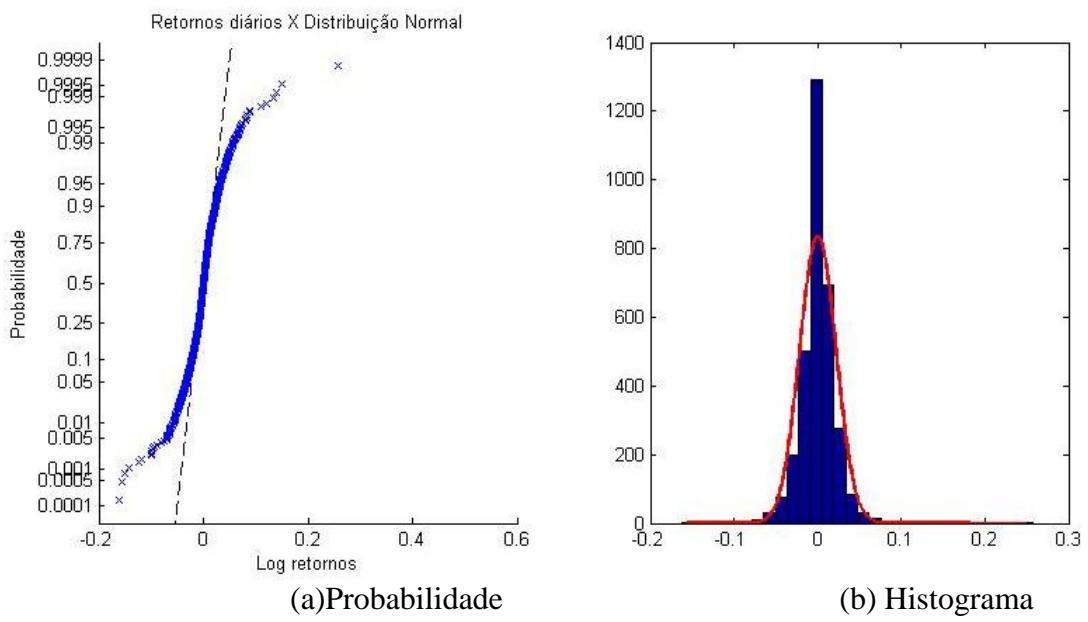


Figura 7: Histograma dos Log Retornos diários do café e retornos diários X distribuição normal.

Nota-se no gráfico da probabilidade, que traz no eixo das abscissas o cálculo dos retornos diários da *commodity* de café para o período analisado e no eixo das ordenadas a distribuição de probabilidade acumulada de acordo com a curva de Gauss. Para que os dados apresentassem uma distribuição normal os log retornos precisariam estar sobre a curva pontilhada, mas como observado na figura 7(a), os dados apresentam-se dispersos da curva pontilhada, o que mostra que a amostra não é uma distribuição normal.

O histograma tem sua parte central bem mais alta do que uma normal e há valores bem afastados da tendência central dos dados. Essas observações concluem que

a série temporal financeira tem um comportamento leptocúrtico com caudas mais pesadas que a normal pelo fato dos valores se afastarem da média a vários múltiplos do desvio padrão. Tais características encontradas coincidem com as características apontadas por Tsay(2005).

Para se analisar numericamente a normalidade da série dos log retornos de acordo com Jarque Bera (1987). Definem-se as seguintes hipóteses:

- a série segue uma distribuição normal.
- a série não segue uma distribuição normal.

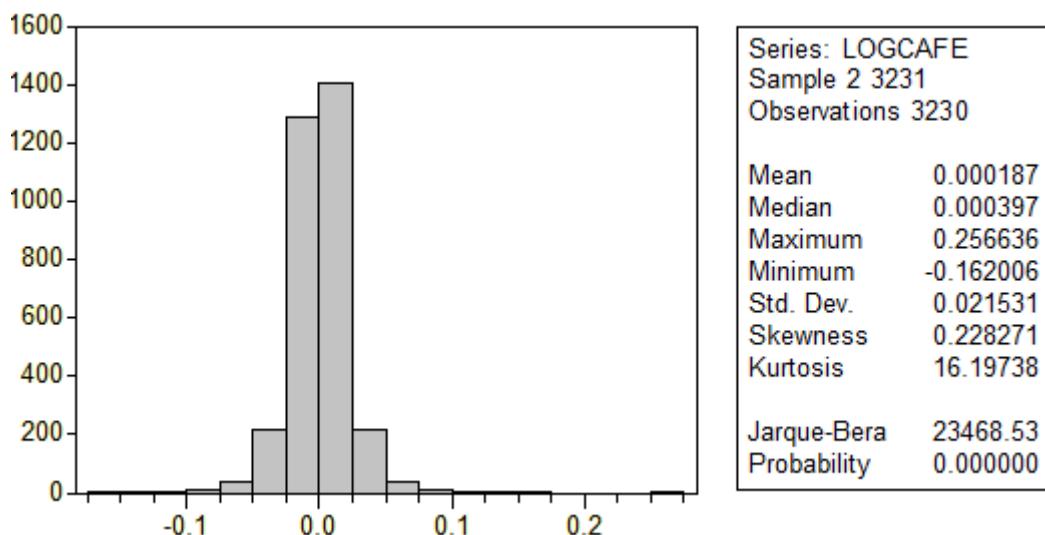


Figura 8 Histograma e estatística de Jarque Bera para os log retornos da *commodity* de café.

Desenvolvendo-se este teste e observando a figura 8, nota-se a estatística de JB=23468,53 com p-valor igual a zero, significando que a distribuição é leptocúrtica, e concentrada em torno da média, com uma alta homogeneidade. O nível de confiança escolhido foi de 95%, sendo sua probabilidade de significância inferior a 5%, indicando a rejeição da hipótese nula que revela que a série não segue uma distribuição normal. Ressalta-se rejeição dessa hipótese também a 99% de confiança.

Analizando as estatísticas descritivas para a *commodity* de café, observa-se que a série dos log retornos apresentam os principais fatos estilizados, conforme aponta Moretin e Toloi (2004, p.12); caudas mais pesadas que a distribuição normal, embora aproximadamente simétrica, leptocúrtica e não linear.

O próximo passo é testar a estacionariedade da série. Na literatura existem diversos tipos de testes para se verificar estacionariedade. O principal teste é o de Dickey Pantula (Dickey e Fuller, 1979). Este teste verifica se a série já se encontra estacionária no nível ou se é necessário fazer diferenças entre ela para se tornar estacionária. As hipóteses são:

H_1 RU: existe raiz unitária, isto é, a série é não estacionária.

H_0 RU: não existe raiz unitária, isto é, a série é estacionária.

A tabela a seguir mostra a estatística de teste para a série dos log retornos da *commoditie* de café com significância de 5%, testando a existência de até três raízes.

Tabela 1 teste ADF para os log retornos da *commoditie* de café

Variável	Coeficiente	Desv. Padrão	Estatística t	Probabilidade
D(LOGCAFE(-1,2))	-0,977110	0,017602	-55,50997	0,0000
D(LOGCAFE(-1,1))	-0,984260	0,040182	-24,49495	0,0000
D(LOGCAFE(-1))	-1,089814	0,032151	-33,89651	0,0000

O teste de Dickey Pantula para as possíveis raízes indicam um p-valor inferior a 5%, revelando a rejeição da hipótese nula. Logo, a série dos log retornos da *commoditie* de café é estacionária.

Para continuar a análise estatística também é preciso verificar se a série temporal segue um comportamento independente e identicamente distribuído, chamado i.i.d., isto é, todos os termos tem a mesma distribuição de probabilidade. O objetivo deste teste é identificar lineariedade ou não lineariedade determinista ou estocástica nos dados dos log retornos da *commoditie* de café.

O melhor teste para isso, segundo Tsay (2005), é o teste desenvolvido por Brock, Dechert e Scheinkman (1996) – BDS. E as hipóteses testadas são:

H_0 : a série é estocasticamente independente

H_1 : a série é estocasticamente dependente

Tabela 2 Valores obtidos pelo teste BDS nas respectivas dimensões para a *commoditie* de café

Dimensão	Estatística BDS	Desvio Padrão	Estatística z	Probabilidade
2	0,014914	0,001728	8,632243	0,0000
3	0,027856	0,002746	10,14501	0,0000
4	0,036884	0,003270	11,27804	0,0000
5	0,043593	0,003410	12,78438	0,0000
6	0,046339	0,003290	14,08603	0,0000

Observando o resultado, percebe-se que em todas as dimensões, os log retornos da *commoditie* de café não seguem um comportamento independente e identicamente distribuído com p-valores menores que 5%, onde se rejeita a hipótese nula. Assim verifica-se que existe uma dependência temporal entre os log retornos, ou seja, os retornos futuros são influenciados pelos retornos passados. A questão da série não ser identicamente distribuída, mostra que os log retornos possuem intervalos de tempos com diferentes distribuições de probabilidade, apresentando assim uma possível não linearidade dos dados.

Para se comprovar a não linearidade dos dados, pode-se aplicar o teste de McLeod-Li(1983). Este teste é baseado na autocorrelação amostral dos quadrados dos resíduos de uma sucessão cronológica estacionária gaussiana. As hipóteses são:

H_0 : a série temporal é linear

H_1 : a série é temporal é não linear

O teste é feito da seguinte forma: após ajustar um modelo autorregressivo - AR(1), que remova a dependência linear dos dados, é feita uma análise dos resíduos dessa regressão.

Tabela 3 Valores obtidos pelo teste de McLeod-Li para a *commoditie* de café

Variável	Coeficiente	Desv. padrão	Estatística t	Probabilidade
D(LOGCAFE(-1,2))	-0,977110	0,017602	-55,50997	0,0000
D(LOGCAFE(-1,1))	-0,984260	0,040182	-24,49495	0,0000
D(LOGCAFE(-1))	-1,089814	0,032151	-33,89651	0,0000
R-squared	0,952065	Mean dependent var		-0,0000205
Adjusted R-squared	0,952036	S.D. dependent var		0,098025
S.E. of regression	0,021468	Akaike info criterion		-4,843545
Sum squared resid	1,485909	Schwarz criterion		-4,837894
Log likelihood	7818,061	Durbin-Watson stat		1,999240

O teste de McLeod-Li, rejeitou a hipótese nula pelo fato da probabilidade ser menor que 5%. Indicando que a série temporal dos log retornos da *commoditie* de café possui comportamento não linear. A figura 9 apresenta a série dos log retornos ao quadrado que apresentam forte autocorrelação o que dá indícios de que o modelo GARCH (autorregressivo de heterocedasticidade condicional generalizada) pode ser utilizado para uma melhor modelagem da série.

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
		1	-0.050	-0.050	8.0606 0.005
		2	-0.060	-0.063	19.701 0.000
		3	0.029	0.023	22.413 0.000
		4	-0.017	-0.018	23.299 0.000
		5	-0.016	-0.015	24.160 0.000
		6	0.038	0.034	28.890 0.000
		7	-0.070	-0.068	44.728 0.000
		8	0.003	0.001	44.763 0.000
		9	0.018	0.008	45.831 0.000
		10	0.005	0.011	45.920 0.000
		11	0.019	0.021	47.148 0.000
		12	0.017	0.016	48.066 0.000
		13	0.010	0.019	48.406 0.000
		14	0.016	0.015	49.237 0.000
		15	0.000	0.003	49.238 0.000
		16	-0.025	-0.022	51.302 0.000
		17	-0.010	-0.012	51.605 0.000
		18	0.024	0.023	53.476 0.000
		19	-0.028	-0.025	56.047 0.000
		20	0.018	0.019	57.081 0.000

Figura 9: Correlograma dos quadrados dos resíduos da *commoditie* de café

Uma característica conhecida para séries temporais de retornos financeiros é o comportamento heterocedástico, ou seja, a variância dos dados não é constante para diferentes intervalos de tempo. Isso acontece devido as fortes oscilações do mercado que acontecem de forma inesperada e inconstante.

Após estabelecer que há não linearidade em uma determinada série temporal, é necessário que se identifique o tipo de não linearidade existente, se na média ou na variância condicionada.

Para isso fez-se o teste Hsieh (1989), com as seguintes hipóteses:

: dependência não linear na variância condicionada

: dependência não linear na média

O teste de Hsieh com um nível de confiança de 95%, não rejeitou a hipótese nula dos coeficientes iguais a zero para nenhum valor, o que indica uma não linearidade na variância condicionada.

Com esta análise estatística mostrada anteriormente pode-se concluir que se tem uma não linearidade nos log retornos da *commoditie* de café. Assim, os melhores modelos existentes hoje na literatura para se fazer previsão com dados com característica não linear são os da família GARCH e Redes Neurais Recorrentes. Neste trabalho o foco serão as previsões com Redes Neurais Recorrentes.

Devido à não linearidade dos dados, o modelo de Redes Neurais é aplicado. Foi utilizada neste trabalho uma rede neural como no trabalho de Lima (2011), onde existe uma conexão de realimentação da rede entre os processadores em uma mesma camada e também em camadas diferentes.

A figura a seguir mostra os resultados encontrados:

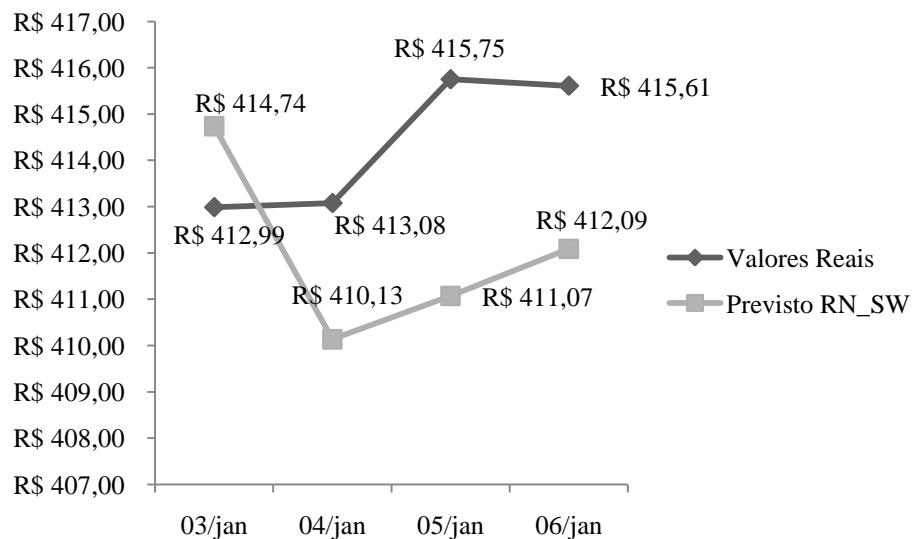


Figura 10: Gráfico da *commoditie* de café real e previsto 4 passos a frente com previsão estatística para Redes Neurais Recorrentes sem uso de ondaletas

A tabela apresenta os indicadores de análise de previsão para as Redes Neurais Recorrentes.

Tabela 4 Estatística de acurácia das RN recorrentes para a *commoditie* de café.

Estatística de acurácia do modelo preditivo	Valores
Correlação	-0,3587
TIC	0,0031
MAPE	0,4541%
RMSE	2,5527

A seguir, de acordo com Lima (2004, 2011), para a série temporal da *commoditie* do café utiliza-se a ondaleta de Daubesch número 1 conforme encontrada na literatura.

A figura 11 mostra os resultados encontrados:

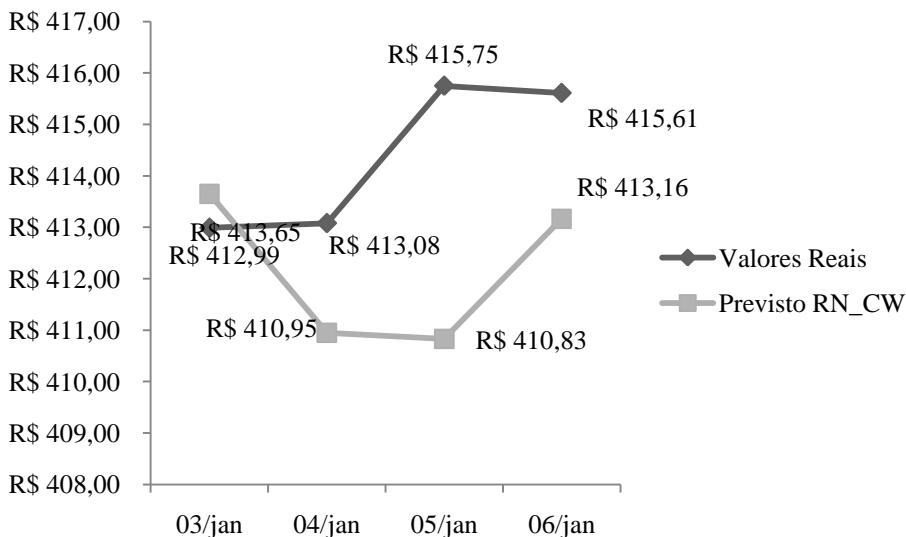


Figura 11: Gráfico da *commodity* de café real e previsto 4 passos a frente com previsão estatística para Redes Neurais Recorrentes com uso das ondaletas.

A tabela apresenta os indicadores de análise de previsão para as Redes Neurais Recorrentes com o uso de ondaletas.

Tabela 5 Estatística de acurácia das RN recorrentes para a *commodity* de café com decomposição de altas e baixas freqüências por *ondaletas*

Estatística de acurácia do modelo preditivo	Valores
Correlação	-0,0103
TIC	0,0024
MAPE	0,3096%
RMSE	1,94814

Os resultados acima apontaram para um desempenho mais ou menos com resultados próximos nos modelos, para a *commodity* de café considerando o efeito do pré-processamento de filtros de altas e baixas freqüências. Tal resultado pode estar relacionado ao período analisado dos dados. Resultados que apontaram para um melhor desempenho das previsões com uso das ondaletas foram encontrados por Lima (2004), Wong et al. (2003) e Homysy, Portugal e Araújo (2000).

Os resultados acima foram inseridos no trabalho apenas para exemplificar como são feitas as previsões com o modelo de Redes Neurais Recorrentes, com e sem uso de filtros de volatilidade como é o caso das ondaletas, para 4 passos previstos, modelos estes, já tradicionais nos trabalhos acadêmicos.

Todavia, sabe-se que as cotações da série inicial testada nessa pesquisa, podem ter pouca, ou nenhuma, influência nas cotações atuais da série do café. Isso também

pode ocorrer com outras séries devido a período de persistência ou não dos dados em relação ao período estudado.

Assim, o próximo passo vem de encontro ao objetivo do trabalho que ilustra os medidores de qualidade em função do expoente de Hurst, estudando justamente a questão da janela de previsão, ou seja, do período de pequenos blocos onde se obteria uma melhor acurácia das métricas de previsão.

Para exemplificar a rotina de apuração da junção do estudo da janela com verificação do expoente de Hurst e a qualidade das previsões, far-se-á inicialmente para uma janela de 500 pontos.

Assim, a série temporal será varrida tomando sempre 500 pontos de cada vez e considerando 496 para estimativa da rede e os 4 últimos para previsão. Como a série do café no período analisado tem 3231 cotações de preços e 3230 log retornos, para a janela de Hurst de 500 pontos, serão feitas 2731 iterações, ou seja, serão calculados 2731 expoentes de Hurst e serão obtidas a mesma quantidade de métricas de acurácia das previsões para cada bloco de 500 pontos.

A seguir, mostra-se o gráfico da série dos preços do café e os respectivos expoentes de Hurst ao longo da série para uma janela de 500 pontos.

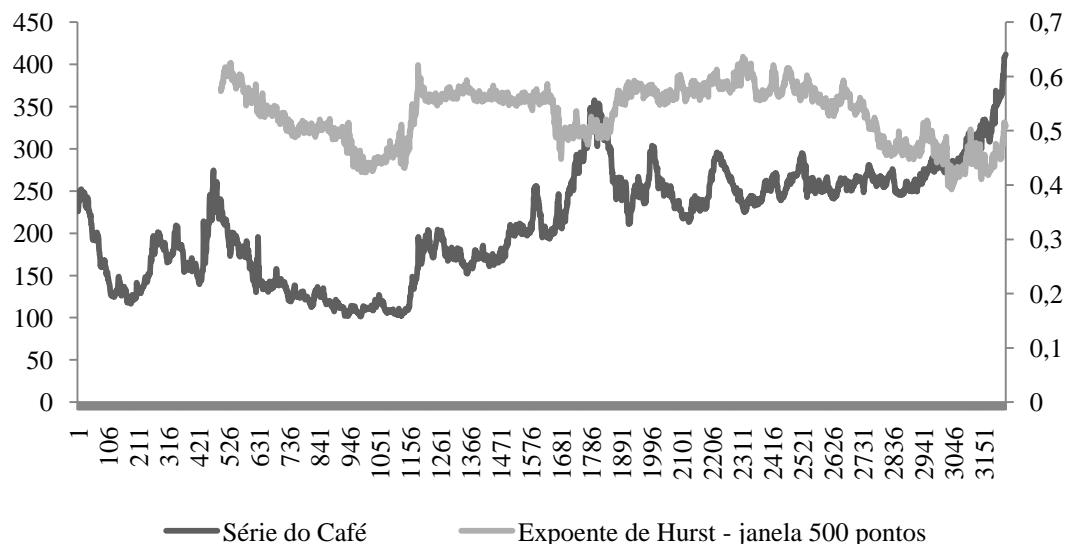


Figura 12: Série do café e os expoentes de Hurst ao longo da série para uma janela de 500 pontos.

Nota-se no gráfico que a série do café apresenta poucos valores do expoente de Hurst abaixo de 0,5. Tais valores ocorreram quando a série dos preços apresentava-se oscilando no mercado sem tomar uma tendência específica. No jargão do mercado

financeiro, diz-se que os preços estão “andando de lado”. Na figura acima, tal fato ocorreu por volta das cotações 410 e 1165 e 2500 e 2700.

Na sequência da análise mostra-se os medidores de qualidade com e sem filtro de ondaletas utilizando o modelo de Redes Neurais Recorrentes, em função do expoente de Hurst.

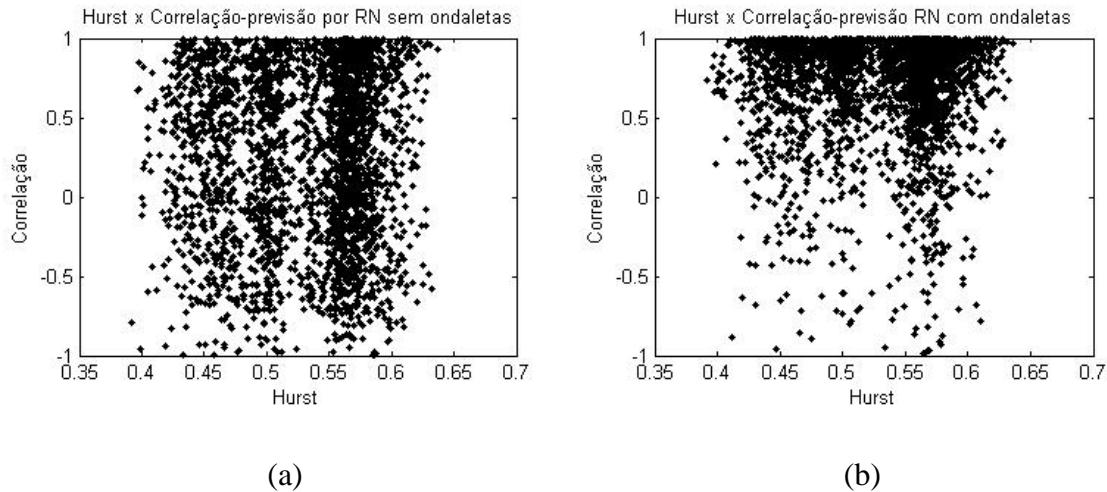


Figura 13: Correlação em função do Expoente de Hurst calculados para janelas de 500 pontos com o modelo de Redes Neurais Recorrentes com e sem filtro de ondaletas para a *commodity* de café.

Na inspeção visual dos pontos destacados na figura acima, nota-se claramente que a correlação aumenta quando se aplica o filtro de ondaletas nas janelas de Hurst como demonstrado na figura 13(b) por estarem mais próximos de um, ao passo que na figura 13(a), os valores das correlações encontram-se mais espalhados.

Em relação ao indicador MAPE, a concentração do erro cai na figura 14(b) com mais pontos próximos ao eixo do que em relação à figura 14(a) como pode ser visto na figura a seguir.

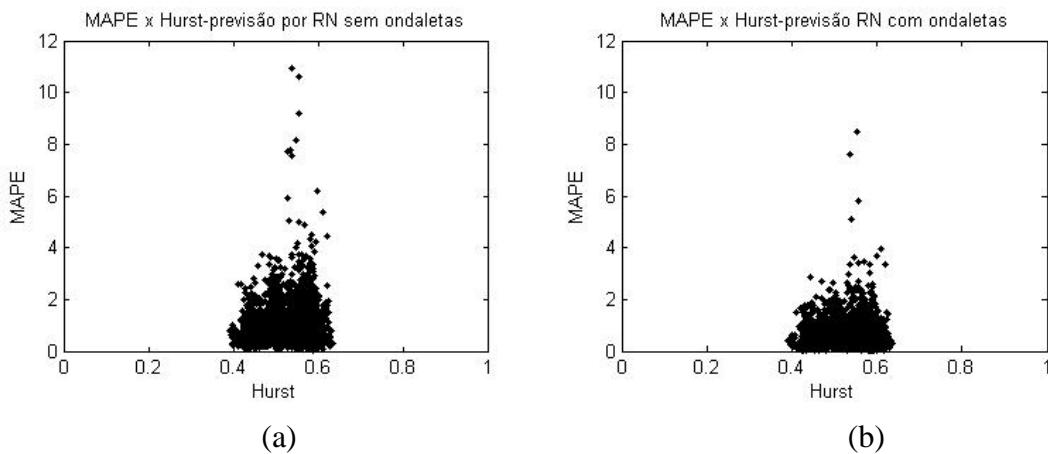


Figura 14: MAPE em função do Exponente de Hurst calculados para uma janela de 500 pontos com o modelo de Redes Neurais Recorrentes com e sem filtro de ondaletas para a *commoditie* de café.

A figura a seguir mostra o medidor de qualidade TIC em função do Exponente de Hurst.

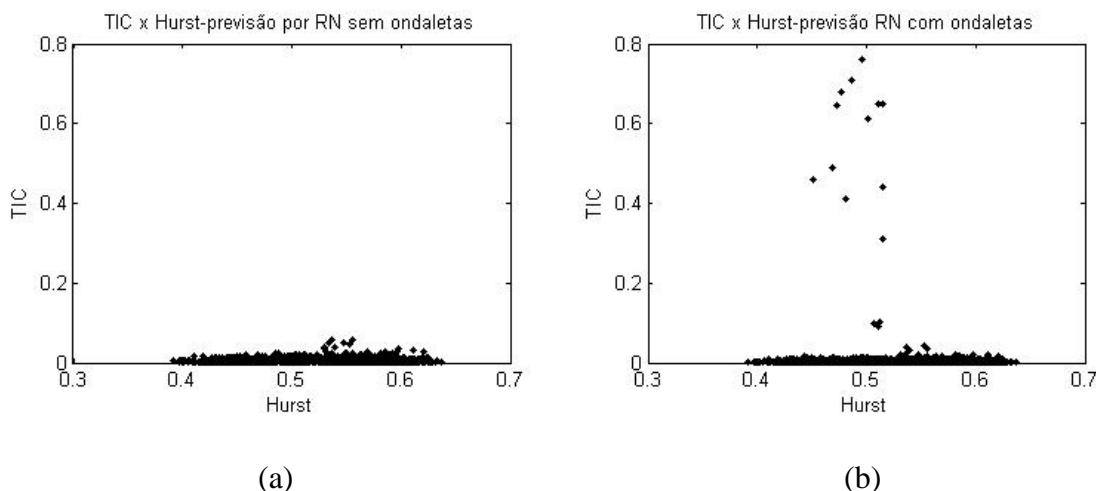


Figura 15: TIC em função do Exponente de Hurst calculados para uma janela de 500 pontos com o modelo de Redes Neurais Recorrentes com e sem filtro de ondaletas para a *commoditie* de café.

E para o RMSE também apresenta uma leve redução do erro a partir da filtragem em alta e baixa frequência e com as previsões feitas nas sub-séries como pode ser visto na figura a seguir.

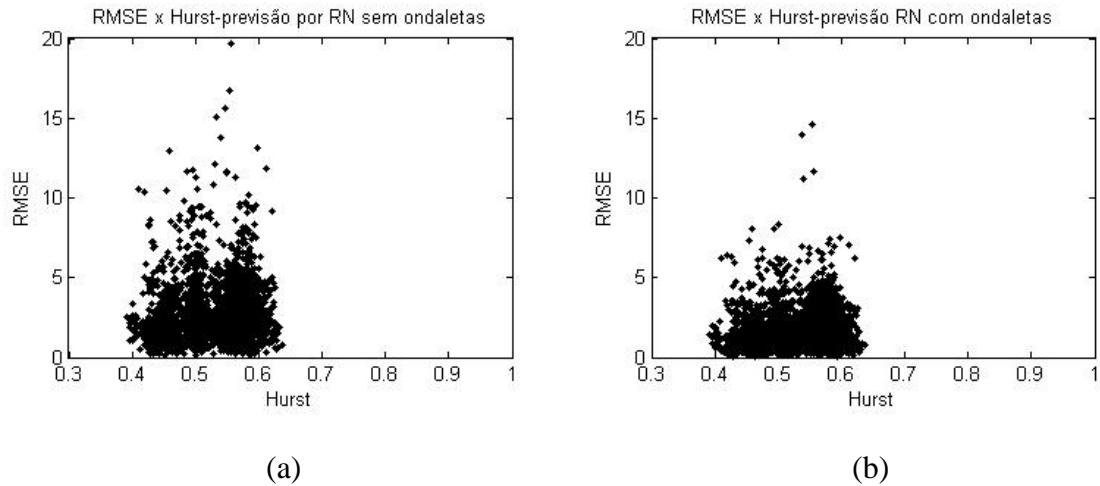


Figura 16: RMSE em função do Exponente de Hurst calculados para uma janela de 500 pontos com o modelo de Redes Neurais Recorrentes com e sem filtro de ondaletas para a *commodity* de café.

Posteriormente, separaram-se os valores do expoente de Hurst onde os valores eram maiores que 0,5 e menores que 0,5. Calculou-se a média dos indicadores de acurácia das previsões em cada conjunto de expoentes e obtiveram-se os seguintes resultados:

Tabela 6 Estatísticas de acurácia das previsões para expoentes de Hurst maiores e menores do que 0,5 para o café.

Hurst < 0,5 – 833 casos					
	CORREL_SW	TIC_SW	MAPE_SW	RMSE_SW	Expoente de Hurst
Média	0,266647	0,005622	0,962319	2,447497	0,462788
DP	0,528131	0,003624	0,618114	1,918933	0,024802

Hurst < 0,5 – 833 casos					
	CORREL_CW	TIC_CW	MAPE_CW	RMSE_CW	Expoente de Hurst
Média	0,681959	0,003607	0,613933	1,522265	0,462788
DP	0,360601	0,055207	0,426428	1,221948	0,024802

Hurst > 0,5 – 1898 casos					
	CORREL_SW	TIC_SW	MAPE_SW	RMSE_SW	Expoente de Hurst
Média	0,268379	0,006749	1,159958	2,831916	0,55948
DP	0,528317	0,004905	0,858277	1,920812	0,028697

Hurst > 0,5 – 1898 casos					
	CORREL_CW	TIC_CW	MAPE_CW	RMSE_CW	Expoente de Hurst
Média	0,703179	0,004174	0,716266	1,7421	0,55948
DP	0,346276	0,028385	0,568874	1,259974	0,028697

De posse das estatísticas acima, aplica-se um teste paramétrico de diferença de médias com as seguintes hipóteses assumindo variâncias diferentes:

Para a correlação, a estatística de teste para a série sem o uso de ondaletas é:

$$\frac{\text{CORREL}_\text{SW}}{\sqrt{\text{MAPE}_\text{SW} \cdot \text{RMSE}_\text{SW}}} \quad (37)$$

O valor p para a estatística acima foi de 0,937. Os demais valores para as outras hipóteses são demonstradas a seguir:

Tabela 7 Teste de hipóteses de diferenças de médias das estatísticas de acurácia com e sem ondaletas para os expoentes de Hurst maiores e menores que 0,5 para o café.

	CORREL_SW	TIC_SW	MAPE_SW	RMSE_SW
t_{teste}	-0,07888	-6,67872	-6,79187	-4,81864
valor p	0,937144	0,0000	0,0000	0,0000
Resultado	Aceita H_0	Rejeita H_0	Rejeita H_0	Rejeita H_0

	CORREL_CW	TIC_CW	MAPE_CW	RMSE_CW
t_{teste}	-1,43302	-4,98092	-5,18982	-4,28753
valor p	0,14555	0,0000	0,0000	0,0000
Resultado	Aceita H_0	Rejeita H_0	Rejeita H_0	Rejeita H_0

De acordo com resultados acima para as previsões feitas com a correlação, nota-se que o expoente de Hurst não mostra indícios de interferir nos resultados das previsões, os testes de hipóteses com 95% de confiança não são conclusivos acerca da influência do expoente de Hurst na capacidade preditiva do café.

Todavia, quando realiza-se o cálculo das previsões e aplica-se os medidores de qualidade TIC, MAPE e RMSE, os indicadores rejeitaram a hipótese nula de igualdade dos indicadores de previsão quando o expoente de Hurst é menor e quando é maior que 0,5. Assim, há indícios significativos de que o expoente de Hurst pode demonstrar influência na janela de resultados usada nas estimativas de previsão de preços futuros de

séries temporais financeiras. Além disso observa-se uma leve melhora nos valores dos indicadores quando a série é filtrada por ondaletas de alta e baixa freqüência.

Agora, tomam-se os mesmos resultados para diferentes configurações de janelas, ou seja, mudam-se os tamanhos das janelas de estimativa do expoente de Hurst mas mantêm-se as previsões de quatro passos a frente para as distintas janelas. Os resultados obtidos são demonstrados nas tabelas a seguir.

Tabela 8 Varredura das diferentes janelas de Hurst e estatísticas de acurácia da correlação para as previsões com e sem ondaletas e teste de significância para a *commoditie* de café

Janela	Expoente de Hurst médio	Número de Exp. Hurst < 0,5	Número de Exp. Hurst > 0,5	CORREL_SW Exp. Hurst < 0,5 - média	CORREL_SW Exp. Hurst > 0,5 - média	T-teste	Valor p	CORREL_CW Exp. Hurst < 0,5 - média	CORREL_CW Exp. Hurst > 0,5 - média	T-teste	Valor p
60	0,555056	1048	2123	0,28173	0,293109	-0,57279	0,567207	0,71607	0,722948	-0,56346	0,574407
80	0,549057	1055	2096	0,285484	0,287408	-0,09737	0,922855	0,714254	0,722447	-0,66089	0,504454
100	0,548192	937	2194	0,277232	0,29007	-0,6321	0,532593	0,720696	0,718341	0,187199	0,852824
120	0,54581	824	2287	0,274552	0,288336	-0,65642	0,519663	0,7068	0,719208	-0,92677	0,353629
150	0,545053	756	2325	0,289108	0,278592	0,485298	0,634149	0,706616	0,714511	-0,55655	0,567526
180	0,547032	701	2350	0,303829	0,273599	1,370887	0,182844	0,69483	0,714913	-1,31802	0,163867
200	0,546238	872	2159	0,275377	0,281766	-0,30265	0,76249	0,693271	0,713797	-1,46836	0,129903
250	0,54385	941	2040	0,283778	0,276086	0,373043	0,711411	0,70112	0,706951	-0,42938	0,663708
300	0,543865	687	2244	0,278384	0,274053	0,19047	0,850705	0,675781	0,706921	-1,97322	0,039107
400	0,532986	850	1981	0,255452	0,278082	-1,03636	0,294949	0,692873	0,702697	-0,68034	0,486903
500	0,529988	833	1898	0,266647	0,268379	-0,07888	0,937144	0,681959	0,703179	-1,43302	0,14555
600	0,537384	586	2045	0,254251	0,262843	-0,35205	0,728808	0,661807	0,709196	-2,65746	0,004179
700	0,531929	488	2043	0,266853	0,258152	0,325007	0,74526	0,684889	0,709742	-1,37097	0,152567

Tabela 9 Varredura das diferentes janelas de Hurst e estatísticas de acurácia do MAPE para as previsões com e sem ondaletas e teste de significância para a *commoditie* de café

Janela	Expoente de Hurst médio	Número de Exp. Hurst < 0,5	Número de Exp. Hurst > 0,5	MAPE_SW Exp. Hurst < 0,5 - média	MAPE_SW Exp. Hurst > 0,5 - média	T-teste	Valor p	MAPE_CW Exp. Hurst < 0,5 - média	MAPE_CW Exp. Hurst > 0,5 - média	T-teste	Valor p
60	0,555056	1048	2123	1,077917	1,128224	-1,61521	0,1302	0,67734	0,707206	-1,4187	0,1701
80	0,549057	1055	2096	1,071661	1,140046	-2,20246	0,0401	0,670394	0,717295	-2,2921	0,0326
100	0,548192	937	2194	1,051774	1,151688	-3,06816	0,0038	0,635003	0,733584	-4,8090	0,0000
120	0,54581	824	2287	1,041089	1,153952	-3,52897	0,0018	0,643377	0,730814	-4,2334	0,0002
150	0,545053	756	2325	1,02158	1,165965	-3,99586	0,0001	0,631246	0,740879	-4,666	0,0000
180	0,547032	701	2350	0,989591	1,175924	-5,72863	0,0000	0,624109	0,745158	-5,7688	0,0000
200	0,546238	872	2159	0,989949	1,192658	-6,82334	0,0000	0,630491	0,75542	-6,2145	0,0000
250	0,54385	941	2040	0,935233	1,231543	-10,1514	0,0000	0,595765	0,777289	-9,2489	0,0000
300	0,543865	687	2244	0,948708	1,196471	-7,91865	0,0000	0,623077	0,749311	-5,9566	0,0000
400	0,532986	850	1981	0,888009	1,248408	-12,0981	0,0000	0,568033	0,780583	-10,499	0,0000
500	0,529988	833	1898	0,962319	1,159958	-6,79187	0,0000	0,613933	0,716266	-5,1898	0,0000
600	0,537384	586	2045	0,978337	1,131637	-4,95497	0,0000	0,620149	0,686763	-3,1185	0,0080
700	0,531929	488	2043	0,962103	1,076145	-3,84671	0,0006	0,585338	0,647663	-3,1427	0,0059

Tabela 10 Varredura das diferentes janelas de Hurst e estatísticas de acurácia do TIC para as previsões com e sem ondaletas e teste de significância para a commodity de café

Janela	Total de Exponentes de Hurst	Número de Exp. Hurst < 0,5	Número de Exp. Hurst > 0,5	TIC_SW Exp. Hurst < 0,5 - média	TIC_SW Exp. Hurst > 0,5 - média	T-teste	Valor p	TIC_CW Exp. Hurst < 0,5 - média	TIC_CW Exp. Hurst > 0,5 - média	T-teste	Valor p
60	0,555056	1048	2123	0,00632	0,006593	-1,49687	0,1589	0,003943	0,004113	-1,3984	0,1750
80	0,549057	1055	2096	0,00627	0,006663	-2,16895	0,0421	0,003901	0,004168	-2,2559	0,0347
100	0,548192	937	2194	0,006142	0,006739	-3,15267	0,0040	0,003726	0,004263	-4,4880	0,0000
120	0,54581	824	2287	0,006113	0,006735	-3,31675	0,0030	0,003781	0,004243	-3,8321	0,0008
150	0,545053	756	2325	0,005981	0,006809	-3,98285	0,0001	0,003687	0,004314	-4,6161	0,0000
180	0,547032	701	2350	0,005808	0,006861	-5,56125	0,0000	0,003655	0,004341	-5,5937	0,0000
200	0,546238	872	2159	0,005818	0,006958	-6,51264	0,0000	0,003686	0,004402	-6,1289	0,0000
250	0,54385	941	2040	0,005501	0,007177	-9,76616	0,0000	0,003494	0,004525	-9,0902	0,0000
300	0,543865	687	2244	0,005592	0,006974	-7,42679	0,0000	0,003647	0,004368	-5,8617	0,0000
400	0,532986	850	1981	0,005202	0,007276	-11,8486	0,0000	0,003347	0,004549	-10,230	0,0000
500	0,529988	833	1898	0,005622	0,006749	-6,67872	0,0000	0,003607	0,004174	-4,9810	0,0000
600	0,537384	586	2045	0,005737	0,006576	-4,61371	0,0001	0,003638	0,003995	-2,8830	0,0130
700	0,531929	488	2043	0,005633	0,00627	-3,68712	0,0010	0,003442	0,003773	-2,8730	0,0112

Tabela 11 Varredura das diferentes janelas de Hurst e estatísticas de acurácia do RMSE para as previsões com e sem ondaletas e teste de significância para a *commoditie* de café

Janela	Total de Exponentes de Hurst	Número de Exp. Hurst < 0,5	Número de Exp. Hurst > 0,5	RMSE_SW Exp. Hurst < 0,5 - média	RMSE_SW Exp. Hurst > 0,5 - média	T-teste	Valor p	RMSE_CW Exp. Hurst < 0,5 - média	RMSE_CW Exp. Hurst > 0,5 - média	T-teste	Valor p
60	0,55506	1048	2123	2,91358	2,57086	4,202	0,0000	1,81462	1,5936	4,079	0,0000
80	0,54906	1055	2096	2,89957	2,59542	3,688	0,0002	1,79594	1,6149	3,368	0,0005
100	0,54819	937	2194	2,89030	2,63688	2,918	0,0024	1,75219	1,6574	1,707	0,0821
120	0,54581	824	2287	2,88486	2,66281	2,526	0,0110	1,78216	1,6630	2,097	0,0365
150	0,54505	756	2325	2,81692	2,71896	1,044	0,2777	1,72555	1,7066	0,313	0,7475
180	0,54703	701	2350	2,68464	2,78026	-1,062	0,3046	1,65970	1,7430	-1,465	0,1703
200	0,54624	872	2159	2,50449	2,87678	-4,727	0,0000	1,55548	1,8030	-4,884	0,0000
250	0,54385	941	2040	2,20641	3,06282	-11,686	0,0000	1,37033	1,9141	-11,525	0,0000
300	0,54387	687	2244	2,18060	2,98858	-10,125	0,0000	1,37708	1,8563	-9,441	0,0000
400	0,53299	850	1981	2,20830	3,07761	-11,119	0,0000	1,37039	1,9129	-10,809	0,0000
500	0,52999	833	1898	2,44750	2,83192	-4,819	0,0000	1,52227	1,7421	-4,288	0,0000
600	0,53738	586	2045	2,33796	2,82324	-5,818	0,0000	1,44383	1,7090	-4,911	0,0000
700	0,53193	488	2043	2,36042	2,74725	-4,356	0,0000	1,45754	1,6478	-3,220	0,0018

Das tabelas anteriores pode-se verificar que, na maioria das vezes, quando o expoente de Hurst é menor que 0,5, a qualidade das previsões melhora significativamente para os indicadores MAPE, TIC e RMSE. A correlação não demonstrou significância em nenhuma das janelas testadas.

As janelas de varredura da série de 100 a 700 pontos, para os medidores MAPE e TIC foram significativas rejeitando-se a hipótese nula. Para o RMSE tivemos janelas significativas de 60 à 120 e de 200 à 700 sem o processamento de ondaletas e 60, 80 e de 200 à 700 com o processamento de ondaletas. Para essas janelas, na maioria dos casos, o erro de previsão apresenta-se menor quando a série é antipersistentes, os únicos casos em que se pode observar que o erro de previsão é menor quando a série é persistente acontece para o medidor de qualidade RMSE para as janelas de 60 à 120 sem o processamento de ondaletas e para as janelas de 60 e 80 com o processamento de ondaletas.

Outras observações importantes, é que os indicadores de qualidade das previsões foram melhores para o uso das previsões feitas com pré-processamento de ondaletas. Até mesmo a correlação entre os pontos reais e os previstos que não demonstrou diferença estatística a 95% de confiança, teve aumento considerável com as previsões feitas com os filtros de ondaletas.

4.2 Análises de outras séries observando o comportamento da qualidade da previsão e o Expoente de Hurst

Nesta etapa do trabalho analisaremos as séries da *commoditie* de soja, taxa de câmbio, índice Ibovespa e as ações com maior volume de negociação na BM&FBOVESPA que são: Petrobrás (PETR4), Vale do Rio Doce (VALE5) e Itaú (ITSA4). No período de 02/01/1998 a 30/12/2010, obtidas através do economática e CEPEA/ESALQ. Como no caso do café, usaremos os medidores de qualidade em função do expoente de Hurst.

4.2.1 Análise *Commoditie* de Soja

A soja é um grão rico em proteínas, cultivado como alimento tanto para humanos quanto para animais. Ela tem origem na China e no Japão e pertence à família *Fabaceae* (leguminosa), assim como o feijão, a lentilha e a ervilha. A palavra soja vem do japonês *shoyu* (KLAUS,2006). Atualmente os maiores produtores do mundo são:

Estados Unidos e Brasil. A safra mundial 2009/2010 foi de 259,89 milhões de toneladas.

Nesta pesquisa analisaram-se os preços diários da *commoditie* de soja, ou seja, o preço à vista da saca de 60kg, no período de 02 de janeiro de 1998 a 30 de dezembro de 2010, perfazendo 3230 observações obtidas junto ao CEPEA-ESALQ.

Como para a *commoditie* de café, primeiramente exemplifica-se a rotina de apuração da junção do estudo da janela com verificação do expoente de Hurst e a qualidade das previsões, inicialmente para uma janela de 500 pontos.

Assim, a série temporal será varrida tomando sempre 500 pontos de cada vez e considerando 496 para estimativa da rede e os 4 últimos para previsão. Como a série da soja no período analisado tem 3230 cotações de preços e 3229 log retornos, para a janela de Hurst de 500 pontos, serão feitas 2730 iterações, ou seja, serão calculados 2730 expoentes de Hurst e serão obtidas a mesma quantidade de métricas de acurácia das previsões para cada bloco de 500 pontos.

A seguir, mostra-se o gráfico da série dos preços da soja e os respectivos expoentes de Hurst ao longo da série para uma janela de 500 pontos.

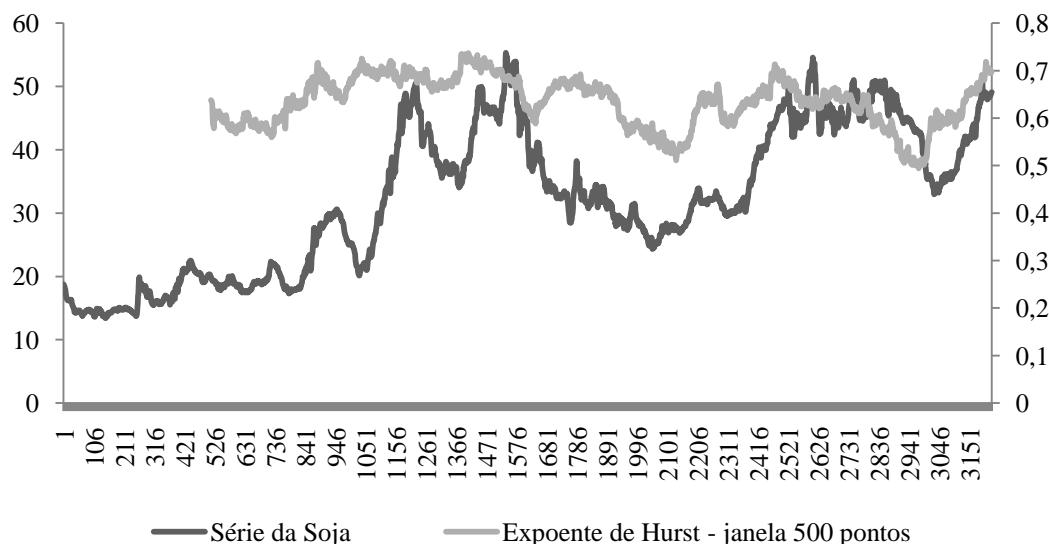


Figura 17: Série da soja e os expoentes de Hurst ao longo da série para uma janela de 500 pontos.

Nota-se no gráfico que a série da soja apresenta poucos valores do expoente de Hurst abaixo de 0,5. Na sequência da análise mostram-se os medidores de qualidade com e sem filtro de ondaletas utilizando o modelo de Redes Neurais Recorrentes, em função do expoente de Hurst.

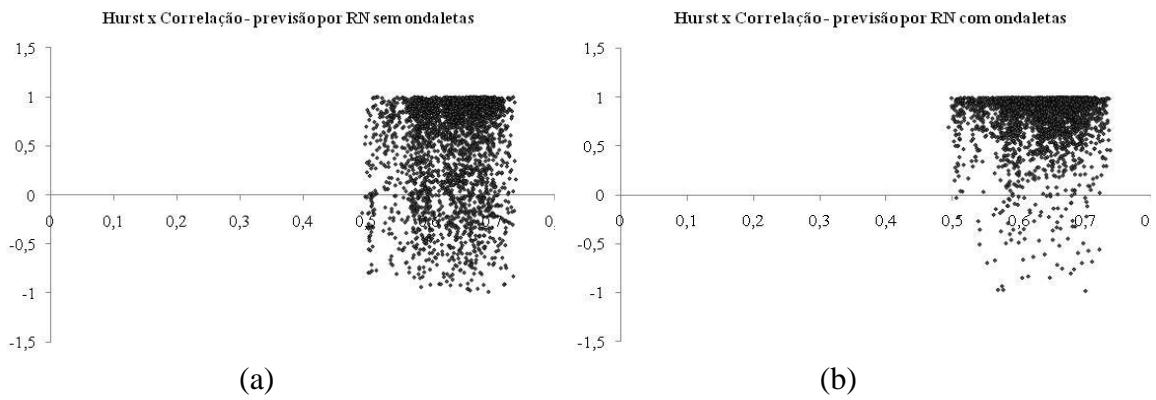


Figura 18: Correlação em função do Exponente de Hurst calculados para uma janela de 500 pontos com o modelo de Redes Neurais Recorrentes com e sem filtro de *ondaletas* para a *commodity* de soja.

Na inspeção visual dos pontos destacados na figura acima, nota-se claramente que a correlação aumenta quando se aplica o filtro de ondaletas nas janelas de Hurst como demonstrado na figura 18(b) por estarem mais próximos de um, ao passo que na figura 18(a), os valores das correlações encontram-se mais espalhados.

Em relação ao indicador MAPE, a concentração do erro cai na figura 19(b) com mais pontos próximos ao eixo do que em relação à figura 19(a) como pode ser visto na figura a seguir.

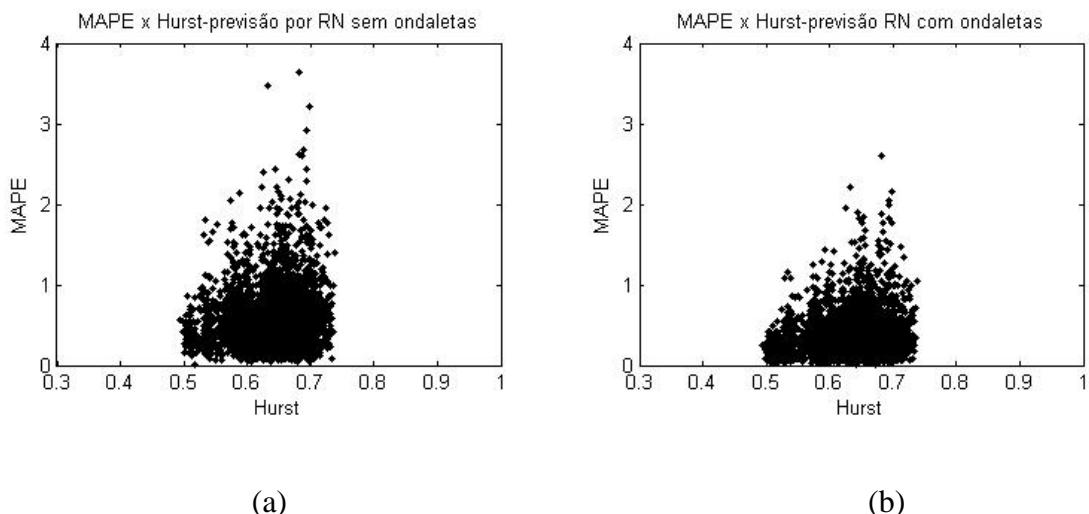


Figura 19: MAPE em função do Expoente de Hurst calculados para uma janela de 500 pontos com o modelo de Redes Neurais Recorrentes com e sem filtro de ondaletas para a *commoditie* de soja.

A figura a seguir apresenta graficamente o TIC em função do expoente de Hurst:

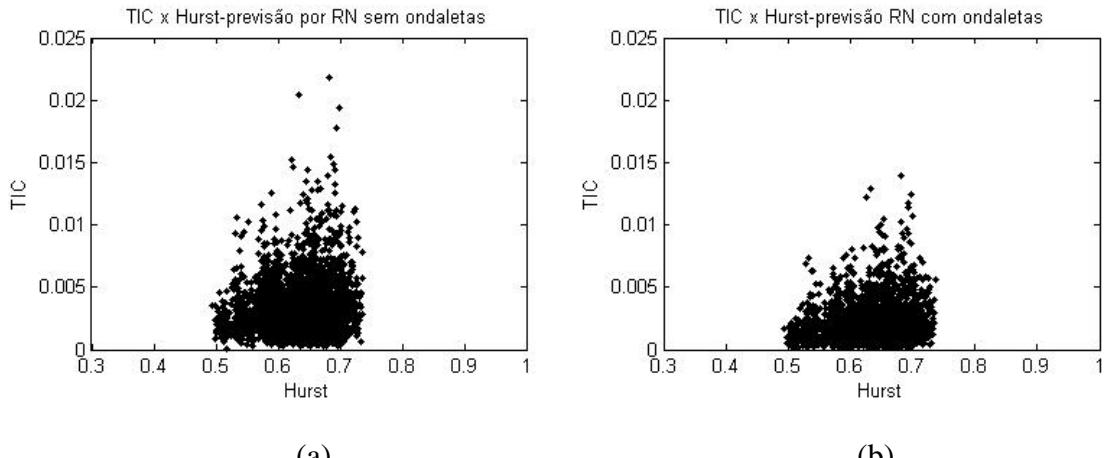


Figura 20: TIC em função do Exponente de Hurst calculados para uma janela de 500 pontos com o modelo de Redes Neurais Recorrentes com e sem filtro de ondaletas para a *commoditie* de soja.

E para o RMSE também apresenta redução do erro a partir da filtragem em alta e baixa freqüência e com as previsões feitas nas sub-séries como pode ser visto na figura a seguir.

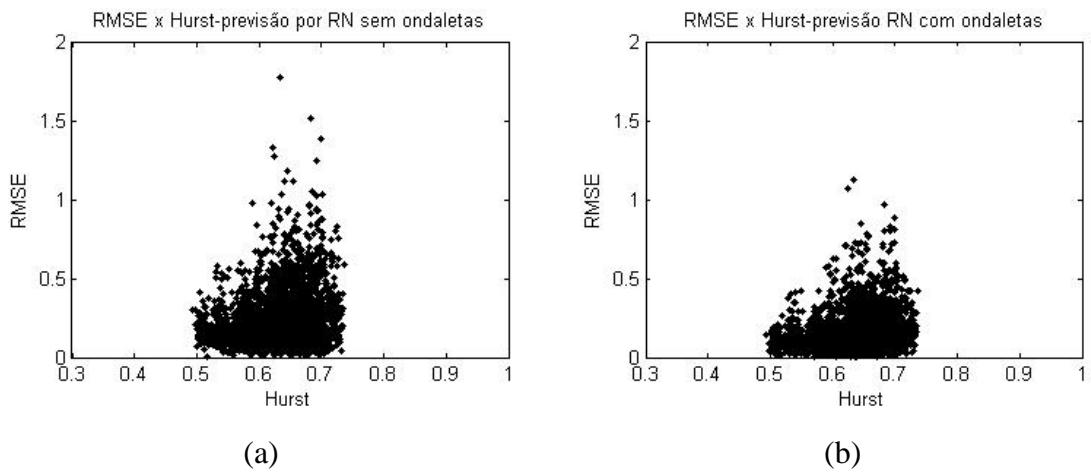


Figura 21: RMSE em função do Exponente de Hurst calculados para uma janela de 500 pontos com o modelo de Redes Neurais Recorrentes com e sem filtro de ondaletas para a *commoditie* de soja.

Posteriormente, separaram-se os valores do expoente de Hurst onde os valores eram maiores que meio e menores que meio. Calculou-se a média dos indicadores de acurácia das previsões em cada conjunto de expoentes e obtiveram-se os seguintes resultados:

Tabela 12 Estatísticas de acurácia das previsões para expoentes de Hurst maiores e menores do que 0,5 para a soja.

Hurst < 0,5 – 4 casos					
	CORREL_SW	TIC_SW	MAPE_SW	RMSE_SW	Expoente de Hurst
Média	0,225768	0,002541	0,421144	0,2197035	0,497694
DP	0,387974	0,001263	0,186744	0,1088669	0,002552

Hurst < 0,5 – 4 casos					
	CORREL_CW	TIC_CW	MAPE_CW	RMSE_CW	Expoente de Hurst
Média	0,876315	0,000908	0,153114	0,07856902	0,497694
DP	0,126792	0,000525	0,084893	0,04526001	0,002552

Hurst > 0,5 – 2726 casos					
	CORREL_SW	TIC_SW	MAPE_SW	RMSE_SW	Expoente de Hurst
Média	0,470801	0,003399	0,586925	0,2417102	0,634995
DP	0,051828	0,002326	0,401116	0,1841337	0,051828

Hurst > 0,5 – 2726 casos					
	CORREL_CW	TIC_CW	MAPE_CW	RMSE_CW	Expoente de Hurst
Média	0,791437	0,002158	0,374907	0,1538865	0,634995
DP	0,289316	0,001638	0,2882	0,1300759	0,051828

De posse das estatísticas acima, aplica-se um teste paramétrico de diferença de médias com as seguintes hipóteses assumindo variâncias diferentes:

Para a correlação, a estatística de teste é:

$$\frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}$$

O valor p para a estatística acima foi de 0,296. Os demais valores para as outras hipóteses são demonstradas a seguir:

Tabela 13 Teste de hipóteses de diferenças de médias das estatísticas de acurácia com e sem ondaletas para os expoentes de Hurst maiores e menores que 0,5 para a soja.

	CORREL_SW	TIC_SW	MAPE_SW	RMSE_SW
t _{teste}	- 1,2615	- 1,355	- 1,768	-0,4034
valor p	0,2962	0,2683	0,1749	0,7137
Resultado	Aceita H ₀	Aceita H ₀	Aceita H ₀	Aceita H ₀

	CORREL_CW	TIC_CW	MAPE_CW	RMSE_CW
t _{teste}	1,3337	-4,723	- 5,1816	- 3,3082
valor p	0,2745	0,018	0,014	0,045
Resultado	Aceita H ₀	Rejeita H ₀	Rejeita H ₀	Rejeita H ₀

De acordo com resultados acima, para as previsões feitas sem pré-processamento de ondaletas em altas e baixas freqüências, nota-se que o expoente de Hurst não mostra indícios de interferir nos resultados das previsões, ou seja, mesmo quando é menor que 0,5 e apresenta comportamento anti-persistente, os testes de hipóteses com 95% de confiança não são conclusivos a cerca da influência do expoente de Hurst na capacidade preditiva da série da soja.

Todavia, quando realiza-se a separação da série em altas e baixas freqüências, e realizam-se as previsões dentro das sub-séries, nota-se que somente a correlação não demonstra diferenças nas estatísticas de acurácia. Todos os demais indicadores rejeitaram a hipótese nula de igualdade dos indicadores de previsão quando o expoente de Hurst é menor e quando é maior que 0,5. Assim, há indícios significativos de que o expoente de Hurst pode demonstrar influência na janela de resultados usada nas estimativas de previsão de preços futuros de séries temporais financeiras.

Agora, tomam-se os mesmos resultados para diferentes configurações de janelas, ou seja, mudam-se os tamanhos das janelas de estimativa do expoente de Hurst mas mantêm-se as previsões de quatro passos à frente para as distintas janelas. Os resultados obtidos são demonstrados nas tabelas a seguir.

Até esse ponto detectou-se que no cruzamento dos dados do expoente de Hurst com o uso de filtragem por ondaletas, o erro reduz quando se aplica o processamento e as previsões dentro de cada série de alta e baixa freqüência.

A partir de agora, será mudada a janela de Hurst e obtidas as previsões e as estatísticas de erro para cada janela e seu respectivo teste de hipótese de diferença de médias para cada métrica visando verificar se a mudança do expoente de Hurst acima e abaixo de 0,5 influenciam nos resultados das previsões.

A seguir, tem-se as tabelas com as respectivas janelas e os indicadores de erro das previsões:

Tabela 14 Varredura das diferentes janelas de Hurst e estatísticas de acurácia da correlação para as previsões com e sem ondaletas e teste de significância para a commodity de soja

Janela	Expoente de Hurst médio	Número de Exp. Hurst < 0,5	Número de Exp. Hurst > 0,5	CORREL_SW Exp. Hurst < 0,5 - média	CORREL_SW Exp. Hurst > 0,5 - média	T-teste	Valor p	CORREL_CW Exp. Hurst < 0,5 - média	CORREL_CW Exp. Hurst > 0,5 - média	T-teste	Valor p
60	0,656143	520	2650	0,4451	0,4926	-2,041	0,0464	0,7825	0,7958	-1,066	0,3108
80	0,657284	349	2801	0,4482	0,4861	-1,3126	0,1773	0,7851	0,7935	-0,5472	0,5924
100	0,659255	260	2870	0,4623	0,4859	-0,7651	0,4631	0,7857	0,7941	-0,4612	0,6439
120	0,656621	214	2896	0,4962	0,4842	0,3633	0,7335	0,7991	0,7943	0,2541	0,8113
150	0,657659	168	2912	0,4545	0,4854	-0,8590	0,4312	0,7750	0,7956	-0,864	0,3556
180	0,657676	51	2999	0,5128	0,4820	0,5447	0,6611	0,7692	0,7941	-0,6159	0,5336
200	0,655109	85	2945	0,4283	0,4836	-1,0847	0,3133	0,7915	0,7934	-0,0652	0,9520
250	0,654922	50	2930	0,3052	0,4774	-2,1605	0,0160	0,7854	0,7924	-0,1605	0,8626
300	0,658501	4	2926	0,0496	0,4753	-1,2522	0,0896	0,9442	0,7938	7,5096	0,2814
400	0,644054	15	2815	0,3900	0,4722	-0,5502	0,5284	0,8525	0,7935	1,3561	0,4225
500	0,634794	4	2726	0,2258	0,4708	-1,2616	0,3298	0,8763	0,7914	1,3338	0,5575
600	0,645325	0	2630	0,3899	0,4740	-	-	0,8525	0,7944	-	-
700	0,642785	0	2530	0,3899	0,4824	-	-	0,8525	0,7967	-	-

Tabela 15 Varredura das diferentes janelas de Hurst e estatísticas de acurácia do MAPE para as previsões com e sem ondaletas e teste de significância para a *commoditie* de soja

Janela	Expoente de Hurst médio	Número de Exp. Hurst < 0,5	Número de Exp. Hurst > 0,5	MAPE_SW Exp. Hurst < 0,5 - média	MAPE_SW Exp. Hurst > 0,5 - média	T-teste	Valor p	MAPE_CW Exp. Hurst < 0,5 - média	MAPE_CW Exp. Hurst > 0,5 - média	T-teste	Valor p
60	0,656143	520	2650	0,6512	0,5630	4,2220	0,0001	0,4182	0,3706	3,2781	0,0005
80	0,657284	349	2801	0,7181	0,5629	5,3070	0,0000	0,4606	0,3682	4,4799	0,0000
100	0,659255	260	2870	0,7175	0,5687	4,6539	0,0000	0,4666	0,3703	4,0639	0,0000
120	0,656621	214	2896	0,7365	0,5693	5,6485	0,0000	0,4767	0,3697	4,6813	0,0000
150	0,657659	168	2912	0,7130	0,5726	4,6649	0,0000	0,4537	0,3715	3,6321	0,0005
180	0,657676	51	2999	0,5950	0,5837	0,2821	0,8484	0,3573	0,3782	-0,7388	0,6176
200	0,655109	85	2945	0,4897	0,5892	-3,7094	0,0315	0,3033	0,3816	-3,8080	0,0165
250	0,654922	50	2930	0,4794	0,5968	-3,3651	0,0501	0,2845	0,3848	-3,6495	0,0178
300	0,658501	4	2926	0,2886	0,5876	-12,9156	0,1418	0,1147	0,3783	-17,491	0,0695
400	0,644054	15	2815	0,4907	0,5821	-1,9723	0,3797	0,2442	0,3727	-4,1797	0,0835
500	0,634794	4	2726	0,4211	0,5869	-1,7695	0,4087	0,1531	0,3749	-5,1816	0,1240
600	0,645325	0	2630	0,4906	0,5712	-	-	0,2442	0,3655	-	-
700	0,642785	0	2530	0,4906	0,5701	-	-	0,2442	0,3651	-	-

Tabela 16 Varredura das diferentes janelas de Hurst e estatísticas de acurácia do TIC para as previsões com e sem ondaletas e teste de significância para a *commoditie* de soja

Janela	Total de Exponentes de Hurst	Número de Exp. Hurst < 0,5	Número de Exp. Hurst > 0,5	TIC_SW Exp. Hurst < 0,5 - media	TIC_SW Exp. Hurst > 0,5 - média	T-teste	Valor p	TIC_CW Exp. Hurst < 0,5 - média	TIC_CW Exp. Hurst > 0,5 - média	T-teste	Valor p
60	0,656143	520	2650	0,0038	0,0033	4,1348	0,0000	0,0024	0,0021	3,3081	0,0005
80	0,657284	349	2801	0,0041	0,0033	5,1805	0,0000	0,0026	0,0021	4,4195	0,0000
100	0,659255	260	2870	0,0042	0,0033	4,6638	0,0000	0,0027	0,0021	3,9700	0,0000
120	0,656621	214	2896	0,0043	0,0033	5,7126	0,0000	0,0027	0,0021	4,6233	0,0000
150	0,657659	168	2912	0,0042	0,0033	4,9279	0,0000	0,0026	0,0021	3,7382	0,0004
180	0,657676	51	2999	0,0035	0,0034	0,2888	0,8481	0,0021	0,0022	-0,5352	0,7178
200	0,655109	85	2945	0,0028	0,0034	-3,9561	0,0260	0,0017	0,0022	-4,1011	0,0150
250	0,654922	50	2930	0,0028	0,0035	-3,3428	0,0567	0,0017	0,0022	-3,3335	0,0262
300	0,658501	4	2926	0,0017	0,0034	-11,6964	0,1608	0,0006	0,0022	-22,458	0,0628
400	0,644054	15	2815	0,0028	0,0034	-2,4313	0,3349	0,0014	0,0021	-4,3374	0,0782
500	0,634794	4	2726	0,0025	0,0034	-1,3554	0,4609	0,0009	0,0022	-4,7238	0,1272
600	0,645325	0	2630	0,0027	0,0033	-	-	0,0014	0,0021	-	-
700	0,642785	0	2530	0,0027	0,0033	-	-	0,0014	0,0021	-	-

Tabela 17 Varredura das diferentes janelas de Hurst e estatísticas de acurácia do RMSE para as previsões com e sem ondaletas e teste de significância para a *commoditie* de soja

Janela	Total de Expoentes de Hurst	Número de Exp. Hurst < 0,5	Número de Exp. Hurst > 0,5	RMSE_SW Exp. Hurst < 0,5 - média	RMSE_SW Exp. Hurst > 0,5 - média	T-teste	Valor p	RMSE_CW Exp. Hurst < 0,5 - média	RMSE_CW Exp. Hurst > 0,5 - média	T-teste	Valor p
60	0,656143	520	2650	0,2541	0,2155	4,2666	0,0000	0,1644	0,1414	3,6301	0,0001
80	0,657284	349	2801	0,2931	0,2147	6,2045	0,0000	0,1882	0,1401	5,4821	0,0000
100	0,659255	260	2870	0,3105	0,2166	6,6131	0,0000	0,1992	0,1407	5,7289	0,0000
120	0,656621	214	2896	0,3274	0,2178	7,8985	0,0000	0,2103	0,1410	6,5901	0,0000
150	0,657659	168	2912	0,3293	0,2206	7,3267	0,0000	0,2070	0,1428	5,8789	0,0000
180	0,657676	51	2999	0,2182	0,2288	-0,6938	0,6815	0,1336	0,1478	-1,2437	0,4316
200	0,655109	85	2945	0,2051	0,23067	-2,3057	0,2010	0,1268	0,1491	-2,7100	0,1131
250	0,654922	50	2930	0,2212	0,2339	-0,6015	0,6223	0,1302	0,1506	-1,3823	0,2611
300	0,658501	4	2926	0,1569	0,2340	-6,1211	0,3959	0,0575	0,1503	-15,522	0,1464
400	0,644054	15	2815	0,2609	0,2367	1,1091	0,6087	0,1310	0,1510	-1,2556	0,5465
500	0,634794	4	2726	0,2197	0,2417	-0,4034	0,8112	0,0786	0,1539	-3,3082	0,2471
600	0,645325	0	2630	26,09	0,2426	-	-	0,1310	15,47	-	-
700	0,642785	0	2530	26,09	0,2474	-	-	0,1310	15,77	-	-

Das tabelas anteriores se pode verificar que, na maioria dos casos, quando o expoente de Hurst é maior que 0,5, a qualidade das previsões melhora significativamente para os indicadores MAPE, TIC e RMSE. A correlação não demonstrou significância em nenhuma das janelas testadas.

As janelas de varredura da série de 60 a 150 pontos foram significativas a 95% de confiança e rejeitam a hipótese nula e há uma melhor capacidade preditiva para a série persistente para os medidores MAPE, TIC e RMSE. A partir de uma janela de 180 pontos o teste não consegue enxergar os parâmetros de médias diferentes, aceitando-se a hipótese nula para o RMSE, já para a janela de 200 pontos com o processamento de ondaletas para o TIC e de 200 e 250 também com o processamento de ondaletas para o MAPE rejeita-se a hipótese nula, e observa-se uma melhora na qualidade das previsões para a série antipersistente.

Outra observação importante, é que os indicadores de qualidade das previsões foram melhores para o uso das previsões feitas com pré-processamento de ondaletas. Até mesmo a correlação entre os pontos reais e os previstos, que não demonstrou diferença estatística a 95% de confiança, teve aumento considerável com a qualidade das previsões feitas com os filtros de ondaletas.

4.2.2 Análise da taxa de câmbio

Taxa de câmbio é o preço de uma moeda estrangeira medido em unidades ou frações (centavos) da moeda nacional. No Brasil, a moeda estrangeira mais negociada é o dólar dos Estados Unidos, fazendo com que a cotação comumente utilizada seja a dessa moeda. Dessa forma, quando dizemos, por exemplo, que a taxa de câmbio é 2,00, significa que um dólar dos Estados Unidos custa R\$ 2,00. A taxa de câmbio reflete, assim, o custo de uma moeda em relação à outra. As cotações apresentam taxas para a compra e para a venda da moeda, as quais são referenciadas do ponto de vista do agente autorizado a operar no mercado de câmbio pelo Banco Central.

Para este trabalho foram analisadas as taxas de cambio tomadas no período de 02/01/1998 quando representava 1,789 até 30/12/2010, quando apresentou um valor de 1,6662. Nesse período, houve uma desvalorização de -6,86%. As taxas foram consideradas de fechamento no período diário num total de 2763 observações.

Inicialmente exemplificamos a rotina de apuração do estudo da janela da taxa de câmbio para verificação do expoente de Hurst e a qualidade das previsões para uma janela de 500 pontos.

Assim, a série temporal será varrida tomando sempre 500 pontos de cada vez e considerando 496 para estimação da rede e os 4 últimos para previsão. Como a série da taxa de cambio no período analisado tem 2763 cotações de preços e 2762 log retornos, para a janela de Hurst de 500 pontos, serão feitas 2263 iterações, ou seja, serão calculados 2263 expoentes de Hurst e serão obtidas a mesma quantidade de métricas de acurácia das previsões para cada bloco de 500 pontos.

A seguir, mostra-se o gráfico da série das taxas de cambio e os respectivos expoentes de Hurst ao longo da série para uma janela de 500 pontos.

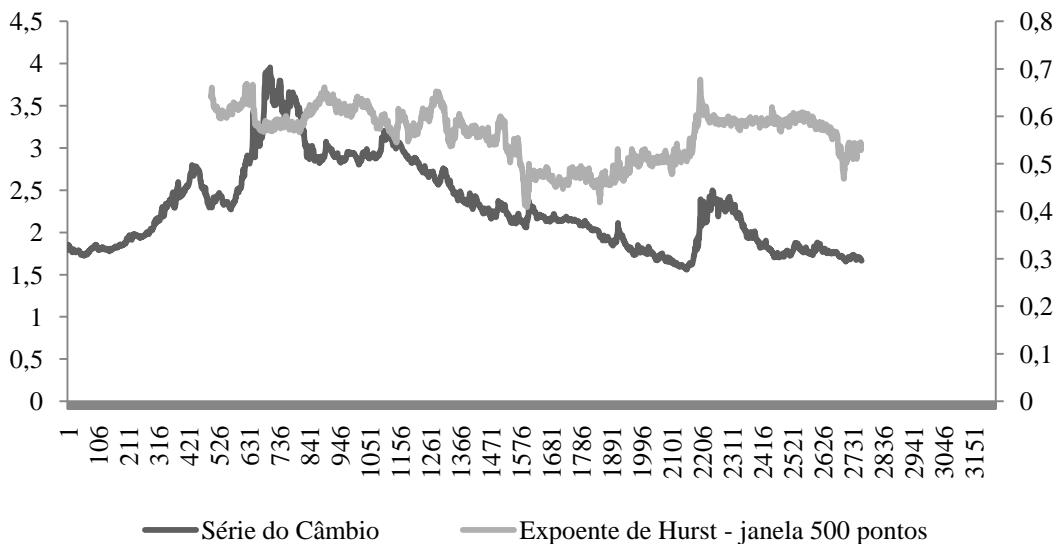


Figura 22: Série da taxa de cambio e os expoentes de Hurst ao longo da série para uma janela de 500 pontos.

Nota-se no gráfico que a série da taxa de câmbio apresenta poucos valores do expoente de Hurst abaixo de 0,5. Na sequência da análise mostram-se os medidores de qualidade com e sem filtro de ondaletas utilizando o modelo de Redes Neurais Recorrentes, em função do expoente de Hurst.

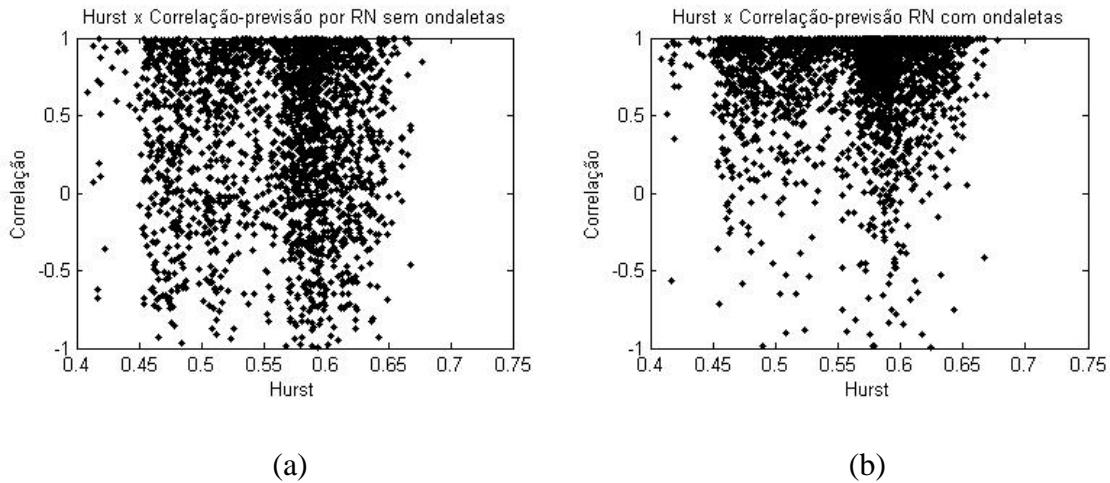


Figura 23: Correlação em função do Exponente de Hurst calculados para uma janela de 500 pontos com o modelo de Redes Neurais Recorrentes com e sem filtro de ondaletas para a taxa de câmbio.

Na inspeção visual dos pontos destacados na figura acima, nota-se claramente que a correlação aumenta quando se aplica o filtro de ondaletas nas janelas de Hurst como demonstrado na figura 23(b) por estarem mais próximos de um, ao passo que na figura 23(a), os valores das correlações encontram-se mais espalhados.

Em relação ao indicador MAPE, a concentração do erro cai na figura 24(b) com mais pontos próximos ao eixo do que em relação à figura 24(a) como pode ser visto na figura a seguir.

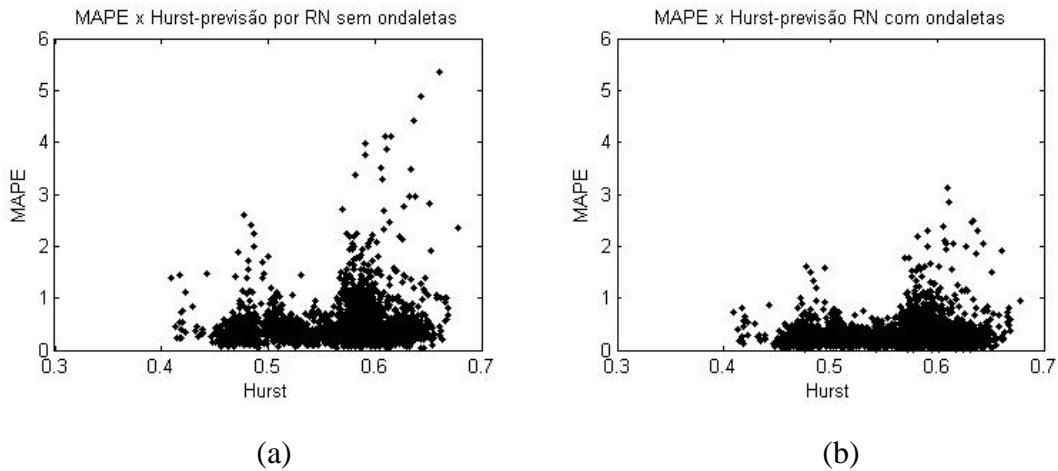


Figura 24: MAPE em função do Exponente de Hurst calculados para uma janela de 500 pontos com o modelo de Redes Neurais Recorrentes com e sem filtro de ondaletas para a taxa de câmbio.

Em relação ao TIC, a análise é muito semelhante, com redução do erro apresentando maior concentração pelo uso do pré-processamento por ondaletas como pode ser visto na figura a seguir.

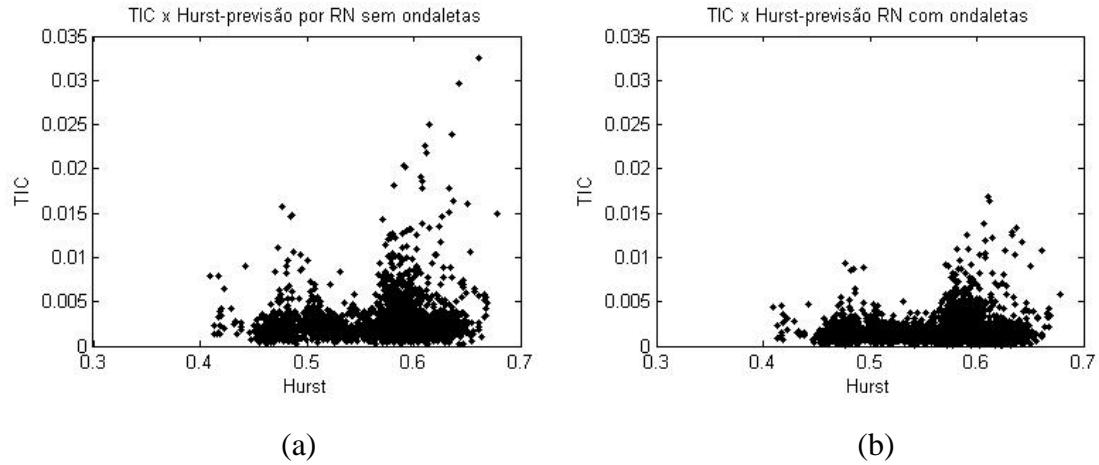


Figura 25: TIC em função do Exponente de Hurst calculados para uma janela de 500 pontos com o modelo de Redes Neurais Recorrentes com e sem filtro de ondaletas para a taxa de câmbio.

E para o RMSE também apresenta redução do erro a partir da filtragem em alta e baixa freqüência e com as previsões feitas nas sub-séries como pode ser visto na figura 26.

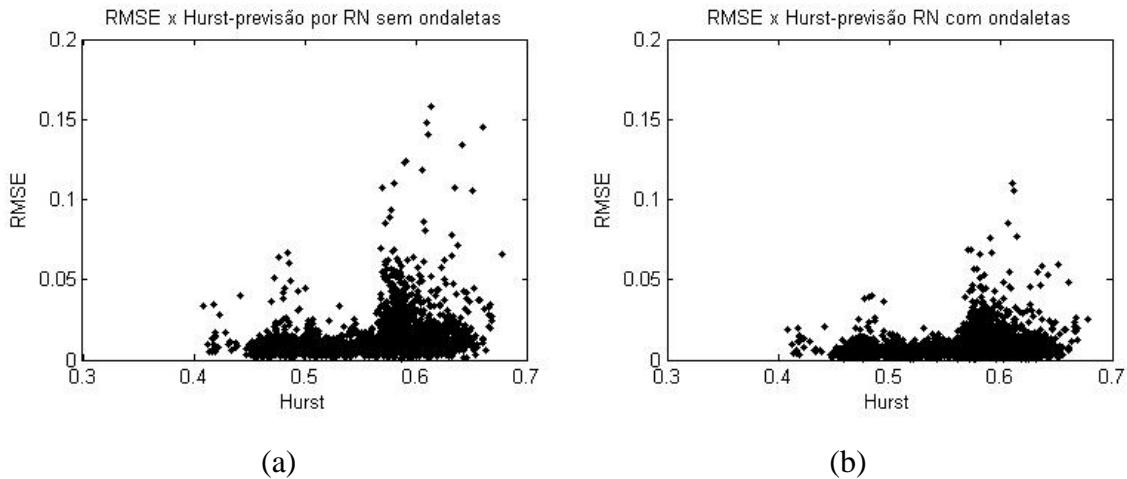


Figura 26: RMSE em função do Exponente de Hurst calculados para uma janela de 500 pontos com o modelo de Redes Neurais Recorrentes com e sem filtro de ondaletas para a taxa de câmbio.

Posteriormente, separaram-se os valores do expoente de Hurst onde os valores eram maiores que meio e menores que meio. Calculou-se a média dos indicadores de

acurácia das previsões em cada conjunto de expoentes e obtiveram-se os seguintes resultados:

Tabela 18 Estatísticas de acurácia das previsões para expoentes de Hurst maiores e menores do que 0,5 para o câmbio.

Hurst < 0,5 – 409 casos					
	CORREL_SW	TIC_SW	MAPE_SW	RMSE_SW	Expoente de Hurst
Média	0,358035	0,002351	0,405345	0,0096317	0,472313
DP	0,522793	0,001962	0,330466	0,0083406	0,016974

Hurst < 0,5 – 409 casos					
	CORREL_CW	TIC_CW	MAPE_CW	RMSE_CW	Expoente de Hurst
Média	0,703697	0,001536	0,264303	0,0062736	0,472313
DP	0,326346	0,001231	0,213854	0,0051791	0,016974

Hurst > 0,5 – 1854 casos					
	CORREL_SW	TIC_SW	MAPE_SW	RMSE_SW	Expoente de Hurst
Média	0,395787	0,003108	0,53697	0,0153477	0,581529
DP	0,505662	0,002677	0,467097	0,0149184	0,036525

Hurst > 0,5 – 1854 casos					
	CORREL_CW	TIC_CW	MAPE_CW	RMSE_CW	Expoente de Hurst
Média	0,723211	0,002096	0,358309	0,0100601	0,581529
DP	0,338718	0,001736	0,306105	0,0094148	0,036525

De posse das estatísticas acima, aplica-se um teste paramétrico de diferença de médias com as seguintes hipóteses assumindo variâncias diferentes:

Para a correlação, a estatística de teste é:

O valor p para a estatística acima foi de 0,1745. Os demais valores para as outras hipóteses são demonstradas a seguir:

Tabela 19 Teste de hipóteses de diferenças de médias das estatísticas de acurácia com e sem ondaletas para os expoentes de Hurst maiores e menores que 0,5 para o câmbio.

	CORREL_SW	TIC_SW	MAPE_SW	RMSE_SW
t _{teste}	-1,33	-6,57411	-6,71094	-10,6119
valor p	0,1745	0,0000	0,0000	0,00000
Resultado	Aceita H ₀	Rejeita H ₀	Rejeita H ₀	Rejeita H ₀

	CORREL_CW	TIC_CW	MAPE_CW	RMSE_CW
t _{teste}	-1,08701	-7,37773	-7,66726	-11,2448
valor p	0,2885	0,0000	0,0000	0,00000
Resultado	Aceita H ₀	Rejeita H ₀	Rejeita H ₀	Rejeita H ₀

De acordo com resultados acima, para as previsões feitas com a correlação, nota-se que o teste t aceita a hipótese nula, não apresentando uma diferenciação entre os parâmetros das médias, os testes de hipóteses com 95% de confiança não são conclusivos cerca da influência do expoente de Hurst na capacidade preditiva da taxa de câmbio.

Todavia, quando realiza-se o cálculo das previsões e aplicam-se os medidores de qualidade TIC, MAPE e RMSE, os indicadores rejeitaram a hipótese nula de igualdade dos indicadores de previsão quando o expoente de Hurst é menor e quando é maior que 0,5. Assim, há indícios significativos de que o expoente de Hurst pode demonstrar influência na janela de resultados usada nas estimativas de previsão de valores futuros da taxa de câmbio. Além disso observa-se uma melhora nos valores dos indicadores quando a série é filtrada por ondaletas de alta e baixa freqüência.

Agora, tomam-se os mesmos resultados para diferentes configurações de janelas, ou seja, mudam-se os tamanhos das janelas de estimativa do expoente de Hurst mas mantêm-se as previsões de quatro passos à frente para as distintas janelas. Os resultados obtidos são demonstrados no quadro a seguir.

A partir de agora, será mudada a janela de Hurst e obtidas as previsões e as estatísticas de erro para cada janela e seu respectivo teste de hipótese de diferença de médias para cada métrica visando verificar se a mudança do expoente de Hurst acima e

abaixo de 0,5 influenciam nos resultados das previsões. O que pode ser observado nas tabelas a seguir.

Tabela 20 Varredura das diferentes janelas de Hurst e estatísticas de acurácia da correlação para as previsões com e sem ondaletas e teste de significância para o câmbio.

Janela	Expoente de Hurst médio	Número de Exp. Hurst < 0,5	Número de Exp. Hurst > 0,5	CORREL_SW Exp. Hurst < 0,5 - média	CORREL_SW Exp. Hurst > 0,5 - média	T-teste	Valor p	CORREL_CW Exp. Hurst < 0,5 - média	CORREL_CW Exp. Hurst > 0,5 - média	T-teste	Valor p
60	0,586096	615	2088	0,384805	0,40486	-0,85517	0,384998	0,747453	0,768706	-1,58128	0,098471
80	0,581374	579	2104	0,40192	0,400378	0,064969	0,947939	0,753988	0,763385	-0,69376	0,482203
100	0,570192	599	2064	0,382114	0,406032	-1,02839	0,306521	0,738646	0,762613	-1,69013	0,076515
120	0,569131	511	2132	0,41172	0,396145	0,633964	0,5316	0,749261	0,758857	-0,63204	0,503127
150	0,558547	618	1995	0,416669	0,390854	1,136238	0,268735	0,761721	0,753335	0,643812	0,533231
180	0,558362	472	2111	0,432791	0,38903	1,773113	0,089467	0,760428	0,751249	0,619035	0,54235
200	0,555529	503	2060	0,440089	0,387184	2,155701	0,035408	0,758646	0,748474	0,685652	0,493367
250	0,561275	380	2133	0,433183	0,397391	1,297955	0,201606	0,756743	0,743576	0,799249	0,438042
300	0,562584	339	2124	0,465858	0,388181	2,723422	0,008328	0,750929	0,734121	0,951164	0,361537
400	0,561508	391	1972	0,374546	0,398393	-0,82719	0,394051	0,705661	0,732998	-1,46313	0,131298
500	0,56179	409	1854	0,358035	0,395787	-1,32965	0,174521	0,703697	0,723211	-1,08701	0,288584
600	0,558801	336	1827	0,36988	0,383301	-0,4376	0,656856	0,707991	0,718368	-0,51638	0,603878
700	0,556415	331	1732	0,367688	0,377902	-0,33138	0,738731	0,725757	0,716322	0,49903	0,638073

Tabela 21 Varredura das diferentes janelas de Hurst e estatísticas de acurácia do MAPE para as previsões com e sem ondaletas e teste de significância para o câmbio.

Janela	Expoente de Hurst médio	Número de Exp. Hurst < 0,5	Número de Exp. Hurst > 0,5	MAPE_SW Exp. Hurst < 0,5 - média	MAPE_SW Exp. Hurst > 0,5 - média	T-teste	Valor p	MAPE_CW Exp. Hurst < 0,5 - média	MAPE_CW Exp. Hurst > 0,5 - média	T-teste	Valor p
60	0,586096	615	2088	0,489285	0,483819	0,291547	0,7830	0,322002	0,307355	1,143864	0,2498
80	0,581374	579	2104	0,535992	0,472763	3,103536	0,0019	0,353854	0,300871	3,725114	0,0000
100	0,570192	599	2064	0,491092	0,48836	0,151133	0,8924	0,316874	0,314638	0,185277	0,8639
120	0,569131	511	2132	0,474673	0,496327	-1,19194	0,3134	0,307759	0,320072	-1,01431	0,3738
150	0,558547	618	1995	0,449939	0,510753	-3,69521	0,0025	0,288752	0,330689	-3,86679	0,0013
180	0,558362	472	2111	0,43321	0,516182	-5,35373	0,0002	0,27605	0,335628	-5,70216	0,0000
200	0,555529	503	2060	0,430459	0,521694	-5,68518	0,0000	0,281158	0,338891	-5,38254	0,0000
250	0,561275	380	2133	0,468316	0,516266	-2,48576	0,0515	0,307259	0,338369	-2,43911	0,0515
300	0,562584	339	2124	0,454368	0,526904	-3,99947	0,0052	0,308475	0,346778	-3,16416	0,0233
400	0,561508	391	1972	0,385048	0,535689	-8,58252	0,0000	0,252274	0,355253	-8,81965	0,0000
500	0,56179	409	1854	0,405345	0,53697	-6,71094	0,0000	0,264303	0,358309	-7,37773	0,0000
600	0,558801	336	1827	0,389735	0,546502	-8,46606	0,0000	0,254417	0,363642	-8,55347	0,0000
700	0,556415	331	1732	0,449183	0,509096	-3,14169	0,0127	0,286347	0,337806	-3,99783	0,0012

Tabela 22 Varredura das diferentes janelas de Hurst e estatísticas de acurácia do TIC para as previsões com e sem ondaletas e teste de significância para o câmbio

Janela	Total de Exponentes de Hurst	Número de Exp. Hurst < 0,5	Número de Exp. Hurst > 0,5	TIC_SW Exp. Hurst < 0,5 - média	TIC_SW Exp. Hurst > 0,5 - média	T-teste	Valor p	TIC_CW Exp. Hurst < 0,5 - média	TIC_CW Exp. Hurst > 0,5 - média	T-teste	Valor p
60	0,586096	615	2088	0,002832	0,002795	0,340547	0,7470	0,001857	0,001773	1,1581	0,2406
80	0,581374	579	2104	0,003102	0,002732	3,16267	0,0016	0,002052	0,001733	3,963396	0,0000
100	0,570192	599	2064	0,002845	0,002822	0,219851	0,8443	0,001841	0,00181	0,441759	0,6799
120	0,569131	511	2132	0,002741	0,00287	-1,23887	0,2985	0,001772	0,001849	-1,13289	0,3248
150	0,558547	618	1995	0,002618	0,002948	-3,47023	0,0045	0,001679	0,001911	-3,77527	0,0015
180	0,558362	472	2111	0,002509	0,002983	-5,27325	0,0002	0,001602	0,001943	-5,75946	0,0000
200	0,555529	503	2060	0,002483	0,003017	-5,74794	0,0000	0,001626	0,001966	-5,60729	0,0000
250	0,561275	380	2133	0,002718	0,002981	-2,32008	0,0642	0,001781	0,001966	-2,52967	0,0405
300	0,562584	339	2124	0,002633	0,003046	-3,91691	0,0057	0,00178	0,002017	-3,49587	0,0130
400	0,561508	391	1972	0,002229	0,003098	-8,50998	0,0000	0,001483	0,002064	-8,56132	0,0000
500	0,56179	409	1854	0,002351	0,003108	-6,57411	0,0000	0,001536	0,002096	-7,66726	0,0000
600	0,558801	336	1827	0,002264	0,003163	-8,30811	0,0000	0,001485	0,002128	-8,88579	0,0000
700	0,556415	331	1732	0,002629	0,002947	-2,78587	0,0232	0,00168	0,001979	-4,05566	0,0016

Tabela 23 Varredura das diferentes janelas de Hurst e estatísticas de acurácia do RMSE para as previsões com e sem ondaletas e teste de significância para o câmbio.

Janela	Total de Exponentes de Hurst	Número de Exp. Hurst < 0,5	Número de Exp. Hurst > 0,5	RMSE_SW Exp. Hurst < 0,5 - média	RMSE_SW Exp. Hurst > 0,5 - média	T-teste	Valor p	RMSE_CW Exp. Hurst < 0,5 - média	RMSE_CW Exp. Hurst > 0,5 - média	T-teste	Valor p
60	0,586096	615	2088	0,012764	0,013482	-1,20234	0,2450	0,008379	0,008559	-0,45569	0,6452
80	0,581374	579	2104	0,015278	0,012851	3,490655	0,0001	0,010103	0,008142	4,128759	0,0000
100	0,570192	599	2064	0,014063	0,013295	1,250004	0,2218	0,009037	0,008528	1,275744	0,2020
120	0,569131	511	2132	0,013692	0,013538	0,236919	0,8184	0,008786	0,008719	0,159165	0,8755
150	0,558547	618	1995	0,012831	0,013983	-1,97727	0,0672	0,008138	0,00907	-2,52879	0,0198
180	0,558362	472	2111	0,01113	0,014461	-6,92325	0,0000	0,007077	0,009377	-7,48735	0,0000
200	0,555529	503	2060	0,011074	0,014636	-7,34724	0,0000	0,007237	0,009478	-7,13518	0,0000
250	0,561275	380	2133	0,012516	0,0144	-3,34746	0,01437	0,00819	0,009401	-3,3517	0,0127
300	0,562584	339	2124	0,011853	0,014773	-5,59473	0,0003	0,007999	0,009652	-4,95273	0,0013
400	0,561508	391	1972	0,009618	0,01517	-11,2078	0,0000	0,006388	0,009903	-10,8945	0,0000
500	0,56179	409	1854	0,009632	0,015348	-10,6119	0,0000	0,006274	0,010060	-11,2448	0,0000
600	0,558801	336	1827	0,008747	0,015675	-14,3307	0,0000	0,005745	0,010211	-14,029	0,0000
700	0,556415	331	1732	0,009623	0,014311	-9,80774	0,0000	0,006154	0,009263	-10,2541	0,0000

Analisando-se os dados para a série de câmbio em um período de janeiro de 1998 à dezembro de 2010, pode-se verificar que quando o expoente de Hurst é menor que 0,5, a qualidade das previsões tem uma melhora para os indicadores MAPE, TIC e RMSE, exceto para a janela de 80 pontos onde a qualidade da previsão é melhor para a série persistente. A correlação não demonstrou significância em nenhuma das janelas testadas.

As janelas de varredura da série de 150 a 200 e de 300 a 700 foram significativas rejeitando-se a hipótese nula, para o MAPE e o TIC e as janelas de 80, 150 a 700 pontos foram significativas para o RMSE. Outra observação importante é que a qualidade das previsões melhoraram significativamente com o uso das ondaletas.

4.2.3 Análise do Ibovespa

O índice da Bolsa de Valores de São Paulo (BM&FBOVESPA), chamado também de IBOVESPA, retrata o comportamento dos principais papéis negociados na BOVESPA. Com isso, é o mais importante indicador do desempenho médio das cotações do mercado de ações brasileiro. (BM&FBOVESPA)

Esse índice foi constituído em 02/01/1968, a partir do valor atual, em moeda corrente, de uma carteira teórica de ações, a partir de uma aplicação hipotética. Supõe-se não ter sido efetuado nenhum investimento adicional desde então, considerando-se somente os ajustes efetuados em decorrência da distribuição de proventos pelas empresas emissoras. Assim, o índice Ibovespa reflete além das variações dos preços das ações, o impacto da distribuição dos proventos, sendo considerado um indicador que avalia o retorno total de suas ações componentes. (BM&FBOVESPA)

A finalidade básica do Ibovespa é a de ser um indicador médio do comportamento do mercado. Para isso em termos de liquidez, as ações integrantes da carteira teórica do índice Bovespa respondem por mais de 80% do número de negócios e do volume financeiro verificados no lote padrão do mercado à vista da BOVESPA. Em termos de capitalização bursátil, as empresas emissoras das ações integrantes da carteira teórica do Índice Bovespa são responsáveis, em média, por aproximadamente 70% do somatório da capitalização bursátil de todas as empresas com ações negociáveis na BOVESPA. (BM&FBOVESPA)

Para este trabalho a série do IBOVESPA foi tomada no período de 02/01/1998 quando representava 10479 pontos até 30/12/2010, quando apresentou 69304 pontos.

Nesse período, a valorização foi de 561,36%. As cotações foram consideradas de fechamento no período diário num total de 3216 observações.

Para a série do Ibovespa no período analisado tem-se 3216 cotações de preços e 3215 log retornos. Para a exemplificação a série será varrida tomando 500 pontos de cada vez e considerando 496 para a estimação da rede e os 4 últimos para a previsão. Para esses dados ter-se-á feitas 2716 iterações, ou seja, serão calculados 2716 expoentes de Hurst e serão obtidas a mesma quantidade de métricas de acurácia das previsões para cada bloco de 500 pontos

A seguir, mostra-se o gráfico da série do Índice Ibovespa e os respectivos expoentes de Hurst ao longo da série para uma janela de 500 pontos.

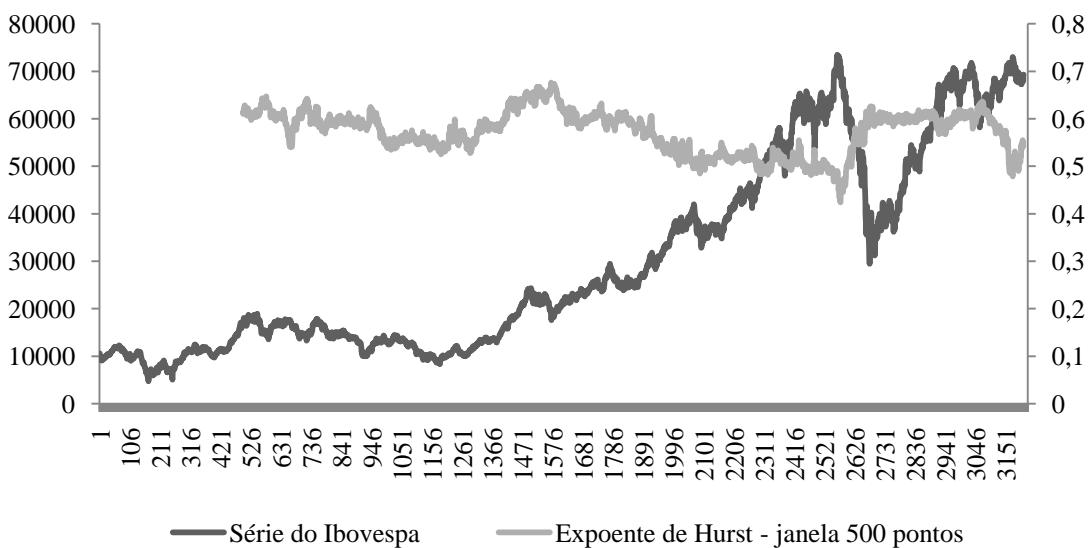


Figura 27: Série do Ibovespa e os expoentes de Hurst ao longo da série para uma janela de 500 pontos.

Nota-se na figura 28 que a série do índice Ibovespa apresenta poucos valores do expoente de Hurst abaixo de 0,5. Na sequência da análise mostra-se os medidores de qualidade com e sem filtro de ondaletas utilizando o modelo de Redes Neurais Recorrentes, em função do expoente de Hurst.

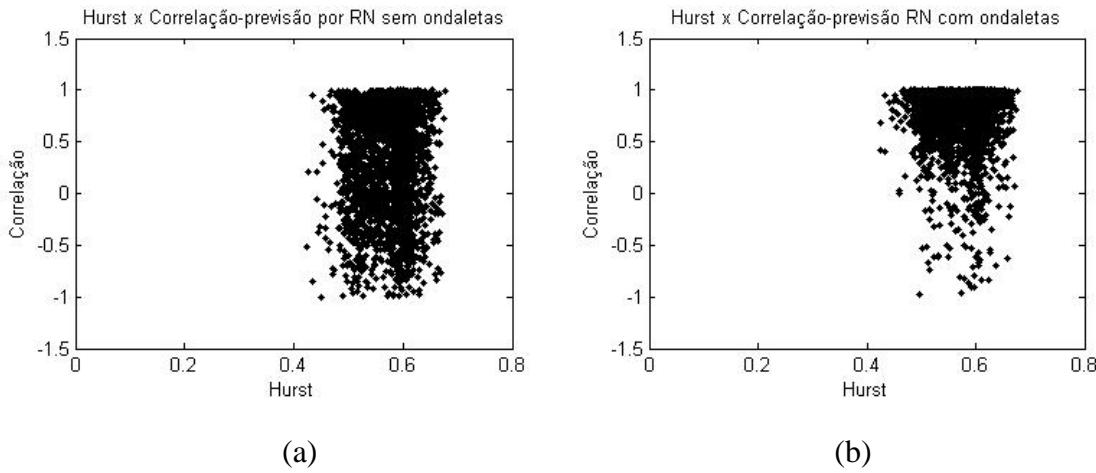


Figura 28: Correlação em função do Exponente de Hurst calculados para uma janela de 500 pontos com o modelo de Redes Neurais Recorrentes com e sem filtro de ondaletas para o Índice Ibovespa.

Na inspeção visual dos pontos destacados na figura acima, nota-se claramente que a correlação aumenta quando se aplica o filtro de ondaletas nas janelas de Hurst como demonstrado na figura 28(b) por estarem mais próximos de um, ao passo que na figura 28(a), os valores das correlações encontram-se mais espalhados.

Em relação ao indicador MAPE, a concentração do erro cai na figura 29(b) com mais pontos próximos ao eixo do que em relação à figura 29(a) como pode ser visto na figura a seguir.

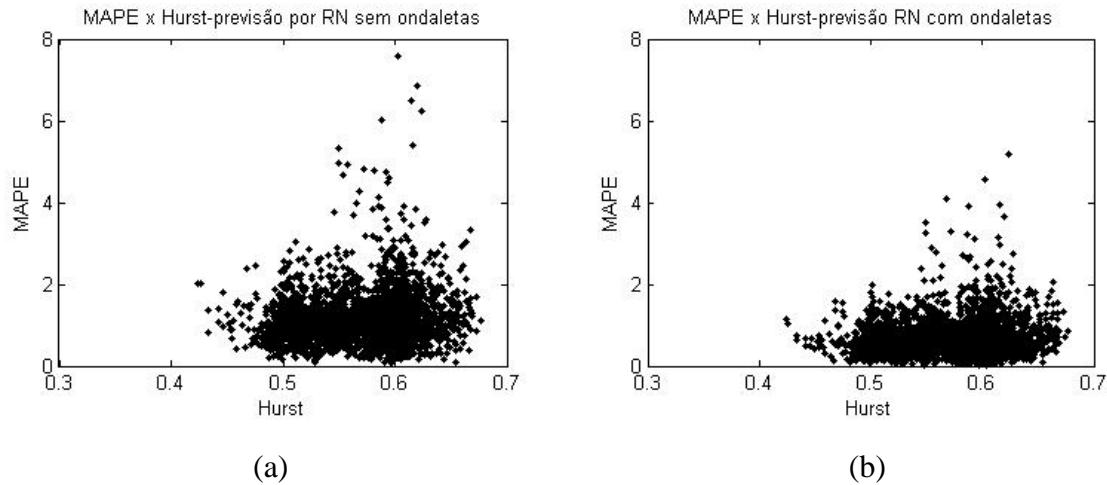


Figura 29: MAPE em função do Exponente de Hurst calculados para uma janela de 500 pontos com o modelo de Redes Neurais Recorrentes com e sem filtro de ondaletas para o Índice Ibovespa.

Em relação ao TIC, a análise é muito semelhante, com redução do erro apresentando maior concentração pelo uso do pré-processamento por ondaletas como pode ser visto na figura a seguir.

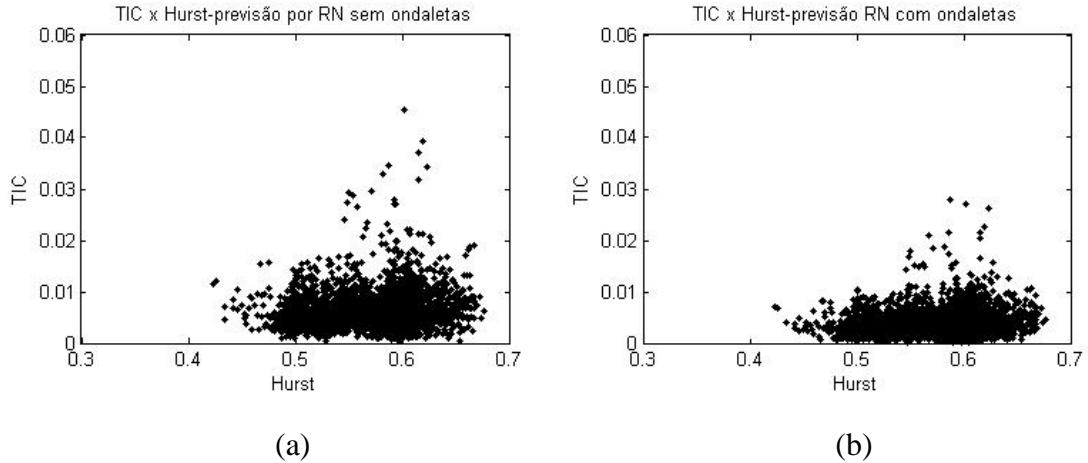


Figura 30: TIC em função do Exponente de Hurst calculados para uma janela de 500 pontos com o modelo de Redes Neurais Recorrentes com e sem filtro de ondaletas para o Índice Ibovespa.

E para o RMSE também apresenta redução do erro a partir da filtragem em alta e baixa freqüência e com as previsões feitas nas sub-séries como pode ser visto na figura a seguir.

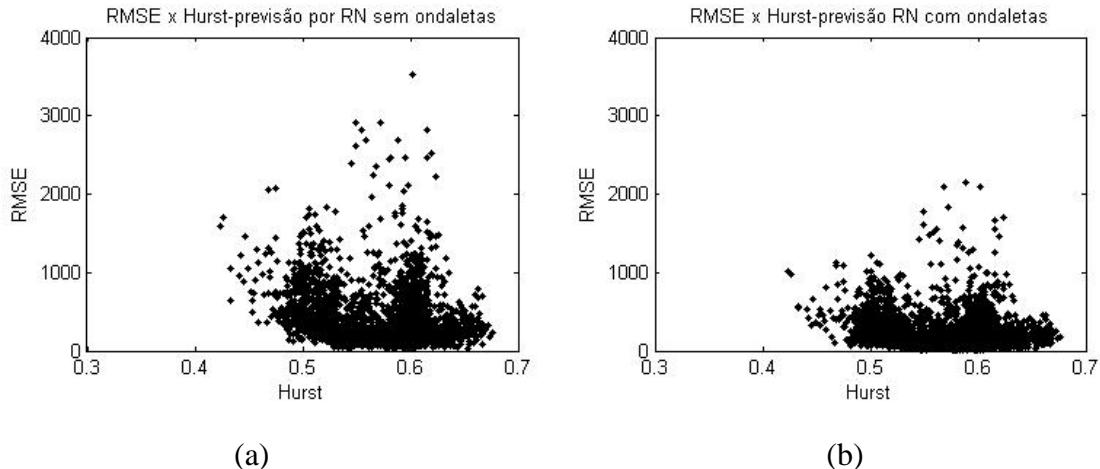


Figura 31: RMSE em função do Exponente de Hurst calculados para uma janela de 500 pontos com o modelo de Redes Neurais Recorrentes com e sem filtro de ondaletas para o Índice Ibovespa.

Posteriormente, separou-se os valores do expoente de Hurst onde os valores eram maiores que meio e menores que meio. Calculou-se a média dos indicadores de

acurácia das previsões em cada conjunto de expoentes e obtiveram-se os seguintes resultados:

Tabela 24 Estatísticas de acurácia das previsões para expoentes de Hurst maiores e menores do que 0,5 para o Ibovespa.

Hurst < 0,5 – 203 casos					
	CORREL_SW	TIC_SW	MAPE_SW	RMSE_SW	Expoente de Hurst
Média	0,34228	0,005679	0,969699	670,0041	0,486723
DP	0,519956	0,002754	0,451436	342,0014	0,015026

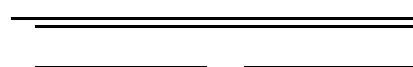
Hurst < 0,5 – 203 casos					
	CORREL_CW	TIC_CW	MAPE_CW	RMSE_CW	Expoente de Hurst
Média	0,803509	0,00316	0,553085	371,6397	0,486723
DP	0,235266	0,001829	0,322363	221,6631	0,015026

Hurst > 0,5 – 2513 casos					
	CORREL_SW	TIC_SW	MAPE_SW	RMSE_SW	Expoente de Hurst
Média	0,509057	0,003958	0,685104	359,0488	0,57817
DP	0,330585	0,00661	1,141498	399,9888	0,039553

Hurst > 0,5 – 2513 casos					
	CORREL_CW	TIC_CW	MAPE_CW	RMSE_CW	Expoente de Hurst
Média	0,743114	0,004052	0,699925	243,8868	0,57817
DP	0,306153	0,002699	0,469625	227,6795	0,039553

De posse das estatísticas acima, aplica-se um teste paramétrico de diferença de médias com as seguintes hipóteses assumindo variâncias diferentes:

Para a correlação, a estatística de teste sem o uso de ondaletas é:



O t teste para a estatística acima foi de 0,3087. Os demais valores para as outras hipóteses são demonstradas a seguir:

Tabela 25 Teste de hipóteses de diferenças de médias das estatísticas de acurácia com e sem ondaletas para os expoentes de Hurst maiores e menores que 0,5 para o índice Ibovespa

	CORREL_SW	TIC_SW	MAPE_SW	RMSE_SW
t _{teste}	0,308742	-4,97874	-4,45961	10,77923
valor p	0,753274	0,0005	0,0010	0,0000
Resultado	Aceita H ₀	Rejeita H ₀	Rejeita H ₀	Rejeita H ₀

	CORREL_CW	TIC_CW	MAPE_CW	RMSE_CW
t _{teste}	3,430444	-5,99637	-6,40712	7,882527
valor p	0,006077	0,0000	0,0000	0,0000
Resultado	Aceita H ₀	Rejeita H ₀	Rejeita H ₀	Rejeita H ₀

De acordo com resultados acima, para as previsões feitas com a correlação, nota-se que teste t aceita a hipótese nula, apresentando uma igualdade entre os parâmetros.

Todavia, quando realiza-se o cálculo das previsões e aplica-se os medidores de qualidade TIC, MAPE e RMSE, os indicadores rejeitaram a hipótese nula de igualdade dos indicadores de previsão quando o expoente de Hurst é menor e quando é maior que 0,5. Assim, há indícios significativos de que o expoente de Hurst pode demonstrar influência na janela de resultados usada nas estimativas de previsão de preços futuros de séries temporais financeiras.

Agora, como nas análises anteriores, tomam-se os mesmos resultados para diferentes configurações de janelas, ou seja, mudam-se os tamanhos das janelas de estimativa do expoente de Hurst mas mantêm-se as previsões de quatro passos a frente para as distintas janelas. Os resultados obtidos são demonstrados nas tabelas a seguir.

E assim, será mudada a janela de Hurst e obtidas as previsões e as estatísticas de erro para cada janela e seu respectivo teste de hipótese de diferença de médias para cada métrica visando verificar se a mudança do expoente de Hurst acima e abaixo de 0,5 influenciam nos resultados das previsões.

Tabela 26 Varredura das diferentes janelas de Hurst e estatísticas de acurácia da correlação para as previsões com e sem ondaletas e teste de significância para o Ibovespa

Janela	Expoente de Hurst médio	Número de Exp. Hurst < 0,5	Número de Exp. Hurst > 0,5	CORREL_SW Exp. Hurst < 0,5 - média	CORREL_SW Exp. Hurst > 0,5 - média	T-teste	Valor p	CORREL_CW Exp. Hurst < 0,5 - média	CORREL_CW Exp. Hurst > 0,5 - média	T-teste	Valor p
60	0,557427	980	2176	0,33507	0,344047	-0,45432	0,649742	0,77156	0,751036	1,964877	0,05978
80	0,569733	785	2351	0,318234	0,349234	-1,44339	0,1431	0,756135	0,754984	0,099579	0,922185
100	0,576074	652	2464	0,332772	0,344068	-0,50387	0,617707	0,744389	0,755796	-0,88859	0,373826
120	0,573299	496	2600	0,322376	0,344226	-0,86566	0,38573	0,75281	0,755606	-0,19971	0,84324
150	0,5789	371	2695	0,298294	0,341462	-1,50931	0,130231	0,761247	0,750648	0,709683	0,51307
180	0,572709	322	2714	0,293735	0,339349	-1,45798	0,132076	0,732597	0,752125	-1,03428	0,261028
200	0,573971	354	2662	0,332522	0,334924	-0,08089	0,934089	0,763285	0,750344	0,807716	0,435442
250	0,575175	239	2727	0,312744	0,330311	-0,48769	0,611194	0,765631	0,748444	0,911289	0,388507
300	0,575098	109	2807	0,349776	0,329617	0,398706	0,68657	0,780661	0,745768	1,418217	0,236725
400	0,569602	173	2643	0,242719	0,336719	-2,0989	0,019102	0,769486	0,747572	0,991778	0,35225
500	0,571335	203	2513	0,34228	0,330585	0,308742	0,753274	0,803509	0,743114	3,430444	0,006077
600	0,57373	200	2416	0,323448	0,328785	-0,13603	0,887157	0,806195	0,74078	4,010131	0,003228
700	0,569878	348	2168	0,304219	0,335686	-1,00982	0,286651	0,778448	0,739173	2,529186	0,026243

Tabela 27 Varredura das diferentes janelas de Hurst e estatísticas de acurácia do MAPE para as previsões com e sem ondaletas e teste de significância para o Ibovespa

Janela	Expoente de Hurst médio	Número de Exp. Hurst < 0,5	Número de Exp. Hurst > 0,5	MAPE_SW Exp. Hurst < 0,5 - média	MAPE_SW Exp. Hurst > 0,5 - média	T-teste	Valor p	MAPE_CW Exp. Hurst < 0,5 - média	MAPE_CW Exp. Hurst > 0,5 - média	T-teste	Valor p
60	0,557427	980	2176	1,136849	1,197232	-1,93925	0,0607	0,68538	0,751897	-3,2418	0,0019
80	0,569733	785	2351	1,105286	1,2092	-3,08254	0,0026	0,677444	0,753795	-3,4916	0,0009
100	0,576074	652	2464	1,081428	1,210366	-3,50547	0,0004	0,672089	0,751215	-3,2551	0,0014
120	0,573299	496	2600	1,049178	1,204639	-4,15758	0,0001	0,641455	0,748186	-4,3037	0,0001
150	0,5789	371	2695	0,971515	1,21446	-7,83989	0,0000	0,574699	0,757306	-8,8443	0,0000
180	0,572709	322	2714	0,857481	1,201641	-11,9338	0,0000	0,512772	0,747152	-12,272	0,0000
200	0,573971	354	2662	0,945914	1,181728	-7,99834	0,0000	0,572156	0,733788	-8,1235	0,0000
250	0,575175	239	2727	0,938206	1,157096	-6,11483	0,0000	0,55672	0,71463	-6,6416	0,0000
300	0,575098	109	2807	0,995062	1,124243	-2,6417	0,0442	0,606109	0,692022	-2,2309	0,0527
400	0,569602	173	2643	1,046069	1,12652	-1,93092	0,1234	0,593514	0,693002	-3,5542	0,0057
500	0,571335	203	2513	0,969699	1,141498	-4,97874	0,0005	0,553085	0,699925	-5,9964	0,0000
600	0,57373	200	2416	1,074232	1,127673	-1,31617	0,2804	0,612325	0,69426	-3,1046	0,0168
700	0,569878	348	2168	1,105167	1,128036	-0,69865	0,5628	0,642889	0,696508	-2,5136	0,0475

Tabela 28 Varredura das diferentes janelas de Hurst e estatísticas de acurácia do TIC para as previsões com e sem ondaletas e teste de significância para o Ibovespa

Janela	Total de Exponentes de Hurst	Número de Exp. Hurst < 0,5	Número de Exp. Hurst > 0,5	TIC_SW Exp. Hurst < 0,5 - média	TIC_SW Exp. Hurst > 0,5 - média	T-teste	Valor p	TIC_CW Exp. Hurst < 0,5 - média	TIC_CW Exp. Hurst > 0,5 - média	T-teste	Valor p
60	0,557427	980	2176	0,00659	0,006942	-1,9458	0,0590	0,003956	0,004341	-3,2481	0,0016
80	0,569733	785	2351	0,006418	0,007005	-2,98559	0,0034	0,003905	0,004351	-3,5189	0,0007
100	0,576074	652	2464	0,006282	0,007017	-3,42244	0,0006	0,003893	0,004328	-3,0714	0,0026
120	0,573299	496	2600	0,006129	0,006981	-3,85477	0,0003	0,003705	0,004311	-4,1607	0,0001
150	0,5789	371	2695	0,00568	0,007047	-7,41702	0,0000	0,003369	0,004363	-8,2369	0,0000
180	0,572709	322	2714	0,004975	0,006975	-11,9254	0,0000	0,002965	0,004311	-12,249	0,0000
200	0,573971	354	2662	0,00552	0,006853	-7,63294	0,0000	0,00331	0,004235	-8,0208	0,0000
250	0,575175	239	2727	0,005482	0,00671	-5,8297	0,0000	0,003199	0,004133	-6,9432	0,0000
300	0,575098	109	2807	0,005843	0,006518	-2,245	0,0701	0,003449	0,004003	-2,6287	0,0296
400	0,569602	173	2643	0,006096	0,006529	-1,77361	0,1530	0,003384	0,004009	-4,0137	0,0026
500	0,571335	203	2513	0,005679	0,00661	-4,45961	0,0010	0,00316	0,004052	-6,4071	0,0000
600	0,57373	200	2416	0,006261	0,006533	-1,12556	0,3427	0,003487	0,004018	-3,5628	0,0069
700	0,569878	348	2168	0,006446	0,006537	-0,47652	0,6897	0,003685	0,004034	-2,8387	0,0249

Tabela 29 Varredura das diferentes janelas de Hurst e estatísticas de acurácia do RMSE para as previsões com e sem ondaletas e teste de significância para o Ibovespa

Janela	Total de Expoentes de Hurst	Número de Exp. Hurst < 0,5	Número de Exp. Hurst > 0,5	RMSE_SW Exp. Hurst < 0,5 - média	RMSE_SW Exp. Hurst > 0,5 - média	T-teste	Valor p	RMSE_CW Exp. Hurst < 0,5 - média	RMSE_CW Exp. Hurst > 0,5 - média	T-teste	Valor p
60	0,557427	980	2176	402,9579	368,9959	2,259517	0,0112	242,6829	229,1895	1,424477	0,1113
80	0,569733	785	2351	407,4224	373,5265	1,98014	0,0185	248,0323	229,7945	1,695318	0,0447
100	0,576074	652	2464	429,3558	371,913	3,190394	0,0002	270,9882	225,7168	3,880535	0,0000
120	0,573299	496	2600	428,997	377,09	2,758035	0,0025	262,4062	229,8843	2,674217	0,0027
150	0,5789	371	2695	408,4939	386,2001	1,265193	0,2526	241,9126	236,654	0,479899	0,6689
180	0,572709	322	2714	401,8353	389,4415	0,796762	0,5520	238,7437	237,7779	0,100121	0,9414
200	0,573971	354	2662	475,0154	380,9605	5,65316	0,0000	285,0785	232,0015	4,899922	0,0000
250	0,575175	239	2727	664,2596	372,7088	11,42296	0,0000	387,8298	226,8004	10,1117	0,0000
300	0,575098	109	2807	724,3726	387,5019	8,971048	0,0000	429,5211	234,6014	7,45237	0,0000
400	0,569602	173	2643	687,8634	391,6691	8,981346	0,0000	386,7904	237,7548	7,428192	0,0000
500	0,571335	203	2513	670,0041	399,9888	10,77923	0,0000	371,6397	243,8868	7,882527	0,0000
600	0,57373	200	2416	744,0188	400,6763	11,29443	0,0000	414,5241	244,7276	9,2191	0,0000
700	0,569878	348	2168	684,3135	395,3799	13,12564	0,0000	387,4293	242,3814	10,54984	0,0000

Analisando-se os cálculos feitos para a série do Ibovespa, no período de janeiro de 1998 à dezembro de 2010, observa-se que, a correlação não demonstra significância em nenhuma das janelas testadas. Para as janelas de varredura da série de 80 à 250 e 500 pontos foram significantes para os medidores MAPE e TIC com e sem o uso de ondaletas. Para as janelas de 60, 400 e 600 pontos houve uma significância apenas com o processamento de ondaletas para os dois medidores e em particular uma significância com ondaletas para o TIC na janela de 700 pontos. Isso significa que para essas janelas, rejeitou-se a hipótese nula e observou-se também que, como para as séries de café e câmbio, para os expoentes de Hurst menor que 0,5 há um menor erro de previsão. Já para o RMSE houve uma rejeição da hipótese nula para as janelas de 60 e 80 pontos apenas sem o processamento de ondaletas e para janelas de 100 a 700 pontos para o caso sem o uso de ondaletas e com o uso de ondaletas, no entanto observou-se um menor erro de previsão para a série persistente.

4.2.4 Análise Da Vale do Rio Doce (VALE5)

A Vale é uma empresa privada de capital aberto brasileira, uma das maiores mineradoras do mundo. Criada em 1942, no governo Getúlio Vargas, é hoje uma grande empresa privada, de capital aberto, com sede na cidade do Rio de Janeiro, e ações negociadas na Bolsa de Valores de São Paulo (BM&FBOVESPA), na bolsa de valores de Paris (NYSE Euronext), na bolsa de valores de Madrid (Latibex), na bolda de valores de Hong Kong (HKEx) e na Bolsa de Valores de Nova York (NYSE), integrando o Dow Jones Sector Titans Composite Index.

A Vale tornou-se, hoje, a maior empresa de mineração diversificada das Américas e a segunda maior do mundo. É a maior produtora de minério de ferro do mundo e a segunda maior de níquel. A Vale destaca-se ainda na produção de manganês, cobre, carvão, cobalto, pelotas, ferroligas e alguns fertilizantes, como os fosfatados (TSP e DCP) e nitrogenados (ureia e amônia).

Em número de negociações, as ações da Vale são as mais negociada na Bolsa de Valores de São Paulo. Para este trabalho foi usada a série das ações preferenciais da Vale do Rio Doce (VALE5) no período de 02/01/1998 quando representava R\$ 0,99 até 30/12/2010, quando apresentou R\$ 47,57. Nesse período, a valorização foi de

4705,05%. As cotações foram consideradas de fechamento no período diário num total de 3216 observações.

Para exemplificar a rotina de apuração da junção do estudo da janela com verificação do expoente de Hurst e a qualidade das previsões, far-se-á inicialmente para uma janela de 500 pontos.

Assim, a série temporal será varrida tomando sempre 500 pontos de cada vez e considerando 496 para estimativa da rede e os 4 últimos para previsão. Como a série da soja no período analisado tem 3216 cotações de preços e 3215 log retornos, para a janela de Hurst de 500 pontos, serão feitas 2716 iterações, ou seja, serão calculados 2716 expoentes de Hurst e serão obtidas a mesma quantidade de métricas de acurácia das previsões para cada bloco de 500 pontos.

A seguir, mostra-se o gráfico da série dos preços da soja e os respectivos expoentes de Hurst ao longo da série para uma janela de 500 pontos.

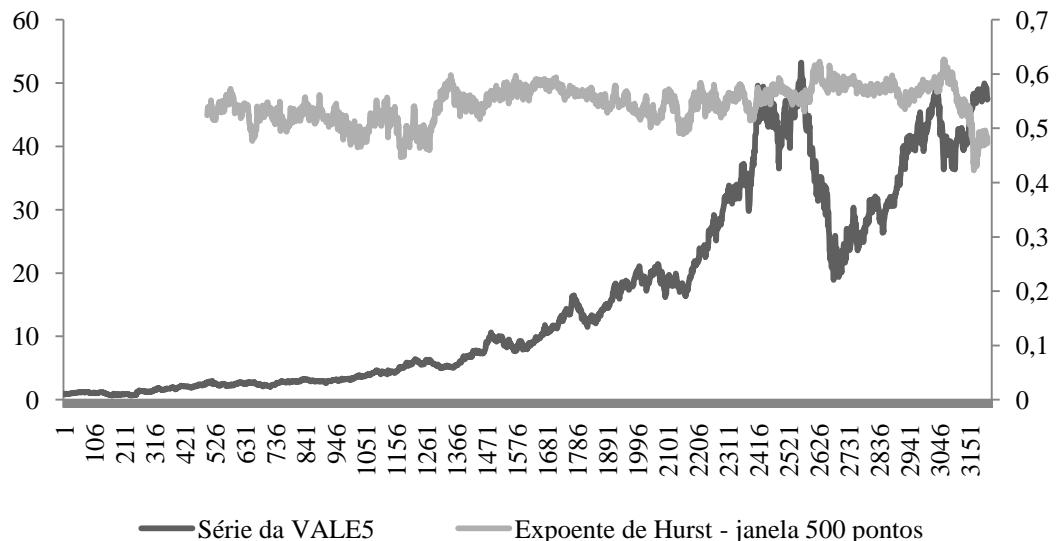


Figura 32: Série da VALE5 e os expoentes de Hurst ao longo da série para uma janela de 500 pontos.

Nota-se no gráfico que a série da soja apresenta poucos valores do expoente de Hurst abaixo de 0,5. Na sequência da análise mostra-se os medidores de qualidade com e sem filtro de ondaletas utilizando o modelo de Redes Neurais Recorrentes, em função do expoente de Hurst.

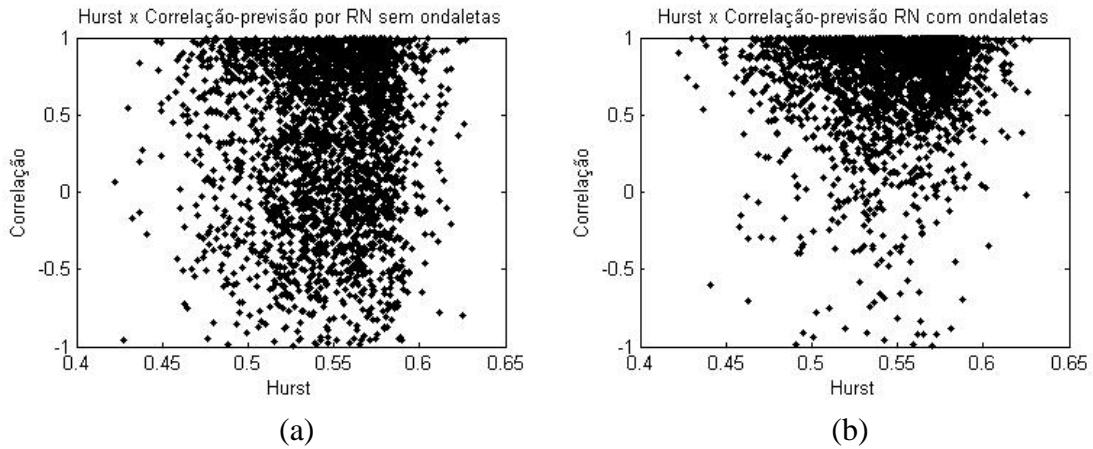


Figura 33: Correlação em função do Exponente de Hurst calculados para uma janela de 500 pontos com o modelo de Redes Neurais Recorrentes com e sem filtro de ondaletas para a VALE5.

Na inspeção visual dos pontos destacados na figura acima, nota-se claramente que a correlação aumenta quando se aplica o filtro de ondaletas nas janelas de Hurst como demonstrado na figura 33(b) por estarem mais próximos de um, ao passo que na figura 33(a), os valores das correlações encontram-se mais espalhados.

Em relação ao indicador MAPE, a concentração do erro cai na figura 34(b) com mais pontos próximos ao eixo do que em relação à figura 34(a) como pode ser visto na figura a seguir.

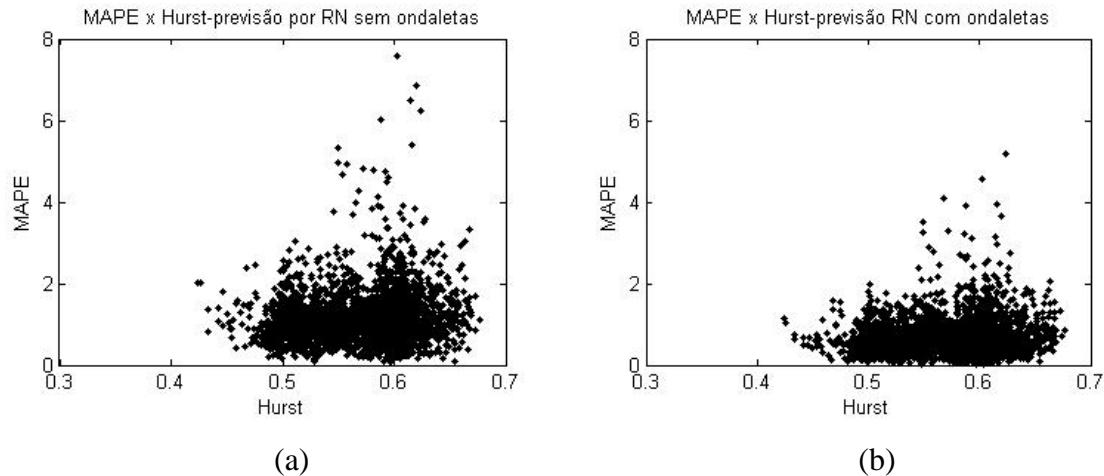


Figura 34: MAPE em função do Exponente de Hurst calculados para uma janela de 500 pontos com o modelo de Redes Neurais Recorrentes com e sem filtro de ondaletas para a VALE5.

Em relação ao TIC, a análise é muito semelhante, com redução do erro apresentando maior concentração pelo uso do pré-processamento por ondaletas como pode ser visto na figura a seguir.

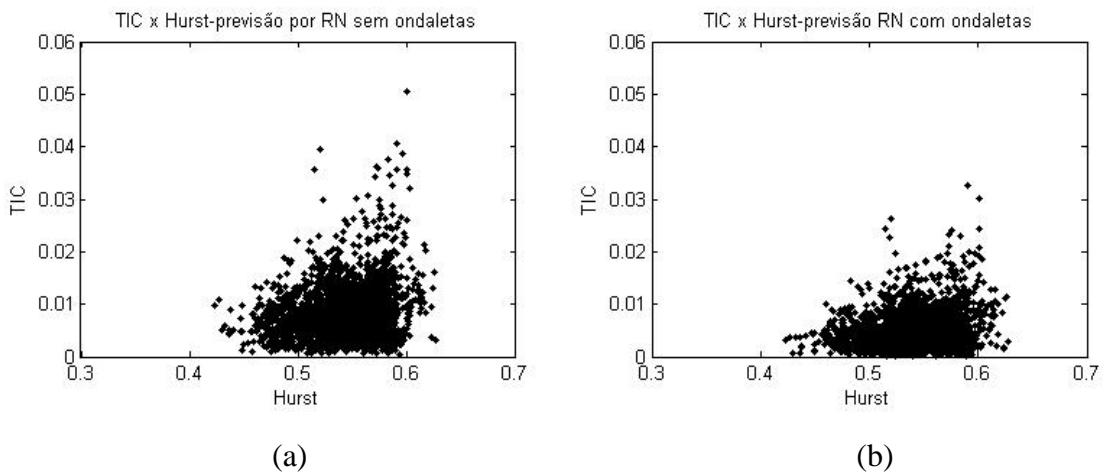


Figura 35: TIC em função do Exponente de Hurst calculados para uma janela de 500 pontos com o modelo de Redes Neurais Recorrentes com e sem filtro de ondaletas para a VALE5.

E para o RMSE também apresenta redução do erro a partir da filtragem em alta e baixa freqüência e com as previsões feitas nas sub-séries como pode ser visto na figura a seguir.

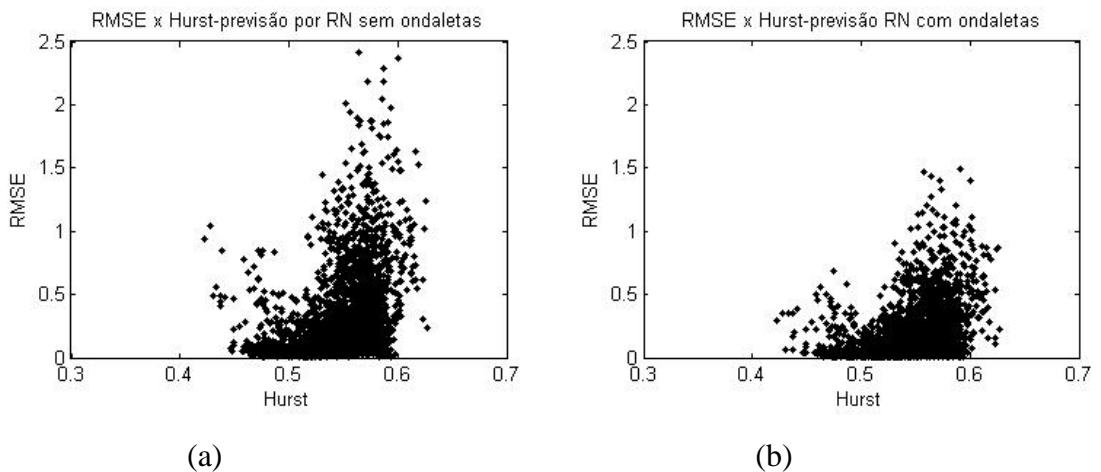


Figura 36: RMSE em função do Exponente de Hurst calculados para uma janela de 500 pontos com o modelo de Redes Neurais Recorrentes com e sem filtro de ondaletas para a VALE5.

Posteriormente, separou-se os valores do expoente de Hurst onde os valores eram maiores que meio e menores que meio. Calculou-se a média dos indicadores de acurácia das previsões em cada conjunto de expoentes e obte-se os seguintes resultados:

Tabela 30 Estatísticas de acurácia das previsões para expoentes de Hurst maiores e menores do que 0,5 para a VALE5.

Hurst < 0,5 – 274 casos					
	CORREL_SW	TIC_SW	MAPE_SW	RMSE_SW	Expoente de Hurst
Média	0,365129	0,006644	1,137288	0,1482305	0,481306
DP	0,534494	0,003623	0,630197	0,1899008	0,014896

Hurst < 0,5 – 274 casos					
	CORREL_CW	TIC_CW	MAPE_CW	RMSE_CW	Expoente de Hurst
Média	0,710558	0,004272	0,742905	0,0945508	0,481306
DP	0,356203	0,002472	0,437659	0,1198473	0,014896

Hurst > 0,5 – 2442 casos					
	CORREL_SW	TIC_SW	MAPE_SW	RMSE_SW	Expoente de Hurst
Média	0,343288	0,008212	1,408497	0,3214464	0,553105
DP	0,518314	0,00531	0,913861	0,3557295	0,024764

Hurst > 0,5 – 2442 casos					
	CORREL_CW	TIC_CW	MAPE_CW	RMSE_CW	Expoente de Hurst
Média	0,742269	0,005083	0,867996	0,1925919	0,553105
DP	0,306294	0,003547	0,601326	0,2166927	0,024764

Tabela 31 Teste de hipóteses de diferenças de médias das estatísticas de acurácia com e sem ondaletas para os expoentes de Hurst maiores e menores que 0,5 para as ações da VALE5.

	CORREL_SW	TIC_SW	MAPE_SW	RMSE_SW
t _{teste}	-0,6593	0,006644	-6,40772	-12,7893
valor p	0,509753	0,008212	0,000002	0,00000
Resultado	Aceita	Rejeita	Rejeita	Rejeita

	CORREL_CW	TIC_CW	MAPE_CW	RMSE_CW
t _{teste}	1,5969	-4,92129	-4,29784	-11,5825
valor p	0,110406	0,000212	0,000833	0,00000
Resultado	Aceita	Rejeita	Rejeita	Rejeita

De acordo com resultados acima, para as previsões feitas sem pré-processamento de ondaletas em altas e baixas freqüências, nota-se que o expoente de Hurst não mostra indícios de interferir nos resultados das previsões para o medidor da correlação, os testes de hipóteses com 95% de confiança não são conclusivos a cerca da influência do expoente de Hurst na capacidade preditiva da série da vale 5. Já para os medidores: TIC, MAPE e RMSE eles apresentam que o expoente de Hurst rejeita a hipótese nula mostrando uma diferenciação entre as duas amostras o que é significativo.

Todavia, quando realiza-se a separação da série em altas e baixas freqüências, e realiza-se as previsões dentro das sub-séries, nota-se que há uma leve melhora nos indicadores TIC e RMSE e uma leve piora no MAPE.

E assim como nos exemplos anteriores tomam-se os mesmos resultados para diferentes configurações de janelas como é mostrado nas tabelas a seguir.

Tabela 32 Varredura das diferentes janelas de Hurst e estatísticas de acurácia da correlação para as previsões com e sem ondaletas e teste de significância para a VALE5.

Janela	Expoente de Hurst médio	Número de Exp. Hurst < 0,5	Número de Exp. Hurst > 0,5	CORREL_SW Exp. Hurst < 0,5 - média	CORREL_SW Exp. Hurst > 0,5 - média	T-teste	Valor p	CORREL_CW Exp. Hurst < 0,5 - média	CORREL_CW Exp. Hurst > 0,5 - média	T-teste	Valor p
60	0,556333	1024	2132	0,315675	0,3635	2,4279	0,01524	0,725777	0,7543	2,4901	0,012822
80	0,550412	1055	2081	0,337917	0,3505	0,6425	0,520552	0,739669	0,7463	0,5785	0,562972
100	0,553932	991	2125	0,348042	0,3461	-0,0961	0,92347	0,748119	0,7406	-0,6434	0,520017
120	0,549921	916	2180	0,359378	0,3413	-0,8839	0,376812	0,749452	0,7389	-0,8718	0,383388
150	0,547497	799	2267	0,376632	0,3295	-2,2052	0,027515	0,755636	0,7336	-1,7130	0,08682
180	0,549157	663	2373	0,391847	0,3281	-2,7966	0,005197	0,751746	0,7324	-1,3757	0,16901
200	0,549671	605	2411	0,398954	0,3299	-2,9356	0,003355	0,756036	0,7330	-1,5939	0,111069
250	0,550006	541	2425	0,331869	0,3319	-2,4646	0,013772	0,735099	0,7351	-1,0360	0,300308
300	0,547671	362	2554	0,432906	0,3305	-3,5215	0,000436	0,762642	0,7354	-1,5401	0,123634
400	0,544979	314	2502	0,37678	0,3368	-1,2842	0,199181	0,742317	0,7368	-0,2949	0,768065
500	0,545862	274	2442	0,365129	0,3429	-0,6593	0,509753	0,710558	0,7423	1,5969	0,110406
600	0,54526	128	2488	0,401055	0,3425	-1,2347	0,217057	0,741211	0,7310	-0,3434	0,731293
700	0,5401	128	2388	0,386706	0,3461	-0,9066	0,364718	0,742869	0,7273	-0,5147	0,606831

Tabela 33 Varredura das diferentes janelas de Hurst e estatísticas de acurácia do MAPE para as previsões com e sem ondaletas e teste de significância para as ações da VALE5.

Janela	Expoente de Hurst médio	Número de Exp. Hurst < 0,5	Número de Exp. Hurst > 0,5	MAPE_SW Exp. Hurst < 0,5 - média	MAPE_SW Exp. Hurst > 0,5 - média	T-teste	Valor p	MAPE_CW Exp. Hurst < 0,5 - média	MAPE_CW Exp. Hurst > 0,5 - média	T-teste	Valor p
60	0,556333	1024	2132	1,519394	1,355539	3,955921	0,000024	0,952512	0,857625	3,647644	0,000155
80	0,550412	1055	2081	1,449102	1,395918	1,354174	0,171362	0,899719	0,886077	0,547319	0,58714
100	0,553932	991	2125	1,379487	1,434708	-1,39731	0,164898	0,865901	0,904865	-1,5335	0,130843
120	0,549921	916	2180	1,350244	1,443848	-2,36753	0,02101	0,840949	0,911224	-2,76878	0,007202
150	0,547497	799	2267	1,27183	1,481098	-5,75137	0,000000	0,796757	0,931521	-5,71101	0,000001
180	0,549157	663	2373	1,267753	1,452887	-5,07344	0,000022	0,802351	0,908055	-4,34945	0,000184
200	0,549671	605	2411	1,237201	1,452008	-5,85239	0,000001	0,776686	0,908431	-5,37997	0,000006
250	0,550006	541	2425	1,250398	1,437863	-4,76102	0,000060	0,795807	0,893234	-3,72082	0,001377
300	0,547671	362	2554	1,061763	1,425657	-10,6034	0,000000	0,665089	0,885511	-9,6412	0,000000
400	0,544979	314	2502	1,162814	1,402051	-5,43853	0,000008	0,717252	0,865624	-4,81057	0,000024
500	0,545862	274	2442	1,137288	1,408497	-6,40772	0,000002	0,742905	0,867996	-4,29784	0,000833
600	0,54526	128	2488	1,169795	1,377724	-3,101	0,008235	0,774871	0,854309	-1,68901	0,126133
700	0,5401	128	2388	1,098296	1,384589	-5,53951	0,000269	0,715308	0,861392	-4,14749	0,004636

Tabela 34 Varredura das diferentes janelas de Hurst e estatísticas de acurácia do TIC para as previsões com e sem ondaletas e teste de significância para a VALE5.

Janela	Total de Exponentes de Hurst	Número de Exp. Hurst < 0,5	Número de Exp. Hurst > 0,5	TIC_SW Exp. Hurst < 0,5 - média	TIC_SW Exp. Hurst > 0,5 - média	T-teste	Valor p	TIC_CW Exp. Hurst < 0,5 - média	TIC_CW Exp. Hurst > 0,5 - média	T-teste	Valor p
60	0,556333	1024	2132	0,008823	0,007897	3,829054	0,00005	0,005515	0,004978	3,528094	0,000243
80	0,550412	1055	2081	0,008376	0,008147	1,003807	0,315841	0,005204	0,005152	0,361945	0,722232
100	0,553932	991	2125	0,008019	0,008355	-1,45043	0,15194	0,005007	0,005274	-1,80323	0,078348
120	0,549921	916	2180	0,007848	0,00842	-2,46544	0,016702	0,004897	0,005304	-2,732	0,008172
150	0,547497	799	2267	0,007391	0,008642	-5,87568	0,00000	0,004612	0,005441	-6,0879	0,000000
180	0,549157	663	2373	0,007373	0,008484	-5,23574	0,00001	0,004672	0,005307	-4,47979	0,000146
200	0,549671	605	2411	0,007203	0,008485	-6,00842	0,000001	0,00453	0,005317	-5,49927	0,000005
250	0,550006	541	2425	0,00728	0,008393	-4,87892	0,000051	0,004609	0,005241	-4,12122	0,000464
300	0,547671	362	2554	0,006171	0,008318	-10,7339	0,000000	0,003844	0,00519	-10,3663	0,000000
400	0,544979	314	2502	0,006822	0,008158	-5,20948	0,000016	0,004201	0,005068	-4,80096	0,000028
500	0,545862	274	2442	0,006644	0,008212	-6,43001	0,000002	0,004272	0,005087	-4,92129	0,000212
600	0,54526	128	2488	0,00678	0,00805	-3,37554	0,005553	0,004362	0,005012	-2,56172	0,032952
700	0,5401	128	2388	0,006353	0,008118	-5,91437	0,000114	0,004093	0,005062	-4,82726	0,001372

Tabela 35 Varredura das diferentes janelas de Hurst e estatísticas de acurácia do RMSE para as previsões com e sem ondaletas e teste de significância para a VALE5.

Janela	Total de Expoentes de Hurst	Número de Exp. Hurst < 0,5	Número de Exp. Hurst > 0,5	RMSE_SW Exp. Hurst < 0,5 - média	RMSE_SW Exp. Hurst > 0,5 - média	T-teste	Valor p	RMSE_CW Exp. Hurst < 0,5 - média	RMSE_CW Exp. Hurst > 0,5 - média	T-teste	Valor p
60	0,556333	1024	2132	0,279153	0,25131	2,0359	0,02703	0,173769	0,154452	2,274	0,013457
80	0,550412	1055	2081	0,285256	0,250846	2,5429	0,006151	0,174904	0,155366	2,3645	0,012164
100	0,553932	991	2125	0,289383	0,252494	2,7776	0,004008	0,180829	0,154239	3,2142	0,000796
120	0,549921	916	2180	0,274953	0,262389	0,9602	0,339509	0,171771	0,160245	1,4084	0,156324
150	0,547497	799	2267	0,25306	0,274929	-1,785	0,113298	0,157511	0,168249	-1,404	0,20959
180	0,549157	663	2373	0,223122	0,285794	-4,945	0,000023	0,141042	0,173676	-4,05	0,000364
200	0,549671	605	2411	0,186703	0,295959	-9,111	0,00000	0,118245	0,179504	-7,996	0,00000
250	0,550006	541	2425	0,169547	0,303185	-10,734	0,00000	0,10606	0,183351	-9,958	0,00000
300	0,547671	362	2554	0,162439	0,300784	-9,991	0,00000	0,105196	0,180916	-8,353	0,00000
400	0,544979	314	2502	0,195293	0,305755	-5,939	0,00000	0,115204	0,18401	-6,139	0,00000
500	0,545862	274	2442	0,148231	0,321446	-12,789	0,00000	0,094551	0,192592	-11,583	0,00000
600	0,54526	128	2488	0,070324	0,326589	-32,536	0,00000	0,045079	0,196161	-30,929	0,00000
700	0,5401	128	2388	0,064588	0,339266	-35,194	0,00000	0,041615	0,203763	-33,725	0,00000

Das tabelas acima pode-se verificar que realmente, o valor p é significativo para os medidores MAPE, TIC e RMSE na maioria de suas janelas. Nas janelas de varredura da série de 60 e de 120 a 700 pontos, sendo que para a janela de 600 pontos há uma significância apenas para a varredura sem o processamento de ondaletas, para o MAPE e o TIC e o erro de previsão diminui para a série antipersistentes. Já para o RMSE há uma significância para as janelas de 60 pontos apenas com o processamento de ondaletas e de 80,100 e de 180 a 700 para as janelas com e sem o processamento de ondaletas. Para as janelas de 60 a 100 pontos há uma melhora na previsão para a série persistente e para as janelas de 180 a 700 pontos há uma melhora na previsão para a série antipersistentes. Outra observação importante é que as janelas filtradas por ondaletas apresentam um menor erro de previsão.

4.2.5 Análise Itaú (ITSA4)

Com mais de seis décadas de história, o Itaú é hoje um dos maiores bancos em valor de mercado da América Latina e possui a marca mais valiosa do Brasil. A história dessa companhia também foi marcada por importantes fusões e aquisições, como parte do processo de crescimento. Uma das estratégias mais fundamentais foi a compra de bancos públicos estaduais em importantes mercados do País – Rio de Janeiro, Minas Gerais, Paraná e Goiás –, aumentando a presença nacional do Itaú e conquistando a liderança em cada um deles. Além disso, com a expansão dos últimos anos, tornou-se o banco nacional com maior volume de ativos no exterior. (Investmais, 2009)

O Itaú é líder em diversos segmentos do mercado brasileiro: cartões de crédito, seguro residencial, *private banking*, alta renda, financiamento de veículos, serviços de custódia de ativos, etc. É também um dos grupos mais rentáveis do sistema financeiro nacional, especialmente no que se refere ao lucro bancário. O banco chegou ao fim de 2007 com uma rede de mais de 3,5 mil agências e postos de serviços e um quadro de mais de 65 mil funcionários, um aumento superior a 20 mil colaboradores agregados apenas nos últimos quatro anos. (Investmais, 2009)

Em novembro de 2008, o Itaú fundiu-se com o Unibanco e tornou-se o maior conglomerado financeiro do hemisfério sul. O novo banco passou a ter um patrimônio líquido de aproximadamente R\$43,7 bilhões, que o coloca na 16^a posição no ranking das maiores empresas do mundo em termos de valor de mercado. Na Bovespa, o Itaú

Unibanco é a segunda maior força do IBrX-50 e a terceira do Ibovespa. (Investmais, 2009)

Nesta pesquisa analisou-se os preços diários da ação preferencial do itaú (itsa4), no período de 02 de janeiro de 1998 à 30 de dezembro de 2010, perfazendo 3216 observações.

Assim, a série temporal será varrida tomando sempre 500 pontos de cada vez e considerando 496 para estimação da rede e os 4 últimos para previsão. Como a série da soja no período analisado tem 3215 cotações de preços e 3214 log retornos, para a janela de Hurst de 500 pontos, serão feitas 2711 iterações, ou seja, serão calculados 2711 expoentes de Hurst e serão obtidas a mesma quantidade de métricas de acurácia das previsões para cada bloco de 500 pontos.

A seguir, mostra-se o gráfico da série dos preços da soja e os respectivos expoentes de Hurst ao longo da série para uma janela de 500 pontos.

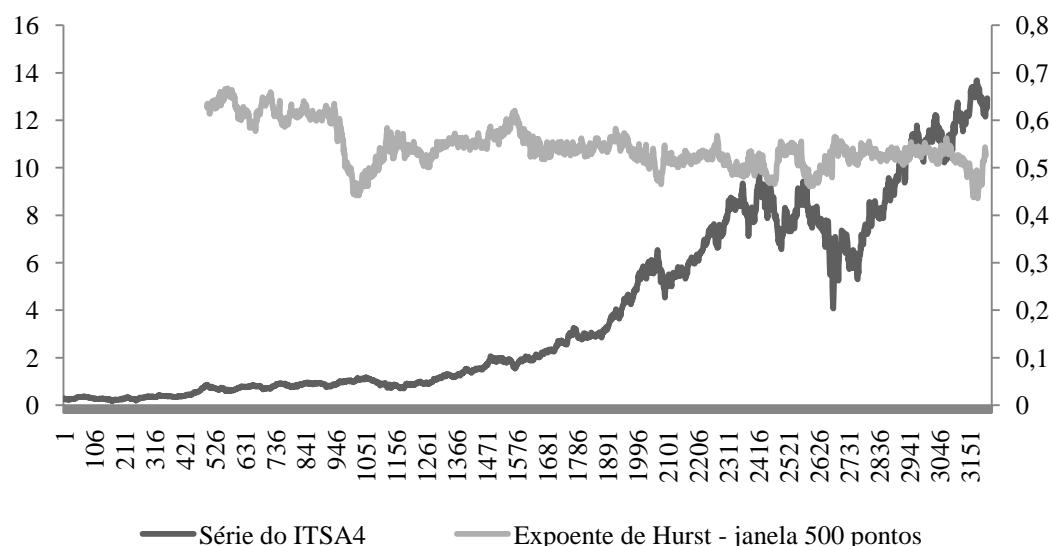


Figura 37: Série do ITSA4 e os expoentes de Hurst ao longo da série para uma janela de 500 pontos

Nota-se no gráfico que a série do itaú apresenta poucos valores do expoente de Hurst abaixo de 0,5. Na sequência da análise mostra-se os medidores de qualidade com e sem filtro de ondaletas utilizando o modelo de Redes Neurais Recorrentes, em função do expoente de Hurst.

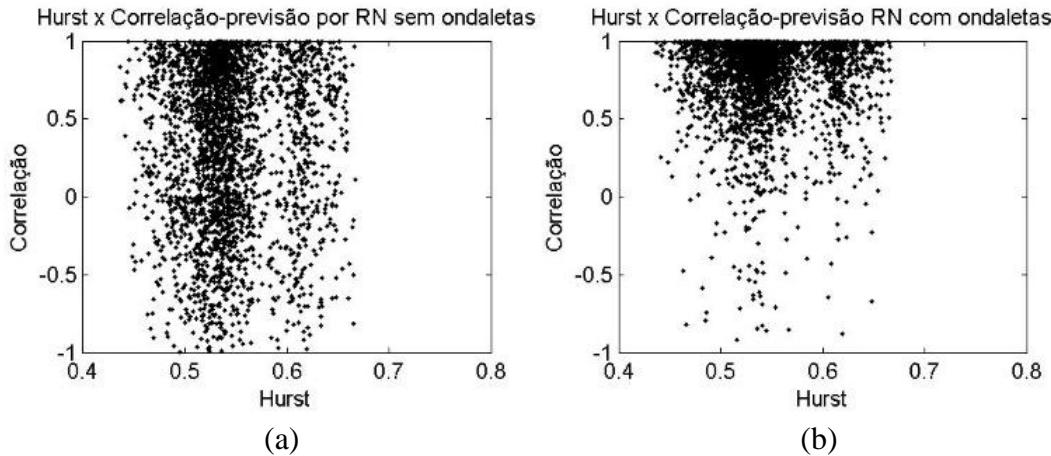


Figura 38: Correlação em função do Exponente de Hurst calculados para uma janela de 500 pontos com o modelo de Redes Neurais Recorrentes com e sem filtro de ondaletas para o ITSA4.

Na inspeção visual dos pontos destacados na figura acima, nota-se claramente que a correlação aumenta quando se aplica o filtro de ondaletas nas janelas de Hurst como demonstrado na figura 38(b) por estarem mais próximos de um, ao passo que na figura 38(a), os valores das correlações encontram-se mais espalhados.

Em relação ao indicador MAPE, a concentração do erro cai na figura 39(b) com mais pontos próximos ao eixo do que em relação à figura 39(a) como pode ser visto na figura a seguir.

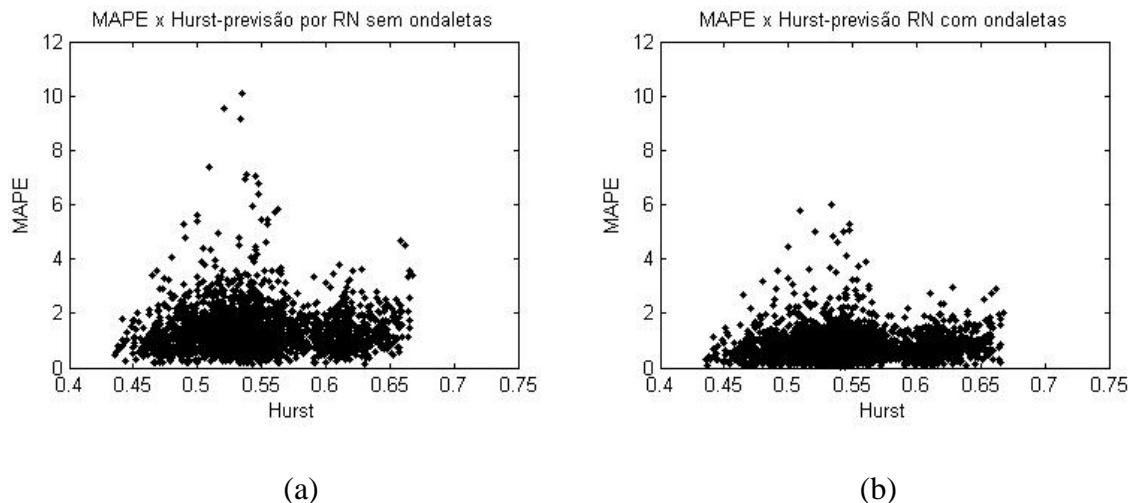


Figura 39: MAPE em função do Exponente de Hurst calculados para uma janela de 500 pontos com o modelo de Redes Neurais Recorrentes com e sem filtro de ondaletas para o ITSA4.

Em relação ao TIC, a análise é muito semelhante, com redução do erro apresentando maior concentração pelo uso do pré-processamento por ondaletas como pode ser visto na figura a seguir.

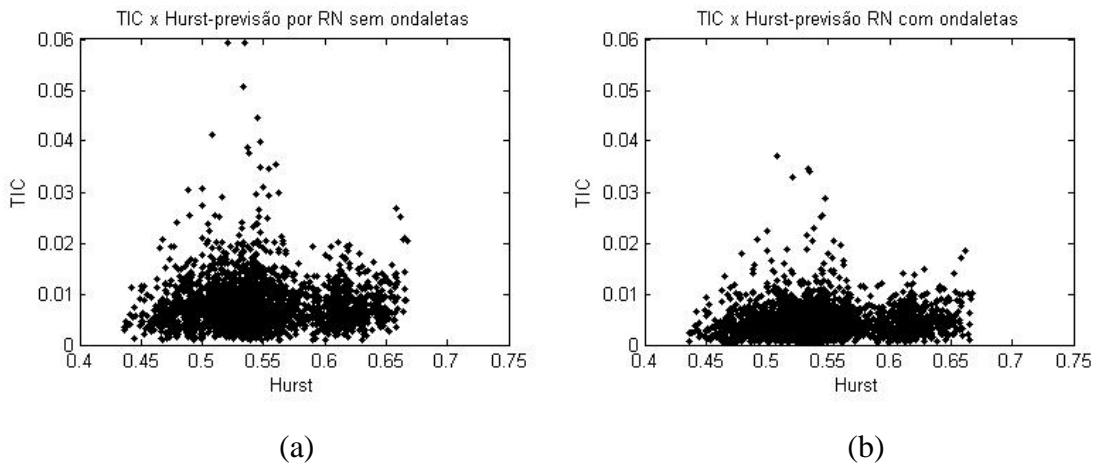


Figura 40: TIC em função do Exponente de Hurst calculados para uma janela de 500 pontos com o modelo de Redes Neurais Recorrentes com e sem filtro de ondaletas do ITSA4.

E para o RMSE também apresenta redução do erro a partir da filtragem em alta e baixa freqüência e com as previsões feitas nas sub-séries como pode ser visto na figura a seguir.

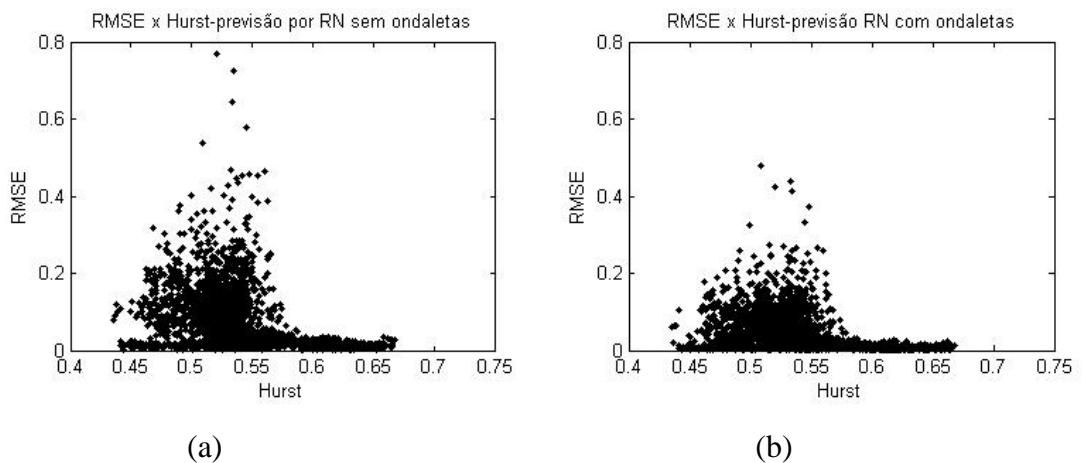


Figura 41: RMSE em função do Exponente de Hurst calculados para uma janela de 500 pontos com o modelo de Redes Neurais Recorrentes com e sem filtro de ondaletas do ITSA4.

Posteriormente, separou-se os valores do expoente de Hurst onde os valores eram maiores que meio e menores que meio. Calculou-se a média dos indicadores de acurácia das previsões em cada conjunto de expoentes e obte-se os seguintes resultados:

Tabela 36 Estatísticas de acurácia das previsões para expoentes de Hurst maiores e menores do que 0,5 para o ITSA4.

Hurst < 0,5 – 374 casos					
	CORREL_SW	TIC_SW	MAPE_SW	RMSE_SW	Expoente de Hurst
Média	0,340074	0,00728	1,267525	0,09954706	0,480078
DP	0,504292	0,004341	0,769773	0,07791958	0,014427

Hurst < 0,5 – 374 casos					
	CORREL_CW	TIC_CW	MAPE_CW	RMSE_CW	Expoente de Hurst
Média	0,769312	0,004302	0,752699	0,05799813	0,480078
DP	0,280653	0,003059	0,550882	0,05129012	0,014427

Hurst > 0,5 – 2341 casos					
	CORREL_SW	TIC_SW	MAPE_SW	RMSE_SW	Expoente de Hurst
Média	0,318429	0,007812	1,352679	0,06709771	0,553778
DP	0,519342	0,004808	0,837047	0,07819586	0,038561

Hurst > 0,5 – 2341 casos					
	CORREL_CW	TIC_CW	MAPE_CW	RMSE_CW	Expoente de Hurst
Média	0,759947	0,004652	0,809619	0,03935987	0,553778
DP	0,279801	0,003223	0,562844	0,04812911	0,038561

Tabela 37 Teste de hipóteses de diferenças de médias das estatísticas de acurácia com e sem ondaletas para os expoentes de Hurst maiores e menores que 0,5 para o ITSA4.

	CORREL_SW	TIC_SW	MAPE_SW	RMSE_SW
t _{teste}	-0,7513	-2,15971	-1,95382	7,433087
valor p	0,450114	0,0449	0,0659	0,0000
Resultado	Aceita H ₀	Rejeita H ₀	Aceita H ₀	Rejeita H ₀

	CORREL_CW	TIC_CW	MAPE_CW	RMSE_CW
t _{teste}	-0,6007	-2,03563	-1,84686	6,529608
valor p	0,536774	0,0499	0,0690	0,0000
Resultado	Aceita H ₀	Rejeita H ₀	Aceita H ₀	Rejeita H ₀

De acordo com resultados acima, para as previsões feitas com e sem pré-processamento de ondaletas em altas e baixas freqüências, nota-se que para os medidores de qualidade: Correlação e MAPE, o expoente de Hurst não mostra indícios de interferir nos resultados das previsões, ou seja, mesmo quando é menor que 0,5 e apresenta comportamento anti-persistente.

Já para os medidores de qualidade TIC e RMSE com e sem o pré processamento de ondaletas o expoente de Hurst mostra indícios de interferir nos resultados das

previsões, ou seja, consegue diferenciar os parâmetros das médias das séries persistentes e antipersistentes.

E assim, mudam-se o tamanho da janela de Hurst e obtêm-se as previsões e as estatísticas de erro para cada janela e seu respectivo teste de hipótese de diferença de médias para cada métrica visando verificar se a mudança do expoente de Hurst acima e abaixo de 0,5 influenciam nos resultados das previsões. Os resultados são mostrados nas tabelas a seguir:

Tabela 38 Varredura das diferentes janelas de Hurst e estatísticas de acurácia da correlação para as previsões com e sem ondaletas e teste de significância para as ações do ITSA4.

Janela	Expoente de Hurst médio	Número de Exp. Hurst < 0,5	Número de Exp. Hurst > 0,5	CORREL_SW Exp. Hurst < 0,5 - média	CORREL_SW Exp. Hurst > 0,5 - média	T-teste	Valor p	CORREL_CW Exp. Hurst < 0,5 - média	CORREL_CW Exp. Hurst > 0,5 - média	T-teste	Valor p
60	0,558223	959	2196	0,322137	0,3445	1,1117	0,267914	0,755671	0,7637	0,7444	0,467793
80	0,557817	926	2209	0,324377	0,3433	0,9296	0,354504	0,752597	0,7677	1,4101	0,164086
100	0,558601	794	2321	0,348243	0,3307	-0,8205	0,410038	0,748455	0,7674	1,6824	0,096085
120	0,556203	805	2290	0,33134	0,3362	0,2275	0,823264	0,756579	0,7645	0,7078	0,488496
150	0,555152	736	2329	0,327678	0,3382	0,4770	0,636906	0,762171	0,7640	0,1557	0,893134
180	0,551548	647	2388	0,315971	0,3412	1,0944	0,27585	0,749971	0,7692	1,6082	0,112692
200	0,549688	677	2338	0,317766	0,3400	0,9796	0,329666	0,764297	0,7659	0,1364	0,909171
250	0,549014	644	2321	0,317356	0,3350	0,7450	0,445366	0,753103	0,7672	1,1074	0,243439
300	0,550068	436	2479	0,284586	0,3359	1,8921	0,056878	0,758594	0,7639	0,2570	0,708019
400	0,544319	390	2425	0,309824	0,3270	0,5933	0,541579	0,755883	0,7651	0,4926	0,537858
500	0,543597	374	2341	0,340074	0,3184	-0,7513	0,450114	0,769312	0,7600	-0,6007	0,536774
600	0,546033	80	2535	0,289977	0,3177	0,4701	0,640118	0,781523	0,7601	-0,6747	0,494095
700	0,539359	242	2273	0,265273	0,3240	1,6793	0,094075	0,753602	0,7593	0,3047	0,774166

Tabela 39 Varredura das diferentes janelas de Hurst e estatísticas de acurácia do MAPE para as previsões com e sem ondaletas e teste de significância para as ações do ITSA4.

Janela	Expoente de Hurst médio	Número de Exp. Hurst < 0,5	Número de Exp. Hurst > 0,5	MAPE_SW Exp. Hurst < 0,5 - média	MAPE_SW Exp. Hurst > 0,5 - média	T-teste	Valor p	MAPE_CW Exp. Hurst < 0,5 - média	MAPE_CW Exp. Hurst > 0,5 - média	T-teste	Valor p
60	0,558223	959	2196	1,365908	1,365383	0,014981	0,9878	0,849409	0,849414	-0,00022	0,9998
80	0,557817	926	2209	1,316516	1,388733	-2,11302	0,0389	0,824211	0,855465	-1,3556	0,1859
100	0,558601	794	2321	1,264366	1,408149	-3,90481	0,0000	0,805357	0,862449	-2,32235	0,0219
120	0,556203	805	2290	1,211736	1,42689	-6,18349	0,0000	0,75836	0,877171	-5,06875	0,0000
150	0,555152	736	2329	1,203116	1,428726	-6,48309	0,0000	0,737316	0,881259	-6,02262	0,0000
180	0,551548	647	2388	1,230934	1,388746	-4,48532	0,0000	0,758427	0,849935	-3,76971	0,0004
200	0,549688	677	2338	1,165428	1,40701	-7,73032	0,0000	0,700591	0,865018	-7,78734	0,0000
250	0,549014	644	2321	1,240411	1,372913	-3,27516	0,0005	0,760066	0,832564	-2,65231	0,0050
300	0,550068	436	2479	1,223752	1,350457	-2,58435	0,0031	0,731909	0,817963	-2,70909	0,0029
400	0,544319	390	2425	1,378372	1,332703	0,810143	0,3106	0,838036	0,797593	1,092504	0,1850
500	0,543597	374	2341	1,267525	1,352312	-1,95382	0,0659	0,752699	0,809519	-1,84686	0,0690
600	0,546033	80	2535	1,132244	1,345809	-2,96016	0,0229	0,636307	0,800963	-3,3115	0,0094
700	0,539359	242	2273	1,431846	1,337485	1,661293	0,0961	0,865139	0,795718	1,723339	0,0692

Tabela 40 Varredura das diferentes janelas de Hurst e estatísticas de acurácia do TIC para as previsões com e sem ondaletas e teste de significância para o ITSA4.

Janela	Total de Exponentes de Hurst	Número de Exp. Hurst < 0,5	Número de Exp. Hurst > 0,5	TIC_SW Exp. Hurst < 0,5 - média	TIC_SW Exp. Hurst > 0,5 - média	T-teste	Valor p	TIC_CW Exp. Hurst < 0,5 - média	TIC_CW Exp. Hurst > 0,5 - média	T-teste	Valor p
60	0,558223	959	2196	0,007918	0,007896	0,109426	0,9109	0,004898	0,004901	-0,02091	0,983
80	0,557817	926	2209	0,007625	0,008034	-2,07667	0,0413	0,004738	0,004933	-1,48389	0,1441
100	0,558601	794	2321	0,007298	0,008156	-4,04808	0,0000	0,004635	0,004965	-2,30991	0,0192
120	0,556203	805	2290	0,007005	0,008257	-6,3027	0,0000	0,004363	0,005043	-5,05903	0,0000
150	0,555152	736	2329	0,00694	0,008264	-6,62707	0,0000	0,004238	0,005065	-6,01961	0,0000
180	0,551548	647	2388	0,007108	0,008028	-4,53243	0,0000	0,004373	0,00488	-3,58111	0,0006
200	0,549688	677	2338	0,00672	0,008135	-7,88811	0,0000	0,004047	0,004963	-7,57142	0,0000
250	0,549014	644	2321	0,007162	0,007926	-3,28077	0,0005	0,004374	0,004781	-2,54404	0,0055
300	0,550068	436	2479	0,007083	0,007798	-2,52004	0,0036	0,004213	0,004702	-2,58967	0,0029
400	0,544319	390	2425	0,007941	0,007701	0,731608	0,3543	0,004861	0,004577	1,289584	0,1019
500	0,543597	374	2341	0,00728	0,00781	-2,15971	0,0449	0,004302	0,004652	-2,03563	0,0499
600	0,546033	80	2535	0,006527	0,007765	-3,03956	0,0212	0,003697	0,004597	-3,15172	0,0127
700	0,539359	242	2273	0,008199	0,007718	1,498447	0,1382	0,004917	0,004577	1,536071	0,1195

Tabela 41 Varredura das diferentes janelas de Hurst e estatísticas de acurácia do RMSE para as previsões com e sem ondaletas e teste de significância para as ações do ITSA4.

Janela	Total de Exponentes de Hurst	Número de Exp. Hurst < 0,5	Número de Exp. Hurst > 0,5	RMSE_SW Exp. Hurst < 0,5 - média	RMSE_SW Exp. Hurst > 0,5 - média	T-teste	Valor p	RMSE_CW Exp. Hurst < 0,5 - média	RMSE_CW Exp. Hurst > 0,5 - média	T-teste	Valor p
60	0,558223	959	2196	0,063794	0,058261	1,828279	0,0523	0,039607	0,035914	1,890561	0,0426
80	0,557817	926	2209	0,059694	0,060701	-0,34614	0,7281	0,036989	0,037014	-0,01351	0,9891
100	0,558601	794	2321	0,070747	0,057466	4,186845	0,0000	0,044555	0,034769	4,722386	0,0000
120	0,556203	805	2290	0,072923	0,057166	5,082847	0,0000	0,04528	0,034602	5,270505	0,0000
150	0,555152	736	2329	0,080177	0,056162	7,236661	0,0000	0,049055	0,034149	6,845479	0,0000
180	0,551548	647	2388	0,087491	0,055877	9,251338	0,0000	0,053712	0,033746	8,810464	0,0000
200	0,549688	677	2338	0,089895	0,055355	10,98084	0,0000	0,054358	0,033485	10,12356	0,0000
250	0,549014	644	2321	0,093409	0,056402	9,571178	0,0000	0,057542	0,033502	9,352824	0,0000
300	0,550068	436	2479	0,104482	0,05894	9,224369	0,0000	0,062372	0,035234	8,49993	0,0000
400	0,544319	390	2425	0,106208	0,062642	8,232034	0,0000	0,065132	0,036528	8,258019	0,0000
500	0,543597	374	2341	0,099547	0,067275	7,433087	0,0000	0,057998	0,039499	6,529608	0,0000
600	0,546033	80	2535	0,105389	0,073586	4,417069	0,0005	0,059738	0,042866	3,526047	0,0027
700	0,539359	242	2273	0,123106	0,072209	9,787686	0,0000	0,073108	0,041895	8,95521	0,0000

Analisando-se a série da ação do Itaú (itsa4), no período de 10 de agosto de 2000 à dezembro de 2010, pode-se verificar que a correlação permanece não apresentando nenhuma significância para nenhuma das janelas testadas. Para os indicadores MAPE e TIC há uma rejeição da hipótese nula para as janelas de 100 à 300 e 600 pontos e o erro de previsão diminui para as séries antipersistentes. Já para o RMSE nas janelas de 100 à 700 pontos o erro de previsão diminui para as séries persistentes. Como para as séries testadas anteriormente a filtragem da série por ondaletas apresenta uma melhora significativa no erro de previsão.

4.2.6 Análise Petrobrás (PETR4)

A Petrobrás foi fundada no dia 3 de outubro de 1953, pelo então presidente Getúlio Vargas, com o objetivo de executar as atividades do setor petrolífero no Brasil em nome da União. Até 1997, a empresa detinha o monopólio das operações de exploração e produção de petróleo, bem como as demais atividades ligadas ao setor, à exceção da distribuição atacadista e da revenda no varejo pelos postos de abastecimento.

Atualmente, a Companhia está presente em 27 países. Em 2007, a Petrobras foi classificada como a 7^a maior empresa de petróleo do mundo com ações negociadas em bolsas de valores, de acordo com a Petroleum Intelligence Weekly (PIW), publicação que divulga anualmente o ranking das 50 maiores e mais importantes empresas de petróleo. Além disso, no início de 2008, a Petrobrás foi reconhecida através de pesquisa da Management & Excellence (M&E) a petroleira mais sustentável do mundo.

No primeiro trimestre de 2011, o lucro líquido foi de R\$ 10,99 bilhões (R\$ 0,84 por ação) 42% superior ao lucro do primeiro trimestre de 2010, que foi de R\$ 7,73 bilhões. O lucro bruto alcançou R\$ 20,20 bilhões no primeiro trimestre do ano, 5% superior ao mesmo período de 2010 (R\$ 19,31 bilhões).

Na bolsa de valores de São Paulo, BM&F Bovespa, as ações da Petrobrás estão em segundo lugar no ranking das mais negociadas, atrás apenas da Vale do Rio Doce. Nesta pesquisa analisou-se os preços diários das ações preferenciais da Petrobrás (PETR4), no período de 02 de janeiro de 1998 à 30 de dezembro de 2010, perfazendo 3215 observações.

Assim, a série temporal será varrida tomando sempre 500 pontos de cada vez e considerando 496 para estimação da rede e os 4 últimos para previsão. Como a série da soja no período analisado tem 3215 cotações de preços e 3214 log retornos, para a janela de Hurst de 500 pontos, serão feitas 2715 iterações, ou seja, serão calculados 2715 expoentes de Hurst e serão obtidas a mesma quantidade de métricas de acurácia das previsões para cada bloco de 500 pontos.

A seguir, mostra-se o gráfico da série dos preços da soja e os respectivos expoentes de Hurst ao longo da série para uma janela de 500 pontos.

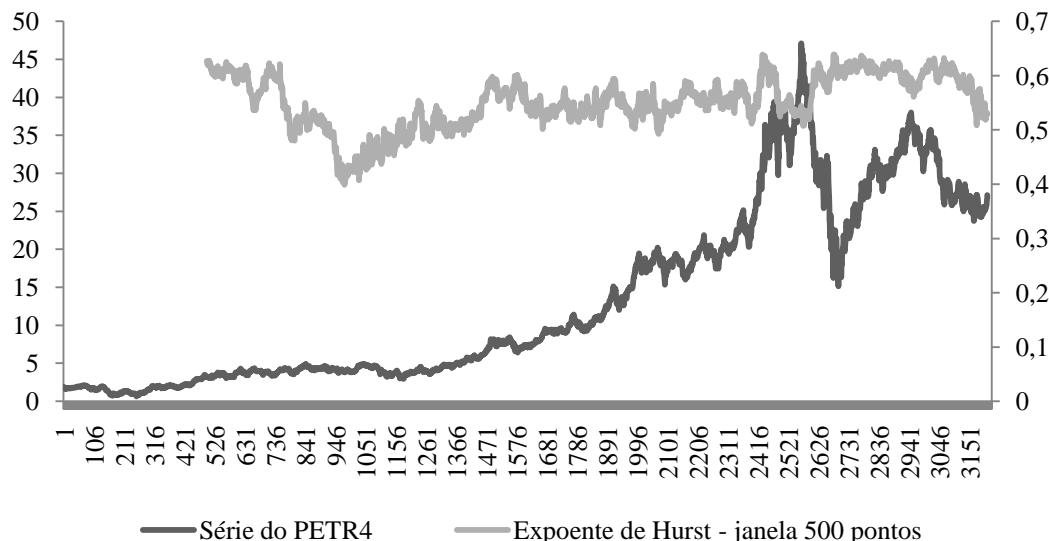


Figura 42: Série da PETR4 e os expoentes de Hurst ao longo da série para uma janela de 500 pontos.

Nota-se no gráfico que a série da Petrobrás apresenta poucos valores do expoente de Hurst abaixo de 0,5. Na sequência da análise mostra-se os medidores de qualidade com e sem filtro de ondaletas utilizando o modelo de Redes Neurais Recorrentes, em função do expoente de Hurst.

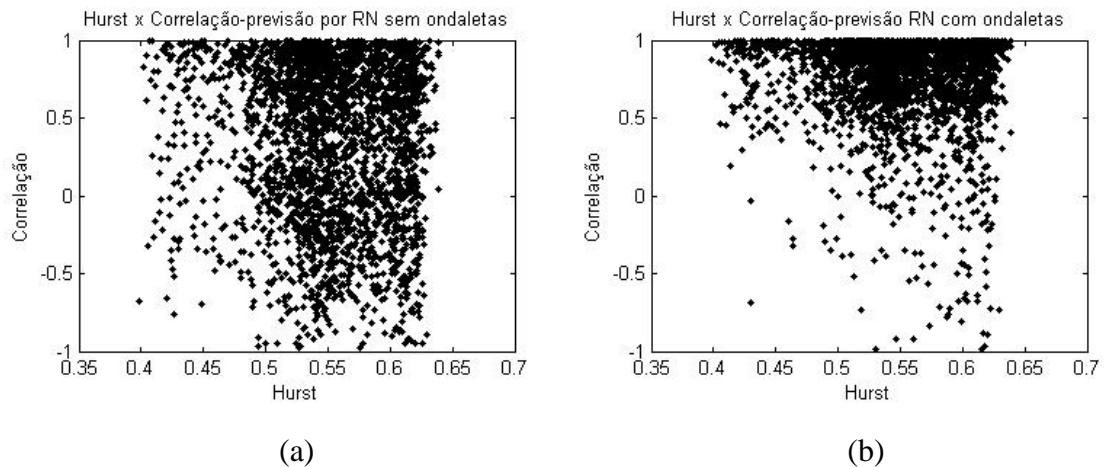


Figura 43: Correlação em função do Exponente de Hurst calculados para uma janela de 500 pontos com o modelo de Redes Neurais Recorrentes com e sem filtro de ondaletas para a PETR4.

Na inspeção visual dos pontos destacados na figura acima, nota-se claramente que a correlação aumenta quando se aplica o filtro de ondaletas nas janelas de Hurst como demonstrado na figura 43(b) por estarem mais próximos de um, ao passo que na figura 43(a), os valores das correlações encontram-se mais espalhados.

Em relação ao indicador MAPE, a concentração do erro cai na figura 44(b) com mais pontos próximos ao eixo do que em relação à figura 44(a) como pode ser visto na figura a seguir.

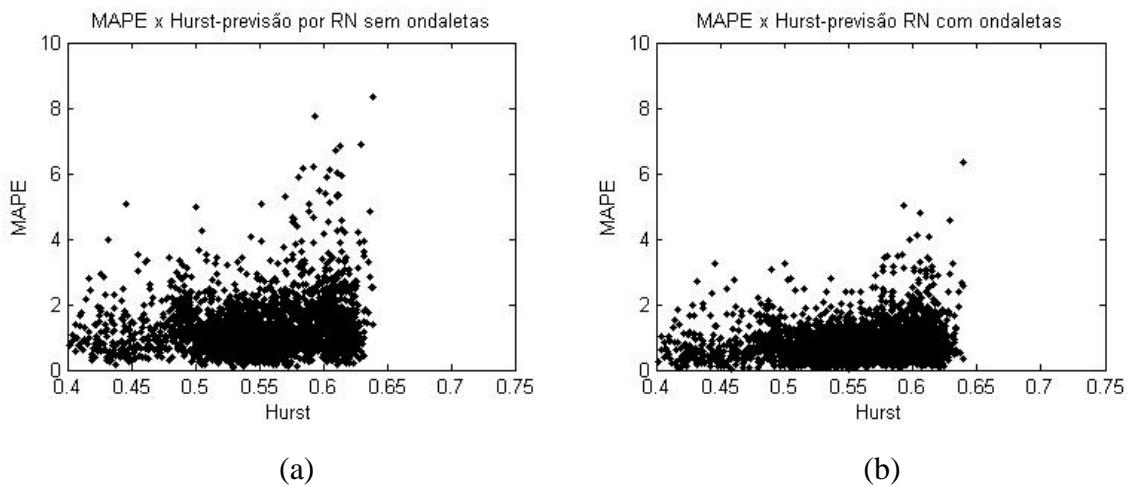


Figura 44: MAPE em função do Exponente de Hurst calculados para uma janela de 500 pontos com o modelo de Redes Neurais Recorrentes com e sem filtro de ondaletas para a PETR4.

Em relação ao TIC, a análise é muito semelhante, com redução do erro apresentando maior concentração pelo uso do pré-processamento por ondaletas como pode ser visto na figura a seguir.

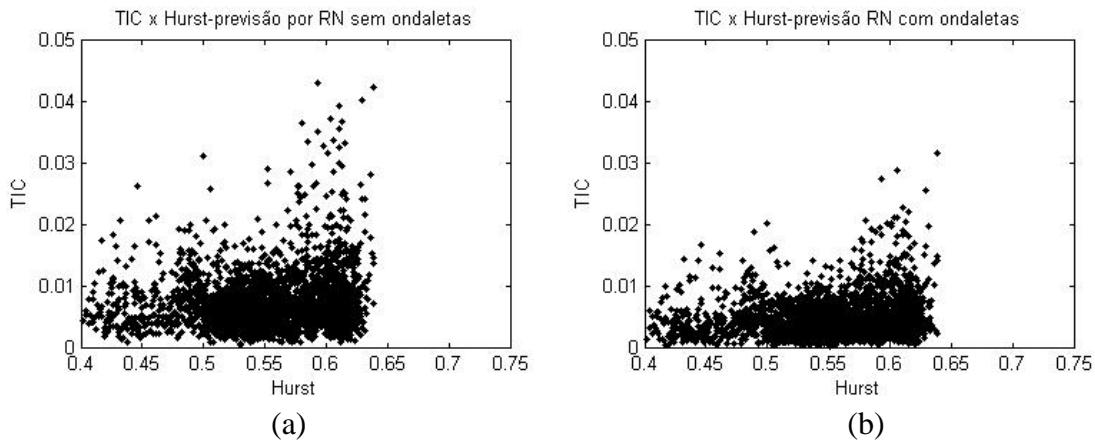


Figura 45: TIC em função do Exponente de Hurst calculados para uma janela de 500 pontos com o modelo de Redes Neurais Recorrentes com e sem filtro de ondaletas para a PETR4.

E para o RMSE também apresenta redução do erro a partir da filtragem em alta e baixa freqüência e com as previsões feitas nas sub-séries como pode ser visto na figura a seguir.

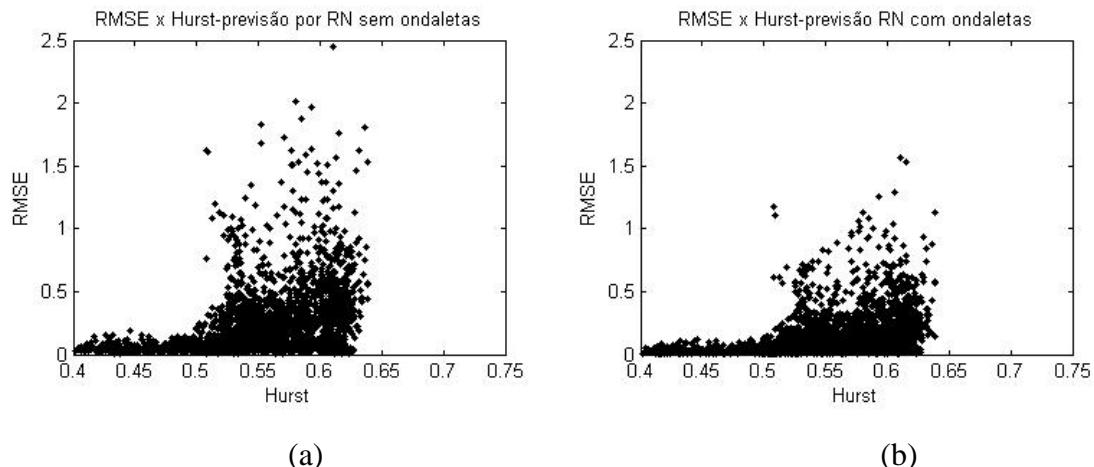


Figura 46: RMSE em função do Exponente de Hurst calculados para uma janela de 500 pontos com o modelo de Redes Neurais Recorrentes com e sem filtro de ondaletas para a PETR4.

Posteriormente, separou-se os valores do expoente de Hurst onde os valores eram maiores que meio e menores que meio. Calculou-se a média dos indicadores de acurácia das previsões em cada conjunto de expoentes e obteve-se os seguintes resultados:

Tabela 42 Estatísticas de acurácia das previsões para expoentes de Hurst maiores e menores do que 0,5 para a PETR4.

Hurst < 0,5 – 349 casos					
	CORREL_SW	TIC_SW	MAPE_SW	RMSE_SW	Expoente de Hurst
Média	0,437812	0,007476	1,28785	0,06045281	0,463631
DP	0,484958	0,004246	0,740645	0,03321633	0,027812

Hurst < 0,5 – 349 casos					
	CORREL_CW	TIC_CW	MAPE_CW	RMSE_CW	Expoente de Hurst
Média	0,783777	0,004643	0,801823	0,03713493	0,463631
DP	0,259942	0,003065	0,534747	0,02278888	0,027812

Hurst > 0,5 – 2366 casos					
	CORREL_SW	TIC_SW	MAPE_SW	RMSE_SW	Expoente de Hurst
Média	0,36521	0,007384	1,27758	0,2618766	0,564834
DP	0,509074	0,00482	0,8415	0,2820053	0,035839

Hurst > 0,5 – 2366 casos					
	CORREL_CW	TIC_CW	MAPE_CW	RMSE_CW	Expoente de Hurst
Média	0,752431	0,004648	0,803942	0,1617862	0,564834
DP	0,304519	0,003274	0,570033	0,179534	0,035839

Tabela 43 Teste de hipóteses de diferenças de médias das estatísticas de acurácia com e sem ondaletas para os expoentes de Hurst maiores e menores que 0,5 para as ações da PETR4.

	CORREL_SW	TIC_SW	MAPE_SW	RMSE_SW
t _{teste}	2,593901	0,370795	0,237427	-33,2155
valor p	0,0124	0,7357	0,8290	0,0000
Resultado	Rejeita H ₀	Aceita H ₀	Aceita H ₀	Rejeita H ₀

	CORREL_CW	TIC_CW	MAPE_CW	RMSE_CW
t _{teste}	2,054432	-0,0288	-0,06849	-32,0661
valor p	0,0678	0,9781	0,9479	0,0000
Resultado	Aceita H ₀	Aceita H ₀	Aceita H ₀	Rejeita H ₀

De acordo com resultados acima, para as previsões feitas sem pré-processamento de ondaletas em altas e baixas freqüências, nota-se a aceitação da hipótese nula para os medidores TIC e MAPE, ou seja, não há uma diferença entre as médias comparadas, os testes de hipóteses com 95% de confiança não são conclusivos a cerca da influência do expoente de Hurst na capacidade preditiva da série da Petrobrás. Já para a Correlação e

RMSE demonstra que o expoente de Hurst demonstra uma capacidade de melhora na previsão ao separar a série em persistente e anti-persistente.

Todavia, quando realiza-se a separação da série em altas e baixas freqüências, e realiza-se as previsões dentro das sub-séries, nota-se que a Correlação, o TIC e o MAPE não demonstram diferenças nas estatísticas de acurácia. Somente o RMSE rejeita a hipótese nula de igualdade dos indicadores de previsão quando o expoente de Hurst é menor e quando é maior que 0,5.

Agora, tomam-se os mesmos resultados para diferentes configurações de janelas, ou seja, mudam-se os tamanhos das janelas de estimativa do expoente de Hurst mas mantêm-se as previsões de quatro passos a frente para as distintas janelas. Os resultados obtidos são demonstrados nas tabelas a seguir.

Tabela 44 Varredura das diferentes janelas de Hurst e estatísticas de acurácia da correlação para as previsões com e sem ondaletas e teste de significância para as ações da PETR4.

Janela	Expoente de Hurst médio	Número de Exp. Hurst < 0,5	Número de Exp. Hurst > 0,5	CORREL_SW Exp. Hurst < 0,5 - média	CORREL_SW Exp. Hurst > 0,5 - média	T-teste	Valor p	CORREL_CW Exp. Hurst < 0,5 - média	CORREL_CW Exp. Hurst > 0,5 - média	T-teste	Valor p
60	0,592507	666	2489	0,368279	0,389263	-0,93128	0,3474	0,767248	0,767413	-0,01376	0,9891
80	0,596101	544	2591	0,393605	0,383769	0,409635	0,6826	0,769185	0,765555	0,277121	0,7831
100	0,590755	482	2633	0,372121	0,387158	-0,60218	0,5516	0,766181	0,767889	-0,12697	0,9021
120	0,584967	329	2766	0,361171	0,388326	-0,8714	0,3606	0,770752	0,767937	0,169331	0,8630
150	0,578463	353	2712	0,396924	0,378963	0,654432	0,5321	0,7784	0,764652	0,951585	0,3887
180	0,575211	279	2756	0,385396	0,378722	0,21177	0,8343	0,778744	0,764206	0,796611	0,4150
200	0,565467	391	2624	0,356451	0,382433	-0,94193	0,3451	0,758382	0,7674	-0,56588	0,5607
250	0,558609	445	2520	0,409854	0,372277	1,437461	0,1507	0,777544	0,763123	0,987398	0,3262
300	0,562981	267	2648	0,438248	0,371012	2,229138	0,0387	0,779318	0,761009	1,121858	0,3284
400	0,553307	371	2444	0,405558	0,373146	1,172215	0,2511	0,779371	0,756121	1,538846	0,1633
500	0,551825	349	2366	0,437812	0,36521	2,593901	0,0124	0,783777	0,752431	2,054432	0,0678
600	0,553376	254	2361	0,43272	0,370401	1,946718	0,0612	0,78098	0,756551	1,284692	0,2040
700	0,548293	240	2275	0,42896	0,372437	1,605527	0,0993	0,796609	0,754546	2,178981	0,0341

Tabela 45 Varredura das diferentes janelas de Hurst e estatísticas de acurácia do MAPE para as previsões com e sem ondaletas e teste de significância para as ações da PETR4.

Janela	Expoente de Hurst médio	Número de Exp. Hurst < 0,5	Número de Exp. Hurst > 0,5	MAPE_SW Exp. Hurst < 0,5 - média	MAPE_SW Exp. Hurst > 0,5 - média	T-teste	Valor p	MAPE_CW Exp. Hurst < 0,5 - média	MAPE_CW Exp. Hurst > 0,5 - média	T-teste	Valor p
60	0,592507	666	2489	1,318448	1,370592	-1,27638	0,2301	0,822343	0,873973	-1,90615	0,0799
80	0,596101	544	2591	1,331319	1,369747	-0,87066	0,4146	0,838879	0,87211	-1,12268	0,3022
100	0,590755	482	2633	1,30478	1,372916	-1,48882	0,1679	0,811727	0,87046	-1,95849	0,0830
120	0,584967	329	2766	1,148933	1,384345	-5,05226	0,0000	0,696257	0,876878	-6,00623	0,0000
150	0,578463	353	2712	1,179428	1,388654	-4,98328	0,0002	0,718746	0,876923	-5,36748	0,0000
180	0,575211	279	2756	0,963872	1,376497	-10,9146	0,0000	0,558892	0,869785	-12,3399	0,0000
200	0,565467	391	2624	1,05279	1,366616	-9,20585	0,0000	0,639036	0,86084	-9,09976	0,0000
250	0,558609	445	2520	1,125953	1,341936	-6,2966	0,0000	0,702331	0,840948	-5,56982	0,0000
300	0,562981	267	2648	1,277835	1,282821	-0,10261	0,9251	0,806871	0,802453	0,130102	0,9026
400	0,553307	371	2444	1,240663	1,278643	-0,94946	0,4069	0,756491	0,80599	-1,71067	0,1133
500	0,551825	349	2366	1,28785	1,27758	0,237427	0,8290	0,801823	0,803942	-0,06849	0,9479
600	0,553376	254	2361	1,326529	1,2682	1,108849	0,2887	0,826443	0,798243	0,76373	0,4480
700	0,548293	240	2275	1,216336	1,269458	-1,11961	0,3481	0,748661	0,802208	-1,59771	0,1630

Tabela 46 Varredura das diferentes janelas de Hurst e estatísticas de acurácia do TIC para as previsões com e sem ondaletas e teste de significância para a PETR4.

Janela	Total de Exponentes de Hurst	Número de Exp. Hurst < 0,5	Número de Exp. Hurst > 0,5	TIC_SW Exp. Hurst < 0,5 - média	TIC_SW Exp. Hurst > 0,5 - média	T-teste	Valor p	TIC_CW Exp. Hurst < 0,5 - média	TIC_CW Exp. Hurst > 0,5 - média	T-teste	Valor p
60	0,592507	666	2489	0,007647	0,007917	-1,14179	0,2775	0,004744	0,005046	-1,94685	0,0693
80	0,596101	544	2591	0,007742	0,007912	-0,66677	0,5300	0,00482	0,005035	-1,29444	0,2360
100	0,590755	482	2633	0,00757	0,007934	-1,38661	0,1978	0,004686	0,005017	-1,94442	0,0826
120	0,584967	329	2766	0,006694	0,008002	-4,86136	0,0000	0,004022	0,005056	-6,15797	0,0000
150	0,578463	353	2712	0,006874	0,008025	-4,65812	0,0004	0,004146	0,005056	-5,37594	0,0000
180	0,575211	279	2756	0,005615	0,007962	-10,6939	0,0000	0,003266	0,005013	-11,9415	0,0000
200	0,565467	391	2624	0,006103	0,007908	-9,27217	0,0000	0,003694	0,00497	-9,48055	0,0000
250	0,558609	445	2520	0,00653	0,007762	-6,23167	0,0000	0,004089	0,004852	-5,39542	0,0000
300	0,562981	267	2648	0,007425	0,007421	0,016302	0,9879	0,004678	0,004646	0,160391	0,8779
400	0,553307	371	2444	0,007177	0,007395	-0,94088	0,4056	0,004393	0,004661	-1,59771	0,1350
500	0,551825	349	2366	0,007476	0,007384	0,370795	0,7357	0,004643	0,004648	-0,0288	0,9781
600	0,553376	254	2361	0,007692	0,007326	1,209098	0,2436	0,004771	0,004614	0,738194	0,4610
700	0,548293	240	2275	0,007031	0,00734	-1,12745	0,3394	0,004331	0,004634	-1,56921	0,1668

Tabela 47 Varredura das diferentes janelas de Hurst e estatísticas de acurácia do RMSE para as previsões com e sem ondaletas e teste de significância para as ações da PETR4.

Janela	Total de Expoentes de Hurst	Número de Exp. Hurst < 0,5	Número de Exp. Hurst > 0,5	RMSE_SW Exp. Hurst < 0,5 - média	RMSE_SW Exp. Hurst > 0,5 - média	T-teste	Valor p	RMSE_CW Exp. Hurst < 0,5 - média	RMSE_CW Exp. Hurst > 0,5 - média	T-teste	Valor p
60	0,592507	666	2489	0,242911	0,194761	3,503332	0,0000	0,148965	0,12149	3,190424	0,0001
80	0,596101	544	2591	0,260657	0,19492	4,473358	0,0000	0,159767	0,121427	4,138179	0,0000
100	0,590755	482	2633	0,234256	0,202799	2,345975	0,0144	0,141754	0,125819	1,8934	0,0502
120	0,584967	329	2766	0,187791	0,211467	-1,82304	0,1188	0,11021	0,130698	-2,73397	0,0326
150	0,578463	353	2712	0,159417	0,218189	-5,20698	0,0000	0,093658	0,134838	-5,9778	0,0000
180	0,575211	279	2756	0,142262	0,220698	-7,0572	0,0000	0,081595	0,136167	-8,18352	0,0000
200	0,565467	391	2624	0,161792	0,22274	-5,72642	0,0000	0,094678	0,137472	-6,61462	0,0000
250	0,558609	445	2520	0,122611	0,235456	-12,4174	0,0000	0,074545	0,144537	-12,5106	0,0000
300	0,562981	267	2648	0,074336	0,23667	-24,1349	0,0000	0,046872	0,145543	-23,063	0,0000
400	0,553307	371	2444	0,065764	0,253394	-29,4657	0,0000	0,039321	0,156526	-29,3292	0,0000
500	0,551825	349	2366	0,060453	0,261877	-33,2155	0,0000	0,037135	0,161786	-32,0661	0,0000
600	0,553376	254	2361	0,059623	0,262578	-33,0745	0,0000	0,03675	0,161902	-31,7606	0,0000
700	0,548293	240	2275	0,054515	0,2705	-34,7237	0,0000	0,033458	0,166596	-33,4693	0,0000

Das tabelas acima se pode verificar que, as janelas de varredura de 120 a 250 pontos são significativas para os medidores TIC e MAPE e o erro de previsão diminui para a série antipersistentes. Para o RMSE as janelas significativas são de 60 a 100 e de 150 a 700 pontos e o erro de previsão diminui nas janelas de 60 a 100 pontos para a série persistente e de 150 a 700 pontos para a série antipersistentes.

CONCLUSÃO

Com a pesquisa observa-se que, o expoente de Hurst, é útil para classificar janelas como persistentes e anti-persistentes, concordando com Cajueiro e Tabak (2005) que afirmam que este expoente é um bom indicador para calcular a existência de memória longa.

Os resultados verificados à partir da análise da interferência do Expoente de Hurst foram os seguintes:

Para várias janelas testadas em cada série analisada, na maioria dos casos, quando o expoente de Hurst é menor do que 0,5, a qualidade das previsões melhoraram significativamente para os indicadores MAPE, TIC e RMSE. A correlação não demonstrou significância para nenhuma série testada.

Para o medidor de qualidade MAPE, existem várias janelas significativas em cada série analisada, e a janela de varredura da série de 150 pontos foi significativa para todas as séries observadas. Portanto para o MAPE a janela ótima foi de 150 pontos.

Para o medidor de qualidade TIC, também existem várias janelas significativas em cada série temporal financeira, e a janela com varredura 150 pontos, como para o MAPE, também foi sua janela ótima.

Para o RMSE observa-se que normalmente as janelas com varreduras maiores são as mais significativas do que as menores e não há uma janela ótima que seja significativa para todas as séries ao mesmo tempo.

Outra observação importante, é que para os medidores MAPE e TIC a qualidade das previsões para as séries antipersistentes foi melhor para todos os casos, exceto para a *commodity* da soja e para a janela de 60 pontos da VALE5. Para o RMSE as janelas de 60, 80 e 100 pontos, quando significativas melhoraram a previsão quando a série é persistente, este fato acontece também para todas as janelas significativas da *commodity* de soja, para a série do Ibovespa e para a série do Itaú. Para as janelas maiores que 120 pontos, na grande maioria dos casos, há uma melhor previsão para as séries antipersistentes.

Com isso pode-se concluir que o expoente de Hurst juntamente com o modelo de previsão de Redes Neurais Recorrentes revela indícios de ser capaz de diminuir o erro previsões para, na maioria das vezes, séries antipersistentes, para os medidores TIC, MAPE e para as janelas maiores que 120 pontos para o RMSE. Já para janelas de 60, 80 e 100 pontos e o medidor RMSE as previsões são melhores para as séries persistentes. Em relação à janela ótima, sabe-se que para todos os modelos analisados não há uma em

comum para todos os medidores. No entanto, para se ter certeza do tamanho da janela ótima para cada medidor é necessário que se teste mais séries temporais financeiras, idealizando um próximo trabalho.

REFERÊNCIAS BIBLIOGRÁFIAS

ABIYEV, Rahib H. Fuzzy wavelet neural network for prediction of electricity consumption. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*. Cambridge University Press. Estados Unidos, 2007. p.109-118.

ADESI, Giovanni Barone. ENGLE, Robert F., MANCINI, Loriano. A GARCH Option Pricing Model with Filtered Historical Simulation. **Review of Financial Studies**. Oxford university. V.21, No. 3. Março-2008.p.1223-1258.

AMINIAN, Farzan; SUAREZ, E. Dante; AMINIAN, Mehran; WALZ, Daniel T. Forecasting Economic Data with Neural Networks. **Computational Economics**. Springer. 2006.p.71-88.

BEUREN, I. K. Como elaborar trabalhos monográficos em contabilidade. 3ed. São Paulo: Atlas, 2006, 195p.

BERNSTEIN, Peter L.; Desafio aos deuses: a fascinante história do risco, Tradução de: Ivo Korylowski, 23^a ed. Rio de Janeiro: Elsevier 1997.

BOLLERSLEV, T. A conditional heteroscedasticity time series model for speculative process and rates of return. **Review of Economics and Statistics**, v.69, p. 542-547, 1987.

BOLLERSLEV, T. Generalized autoregressive conditional heteroscedasticity. **Journal of Econometrics**. v.31, p.307-327, 1986.

BOLLERSLEV, T; CHOU, R.; KRONER, K. ARCH modeling in finance: a review of the theory and empirical evidence. **Journal of Econometrics**. v.52, p.5-59, 1992.

BOX, G. E. P., JENKINS, G. M., REINSEL, G. C. Time series analysis: forecasting and control. 3. ed. New York: Prentice Hall, 1994.

BOX, G. E. P., JENKINS, G. M. Time series analysis: forecasting and control. 3 ed. New York: Prentice Hall, 1994.

BRAGANÇA, G.G.F.. Poder de Mercado do café brasileiro nos EUA: abordagem via demanda residual. Dissertação (Mestrado em Economia) – Fundação Getúlio Vargas. Rio de Janeiro, 2003.

BROCK, W.; DECHERT,W.;SCHEINKMAN, J.; LE BARON, B. A test for independence based on the correlation dimension. **Econometric Reviews**. V.15. 1996. P.197-235.

CALDEIRA, André Machado; SOUZA, Reinaldo Castro; MACHADO, Maria Augusta Soares. Identificação automática das ordens dos modelos GARCH utilizando redes neurais. **ENGEVISTA**. V.11, No.2. Dezembro-2009. p.82-89.

CAJUEIRO, Daniel O.; TABAK, Benjamin M. Long-range dependence and market structure. **Chaos Solitons & Fractals**. Elsevier. Outubro-2005. p.995-1000.

CAJUEIRO, Daniel O.; TABAK, Benjamin M. The Hurst exponent over time: testing the assertion that emerging markets are becoming more efficient. **Physica A**. Elsevier. Outubro-2003.p.521-537.

CARDOSO, Mayra Moutinho. Simulação de modelos Garch para séries temporais univariadas de demanda de energia elétrica para consumidores livres em regime de curto prazo. Dissertação Mestrado. Universidade Federal de Itajubá. Itajubá, 2005.

CAPORIN, Massimiliano; McALLER, Michael. Dynamic Asymmetric GARCH. **Journal of Financial Econometrics**. Vol.4, No.3. 2006.p.385-412.

CHARLES, Amélie. Forecasting Volatility with Outliers in GARCH Models. **Journal of Forecasting**. Setembro-2008. p.551-565.

CORRAR, Luiz J. et al, Análise Multivariada para os cursos de Administração, Ciências Contábeis e Economia, São Paulo: Editora Atlas 2007.

DAMODARAN, Aswath. Avaliação de Investimentos. 7. Reimpressão São Paulo : Qualitymark Editora 2005.

DE OLIVEIRA, M. A. Previsão de Sucessões Cronológicas econômico-financeiras por meio de redes neurais artificiais recorrentes de tempos real e de processos ARMA GARCH: um estudo comparativo quanto a eficiência de previsão. Dissertação Mestrado. FEA – Faculdade de Economia Administração e Contabilidade, São Paulo, 2003.

DICKEY, D. A.; FULLER,W.A. Distribution of the estimators for autoregressive time series with a unit root. **Journal of the American Statistical Association**. V.74,p.427-431, 1979.

DINIZ, H. A. L.; CARVALHO, A.; ANDRADE, M. Previsão de séries temporais utilizando redes neurais artificiais e modelos de Box e Jenkins. Procedings of the **V Brazilian Symposium on Neural Networks**, Belo Horizonte: p. 173-178, dez. 1998.

DONOHO, David L. JOHNSTONE, Iain M. Minimax estimation via wavelet shrinkage. **Annals of statistics**. Vol. 26, number 3, 1998, 879-921.

EOM, Cheoljun; CHOI, Sunghoon; OH, Gabjin; JUNG, Woo-Sung. Hurst exponent and prediction based on weak-form efficient market hypothesis of stock markets. **Physica A**. Elsevier. Abril-2008. p.4630-4636.

ENDERS, W. Applied econometric time series. 2. ed. John Wiley & Sons, 2003.

FENG, Yuanhua. Simultaneously modeling conditional heteroskedasticity and scale change. **Econometric Theory**. Vol.20. Estados Unidos, 2004.p.563-596.

FRANSES, P. H. GHYSELS, E. Additive outliers, GARCH and forecasting volatility; **International journal of forecasting**. Vol. 15. 1999, p.1-9.

GALDI, Fernando Caio; PEREIRA, Leonel Molero. Valor em Risco (VaR) utilizando modelos de previsão de volatilidade: EWMA, GARCH e Volatilidade Estocástica. **Brazilian Business Review**. Vol.4, No.1. Jan/Ab-2007. Vitória-ES, Brasil. 2007. p.74-95.

GOOIJER, J. G.; HYNDMAN, R. J. 25 years of time series forecasting, **International Journal of Forecasting**, vol. 22, 2006.

GOSSE, Laurent. Analysis and short-time extrapolation of stock market indexes through projection onto discrete wavelet subspaces. **Nonlinear Analysis: Real World Applications**. Elsevier, Itália, 2009. p. 3139-3154.

GRANÉ, Aurea; VEIGA, Helena. Wavelet-based detection of outliers in financial time series. **Computational Statistics and Data Analysis**. Elsevier. Janeiro-2010.p.2580-2593.

GRIMMENT, G. R., STIRZAKER D. R. Probability and Random Process. **Oxford**. University Press: Nova York, 1982.

HANSELMAN, D; LITTLEFIELD, B. MatLab 6 – Curso Completo. São Paulo: Prentice Hall, 2003. 676 p.

HAYKIN, S. Redes Neurais: princípios e prática.2.ed. Porto Alegre. Bookman, 2001.

HENDRIKSEN, Eldon S.; VAN BREDA, Michael F. Teoria da Contabilidade. Tradução por Antonio Zoratto Sanvicente. 5^a ed. São Paulo: Atlas, 1999.

HISIEH, D. Testing for nonlinear dependence in daily foreign Exchange rates. **Journal of Business**. V.62, n.3, 1989.p.339-368.

HOMSY, G. V.; PORTUGAL, M. S.; ARAÚJO, J. P.. Ondaletas e Previsão de Séries de Tempo: uma análise empírica. **XXII Encontro Brasileiro de Econometria**. Campinas, 2000.

HONG, Yongmiao; LEE, Jin. One-sided testing for ARCH effects using wavelets. **Econometric Theory**. Cambridge University Press. 2001.p.1051-1081.

HUANG, Shian-Chang; WU, Tung-Kuang. Combining wavelet-based feature extractions with relevance vector machines for stock index forecasting. In: **Expert Systems**. Vol. 25 No. 2, Taiwan, Maio-2008. p. 133-149.

HURST, E. Long term storage capacity of reservoirs. **Transactions of the American society of civil engineers**. V.116, p.770-799, 1951.

IUDÍCIBUS, Sergio de; LOPES, Alexsandro Broedel, Teoria Avançada da Contabilidade. São Paulo, Editora Atlas 2008.

JACOBSEN, B. Long Term Dependence in Stock Returns. **Journal of empirical Finance**, p. 393-417 Amsterdam, 1996.

JARQUE, C.;BERA A. A test for normality of observations and regression residuals. **International statistical review**, v.55, p.163-172, 1987.

KAMITSUJI, Shigeo; SHIBATA, Ritei. Effectiveness of Stochastic Neural Network for Prediction of Fall or Rise of TOPIX. **Asia-Pacific Financial Markets**. Springer. 2005.p.187-204.

KARLAFTIS, Matthew G.; VLAHOGIANNI, Eleni I. Memory properties and fractional integration in transportation time-series. **Transportation Research Part C**. Elsevier. Março-2009.p.444-453.

KYAW, Nyonyo A.; LOS, Cornelis A.; ZONG, Sijing. Persistence Characteristics of Latin American financial markets. **Journal Multinational Financial Management**. Elsevier. Agosto-2005.p.269-290.

KLEIN, J.L. Statistical visions in time: a history of time series analysis, 1662-1938. New York: **Cambridge University Press**, 1997.

LAMOUNIER, W. M. Comportamento dos preços no Mercado spot de café no Brasil: Análise dos domínios do tempo e da freqüência. BM&F. São Paulo, 2003. 125p.

LEAL, Ricardo P.C., SAITO, Richard. Finanças Corporativas no Brasil. RAE-eletrônica, v. 2, n. 2, jul-dez/2003.

LIMA, Fabiano Guasti. Um método de análise e previsão de sucessões cornológicas unidimensionais lineares e não lineares. Tese Doutorado. Universidade de São Paulo. Ribeirão Preto, 2004.

LIMA, Fabiano Guasti. Modelos de previsão de séries temporais financeiras com combinação de filtros de kalman e wavelets. Tese Livre Docência. FEARP/USP. Ribeirão Preto, 2011.

LIN, Xiaowei; YANG, Zehong; SONG, Yixu. The Application of Echo State Network in Stock Data Mining. **Verlag Berlin Heidelberg**. Springer. Berlin. 2008. p.932-937.

LOS, Cornelis A.; YU, Bing. Persistence Characteristics of the Chinese stock markets. **IRFA International Review of Financial Analysis**. Elsevier. 2008.p.64-82.

LOPES, Alexsandro Broedel, SANT ANNA, Dimitri Pinheiro, COSTA, Fábio Morais. A Relevância das Informações contábeis na BOVESPA a partir do arcabouço teórico de Ohlson: avaliação dos modelos de Residual Income Valuation e Abnormal Earnings Growth. **Revista de Administração USP (RAUSP)**, vol.42, n.4 pp.497-510, 4º.Trimestre 2007.

LU, Chi-Jie. Integrating independent component analysis-based denoising scheme with neural network for stock price prediction. **Expert Systems with Applications**. Elsevier. Março-2010.p.7056-7064.

_____. A Informação Contábil e o Mercado de Capitais. São Paulo: Pioneira **Thompson Learning**, 2002.

MAKRIDAKIS, Spyros G.; WHEELWRIGHT, Steven C.; HYNDMAN, Rob J. **Forecasting: methods and applications**. 3ed. New York: John Wiley & Sons, 1998.

MANDELBROT, B. B.; WALLIS, J. R. Robustness of the rescaled range R/S in the measurement of noncyclic log run statistical dependence. **Water Resources Research**, v.5, p.967-988, 1969.

MANIMARAN, P.; PANIGRAHI, Prasanta K.; PARIKN, jitendra C. Difference in nature of correlation between NASDAQ and BSE indices. **Physica A**. Elsevier. Junho-2008. p.5810-5817.

MATIAS, M. A. Análise do Comportamento de Preços da Commodity Cobre: Uma abordagem sob a ótica da teoria dos fractais. Dissertação de Mestrado. Universidade de Brasília, UnB. Brasília. 2006. 94p.

MATHWORKS. Garch Toolbox User's Guide. Disponível em: www.mathworks.com/access/helpdesk/help/pdf_doc/garch/garch.pdf. Acesso em 20 de Novembro de 2009.

MARCEK, Dusan. Forecasts Generating for ARCH-GARCH Processes Using the MATLAB Procedures. **Institute of Computer Science**, Faculty of Philosophy and Science, The Silesian University Opava. 2000.

MARTINS, Caroline M. F. A volatilidade nos preços futuros do café brasileiro e seus principais elementos causadores. Dissertação (Mestrado em Administração). UFLA. Lavras-MG, 2005.

MENTZER, John T.; BIENSTOCK, Carol C. **Sales Forecasting Management**. Thousand Oaks: Sage, 1998.

MCLEOD,A.J.; LI W.K. Diagnostic cheking ARMA time series models using squared residuals correlations. **Journal of time series analysis**, n.4, 1983,p.269-273.

MILLER, M. B., Y MODIGLIANI, F. (1961): Dividend Policy, Growth, and the Valuation Shares, **Journal of Business**, vol. 34, n." 4, October 1961, pp. 41 1-433.

MISITI, M. et al. Ondaleta Toolbox: for use with MATLAB. The Math Works, Inc. 1997.

MODIGLIANI, Franco; MILLER, Merlon, "The Cost of Capital, Corporate Finance, and the Theory of Investment: Reply "**Américan Economics Review**, Vol 49, n.4, 1958 pp 655 –669.

MORETTIN, P. A. Ondas e Ondaletas. São Paulo: EDUSP, 1999.

MORETTIN, P. A. Séries temporais em finanças. São Paulo, 2002. 156p. (Texto para um curso na Universidade Católica de Lima, Peru).

MORETTIN, Pedro A., TOLOI, Clelia M. C. Análise de Séries Temporais. 1^a ed. São Paulo: Egard Blucher, 2004.

MORETTIN, Pedro A., TOLOI, Clelia M. C. Análise de Séries Temporais. 2^a ed. São Paulo: Egard Blucher, 2006.

MOTA, Bernardo de Sá; FERNANDES, Marcelo. Desempenho de Estimadores de Volatilidade na Bolsa de Valores de São Paulo. **Revista Brasileira de Economia**, v.58, n.3, 2004.

NAGHETTINI, Mauro; PINTO, Éber José de Andrade: Hidrologia Estatística, Executado pela CPRM – Serviço Geológico do Brasil, Superintendência Regional de Belo Horizonte, CPRM, 2007.

NGUYEN, Hang T.; NABBNEY, Ian T. Short-term electricity demand and gas price forecasts using wavelet transforms and adaptive models. **Energy**. Elsevier. Junho-2010.p.3674-3685.

_____ ; LOPES, Aleksandro Broedel. Avaliação de Empresas com Base em Números Contábeis. **Brazilian Business Review** Vol. 4, n.2. p.96-103, 2007.

PELLEGRINI, Fernando R.; FOGLIATTO, Flávio, S.; Passos para implementação de Sistemas de previsão de demanda – **Técnicas e Estudos de Caso, Revista Produção**, v.11, n.1, Novembro de 2001.

PEÑA, D.; TIAO, G. C.; TSAY, R. S. A Course in Time Series Analysis, New York: John Wiley and Sons, 2001.

PINDYCK, Robert S., RUBINFELD, Daniel L.; Econometria, Modelos & Previsões. 4a. ed., Rio de Janeiro, 2004.

PINDYCK, Robert S., RUBINFELD, Daniel L. Econometric models and economic forecasting. 4.ed. New York: McGraw-Hill International, 1997.

PINDYCK, Robert S., RUBINFELD, Daniel L. Econometria. _____ ed. Rio de Janeiro: Elsevier, 2004.

POWER, Gabriel J.; TURVEY, Calum G. Long-range dependence in the volatility of commodity futures prices: Wavelet-based evidence. **Physica A**. Elsevier. Setembro-2009. p.79-90.

RIBEIRO, Luiz Cláudio; PAULA, Anaparecida Vieira. Previsão de população através dos modelos de ARIMA de Box e Jenkins. In: **Encontro Nacional de Estudos Populacionais, Universidade Federal de Minas Gerais**, Caxambu, 2000. Disponível em: www.abep.nepo.unicamp.br/anais_2000.htm.

SALCEDO-SANZ, Sancho; PÉREZ-BELLIDO, Ángel M; ORTIZ-GARCÍA, Emilio G.; PORTILLA-FIGUERAS, Antonio; PRIETO, Luis; PAREDES, Daniel. Hybridizing the fifth generation mesoscale model with artificial neural networks for short-term Wind speed prediction. **Renewable Energy**. Elsevier. Novembro-2008. p.1451-1457.

SELVARATNAM, Somesh; KIRLEY, Michael. Predicting Stock Market Time Series Using Evolutionary Artificial Neural Networks with Hurst Exponent Input Windows. **Verlag Berlin Heidelberg**. Springer. Berlin. 2006.p.617-626.

SOUZA, Sergio Rubens Stancato de; TABAK, Benjamin Miranda; CAJUEIRO, Daniel O. Investigação de Memória de Longo Prazo da Taxa de Câmbio do Brasil. **Departamento de Estudos e Pesquisas do Banco Central do Brasil**, Trabalhos para discussão. Agosto-2006. p.37.

STENGEL, Robert F. Stochastic Optimal Control: Theory and Application. 1ed. New York: John Wiley & Sons Inc, 1986.

ROSS, Stephen A., WESTERFIELD, Randolph W., JAFFE, Jeffrey F. Administração Financeira. 2.ed – 7.reimp – São Paulo. Atlas, 2008.

TABAK, Benjamin M.; CAJUEIRO, Daniel O. The long-range dependence behavior of the term structure of interest rates in Japan. **Physica A**. Elsevier. Dezembro-2004.p.418-426.

TAN, Zhongfu; ZHANG, Jinliang; WANG, Jianhui; XU, Jun. Day-ahead electricity price forecasting using wavelet transform combined with ARIMA and GARCH Models. **Applied Energy**. Elsevier. Maio-2010. p.3606-3010.

TITTERINGTON, D. M. Bayesian Methods for Neural Networks and Related Models. **Statistical Science**. Vol. 19, No.1, Fev-2004. p.128-139.

TSAY, R. Analysis of financial time series. Wiley series in probability and statistics. 2.ed, 2005.

TSAY, R. S. Analysis of financial time series. Wiley series in probability and statistics, 2002.

TUBINO, Dalvio, F. Manual de Planejamento e Controle da Produção. 2ed. São Paulo: Atlas, 2000.

VASCONCELLOS, M. A. S., ALVES, D. Manual de econometria. São Paulo: Atlas, 2000.

VISSEUR, Marcel P. GARCH Parameter Estimation Using High-Frequency Data. **Journal of Financial Econometrics**. Junho-2010.p.1-36.

XU, Jianhua; CHEN, Yaning; LI, Weihong; DONG, Shan. Long-Term Trend and Fractal of Annual Runoff Process in Mainstream of Tarim River. Chin. Geogra. Sci. Shanghai Academic Discipline Project. Shanghai, China-2007.p.77-84.

XU, Xiaoyan; QI, Yuqing; HUA, Zhongsheng. Forecasting demand of commodities after natural disasters. **Expert Systems with Applications**. Elsevier. 2009.p.4313-4317.

ZUNINO, L.; TABAK, Benjamin M.; PÉREZ, D.G.; GARAVAGLIA, M.; ROSSO, O. A. Inefficiency in Latin-American market índices. **The European Physical Journal B.** Springer-Verlag. 2007.p.111-121.

YIM, Juliana. A Comparison of Neural Networks with Time Series Models for Forecasting Returns on a Stock Market Index. **Springer-Verlag Berlin Heidelberg.** Berlin, 2002. p. 25-35.

WANG, Jui-Lin; KUO, Ju-Hsiu; CHOU, Shih-Yu; WANG, Sheng-Zhi. A comparison of bullwhip effect in a single-stage supply chain for autocorrelated demands when using correct, MA, and EWMA methods. **Expert Systems with Applications.** Elsevier. Setembro-2009.p.4726-4736.

WANG, Peng; VACHTSEVANOS, George. Fault Prognostics using dynamic wavelet neural networks. Artificial Intelligence for Engineering Design, Analysis and Manufacturing. **Cambridge University Press.** Estados Unidos, 2001. p.349-365.

WANG, Kehluh; HUANG, Szuwei. Using fast adaptive neural network classifier for mutual fund performance evaluation. **Expert Systems with Applications.** Elsevier. Fevereiro-2010.p.6007-6011.

WERNER, Liane; RIBEIRO, José Luis Duarte. Previsão de demanda: uma aplicação dos modelos Box-Jenkins na área de assistência técnica de computadores pessoais. Gestão e Produção. V.10, No.1, Abril-2003.p.47-67.

WONG, H.; IP, W.;XIE,Z.; LUI, X. Modelling and forecasting by wavelets and the application to exchange rates. **Journal of Applied Statistics**, v.30, n.5, 2003. P537-553.