

PROPRIEDADES ESTATÍSTICAS DAS SÉRIES DE RETORNOS DAS PRINCIPAIS AÇÕES BRASILEIRAS

Paulo Henrique Soto Costa

Tara Keshar Nanda Baidya

Departamento de Engenharia Industrial

PUC-Rio

Resumo

O artigo analisa seis séries de retornos, escolhidas entre as mais líquidas do mercado e de setores diferentes da economia. São estudadas a estacionariedade, a distribuição incondicional e a independência; de maneira geral as séries podem ser consideradas estacionárias, com distribuição não normal (leptocúrtica) e dependentes no tempo. A estacionariedade é analisada através do teste ADF, através dos coeficientes de modelos GARCH ajustados aos dados, pelos coeficientes de bicorrelação e com o uso de regressão localmente ponderada. A normalidade é rejeitada pelo teste de Jarque e Bera. A dependência (linear e não linear) é constatada pelas autocorrelações dos retornos e dos seus quadrados: tenta-se modelar a dependência com modelos ARMA, de amortecimento exponencial e GARCH mas os resíduos dos modelos, testados pelo teste BDS, mostram que nenhum deles representa bem o processo gerador dos dados.

Palavras-chave: séries financeiras, estacionariedade, independência.

Abstract

This paper studies the returns of six Brazilian stocks, chosen among the most liquid and from different economic sectors. Stationarity, unconditional distribution and independence are studied, and the conclusion is that the series are stationary, have non-normal (leptokurtic) distribution and are dependent. Stationarity is studied with the help of the ADF test, of the coefficients of GARCH models adjusted to the data, of bicorrelation coefficients and of locally weighted regression. Normality was rejected using the Jarque and Bera test. Dependence (linear and non-linear) was indicated by the autocorrelations of returns and squared returns: ARMA, exponential smoothing, and GARCH models are used to try to capture the dependence, but the BDS tests of the model's residuals show that they do not represent well the data generating process.

Keywords: financial series, stationarity, independence.

1. Introdução

A caracterização das propriedades estatísticas das séries de retornos das ações é importante para permitir a correta aplicação de modelos aos dados, que permitam inferir a respeito das características dos retornos, sobretudo no que diz respeito a média e variância, que vão determinar o retorno esperado e a previsão da volatilidade para os próximos períodos. Esta determinação é fundamental para a decisão de investimento não só na ação como também em seus derivativos.

Este trabalho analisa estacionariedade, distribuição e independência em seis séries de retornos de ações: Telebrás, Petrobrás, Eletrobrás, Bradesco, Vale do Rio Doce e Brahma, todas preferenciais. A escolha foi feita entre as mais líquidas do mercado, de setores diferentes da economia. Todas as séries têm 1200 observações de retornos diários, num período de aproximadamente cinco anos, de julho de 1994 a julho de 1999.

As séries temporais são analisadas quanto a estacionariedade. Inicialmente ela é estudada supondo os dados gerados por processo linear (teste ADF) ou por processo GARCH, e se conclui pela estacionariedade. A seguir é estudada a estacionariedade da média sem premissas a respeito do processo gerador dos dados, através do coeficiente de bicorrelação (que indica que as séries estudadas são estacionárias) e através de regressão localmente ponderada (que indica estacionariedade em quatro séries e é inconclusiva para as outras duas). Pelo conjunto dos testes, conclui-se que as séries são estacionárias.

A análise da distribuição incondicional dos retornos é bastante simples [teste de Jarque & Bera (1987)] pois se restringe à constatação da evidente não normalidade dos retornos diários, que apresentam leptocurtose. Uma possível explicação para a não normalidade seria distribuição não independente e idêntica dos retornos diários, e isto motiva o restante do trabalho.

A independência entre os retornos consecutivos é estudada inicialmente através da função de autocorrelação, que indica existência de dependência (linear) nas séries. Ela é modelada através de um ARMA(1,1) mas os resíduos continuam apresentando correlação, o que indica uma possível dependência também não linear, o que é confirmado pelo fato de os quadrados dos resíduos serem fortemente correlacionados.

Constatado que os retornos não são IID, são aplicados três modelos aos dados, para tentar explicar o tipo de dependência existente: amortecimento exponencial do desvio padrão condicional, GARCH(1,1) puro, e GARCH(1,1) com modelagem ARMA(1,1) da média condicional. O poder explicativo dos modelos é estudado através dos resíduos, que devem ser IID normal(0,1).

Os resíduos dos três modelos são não normais (são leptocúrticos). Eles são analisados quanto a autocorrelação (também dos resíduos quadrados) que em geral indica alguma dependência, e é aplicado o teste BDS aos resíduos para verificar distribuição independente e idêntica; é feita a correção dos valores críticos da estatística BDS para os resíduos GARCH e ARMA-GARCH. Conclui-se que nenhum dos três modelos produz resíduos IID, e que o modelo que melhor explica os dados é o ARMA-GARCH.

A maior parte das análises foi feita com o auxílio do programa EVIEWS 3.0; os testes BDS empregaram programa específico do Departamento de Economia da Universidade de Houston, EUA, (copyright 1987, 1988, por W.D. Dechert); as análises de bicorrelação e a regressão localmente ponderada foram feitas em programas desenvolvidos pelos autores, que se colocam à disposição para fornecer maiores informações sobre suas características.

2. Referências

O estudo das propriedades estatísticas das séries financeiras inicia-se com Bachelier (1900), em um trabalho notável que implicava em que os retornos tivessem distribuição normal independente e idêntica. Alguns outros trabalhos se seguiram, entre os quais se pode citar o de Working (1934) que estudou séries de preços de futuros de *commodities*, e o de Kendall (1953) que estudou séries de preços de ações e de produtos agrícolas. No entanto, foi com Fama (1965) que se iniciou a publicação de um grande número de trabalhos sobre a estatística das séries financeiras; ele estudou em detalhe as séries de preços das 30 ações do *Dow Jones Industrial Average Index*, do mercado de ações norte-americano, num longo e fundamental artigo onde ele conclui:

- que as séries têm distribuições com caudas pesadas incompatíveis com as da normal.
- que as distribuições das séries podem ser representadas por distribuições estáveis paretianas com expoente característico menor que dois.
- que a dependência entre os retornos, medida pelos coeficientes de autocorrelação e pelo número e comprimento dos *runs* é muito pequena para ter significado econômico, e que não é possível obter ganhos continuados usando a técnica de “filtro” proposta por Alexander (1961).
- que o processo estocástico dos preços se aproxima do modelo de “passeio aleatório”, não sendo portanto possível usar preços passados para prever preços futuros.

Nos muitos trabalhos que se seguiram, embora a constatação da não normalidade dos retornos seja amplamente aceita, o mesmo não acontece com as distribuições estáveis paretianas.

É interessante observar que o uso destas duas distribuições para os retornos têm uma justificativa econômica, pois o retorno (logarítmico) em um dado período de tempo é a soma dos retornos verificados em períodos menores, ou mesmo transação a transação, e tanto a normal como as distribuições estáveis paretianas se preservam quando se somam variáveis aleatórias. Ao contrário, supor distribuição t (por exemplo) para os retornos diários, significa supor também que os retornos horários ou semanais têm outras distribuições diferentes da t .

A explicação mais aceita atualmente para as caudas pesadas das distribuições de retornos é a de mistura de distribuições; muitos autores, como Praetz (1972) e Clark (1973) argumentam que esta mistura é de distribuições normais. A idéia da mistura de normais deu origem a muitos modelos não lineares para as séries de retornos e os mais difundidos nesta linha são os da família GARCH, introduzida por Engel (1982) e generalizada por Bollerslev (1986), onde fica caracterizada uma dependência não linear entre os retornos, função da dependência serial da variância condicional.

O modelo do passeio aleatório também foi muito estudado, dando origem a grande número de artigos, entre os quais podem-se destacar Granger & Morgenstern (1970), Cunningham (1973), e Cooper (1982), entre muitos outros voltados para retornos de ações. De maneira geral os autores concordam com a validade deste modelo (ou de variações dele), pelo menos no que diz respeito à imprevisibilidade dos retornos. Taylor (1986), por exemplo, testa a hipótese de passeio aleatório e a previsibilidade de 40 séries de preços, concluindo que os preços de ações seguem o passeio aleatório, mas encontra alguma previsibilidade em preços de futuros de *commodities*.

Mais recentemente, aceita a impossibilidade de obter lucro usando séries de retornos para prever retornos, a ênfase dos trabalhos com séries temporais financeiras deslocou-se para a modelagem da variância: os retornos esperados são previstos através de outros tipos de modelos (análise de fatores, por exemplo) e apenas o risco vem das séries temporais de retornos.

Nesta linha, os modelos GARCH foram expandidos e generalizados em várias direções: por exemplo, o GARCH exponencial [de Nelson (1991)], o de limiar [Zakoian (1991)], a generalização de Hentschel (1995) que aninha vários modelos em uma só especificação, e os modelos com parâmetros variáveis no tempo de acordo com processos de Markov [Hamilton & Susmel (1994) e Gray (1996)]. Outras técnicas também foram empregadas, como por exemplo os modelos de volatilidade estocástica [ver Hsieh (1993), que emprega modelo baseado na volatilidade de Parkinson (1980), ou Kim, Shephard & Chib (1998) que comparam volatilidade estocástica e modelos da família GARCH] e também os modelos empregando distribuição de valores extremos [ver por exemplo Embrechts, Resnick & Samorodnitsky (1998) ou McNeil (1998)].

Encontram-se também, nos periódicos especializados, diversos trabalhos sobre a modelagem de séries temporais financeiras em países emergentes e, em particular no Brasil. Entre os trabalhos sobre modelos de volatilidade em países emergentes, pode-se citar Bekaert & Harvey (1997), Duarte & Mendes (1997), Campbell (1995) e Susmel (1998). Séries brasileiras foram estudadas, entre outros, por Duarte, Heil & Pinheiro (1996), que comparam modelos de volatilidade de preço e volatilidade implícita; por Herencia, Hotta & Pereira (1998) que também comparam modelos de volatilidade, usando séries reais e simuladas; por Issler (1999), que compara os resultados obtidos usando modelos da família GARCH em séries de retornos de ativos de diferentes tipos (ação, bônus, moeda e *commodity*); por Pereira, Hotta, Souza & Almeida (1999), que comparam modelos da família GARCH, volatilidade estocástica e GARCH com parâmetros variantes no tempo, usando três das quatro séries usadas por Issler (1999).

Este trabalho pretende estudar as séries de retornos de ações brasileiras, no que diz respeito a suas propriedades estatísticas em geral, não tendo o objetivo de se aprofundar na modelagem de estruturas que possam existir nos dados: como foi detectada dependência linear e não linear nas séries, são aplicados modelos simples (ARMA, amortecimento exponencial e GARCH) objetivando verificar, através da análise dos resíduos, se a dependência é capturada por estes modelos ou se é necessário, em estudo posterior, o uso de modelos mais elaborados.

Também foge ao objetivo deste estudo o uso de técnicas de análise espectral nas séries [vide, por exemplo, Koopmans (1974) e Ventura (2000)] que poderia, inclusive, ser útil na análise da estacionariedade; o emprego destas técnicas fica como sugestão para possíveis continuações deste trabalho.

Os dados cobrem período recente, que se inicia logo após o “Plano Real”, onde a baixa taxa de inflação dispensa a necessidade de usar deflatores que poderiam introduzir algum ruído nas séries. Na seção 3, a seguir, são apresentadas as séries e suas características mais gerais; na seção 4 estuda-se a estacionariedade; na seção 5 é analisada a distribuição incondicional, e em 6 e 7 a independência dos retornos, concluindo-se na seção 8.

3. As Séries

Os preços das ações da Telebrás estão mostrados, para ilustração, na Figura 1. As figuras com os preços e retornos das outras ações não estão apresentadas.



Figura 1 – Preços das ações da Telebrás

Os retornos foram calculados através do logaritmo neperiano do quociente entre os preços de fechamento em dois dias consecutivos. Na Figura 2 estão os retornos das ações da Telebrás.

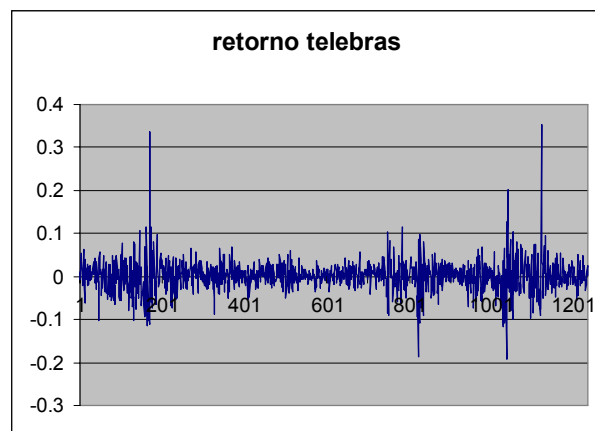


Figura 2 – Retornos das ações da Telebrás

Foi ajustada na série de preços uma função exponencial, que ilustra a existência de uma tendência de crescimento do preço no longo prazo. Esta tendência faz com que o retorno médio diário seja positivo; para uniformização e facilidade de comparação entre as séries elas foram padronizadas, ou seja, de cada observação foi subtraída a média, e o resultado foi dividido pelo desvio padrão, resultando séries de média zero e variância um. Todas as análises deste trabalho se referem às séries de retornos padronizados das ações. Na Figura 3 estão os retornos padronizados das ações da Telebrás.

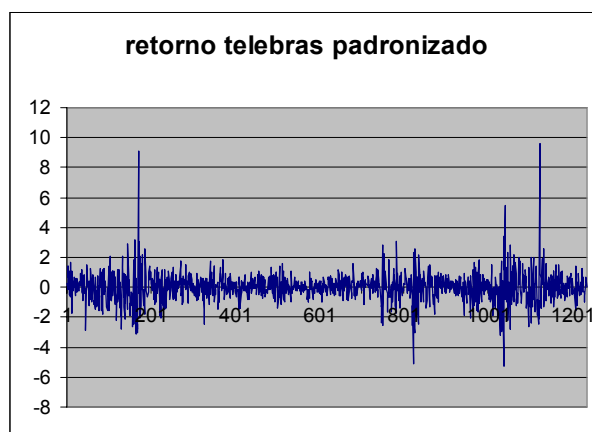


Figura 3 – Retornos padronizados das ações da Telebrás

Observando a série dos retornos padronizados da Telebrás, nota-se que existem algumas observações aberrantes, com valores da ordem de até dez desvios padrão. Estes valores extremos de retornos ocorreram em dias atípicos; em função disto foi formada uma nova série de retornos padronizados para cada ação, da qual foram retirados cinco valores de retorno, correspondentes aos dias 10/3/95 (crise do México), 27/10/97 (crise da Ásia), 10/9/98 e 15/9/98 (crise da Rússia), e 15/1/99 (crise do Real) e foram acrescentados mais cinco dias no final. Desta forma, há duas séries de retornos padronizados para cada ação: dados *originais* e dados *modificados* (sem os cinco dias de crise).

As séries de retornos padronizados receberam siglas, conforme a Tabela 1.

Tabela 1 – Siglas das séries de retornos padronizados

AÇÃO	DADOS ORIGINAIS	DADOS MODIFICADOS
TELEBRÁS	OTELB	MTELB
PETROBRÁS	OPETR	MPETR
ELETOBRÁS	OELET	MELET
BRADESCO	OBBDC	MBBDC
VALE DO RIO DOCE	OVALE	MVALE
BRAHMA	OBRHA	MBRHA

4. Estacionariedade

4.1 Introdução

Um processo estocástico representado por uma seqüência de variáveis aleatórias $\{... X_1, X_2, ..., X_n, ...\}$ é estritamente estacionário se, para todos os inteiros i, j e para todos os inteiros positivos k , a distribuição multivariada de $(X_i, X_{i+1}, ..., X_{i+k})$ é idêntica à de $(X_j, X_{j+1}, ..., X_{j+k})$. Na prática só é possível testar algumas das conseqüências da premissa de que um processo é estritamente estacionário; em particular, se as distribuições multivariadas são idênticas, então $E(X_t)$ é igual a alguma constante μ e $cov(X_t, X_{t+\tau})$ é igual a uma constante λ_τ que depende apenas de τ (supondo, evidentemente, que as distribuições têm

média e covariância finitas). Um processo que atenda a estas duas condições é dito *estacionário de segunda ordem*, que será aqui abreviado por *estacionário*.

Neste trabalho foi estudada a estacionariedade das séries inicialmente dentro da suposição de linearidade, com o teste ADF, e depois dentro da suposição que os dados são gerados por processo GARCH(1,1); a seguir foi estudada a estacionariedade da média incondicional através de testes com a média condicional, sem suposições a respeito do processo dos dados. A conclusão é que, de maneira geral, pode-se considerar que as séries são estacionárias.

4.2 Teste ADF

O primeiro teste de estacionariedade aqui usado é o de Dickey-Fuller aumentado (ADF). Um processo linear auto-regressivo de ordem p , $AR(p)$, pode ser escrito como

$$A(L)y_t = \mu + \varepsilon_t$$

Onde $A(L)$ é um polinômio de grau p no operador L , de *lag*:

$$A(L) = 1 - \alpha_1 L - \alpha_2 L^2 - \dots - \alpha_p L^p$$

O processo é estacionário se $A(L)$ tiver todas suas raízes fora do círculo unitário; se houver raiz dentro do círculo unitário, o processo é explosivo e evidentemente não estacionário. No entanto, se houver raiz sobre o círculo, ele é não estacionário mas não é explosivo. No caso do $AR(1)$ a raiz unitária corresponde ao processo

$$y_t = \mu + y_{t-1} + \varepsilon_t$$

que é um “passeio aleatório” com tendência determinística μ .

O teste ADF tem como hipótese nula a presença de raiz unitária no processo $AR(p)$.

Neste trabalho considerou-se $p=4$ e os resultados obtidos levam à rejeição da hipótese da raiz unitária para todas as séries, o que é consistente com a estacionariedade. Para um nível de 5% e 1200 observações, o valor crítico da estatística do teste para a rejeição da hipótese de existência de raiz unitária é 2,86 (em valor absoluto). Na Tabela 2 estão os valores da estatística obtidos para as doze séries.

Tabela 2 – Teste de Dickey-Fuller aumentado

	ORIGINAL	MODIFICADA
TELB	-17,73	-17,18
PETR	-17,35	-16,27
ELET	-18,60	-17,94
BBDC	-17,03	-16,48
VALE	-16,87	-17,56
BRHA	-16,51	-16,29

4.3 Modelagem GARCH

Como já foi dito anteriormente, a verificação da estacionariedade de uma série é sempre feita dentro de hipóteses que limitam a generalidade do problema; um processo estocástico GARCH(1,1), é estacionário se $(\alpha + \beta) < 1$, com α e β sendo os coeficientes do quadrado da inovação e da variância no período anterior, respectivamente, na equação da variância do

modelo. Este modelo será novamente abordado na seção 7; ele é muito empregado para representar séries de retornos e foi aqui usado como um possível processo gerador dos retornos das ações, para efeito de verificação de estacionariedade.

Foi então ajustado um modelo GARCH(1,1) a cada uma das séries e os coeficientes α e β foram sempre significativos, num nível de 5%. Os resultados estão na Tabela 3.

Tabela 3 – Parâmetros do GARCH(1,1)

	α	β	$\alpha+\beta$
otelb	0,159	0,816	0,975
mtelb	0,126	0,850	0,977
opetr	0,076	0,915	0,990
mpetr	0,066	0,926	0,992
oelet	0,106	0,881	0,987
melet	0,096	0,893	0,988
obbdc	0,420	0,497	0,917
mbbdc	0,361	0,560	0,921
ovale	0,225	0,732	0,957
mvale	0,207	0,750	0,957
obrha	0,070	0,911	0,981
mbrha	0,071	0,910	0,981

Embora as somas dos coeficientes sejam elevadas, indicando alto grau de persistência, os valores de $(\alpha+\beta)$ são todos menores que um, ou seja, todas as séries podem ser modeladas por GARCH(1,1) estacionário. Para algumas séries o valor de $\alpha+\beta$ está bastante próximo da unidade, sugerindo que o modelo apropriado poderia ser o IGARCH(1,1), que é o modelo GARCH(1,1) com $\alpha+\beta=1$. Este modelo também é estacionário, embora não tenha variância incondicional.

Concluindo, supondo-se que os dados de cada série são gerados por processo GARCH(1,1), este processo é estacionário, mas não necessariamente tem variância finita.

4.4 Média Condicional

Uma outra maneira de abordar a estacionariedade é através da média condicional:

$$\mu_t = E(y_t | y_{t-1}, y_{t-2}, \dots).$$

Se a média condicional não depende do tempo ($\mu_t = \text{constante}$) então a média incondicional é igual à condicional e não depende do tempo, logo há estacionariedade na média; como a média amostral de todas as séries é zero, se a média condicional for constante ela será também nula, portanto será testado $\mu_t = 0$.

4.4.1 Teste de Bicorrelação

Hsieh (1989) propõe um teste para a hipótese nula de média condicional igual a zero, que usa o fato de que, se a média condicional de y_t é zero, então os coeficientes de bicorrelação dados

por $E(y_t \cdot y_{t-i} \cdot y_{t-j}) / \text{var}(y_t)^{3/2}$ são zero para $i, j \geq 1$. Nas Tabelas 4 e 5 estão, para quinze pares i, j , os valores dos coeficientes de bicorrelação para as seis séries originais e para as seis séries modificadas, e também as estatísticas do teste de significância de cada coeficiente isoladamente e dos quinze em conjunto.

Para um nível de 5%, o valor crítico da estatística para cada coeficiente isoladamente é igual a 1,96 e para os quinze coeficientes em conjunto é 25,0 e, nas tabelas, os valores significantes estão marcados com um asterisco.

Na Tabela 4 observa-se que para nenhuma série houve significância dos 15 coeficientes em conjunto, OBBDC apresenta três coeficientes significantes e OVALE apresenta um, podendo-se dizer que o teste de maneira geral não rejeita a hipótese de média condicional zero para as séries originais.

Na Tabela 5 estão os mesmos valores para as séries modificadas; embora os números tenham mudado um pouco, continua a não haver significância conjunta dos 15 coeficientes (embora MBBDC tenha se aproximado do valor crítico), e poucos coeficientes são individualmente significantes.

Concluindo, de maneira geral os coeficientes de bicorrelação não são significantemente diferentes de zero, o que é consistente com a média condicional nula e com a estacionariedade da média incondicional, para todas as séries.

Tabela 4 – Coeficientes de bicorrelação: séries originais

i, j		otelb	opetr	oelet	obbdc	ovale	obrha
1.1	coeficiente	0,28	0,20	0,33	0,67 (*)	-0,27	0,22
	teste	1,43	0,99	1,80	2,47	-1,02	1,11
1.2	coeficiente	0,14	0,09	0,07	0,16	0,07	-0,02
	teste	1,52	0,73	0,55	0,87	0,86	-0,27
2.2	coeficiente	0,32	0,18	0,32	0,58 (*)	0,29(*)	0,33
	teste	1,24	0,99	1,10	2,04	2,15	1,48
1.3	coeficiente	-0,03	0,02	-0,03	-0,16	-0,12	-0,16
	teste	-0,25	0,19	-0,20	-1,41	-1,47	-1,10
2.3	coeficiente	-0,07	-0,09	-0,26	-0,04	-0,13	-0,01
	teste	-0,54	-0,72	-1,61	-0,28	-1,62	-0,06
3.3	coeficiente	0,12	0,12	0,16	0,08	0,24	0,32
	teste	0,56	0,67	1,00	0,40	0,88	1,31
1.4	coeficiente	0,05	0,02	-0,01	0,03	0,01	-0,12
	teste	0,62	0,17	-0,08	0,29	0,11	-1,23
2.4	coeficiente	-0,01	-0,01	-0,05	-0,08	-0,06	-0,01
	teste	-0,08	-0,08	-0,34	-0,91	-0,90	-0,13
3.4	coeficiente	0,04	0,01	0,08	0,06	0,08	-0,05
	teste	0,44	0,07	0,81	0,68	1,30	-0,34
4.4	coeficiente	-0,03	-0,06	-0,02	-0,05	-0,03	0,21
	teste	-0,14	-0,40	-0,11	-0,55	-0,23	1,41
1.5	coeficiente	-0,03	0,01	-0,10	-0,07	0,04	-0,10
	teste	-0,42	0,14	-1,26	-0,83	0,71	-1,39
2.5	coeficiente	0,08	0,08	-0,05	0,24 (*)	0,01	0,06
	teste	0,85	0,58	-0,34	2,44	0,15	0,60
3.5	coeficiente	0,02	0,14	0,02	0,00	-0,03	0,08
	teste	0,22	1,46	0,20	0,02	-0,33	1,03
4.5	coeficiente	-0,11	0,07	-0,04	-0,01	-0,01	0,07
	teste	-1,41	0,65	-0,50	-0,11	-0,13	1,02
5.5	coeficiente	-0,14	-0,04	-0,08	-0,18	0,09	0,02
	teste	-0,81	-0,22	-0,49	-1,21	0,94	0,15
teste 15 coefic.		10,72	6,65	11,42	23,07	16,07	14,47

Tabela 5 – Coeficientes de bicorrelação: séries modificadas

i,j		mtelb	mpetr	melet	mbbdc	mvale	mbrha
1.1	coeficiente	-0,06	-0,06	0,07	0,35 (*)	-0,02	0,32
	teste	-0,59	-0,47	0,55	2,33	-0,23	1,18
1.2	coeficiente	0,08	0,03	-0,08	0,02	0,01	-0,02
	teste	1,14	0,39	-0,94	0,28	0,18	-0,09
2.2	coeficiente	0,05	-0,05	-0,05	0,24	0,16	0,37
	teste	0,43	-0,39	-0,30	1,60	1,72	1,77
1.3	coeficiente	-0,02	0,03	-0,11	-0,04	0,03	-0,22
	teste	-0,39	0,34	-1,46	-0,46	0,62	-1,66
2.3	coeficiente	0,00	-0,06	-0,08	0,00	-0,04	0,14
	teste	0,07	-0,86	-0,96	0,03	-0,63	1,06
3.3	coeficiente	-0,05	0,07	-0,10	0,13	0,02	0,20
	teste	-0,59	0,57	-0,80	0,93	0,21	1,43
1.4	coeficiente	0,03	-0,08	-0,13 (*)	-0,09	-0,02	-0,07
	teste	0,44	-1,29	-2,04	-1,16	-0,30	-0,71
2.4	coeficiente	0,02	0,06	0,08	0,05	-0,05	0,03
	teste	0,25	0,78	0,68	0,72	-0,76	0,39
3.4	coeficiente	0,02	0,05	0,11	0,04	0,04	0,05
	teste	0,27	0,60	1,59	0,58	0,65	0,64
4.4	coeficiente	0,01	0,04	0,04	-0,04	-0,01	0,18
	teste	0,13	0,35	0,34	-0,35	-0,11	1,45
1.5	coeficiente	0,02	-0,01	-0,12	-0,19 (*)	0,06	-0,12
	teste	0,29	-0,16	-1,33	-2,77	1,12	-1,45
2.5	coeficiente	0,10	0,04	0,16	0,11	0,02	0,07
	teste	1,42	0,57	1,61	1,55	0,35	0,80
3.5	coeficiente	0,07	0,08	0,15	0,03	-0,12 (*)	0,04
	teste	0,91	1,21	1,59	0,46	-2,22	0,49
4.5	coeficiente	-0,10	-0,01	-0,04	-0,05	0,02	0,06
	teste	-1,55	-0,13	-0,49	-0,91	0,40	0,85
5.5	coeficiente	-0,11	-0,11	-0,08	-0,13	0,05	-0,01
	teste	-1,09	-1,07	-0,54	-1,29	0,54	-0,06
teste 15 coefic.		9,21	7,41	19,67	24,27	11,76	17,30

4.4.2 Regressão Localmente Ponderada

Segundo Pemberton & Tong (1981) existem modelos não lineares com coeficientes de bicorrelação nulos e médias condicionais diferentes de zero; por isto será feita uma outra abordagem à questão da média condicional, por meio de sua estimação direta por método não paramétrico. Se y_t é gerado por um modelo do tipo $y_t = g(y_{t-1}, y_{t-2}, \dots) + \varepsilon_t$ onde ε_t é IID e a função g é suficientemente bem comportada, Stone (1977) mostra que métodos de regressão não paramétricos podem ser usados para estimar a função g consistentemente.

O método da regressão localmente ponderada (RLP) foi desenvolvido por Cleveland & Devlin (1988) e foi usado em séries financeiras por Diebold & Nason (1990), Meese & Rose (1991) e Hsieh (1993). Hsieh (1991) usa simulação para mostrar que a RLP pode detectar todos os modelos de dinâmica não linear mais citados na literatura de séries temporais.

O método foi aqui empregado usando apenas um regressor (um *lag* de y_t), ou seja, foi suposta uma função de média condicional do tipo $y_t = a_{t-1} + b_{t-1} y_{t-1} + \varepsilon_t$; ele consiste em estimar os coeficientes da regressão, para cada valor de y_{t-1} , usando como amostra apenas os r vizinhos mais próximos (em valor) de y_{t-1} e empregando alguma função de ponderação que dê mais peso às observações mais próximas (em valor) de y_{t-1} .

A RLP foi usada aqui de forma semelhante ao que foi feito em Hsieh (1993), no que diz respeito à escolha das condições em que são feitas as regressões: o número de vizinhos mais

próximos foi arbitrado em 50, 100 e 200 observações; foi usado o sistema de ponderação com pesos tricúbicos, proposto em Cleveland & Devlin (1988) e foi feita a previsão dos retornos no terço final das séries (400 valores), usando como amostra os 800 valores iniciais; foi então determinado o erro quadrático médio destas previsões, e o mesmo foi feito para as previsões baseadas na hipótese de média condicional nula (previsão de y_t igual a zero), cujo erro quadrático médio foi tomado como referência.

Na Tabela 6, nas colunas de “quocientes” estão os valores dos quocientes entre os erros quadráticos médios obtidos pelos dois processos; os valores maiores que um indicam que as previsões do método de RLP foram piores que as de referência.

Observando-se os resultados nota-se que, na média, os dois métodos de previsão têm erros equivalentes mas que em algumas séries a RLP consegue erros menores, indicando que, para elas, talvez a média condicional não seja nula (especialmente para BBDC, que também apresentou coeficientes de bicorrelação elevados).

Harvey, Leybourne & Newbold (1977) propõem uma variante do teste de Morgan (1939), Granger & Newbold (1977) para testar igualdade de erros de previsão um passo à frente. Esta variante, ao contrário do teste original, é robusta à não normalidade dos erros, ou seja, ela apresenta bons resultados ainda que, como acontece nas séries de retornos, os erros de previsão apresentem caudas mais pesadas que as da normal. A hipótese nula é a igualdade dos erros quadráticos médios das duas previsões; sob esta hipótese e para amostras grandes, a estatística de teste tem distribuição normal padronizada. Como hipótese alternativa considerou-se que os erros da RLP são menores que os de referência.

Na Tabela 6, nas colunas de “estatísticas”, estão os valores calculados para o teste de Morgan, Granger e Newbold modificado; o valor crítico para um nível de 5% é 1,65. As séries onde os erros de previsão da RLP são significativamente menores que os do modelo de referência estão com seus quocientes marcados com um asterisco.

Tabela 6 – Regressão Localmente Ponderada – comparação de erros de previsão

vizinhos:	50		100		200	
	quocientes	estatísticas	quocientes	estatísticas	quocientes	estatísticas
otelb	0,97	1,07	0,96 (*)	2,12	0,96 (*)	3,76
mtelb	1,02	-0,86	1,00	0,24	1,00	0,31
opetr	1,10	-2,82	1,06	-2,95	1,03	-2,54
mpetr	1,03	-0,51	1,01	-0,19	0,99	0,22
oelet	0,94 (*)	1,67	0,95	1,64	0,96	1,52
melet	1,01	-0,29	1,00	-0,06	1,00	0,07
obbdc	0,96	0,35	0,96	0,44	0,95	0,68
mbbdc	0,95	1,33	0,95	1,79	0,95 (*)	1,99
ovale	1,03	-1,60	1,01	-1,41	1,01	-2,28
mvale	1,07	-2,78	1,03	-2,68	1,04	-3,58
obrha	1,00	-0,01	0,99	0,42	0,98	0,78
mbrha	0,98	0,67	0,97	1,38	0,96 (*)	2,03
média	1,00		0,99		0,99	

Na Tabela 6 observa-se que as reduções de erro para TELB e ELET originais desaparecem para as séries modificadas, sugerindo que as reduções estão mais ligadas aos elevados valores dos retornos (e dos erros quadrados de previsão) nos dias atípicos do que à existência de estrutura nos retornos diários, supostamente modelada pela RLP.

Para BBDC e BRHA há redução de erro nas séries originais e nas modificadas, mais acentuada nas modificadas e significativa apenas na RLP de 200 vizinhos.

4.4.3 Estacionariedade da Média: Conclusão

Como BBDC apresentou alguns coeficientes de bicorrelação individualmente significantes e, nos quinze coeficientes tomados em conjunto, a estatística de teste ficou próxima do valor crítico, há indicação de que a média condicional talvez não seja nula para esta série, já que a RLP também indica alguma estrutura.

Para as demais séries, o conjunto dos dois testes indica média condicional nula e conseqüentemente média estacionária.

Observe-se que a média condicional não nula não indica que necessariamente a média da série não é estacionária e, portanto, BBDC pode ter média estacionária, mesmo que se conclua que sua média condicional varia.

5. Distribuição Incondicional

Observando os histogramas, os valores máximo e mínimo, e os coeficientes de assimetria e curtose dos dados de retorno das séries (não apresentados neste texto), pode-se inicialmente notar a grande influência das observações aberrantes na forma da distribuição: as séries modificadas têm coeficiente de assimetria, em geral, mais próximo de zero e coeficiente de curtose menor que nas séries originais.

Mesmo após a retirada dos dias atípicos, as séries modificadas têm as caudas bem mais pesadas que as da distribuição normal, como sempre acontece com as séries de retornos de ações. Ao contrário do observado por Taylor (1986) não se nota predominância de valores positivos para o coeficiente de assimetria.

O modelo de Bachelier (1900) e Osborne (1959) assume que o preço das ações varia em decorrência do somatório de um número grande de mudanças que ocorrem transação a transação; se estas mudanças são independentes, o teorema do limite central leva à conclusão que as variações diárias (ou em outro intervalo de tempo maior) têm distribuição normal.

Para testar normalidade das distribuições foi usado o teste de Jarque-Bera (1987), cuja hipótese nula é coeficiente de assimetria nulo e de curtose igual a três, simultaneamente. A estatística de teste tem distribuição qui-quadrado com dois graus de liberdade e o valor crítico, para um nível de 5%, é igual a 5,99. Na Tabela 7 estão os valores da estatística para as doze séries, que indicam que as distribuições estão bem longe de serem normais.

Tabela 7 – Teste de Jarque-Bera

	ORIGINAL	MODIFICADA
TELB	10745	105
PETR	1142	518
ELET	2986	703
BBDC	5569	874
VALE	11159	127
BRHA	3078	2582

6. Independência

6.1 Introdução

A não normalidade incondicional das séries de retornos é muitas vezes explicada pelo fato de os retornos não serem independentes e identicamente distribuídos (IID), levando naturalmente ao estudo da independência.

A questão da existência ou não de dependência em séries de retornos resultou em inúmeras publicações, nem todas chegando às mesmas conclusões; o assunto é importante porque está relacionado com a eficiência dos mercados e com a possibilidade de obtenção de ganhos continuados, explorando possíveis padrões de comportamento dos preços através da modelagem da média condicional.

A maior parte dos autores concorda que é muito difícil prever preços a partir de preços passados, mas a verificação da existência de dependência não linear entre retornos consecutivos e a modelagem desta dependência são importantes para a quantificação (e gerência) do risco.

Nesta seção, a dependência temporal entre as distribuições de retornos é inicialmente estudada através do teste da existência de dependência linear, que é aquela que ocorre quando o retorno pode ser representado por uma distribuição que é combinação linear de retornos passados. Verifica-se que, embora exista alguma dependência linear nas séries estudadas, ela não explica a contento o comportamento dos retornos; na seção 7, caracterizada a existência de dependência não linear, são ajustados aos dados três modelos não lineares, sendo seu poder de explicação investigado através das características dos resíduos.

6.2 Autocorrelações

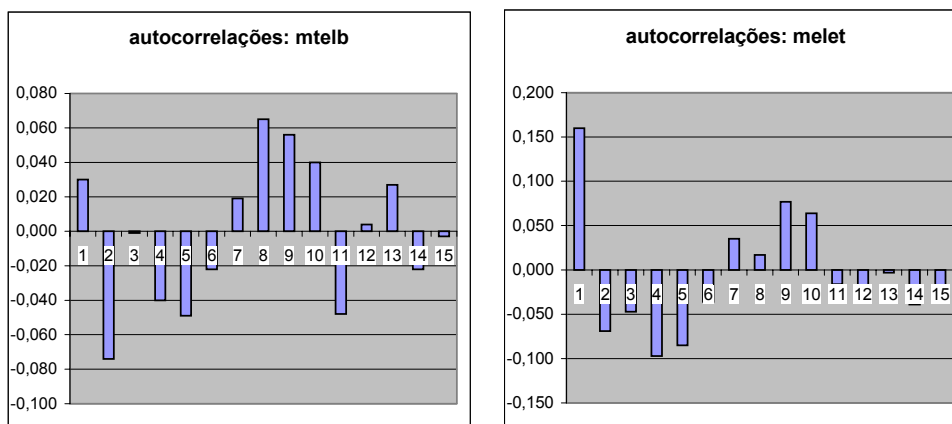
Na Tabela 8 estão os coeficientes de autocorrelação para as séries originais. O tipo de função de autocorrelação é o mesmo para as séries originais e modificadas; entretanto as observações aberrantes parecem introduzir nas séries um *ruído* que atenua as características das funções de autocorrelação, que são mais acentuadas nas séries modificadas. Por isto, toda a análise de dependência será feita apenas para as séries modificadas; as séries originais apresentam as mesmas características, porém de forma menos nítida.

Tabela 8 – Função de autocorrelação: séries modificadas

Lag	mtelb	mpetr	melet	mbbdc	mvale	mbrha
1	0,030	0,160	0,160	0,171	0,044	0,147
2	-0,074	-0,066	-0,069	-0,082	-0,074	-0,025
3	-0,001	-0,034	-0,047	-0,005	-0,055	-0,038
4	-0,040	-0,037	-0,097	-0,023	-0,031	-0,038
5	-0,049	-0,014	-0,085	-0,057	-0,050	-0,034
6	-0,022	-0,005	-0,036	-0,045	-0,011	-0,083
7	0,019	0,020	0,035	0,000	-0,013	-0,020
8	0,065	0,022	0,017	0,068	0,009	0,022
9	0,056	0,035	0,077	0,098	0,052	0,033
10	0,040	0,006	0,064	-0,012	0,021	0,013
11	-0,048	-0,018	-0,016	-0,109	-0,044	0,012
12	0,004	0,034	-0,019	0,004	-0,003	0,078
13	0,027	0,022	-0,003	0,007	0,061	0,036
14	-0,022	-0,022	-0,039	-0,051	0,011	-0,034
15	-0,003	0,031	-0,023	-0,072	-0,050	-0,063

Observando a Tabela 8, pode-se notar que a autocorrelação de *lag* um é positiva para todas as séries; num nível de 5%, ela não é significativa apenas para MTELB e MVALE. Do *lag* 2 ao *lag* 6 todas as autocorrelações são negativas, para todas as séries. Tomadas em conjunto, as autocorrelações dos *n* primeiros *lags* (*n* menor ou igual a quinze) são significantes, sempre num nível de 5%, para todas as séries (exceto MTELB e MVALE para *n*=1), o que implica na existência de dependência.

Na Figura 4 pode-se visualizar, para duas das séries, o padrão de senóide amortecida das funções de autocorrelação, característico de modelos lineares autorregressivos; este padrão se repete para todas as séries. As funções de autocorrelação parcial também apresentam padrão de senóide amortecida, característico de modelos de médias móveis.

**Figura 4** – Gráfico das autocorrelações para duas das séries

6.3 Modelagem ARMA

As funções de autocorrelação sugerem que um modelo ARMA possa representar as séries em estudo; foram estimados modelos ARMA(p,q) para as seis séries modificadas. Embora o formato das funções de autocorrelação possam sugerir valores de p e q maiores, o critério de informação de Schwarz indicou, para a maior parte das séries $p=q=1$. O modelo considerado para as séries foi $y_t = a + by_{t-1} + c\varepsilon_{t-1} + \varepsilon_t$ onde os valores obtidos para as constantes a , b e c estão na Tabela 9.

Tabela 9 – Coeficientes do modelo linear

	a	b	c
mtelb	0,000 <i>0,030</i>	-0,676 <i>0,179</i>	0,724 <i>0,168</i>
mpetr	0,000 <i>0,032</i>	-0,277 <i>0,136</i>	0,454 <i>0,126</i>
melet	0,000 <i>0,032</i>	-0,312 <i>0,130</i>	0,491 <i>0,119</i>
mbbdc	0,001 <i>0,032</i>	-0,363 <i>0,107</i>	0,558 <i>0,095</i>
mvale	0,001 <i>0,030</i>	-0,351 <i>0,373</i>	0,409 <i>0,363</i>
mbrha	0,000 <i>0,031</i>	-0,535 <i>0,107</i>	0,655 <i>0,095</i>

Em itálico está o erro padrão das estimativas; num nível de 5%, o coeficiente a não é significativo, enquanto que os coeficientes b e c são significantes, para a maioria das séries. Se a modelagem foi eficaz, toda dependência foi capturada pelo modelo e os resíduos devem ser independentes. Entretanto, os resíduos deste modelo apresentam ainda alguma autocorrelação, nos *lags* mais altos, e os quadrados dos resíduos são fortemente correlacionados, indicando dependência não linear nos resíduos e, conseqüentemente, nos retornos.

7. Modelagem Não Linear

7.1 Introdução

Como na seção anterior ficou caracterizada dependência não linear entre os retornos, tenta-se aqui modelar as séries através de três modelos comuns para este tipo de dependência: o desvio padrão condicional com amortecimento exponencial (EWMA), o GARCH(1,1) e o GARCH(1,1) com modelagem ARMA(1,1) da média condicional, que será chamado de ARMA-GARCH. Os resíduos destes modelos foram analisados quanto a normalidade, quanto a autocorrelação e quanto a distribuição independente e idêntica.

Os resultados mostram que nenhum dos modelos produz resíduos com distribuição normal. Os resíduos do EWMA e do GARCH são correlacionados mas os do ARMA-GARCH não o são; quanto aos quadrados dos resíduos, os do GARCH não são correlacionados, os do

ARMA-GARCH o são para duas das séries e os de EWMA para três. Quanto ao comportamento IID, verifica-se, através do teste BDS, que os resíduos mais próximos deste comportamento são os do ARMA-GARCH e os mais afastados são os do EWMA.

7.2 Teste BDS

Este teste foi desenvolvido por Brock, Dechert & Scheinkman (1987) e é usado para testar se as variáveis aleatórias que compõem uma série são IID; ele pode detectar várias situações em que as variáveis não são IID, tais como não estacionariedade, não linearidade e caos determinístico. Ele vai ser empregado mais adiante para testar os resíduos dos três modelos não lineares; a seguir é feita uma rápida descrição do teste.

O teste se baseia no conceito de *integral de correlação* $C_m(\delta)$.

Seja $\{y_t, t = 1, 2, \dots, T\}$ uma série temporal e $y_t^m = (y_t, y_{t-1}, \dots, y_{t-m+1})$ um ponto no espaço euclidiano a m dimensões; a integral de correlação, definida na equação a seguir, mede o quociente entre o número de pares de pontos (y_t^m, y_s^m) cuja distância entre os pontos é menor que δ , e o número total T^2 de pares (y_t^m, y_s^m) . Formalmente:

$$C_m(\delta) = p \lim_{T \rightarrow \infty} \left[\frac{\#\left\{ (t, s), 0 < t < T, 0 < s < T : \max_{i=0, \dots, m-1} |y_{t-i} - y_{s-i}| < \delta \right\}}{T^2} \right]$$

Note-se que o conceito de distância usado é o de máxima norma e não a distância euclidiana. Se $\{y_t\}$ é IID, então $C_m(\delta) = [C_1(\delta)]^m$; Brock, Dechert & Scheinkman (1987) desenvolveram uma estatística para testar a hipótese nula de validade desta igualdade e mostraram que esta estatística tem distribuição assintoticamente normal padronizada.

Neste trabalho foram arbitrados seis valores para δ , entre meio e três desvios padrão, e cinco valores para o número de dimensões, de 2 a 6, resultando 30 testes para cada série.

O valor crítico para rejeição da hipótese nula (resíduos IID), com nível de 5%, é igual a 1,96 (em valor absoluto). Entretanto, o teste BDS aplicado a resíduos de modelos da família GARCH apresenta tendência em favor da hipótese nula. Os valores críticos corretos são menores que os teóricos (da distribuição normal) e são função do modelo, do valor dos parâmetros, de m e de T , podem chegar a ser da ordem de 50% do valor teórico e não convergem para o valor teórico, nem para amostras grandes. Muitos autores usam valores críticos obtidos por simulação; neste trabalho são usadas as fórmulas empíricas de correção obtidas por Fernandes & Preumont (1996) que, também através de simulação, usando superfícies de resposta, determinaram a forma da dependência entre os valores críticos e os fatores que os influenciam.

7.3 Modelagem e Resultados

7.3.1 Amortecimento Exponencial

O modelo supõe que os retornos seguem um processo produto $X_t = V_t U_t$, onde U_t tem distribuição independente normal padronizada, e também independente de V_t , que é o processo do desvio padrão condicional. Em valor absoluto vem $|X_t| = |V_t| |U_t|$ e também

$E[|X_t|] = E[V_t]E[|U_t|]$ ou seja, chamando de M_t o valor absoluto do retorno, a previsão do desvio padrão condicional é dada por $\hat{v}_t = \hat{m}_t / \sqrt{2/\pi}$. Desta forma, obtendo-se a previsão do valor absoluto do retorno, obtém-se também a previsão do desvio padrão condicional. Uma boa previsão do retorno absoluto (Taylor, 1986) é obtida por médias móveis com amortecimento exponencial: $\hat{m}_t = \theta \hat{m}_{t-1} + (1-\theta)m_{t-1}$, sendo o parâmetro θ estimado por mínimos quadrados. O retorno padronizado $u_t = x_t/v_t$ não é observável mas o retorno *re-escalado* $y_t = x_t/\hat{v}_t$ é uma boa aproximação, se \hat{v}_t é uma boa previsão. Este retorno re-escalado é o resíduo que será analisado a seguir.

Quanto à distribuição incondicional, os resíduos apresentaram excesso de curtose incompatível com a distribuição normal, e o desvio padrão foi maior que um, conforme a Tabela 10.

Tabela 10 – Resíduos do amortecimento exponencial – parâmetros da distribuição incondicional

	CURTOSE	DESVIO PADRÃO
MTELB	4,11	1,05
MPETR	4,12	1,06
MELET	4,49	1,08
MBBDC	4,66	1,13
MVALE	5,07	1,10
MBRHA	27,45	1,36

Quanto à autocorrelação, os resíduos de cinco das seis séries se mostraram correlacionados e seus quadrados apresentaram correlação significativa apenas para três séries. Na Tabela 11 estão os *p-valores* do teste de significância conjunta dos coeficientes de autocorrelação para os cinco primeiros *lags*; os que correspondem a significância (existência de autocorrelação), para um nível de 5%, estão marcados com um asterisco.

Tabela 11 – Resíduos do amortecimento exponencial – *p-valores* para significância de autocorrelação

	RESÍDUOS	RESÍD.QUADRADOS
MTELB	0,126	0,023 (*)
MPETR	0,000 (*)	0,000 (*)
MELET	0,000 (*)	0,448
MBBDC	0,000 (*)	0,027 (*)
MVALE	0,003 (*)	0,111
MBRHA	0,049 (*)	0,989

O teste BDS indica que apenas MELET tem resíduos IID. Na Tabela 12 estão as estatísticas de teste, e os valores significantes estão marcados com um asterisco; para um nível de 5%, o valor crítico é 1,96.

Tabela 12 – Estatística do teste BDS para resíduos do modelo EWMA

δ	m	mtelb	mpetr	melet	mbbdc	mvale	mbrha
3,0	2	2,99 (*)	1,33	1,35	5,27 (*)	2,71 (*)	5,28 (*)
3,0	3	3,56 (*)	1,86	0,71	4,91 (*)	2,99 (*)	4,92 (*)
3,0	4	3,38 (*)	2,55 (*)	0,21	4,65 (*)	2,74 (*)	4,66 (*)
3,0	5	3,39 (*)	2,41 (*)	0,08	4,22 (*)	2,78 (*)	4,23 (*)
3,0	6	3,19 (*)	2,23 (*)	-0,24	3,89 (*)	2,58 (*)	3,91 (*)
2,5	2	2,84 (*)	1,46	1,77	5,64 (*)	3,04 (*)	5,65 (*)
2,5	3	3,28 (*)	2,24 (*)	1,04	5,67 (*)	3,39 (*)	5,68 (*)
2,5	4	2,90 (*)	2,61 (*)	0,50	5,35 (*)	3,14 (*)	5,36 (*)
2,5	5	2,91 (*)	2,34 (*)	0,38	4,85 (*)	3,05 (*)	4,85 (*)
2,5	6	2,67 (*)	2,19 (*)	-0,03	4,48 (*)	2,81 (*)	4,48 (*)
2,0	2	1,73	1,34	1,93	5,80 (*)	3,29 (*)	5,79 (*)
2,0	3	2,18 (*)	2,19 (*)	1,19	6,12 (*)	3,60 (*)	6,12 (*)
2,0	4	1,78	2,55 (*)	0,74	5,85 (*)	3,52 (*)	5,85 (*)
2,0	5	1,84	2,29 (*)	0,64	5,43 (*)	3,42 (*)	5,43 (*)
2,0	6	1,71	2,13 (*)	0,24	5,07 (*)	3,25 (*)	5,07 (*)
1,5	2	1,06	1,03	1,79	5,57 (*)	3,04 (*)	5,56 (*)
1,5	3	1,45	1,90	1,20	6,29 (*)	3,25 (*)	6,27 (*)
1,5	4	1,05	2,24 (*)	0,73	6,14 (*)	3,25 (*)	6,11 (*)
1,5	5	1,21	1,90	0,54	5,77 (*)	3,17 (*)	5,74 (*)
1,5	6	1,18	1,65	0,13	5,30 (*)	3,06 (*)	5,28 (*)
1,0	2	0,11	0,43	1,38	4,92 (*)	2,94 (*)	4,93 (*)
1,0	3	0,64	1,30	0,47	6,17 (*)	2,99 (*)	6,18 (*)
1,0	4	0,40	1,60	-0,11	6,26 (*)	2,96 (*)	6,26 (*)
1,0	5	0,68	1,22	-0,23	6,03 (*)	2,95 (*)	6,04 (*)
1,0	6	0,73	0,92	-0,65	5,55 (*)	2,88 (*)	5,55 (*)
0,5	2	0,13	0,35	1,14	3,62 (*)	2,58 (*)	3,62 (*)
0,5	3	0,60	1,09	0,31	5,32 (*)	2,53 (*)	5,34 (*)
0,5	4	0,68	1,64	-0,37	5,28 (*)	2,45 (*)	5,30 (*)
0,5	5	0,79	1,52	-0,19	5,52 (*)	2,69 (*)	5,58 (*)
0,5	6	1,26	1,27	-0,28	4,63 (*)	2,30 (*)	4,75 (*)

7.3.2 GARCH(1,1)

Este modelo supõe que os retornos são gerados por um processo $x_t = \hat{v}_t u_t$, mas agora o desvio padrão condicional é estimado pela equação $\hat{v}_t^2 = C + \alpha x_{t-1}^2 + \beta \hat{v}_{t-1}^2$ onde os parâmetros C , α e β são estimados por máxima verossimilhança. Os resultados para a análise de normalidade dos retornos re-escalados (u_t) são estão na Tabela 13.

Tabela 13 – Resíduos do GARCH(1,1) – parâmetros da distribuição incondicional

	CURTOSE	DESVIO PADRÃO
MTELB	3,74	1,00
MPETR	3,42	1,00
MELET	4,16	1,00
MBBDC	4,17	1,00
MVALE	4,64	1,00
MBRHA	6,87	1,00

Observa-se também excesso de curtose que leva à rejeição de normalidade, mas o desvio padrão de todas as séries é igual a um. Na Tabela 14 estão os *p-valores* do teste de significância conjunta dos coeficientes de autocorrelação para os cinco primeiros *lags*; os que correspondem a significância (existência de autocorrelação), para um nível de 5%, estão marcados com um asterisco. Esta tabela mostra que os resíduos de todas as séries – exceto MTELB – são correlacionados, mas os resíduos quadrados não o são, indicando que o GARCH(1,1) modelou bem a dependência não linear, mas passou dependência linear para os resíduos.

Tabela 14 – Resíduos do GARCH(1,1) – *p-valores* para significância de autocorrelação

	RESÍDUOS	RESÍDUOS QUADRADOS
MTELB	0,144	0,529
MPETR	0,001 (*)	0,068
MELET	0,000 (*)	0,612
MBBDC	0,003 (*)	0,531
MVALE	0,000 (*)	0,938
MBRHA	0,012 (*)	0,058

O teste BDS, cujas estatísticas estão na Tabela 15, mostra que em relação ao modelo EWMA, os resíduos não estão mais próximos de serem IID: todas as séries apresentam alguns valores significantes da estatística e, a rigor, nenhuma série pode ser considerada IID. A que mais se aproxima é MBBDC, que tem apenas três valores significantes.

Para maior clareza, as correções de Fernandes & Preumont foram aplicadas às estatísticas BDS e não aos valores críticos, para evitar que o valor crítico fosse diferente em cada célula da Tabela 15, onde os valores maiores que 1,96 (significantes) estão marcados com um asterisco.

Tabela 15 – Estatística do teste BDS para resíduos do modelo GARCH

δ	m	mtelb	mpetr	melet	mbbdc	mvale	mbrha
3,0	2	2,97 (*)	4,18 (*)	2,64 (*)	0,83	0,25	5,88 (*)
3,0	3	2,96 (*)	4,94 (*)	1,96 (*)	-0,26	0,22	4,47 (*)
3,0	4	2,51 (*)	4,79 (*)	1,41	-1,03	-0,28	3,70 (*)
3,0	5	2,13 (*)	4,02 (*)	1,24	-1,44	-0,25	3,03 (*)
3,0	6	1,60	3,48 (*)	0,92	-1,57	-0,49	2,51 (*)
2,5	2	2,93 (*)	3,92 (*)	3,93 (*)	1,56	1,12	8,48 (*)
2,5	3	3,32 (*)	4,81 (*)	3,27 (*)	0,46	0,82	6,73 (*)
2,5	4	2,75 (*)	4,78 (*)	2,57 (*)	-0,42	0,29	5,61 (*)
2,5	5	2,54 (*)	4,11 (*)	2,34 (*)	-0,89	0,28	4,83 (*)
2,5	6	2,11 (*)	3,59 (*)	1,87 (*)	-1,11	0,02	4,29 (*)
2,0	2	2,27 (*)	2,82 (*)	4,53 (*)	2,03 (*)	1,85	8,54 (*)
2,0	3	3,09 (*)	4,24 (*)	3,93 (*)	1,60	1,33	6,86 (*)
2,0	4	2,60 (*)	4,52 (*)	3,28 (*)	0,85	1,01	5,72 (*)
2,0	5	2,53 (*)	3,89 (*)	3,01 (*)	0,35	1,11	5,17 (*)
2,0	6	2,30 (*)	3,38 (*)	2,46 (*)	0,00	0,95	4,82 (*)
1,5	2	1,41	2,39 (*)	4,64 (*)	2,35 (*)	2,13 (*)	8,63 (*)
1,5	3	2,58 (*)	3,70 (*)	4,03 (*)	2,14 (*)	1,47	7,09 (*)
1,5	4	2,28 (*)	4,02 (*)	3,49 (*)	1,40	1,36	5,97 (*)
1,5	5	2,60 (*)	3,37 (*)	3,12 (*)	0,81	1,57	5,57 (*)
1,5	6	2,62 (*)	2,75 (*)	2,45 (*)	0,32	1,55	5,24 (*)
1,0	2	0,60	1,00	4,20 (*)	1,56	2,85 (*)	9,43 (*)
1,0	3	1,75	2,39 (*)	3,08 (*)	1,81	2,26 (*)	8,65 (*)
1,0	4	1,61	2,97 (*)	2,34 (*)	1,27	2,16 (*)	7,47 (*)
1,0	5	1,98 (*)	2,34 (*)	2,07 (*)	0,92	2,38 (*)	6,99 (*)
1,0	6	2,18 (*)	1,74	1,42	0,67	2,37 (*)	6,50 (*)
0,5	2	0,57	1,09	4,33 (*)	0,60	3,72 (*)	9,87 (*)
0,5	3	1,38	2,04 (*)	3,33 (*)	1,31	3,04 (*)	9,92 (*)
0,5	4	1,23	2,76 (*)	1,70	0,90	2,53 (*)	9,85 (*)
0,5	5	1,57	2,41 (*)	1,51	1,54	3,79 (*)	10,55 (*)
0,5	6	2,74 (*)	2,23 (*)	0,15	1,83	4,62 (*)	11,71 (*)

7.3.3 GARCH(1,1) com Modelagem ARMA(1,1) da Média Condicional

Como os resíduos do GARCH(1,1) apresentaram autocorrelação, sugerindo dependência linear entre eles (e conseqüentemente entre os retornos), foi ajustado aos retornos modelo GARCH(1,1) com modelagem ARMA(1,1) para a média condicional. A equação do retorno é $x_t = C_1 + a x_{t-1} + b u_{t-1} + u_t$, onde u_t segue um processo GARCH(1,1) do tipo $u_t = \hat{v}_t z_t$ com z_t IID $N(0,1)$. O desvio padrão condicional de u_t é dado por $\hat{v}_t^2 = C + \alpha z_{t-1}^2 + \beta \hat{v}_{t-1}^2$.

Novamente os parâmetros são estimados por verossimilhança, e os resultados para análise de normalidade dos resíduos z_t são mostrados na Tabela 16, onde os resultados são semelhantes aos do modelo anterior, com uma pequena diminuição dos coeficientes de curtose.

Tabela 16 – Resíduos do GARCH(1,1) com modelagem ARMA(1,1) da média – parâmetros da distribuição incondicional

	CURTOSE	DESVIO PADRÃO
MTELB	3,61	1,00
MPETR	3,31	1,00
MELET	4,05	1,00
MBBDC	4,12	1,00
MVALE	4,52	1,00
MBRHA	6,83	1,00

Entretanto, para as autocorrelações há algumas mudanças importantes, conforme pode ser visto na Tabela 17, onde estão os *p-valores* do teste de significância conjunta dos coeficientes de autocorrelação para os cinco primeiros *lags*; os que correspondem a significância (existência de autocorrelação), para um nível de 5%, estão marcados com um asterisco. Nesta tabela, pode-se observar que desapareceu a autocorrelação dos resíduos, mas aumentou um pouco a dos resíduos quadrados, com relação ao modelo anterior.

Tabela 17 – Resíduos do GARCH(1,1) com modelagem ARMA(1,1) da média – *p-valores* para significância de autocorrelação

	RESÍDUOS	RESÍDUOS QUADRADOS
MTELB	0,067	0,218
MPETR	0,207	0,050 (*)
MELET	0,210	0,285
MBBDC	0,254	0,272
MVALE	0,737	0,660
MBRHA	0,186	0,008 (*)

O teste BDS, cujas estatísticas estão na Tabela 18, indica que este modelo produz resíduos mais próximos do comportamento IID do que os modelos anteriores: para valor crítico 1,96, MVALE não tem estatística significativa, MBBDC tem duas e MTELB tem quatro. Entretanto, como um todo, o ARMA-GARCH também não conseguiu representar bem o processo gerador dos dados.

Aqui também, para maior clareza, as correções de Fernandes & Preumont foram aplicadas às estatísticas BDS e não aos valores críticos, para evitar que o valor crítico fosse diferente em cada célula da Tabela 18, onde os valores maiores que 1,96 (significantes) estão marcados com um asterisco.

Tabela 18 – Estatística do teste BDS para resíduos do modelo ARMA-GARCH

δ	m	mtelb	mpetr	melet	mbbdc	mvale	mbrha
3,0	2	1,91	2,86 (*)	0,95	0,65	-0,13	5,71 (*)
3,0	3	2,08 (*)	3,92 (*)	0,61	-0,05	-0,11	4,62 (*)
3,0	4	1,29	3,45 (*)	0,18	-0,80	0,02	3,48 (*)
3,0	5	0,85	2,74 (*)	0,14	-1,23	-0,01	2,68 (*)
3,0	6	0,30	2,16 (*)	-0,11	-1,36	-0,26	2,15 (*)
2,5	2	1,85	2,01 (*)	2,16 (*)	1,40	-0,86	8,41 (*)
2,5	3	2,54 (*)	3,40 (*)	1,78	0,97	-0,90	7,15 (*)
2,5	4	1,73	3,41 (*)	1,42	0,19	-0,59	5,62 (*)
2,5	5	1,49	2,88 (*)	1,26	-0,33	-0,41	4,71 (*)
2,5	6	1,05	2,35 (*)	0,82	-0,59	-0,54	4,14 (*)
2,0	2	1,13	1,67	2,94 (*)	1,79	-0,62	8,58 (*)
2,0	3	2,27 (*)	3,03 (*)	2,46 (*)	1,86	-0,82	7,37 (*)
2,0	4	1,52	3,35 (*)	2,12 (*)	1,20	-0,42	5,88 (*)
2,0	5	1,48	2,86 (*)	1,82	0,79	-0,10	5,18 (*)
2,0	6	1,31	2,38 (*)	1,28	0,42	-0,15	4,77 (*)
1,5	2	0,28	1,07	3,35 (*)	1,79	-0,14	8,64 (*)
1,5	3	1,66	2,34 (*)	2,96 (*)	2,42 (*)	-0,45	7,68 (*)
1,5	4	1,17	2,88 (*)	2,70 (*)	1,82	0,10	6,33 (*)
1,5	5	1,47	2,30 (*)	2,26 (*)	1,31	0,49	5,83 (*)
1,5	6	1,40	1,75	1,66	0,74	0,55	5,45 (*)
1,0	2	-0,18	1,05	3,00 (*)	1,49	0,48	9,67 (*)
1,0	3	1,12	2,42 (*)	2,14 (*)	2,19 (*)	0,09	9,21 (*)
1,0	4	0,99	3,53 (*)	1,87	1,77	0,60	7,94 (*)
1,0	5	1,60	3,07 (*)	1,55	1,40	0,92	7,49 (*)
1,0	6	1,82	2,30 (*)	1,03	0,96	0,96	7,01 (*)
0,5	2	-1,06	-0,18	4,16 (*)	0,05	1,49	10,03 (*)
0,5	3	0,00	1,86	3,03 (*)	1,29	1,13	10,11 (*)
0,5	4	0,33	2,53 (*)	1,74	1,30	1,58	10,13 (*)
0,5	5	1,15	2,12 (*)	1,54	1,56	1,66	11,11 (*)
0,5	6	2,19	0,75	0,69	1,70	1,23	12,64 (*)

É interessante notar que os parâmetros da equação de variância nos dois modelos GARCH são muito parecidos, e que a melhora dos resíduos corresponde à dependência linear capturada pela equação da média do último modelo.

7.4 Modelagem Não Linear: Conclusão

Do que foi visto nesta seção, pode-se concluir que o modelo mais sofisticado (ARMA-GARCH) foi o que melhor representou o processo gerador dos retornos, no que diz respeito às propriedades dos resíduos. Entretanto, esta representação tem ainda alguns problemas: os

resíduos têm excesso de curtose, há correlação em seus quadrados para duas séries (MPETR e MBRHA) e estas mesmas duas séries (e também MELET) têm resíduos claramente não IID, pelo resultado do teste BDS. Na verdade, o processo gerador da série MBRHA está longe de ser bem representado por qualquer dos três modelos.

Na Tabela 19 está um resumo da análise dos resíduos dos modelos. Os *p-valores* marcados com um asterisco correspondem a autocorrelações significantes. Para resumir o teste BDS está apresentada, apenas a título de ilustração, a média das 30 estatísticas calculadas para os resíduos de cada série; os valores maiores que 1,96 estão marcados com um asterisco, mas um valor da média das estatísticas maior que 1,96 não necessariamente significa que a série não é IID, e vice-versa.

Tabela 19 – Resumo da análise dos resíduos dos modelos

		MTELB	MPETR	MELET	MBBDC	MVALE	MBRHA
Curtose	EWMA	4.11	4.12	4.49	4.66	5.07	27.45
	GARCH	3.74	3.42	4.16	4.17	4.64	6.87
	ARMA-GARCH	3.61	3.31	4.05	4.12	4.52	6.83
p-valor da correlação do resíduo	EWMA	0.126	0.000 (*)	0.000 (*)	0.000 (*)	0.003 (*)	0.049 (*)
	GARCH	0.144	0.001 (*)	0.000 (*)	0.003 (*)	0.000 (*)	0.012 (*)
	ARMA-GARCH	0.067	0.207	0.210	0.254	0.737	0.186
p-valor da correlação do resíduo ²	EWMA	0.023 (*)	0.000 (*)	0.448	0.027 (*)	0.111	0.989
	GARCH	0.529	0.068	0.612	0.531	0.938	0.058
	ARMA-GARCH	0.218	0.050 (*)	0.285	0.272	0.660	0.008 (*)
Média das estatísticas BDS	EWMA	1.74	1.73	0.55	5.32 (*)	2.98 (*)	5.33 (*)
	GARCH	2.19 (*)	3.29 (*)	2.71 (*)	0.68	1.53	6.80 (*)
	ARMA-GARCH	1.23	2.41 (*)	1.78	0.89	0.16	6.99 (*)

8. Conclusão do Trabalho

As seis ações aqui estudadas representam, em termos de volume, cerca de 60% do mercado acionário brasileiro e as suas características gerais mostraram-se bastante homogêneas, podendo-se considerar que estas são, de modo geral, as características do grupo de ações mais líquidas de nosso mercado.

Quanto à estacionariedade, o teste ADF indicou a inexistência de raiz unitária, raiz esta característica de modelos lineares não estacionários, mas como existe dependência não linear entre os retornos, a indicação de estacionariedade do teste ADF não é válida; as séries puderam ser modeladas por GARCH(1,1) estacionário, embora de alta persistência. Embora o teste de bicorrelação indique estacionariedade da média, a regressão localmente ponderada tem resultados inconclusivos quanto a esta estacionariedade em algumas séries.

O conjunto das análises, porém, leva à conclusão de que a média das séries pode ser considerada estacionária. Quanto à variância, não há uma conclusão: não foram feitos testes de estacionariedade fora das premissas de linearidade (que não se mostrou verdadeira) ou de modelo GARCH (1,1), cujos coeficientes sugerem que o processo gerador de algumas séries pode ser o IGARCH (1,1), sem variância finita.

Como na maioria dos estudos deste tipo, as distribuições dos retornos apresentaram caudas pesadas incompatíveis com a normal.

As autocorrelações dos retornos indicam a existência de dependência linear, modelada pelo ARMA(1,1), que reduziu bastante a autocorrelação nos *lags* mais baixos; entretanto, passou para os resíduos autocorrelação nos *lags* mais altos, indicando dependência também não linear nos retornos. Os quadrados dos resíduos se mostraram altamente correlacionados, o que também é indício de dependência não linear.

O próximo passo foi tentar modelar a dependência detectada, usando as propriedades dos resíduos para checar o ajuste dos modelos. Neste trabalho considerou-se válida a idéia de que os retornos são misturas de distribuições (condicionalmente) normais; os modelos foram estimados dentro desta hipótese e, assim, seus resíduos devem ser IID normal (0,1). O modelo que melhor se ajustou aos dados foi o GARCH(1,1) com modelagem ARMA(1,1) da média, que apresentou resíduos sempre descorrelacionados; para três das seis séries, o teste BDS indicou resíduos deste modelo próximos de IID, mas eles foram sempre não normais.

Os outros dois modelos analisados (EWMA e GARCH puro) tiveram ajuste pior; como eles são muito utilizados pelas instituições financeiras para modelagem de volatilidade, seria interessante verificar até que ponto suas previsões diferem das do melhor modelo.

Outras possíveis continuações deste trabalho seriam:

- aplicar outros modelos aos dados, em busca de resíduos normais.
- aplicar aos dados modelos onde a distribuição teórica dos resíduos não seja a normal.
- aplicar aos dados modelos no domínio da frequência.
- comparação das previsões de média e volatilidade dos diversos modelos com a realidade.

Referências Bibliográficas

- (1) Alexander, S.S. (1961). Price movements in speculative markets: trends or random walks?, *Industrial Management Review*, **2**, 7-26.
- (2) Bachelier, L. (1964). Theory of speculation (1900), reprinted in *The random character of stock market prices* [edited by Cootner, P.], MIT Press, Cambridge, USA, 17-78.
- (3) Baidya, T.K.N. & Costa, P.H.S. (1999). Modelagem de séries financeiras brasileiras: previsão de preços de alguns ativos. *Revista da SOBRAPO*.
- (4) Bekaert, G. & Harvey, C.R. (1997). Emerging markets volatility. *Journal of Financial Economics*.
- (5) Bollerslev, T. (1986). Generalized conditional heteroskedasticity. *Journal of Econometrics*, **31**, 307-328.
- (6) Bollerslev, T.; Ray, Y.C. & Kenneth, F.K. (1992). ARCH Modeling in Finance: A Review of the Theory and Empirical Evidence. *Journal of Econometrics*, **52**, 5-59.
- (7) Brock, W.; Dechert, W. & Scheinkman, J. (1987). A test for independence based on the correlation dimension. Working paper, Univ. of Wisconsin, Univ. of Houston and Univ. of Chicago.
- (8) Campbell, H. (1995). Predictable risk and returns in emerging markets. *Review of Financial Studies*.

- (9) Chu, S. & Freund, S. (1996). Volatility estimation for stock index options: a GARCH approach. *The Quarterly Review of Economics and Finance*, **36**(4), 431-450.
- (10) Clark, P.K. (1973). A subordinated stochastic process model with finite variance for speculative prices. *Econometrica*, **41**, 135-155.
- (11) Cleveland, W.S. & Devlin, S.J. (1988). Locally weighted regression: an approach to regression analysis by local fitting. *Journal of the American Statistical Association*, **83**, 596-610.
- (12) Cooper, J.C.B. (1982). World stock markets: some random walk tests. *Applied Economics*, **14**, 515-531.
- (13) Cunningham, S.W. (1973). The predictability of British stock market prices. *Applied Statistics*, **22**, 315-331.
- (14) Diebold, F. & Nason, J. (1990). Nonparametric exchange rate prediction. *Journal of International Economics*, **28**, 315-332.
- (15) Duarte Jr., A.M.; Heil, T.B.B. & Pinheiro, M.A. (1996). Previsão de volatilidade de ativos e índices brasileiros. *Resenha da BMF*, **112**, 15-27.
- (16) Duarte Jr., A.M. & Mendes, B.V.M.M. (1997). Robust estimation of systematic risk in emerging stock markets. *Emerging Markets Quarterly*, **1**(4), 85-95.
- (17) Embrechts, P.; Resnick, S. & Samorodnitsky, G. (1998). Extreme value theory as a risk management tool. *North American Actuarial Journal*.
- (18) Engel, R.F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, **50**, 987-1007.
- (19) Fama, E.F. (1965). The behaviour of stock market prices. *Journal of Business*, **38**, 34-105.
- (20) Fernandes, M. & Preumont, P.Y. (1996). The size of the BDS test on GARCH standardized residuals. Working paper, Univ. Libre de Bruxelles.
- (21) Granger, C. & Morgenstern, O. (1970). Predictability of stock market prices. *Heath Lexington, Massachusetts, USA*.
- (22) Granger, C.W.J. & Newbold, P. (1977). *Forecasting Economic Time Series*. Academic Press, New York.
- (23) Gray, S.F. (1996). Modeling the conditional distribution of interest rates as a regime-switching process. *Journal of Financial Economics*, **42**, 27-62.
- (24) Hamilton, J.D. & Susmel, R. (1994). Autoregressive conditional heteroskedasticity and changes in regime. *Journal of Econometrics*, **64**, 307-333.
- (25) Harvey, D.; Leybourne, S. & Newbold, P. (1977). Testing the equality of mean squared errors. *International Journal of Forecasting*, **13**, 281-291.
- (26) Hentschel, L. (1995). All in the family: Nesting symmetric and asymmetric GARCH models. *Journal of Financial Economics*, **39**, 71-104.
- (27) Herencia, M.Z.; Hotta, L.K. & Pereira, P.L.V. (1998). Filtragem e previsão com modelos de volatilidade estocástica versus GARCH. *Revista Brasileira de Economia*, **52**(2), 241-278.

-
- (28) Hsieh, D. (1989). Testing for nonlinearity in daily foreign exchange rates. *Journal of Business*, **62**, 339-368.
- (29) Hsieh, D. (1991). Chaos and nonlinear dynamics: application to financial markets. *Journal of Finance*, **46**, 1839-1877.
- (30) Hsieh, D. (1993). Implications of nonlinear dynamics for financial risk management. *Journal of Financial and Quantitative Analysis*, **28**(1), 41-64.
- (31) Issler, J.V. (1999). Estimating and forecasting the volatility of Brazilian finance series using ARCH models. *Revista de Econometria*, **19**(1), 5-56.
- (32) Jarque, C. & Bera, A. (1987). A test for normality of observations and regression residuals. *International Statistical Review*, **55**, 163-172.
- (33) Kendall, M. (1953). The analysis of economic time series, Part I: Price. *Journal of the Royal Statistical Society*, **96A**, 11-25.
- (34) Koopmans, L.H. (1974). *The spectral analysis of time series*. Academic Press Inc., New York.
- (35) Kim, S.; Shephard, N. & Chib, S. (1998). Stochastic volatility: likelihood inference and comparison with ARCH models. *Review of Economic Studies*, **65**, 361-393.
- (36) Li, W.K. & Mark, T.K. (1994). On the squared residual autocorrelation in non-linear time series with conditional heteroskedasticity. *Journal of Time Series Analysis*, **15**, 627-636.
- (37) McNeil, A.J. (1998). Calculating quantile risk measures for financial return series using extreme value theory. Department Mathematic, EHT Zentrum, Zurich.
- (38) Meese, R. & Rose, A. (1991). Nonlinear, nonparametric, nonessential exchange rate estimation. *American Economic Review*, **80**, 192-196.
- (39) Mendes, B.V.M.M. & Duarte Jr., A.M. (1999). Robust estimation of ARCH models. *Revista de Econometria*, **19**(1), 139-180.
- (40) Morgan, W.A. (1939). A test for significance of the difference between two variances in a sample from a normal bivariate population. *Biometrika*, **31**, 13-19.
- (41) Nelson, D.B. (1991). Conditional Heteroskedasticity in Asset Returns: A New Approach. *Econometrica*, **59**, 347-370.
- (42) Osborne, M.M. (1959). Brownian motion in the stock market. *Operations Research*, **7**, 145-173.
- (43) Parkinson, M. (1980). The extreme value method of estimating the variance of the rate of return. *Journal of Business*, **53**, 61-65.
- (44) Pemberton, J. & Tong, H. (1981). A note on the distribution of non-linear autoregressive stochastic models. *Journal of Time Series Analysis*, **2**, 49-52.
- (45) Pereira, P.L.V.; Hotta, L.K.; Souza, L.A.R. & Almeida, N.M.C.G. (1999). Alternative models to extract asset volatility: a comparative study. *Revista de Econometria*, **19**(1), 57-109.
- (46) Praetz, P.D. (1972). The distribution of share price changes. *Journal of Business*, **45**, 49-55.

-
- (47) Stone, C.J. (1977). Consistent nonparametric regressions. *Annals of Statistics*, **5**, 595-620.
- (48) Susmel, R. (1998). Switching volatility in Latin America equity markets. *Emerging Markets Quarterly*, **2**(1), 44-56.
- (49) Taylor, S.J. (1986). *Modelling Financial; Time Series*. John Wiley & Sons, USA.
- (50) Ventura, A.M. (2000). Alguns testes de linearidade para séries temporais nos domínios do tempo e da frequência. Tese de mestrado, IME-USP.
- (51) Working, H. (1934). A random difference series for use in the analysis of time series. *Journal of the American Statistical Association*, **29**, 11-24.
- (52) Zakoian, J.M. (1991). Threshold heteroskedastic models. Não publicado, Institut National de la Statistique et des Etudes Economics, Paris.