

Modelo GARCH com
coeficientes
variando no tempo

Tiago Pilan Ferreira

DISSERTAÇÃO APRESENTADA
AO
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA
DA
UNIVERSIDADE DE SÃO PAULO
PARA
OBTENÇÃO DO TÍTULO
DE
MESTRE EM CIÊNCIAS

Programa: Estatística
Orientadora: Profa. Dra. Chang Chiann

- São Paulo, fevereiro de 2009 -

BRUNO
BRUNO
BRUNO

BRUNO
BRUNO
BRUNO

*À minha mãe e em memória ao meu pai,
Irmãos, familiares,
amigos e professores.*

*“Eu o livrarei, porque a mim se apegou
Eu o protegerei, pois conhece o meu nome
Ele me invocará, e eu responderei
Na angústia estarei com ele
Eu o livrarei e glorificarei
Vou saciá-lo de longos dias
E lhe farei ver a minha salvação”
(Salmo 91, Versículo 14)*

Agradecimentos

Gostaria de agradecer primeiramente a Deus que me deu a vida e possibilitou que eu estivesse vivo até o dia de hoje. Agradeço a Ele por ter me dado forças nos momentos mais difíceis e ter iluminado minha vida em todos os meus caminhos.

À minha mãe que desde a minha infância, com a morte de meu pai, fez o papel de pai e mãe de forma inigualável. Ela sempre buscou me educar da melhor forma possível e ensinou-me alguns dos preceitos mais valiosos que sigo hoje, grande parte do que tenho e sou devo a ela.

Às minhas irmãs que sempre me incentivaram e ajudaram. Obrigada por existirem e serem minhas irmãs.

Aos meus amigos de trabalhos, que por muitas vezes, seguraram as pontas enquanto eu tinha que me ausentar por conta do mestrado. Agradeço em especial a minha chefe, Karin Ayumi Tamura e ao nosso diretor Marcelo Souza por permitirem que eu fizesse este curso.

À minha professora Chang Chiann, pelo jeito atencioso e prestativo com que sempre me tratou, me orientando da forma mais sensata possível nos momentos de dúvida.

Enfim, agradeço a todos que de maneira direta ou indireta me ajudaram na realização deste projeto.

Resumo

O objetivo deste trabalho é propor um novo método para obtenção das estimativas dos coeficientes do modelo GARCH, considerando que eles variam ao longo do tempo. Além da metodologia proposta também foram apresentados modelos GARCH tradicionais com coeficientes fixos no tempo. Ambos os modelos foram testados em 3 séries reais: retornos do Ibovespa, Vale e Itau com o objetivo de verificar se o fato de considerar coeficientes variando no tempo ajuda a melhorar a capacidade preditiva do modelo.

Palavras-chave: coeficientes variando no tempo, GARCH, volatilidade, séries temporais não lineares

Abstract

The goal of this project is to propose a new method for obtaining estimates of the coefficients of the GARCH model, considering that they vary along the time. Beyond the proposed methodology we also adjusted traditional GARCH model with fixed coefficients in time. Both models were tested in 3 real series: Ibovespa, Vale and Itau in order to verify if the fact of considers coefficients varying in time help to improve the predictive ability of the model.

Key-words: coefficients varying in time, GARCH, volatility, non-linear time series

Sumário

1. Introdução	8
2. Preliminares	11
2.1 Séries Temporais Estacionárias e Ergódicas	11
2.2 Processos Lineares	13
2.3 Modelos da família ARIMA	14
2.4 Martingal e Sequência de Diferenças de Martingal	18
2.5 Variância de Longo Prazo	20
2.6 Séries Temporais Não Estacionárias	22
2.7 Retornos.....	23
3. Modelos GARCH	27
3.1 Introdução	27
3.2 Modelos ARCH	28
3.3 O Modelo GARCH e suas Propriedades	30
3.3.1 Representação ARMA do Modelo GARCH	31
3.3.2 Testes para heteroscedasticidade condicional	33
3.3.3 Modelo GARCH e Fatos Estilizados	33
3.4 Estimação de Modelos GARCH.....	34
3.4.1 Estimador da Máxima Verossimilhança Condicional	35
3.4.2 Estimadores de Mínimos Desvios Absolutos	38
3.5 Diagnóstico de Modelos GARCH.....	41
3.6 Extensões dos Modelos GARCH.....	44
3.6.1 Efeitos de Alavancagem Assimétrica e Informações Externas	44
3.6.2 Variáveis Exógenas na Equação Geral da Média Condicional	46
3.6.3 Distribuição de Erros Não-Gaussianos	47
3.7 Seleção e Comparação de Modelos GARCH.....	49
3.8 Predição de Modelos GARCH.....	51
4. Rolling Analysis de Séries Temporais	53

4.1	Introdução	53
4.2	Estatísticas Descritivas	54
4.2.1	Estatísticas univariadas	54
4.2.2	Estatísticas bivariadas	57
4.2.3	Médias móveis ponderadas exponencialmente	58
4.3	<i>Rolling Regression</i>	60
4.4	<i>Rolling Predictions e Backtesting</i>	63
4.4.1	Comparação de modelos	64
5.	Metodologia Proposta	66
5.1	Introdução	66
5.2	Metodologia Utilizada	68
5.3	Previsão utilizando a metodologia proposta	70
5.4	Avaliação do Desempenho do Modelo	73
5.5	Considerações Finais	74
6.	Aplicações em Séries Reais	76
6.1	Introdução	76
6.2	Preços da Ação da Bovespa	76
6.3	Preços da Ação do Itaú (Itau4)	105
6.4	Preços da Ação da Vale (Vale5)	126
6.4	Conclusão	147
7.	Conclusão e Trabalhos Futuros	148
	Apêndice A	150
	Referências Bibliográficas	160

Capítulo 1

Introdução

O processo de desenvolvimento de um modelo para uma série de valores no tempo é composto de diversas etapas, como por exemplo:

- Identificação do modelo;
- Ajuste do modelo;
- Análise dos resíduos para verificação do ajuste do modelo;
- Utilização do modelo na realização de previsões.

Dentre todas essas etapas, uma que chama muito a atenção no mercado financeiro é a referente a previsões. Diferente de outras áreas, no mercado financeiro o principal objetivo do ajuste de modelos é a previsão. Todos os agentes envolvidos no processo estão interessados em saber qual será o comportamento do mercado no curto e no médio prazo de forma à reduzir os riscos e realizar uma alocação eficiente de capital.

A realização de uma boa previsão passa por identificar um bom modelo de modo a conseguir capturar as principais características presentes na série de dados. Por exemplo, sabe-se que modelos lineares do tipo ARIMA não são adequados para capturar a característica de evolução da variância condicional nas séries financeiras. Um modelo não linear bastante utilizado para representar essa evolução na volatilidade das séries temporais financeiras é o modelo GARCH (“Generalized Autoregressive Conditional Heterocedasticity”) de Engle (1982) e Bollerslev (1986).

O mercado financeiro é muito dinâmico e a visão dos agentes envolvidos pode mudar da noite para o dia conforme surjam novos fatos ou haja alteração na expectativa do mercado com relação ao ativo considerado. Dessa forma, uma questão que se coloca é se a utilização de modelos mais dinâmicos, com coeficientes se alterando no tempo, não melhorariam a qualidade do ajuste da série

e por consequência possibilitariam a realização de melhores previsões. Alguns modelos com essa característica são: os modelos *exponenciais autoregressivos* (EXPAR) de Haggan e Ozaki (1981), o modelo *threshold autoregressivo* (TAR) de Tong (1990) e o modelo *funcional autoregressivo* (FAR) de Chen e Tsay (1993). Todos esses modelos são casos particulares de uma classe mais geral de modelos não lineares: "*functional coefficient regression model*". Nestes modelos, os coeficientes são ajustados como função de observações passadas e focam na modelagem da média e não da variância condicional (*volatilidade*).

Esse trabalho tem como objetivo propor um método de ajuste do modelo GARCH, supondo que seus coeficientes se alteram ao longo tempo. Espera-se que essa alteração do modelo GARCH tradicional possibilite uma melhora previsão na variância condicional (*volatilidade*).

No Capítulo 2 apresentamos alguns conceitos básicos que serão utilizados no decorrer de todo o trabalho.

O Capítulo 3 aborda o modelo GARCH e suas extensões. Apresentamos testes para efeito ARCH, métodos de estimação dos parâmetros e algumas ferramentas de diagnóstico do modelo ajustado. Algumas análises pontuais foram realizadas tendo como objeto de estudo o modelo GARCH(1,1), porém a maioria dos resultados apresentados no capítulo são gerais e podem se aplicados à qualquer modelo GARCH.

No Capítulo 4 é introduzido o método "*Rolling Regression*". São apresentados alguns conceitos básicos sobre este método, sua forma de funcionamento, os procedimentos utilizados na estimação dos parâmetros e algumas aplicações.

O método proposto neste trabalho é discutido no Capítulo 5. Apresentamos um pequeno resumo da forma de funcionamento do método, comentamos sobre o processo de estimação dos parâmetros e realizamos uma comparação com outros métodos do gênero focando nas vantagens e desvantagens de cada um.

Aplicações a dados reais podem ser encontradas no Capítulo 6. As séries utilizadas nas aplicações referem-se às ações do Ibovespa, Vale (VALE5) e Itaú

(ITAU4). As aplicações foram feitas utilizando o método tradicional e o método proposto neste trabalho.

Finalmente apresentamos algumas conclusões a respeito do trabalho realizado e pontos que podem ser melhorados, bem como, sugestões para trabalhos futuros.

Capítulo 2

Preliminares

Neste capítulo serão apresentados alguns conceitos e definições que servirão de base para o desenvolvimento do restante dos capítulos. Mais detalhes podem ser encontrados em Morettin e Toloi (2006), Fuller (1996) e Hamilton (1994).

2.1 Séries Temporais Estacionárias e Ergódicas

Um processo estocástico $y = \{y(t, w), t \in T, w \in \Omega\}$ é uma família de variáveis aleatórias (v.a.) definidas num espaço de probabilidade $\{\Omega, \Lambda, P\}$ e indexadas pelos elementos de um conjunto de parâmetros T .

Na realidade, um processo estocástico é uma função de dois argumentos, $y(t, w)$, $t \in T$, $w \in \Omega$. Para cada $t \in T$, a função $y(t, \cdot)$ é mensurável relativamente à Λ . Por outro lado, para cada $w \in \Omega$ fixado, obtemos uma função $y(\cdot, w)$ de $t : T \rightarrow \mathfrak{R}$, que é chamada trajetória, realização, função amostral do processo ou *série temporal*.

Para facilitar a notação, representaremos $y(t, w)$ por y_t .

A série temporal y_t é *estacionária de segunda ordem, fracamente estacionária*, ou ainda, *estacionária em sentido amplo* se

- $E(y_t) = \mu$, constante, para todo t ;
- $\text{cov}(y_t, y_{t-j}) = E[(y_t - \mu, y_{t-j} - \mu)] = \gamma_j$, só depende da defasagem j , para todo t e qualquer j ;
- $E(y_t^2) < \infty$ para qualquer t .

Seja y_t um processo estacionário, o valor γ_j é chamado de *autocovariância* de y_t de j -ésima ordem ou de lag j . A função de autocovariância, γ_j , calculada para todo j , satisfaz as seguintes propriedades

- $\gamma_0 = \text{var}(y_t) > 0$;
- $\gamma_0 \geq |\gamma_j|$;
- $\gamma_{-j} = \gamma_j$;
- γ_j é não negativa definida, no sentido que:

$$\sum_{l=1}^n \sum_{k=1}^n a_l a_k \gamma_{j_l - j_k} \geq 0,$$

para quaisquer números reais a_1, \dots, a_n e j_1, \dots, j_n .

A função de autocorrelação (FAC) de y_t é definida por

$$\rho_j = \frac{\text{cov}(y_t, y_{t-j})}{\sqrt{\text{var}(y_t) \text{var}(y_{t-j})}} = \frac{\gamma_j}{\gamma_0}, \quad j = 0, \pm 1, \dots$$

Intuitivamente, uma série temporal estacionária é definida por sua média, variância, função de autocorrelação e função de distribuição de probabilidade.

Qualquer função de uma série temporal estacionária também é uma série temporal estacionária, isto é, se y_t é estacionária, então $z_t = g(y_t)$ também é estacionária, para qualquer função $g(\cdot)$.

A autocovariância amostral de lag j e a autocorrelação amostral de lag j são definidas por

$$\hat{\gamma}_j = \frac{1}{T} \sum_{t=j+1}^T (y_t - \bar{y})(y_{t-j} - \bar{y}), \quad (2.1)$$

$$\hat{\rho}_j = \frac{\hat{\gamma}_j}{\hat{\gamma}_0}, \quad (2.2)$$

em que $\bar{y} = \frac{1}{T} \sum_{i=1}^T y_i$ é a média amostral.

Dizemos que uma série temporal y_t é *ergódica* se os momentos amostrais convergirem em probabilidade para os momentos populacionais, ou seja, $\bar{y} \xrightarrow{P} \mu$, $\hat{\gamma}_j \xrightarrow{P} \gamma_j$ e $\hat{\rho}_j \xrightarrow{P} \rho_j$.

2.2 Processos Lineares

Seja y_t uma série temporal estacionária. Conforme Zivot (2003), a série y_t pode ser representada por um *processo linear*, ou processo de médias móveis de ordem infinita, da seguinte forma:

$$y_t = \mu + \sum_{k=0}^{\infty} \varphi_k \varepsilon_{t-k}, \quad (2.3)$$

$$\varphi_0 = 1, \quad \sum_{k=0}^{\infty} \varphi_k^2 < \infty,$$

$\varepsilon_t \sim RB(0, \sigma^2)$, em que RB é ruído branco.

Para o processo linear (2.3), pode-se mostrar que:

$$E(y_t) = \mu$$

$$\gamma_0 = \text{var}(y_t) = \sigma^2 \sum_{k=0}^{\infty} \varphi_k^2,$$

$$\gamma_j = \text{cov}(y_t, y_{t-1}) = \sigma^2 \sum_{k=0}^{\infty} \varphi_k \varphi_{k+j},$$

$$\text{e } \rho_j = \frac{\sum_{k=0}^{\infty} \varphi_k \varphi_{k+j}}{\sum_{k=0}^{\infty} \varphi_k^2}.$$

Portanto, as autocorrelações de qualquer processo estacionário e ergódico y_t são determinadas pelos pesos $\{\phi_j\}$ da representação (2.3).

2.3 Modelos da família ARIMA

Modelos AR(p)

O modelo autorregressivo de ordem p (AR(p)) é dado por:

$$y_t - \mu = \phi_1(y_{t-1} - \mu) + \dots + \phi_p(y_{t-p} - \mu) + \varepsilon_t,$$

ou, utilizando-se a notação de operadores,

$$\phi(L)(y_t - \mu) = \varepsilon_t,$$

em que L é o *operador de translação para o passado*, definido como $L^j y_t = y_{t-j}$ e $\phi(L) = 1 - \phi_1 L - \dots - \phi_p L^p$ e $\mu = E(y_t)$.

Pode-se mostrar que um processo AR(p) é sempre invertível, e é estacionário e ergódico sob a condição de que as raízes da *equação característica*,

$$\phi(z) = 1 - \phi_1 z - \phi_2 z^2 - \dots - \phi_p z^p = 0, \quad (2.4)$$

estejam fora do círculo unitário. Uma condição necessária para estacionariedade muito útil na prática é $|\phi_1 + \dots + \phi_p| < 1$.

A função de autocovariância de ordem j de um processo AR(p) estacionário é dada por:

$$\gamma_j = \phi_1 \gamma_{j-1} + \phi_2 \gamma_{j-2} + \dots + \phi_p \gamma_{j-p},$$

com $\gamma_0 = \text{var}(y_t)$.

A função de autocorrelação de ordem j do processo é dada por

$$\rho_j = \phi_1 \rho_{j-1} + \phi_2 \rho_{j-2} + \dots + \phi_p \rho_{j-p},$$

que apresenta comportamento de decaimento exponencial senoidal.

Função de Autocorrelação Parcial

A *função de autocorrelação parcial* (FACP) é um importante instrumento para ajudar na identificação de modelos AR(p). Ela é uma função que associa para cada lag K a correlação parcial entre as variáveis y_t e y_{t-k} ajustadas às variáveis intermediárias $y_{t-1}, \dots, y_{t-k+1}$, ou seja, ela mede a correlação remanescente entre y_t e y_{t-k} depois de eliminada a influência de $y_{t-1}, \dots, y_{t-k+1}$.

A obtenção da FACP baseia-se na estimação da seqüência de modelos AR

$$\begin{aligned} z_t &= \phi_{11} z_{t-1} + \varepsilon_{1t} \\ z_t &= \phi_{21} z_{t-1} + \phi_{22} z_{t-2} + \varepsilon_{2t} \\ &\dots \\ z_t &= \phi_{p1} z_{t-1} + \phi_{p2} z_{t-2} + \dots + \phi_{pp} z_{t-p} + \varepsilon_{pt}, \end{aligned}$$

em que $z_t = y_t - \mu$. Os últimos coeficientes em cada AR(p) (ϕ_{jj} , para $j=1, \dots, p$) são chamados *coeficientes de autocorrelação parcial*.

Para um processo AR(p), o p-ésimo coeficiente de autocorrelação parcial é diferente de zero, e o restante é igual a zero para $j > p$.

Modelos MA(q)

O *modelo de médias móveis de ordem q* (MA(q)) tem a forma:

$$y_t - \mu = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q},$$

em que $\varepsilon_t \sim RB(0, \sigma^2)$.

Em notação de operadores temos

$$(y_t - \mu) = \theta(L)\varepsilon_t,$$

em que $\theta(L) = 1 + \theta_1 L + \dots + \theta_q L^q$.

O modelo MA(q) é estacionário e ergódico desde que $\theta_1, \dots, \theta_q$ sejam finitos, e é invertível se todas as raízes da equação característica,

$$\theta(z) = 1 + \theta_1 z + \dots + \theta_q z^q = 0, \quad (2.5)$$

estiverem fora do círculo unitário.

A função de autocovariância de um processo MA(q) é dada por

$$\gamma_0 = \sigma^2(1 + \theta_1^2 + \dots + \theta_q^2);$$

$$\gamma_j = \begin{cases} \sigma^2(\theta_j + \theta_{j+1}\theta_1 + \theta_{j+2}\theta_2 + \dots + \theta_{q-j}\theta_q), & j = 1, \dots, q \\ 0, & j > q \end{cases}$$

Observa-se que a FAC de um processo MA(q) é diferente de zero até o lag q e igual a zero para $j > q$. Dessa forma, a FAC pode ser usada como um importante instrumento na identificação deste tipo de processo.

A função de autocorrelação do processo pode ser escrita como:

$$\rho_j = \frac{(\theta_j + \theta_{j+1}\theta_1 + \theta_{j+2}\theta_2 + \dots + \theta_{q-j}\theta_q)}{(1 + \theta_1^2 + \dots + \theta_q^2)}.$$

A função de autocorrelação parcial de um processo MA(q) invertível apresenta decaimento exponencial.

Modelos ARMA(p, q)

O modelo *autorregressivo e de médias móveis* (ARMA(p,q)) tem a forma:

$$y_t - \mu = \phi_1(y_{t-1} - \mu) + \dots + \phi_p(y_{t-p} - \mu) + \varepsilon_t + \theta_1\varepsilon_{t-1} + \dots + \theta_q\varepsilon_{t-q}, \quad (2.6)$$

$$\varepsilon_t \sim RB(0, \sigma^2).$$

Pode ser escrito na forma da regressão

$$y_t = c + \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q},$$

ou representado utilizando-se o operador L

$$\phi(L)(y_t - \mu) = \theta(L)\varepsilon_t,$$

e é estacionário e ergódico se as raízes da equação característica $\phi(z) = 0$ estiverem fora do círculo unitário, e é invertível se as raízes da equação característica $\theta(z) = 0$ estiverem fora do círculo unitário. Assume-se a condição de que os polinômios $\phi(z) = 0$ e $\theta(z) = 0$ não tenham fatores em comum ou que se cancelam.

A média de um processo estacionário e ergódico ARMA(p,q) é igual a:

$$\mu = \frac{c}{1 - \phi_1 - \dots - \phi_p},$$

e suas autocovariâncias e autocorrelações satisfazem as seguintes relações recursivas

$$\gamma_j = \phi_1 \gamma_{j-1} + \phi_2 \gamma_{j-2} + \dots + \phi_p \gamma_{j-p}, j > q, \quad (2.7)$$

$$\rho_j = \phi_1 \rho_{j-1} + \phi_2 \rho_{j-2} + \dots + \phi_p \rho_{j-p}, j > q.$$

A forma geral da FAC de um processo ARMA(p,q) é complicada. Para detalhes, ver Hamilton (1994). Em geral, para o processo ARMA(p,q), a FAC tem o mesmo comportamento da FAC de um processo AR(p) para $p > q$, e a FACP tem comportamento parecido com a FACP de um processo MA(q) para $q > p$. Dessa forma, ambas FAC e FACP podem apresentar decaimento exponencial.

Modelos ARIMA(p,d,q)

A especificação do modelo ARMA(p,q) (2.6) assume que y_t é estacionário e ergódico. Se y_t possuir alguma tendência, então podem ser aplicadas algumas transformações com o objetivo de eliminar esta tendência.

O modelo ARIMA (*autorregressivo integrado e de médias móveis*) foi proposto para modelagem de processos nos quais pelo menos uma raiz da equação característica $\phi(z) = 0$ se encontra sobre o círculo unitário. Neste caso o processo é não estacionário na média, o que prejudica a estimação dos parâmetros.

Se houver uma tendência linear em y_t , então a primeira diferença $\Delta y_t = y_t - y_{t-1}$ elimina tal tendência, em que $\Delta = 1 - L$ é o *operador da diferença*. Se houver tendência quadrática em y_t , então haverá tendência linear em Δy_t , mas a segunda diferença $\Delta^2 y_t = (1 - 2L + L^2)y_t = y_t - 2y_{t-1} + y_{t-2}$ não apresentará tendência. O modelo ARIMA(p,d,q) é resultante de d diferenças para eliminar as tendências.

O modelo ARIMA(p,d,q) pode ser representado por

$$\phi(L)\Delta^d(y_t - \mu) = \theta(L)\varepsilon_t,$$

em que $\Delta^d = (1 - L)^d$.

2.4 Martingal e Sequência de Diferenças de Martingal

Seja $\{y_t\}$ uma sequência de variáveis aleatórias e seja $I_t = \{y_t, y_{t-1}, \dots\}$, um conjunto de informações baseadas no passado histórico de $\{y_t\}$. A sequência $\{y_t, I_t\}$ é chamada de um *martingal* se:

- $I_{t-1} \subset I_t$;
- $E[|y_t|] < \infty$;
- $E[y_t | I_{t-1}] = y_{t-1}$ (propriedade martingal).

Um exemplo comum de um martingal é o modelo de passeio aleatório

$$y_t = y_{t-1} + \varepsilon_t, \varepsilon_t \sim RB(0, \sigma^2),$$

em que y_0 é um valor inicial fixado. Se $I_t = \{y_t, \dots, y_0\}$ então $E[y_t | I_{t-1}] = y_{t-1}$, dado que $E[\varepsilon_t | I_{t-1}] = 0$.

Seja $\{\varepsilon_t\}$ uma sequência de variáveis aleatórias com um conjunto de informações associadas I_t . A sequência $\{\varepsilon_t, I_t\}$ é chamada de uma *seqüência de diferenças de martingal (martingale difference sequence (MDS))* se

- $I_{t-1} \subset I_t$;
- $E[\varepsilon_t | I_{t-1}] = 0$ (propriedade MDS)

Se $\{y_t, I_t\}$ é um martingal, um MDS $\{\varepsilon_t, I_t\}$ pode ser construído definindo-se

$$\varepsilon_t = y_t - E[y_t | I_{t-1}].$$

Por construção, um MDS é um processo não correlacionado. Para $k > 0$ temos:

$$E[\varepsilon_t \varepsilon_{t-k}] = E[E[\varepsilon_t \varepsilon_{t-k} | I_{t-1}]] \quad (2.8)$$

$$= E[\varepsilon_{t-k} E[\varepsilon_t | I_{t-1}]]$$

$$= 0.$$

Embora um processo MDS seja não correlacionado, ele não é necessariamente independente, isto é, pode existir dependência em momentos de maior ordem de ε_t .

Processos MDS possuem muitos resultados úteis de convergência (lei dos grandes números, teorema central do limite, etc.). White (1984), Hamilton (1994) e Hayashi (2000) apresentam muitos destes resultados para a análise de séries temporais financeiras.

2.5 Variância de Longo Prazo

Seja y_t uma série temporal estacionária e ergódica. Pelo Teorema Central do Limite

$$\sqrt{T}(\bar{y} - \mu) \xrightarrow{d} N(0, \sum_{j=-\infty}^{\infty} \gamma_j),$$

ou

$$\bar{y} \stackrel{A}{\sim} N(\mu, \frac{1}{T} \sum_{j=-\infty}^{\infty} \gamma_j),$$

em que T é o tamanho da amostra, \bar{y} é a média amostral de y_t e γ_j ($-\infty < j < +\infty$) são as autocovariâncias.

A *variância de longo prazo* (vlp) de y_t é dada pela multiplicação de T pela *variância assintótica* (avar) da média amostral:

$$vlp(y_t) = T \cdot a \text{ var}(\bar{y}) = \sum_{j=-\infty}^{\infty} \gamma_j.$$

Dado que $\gamma_{-j} = \gamma_j$, a variância de longo prazo de y_t pode ser escrita como:

$$vlp(y_t) = \gamma_0 + 2 \sum_{j=1}^{\infty} \gamma_j.$$

Estimando a Variância de Longo Prazo

Se y_t é um processo linear temos que

$$\sum_{j=-\infty}^{\infty} \gamma_j = \sigma^2 \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \phi_k \phi_{k+j} = \sigma^2 \phi(1)^2,$$

e, portanto

$$vlp(y_t) = \sigma^2 \phi(1)^2. \quad (2.9)$$

Se $y_t \sim \text{ARMA}(p,q)$ então:

$$\phi_t(1) = \frac{1 + \theta_1 + \dots + \theta_q}{1 - \phi_1 - \dots - \phi_q} = \frac{\theta(1)}{\phi(1)},$$

e, portanto

$$vlp(y_t) = \frac{\sigma^2 \theta(1)^2}{\phi(1)^2}. \quad (2.10)$$

Uma estimativa consistente para $vlp(y_t)$ pode ser calculada pela substituição de parâmetros estimados do modelo $\text{ARMA}(p,q)$ apropriado em (2.10). De forma alternativa, pode-se aproximar o modelo $\text{ARMA}(p,q)$ por um processo $\text{AR}(p^*)$

$$y_t = c + \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} + \varepsilon_t,$$

em que p^* é escolhido de tal forma que ε_t seja não correlacionado. Assim, a estimativa da *variância de longo prazo autorregressiva* fica dada por

$$vlp_{AR}(y_t) = \frac{\sigma^2}{\phi^*(1)^2} \quad (2.11)$$

Um estimador popularizado por Newey e West (1987)¹ é um estimador ponderado pelas autocovariâncias:

$$\overline{vlp}_{NW}(y_t) = \hat{\gamma}_0 + 2 \sum_{j=1}^{M_T} w_{j,T} \cdot \hat{\gamma}_j,$$

em que $w_{j,T}$ são pesos que somam 1 e M_T é um parâmetro de truncamento do lag que satisfaz $M_T = O(T^{1/3})$.

Para um processo MA(q), $\gamma_j = 0$ para $j > q$. Newey e West (1987) sugerem o uso de pesos *retangulares*:

$$\begin{cases} w_{j,T} = 1, \text{ para } j \leq M_T = q; \\ w_{j,T} = 0, \text{ caso contrário.} \end{cases} \quad (2.12)$$

Para processos lineares em geral, Newey e West (1987) sugerem o uso dos pesos *Bartlett*, $w_{j,T} = 1 - \frac{j}{M_T + 1}$, com $M_T = [4(T/100)^{2/9}]$.

2.6 Séries Temporais Não Estacionárias

Uma vez que um processo estacionário possui momentos variantes, um processo *não estacionário* deve ser considerado. As formas mais comuns de não

¹ A estimativa de Newey-West para a variância de longo prazo será utilizada para o cálculo da Estimativa R/S. Ver subseção 3.3.1.

estacionariedade são causadas pela dependência temporal na média e na variância.

Processos Integrados

Chamamos y_t de um *processo integrado* de ordem 1 ($y_t \sim I(1)$) se possui a forma

$$y_t = y_{t-1} + u_t, \quad (2.13)$$

em que u_t é uma série temporal estacionária. Desta forma, tem-se que a primeira diferença de y_t

$$\Delta y_t = u_t.$$

é estacionária.

Devido a esta propriedade, o processo $I(1)$ é chamado também de *processo de diferença estacionária*.

Uma vez que o processo estacionário u_t não precisa sofrer diferenças, é chamado de processo integrado de ordem zero ($u_t \sim I(0)$).

O processo y_t é $I(d)$ se $\Delta^d y_t \sim I(0)$. Em séries temporais financeiras raramente são modelados processos $I(d)$ com $d > 2$. Além dos processos $I(d)$ já citados, existem outros processos não estacionários (não lineares), como por exemplo, os casos em que d é fracionário.

2.7 Retornos

Para avaliar o risco de uma carteira de ativos no mercado financeiro é comum medir as variações de preços destes ativos. Na prática, é preferível trabalhar com retornos do que com preços pois os retornos são livres de escala e têm propriedades estatísticas interessantes (como estacionariedade e

ergodicidade). Os modelos apresentados neste trabalho têm como objetivo modelar a volatilidade dos retornos de séries financeiras, como retornos de preços de ações, por exemplo.

Seja P_t o preço de um ativo no instante t . A variação de preços entre os instantes $t-1$ e t é dada por $\Delta P_t = P_t - P_{t-1}$ e a variação relativa de preços ou *retorno simples* deste ativo é definido por:

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}} = \frac{\Delta P_t}{P_{t-1}}, \quad (2.14)$$

De (2.14) temos que $R_t = \frac{P_t}{P_{t-1}} - 1$. Chamamos $1 + R_t = \frac{P_t}{P_{t-1}}$ de *retorno bruto simples* ou *taxa de retorno*.

Definimos o *retorno composto continuamente*, ou *log-retorno* ou simplesmente *retorno*, como:

$$r_t = \ln \frac{P_t}{P_{t-1}} = \ln(1 + R_t) = \ln(P_t) - \ln(P_{t-1}) = p_t - p_{t-1}, \quad (2.15)$$

em que $p_t = \ln(P_t)$.

Podemos definir também retornos multiperíodos. Dado um período k , o retorno simples entre os instantes $t-k$ e t é dado por:

$$R_t(k) = \frac{P_t - P_{t-k}}{P_{t-k}}. \quad (2.16)$$

Em termos de retornos de um mesmo período podemos escrever

$$\begin{aligned} 1 + R_t(k) &= (1 + R_t)(1 + R_{t-1}) \dots (1 + R_{t-k+1}) \\ &= \frac{P_t}{P_{t-1}} \frac{P_{t-1}}{P_{t-2}} \dots \frac{P_{t-k+1}}{P_{t-k}} = \frac{P_t}{P_{t-k}}, \end{aligned}$$

portanto temos

$$R_t(k) = \frac{P_t}{P_{t-k}} - 1. \quad (2.17)$$

O log-retorno de período k fica dado por:

$$r_t(k) = \ln\left(\frac{P_t}{P_{t-k}}\right) = \ln(1 + R_t(k)) = p_t - p_{t-k}. \quad (2.18)$$

Se houver pagamento de dividendos D_t no período, o retorno simples (R_t) e o log-retorno (r_t) são dados por:

$$R_t = \frac{P_t + D_t}{P_{t-1}} - 1, \quad (2.19)$$

e

$$r_t = \ln(1 + R_t) = \ln(P_t + D_t) - \ln(P_{t-1}). \quad (2.20)$$

Fatos Estilizados sobre os Retornos

Séries financeiras apresentam algumas características que são comuns a outras séries temporais, como tendências, sazonalidade, pontos influentes (atípicos), heteroscedasticidade condicional e não-linearidade. Os retornos financeiros, por outro lado, apresentam outras características peculiares que outras séries não apresentam. Retornos raramente apresentam tendências ou sazonalidade, com exceção eventualmente de retornos intra-diários, e séries de taxas de câmbio ou taxa de juros que podem apresentar tendências que variam com o tempo.

Podemos resumir os principais *fatos estilizados* relativos a retornos financeiros:

1. retornos são em geral não autocorrelacionados;
2. os quadrados dos retornos são autocorrelacionados;
3. séries de retornos apresentam agrupamentos de volatilidades ao longo do tempo;
4. a distribuição (não-condicional) dos retornos apresenta caudas mais pesadas do que uma distribuição normal;
5. algumas séries de retornos são não-lineares (respondem de maneira diferente a choques grandes ou pequenos, ou a choques negativos ou positivos).

Capítulo 3

Modelos GARCH

3.1 Introdução

Para os agentes envolvidos no mercado financeiro, o grande objetivo é o lucro, sejam estes agentes pessoas físicas interessadas em aumentar seu patrimônio, garantir uma melhor aposentadoria ou simplesmente buscar alternativas mais rentáveis do que as presentes hoje, sejam eles pessoas jurídicas, como por exemplo, fundos de pensão interessados em ter opções de investimento mais atrativas.

Para que o lucro seja alcançado e, além disso, seja maximizado, é fundamental que seja realizado um gerenciamento eficiente de alguns aspectos do negócio, como por exemplo:

- gerenciamento eficiente de risco;
- elaboração seletiva de portfólio;
- etc...

Uma boa gestão dos aspectos citados acima está intimamente relacionada com o entendimento e controle da volatilidade presente no mercado. Os ganhos ou perdas no mercado são obtidos através das variações ocorridas nos preços dos ativos, estas variações são influenciadas pela situação macroeconômica do país e do mundo, pela situação da empresa responsável pelo ativo e pela percepção dos agentes envolvidos com relação ao potencial do ativo. Todos esses fatores tornam o mercado financeiro muito volátil, e é através dessa volatilidade que as oportunidades de lucro se apresentam. Por isso é fundamental entendê-la. Em termos estatísticos, a volatilidade é medida pela variância, ou pelo desvio padrão dos preços do ativo.

Nesse capítulo vamos introduzir a classe dos modelos GARCH (*Generalized Autoregressive Conditional Heteroskedasticity*) desenvolvidos por Engle (1982), Bollerslev (1986), Nelson (1991) e outros, que são capazes de modelar a volatilidade e de capturar muitos dos fatos estilizados do comportamento da volatilidade observada em séries temporais financeiras.

3.2 Modelos ARCH

Na Figura 3.1 temos os retornos diários da série de preços da ação da Vale, no período de 19 de agosto de 1994 a 23 de janeiro de 2009, totalizando 1.147 observações. Pode-se perceber que tanto as pequenas mudanças quanto as grandes mudanças parecem estar agrupadas. Essa tendência é típica de muitas séries temporais macroeconômicas e financeiras, ou seja, quando ocorrem grandes mudanças no mercado, os agentes ficam em alerta o que implica em maior volatilidade. Para confirmar este comportamento são apresentadas na Figura 3.2 as funções de autocorrelação dos retornos e dos retornos ao quadrado da série de retornos da Vale.

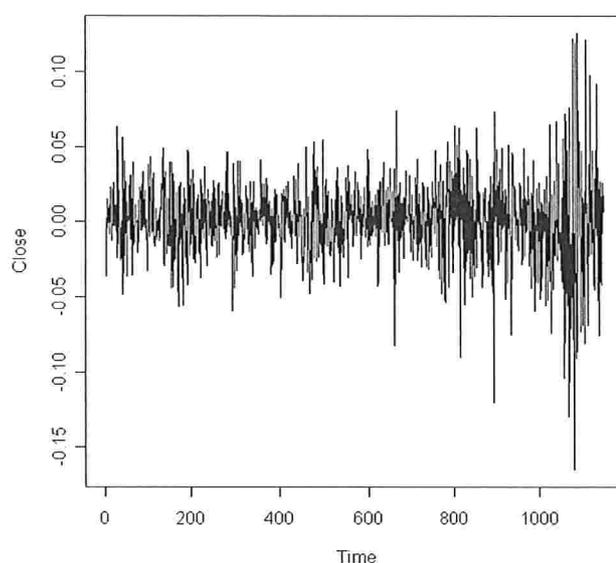


Figura 3.1: Retornos da série de preços diários das ações da Vale referentes ao período de 19 de agosto de 1994 à 23 de janeiro de 2009.

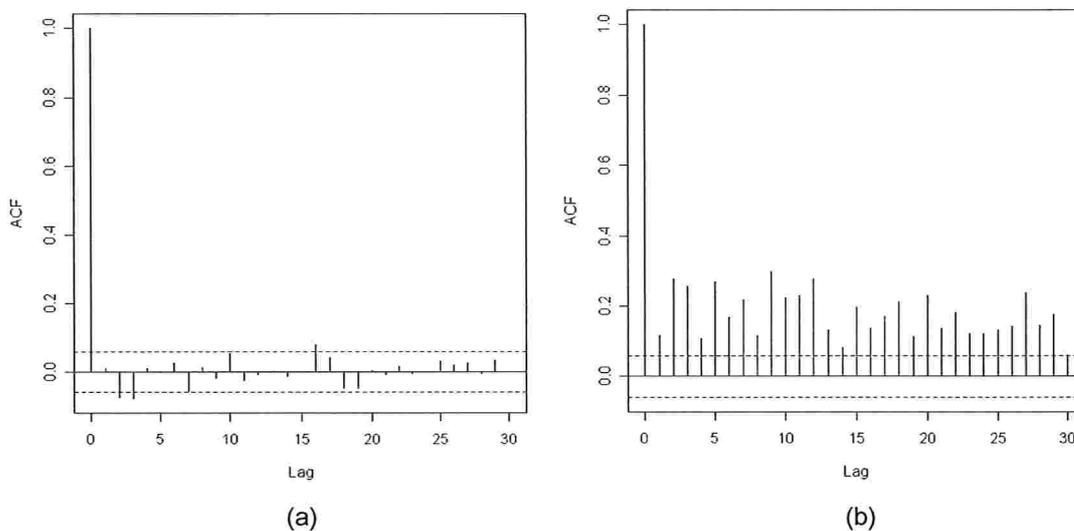


Figura 3.2: (a) FAC dos retornos e (b) FAC do quadrado dos retornos da série de preços da Vale

Observando os gráficos das funções de autocorrelação percebemos que há autocorrelações significantes apenas para alguns lags da série de retornos. Por outro lado, a autocorrelação da série dos retornos ao quadrado se mostra significativa pelo menos até o lag 30. Uma vez que os retornos ao quadrado medem o momento de segunda ordem da série original, este resultado indica que a variância dos retornos da série de preços da Vale condicionada ao passado histórico pode apresentar alterações ao longo do tempo, ou seja, a série de retornos pode mostrar a presença de *heteroscedasticidade condicional ao tempo*.

A heteroscedasticidade condicional da série dos retornos ao quadrado pode ser modelada usando um processo autoregressivo simples (AR) para resíduos quadráticos. Seja y_t uma série temporal estacionária tal como retornos financeiros, então y_t pode ser expresso como sua média acrescida de um ruído branco, caso não exista autocorrelação significativa em y_t :

$$y_t = c + \varepsilon_t, \quad (3.1)$$

em que c é a média de y_t e ε_t é *iid* com média zero.

Para incluir o efeito da heteroscedasticidade condicional, $Var_{t-1}(\varepsilon_t) = \sigma_t^2$, em que $Var_{t-1}(\cdot)$ é a variância condicional com informação até o tempo $t-1$, e

$$\sigma_t^2 = a_0 + a_1\varepsilon_{t-1}^2 + \dots + a_p\varepsilon_{t-p}^2, \quad (3.2)$$

em que p indica o lag mais distante do instante t , de modo que as informações em p ainda influenciam σ_t^2 .

Se ε_t tiver média zero, $Var_{t-1}(\varepsilon_t) = E_{t-1}(\varepsilon_t^2) = \sigma_t^2$, a equação acima pode ser reescrita como:

$$\varepsilon_t^2 = a_0 + a_1\varepsilon_{t-1}^2 + \dots + a_p\varepsilon_{t-p}^2 + u_t, \quad (3.3)$$

em que $u_t = \varepsilon_t^2 - E_{t-1}(\varepsilon_t^2)$ é um processo diferença de martingal com média zero, mas com variância não constante.

A equação (3.3) representa um processo AR(p) para ε_t^2 , e o modelo em (3.1) e (3.2) é conhecido como o modelo *autorregressivo com heteroscedasticidade condicional* de Engle (1982), que é usualmente chamado de modelo ARCH(p).

Uma formulação alternativa para o modelo ARCH(p) é

$$\begin{aligned} y_t &= c + \varepsilon_t \\ \varepsilon_t &= z_t \sigma_t \\ \sigma_t^2 &= a_0 + a_1\varepsilon_{t-1}^2 + \dots + a_p\varepsilon_{t-p}^2, \end{aligned}$$

em que z_t é uma variável aleatória *iid* com uma distribuição especificada. No modelo ARCH básico a distribuição de z_t é a normal padrão *iid*.

3.3 O Modelo GARCH e suas Propriedades

Conforme apresentado anteriormente, uma forma de modelar a volatilidade σ_t é a utilização de modelos ARCH. No entanto, tem-se que na prática geralmente é preciso um número grande de lags p para se obter um bom ajuste do modelo, o que implica em um número grande de parâmetros. Um modelo mais parcimonioso foi proposto por Bollerslev (1986) e substitui o modelo AR em (3.2) pela formulação:

$$\sigma_t^2 = a_0 + \sum_{i=1}^p a_i \varepsilon_{t-i}^2 + \sum_{j=1}^q b_j \sigma_{t-j}^2, \quad (3.4)$$

em que os coeficientes a_i ($i = 0, \dots, p$) e b_j ($j = 1, \dots, q$) são todos positivos para garantir que a variância condicional σ_t^2 seja sempre positiva². O modelo (3.4) junto com (3.1) é conhecido como o modelo GARCH(p, q) ("*generalized ARCH*"). Quando $q=0$, o modelo GARCH se reduz ao modelo ARCH.

No modelo GARCH(p, q) a variância condicional de ε_t , σ_t^2 , depende do quadrado dos resíduos nos p períodos anteriores, e da variância condicional nos q períodos anteriores.

3.3.1 Representação ARMA do Modelo GARCH

Assim como um modelo ARCH pode ser escrito como um modelo AR em termos de resíduos quadráticos, um modelo GARCH pode ser escrito na forma de um modelo ARMA de resíduos quadráticos. Considere o modelo GARCH(1, 1):

$$\sigma_t^2 = a_0 + a_1 \varepsilon_{t-1}^2 + b_1 \sigma_{t-1}^2, \quad (3.5)$$

sendo $E_{t-1}(\varepsilon_t^2) = \sigma_{t-1}^2$. A equação acima pode ser escrita como:

² Coeficientes positivos são condições suficientes mas não necessárias para garantir que a variância condicional seja positiva. Ver Nelson e Cao (1992).

$$\varepsilon_t^2 = a_0 + (a_1 + b_1)\varepsilon_{t-1}^2 + u_t - b_1u_{t-1}, \quad (3.6)$$

que é um modelo ARMA(1,1) em que $u_t = \varepsilon_t^2 - E_{t-1}(\varepsilon_t^2)$ é ruído branco de média zero e variância não constante.

Dada a representação ARMA de um modelo GARCH, muitas propriedades do modelo GARCH seguem facilmente da correspondência de um processo ARMA para ε_t^2 . Por exemplo, para o modelo GARCH(1,1) ser estacionário é preciso que $a_1 + b_1 < 1$. Assumindo a estacionariedade do modelo GARCH(1,1), pode-se mostrar que a variância incondicional de ε_t é dada por $Var(\varepsilon_t) = E(\varepsilon_t^2) = a_0/(1 - a_1 - b_1)$, pois de (3.6) temos:

$$E(\varepsilon_t^2) = a_0 + (a_1 + b_1)E(\varepsilon_{t-1}^2),$$

e portanto

$$E(\varepsilon_t^2) = a_0 + (a_1 + b_1)E(\varepsilon_t^2),$$

baseado na suposição de que ε_t^2 é estacionário.

Para o modelo GARCH(p,q) (3.4), os resíduos quadráticos ε_t^2 têm comportamento de um processo ARMA(max(p,q),q). A estacionariedade requer que

$\sum_{i=1}^p a_i + \sum_{j=1}^q b_j < 1$ e a variância incondicional de ε_t é dada por

$$\bar{\sigma}^2 = Var(\varepsilon_t) = \frac{a_0}{1 - (\sum_{i=1}^p a_i + \sum_{j=1}^q b_j)}. \quad (3.7)$$

3.3.2 Testes para heteroscedasticidade condicional

Antes de estimar um modelo ARCH para uma série temporal financeira, é importante testar a presença de correlação nos resíduos ao quadrado. Se não houver efeitos de correlação nos resíduos ao quadrado, então o modelo ARCH é desnecessário e mal especificado.

Um teste muito utilizado para avaliar a correlação entre os resíduos ao quadrado é o teste de *Box-Pierce-Ljung* (Box e Pierce, 1970). Apesar deste teste não detectar quebras específicas no comportamento de ruído branco, pode indicar se esses valores são muito altos. Uma modificação deste teste foi proposta por Ljung e Box (1978) e é apresentada a seguir:

Se modelo for apropriado, a estatística:

$$Q(K) = n(n+2) \sum_{k=1}^K \frac{\hat{r}_k^2}{(n-k)},$$

terá uma distribuição de χ^2 com $K-p-q$ graus de liberdade, em que, p e q indicam as ordens do modelo GARCH(p,q) e K refere-se as primeiras K correlações amostrais. A hipótese de ruído branco para os resíduos é rejeitada para valores altos de $Q(\cdot)$.

Outro teste que pode ser utilizado para verificar se a série apresenta heteroscedasticidade condicional é o teste de Multiplicador de Lagrange.

3.3.3 Modelo GARCH e Fatos Estilizados

Na prática, existem alguns “fatos estilizados” sobre a volatilidade em séries temporais financeiras. Bollerslev, Engle e Nelson (1994) dão uma visão completa sobre esses fatos. Usando a representação ARMA dos modelos GARCH é possível mostrar que o modelo GARCH é capaz de explicar muitos desses fatos estilizados. Nessa seção serão detalhados dois importantes fatos estilizados: agrupamento de volatilidades (*volatility clustering*) e caudas pesadas.

Agrupamento de Volatilidades

No modelo GARCH(1,1) dado em (3.5) o coeficiente b_1 encontrado é normalmente um número próximo de 0.9 para séries temporais financeiras diárias ou semanais. Dessa maneira, temos que valores grandes de σ_{t-1}^2 são seguidos por valores grandes de σ_t^2 , e valores pequenos de σ_{t-1}^2 , por sua vez, são seguidos por valores pequenos de σ_t^2 . Seguindo o mesmo raciocínio para a representação ARMA em (3.6), temos que grandes/pequenas mudanças em ε_{t-1}^2 são seguidas por grandes/pequenas mudanças em ε_t^2 .

Caudas pesadas

É fato conhecido que a distribuição de muitas séries temporais financeiras, geralmente possui caudas mais pesadas que a distribuição normal. Ou seja, grandes mudanças ocorrem com mais frequência do que em uma distribuição normal. Bollerslev (1986) descreve as condições para a existência do momento de quarta ordem de um processo GARCH(1,1). Assumindo que o momento de quarta ordem existe, Bollerslev (1986) mostra que a curtose de um processo GARCH(1,1) é maior do que 3, que é a curtose de uma distribuição normal. He e Terasvirta (1999a, 1999b) estenderam o resultado de Bollerslev para os modelos GARCH(p,q). Portanto, um modelo GARCH pode reproduzir as caudas pesadas observadas em séries temporais financeiras.

3.4 Estimação de Modelos GARCH

O objetivo dessa seção é estimar um modelo GARCH. Conforme apresentado anteriormente, o modelo geral GARCH(p,q) é da forma:

$$\begin{aligned}y_t &= c + \varepsilon_t, \\ \varepsilon_t &= z_t \sigma_t,\end{aligned}\tag{3.8}$$

$$\sigma_t^2 = a_0 + \sum_{i=1}^p a_i \varepsilon_{t-i}^2 + \sum_{j=1}^q b_j \sigma_{t-j}^2,$$

para $t = 1, \dots, T$, em que $\sigma_t^2 = \text{Var}_{t-1}(\varepsilon_t)$.

Serão introduzidos dois tipos de estimadores para os parâmetros a_0 , a_i e b_j : *estimador de máxima verossimilhança condicional* e *estimador de mínimos desvios absolutos*. O primeiro é o estimador mais conhecido e é o que vem sendo bastante usado na área de operações financeiras. O segundo é interessante quando tratamos de erros com caudas pesadas. Nas simulações utilizaremos apenas o *estimador de máxima verossimilhança condicional*.

3.4.1 Estimador da Máxima Verossimilhança Condicional

Assim como na estimação dos modelos ARMA, os estimadores mais usados para modelos ARCH/GARCH são aqueles derivados de uma função Gaussiana de máxima verossimilhança (condicional). Por exemplo, se z_t em (3.8) é normal e $q=0$ (ou seja, um modelo ARCH puro), a log-verossimilhança condicional negativa baseada nas observações $\varepsilon_1, \dots, \varepsilon_T$, ignorando-se as constantes, é igual a

$$-l(\sigma_t^2 / \varepsilon_1, \dots, \varepsilon_T) = \sum_{t=p+1}^T (\log \sigma_t^2 + \varepsilon_t^2 / \sigma_t^2), \quad (3.9)$$

em que $\sigma_t^2 = a_0 + \sum_{i=1}^p a_i \varepsilon_{t-i}^2$. Os estimadores de máxima verossimilhança são obtidos minimizando-se a função em (3.9). Podemos perceber que esta função de verossimilhança é baseada na função de densidade de probabilidades condicional de $\varepsilon_{p+1}, \dots, \varepsilon_T$, dados $\varepsilon_1, \dots, \varepsilon_p$, uma vez que a função de densidade de probabilidades incondicional, que envolve a densidade conjunta de $\varepsilon_1, \dots, \varepsilon_p$, é difícil de ser obtida.

Para o modelo GARCH geral (ou seja, $q > 0$ no modelo (3.4)) a variância condicional σ_t^2 não pode ser expressa em termos de um número finito de observações passadas $\varepsilon_{t-1}, \varepsilon_{t-2}, \dots$. Por indução, podemos escrever:

$$\sigma_t^2 = \frac{a_0}{1 - \sum_{j=1}^q b_j} + \sum_{i=1}^p a_i \varepsilon_{t-i}^2 + \sum_{i=1}^p a_i \sum_{k=1}^{\infty} \sum_{j_1=1}^q \dots \sum_{j_k=1}^q b_{j_1} \dots b_{j_k} \varepsilon_{t-i-j_1-\dots-j_k}^2,$$

(3.10)

em que a soma múltipla desaparece quando $q=0$. Nota-se que a soma múltipla acima converge com probabilidade 1 desde que cada a_i e b_j sejam não negativos, e desde que o valor esperado da série múltipla seja finito. Na prática, a expressão (3.10) é substituída por uma versão truncada:

$$\tilde{\sigma}_t^2 = \frac{a_0}{1 - \sum_{j=1}^q b_j} + \sum_{i=1}^p a_i \varepsilon_{t-i}^2 + \sum_{i=1}^p a_i \sum_{k=1}^{\infty} \sum_{j_1=1}^q \dots \sum_{j_k=1}^q b_{j_1} \dots b_{j_k} x \varepsilon_{t-i-j_1-\dots-j_k}^2 I(t-i-j_1-\dots-j_k \geq 1), t > p$$

(3.11)

Temos que, quando $q=0$, $\tilde{\sigma}_t^2 = \sigma_t^2 = a_0 + \sum_{i=1}^p a_i \varepsilon_{t-i}^2$. Seja $a = (a_1, \dots, a_p)'$ e $b = (b_1, \dots, b_q)'$. O estimador de máxima verossimilhança (condicional) $(\hat{a}_0, \hat{a}, \hat{b})$ é definido, minimizando-se a expressão:

$$l_v(a_0, a, b) = \sum_{t=v}^T (\log \tilde{\sigma}_t^2 + \varepsilon_t^2 / \tilde{\sigma}_t^2),$$

(3.12)

em que $v > p$ é um número inteiro.

Supondo que $f(\cdot)$ seja a função densidade de probabilidades de z_t , conhecida, temos que os estimadores de máxima verossimilhança são obtidos minimizando-se:

$$l_v(a_o, a, b) = \sum_{i=v}^T \{ \log \tilde{\sigma}_i^2 + 2 \log f(\varepsilon_i / \tilde{\sigma}_i) \}, \quad (3.13)$$

em vez de minimizar a função (3.12).

Além da distribuição normal, algumas formas de $f(\cdot)$ frequentemente usadas são:

- distribuição t com ν graus de liberdade:

$$f(x) = \frac{\Gamma((\nu+1)/2)}{(\pi\nu)^{1/2} \Gamma(\nu/2)} \left(\frac{\nu}{\nu-2} \right)^{1/2} \left(1 + \frac{x^2}{\nu-2} \right)^{-\frac{(\nu+1)}{2}},$$

em que $\nu > 2$ pode ser tratado como um parâmetro contínuo.

- distribuição gaussiana generalizada:

$$\lambda = \nu \{ \lambda 2^{1+1/\nu} \Gamma(1/\nu) \}^{-1} \exp \left\{ -\frac{1}{2} \left| \frac{x}{\lambda} \right|^\nu \right\},$$

em que $\lambda = \{ 2^{-2/\nu} \Gamma(1/\nu) / \Gamma(3/\nu) \}^{1/2}$ e $0 < \nu < 2$.

Quando $\nu = 1$, a distribuição Gaussiana generalizada se reduz para a função $f(x) = \exp\{-\sqrt{2}|x|\} / \sqrt{2}$.

Todas as distribuições acima foram normalizadas para ter média 0 e variância 1, e todas elas possuem caudas mais pesadas do que a distribuição normal.

Para comparar dois ou mais modelos ajustados, usamos os valores AIC e BIC para cada um dos modelos ajustados, que são calculados da seguinte maneira:

$$AIC = l_v(\hat{a}_0, \hat{a}, \hat{b}) + 2(p+q+1), \quad (3.14)$$

e

$$BIC = l_v(\hat{a}_0, \hat{a}, \hat{b}) + (p + q + 1) \log(T - v + 1), \quad (3.15)$$

em que $l_v(\cdot)$ segue de (3.13).

3.4.2 Estimadores de Mínimos Desvios Absolutos

Os estimadores apresentados anteriormente são derivados da maximização da verossimilhança Gaussiana ou de uma aproximação da verossimilhança Gaussiana. Dessa forma, estes são L_2 -estimadores. Sabe-se que os L_1 -estimadores são mais robustos do que os L_2 -estimadores, no que diz respeito às distribuições com caudas pesadas. Algumas evidências empíricas sugerem que este método é mais apropriado para séries financeiras pois estas apresentam caudas mais pesadas que a distribuição normal. Para detalhes ver Mandelbrot (1963), Fama (1965), Rachev and Paoletta (1998), e Mittnik e Rachev (2000).

A idéia deste estimador implica numa reparametrização do modelo (3.4), em que $E(z_t) = 0$ e a mediana de z_t^2 , ao invés da variância de z_t , é igual a 1. Sob essa nova formulação, os parâmetros a_0 e a_i 's diferem daqueles da formulação original por um fator constante, enquanto que os parâmetros b_j 's não se alteram.

Seja:

$$\varepsilon_t^2 / \sigma_t^2 = 1 + e_{t,1}, \quad (3.16)$$

em que $e_{t,1} = (z_t^2 - 1)$, que possui mediana igual a zero. Assim, temos que o estimador de mínimos desvios absolutos $\hat{\theta}_1$, que é um L_1 -estimador baseado na regressão (3.16) pode ser obtido minimizando-se

$$\sum_{t=v}^T |\varepsilon_t^2 / \tilde{\sigma}_t^2 - 1|, \quad (3.17)$$

em que $\tilde{\sigma}_i^2$ é definido em (3.11) e $\nu = p + 1$ se $q=0$ e $\nu > p + 1$ se $q>0$. Peng e Yao (2003) mostraram que este estimador é viesado. Para resolver este problema, Peng e Yao (2003) definiram uma forma modificada para o estimador de mínimos desvios absolutos (θ_2), que resulta da minimização da expressão:

$$\sum_{i=\nu}^T \left| \log(\varepsilon_i^2) - \log(\tilde{\sigma}_i^2) \right|, \quad (3.18)$$

motivados pelo modelo de regressão:

$$\log(\varepsilon_i^2) = \log(\sigma_i^2) + e_{i,2}, \quad (3.19)$$

em que os erros $e_{i,2} = \log(z_i^2)$, são i.i.d com mediana zero.

Peng e Yao (2003) demonstraram que sob condições muito amenas, os estimadores de mínimos desvios absolutos são assintoticamente normais com razão de convergência padrão $T^{1/2}$, independentemente do fato de a distribuição de z_i possuir caudas pesadas ou não. Esta é uma diferença marcante entre estes estimadores e os estimadores de máxima verossimilhança condicional derivados de (3.12), que podem apresentar convergência lenta quando z_i possui caudas pesadas.

Outro estimador proposto por Peng e Yao (2003) foi motivado pela equação de regressão:

$$\varepsilon_i^2 = \sigma_i^2 + e_{i,3}, \quad (3.20)$$

em que $e_{i,3} = \sigma_i^2(z_i^2 - 1)$, possui mediana igual a zero. Assim, o estimador ($\hat{\theta}_3$), é obtido minimizando-se:

$$\sum_{i=v}^T |\varepsilon_i^2 - \tilde{\sigma}_i^2|. \quad (3.21)$$

Intuitivamente, é preferível usar o estimador θ_2 ao θ_3 pois os termos de erros $e_{i,2}$ do modelo de regressão (3.19) são independentes e identicamente distribuídos enquanto que os erros $e_{i,3}$ do modelo (3.20) não são independentes.

Em Peng e Yao (2003) foram realizadas comparações entre os três estimadores de mínimos desvios absolutos com o estimador de máxima verossimilhança Gaussiana, $\hat{\theta}_{ML}$, para modelos simulados ARCH(2) e GARCH(1,1), tomando os erros z_i com distribuição normal padrão ou t de Student padronizada com $d=3$ ou $d=4$ graus de liberdade. Foram realizadas 500 réplicas de tamanho 300, com parâmetros: *constante* = 0.3, $a_1 = 0.5$, $a_2 = b_1 = 0.4$ e $v = 20$. A Figura 3.3 apresenta os boxplots dos erros absolutos médios destas simulações. Pode-se perceber que, para modelos com erros de caudas muito pesadas, ou seja, $u_i \sim t_3$, o estimador de mínimos desvios absolutos $\hat{\theta}_2$, obteve a melhor performance. Para $u_i \sim t_4$ podemos perceber que os estimadores $\hat{\theta}_2$ e $\hat{\theta}_{ML}$ possuem comportamento parecido, e são melhores que os estimadores $\hat{\theta}_1$ e $\hat{\theta}_3$. E no caso em que u_i é normal, temos que $\hat{\theta}_{ML}$ possui a melhor performance. De fato, a performance do estimador de máxima verossimilhança Gaussiana piora conforme as caudas da distribuição dos erros tendem a ser mais pesadas. Entretanto, esse comportamento não ocorre para os estimadores de mínimos desvios absolutos uma vez que estes apresentam-se mais robustos para caudas pesadas.

Para mais detalhes, ver Peng e Yao (2003).

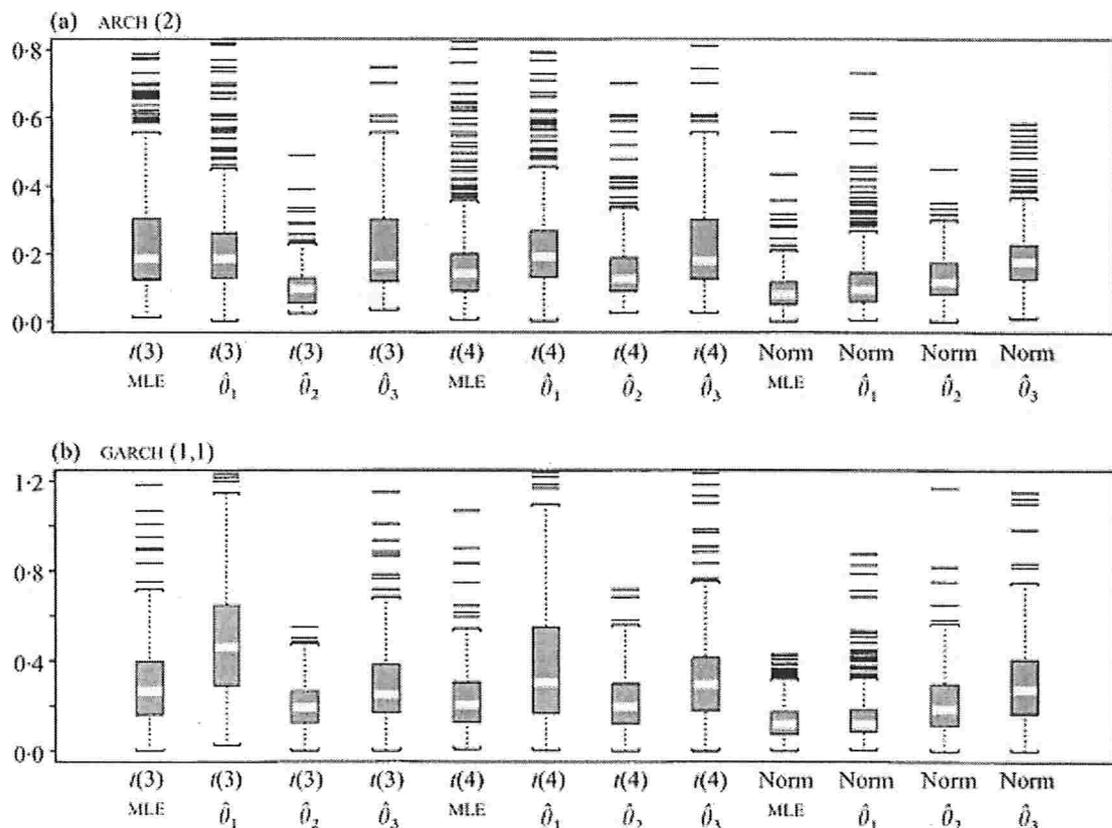


Figura 3.3: Boxplots dos erros absolutos médios das estimativas de máxima verossimilhança, MLE, e das estimativas de mínimos desvios absolutos, $\hat{\theta}_1$, $\hat{\theta}_2$ e $\hat{\theta}_3$. As distribuições dos erros indicadas por $t(3)$, $t(4)$ e Norm são, respectivamente, as distribuições t de Student com 3 e 4 graus de liberdade e a distribuição normal.

3.5 Diagnóstico de Modelos GARCH

As seções anteriores apresentaram métodos para estimação dos modelos GARCH. Para avaliar os modelos ajustados são usados basicamente dois métodos: análise gráfica e análise de algumas estatísticas resumo.

Na análise das estatísticas resumo, podemos observar os erros padrão e os níveis descritivos (p valores) das estatísticas para avaliar se os coeficientes do modelo são estatisticamente iguais a zero. Além disso, existem testes variados para os resíduos padronizados $\hat{\varepsilon}_t / \hat{\sigma}_t$. Para a série de retornos dos preços de ações da Vale, as Tabelas 3.1 e 3.2 apresentam os coeficientes ajustados por meio do método de máxima verossimilhança condicional e alguns testes para resíduos

padronizados do modelo GARCH(1,1): testes de normalidade Jarque-Bera e Shapiro-Wilks, teste de Ljung-Box para resíduos padronizados e para o quadrado dos resíduos padronizados.

No teste Ljung-Box para os resíduos padronizados, não rejeitamos a hipótese nula H_0 da não existência de autocorrelação. O mesmo comportamento pode ser observado no teste Ljung-Box para o quadrado dos resíduos padronizados. Portanto, o modelo capturou com sucesso a estrutura de correlação serial tanto na variância condicional como na média condicional.

Tabela 3.1: Coeficientes estimados do modelo GARCH(1,1)

Coeficiente	Valor			
	Estimado	Erro Padrão	Valor t	Pr(> t)
C	0,0015960	0,0006413	2,488	0,0128000
A	0,0000145	0,0000061	2,387	0,0170000
ARCH(1)	0,0775400	0,0184700	4,199	0,0000268
GARCH(1)	0,9006000	0,0249600	36,089	0,0000000

O modelo GARCH básico assume distribuição normal para os erros ε_t . Se o modelo está corretamente especificado então os resíduos padronizados $\hat{\varepsilon}_t / \hat{\sigma}_t$ devem se comportar como uma variável aleatória normal padrão. Os testes de normalidade Jarque-Bera e Shapiro-Wilks da Tabela 3.2 levam a conclusão de que os dados parecem destoar da distribuição normal. Uma outra ferramenta que pode ser utilizada na avaliação da normalidade é gráfico QQ-plot³ (quantis da distribuição gaussiana x quantis dos resíduos padronizados). Para a série da Vale, temos como resultado o gráfico QQ-plot apresentado na Figura 3.4. Podemos perceber que nas caudas há um significativo desvio da linha QQ normal, e portanto a suposição de normalidade dos resíduos não é apropriada.

³ QQ-plot de y_t é o gráfico de dispersão dos quantis padronizados da distribuição empírica de y_t versus os quantis da distribuição normal padrão. Se y_t seguir uma distribuição normal, então os pontos do gráfico serão marcados sobre a linha de inclinação de 45 graus.

Tabela 3.2: Testes para diagnóstico do modelo GARCH(1,1) ajustado

Teste de Normalidade	
Jarque-Bera	P-valor
68,97517	9,99201E-16
Shapiro-Wilk	P-valor
0,9898691	4,16571E-07

Teste de Ljung-Box para os resíduos padronizados:	
Estatística	P-valor
18,44	0,55813

Teste de Ljung-Box para o quadrado dos resíduos padronizados:	
Estatística	P-valor
17,13	0,64459

Outros gráficos também podem ser utilizados para visualizar o ajuste do modelo. Por exemplo, na Figura 3.5 apresentamos a FAC dos resíduos padronizados e do quadrado dos resíduos padronizados. Observa-se que a autocorrelação foi removida, ou seja, o modelo proposto ajustou bem aos dados.

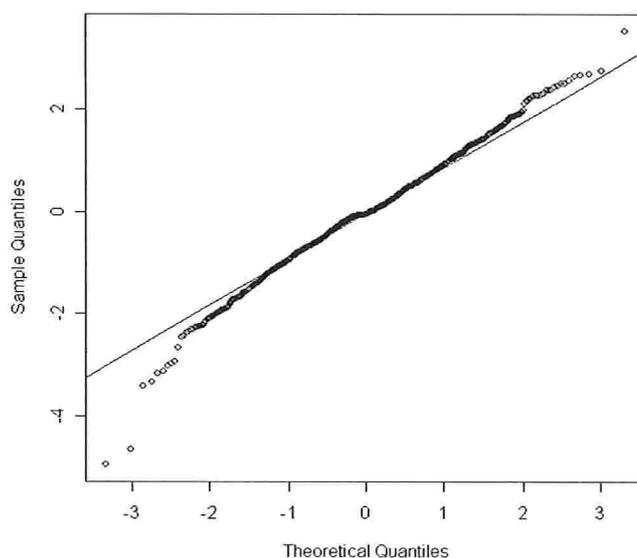


Figura 3.4: Gráfico QQ-plot dos resíduos padronizados – modelo GARCH(1,1) da série de retornos da Vale

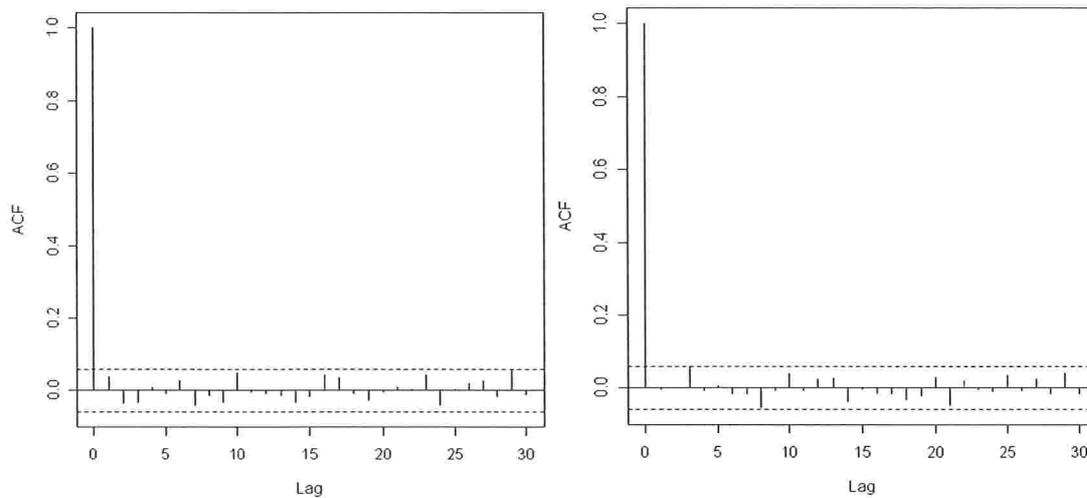


Figura 3.5: FÁC dos resíduos padronizados e do quadrado dos resíduos padronizados – modelo GARCH(1,1) da série de retornos da Vale

3.6 Extensões dos Modelos GARCH

Em muitos casos, o modelo GARCH básico (3.4) fornece um modelo razoavelmente bom para analisar séries temporais financeiras e para estimar a volatilidade condicional. No entanto, existem alguns aspectos do modelo que podem ser melhorados, para que assim, ocorra um melhor ajuste das características e da dinâmica de uma determinada série temporal. Esta seção introduz uma série de extensões do modelo GARCH básico que fazem da modelagem GARCH mais flexível.

3.6.1 Efeitos de Alavancagem Assimétrica e Informações Externas

No modelo GARCH básico (3.4), uma vez que apenas os resíduos ao quadrado ε_{t-1}^2 entram na equação, os sinais dos resíduos ou choques não têm efeitos na volatilidade condicional. Entretanto, um fato estilizado da volatilidade financeira é que más notícias (choques negativos) tendem a oferecer maior impacto na volatilidade do que boas notícias (choques positivos). Black (1976) atribui esse efeito ao fato de que más notícias tendem a abaixar os preços das ações, e portanto aumentam a alavancagem da ação causando maior volatilidade

nessa ação. Baseado nessa conjectura, o impacto assimétrico das notícias, das informações externas, é normalmente chamado de *efeito de alavancagem* ou *efeito alavanca* (*leverage effect*). Nesta subseção serão apresentados os modelos EGARCH, TGARCH e PGARCH que são capazes de incorporar este efeito.

Modelo EGARCH

Nelson (1991) propôs o seguinte modelo *GARCH exponencial* (EGARCH) para capturar os efeitos de alavancagem:

$$h_t = a_0 + \sum_{j=1}^q b_j \frac{|\varepsilon_{t-j}| + \gamma_j \varepsilon_{t-j}}{\sigma_{t-j}} + \sum_{i=1}^p a_i h_{t-i}, \quad (3.22)$$

em que $h_t = \log(\sigma_t^2)$ ou $\sigma_t^2 = \varepsilon^{h_t}$. Note que quando ε_{t-j} é positivo ou se houve "boas notícias", o efeito total de ε_{t-j} é $(1 + \gamma_j)|\varepsilon_{t-j}|$; por outro lado, quando ε_{t-j} é negativo ou se houve "más notícias", o efeito total de ε_{t-j} é $(1 - \gamma_j)|\varepsilon_{t-j}|$. Más notícias podem ter um impacto maior na volatilidade, e é esperado que o valor de γ_j seja negativo.

Outra vantagem do modelo EGARCH sobre o modelo GARCH básico é a garantia de que a variância condicional σ_t^2 é positiva, independentemente dos valores dos coeficientes de (3.22), porque é modelado o logaritmo de σ_t^2 em vez de se modelar σ_t^2 .

Modelo TGARCH

Outra variação do modelo GARCH que é capaz de modelar o efeito alavanca é o modelo TGARCH (*threshold GARCH*), que possui a seguinte forma:

$$\sigma_t^2 = a_0 + \sum_{i=1}^p a_i \varepsilon_{t-i}^2 + \sum_{i=1}^p \gamma_i S_{t-i} \varepsilon_{t-i}^2 + \sum_{j=1}^q b_j \sigma_{t-j}^2, \quad (3.23)$$

em que $S_{t-i} = 1$, se $\varepsilon_{t-i} < 0$ "más notícias" e $S_{t-i} = 0$, se $\varepsilon_{t-i} \geq 0$ "boas notícias", ou seja, ε_{t-i}^2 apresenta diferentes efeitos sobre a variância condicional σ_t^2 : quando ε_{t-i} é positivo, os efeitos totais são dados por $a_i \varepsilon_{t-i}^2$; quando ε_{t-i} é negativo, os efeitos totais são dados por $(a_i + \gamma_i) \varepsilon_{t-i}^2$. Dessa forma, espera-se um valor positivo para γ_i no caso de más notícias para que haja impactos maiores.

Este modelo também é conhecido como o modelo GJR, pois Glosten, Jagannathan e Runkle (1993) propuseram essencialmente o mesmo modelo.

Modelo PGARCH

O modelo GARCH básico também pode ser estendido para permitir a presença do efeito alavanca. Isto é possível se o modelo GARCH básico for tratado como um caso especial do modelo GARCH (PGARCH) proposto por Ding, Granger e Engle (1993):

$$\sigma_t^d = a_0 + \sum_{i=1}^p (a_i | \varepsilon_{t-i} | + \gamma_i \varepsilon_{t-i})^d + \sum_{j=1}^q b_j \sigma_{t-j}^d, \quad (3.24)$$

em que d é positivo e γ_i denota o coeficiente do efeito alavanca. Nota-se que para $d=2$, o modelo (3.24) se reduz ao modelo GARCH básico com efeito alavanca.

O expoente d do modelo PGARCH também pode ser fixado com algum valor diferente de 2. Por exemplo, uma escolha usual é fixar $d=1$ que faz com que o modelo GARCH seja robusto para outliers.

3.6.2 Variáveis Exógenas na Equação Geral da Média Condicional

Até o momento a equação da média condicional ficou restrita a uma constante nos modelos GARCH, porém no caso do modelo GARCH-M a volatilidade leva em conta a equação da média como uma variável explicativa.

A forma geral da média condicional é dada por

$$y_t = c + \sum_{i=1}^r \phi_i y_{t-i} + \sum_{j=1}^s \theta_j \varepsilon_{t-j} + \sum_{l=1}^L \beta'_l x_{t-l} + \varepsilon_t, \quad (3.25)$$

em que x_t é um vetor $k \times 1$ de variáveis exógenas, e β_l é o vetor $k \times 1$ de coeficientes.

3.6.3 Distribuição de Erros Não-Gaussianos

Até o momento, foi utilizada a suposição da distribuição normal para os erros. No entanto, é conhecido o fato de que séries temporais financeiras possuem caudas pesadas, portanto é de interesse usar distribuições que possuem caudas mais pesadas do que as da distribuição normal. Descreveremos aqui três possíveis distribuições de erros com caudas pesadas para o ajuste de modelos GARCH: a *distribuição t de Student*, a *distribuição exponencial dupla* e a *distribuição do erro generalizado*.

Distribuição t de Student

Se uma variável aleatória u_t tem distribuição t de Student com ν graus de liberdade e com parâmetro de escala s_t , a função densidade de probabilidade (fdp) de u_t é dada por:

$$f(u_t) = \frac{\Gamma[(\nu+1)/2]}{(\pi\nu)^{1/2} \Gamma(\nu/2)} \frac{s_t^{-1/2}}{[1 + u_t^2 / (s_t \nu)]^{(\nu+1)/2}},$$

em que $\Gamma(\cdot)$ é a função gama. A variância de u_t é dada por:

$$Var(u_t) = \frac{s_t v}{v-2}, \quad v > 2.$$

Se o termo de erro ε_t em um modelo GARCH segue uma distribuição t de Student com ν graus de liberdade e $Var_{t-1}(\varepsilon_t) = \sigma_t^2$, o parâmetro de escala s_t pode ser escrito como

$$s_t = \frac{\sigma_t^2 (\nu - 2)}{\nu}.$$

Portanto, a função de log verossimilhança de um modelo GARCH com distribuição t de Student para os erros pode ser facilmente construída baseada na fdp acima.

Distribuição do Erro Generalizado e Exponencial Dupla

Nelson (1991) propôs o uso da distribuição do erro generalizado (GED) para capturar as caudas pesadas geralmente observadas na distribuição de séries temporais financeiras. Se a variável aleatória u_t é uma GED com média zero e variância unitária, a fdp de u_t é dada por:

$$f(u_t) = \frac{\nu \exp[-(1/2) |u_t / \lambda|^\nu]}{\lambda 2^{(v+1)/v} \Gamma(1/\nu)},$$

em que

$$\lambda = \left[\frac{2^{-2/\nu} \Gamma(1/\nu)}{\Gamma(3/\nu)} \right]^{1/2}$$

e ν é um parâmetro positivo que determina o comportamento das caudas dessa distribuição. Quando $\nu = 2$, a fdp acima se reduz a fdp normal padrão; quando $\nu <$

2, a densidade possui caudas mais densas do que da distribuição normal; quando $\nu > 2$, a densidade possui caudas mais leves do que da densidade normal.

Quando $\nu = 1$, a fdp de GED se reduz à fdp da *distribuição exponencial dupla*:

$$f(u_t) = \frac{1}{\sqrt{2}} e^{-\sqrt{2}|u_t|}.$$

Baseado na fdp acima, a função de log verossimilhança do modelo GARCH com GED ou de erros com distribuição exponencial dupla pode ser facilmente construída.

3.7 Seleção e Comparação de Modelos GARCH

As seções anteriores ilustraram extensões dos modelos GARCH. Selecionar o melhor modelo para determinado conjunto de dados pode ser uma tarefa muito complicada. O diagnóstico de modelos baseado nos resíduos padronizados pode ser usado para comparar a eficácia de diferentes aspectos dos modelos GARCH. Além disso, os critérios tradicionais de seleção de modelos como o critério da informação de Akaike (AIC) e o critério da informação Bayesiana (BIC) também podem ser usados para selecionar os melhores modelos.

Considerando o ajuste de um modelo GARCH(1,1) simples com distribuição normal e um modelo com distribuição t de Student para a série de retornos dos preços das ações da Vale, temos que o BIC do modelo com distribuição normal é levemente menor do que o do modelo com distribuição t de Student, o que sugere que a distribuição normal seja melhor que a distribuição t de Student. Os valores de BIC, AIC e log da verossimilhança dos modelos ajustados são apresentados na Tabela 3.3.

Tabela 3.3: AIC, BIC e log da verossimilhança dos modelos GARCH(1,1) ajustados para a série de retornos da Vale

	Dist. Normal	Dist. t de Student
AIC	4,6685	4,6805
BIC	4,6861	4,6981
log verossimilhança	-2671	-2677

Podemos comparar graficamente os ajustes dos modelos. Na Figura 3.6 são apresentadas as FAC do quadrado dos resíduos padronizados dos modelos ajustados. Este gráfico sugere que os dois modelos são apropriados para modelar a volatilidade condicional. Podemos comparar também os gráficos QQ-plots dos resíduos padronizados, que estão apresentados na Figura 3.7. A análise gráfica mostra que os dados destoam da distribuição normal e da distribuição t de Student.

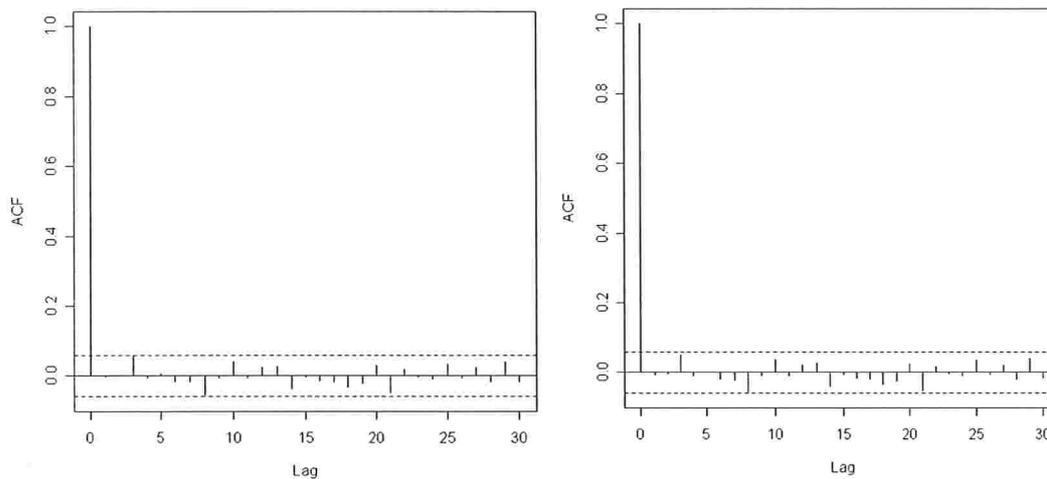


Figura 3.6: Comparação das FACs do Quadrado dos Resíduos Padronizados - (a) Distribuição Normal e (b) Distribuição t de Student

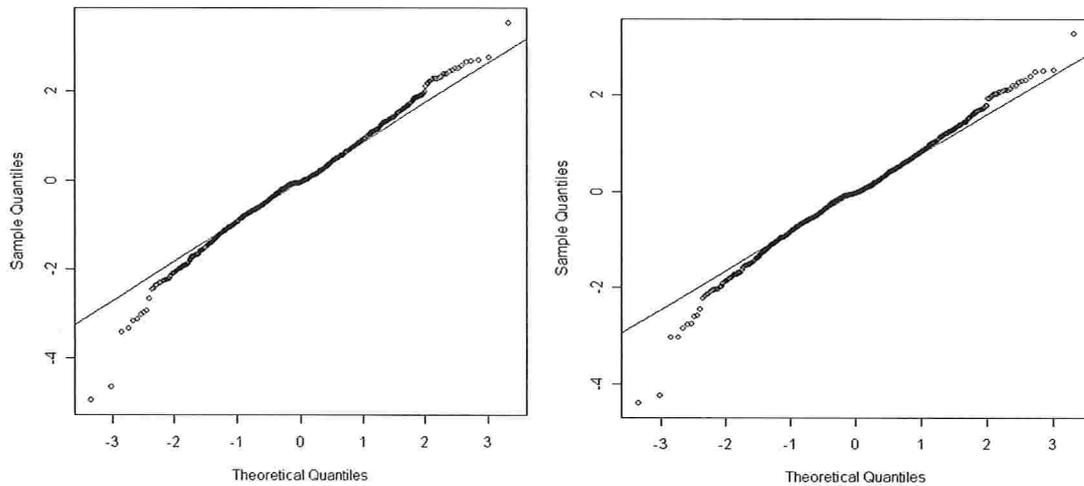


Figura 3.7: Comparação dos QQ-plots dos Resíduos Padronizados - (a) Modelo GARCH(1,1) com Distribuição Normal e (b) Modelo GARCH(1,1) com Distribuição t de Student (foi utilizada a opção de g.l default do R-Plus).

3.8 Predição de Modelos GARCH

Uma tarefa importante na modelagem da volatilidade condicional é a geração de previsões tanto para valores futuros de uma série temporal quanto para sua volatilidade condicional. Uma vez que a média condicional do modelo geral GARCH dado em (3.25) assume a forma de um ARMA tradicional, a previsão de valores futuros de uma série temporal pode ser obtida seguindo-se a abordagem tradicional para predição de modelos ARMA. Portanto, levando-se em conta também a variância condicional, modelos GARCH podem gerar previsões com menor erro para os valores futuros, especialmente sobre horizontes pequenos. Esta seção ilustra como prever a volatilidade usando modelos GARCH.

Para simplificar, consideremos o modelo básico GARCH(1,1):

$$\sigma_t^2 = a_0 + a_1 \varepsilon_{t-1}^2 + b_1 \sigma_{t-1}^2,$$

com $t=1, 2, \dots, T$. Para obter $E_T[\sigma_{T+k}^2]$, que é a previsão da volatilidade futura σ_{T+k}^2 , para $k>0$, dada a informação até o tempo T, a partir da equação anterior temos:

$$E_T[\sigma_{T+1}^2] = a_0 + a_1 E_T[\varepsilon_T^2] + b_1 E_T[\sigma_T^2] = a_0 + a_1 \varepsilon_T^2 + b_1 \sigma_T^2,$$

em que ε_T^2 e σ_T^2 são os valores obtidos depois da estimação⁴. Para $T+2$, temos:

$$E_T[\sigma_{T+2}^2] = a_0 + a_1 E_T[\varepsilon_{T+1}^2] + b_1 E_T[\sigma_{T+1}^2] = a_0 + (a_1 + b_1) E_T[\sigma_{T+1}^2],$$

em que $E_T[\varepsilon_{T+1}^2] = E_T[\sigma_{T+1}^2]$. Seguindo o mesmo raciocínio, pode-se obter a equação de previsão da volatilidade condicional:

$$E_T[\sigma_{T+k}^2] = a_0 \sum_{i=1}^{k-2} (a_1 + b_1)^i + (a_1 + b_1)^{k-1} E_T[\sigma_{T+1}^2], \quad (3.26)$$

para $k \geq 2$. Para $k \rightarrow \infty$, a previsão da volatilidade em (3.26) se aproxima da variância incondicional $a_0 / (1 - a_1 - b_1)$ se o processo GARCH for estacionário (ou seja, se $a_1 + b_1 < 1$).

O algoritmo de predição (3.26) produz uma previsão para a variância condicional σ_{T+k}^2 . A previsão para a volatilidade condicional, σ_{T+k} , é definida pela raiz quadrada da previsão de σ_{T+k}^2 .

⁴ A notação utilizada se refere aos valores ajustados em vez de representar os valores "reais" não observados

Capítulo 4

Rolling Analysis de Séries Temporais

4.1 Introdução

Rolling Analysis está sendo discutida pois apesar de não ser utilizada diretamente no presente trabalho, alguns de seus princípios serviram de base na construção do método proposto, conforme será visto nos capítulos posteriores.

Rolling Analysis em séries temporais é muito utilizada para avaliar a estabilidade do modelo ao longo do tempo. Quando analisamos séries temporais financeiras usando modelos estatísticos, uma das premissas assumidas é que os parâmetros são constantes ao longo do tempo. De qualquer forma, o ambiente econômico é muito dinâmico e esta suposição pode ser falsa. Uma técnica comum para avaliar a estabilidade dos parâmetros do modelo é estimá-los através de sub-amostras (*rolling window*) de tamanho fixo à partir do conjunto de dados completo, em que estas sub-amostras refere-se aos dados de intervalos de tempos distintos. Se os parâmetros são realmente constantes na amostra toda, então as estimativas baseadas nas sub-amostras também devem apresentar o mesmo comportamento. Se os parâmetros mudam em algum instante de tempo, então as estimativas devem capturar esta mudança.

Rolling Analysis também é muito utilizada em *backtesting* com base nos dados históricos para avaliar estabilidade e acurácia de previsão do modelo. *Backtesting* geralmente é usado da seguinte forma: os dados históricos são separados em amostra de estimação e amostra de previsão, o modelo é então ajustado na amostra de estimação e utilizado para realizar previsão *h-passos* à frente. Como os dados para os quais são feitas as previsões já foram observados (*amostra de previsão*) então os erros podem ser calculados. A amostra de estimação então é deslocada para frente em um certo incremento e o processo de estimação e previsão é repetido até que não seja mais possível realizar previsões.

Apesar de *Backtesting* ser considerado em *rolling analysis*, ele não foi utilizado no presente trabalho.

Métodos de médias móveis também são muito utilizados em *rolling analysis* e desempenham um papel importante na análise técnica de séries financeiras. Uma forma de pensar sobre estes métodos é que eles são modelos mais simples com parâmetros variando no tempo. Porém, algumas vezes modelos de médias móveis não são adequados, sendo necessário um modelo mais geral com parâmetros variando no tempo.

4.2 Estatísticas Descritivas

A seguir são apresentadas algumas estatísticas descritivas univariadas e bivariadas muito comuns dentro do contexto de *rolling analysis*.

4.2.1 Estatísticas univariadas

Considere a análise de uma série temporal $y_t, t=1, \dots, T$. Uma das questões de interesse é saber se a média e a variância (ou desvio padrão) referente à distribuição de y_t são constantes sobre a amostra inteira. Para avaliar a estabilidade dos parâmetros, seja n o tamanho de algumas sub-amostras (*window*) e defina a média amostral (*rolling sample means*), variância e desvio padrão:

$$\hat{\mu}_t(n) = \frac{1}{n} \sum_{i=0}^{n-1} y_{t-i}, \quad (4.1)$$

$$\hat{\sigma}_t^2(n) = \frac{1}{n-1} \sum_{i=0}^{n-1} (y_{t-i} - \hat{\mu}_t(n))^2, \quad (4.2)$$

$$\hat{\sigma}_t(n) = \sqrt{\hat{\sigma}_t^2(n)}. \quad (4.3)$$

Para as sub-amostras, $t=n, \dots, T$, as estimativas amostrais da média e da variância no tempo t assumindo tamanho amostral n , são calculadas tomando as n -observações mais recentes. Calculada as estimativas, os dados são deslocados

em uma posição de forma que no final do processo haja $T-n+1$ estimativas (*rolling estimates*) dos parâmetros.

Considere os dados mensais referente aos retornos das ações da Vale no período de 19 de Agosto de 2004 até 23 de Janeiro de 2009. Foram utilizados os preços de fechamento. Na figura 4.1 pode ser observado o gráfico contendo os retornos e as estimativas amostrais (*rolling window*) da média e do desvio padrão. Foi utilizado $n=24$ (*tamanho amostral*).

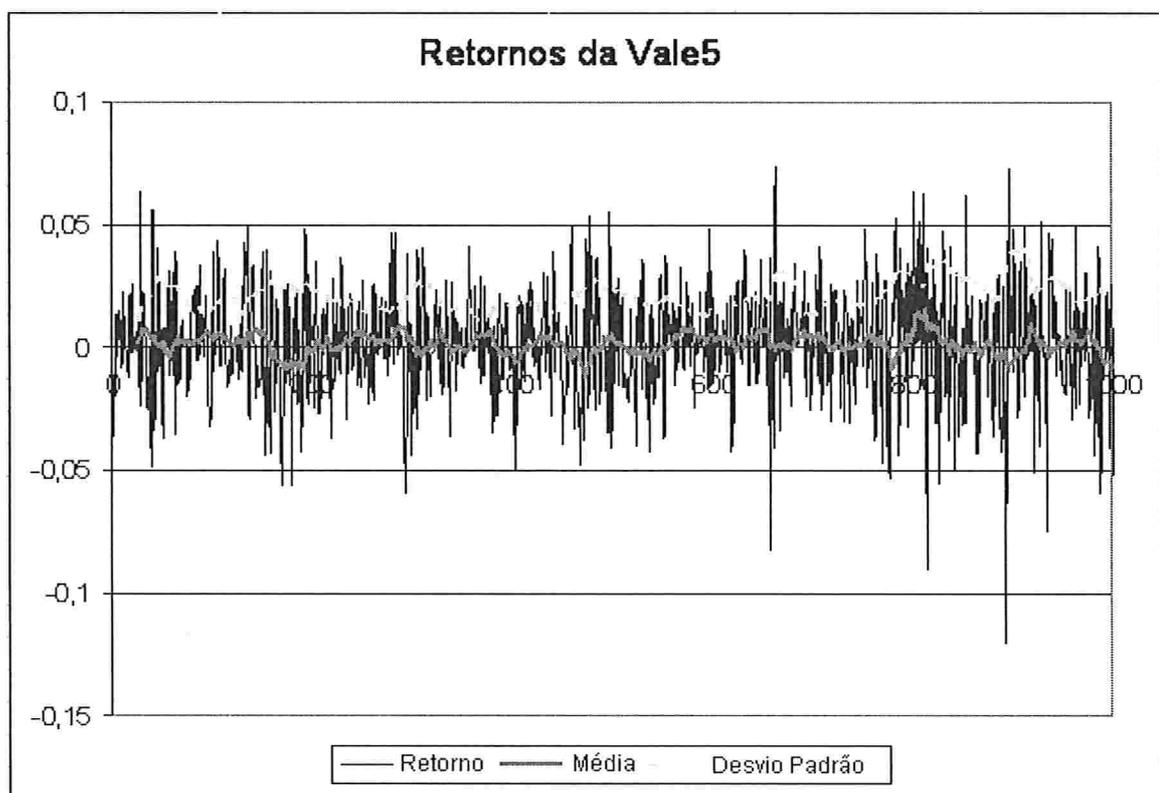


Figura 4.1 Retornos mensais, média e desvio padrão amostral referentes à ação da Vale5.

As estimativas amostrais de $\hat{\mu}_t(24)$ e principalmente $\hat{\sigma}_t(24)$ (*rolling estimates*) variam ao longo do tempo. A média começa positiva e ao longo do tempo ela alterna valores negativos e positivos. Os valores do desvio padrão $\hat{\sigma}_t(24)$, começam por volta de 2%, e próximo do fim do período eles chegam a atingir o dobro do valor inicial.

Os erros padrão das estimativas $\hat{\mu}_t(24)$ e $\hat{\sigma}_t(24)$ podem ser computados utilizando as seguintes fórmulas assintóticas:

$$\overline{SE}(\hat{\mu}_t(n)) = \frac{\hat{\sigma}_t(n)}{\sqrt{n}}, \quad \overline{SE}(\hat{\sigma}_t(n)) = \frac{\hat{\sigma}_t(n)}{\sqrt{2n}}.$$

Nas figura 4.2 e 4.3 seguem os gráficos das estimativas da média e do desvio padrão com suas respectivas bandas de confiança. As bandas foram construídas com base nas fórmulas apresentadas acima e assumindo 95% de confiança. Em geral, $\hat{\sigma}_t(24)$ é estimado mais precisamente que $\hat{\mu}_t(24)$.

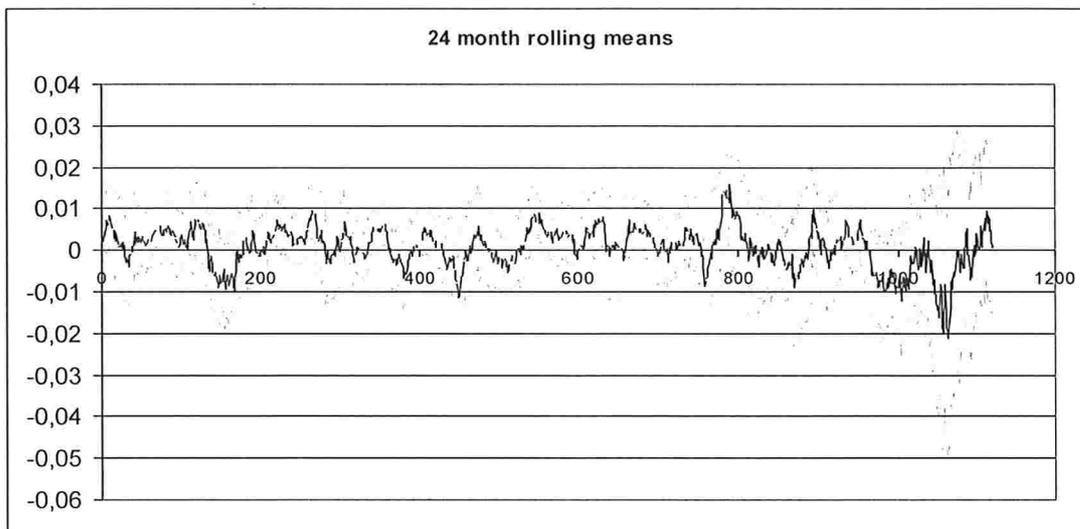


Figura 4.2 Banda de confiança para os valores da média amostral $\hat{\mu}_t(24)$.

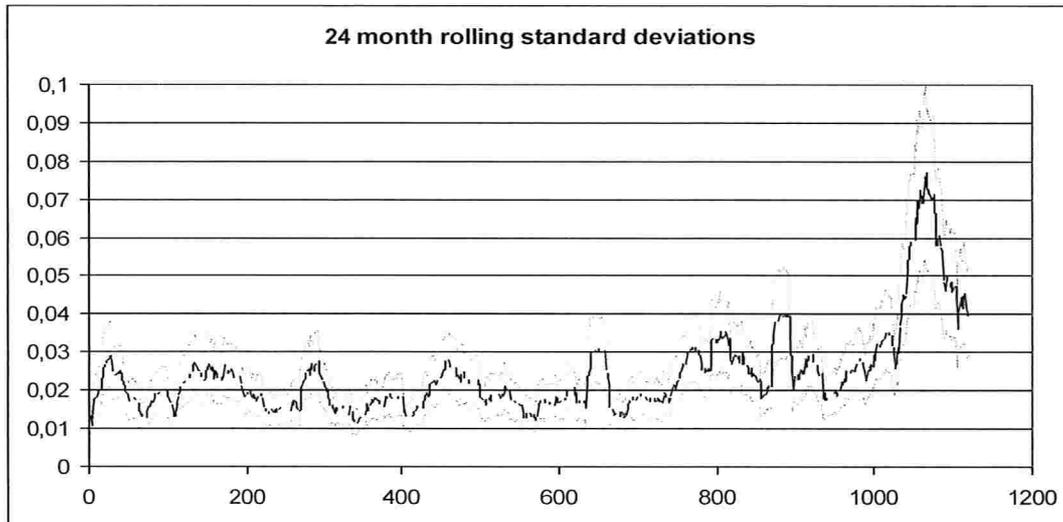


Figura 4.3 Banda de confiança para os valores do desvio padrão $\hat{\sigma}_t(24)$ amostral.

Conforme observado nas Figuras 4.2 e 4.3, não há evidências de que a média se altere ao longo do tempo, porém o desvio padrão apresenta um significativo aumento no final do período considerado.

4.2.2 Estatísticas bivariadas

Considere agora a análise de duas séries temporais y_{1t} e y_{2t} com amostra $t=1, \dots, T$. Para avaliar se a covariância e correlação entre y_{1t} e y_{2t} são constantes ao longo do tempo, sub-amostras de tamanho n podem ser retiradas e a covariância e correlação mensuradas ao longo do tempo:

$$\hat{\sigma}_{12,t}(n) = \frac{1}{n-1} \sum_{i=0}^{n-1} (y_{1t-i} - \hat{\mu}_{1t}(n))(y_{2t-i} - \hat{\mu}_{2t}(n)),$$

$$\hat{\rho}_{12,t}(n) = \frac{\hat{\sigma}_{12,t}(n)}{\hat{\sigma}_{1t}(n)\hat{\sigma}_{2t}(n)}.$$

Analisando as séries mensais de retornos da Vale e do índice Bovespa no período de 03 de Janeiro de 2005 até 15 de Maio de 2008, considerando $n = 24$,

pode-se perceber na Figura 4.4, que no início do período, a correlação entre as duas séries está em torno de 0.4, caindo logo em seguida para 0.2. Depois da queda a correlação subiu consideravelmente oscilando em torno de 0.8, indicando que a correlação parece variar com o tempo.

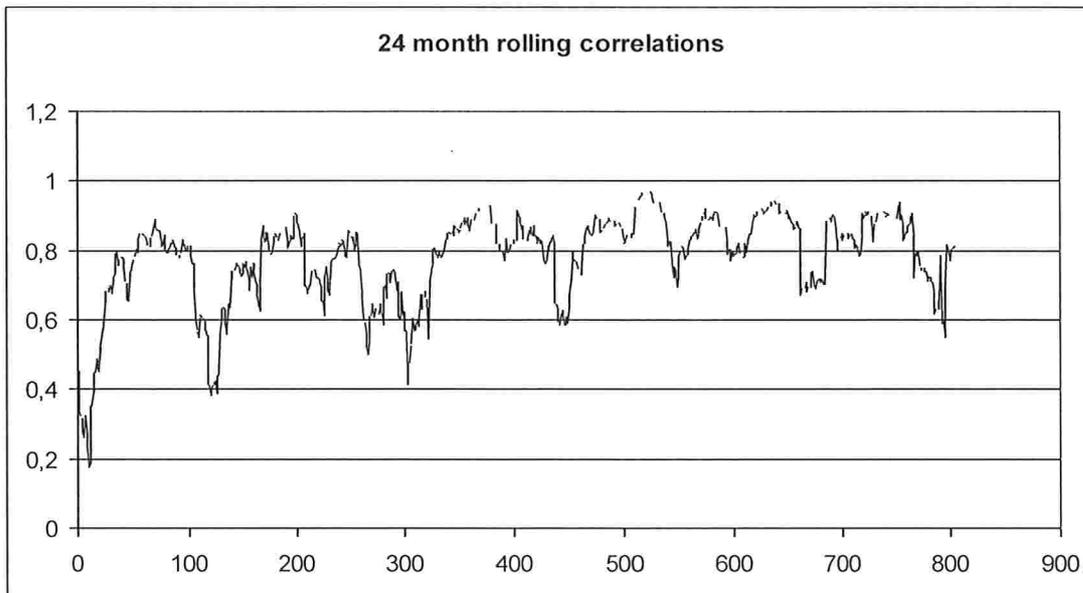


Figura 4.4 Correlação amostral entre os retornos da Vale e do índice do Bovespa (n=24).

4.2.3 Médias móveis ponderadas exponencialmente

As estatísticas descritivas descritas nas seções anteriores são baseadas em médias móveis ponderadas igualmente da série y_t observada. Médias móveis igualmente ponderadas são úteis em algumas situações, mas podem produzir resultados viesados para previsões de curto prazo. Isto é causado pois previsões utilizando a média igualmente ponderada são mais sensíveis à valores extremos. Para ilustrar este fato, considere $T=100$, observações de uma série temporal simulada $y_t \sim GWN(0,1)$, com um outlier inserido em $t=60$: ou seja, $y_{60} = 10$. Os dados e os valores estimados de $\hat{\mu}_t(50)$ e $\hat{\sigma}_t(50)$ são ilustrados na Figura 4.5. Note que o outlier no ponto $t=60$ inflaciona as estimativas $\hat{\mu}_t(50)$ e $\hat{\sigma}_t(50)$ por 9 períodos.

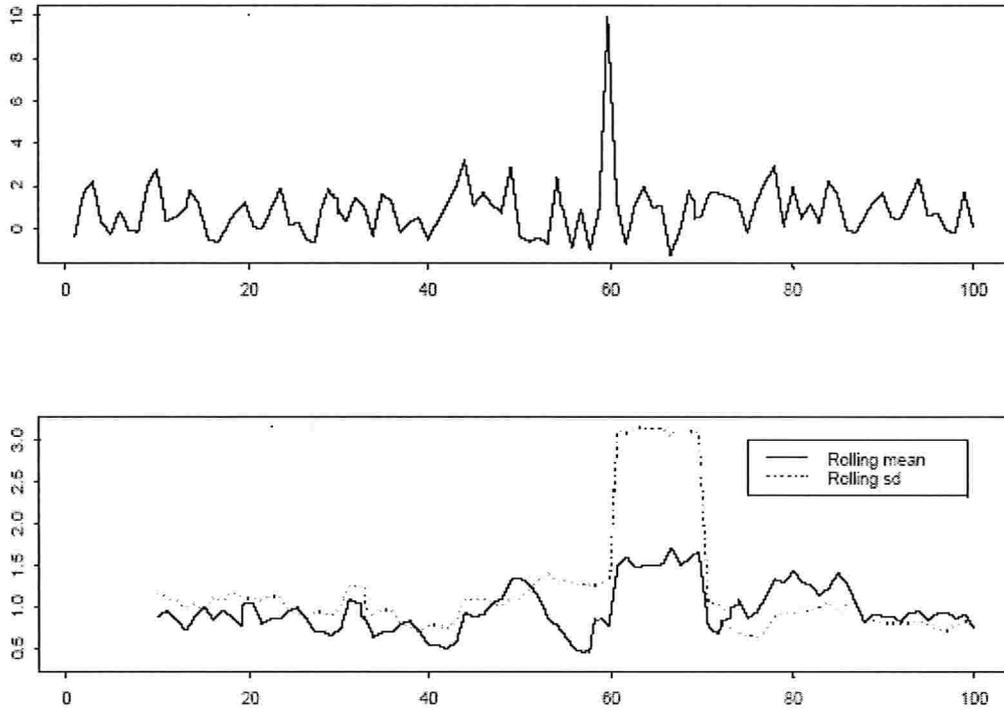


Figura 4.5 Dados originais, média e desvio padrão amostral da série simulada assumindo $n=24$

Para reduzir os efeitos de valores extremos nas estimativas, as observações utilizadas no cálculo podem ter pesos diferentes. Um esquema muito utilizado de ponderação que coloca maior peso nas observações mais recentes é baseado em pesos que decaem exponencialmente (*exponentially weighted moving average-EWMA*). Um *EWMA* em uma série temporal y_t é definido como:

$$\tilde{\mu}_t(n) = \sum_{i=0}^{n-1} w_i \cdot y_{t-i}, \quad w_i = \frac{\lambda^{i-1}}{\sum_{i=0}^{n-1} \lambda^{i-1}},$$

em que $0 < \lambda < 1$ é o parâmetro de decaimento. Quando $n \rightarrow \infty, \lambda^n \rightarrow 0, w^n \rightarrow 0$, *EWMA* converge para:

$$\tilde{\mu}_t(\lambda) = (1 - \lambda) \sum_{i=0}^{\infty} \lambda^i \cdot y_{t-i}, \quad (4.4)$$

Dessa forma, *EWMA* pode ser definido independentemente do tamanho de n . *EWMA* em (4.4) pode ser eficientemente computado usando o processo recursivo:

$$\tilde{\mu}_t(\lambda) = (1 - \lambda)y_t + \lambda\tilde{\mu}_{t-1}(\lambda). \quad (4.5)$$

De (4.5), está claro que quanto mais perto de 1 λ estiver, maior é o peso dado às estimativas de períodos anteriores em relação ao período atual o que pode ser observado na Figura 4.6.

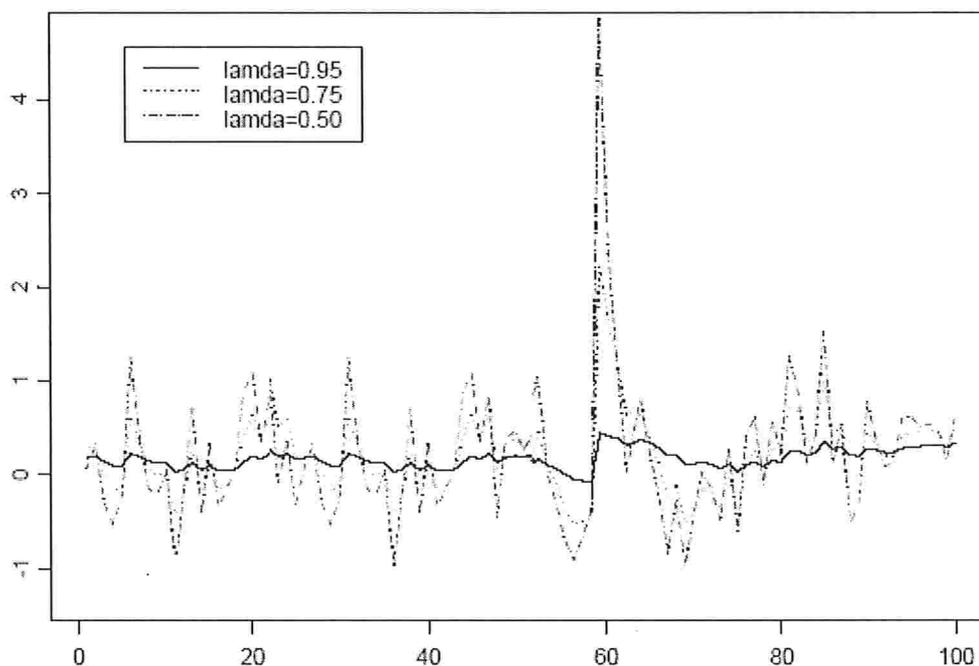


Figura 4.6 Estimativas de μ usando EWMA.

4.3 Rolling Regression

Para o modelo de regressão linear, *rolling analysis* pode ser utilizada para avaliar a estabilidade dos parâmetros do modelo e fornecer um simples modelo com parâmetros variando no tempo. Para uma sub-amostra (*window*) de tamanho $n < T$, o modelo de regressão com parâmetros variando no tempo (*rolling linear regression model*) pode ser expresso como:

$$y_t(n) = X_t(n)\beta_t(n) + \varepsilon_t(n), \quad (4.6)$$

em que $y_t(n)$ é o vetor ($n \times 1$) da variável de interesse, $X_t(n)$ é uma matriz ($n \times k$) de variáveis exploratórias, $\beta_t(n)$ é um vetor ($k \times 1$) dos parâmetros da regressão e $\varepsilon_t(n)$ é um vetor ($n \times 1$) com os termos de erro. As n observações em $y_t(n)$ e $X_t(n)$ são os n valores mais recentes referentes aos tempos $t-n+1$ até t , em que $n > k$. As estimativas de mínimo quadrado são:

$$\begin{aligned} \hat{\beta}_t(n) &= [X_t(n)' X_t(n)]^{-1} X_t(n)' y_t(n), \\ \hat{\sigma}_t^2(n) &= \frac{1}{n-k} \hat{\varepsilon}_t(n)' \hat{\varepsilon}_t(n) \\ &= \frac{1}{n-k} [y_t(n) - X_t(n)\hat{\beta}_t(n)]' [y_t(n) - X_t(n)\hat{\beta}_t(n)], \\ \overline{\text{var}}(\hat{\beta}_t(n)) &= \hat{\sigma}_t^2(n) \cdot [X_t(n)' X_t(n)]^{-1}. \end{aligned}$$

Considere a estimação do seguinte modelo:

$$V_t = \alpha + \beta * I_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim RB(0, \sigma^2) \quad (4.7)$$

em que V_t denota o retorno diário da Vale e I_{t-1} é o retorno diário do Ibovespa no dia anterior. O coeficiente β mede o quanto as variações nos retornos do Ibovespa impactam nos retornos da Vale no dia seguinte e α indica o retorno médio da Vale dado que o Ibovespa foi estável no dia anterior, ou seja, sem nenhuma variação, positiva ou negativa. *Rolling Regression* pode ser utilizada para avaliar a estabilidade do modelo apresentado acima ao longo do tempo.

Ajustando um modelo para a série completa, utilizando o período de 04 de Janeiro de 2005 até 15 de Maio de 2008, foi obtido um valor de 0.4336 para β , indicando relação entre as séries. A estimativa de α foi significativamente diferente

de 0, indicando que mesmo que não haja variação no Ibovespa no dia anterior pode ser que a ação da Vale apresente variações.

Considerando agora *Rolling Regression* com $n=24$, foram obtidas 803 estimativas para cada parâmetro. Na Figuras 4.7, são apresentados os gráficos das estimativas dos parâmetros. Observando o gráfico de $\hat{\alpha}$, pode-se ver que ele oscila bastante, porém está sempre em torno de zero, indicando que variações nulas no Ibovespa no dia anterior podem implicar em pouca variação na ação da Vale, destoando dos resultados observados para o ajuste global (modelo 4.7). Os valores de $\hat{\beta}$, também sobem e descem bastante dificultando encontrar um padrão de comportamento.

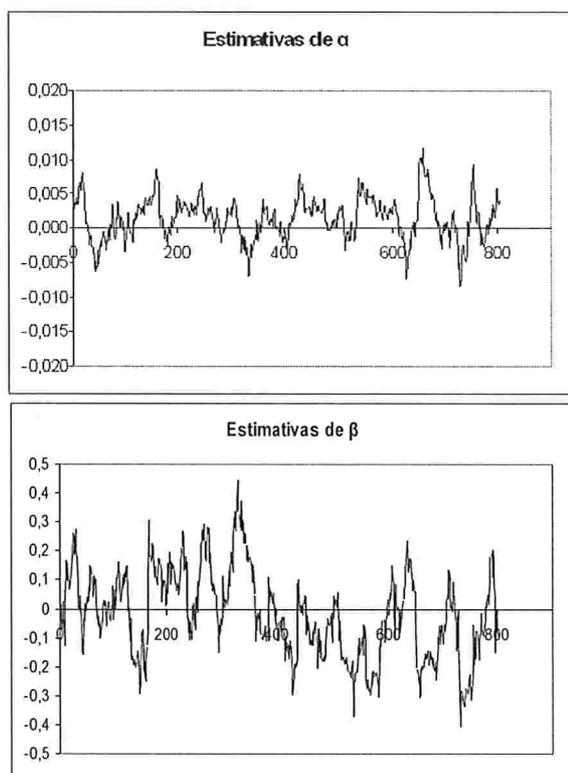


Figura 4.7 Estimativas dos parâmetros α e β .

4.4 Rolling Predictions e Backtesting

Rolling Regression pode ser utilizada para avaliar o desempenho preditivo de modelos com base em dados históricos utilizando a técnica de *backtesting*. Para ilustrá-la, considere o modelo de regressão (4.6). O desempenho preditivo de (4.6) na amostra de validação é baseada nas previsões (*rolling predictions*) e erros de previsão:

$$\hat{y}_{t+h|t} = x'_{t+h} \hat{\beta}_t(n), \quad (4.8)$$

$$\hat{\varepsilon}_{t+h|t} = y_{t+h} - \hat{y}_{t+h|t} = y_{t+h} - x'_{t+h} \hat{\beta}_t(n). \quad (4.9)$$

As previsões de $\hat{\beta}_t(n)$ utilizam os dados até o tempo t para sua estimação, e as previsões são realizadas com base nas observações nos tempos $t+h$, para $h > 0$. As previsões (*rolling predictions*) são atualizadas já que $\hat{\beta}_t(n)$ se altera quando t aumenta. Quando $h=1$ há $T-n$ previsões de 1 passo, quando $h=2$ há $T-n-1$ previsões de 2 passos e assim por diante.

As previsões (4.8) podem ser avaliadas examinando algumas propriedades dos erros de previsão (*rolling forecast errors*) (4.9). Normalmente são utilizadas as seguintes estatísticas:

$$ME = \frac{1}{T-n-h+1} \sum_{t=n}^{T-h} \hat{\varepsilon}_{t+h|t}, \quad (4.10)$$

$$MSE(h) = \frac{1}{T-n-h+1} \sum_{t=n}^{T-h} \hat{\varepsilon}_{t+h|t}^2,$$

$$RMSE(h) = \sqrt{MSE(h)},$$

$$MAE(h) = \frac{1}{T-n-h+1} \sum_{t=n}^{T-h} |\hat{\varepsilon}_{t+h|t}|,$$

$$MAPE(h) = \frac{1}{T-n-h+1} \sum_{t=n}^{T-h} \left| \frac{\hat{\varepsilon}_{t+h|t}}{y_{t+h}} \right|.$$

A primeira medida avalia o viés de previsão e as outras medidas avaliam o viés e a precisão.

4.4.1 Comparação de modelos

Backtesting é frequentemente utilizada para comparar a acurácia de 2 ou mais modelos. Normalmente, as estatísticas utilizadas para avaliar a previsão (4.10), são calculadas para cada modelo, e o modelo que apresentar o menor conjunto de medidas é considerado o melhor. Recentemente, Diebold and Mariano (1995), propuseram um simples algoritmo utilizando os erros de previsão de forma a avaliar estatisticamente se um modelo é melhor do que os outros. Seja $\hat{\varepsilon}_{t+h|t}^1$ e $\hat{\varepsilon}_{t+h|t}^2$, os erros de previsão de 2 modelos e N denota o número de previsões. A acurácia de cada de previsão é medida pela função de perda:

$$L(\hat{\varepsilon}_{t+h|t}^i), i=1,2,\dots$$

Duas funções de perda muito populares são: *square error loss* $L(\hat{\varepsilon}_{t+h|t}^i) = (\hat{\varepsilon}_{t+h|t}^i)^2$ e *absolute error loss* $L(\hat{\varepsilon}_{t+h|t}^i) = |\hat{\varepsilon}_{t+h|t}^i|$. Para determinar se um modelo é melhor do que outro, Diebold and Mariano(1995) sugeriram calcular a função:

$$d_t = L(\hat{\varepsilon}_{t+h|t}^1) - L(\hat{\varepsilon}_{t+h|t}^2),$$

para testar a hipótese nula de igual acurácia:

$$H_0 : E[d_t] = 0.$$

A estatística de teste de Diebold and Mariano (1995) é simplesmente a razão

$$DM = \frac{\bar{d}}{\overline{\text{var}}(\bar{d})^{1/2}} \quad (4.11)$$

em que

$$\bar{d} = \frac{1}{N} \sum_{i=1}^N d_i$$

é a média da função d_i , e $\overline{\text{var}}(\bar{d})$ é uma estimativa consistente da variância assintótica de \bar{d} . Diebold and Mariano (1995) sugeriram calcular $\overline{\text{var}}(\bar{d})$ usando o estimador não paramétrico de Newey-West com uma função de peso retangular e um parâmetro de truncamento igual ao tamanho da previsão, h , menos 1. Diebold and Mariano (1995) mostraram que sob a hipótese nula a estatística DM é assintoticamente $N(0,1)$.

Capítulo 5

Metodologia Proposta

5.1 Introdução

Um dos grandes sonhos da humanidade sempre foi conseguir antecipar o acontecimento de fatos, ou seja, tentar mesmo que de forma empírica prever o futuro. Com o desenvolvimento dos modelos ARMA este sonho aproximou-se um pouco mais da realidade, porém percebeu-se durante o caminho que muita coisa poderia ser melhorada e que existiam muitos aspectos não lineares que estão além da capacidade dos modelos ARMA. Por exemplo, vários fenômenos não lineares como não normalidade, ciclos assimétricos, bi-modalidade, relação não linear entre as variáveis do modelo, variação de desempenho na predição sobre o espaço de estados, não reversibilidade e sensibilidade para as condições iniciais tem sido observadas em muitas séries temporais, por exemplo: *Tong* (1990, 1995) e *Tjøstheim* (1994). Além do domínio linear, há uma infinidade de formas não lineares que podem ser exploradas. O desenvolvimento inicial de séries temporais não lineares focou em várias formas não lineares paramétricas. Exemplos de sucesso incluem: o modelo ARCH para a estrutura de flutuação em séries temporais financeiras (*Engel* 1982 e *Bollerslev* 1986), dentre outros, como por exemplo: os modelos *exponenciais autoregressivos* (EXPAR) de Haggan e Ozaki (1981), o modelo *threshold autoregressivo* (TAR) de Tong (1990) e o modelo *funcional autoregressivo* (FAR) de Chen e Tsay (1993). Todos esses modelos são casos particulares de uma classe mais geral de modelos não lineares: "*functional coefficient regression model*".

Uma forma geral de representação dos modelos "*functional coefficient regression model*" é: seja $\{Y_t, X_t, U_t\}_{t=-\infty}^{\infty}$ um processo estacionário conjunto com $X_t = (X_{t1}, \dots, X_{td})$ tomando valores em \mathbb{R}^d e U_t em \mathbb{R} . Seja $E(Y_t) < \infty$. A função de regressão multivariada é definida como:

$$f(x, u) = E(Y_t | X_t = x, U_t = u)$$

No contexto de séries temporais, ambos X_t e U_t consistem de alguns lags de Y_t . O modelo "*functional coefficient regression model*" requer que a função de regressão tenha a seguinte forma:

$$f(x, u) = \sum_{j=1}^d a_j(u) x_j,$$

em que os $a_j(\cdot)$ são funções mensuráveis de \mathfrak{R} em \mathfrak{R} e $x = (x_1, \dots, x_d)^T$.

Os modelos não lineares "*functional coefficient regression model*", desde sua concepção até hoje vem ganhando muita atenção por serem modelos dinâmicos e dessa forma acredita-se que eles possam representar de forma mais fiel a realidade. Porém eles são modelos mais complexos e a própria estimação das funções $a_j(\cdot)$ se torna difícil. Alguns métodos de estimação, como por exemplo: métodos de alisamento baseado em polinomial *splines* (Cai e Yao) ou técnicas de regressão linear local (Shen e Huang) são utilizadas mas em geral não são simples.

O objetivo deste capítulo é apresentar um método alternativo mais simples para identificar o padrão de comportamento dos parâmetros $a_j(\cdot)$ ao longo do tempo do que os citados anteriormente. Além disso, na formulação sugerida neste trabalho X_t refere-se á alguns lags de Y_t^2 e de alguns lags dos choques aleatórios ao quadrado ε_t^2 , ou seja, será apresentada uma formulação de modelos GARCH com coeficientes variando no tempo.

Os modelos GARCH por si só já conseguem capturar alguns aspectos não lineares além dos modelos ARMA, porém em algumas situações como na modelagem de dados do mercado de ações que é muito dinâmico, a suposição de variação dos coeficientes ao longo do tempo pode ser razoável e pode consequentemente possibilitar a obtenção de um modelo com melhor ajuste.

5.2 Metodologia Utilizada

A metodologia apresentada a seguir é o coração deste trabalho, ela é baseada em outras duas técnicas muito utilizadas em séries temporais: “*Rolling Analysis*” e modelagem ARMA. Até onde foi pesquisado, a forma apresentada aqui parece ser inovadora, por isso muitos dos resultados são empíricos e seria necessário um aprofundamento teórico dos mesmos, porém isto foge do escopo deste trabalho. Conforme já discutido anteriormente, um modelo muito utilizado no mercado financeiro é o GARCH, mais especificamente o GARCH(1,1). Este modelo possibilita que muitos aspectos observados no mercado de ações sejam captados de maneira satisfatória, porém como o mercado financeiro é muito dinâmico, talvez uma adaptação no modelo GARCH, assumindo que os parâmetros ao invés de fixos fossem funções do tempo pudesse resultar em um modelo mais eficiente.

Consideremos o modelo GARCH(1,1) tradicional, para facilitar o entendimento da metodologia:

$$\sigma_t^2 = a_0 + a_1 \varepsilon_{t-1}^2 + b_1 \sigma_{t-1}^2.$$

Nesta formulação os parâmetros a_0 , a_1 e b_1 são fixos, porém suponha que os parâmetros a_0 , a_1 e b_1 sejam funções do tempo, neste caso a formulação do modelo seria:

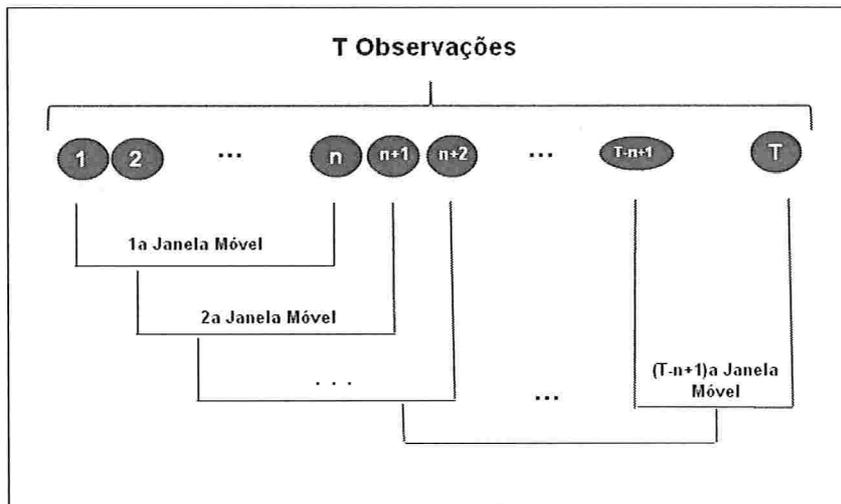
$$\sigma_t^2 = a_0(t) + a_1(t) \varepsilon_{t-1}^2 + b_1(t) \sigma_{t-1}^2.$$

Como $a_0(t)$, $a_1(t)$ e $b_1(t)$ não são fixos, é necessário caracterizarmos qual é o comportamento deles ao longo do tempo. Isto foi feito da seguinte forma:

1º Passo: Para cada uma das 3 séries analisadas, foi definida uma janela móvel (*rolling window*) de tamanho $n < T$, em que T é o tamanho da série. Dado a janela móvel de tamanho n , foi possível obter $T-n+1$ sub-amostras e em cada uma

delas (i-ésima sub-amostra) foi estimado um modelo GARCH(1,1): $\sigma_i^2 = a_{0i} + a_{1i}\varepsilon_{i-1}^2 + b_{1i}\sigma_{i-1}^2$. Os parâmetros: a_{0i} , a_{1i} e b_{1i} estimados em cada sub-amostra são fixos. Dessa forma, serão estimados $T-n+1$ valores para os parâmetros a_{0i} , a_{1i} e b_{1i} , sendo $\hat{a}_{0i} = \hat{a}_0(i+n-1)$, $\hat{a}_{1i} = \hat{a}_1(i+n-1)$ e $\hat{b}_{1i} = \hat{b}_1(i+n-1)$, $i = 1, 2, \dots, T-n+1$, conforme Figura 5.1:

Figura 5.1: Esquema de seleção das sub-amostras



2º Passo: Após a execução de um modelo GARCH(1,1) para cada janela móvel, foi obtido um conjunto de estimativas para os parâmetros a_{0i} , a_{1i} e b_{1i} . Sejam os vetores $A_0 = (a_{01}, a_{02}, \dots, a_{0(T-n+1)})$, $A_1 = (a_{11}, a_{12}, \dots, a_{1(T-n+1)})$ e $B_1 = (b_{11}, b_{12}, \dots, b_{1(T-n+1)})$. Como eles são séries temporais, então podemos ajustar um novo modelo para cada um deles. Como à princípio estamos interessados nas estimativas médias de a_{0i} , a_{1i} e b_{1i} , $i=1, \dots, T-n+1$ e não na variabilidade, foi sugerido o ajuste do modelo ARIMA para as séries $A_0 = (a_{01}, a_{02}, \dots, a_{0(T-n+1)})$, $A_1 = (a_{11}, a_{12}, \dots, a_{1(T-n+1)})$ e $B_1 = (b_{11}, b_{12}, \dots, b_{1(T-n+1)})$. Assim a formulação do modelo será:

$$\sigma_i^2 = a_0(t) + a_1(t)\varepsilon_{i-1}^2 + b_1(t)\sigma_{i-1}^2, \quad (5.1)$$

em que $t=n, \dots, T$. Sendo:

$$\begin{aligned} a_0(t) &\cong ARIMA(p, d, q) & \phi(L)\Delta^d (a_0(t) - \mu_{a_0}) &= \theta(L)\varepsilon_{0,i}, \\ a_1(t) &\cong ARIMA(p_1, d_1, q_1) & \phi(L)\Delta^{d_1} (a_1(t) - \mu_{a_1}) &= \theta(L)\varepsilon_{1,i}, \\ b_1(t) &\cong ARIMA(p_2, d_2, q_2) & \phi(L)\Delta^{d_2} (b_1(t) - \mu_{b_1}) &= \theta(L)\varepsilon_{1,i}. \end{aligned}$$

com $a_0(t) > 0$, $a_1(t) \geq 0$, $b_1(t) \geq 0$ e $a_1(t) + b_1(t) \geq 0$.

No modelo proposto em (5.1), está sendo considerado que os parâmetros variam ao longo do tempo e além disso o padrão de variação pode ser explicado por um modelo ARIMA. Essa formulação é diferente do que é proposto na classe de modelos: “*functional coefficient regression model*”. Nessa classe de modelos os coeficientes em geral são funções dos valores observados no passado. Na sugestão apresentada neste trabalho os parâmetros são funções deles mesmos porém observados no passado, ou seja, é uma estrutura autoregressiva integrada.

A grande vantagem do modelo apresentado é que ele permite que o comportamento de variação dos parâmetros ao longo do tempo seja capturado, sem que sejam utilizados modelos muito complexos ou mesmo métodos de estimação difíceis de serem implementados.

5.3 Previsão utilizando a metodologia proposta

Um dos principais objetivos deste trabalho é apresentar uma forma alternativa do modelo GARCH tradicional de modo que fosse possível obter previsões mais eficientes. Conforme já discutido anteriormente, isso foi feito através da proposta de um modelo GARCH com coeficientes dependentes do tempo. Uma vez definido o modelo (5.1) e realizada a estimação dos parâmetros é necessário definir uma forma de realizar as previsões, já que esta é uma das principais utilizações do modelo no mercado financeiro.

Como este modelo é composto por duas partes:

- 1 - Um modelo GARCH com coeficientes dependentes do tempo;
- 2 - Modelo ARIMA para os coeficientes,

inicialmente serão realizadas as estimativas para o modelo ARIMA e a seguir as estimativas dos parâmetros do modelo GARCH para h passos à frente. Uma vez tendo estas estimativas, os modelos GARCH ficam configurados e estimação da volatilidade ocorre naturalmente.

1º Passo: Seja um modelo ARIMA(p,q,d) para $b_1(t)$:

$$b_1(t) - \mu = \phi_1(b_1(t-1) - \mu) + \dots + \phi_p(b_1(t-p) - \mu) + \varepsilon_t + \theta_1\varepsilon_{t-1} + \dots + \theta_q\varepsilon_{t-q},$$

$$\varepsilon_t \sim RB(0, \sigma^2).$$

Neste caso, estamos interessados em prever $b_1(T+h|T)$, tendo-se observações até o instante T. Chamemos de $\hat{b}_1(T+h|T)$, a previsão de $b_1(T+h|T)$ de origem T e horizonte h. Pode-se provar que a previsão de EQMM (erro quadrático médio mínimo) é dada pela esperança condicional de $b_1(T+h)$ dado o passado $b_1(T)$, $b_1(T-1)$, ..., ou seja,

$$\hat{b}_1(T+h|T) = E(\phi_1 b_1(T+h-1) + \dots + \phi_{p+d} b_1(T+h-p-d) + \theta_0 + \varepsilon_{T+h} - \theta_1 \varepsilon_{T+h-1} - \dots - \theta_q \varepsilon_{T+h-q} | b_1(T), b_1(T-1), \dots).$$

Para calcular as previsões serão utilizados dois fatos:

a) $E(b_1(T+j) | b_1(T), b_1(T-1), \dots) = b_1(T+j)$ se $j \leq 0$ ou

$E(b_1(T+j) | b_1(T), b_1(T-1), \dots) = \hat{b}_1(T+j|T)$ se $j > 0$.

b) $E(\varepsilon_{T+j} | b_1(T), b_1(T-1), \dots) = \varepsilon_{T+j}$ se $j \leq 0$ ou

$$E(\varepsilon_{1T+j} | b_1(T), b_1(T-1), \dots) = 0 \quad \text{se } j > 0.$$

Logo, para calcular previsões temos que:

- (a) substituir esperanças passadas $j \leq 0$ por valores conhecidos, $b_1(T+j)$ e ε_{T+j} .
- (b) Substituir esperanças futuras $j > 0$ por previsões $\hat{b}(T+j|T)$ e 0.

Por exemplo, suponha o modelo AR(2),

$$(1 - \phi_1 B - \phi_2 B^2)X_t = \phi_0 + \varepsilon_t.$$

Temos que

$$X_{T+h} = \phi_1 X_{T+h-1} + \phi_2 X_{T+h-2} + \phi_0 + \varepsilon_{T+h}.$$

Logo:

- i) Para $h=1$, temos $\hat{X}_T(1) = \phi_1 X_T + \phi_2 X_{T-1} + \phi_0$;
- ii) Para $h=2$, temos $\hat{X}_T(2) = \phi_1 \hat{X}_T(1) + \phi_2 X_T + \phi_0$;
- iii) Para $h>2$, temos $\hat{X}_T(h) = \phi_1 \hat{X}_T(h-1) + \phi_2 \hat{X}_T(h-2) + \phi_0$.

As previsões de $a_0(t)$ e $a_1(t)$ são realizadas de modo análogo.

2º Passo: Após a realização das previsões dos coeficientes do modelo GARCH através dos modelos ARIMA, estes valores são utilizados para realizar as previsões da volatilidade.

Seja o modelo proposto em (5.1):

$$\sigma_t^2 = a_0(t) + a_1(t)\varepsilon_{t-1}^2 + b_1(t)\sigma_{t-1}^2, \quad t=n, \dots, T$$

Neste caso, pode-se provar que o EQMM (erro quadrático médio mínimo) de $\hat{\sigma}_t^2$ é dado pela esperança condicional de σ_{T+k}^2 , $E_T[\sigma_{T+k}^2]$, que é a previsão da volatilidade futura σ_{T+k}^2 . Por exemplo, para obtê-la à partir da equação anterior considerando $k=1$, dada a informação até o tempo T , temos:

$$\begin{aligned} E_T[\sigma_{T+1}^2] &= a_0(T+1) + a_1(T+1)E_T[\varepsilon_T^2] + b_1(T+1)E_T[\sigma_T^2] \\ &= a_0(T+1) + a_1(T+1)\varepsilon_T^2 + b_1(T+1)\sigma_T^2, \end{aligned}$$

em que ε_T^2 e σ_T^2 são os valores obtidos depois da estimação (a notação utilizada se refere aos valores ajustados em vez de representar os valores “reais” não observados). Para $T+2$ temos:

$$\begin{aligned} E_T[\sigma_{T+2}^2] &= a_0(T+2) + a_1(T+2)E_T[\varepsilon_{T+1}^2] + b_1(T+2)E_T[\sigma_{T+1}^2] \\ &= a_0(T+2) + (a_1(T+2) + b_1(T+2))E_T[\sigma_{T+1}^2], \end{aligned}$$

em que $E_T[\varepsilon_{T+1}^2] = E_T[\sigma_{T+1}^2]$. Seguindo o mesmo raciocínio, pode-se obter a equação de previsão da volatilidade condicional:

$$E_T[\sigma_{T+k}^2] = a_0(T+k) \sum_{i=1}^{k-2} (a_1(T+i) + b_1(T+i))^i + (a_1(T+i) + b_1(T+i))^{k-1} E_T[\sigma_{T+1}^2],$$

para $k \geq 2$. Na equação acima, os parâmetros $a_0(T+i)$, $a_1(T+i)$ e $b_1(T+i)$ são as previsões obtidas à partir do modelo estimado ARIMA(p,d,q).

5.4 Avaliação do Desempenho do Modelo

Um dos principais objetivos na concepção deste trabalho foi tentar desenvolver um modelo mais preciso que o GARCH com relação à predição da volatilidade. Uma boa predição da volatilidade é fundamental no mercado

financeiro pois esta é uma boa forma de controle dos riscos envolvidos e da possibilidade de lucratividade do negócio.

Uma parte da volatilidade no mercado financeiro é causada pela ocorrência de fatos pontuais ou indeterminados. Este tipo de volatilidade é muito mais difícil de ser previsto pois esta condicionado à ocorrência de fatos isolados difíceis de serem previstos e que em geral causam grandes impactos no mercado. Este tipo de volatilidade não está contemplado no modelo sugerido assim como também não está nos modelos GARCH, porém acredita-se que probabilidade de ocorrência desses eventos é muito pequena.

Para avaliar se o modelo proposto apresenta um desempenho melhor que os modelos convencionais foi utilizada a seguinte função de perda:

$$\sum_{t=1}^N (\varepsilon_t^2 - \sigma_t^2)^2, \quad (5.2)$$

conforme Mills(1999), Pagan e Schwert (1990) e Bollerslev et al. (1994), em que:

$$\varepsilon_t = z_t \sigma_t$$

$$\sigma_t^2 = a_0 + \sum_{i=1}^r a_i \varepsilon_{t-i}^2 + \sum_{j=1}^s \sigma_{t-j}^2$$

sendo que, σ_t^2 é obtido das previsões da volatilidade n-passos à frente e ε_t é obtido com base nos valores observados da série de dados utilizada.

5.5 Considerações Finais

Conforme observado nas seções anteriores deste capítulo, o modelo proposto possibilita que a dinâmica do mercado seja modelada através da relação dos coeficientes do modelo com o tempo, sem que sejam necessárias formulações muito complexas, já que na sua essência o modelo apresenta uma concepção simples, combinando os modelos ARIMA e GARCH.

O foco deste trabalho foi na concepção do modelo, e na sua avaliação de forma empírica. Sabemos que ainda há muitos pontos que precisam ser desenvolvidos, principalmente na sua concepção teórica, porém eles fogem do escopo do trabalho. Desde o início, o objetivo deste trabalho foi encontrar uma alternativa eficiente em termos de previsão aos modelos GARCH convencionais, posto isto temos que o desenvolvimento teórico do mesmo foi desconsiderado porém o ideal é que ele seja retomado em trabalhos futuros.

Capítulo 6

Aplicações em Séries Reais

6.1 Introdução

Neste capítulo será aplicado o método proposto no capítulo 5 em 3 conjuntos de dados reais. Além do método apresentado também será considerado o ajuste de modelos GARCH, usualmente utilizado nestes tipos de séries. As aplicações foram realizadas em séries de retornos de preços de 3 ações muito conhecidas no mercado: Ibovespa, Vale e Itaú. Estas séries foram obtidas do site da Bovespa (Bolsa de Valores de São Paulo).

Foram consideradas 3 ações de segmentos bem diferentes pois dessa forma é possível avaliar com mais segurança a eficiência do método. Como o método está sendo avaliado apenas de forma empírica, com pouco rigor matemático, achamos melhor testá-lo em conjuntos de dados de segmentos diferentes e também na série do Ibovespa que é bem conhecida.

O software utilizado foi o R-Plus 2.6.0 e apresentaremos os códigos ao longo do capítulo. Muitos dos comandos utilizados fazem parte do módulo FinMetrics deste software.

6.2 Preços da Ação da Bovespa

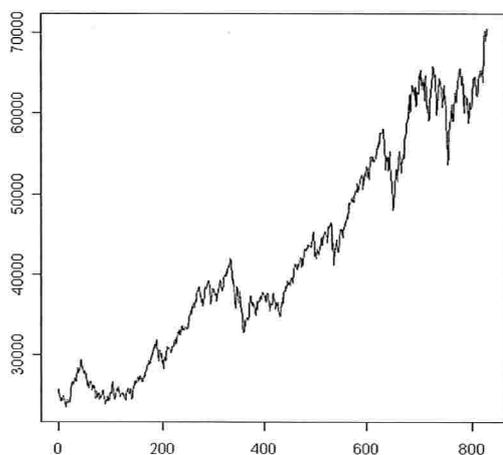
Os dados utilizados nesta aplicação referem-se aos valores diários do índice Bovespa (série IBOVESPA), de 03 de janeiro de 2005 a 12 de maio de 2008, totalizando 828 observações, além disso, foi utilizado o período de 13 de maio de 2008 à 27 de maio de 2008 na validação do modelo. Os dados apresentados indicam a pontuação de fechamento diário do Ibovespa. O ideal seria utilizar um período maior de análise, porém só foi possível ter acesso aos dados do período indicado acima.

Na figura 6.1, apresentamos (a) a série do ibovespa, (b) a série de retornos do ibovespa, (c) o histograma dos retornos do Ibovespa e (d) o gráfico QQ-plot dos retornos do Ibovespa.

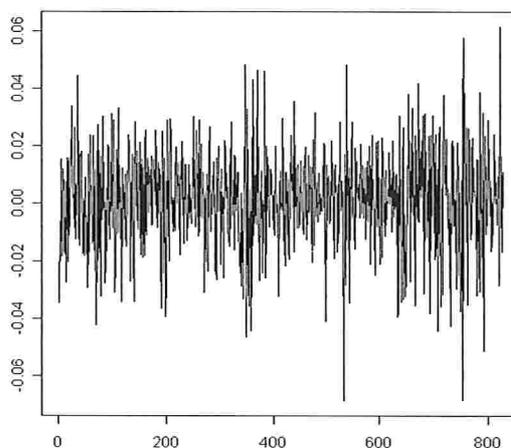
```

> dat<-read.table("C://Tiago/Series Temporais/Mestrado/Ibovespa_atualizado.txt", header=T)
> data<-ts(dat)
> plot(data,main="Série do Ibovespa",xlab="Tempo {03/01/2005 a 12/05/2008}",ylab="Preço")
> dif<-diff(log(data))
> plot(dif)
> hist(dif,probability=T,nclass=100,main="Histograma dos Retornos",xlab="Retornos",ylab="Densidade")
> dens=density(dif,n=200)
> points(dens,type="l")
> qqnorm(dif)
> qqline(dif)

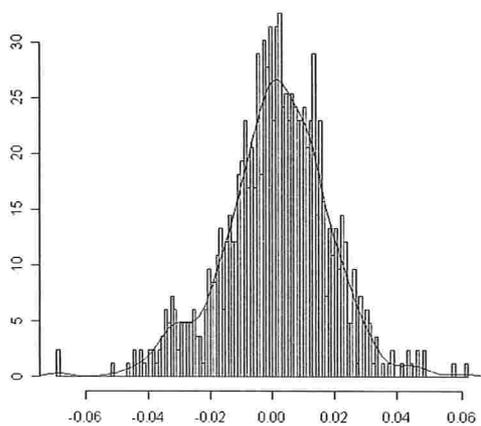
```



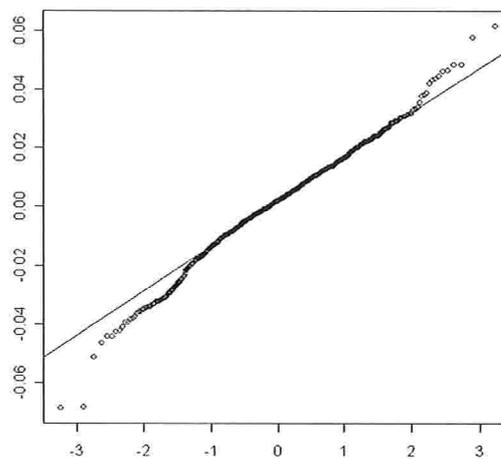
(a)



(b)



(c)



(d)

Figura 6.1: (a) Série do Ibovespa, (b) Série de Retornos do Ibovespa, (c) Histograma dos retornos do Ibovespa e (d) QQ-Plot dos Retornos do Ibovespa.

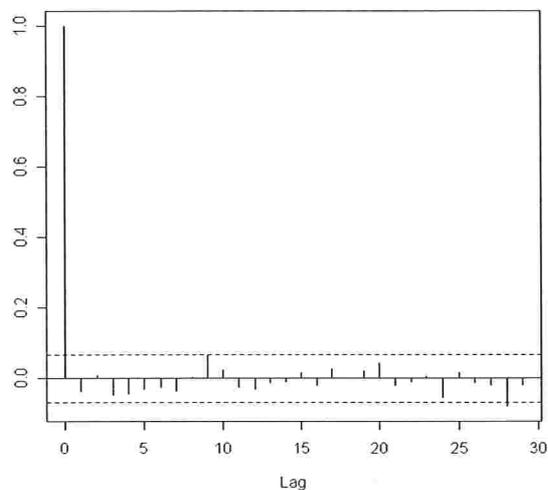
Nota-se pela Figura 6.1 (b) que a série apresenta alguns grupos de volatilidade mostrando indícios de que a volatilidade não seja constante ao longo do tempo. Em (d) pode-se observar que a distribuição da série apresenta caudas mais pesadas que a distribuição Normal.

(a) Modelo GARCH

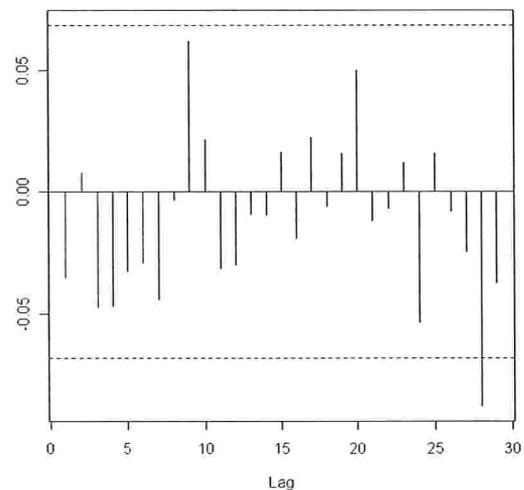
Inicialmente foi ajustado um modelo GARCH. O objetivo é comparar o desempenho deste método que é usualmente utilizado no mercado com o método proposto neste trabalho.

Na Figura 6.2, temos os gráficos das funções de autocorrelações (FAC) e autocorrelações parciais (FACP) da série de retornos do Ibovespa.

```
> acf(dif)
> pacf(dif)
> ret = dif*dif
> acf(ret)
> pacf(ret)
```



(a)



(b)

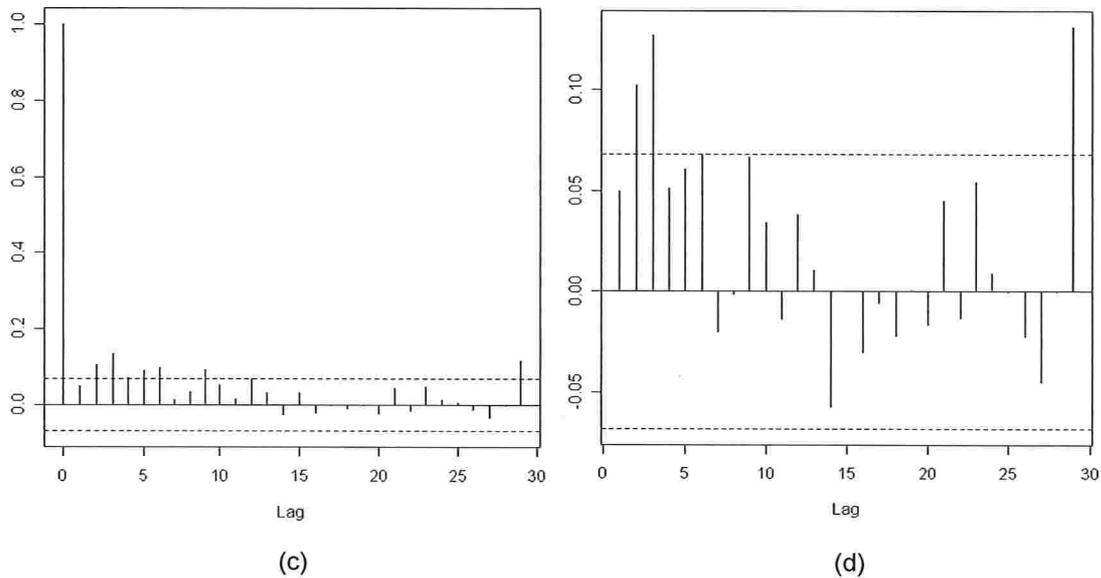


Figura 6.2: (a) FAC dos retornos, (b) FACP dos retornos, (c) FAC dos retornos ao quadrado, (d) FACP dos retornos ao quadrado.

Conforme observado nos gráficos acima, pode-se notar que não há muitas evidências de que haja autocorrelação entre os retornos porém entre os retornos ao quadrado os gráficos indicam fortes indícios da existência de autocorrelação, o que sugere o ajuste de um modelo GARCH.

Foi ajustado um modelo GARCH(1,1) assumindo inicialmente distribuição Normal para os erros, através do seguinte comando:

```
> obj<-garchFit(~garch(1,1),dif)
```

Conforme observado, todos os parâmetros são significantes ao nível de 5%, além disso, pode ser observado através do teste de Ljung Box tanto para os resíduos padronizados ε_t / σ_t quanto para os resíduos padronizados ao quadrado $(\varepsilon_t / \sigma_t)^2$ que ao nível de 5% não rejeitamos a hipótese H_0 de que eles são não correlacionados, ou seja, o modelo capturou com sucesso as estruturas de correlação serial na variância condicional.

Os testes de normalidade Jarque-Bera (p-valor<0,001) e Shapiro-Wilks (p-valor<0,01) para os resíduos padronizados indicam não normalidade.

```

> summary(obj)

Title:
  GARCH Modelling

Call:
  garchFit(formula = ~garch(1, 1), data = dif)

Mean and Variance Equation:
  ~arma(0, 0) + ~garch(1, 1)

Conditional Distribution:
  dnorm

Coefficient(s):
      mu      omega    alpha1    beta1
1.58931e-03  1.99178e-05  6.95483e-02  8.59273e-01

Error Analysis:
      Estimate Std. Error t value Pr(>|t|)
mu      1.589e-03  5.499e-04   2.890  0.00385 **
omega   1.992e-05  9.167e-06   2.173  0.02979 *
alpha1  6.955e-02  2.217e-02   3.137  0.00171 **
beta1   8.593e-01  4.732e-02  18.159 < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Log Likelihood:
-2225.631   normalized: -2.691211

```

```

Standardized Residuals Tests:

Jarque-Bera Test  R      Chi^2  50.22588  1.240474e-11
Shapiro-Wilk Test R      W      0.9894615  1.173239e-05
Ljung-Box Test   R      Q(10)  9.486548  0.4866325
Ljung-Box Test   R      Q(15)  11.02742  0.7506484
Ljung-Box Test   R      Q(20)  13.76853  0.8420354
Ljung-Box Test   R^2    Q(10)  6.346684  0.7853441
Ljung-Box Test   R^2    Q(15)  11.02388  0.7508994
Ljung-Box Test   R^2    Q(20)  14.30689  0.8146013
LM Arch Test     R      TR^2    7.489912  0.8236153

Information Criterion Statistics:
      AIC      BIC      SIC      HQIC
5.392095  5.414914  5.392049  5.400847

Description:
Sat Dec 26 15:36:44 2009 by user: Acer

```

A Figura 6.3 apresenta o gráfico QQ-plot dos resíduos padronizados e o gráfico das autocorrelações do quadrado dos resíduos padronizados. Observamos no gráfico QQ-plot que a suposição de normalidade não é apropriada, confirmando os resultados dos testes apresentados anteriormente, assim será ajustado a seguir um novo modelo GARCH(1,1) assumindo distribuição t de Student para os erros.

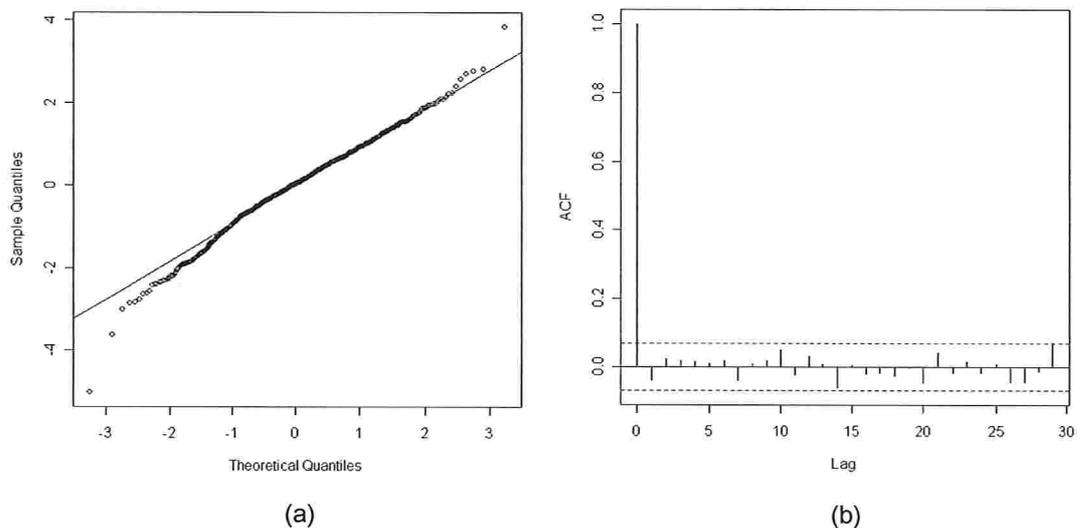


Figura 6.3: (a) QQ-Plot dos resíduos padronizados, (b) Autocorrelação do quadrado dos resíduos padronizados referentes ao modelo com distribuição Normal.

Agora vamos ajustar um modelo GARCH(1,1) supondo distribuição t de Student para os erros (foi considerada a opção `include.shape=FALSE`, ou seja, os graus de liberdade foram fixados durante o processo de estimação e eles são default do R-Plus):

```
> obj<-garchFit(~garch(1,1), dif, cond.dist="dstd", include.shape=FALSE)
```

Observa-se pelos critérios BIC e AIC que o modelo assumindo distribuição normal para os erros é melhor. Apesar dos valores das medidas BIC e AIC estarem bem próximos para ambos os modelos, o modelo com erros normais apresentou valores levemente menores.

Através da Figura 6.4 (a), pode-se notar que a hipótese de que os dados seguem uma distribuição t-Student não é válida, o que pode ser confirmado através dos testes de Jarque-Bera ($p\text{-valor} < 0,001$) e Shapiro-Wilks ($p\text{-valor} < 0,001$).

Assim como no caso do modelo com distribuição normal, aqui também foi possível capturar com sucesso as estruturas de correlação serial na variância condicional conforme pode ser observado na Figura 6.4 (b).

Dado os resultados obtidos, foi selecionado o modelo assumindo distribuição normal, conforme segue:

$$\hat{X}_t = \hat{\sigma}_t \varepsilon_t,$$

$$\hat{\sigma}_t^2 = 0,00002215 + 0,1003 * (X_{t-1} - 0,00208)^2 + 0,8593 * \hat{\sigma}_{t-1}^2. \quad (6.1)$$

```
> summary(obj)

Title:
GARCH Modelling

Call:
garchFit(formula = ~garch(1, 1), data = dif, cond.dist = "dstd",
include.shape = FALSE)

Mean and Variance Equation:
~arma(0, 0) + ~garch(1, 1)

Conditional Distribution:
dstd

Coefficient(s):
      mu      omega      alpha      beta1
2.08009e-03  2.21513e-05  1.00287e-01  8.59279e-01

Error Analysis:
      Estimate Std. Error t value Pr(>|t|)
mu      2.080e-03  5.172e-04  4.022 5.78e-05 ***
omega   2.215e-05  1.224e-05  1.810 0.07022 .
alpha   1.003e-01  3.514e-02  2.854 0.00431 **
beta1   8.593e-01  5.128e-02  16.756 < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Log Likelihood:
-2226.147      normalized: -2.691834
```

```
Standardized Residuals Tests:
      Statistic p-Value
Jarque-Bera Test  R  Chi^2  55.27254  9.947598e-13
Shapiro-Wilk Test  R  W      0.989093  8.03358e-06
Ljung-Box Test     R  Q(10)  9.560293  0.4798786
Ljung-Box Test     R  Q(15)  11.15907  0.741239
Ljung-Box Test     R  Q(20)  13.80381  0.840302
Ljung-Box Test     R^2 Q(10)  6.175253  0.8003305
Ljung-Box Test     R^2 Q(15)  11.38261  0.7250306
Ljung-Box Test     R^2 Q(20)  14.98179  0.7774486
LM Arch Test       R  TR^2   7.163644  0.8466124

Information Criterion Statistics:
      AIC      BIC      SIC      HQIC
5.393341  5.416160  5.393295  5.402093

Description:
Thu Dec 31 09:05:18 2009 by user: Acer
```

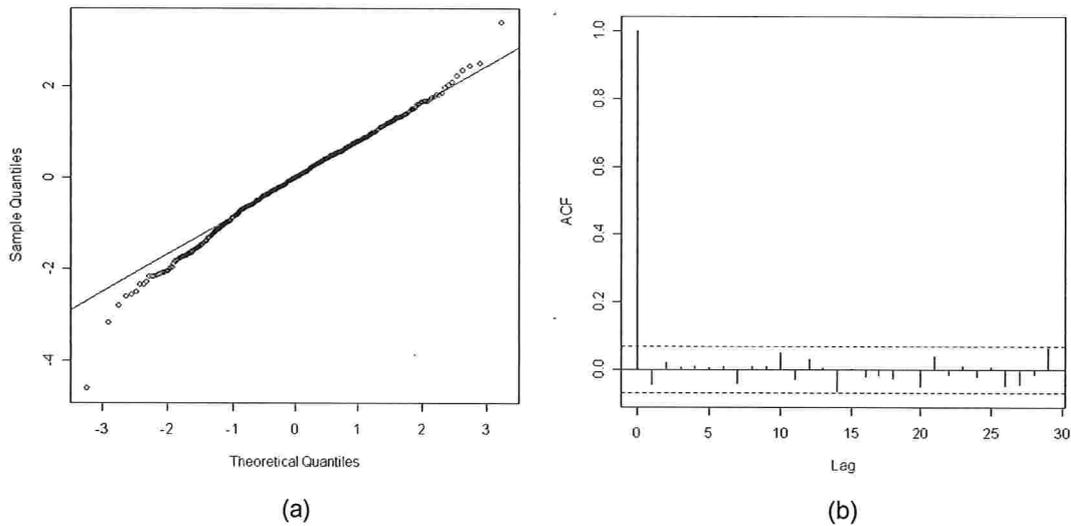


Figura 6.4: (a) QQ-Plot dos resíduos padronizados, (b) Autocorrelação do quadrado dos resíduos padronizados referentes ao modelo com distribuição t de Student.

(b) Modelo com coeficientes variando no tempo

Após o ajuste dos modelos GARCH apresentados anteriormente foi ajustado um modelo com coeficientes variando no tempo conforme a metodologia apresentada no capítulo 5.

A série de retornos do Ibovespa utilizada refere-se à 827 observações e foi considerada no ajuste desse modelo uma janela móvel de tamanho 300 o que possibilitou rodar 528 modelos distintos. Esta quantidade foi definida de maneira arbitrária já que este é um dos pontos de melhoria que podem ser pesquisados no futuro. O modelo considerado foi o GARCH(1,1) com distribuição normal, dessa forma, para cada um dos 528 modelos foram obtidos quatro parâmetros: $\alpha(t)$, $\beta(t)$ (coeficientes do modelo), $\mu(t)$ (a média geral do modelo) e $\omega(t)$ (o intercepto do modelo). Ao final desse processo, foi obtido um vetor de parâmetros com 4 colunas e 528 linhas, onde as colunas indicam os 4 parâmetros e as linhas indicam os valores dos parâmetros para cada um dos 528 modelos. O vetor de parâmetros foi obtido com a utilização dos seguintes comandos:

```
e <- matrix(0,528,4)
m<-matrix(0,827)
m<-1:827
aux<-cbind(m,dif)

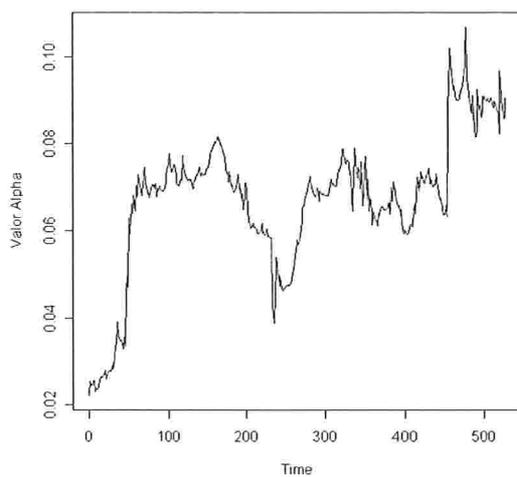
for (i in 1:528) {
  dados<-subset(aux,aux[,1]>(i-1) & aux[,1]<(i+300))
  dados<-dados[301:600,drop=TRUE]
  options(show.error.messages = FALSE)
  options(warn = -1)
  obj<-garchFit(~garch(1,1),dados,trace=FALSE)
  t<-obj@fit$par
  e[i,]<-t
}
```

Após a obtenção deste vetor de parâmetros, cada coluna foi tratada com uma série temporal e foi ajustado um modelo autoregressivo integrado para cada uma. O objetivo é tentar identificar algum padrão de comportamento dos parâmetros ao longo do tempo e tentar capturá-lo através de modelos autoregressivos integrados.

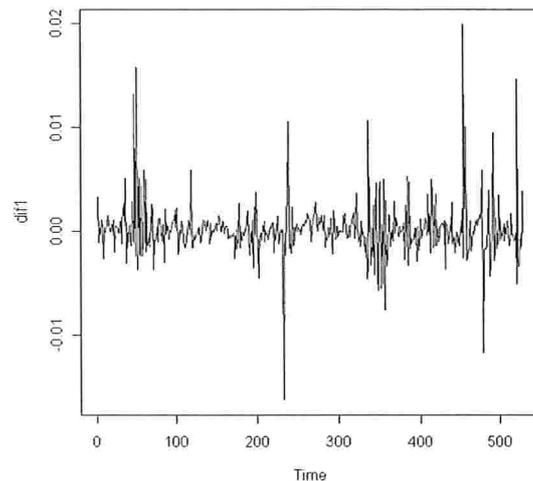
Abaixo seguem os modelos ajustados para cada um dos parâmetros:

Ajuste do modelo para a série de parâmetros $\alpha(t)$

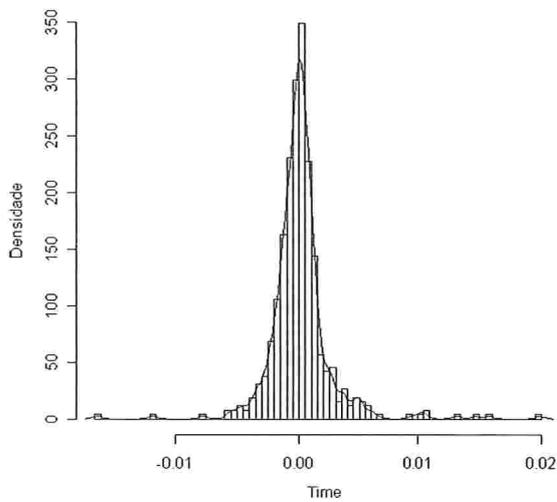
Na figura 6.5, (a) apresentamos a série original dos $\alpha(t)$, (b) a série da 1ª diferença de $\alpha(t)$, (c) o histograma da 1ª diferença de $\alpha(t)$ e (d) o gráfico QQ-plot da 1ª diferença de $\alpha(t)$.



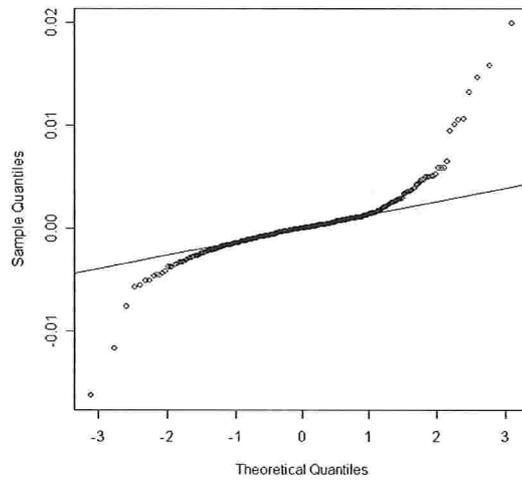
(a)



(b)



(c)



(d)

Nota-se pela Figura 6.5 (b) que tomando a 1ª diferença da série dos $\alpha(t)$, a mesma se torna estacionária, porém a série apresenta alguns valores muito altos ou baixos, referentes aos saltos que podem ser visto na série original em 6.5 (a). Estes saltos podem ser causados pelo fato da série considerada ser pequena ou devido à escala apresentar um range pequeno de variação. Estes aspectos são importante pois podem prejudicar o ajuste do modelo. Em (d) pode-se observar que a distribuição da série destoa da Normal já que as caudas são um pouco mais pesadas.

Na Figura 6.6, temos os gráficos das funções de autocorrelações (FAC) e autocorrelações parciais (FACP) da série da 1ª diferença de $\alpha(t)$ e da 1ª diferença ao quadrado.

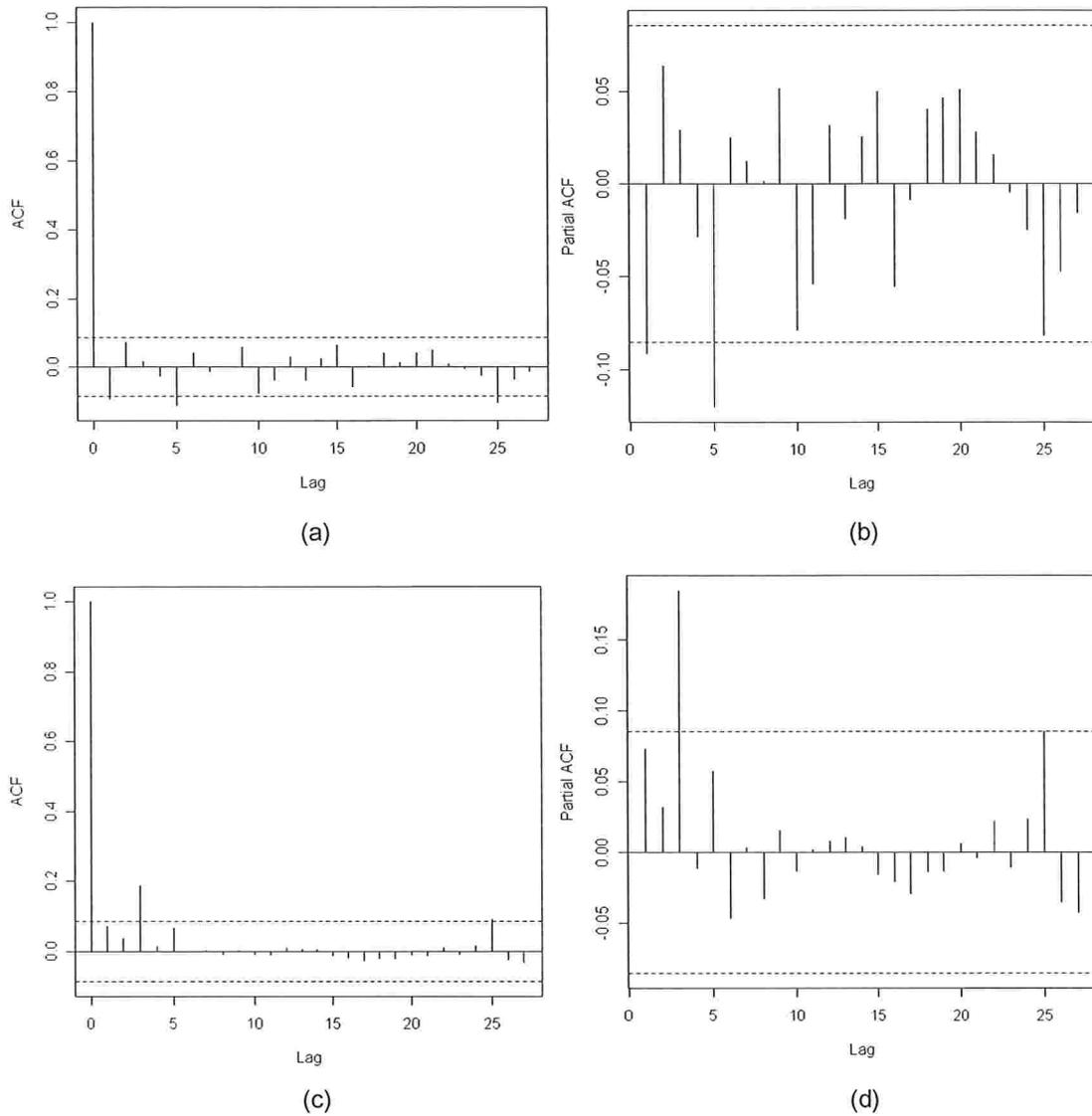


Figura 6.6: (a) FAC da 1ª diferença de $\alpha(t)$, (b) FACP da 1ª diferença de $\alpha(t)$, (c) FAC da 1ª diferença de $\alpha(t)$ ao quadrado, (d) FACP da 1ª diferença de $\alpha(t)$ ao quadrado.

Conforme observado na Figura 6.6 (a) e (b), pode-se notar que há evidências de que exista autocorrelação de lag 1 e principalmente de lag 5. Observando a série referente à 1ª diferença dos $\alpha(t)$ ao quadrado, pode-se notar que também há fortes evidências da existência de autocorrelação o que implicaria no ajuste de um modelo GARCH, porém como estamos interessados nas

estimativas de $\alpha(t)$ e não na sua variabilidade, nos restringiremos em modelar apenas a média da série e não sua volatilidade.

Foi ajustado um modelo AR(5) sendo que os parâmetros AR=2, 3 e 4 foram desconsiderados. Foi utilizado o seguinte comando no ajuste do modelo:

```
> Alpha<- arma(dif1, lag=list(ar=c(1,5)), include.intercept=FALSE)
```

```
> summary(Alpha)
Call:
arma(x = dif1, lag = list(ar = c(1, 5)), include.intercept = FALSE)

Model:
ARMA(5,0)

Residuals:
    Min       1Q   Median       3Q      Max
-1.631e-02 -9.155e-04 -2.299e-05  8.733e-04  1.949e-02

Coefficient(s):
      Estimate Std. Error t value Pr(>|t|)
ar1  -0.09074    0.04322   -2.099   0.0358 *
ar5  -0.11052    0.04321   -2.558   0.0105 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Fit:
sigma^2 estimated as 6.507e-06, Conditional Sum-of-Squares = 0, AIC = -4794.24
```

Conforme observado, todos os parâmetros são significantes ao nível de 5%. O teste de Box-Pierce (p-valor=0,9009) para os resíduos padronizados ε_t / σ_t , indica que ao nível de 5% não rejeitamos a hipótese H_0 de que os dados sejam independentes, ou seja, o modelo capturou com sucesso as estruturas de correlação serial o que também pode ser observado pela Figura 6.7.

```
> Box.test(res_Alpha)

      Box-Pierce test

data:  res_Alpha
X-squared = 0.0155, df = 1, p-value = 0.9009
```

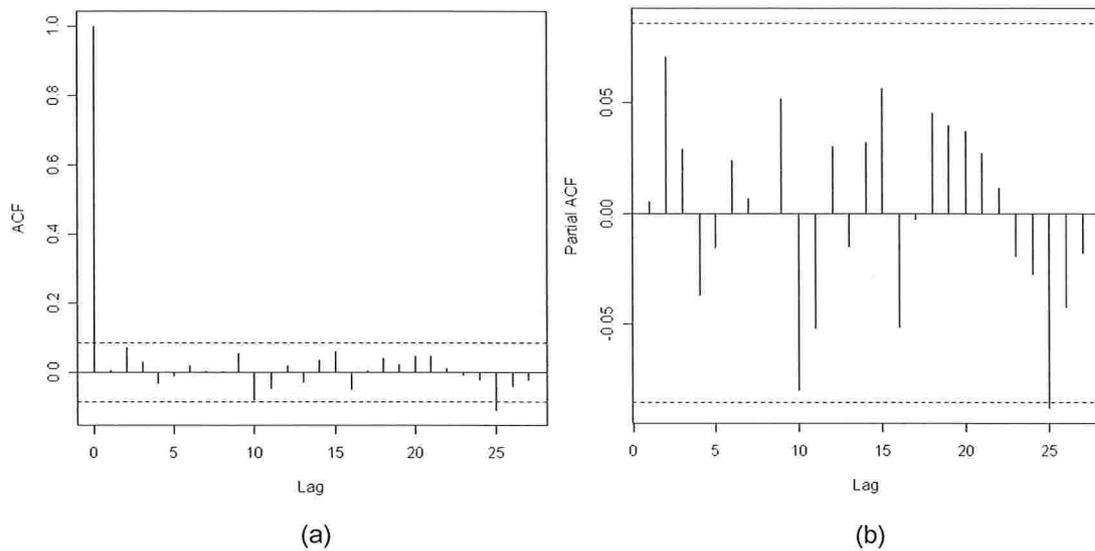


Figura 6.7: (a) FAC dos resíduos padronizados, (b) FACP dos resíduos padronizados

Os testes de normalidade Jarque-Bera (p-valor=0) e Shapiro-Wilks (p-valor=0) para os resíduos padronizados indicam não normalidade.

```
> jarque.bera.test(res_Alpha_norm)

      Jarque Bera Test

data:  res_Alpha_norm
X-squared = 5647.962, df = 2, p-value < 2.2e-16

> shapiro.test(res_Alpha_norm)

      Shapiro-Wilk normality test

data:  res_Alpha_norm
W = 0.7821, p-value < 2.2e-16
```

A Figura 6.8 apresenta o gráfico QQ-plot dos resíduos padronizados e o gráfico das autocorrelações dos resíduos padronizados. Observamos no gráfico QQ-plot que a suposição de normalidade não é apropriada, confirmando os resultados dos testes apresentados anteriormente.

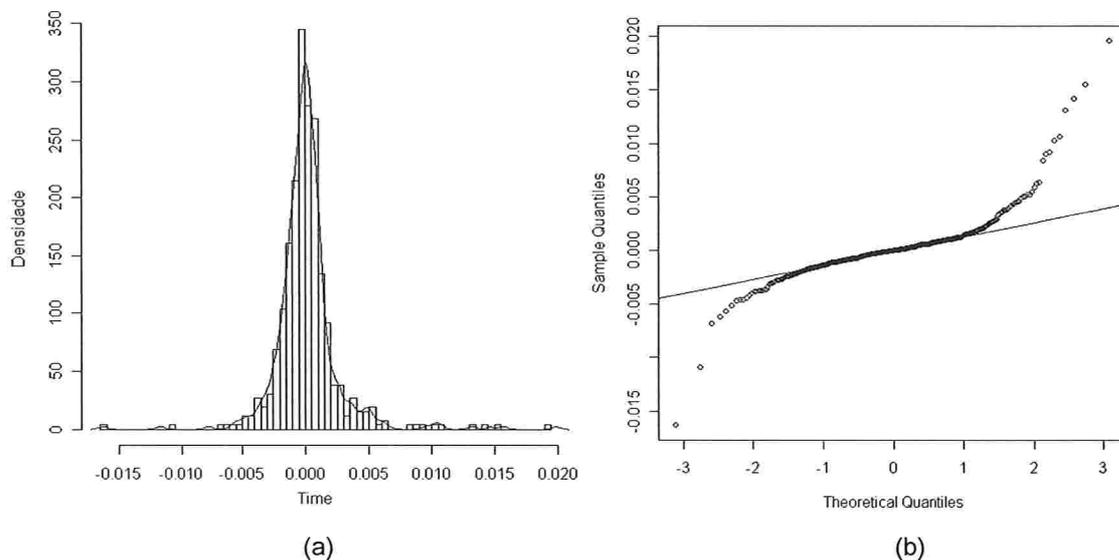


Figura 6.8: (a) Histograma dos resíduos padronizados, (b) QQ-Plot dos resíduos padronizados

Dados os resultados obtidos no ajuste, o modelo final é:

$$(\hat{\alpha}(t) - \hat{\alpha}(t-1)) = -0,09074 * (\hat{\alpha}(t-1) - \hat{\alpha}(t-2)) - 0,11052 * (\hat{\alpha}(t-5) - \hat{\alpha}(t-6)) \quad (6.2)$$

Apesar do modelo ter conseguido capturar com sucesso a estrutura de correlação dos dados, a série apresentou alguns pontos extremos conforme observado na Figura 6.5 (b), dessa forma, é importante ter-se cuidado ao analisar o modelo.

Ajuste do modelo para a série de parâmetros $\beta(t)$

Na figura 6.9, apresentamos (a) a série original dos $\beta(t)$, (b) a série da 1ª diferença de $\beta(t)$, (c) o histograma da 1ª diferença de $\beta(t)$ e (d) o gráfico QQ-plot da 1ª diferença de $\beta(t)$.

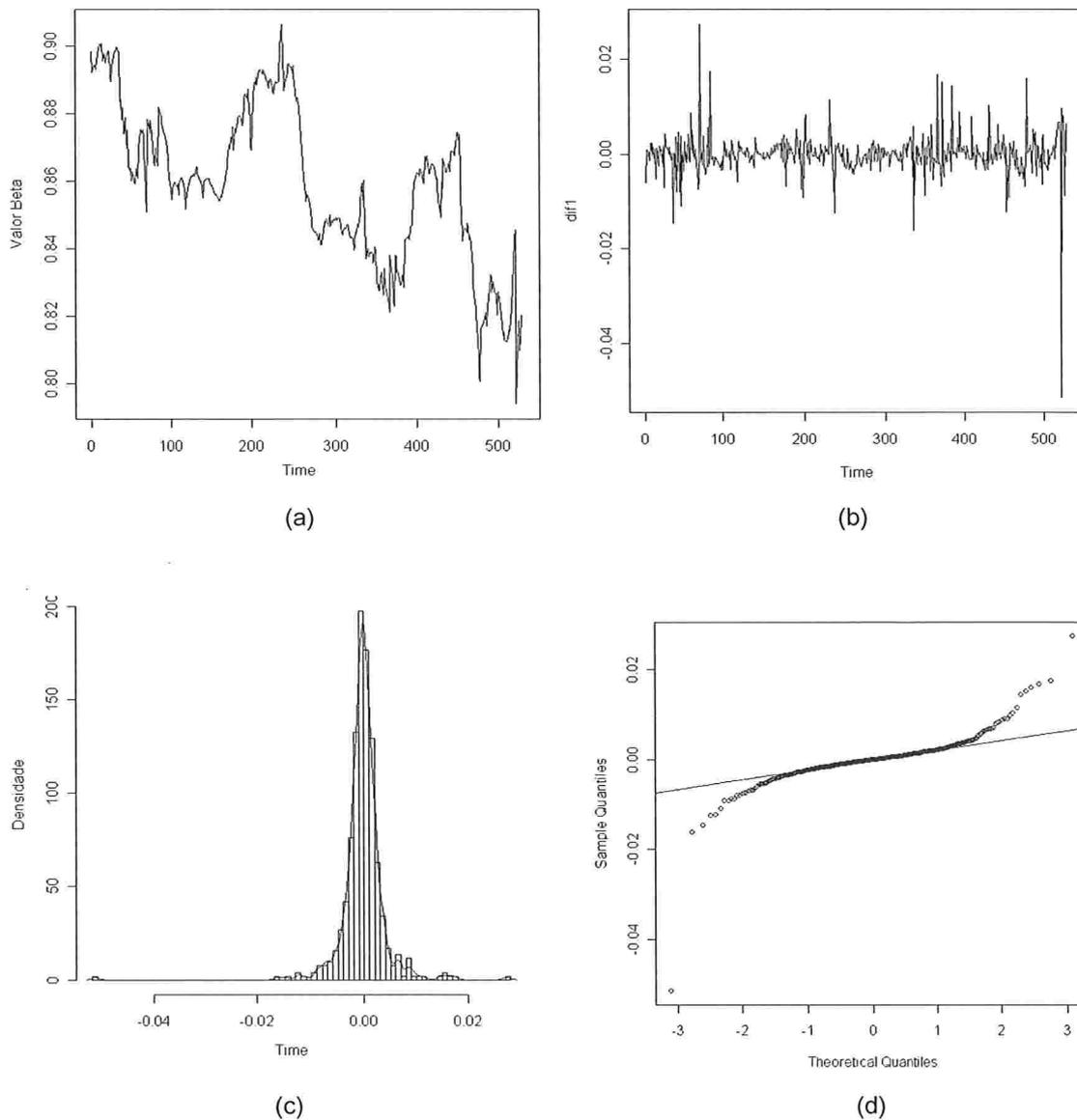


Figura 6.9: (a) Série original de $\beta(t)$, (b) Série da 1ª diferença de $\beta(t)$, (c) Histograma da 1ª diferença de $\beta(t)$ e (d) QQ-Plot da 1ª diferença de $\beta(t)$.

Na Figura 6.9 (b) pode-se ver que tomando a 1ª diferença da série dos $\beta(t)$, a mesma se torna estacionária, porém a série apresenta alguns valores muito altos principalmente no final do período o que pode prejudicar o ajuste do modelo. Em (d) pode-se observar que a distribuição da série parece destoar da distribuição Normal principalmente nas caudas.

Na Figura 6.10, são apresentados os gráficos das funções de autocorrelações (FAC) e autocorrelações parciais (FACP) da série da 1ª diferença dos $\beta(t)$ e da 1ª diferença ao quadrado.

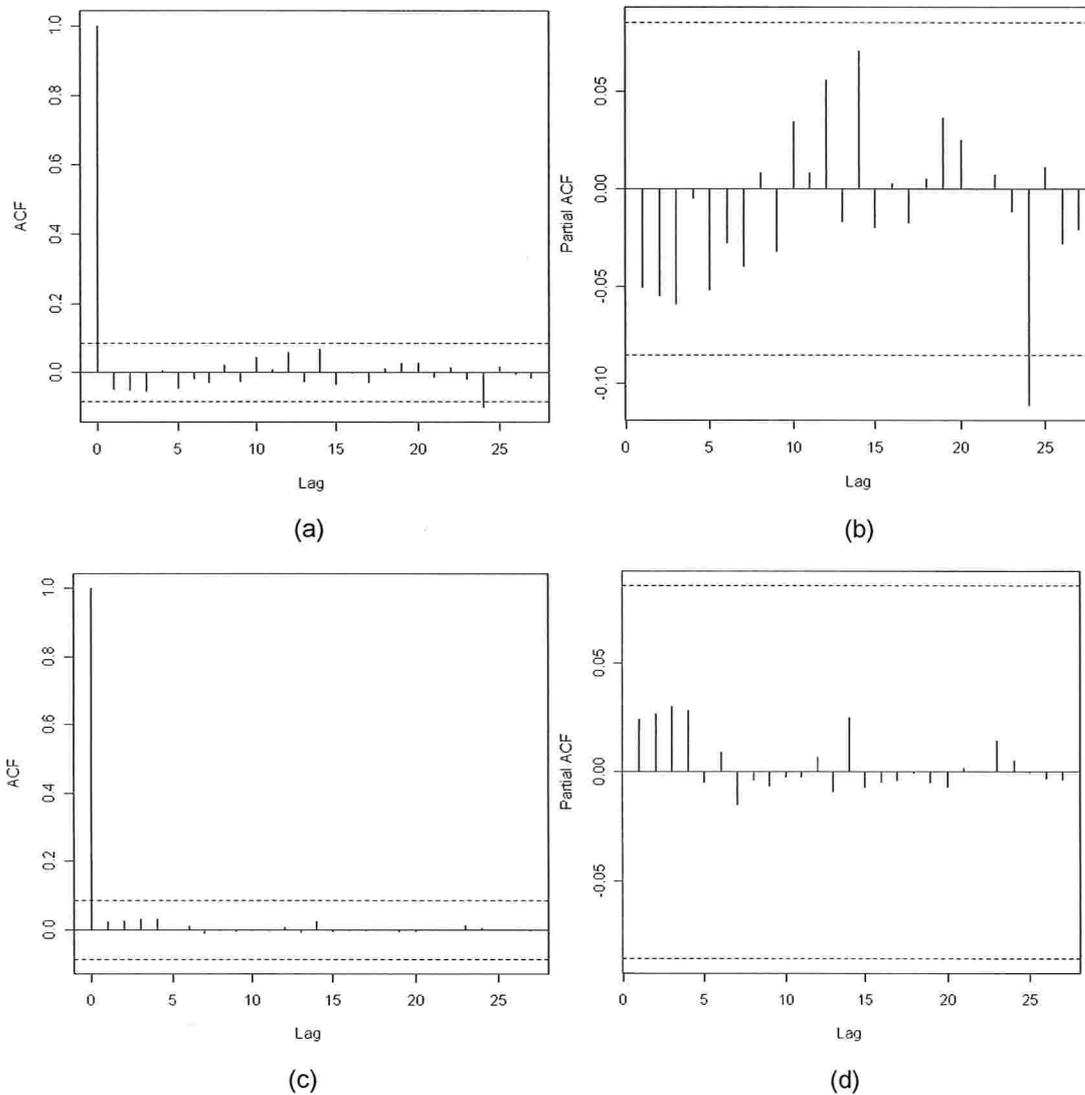


Figura 6.10: (a) FAC da 1ª diferença de $\beta(t)$, (b) FACP da 1ª diferença de $\beta(t)$, (c) FAC da 1ª diferença de $\beta(t)$ ao quadrado, (d) FACP da 1ª diferença de $\beta(t)$ ao quadrado.

Na Figura 6.10 (a) e (b), pode-se notar que há evidências de que exista autocorrelação de lag 24 na série de $\beta(t)$. Observando a série referente à 1ª diferença de $\beta(t)$ ao quadrado, Figura 6.10 (c) e (d), pode-se notar que neste caso não evidências da existência de autocorrelação, dessa forma, será ajustado um

modelo AR(24) sendo que todos os parâmetros diferentes de AR=24 foram desconsiderados. Foi utilizado o seguinte comando no ajuste do modelo:

```
> Beta<- arma(dif1, lag=list(ar=c(24)), include.intercept=FALSE)
```

```
> summary(Beta)

Call:
arma(x = dif1, lag = list(ar = c(24)), include.intercept = FALSE)

Model:
ARMA(24,0)

Residuals:
    Min       1Q   Median       3Q      Max
-0.0514907 -0.0016071 -0.0001607  0.0013186  0.0257599

Coefficient(s):
      Estimate Std. Error t value Pr(>|t|)
ar24   -0.1459    0.0523   -2.789  0.00528 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Fit:
sigma^2 estimated as 1.872e-05, Conditional Sum-of-Squares = 0.01, AIC = -4239.29
```

Conforme observado, o parâmetro estimado referente ao lag 24 é significativo ao nível de 5%, além disso, pode ser observado através do teste de Box-Pierce (p-valor=0,2724) para os resíduos padronizados ε_t / σ_t , que ao nível de 5% não rejeitamos a hipótese H_0 de que eles são independentes, ou seja, o modelo capturou com sucesso as estruturas de correlação serial o que também pode ser observado pela Figura 6.11.

Não é normal a utilização de valores de lags tão altos como este de 24. Isto pode ser reflexo do tamanho da série ser pequeno, por isso deve-se ter cuidado ao analisar este modelo.

```
> Box.test(res_beta)

      Box-Pierce test

data:  res_beta
X-squared = 1.2044, df = 1, p-value = 0.2724
```

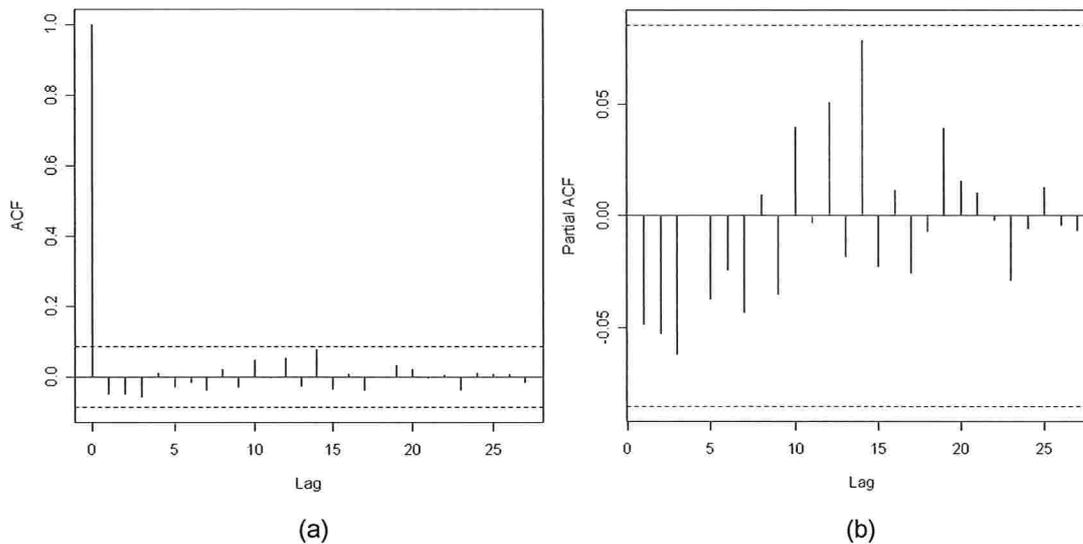


Figura 6.11: (a) FAF dos resíduos padronizados, (b) FACP dos resíduos padronizados

Os testes de normalidade Jarque-Bera (p-valor=0) e Shapiro-Wilks (p-valor=0) para os resíduos padronizados indicam não normalidade.

```
> jarque.bera.test(res_beta_norm)

      Jarque Bera Test

data:  res_beta_norm
X-squared = 483.8456, df = 2, p-value < 2.2e-16

> shapiro.test(res_beta_norm)

      Shapiro-Wilk normality test

data:  res_beta_norm
W = 0.8727, p-value = 5.962e-16
```

A Figura 6.12 apresenta o gráfico QQ-plot dos resíduos padronizados e o gráfico das autocorrelações dos resíduos padronizados. Observamos no gráfico QQ-plot que a suposição de normalidade não é apropriada, conforme já visto anteriormente.

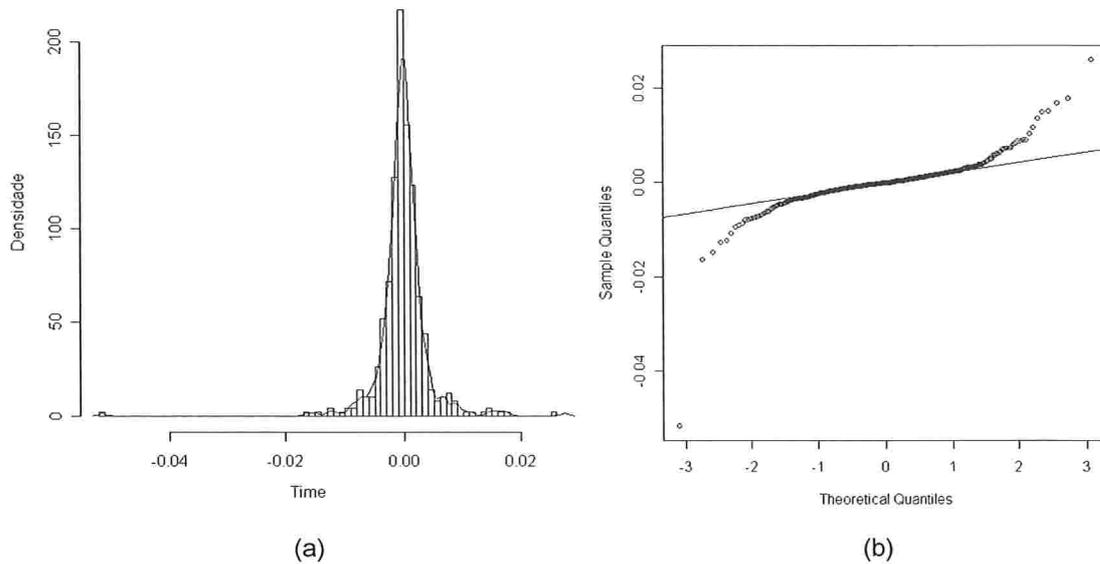


Figura 6.12: (a) Histograma dos resíduos padronizados, (b) QQ-Plot dos resíduos padronizados

Dados os resultados obtidos no ajuste, o modelo final é:

$$(\hat{\beta}(t) - \hat{\beta}(t-1)) = -0,1459 * (\hat{\beta}(t-24) - \hat{\beta}(t-25)) \quad (6.3)$$

Ajuste do modelo para a série de parâmetros $\mu(t)$

Na figura 6.13, apresentamos (a) a série original de $\mu(t)$, (b) a série da 1ª diferença de $\mu(t)$, (c) o histograma da 1ª diferença de $\mu(t)$ e (d) o gráfico QQ-plot da 1ª diferença de $\mu(t)$.

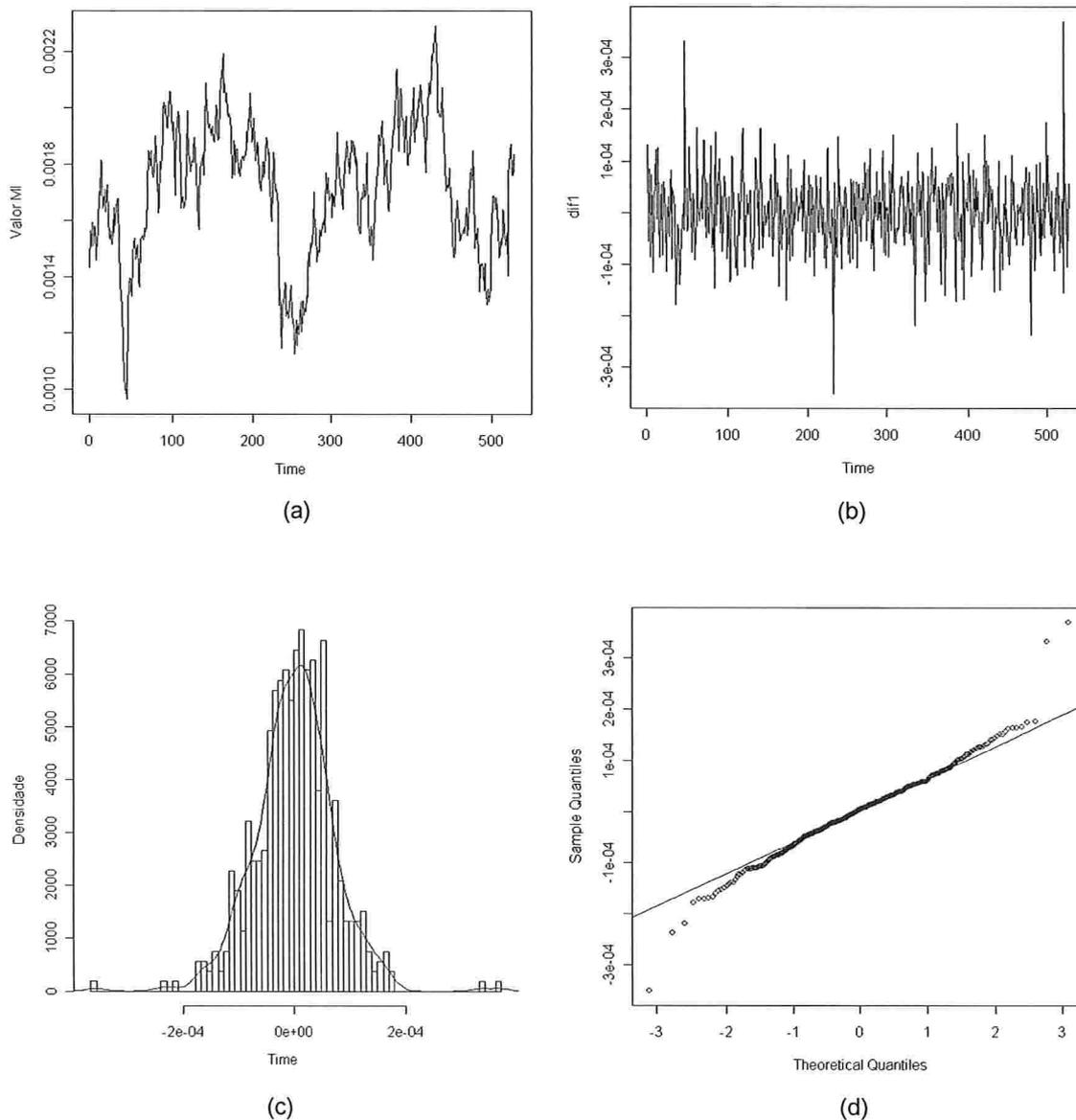


Figura 6.13: (a) Série original dos $\mu(t)$, (b) Série da 1ª diferença dos $\mu(t)$, (c) Histograma da 1ª diferença dos $\mu(t)$ e (d) QQ-Plot da 1ª diferença dos $\mu(t)$.

Na Figura 6.13 (b) pode-se ver que tomando a 1ª diferença da série de $\mu(t)$, a mesma se torna estacionária. Em d) pode-se observar que a distribuição da série parece se aproximar de uma normal com exceção de alguns pontos.

Na Figura 6.14, são apresentados os gráficos das funções de autocorrelações (FAC) e autocorrelações parciais (FACP) da série da 1ª diferença dos $\mu(t)$ e da 1ª diferença ao quadrado.

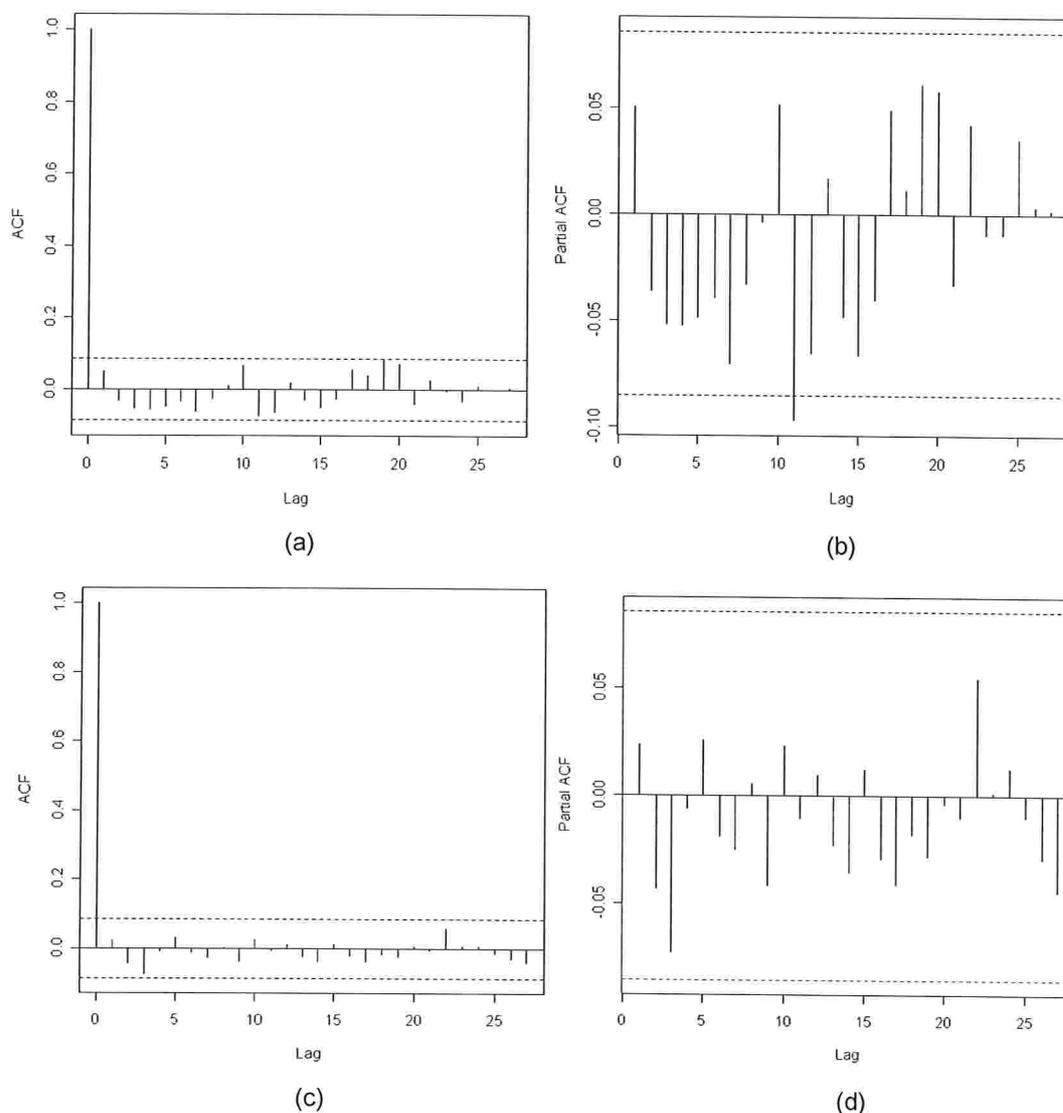


Figura 6.14: (a) FAC da 1ª diferença de $\mu(t)$, (b) FACP da 1ª diferença de $\mu(t)$, (c) FAC da 1ª diferença de $\mu(t)$ ao quadrado, (d) FACP da 1ª diferença de $\mu(t)$ ao quadrado.

Na Figura 6.14, pode-se notar que há pouca evidência de que exista autocorrelação na série referente à 1ª diferença de $\mu(t)$, de qualquer forma, será ajustado um modelo AR(11) em que todos os parâmetros diferentes de AR=11 serão desconsiderados de modo a confirmar esta hipótese.

```
> MI<- arma(dif1, lag=list(ar=c(11)), include.intercept=FALSE)
```

```
> summary(MI)

Call:
arma(x = dif1, lag = list(ar = c(11)), include.intercept = FALSE)

Model:
ARMA(11,0)

Residuals:
    Min       1Q   Median       3Q      Max
-3.520e-04 -3.988e-05  3.448e-06  4.311e-05  3.699e-04

Coefficient(s):
            Estimate Std. Error t value Pr(>|t|)
ar11 -0.08129      0.04430   -1.835  0.0665 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Fit:
sigma^2 estimated as 5.172e-09, Conditional Sum-of-Squares = 0, AIC = -8557.61
```

Conforme observado, adotando um nível de significância de 5% não rejeitamos a hipótese H_0 de que o coeficiente do modelo ajustado seja 0, conforme já visto na Figura acima.

Neste caso, como os dados não apresentaram correlação serial será considerado o seguinte modelo:

$$\mu(t) = \mu(t-1) + \varepsilon_t \quad (6.4)$$

Ajuste do modelo para a série de parâmetros $\alpha_0(t)$

Na Figura 6.15, apresentamos (a) a série original de $\alpha_0(t)$, (b) a série da 1ª diferença de $\alpha_0(t)$, (c) o histograma da 1ª diferença de $\alpha_0(t)$ e (d) o gráfico QQ-plot da 1ª diferença de $\alpha_0(t)$.

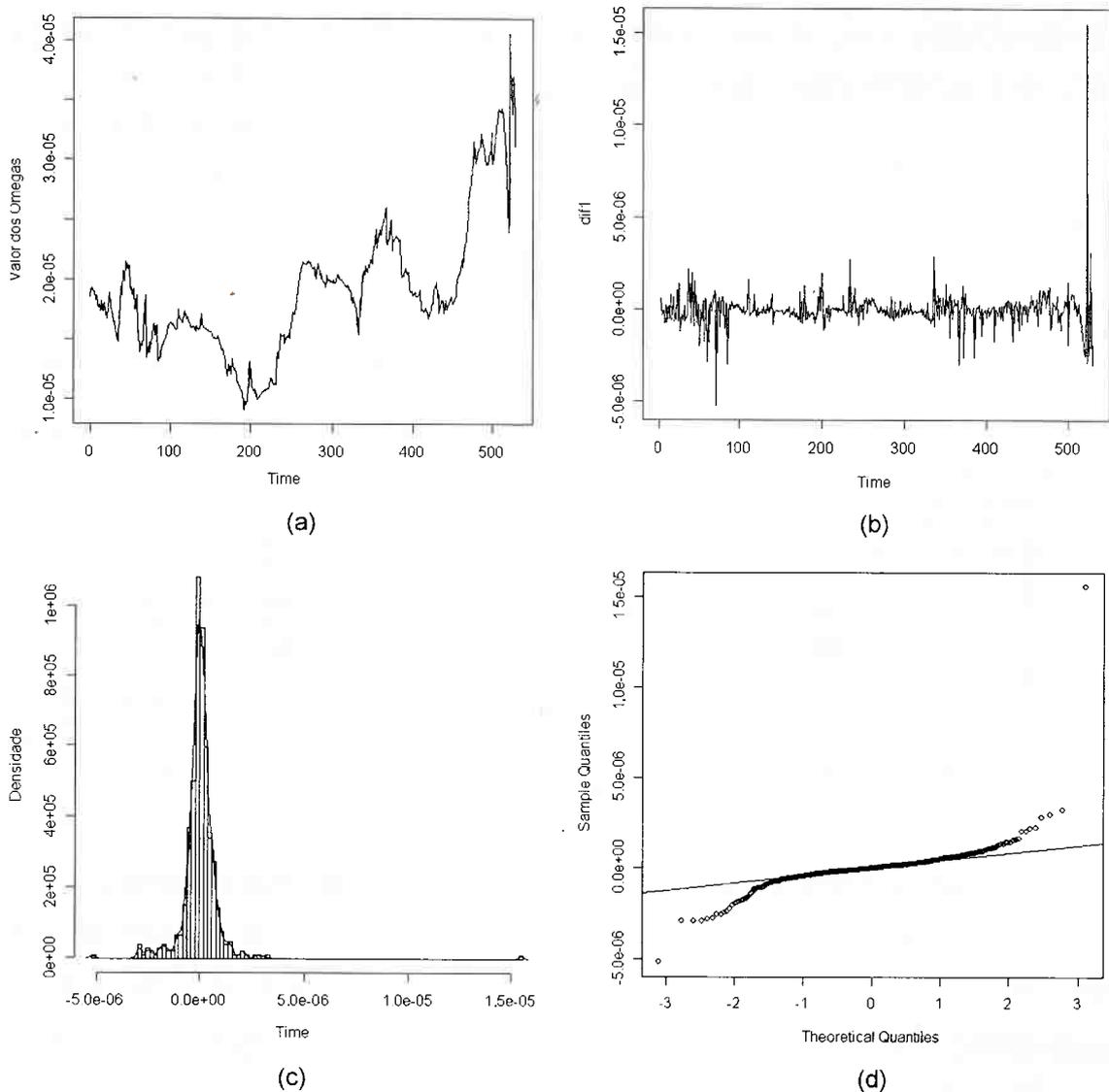


Figura 6.15: (a) Série original de $\alpha_0(t)$, (b) Série da 1ª diferença de $\alpha_0(t)$, (c) Histograma da 1ª diferença de $\alpha_0(t)$ e (d) QQ-Plot da 1ª diferença de $\alpha_0(t)$.

Na Figura 6.15 (b) pode-se ver que tomando a 1ª diferença da série de $\beta(t)$, a mesma se torna estacionária e também pode-se observar alguns agrupamentos de volatilidade indicando que a mesma não seja constante ao longo do tempo, além disso, em (d) pode-se observar que a distribuição da série parece destoar da distribuição Normal.

Na Figura 6.16, são apresentados os gráficos das funções de autocorrelações (FAC) e autocorrelações parciais (FACP) da série da 1ª diferença de $\alpha_0(t)$ e da da 1ª diferença ao quadrado.

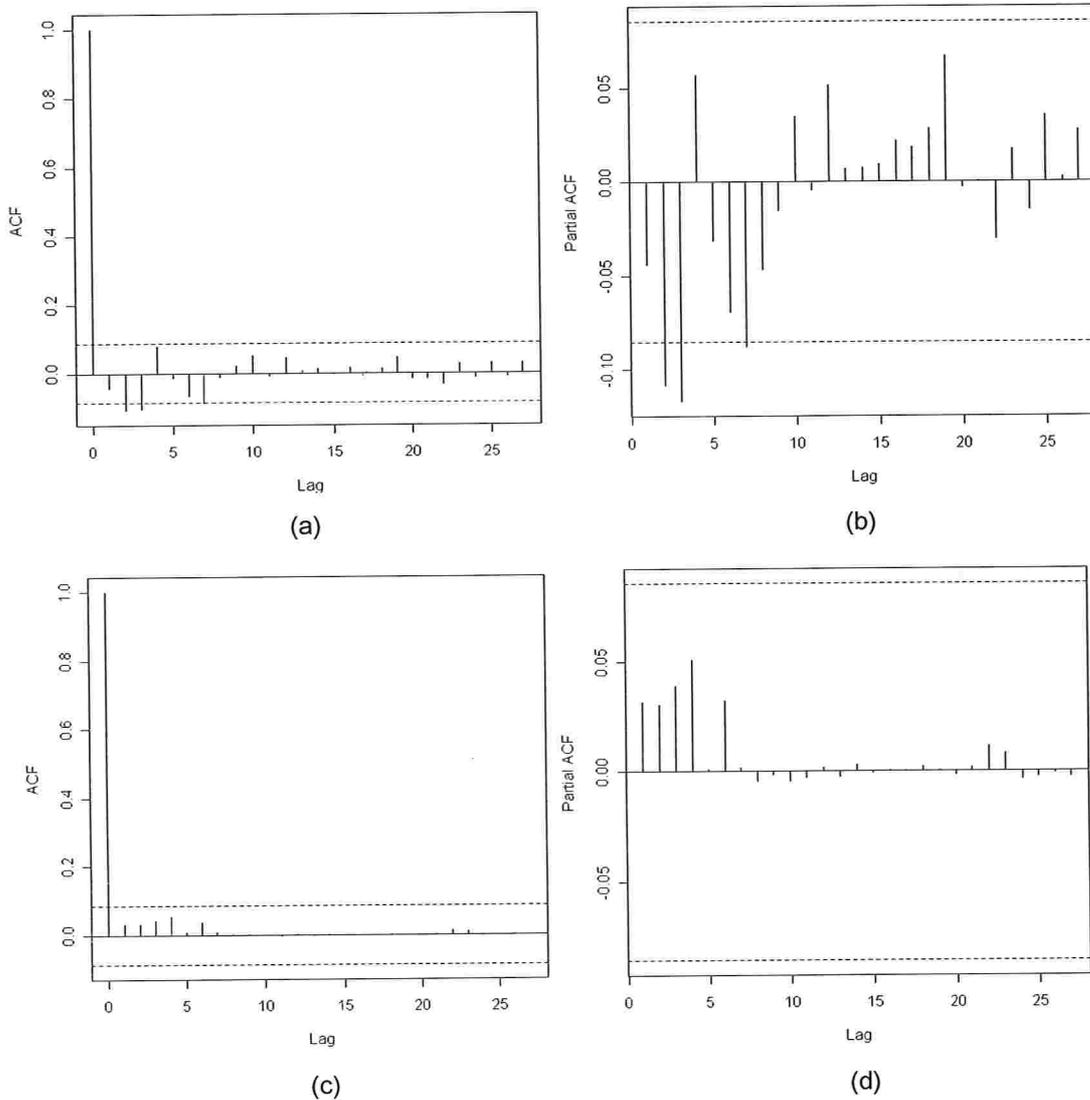


Figura 6.16: (a) FAC da 1ª diferença de $\alpha_0(t)$, (b) FACP da 1ª diferença de $\alpha_0(t)$, (c) FAC da 1ª diferença de $\alpha_0(t)$ ao quadrado, (d) FACP da 1ª diferença de $\alpha_0(t)$ ao quadrado.

Na Figura 6.16 (a) e (b), pode-se notar que há evidências de que exista autocorrelação de lag 2 e lag 3. Observando a série referente à 1ª diferença dos $\alpha_0(t)$ ao quadrado, Figura 6.16 (c) e (d), o mesmo comportamento não se repete já

que todas as autocorrelações estimadas caem dentro do intervalo indicando a não existência de autocorrelação, dessa forma, será ajustado um modelo AR(3), sendo que o parâmetro AR=1 foi desconsiderado. Foi utilizado o seguinte comando no ajuste do modelo:

```
> Omega<- arma(dif1, lag=list(ar=c(2,3)),include.intercept=FALSE)
```

```
> summary(Omega)

Call:
arma(x = dif1, lag = list(ar = c(2, 3)), include.intercept = FALSE)

Model:
ARMA(3,0)

Residuals:
    Min       1Q   Median       3Q      Max
-4.953e-06 -2.653e-07  4.821e-09  3.357e-07  1.497e-05

Coefficient(s):
      Estimate Std. Error t value Pr(>|t|)
ar2  -0.11482    0.04361   -2.633  0.00847 **
ar3  -0.11378    0.04365   -2.607  0.00914 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Fit:
sigma^2 estimated as 9.826e-13, Conditional Sum-of-Squares = 0, AIC = -13071.23
```

Conforme observado, o parâmetros estimados referentes ao lag 2 e lag 3 são significantes ao nível de 5%, além disso pode ser observado através do teste de Box-Pierce (p-valor=0,1864) para os resíduos padronizados ε_t / σ_t , que ao nível de 5% não rejeitamos a hipótese H_0 de que eles são independentes o que também pode ser observado pela Figura 6.17.

```
> Box.test(res_Omega)

      Box-Pierce test

data:  res_Omega
X-squared = 1.7459, df = 1, p-value = 0.1864
```

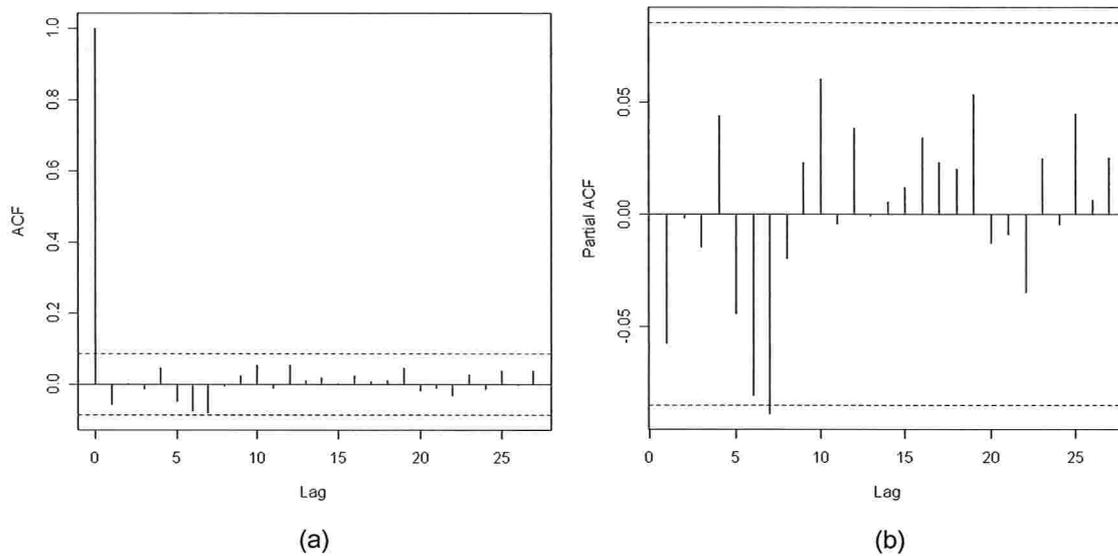


Figura 6.17: (a) FAC dos resíduos padronizados, (b) FACP dos resíduos padronizados

Os testes de normalidade Jarque-Bera (p -valor=0) e Shapiro-Wilks (p -valor=0) para os resíduos padronizados indicam não normalidade.

```

> jarque.bera.test(res_Omega_norm)

      Jarque Bera Test

data:  res_Omega_norm
X-squared = 217792.2, df = 2, p-value < 2.2e-16

> shapiro.test(res_Omega_norm)

      Shapiro-Wilk normality test

data:  res_Omega_norm
W = 0.6218, p-value < 2.2e-16

```

Na Figura 6.18, observamos no gráfico QQ-plot que a suposição de normalidade não é apropriada, confirmando os resultados dos testes apresentados anteriormente.

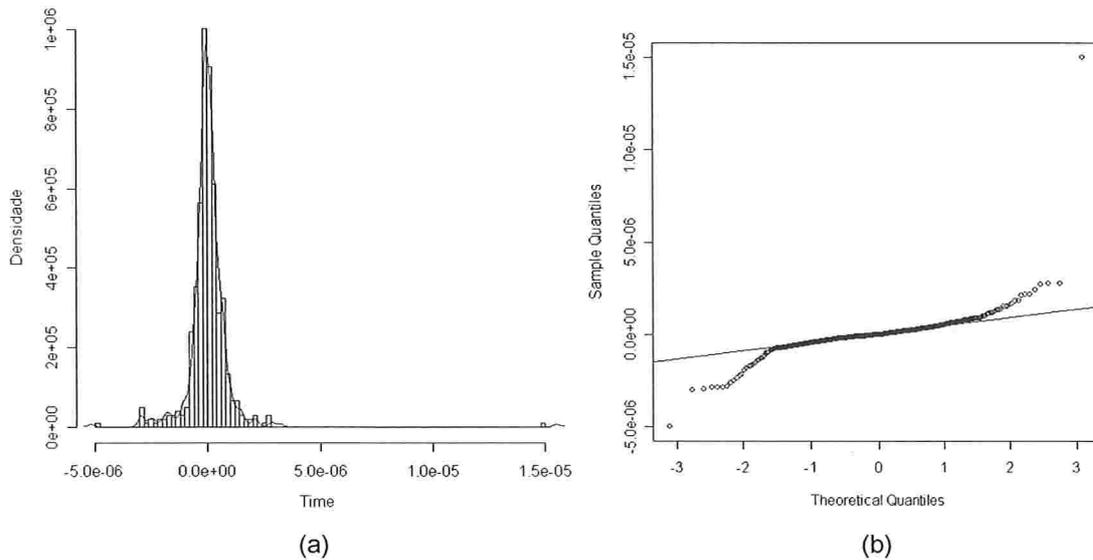


Figura 6.18: (a) Histograma dos resíduos padronizados, (b) QQ-Plot dos resíduos padronizados

Dados os resultados obtidos no ajuste, o modelo final é:

$$(\hat{\alpha}_0(t) - \hat{\alpha}_0(t-1)) = -0,11482 * (\hat{\alpha}_0(t-2) - \hat{\alpha}_0(t-3)) - 0,11378 * (\hat{\alpha}_0(t-3) - \hat{\alpha}_0(t-4)) \quad (6.5)$$

Previsões

Após o desenvolvimento dos modelos citados anteriormente, eles foram utilizados para previsão. Foram realizadas previsões e os valores previstos foram comparados com os valores reais da série referentes ao período 13 de maio de 2008 à 27 de maio de 2008. Para cada metodologia (modelo GARCH convencional e modelo com coeficientes variando no tempo), foi calculado e comparado o EQM (Erro Quadrático Médio). O objetivo é verificar se houve algum ganho na utilização da metodologia proposta, ou seja, se conseguimos reduzir o EQM. Seguem os resultados obtidos:

Cálculo do EQM – Modelo GARCH(1,1) com distribuição Normal

Date	Preço Fechamento	Log(P_t / P_{t-1})	Predito - Garch(1,1) Dist Normal	Erro Quadrático Médio
			Previsão σ_t^2	Garch(1,1)
9/5/2008	69.646	-	-	
12/5/2008	70.416	0,0048	-	
13/5/2008	70.503	0,0005	0,0003814	1,45239E-07
14/5/2008	70.027	-0,0029	0,0003881	1,43992E-07
15/5/2008	71.492	0,0090	0,0003946	9,84219E-08
16/5/2008	72.767	0,0077	0,0004008	1,16853E-07
19/5/2008	73.439	0,0040	0,0004067	1,52711E-07
20/5/2008	73.517	0,0005	0,0004124	1,69921E-07
21/5/2008	72.295	-0,0073	0,0004179	1,3316E-07
23/5/2008	71.452	-0,0051	0,0004232	1,57775E-07
26/5/2008	71.629	0,0011	0,0004282	1,82366E-07
27/5/2008	70.992	-0,0039	0,0004330	1,74711E-07
				1,47515E-06

Tabela 6.1: Cálculo do EQM para o modelo GARCH(1,1)

A previsão da variância condicional σ_t^2 foi obtida com base no modelo GARCH(1,1) com distribuição Normal ajustado. O erro quadrático médio foi calculado segundo (5.2):

$$\sum_{t=1}^N (X_t^2 - \sigma_t^2)^2 \quad (6.6)$$

em que: $X_t = \log(P_t / P_{t-1})$ e $N = 10$ (dia 13/052008 à 27/05/2008)

Somando os valores para cada um dos 10 dias obtemos um EQM igual a $1,47515 e^{-6}$ para o modelo GARCH(1,1).

Cálculo do EQM – Modelo com coeficientes variando no tempo

Date	Preço Fecham.	Log (P _t / P _{t-1})	Metodologia Proposta					Previsão	Erro Quadrático Médio
			α ₁ (t)	β(t)	μ(t)	α ₀ (t)	σ _t ²	Metodologia Proposta	
9/5/2008	69.646	-	-	-	-	-	-	-	
12/5/2008	70.416	0,0048	-	-	-	-	-	-	
13/5/2008	70.503	0,0005	0,0906	0,8234	0,0018	0,0000317	0,0003466	1,199E-07	
14/5/2008	70.027	-0,0029	0,0907	0,8233	0,0018	0,0000322	0,0003177	9,54888E-08	
15/5/2008	71.492	0,0090	0,0908	0,8236	0,0018	0,0000324	0,0002961	4,63142E-08	
16/5/2008	72.767	0,0077	0,0907	0,8236	0,0018	0,0000322	0,0002807	4,91943E-08	
19/5/2008	73.439	0,0040	0,0903	0,8236	0,0018	0,0000322	0,0002665	6,27676E-08	
20/5/2008	73.517	0,0005	0,0903	0,8237	0,0018	0,0000322	0,0002521	6,3433E-08	
21/5/2008	72.295	-0,0073	0,0903	0,8235	0,0018	0,0000322	0,0002399	3,49514E-08	
23/5/2008	71.452	-0,0051	0,0903	0,8225	0,0018	0,0000322	0,0002371	4,45651E-08	
26/5/2008	71.629	0,0011	0,0903	0,8221	0,0018	0,0000322	0,0002314	5,30174E-08	
27/5/2008	70.992	-0,0039	0,0903	0,8215	0,0018	0,0000322	0,0002223	4,29708E-08	
								6,12603E-07	

Tabela 6.2: Cálculo do EQM para o modelo com coeficientes variando no tempo

As previsões 10 passos à frente para $\alpha_0(t)$, $\alpha_1(t)$, $\beta(t)$ e $\mu(t)$ mostradas na Tabela 6.2 foram obtidas com base nos modelos (6.2, 6.3, 6.4 e 6.5) ajustados para cada parâmetro apresentados anteriormente.

A previsão da variância condicional σ_t^2 foi obtida conforme segue:

$$\hat{\sigma}_t^2 = \hat{\alpha}_0(t) + \hat{\beta}(t) * \hat{\sigma}_{t-1}^2 + \hat{\alpha}_1(t) * (X_{t-1} - \hat{\mu}(t))^2 \quad (6.7)$$

em que: $X_{t-1} = \log(P_{t-1} / P_{t-2})$

O EQM foi calculado da mesma forma que no modelo GARCH(1,1) e conforme observado na Tabela 6.2 o valor obtido foi $6,12603 e^{-7}$, dessa forma podemos ver que a utilização da metodologia proposta reduziu consideravelmente o EQM.

Calculando a razão entre os EQMs, obtemos:

$$\text{Razão} = \frac{6,12603e^{-7}}{1,47515e^{-6}} = 0.415,$$

indicando que foi possível reduzir o EQM em mais da metade com a metodologia proposta.

6.3 Preços da Ação do Itaú (Itau4)

Os dados utilizados nesta aplicação referem-se aos valores diários de fechamento da ação do Itaú (ITAU4), de 03 de outubro de 2005 à 20 de março de 2008, totalizando 644 observações. Em 03 de outubro de 2005 houve uma alteração nos valores da ação do Itau4, conforme pode ser observado na Figura 6.19 o que fez com que o preço da ação caísse bruscamente inviabilizando a utilização dos dados referentes ao período anterior à esta data. Foi utilizado o período de 21 de março de 2008 à 04 de abril de 2008 na validação do modelo.

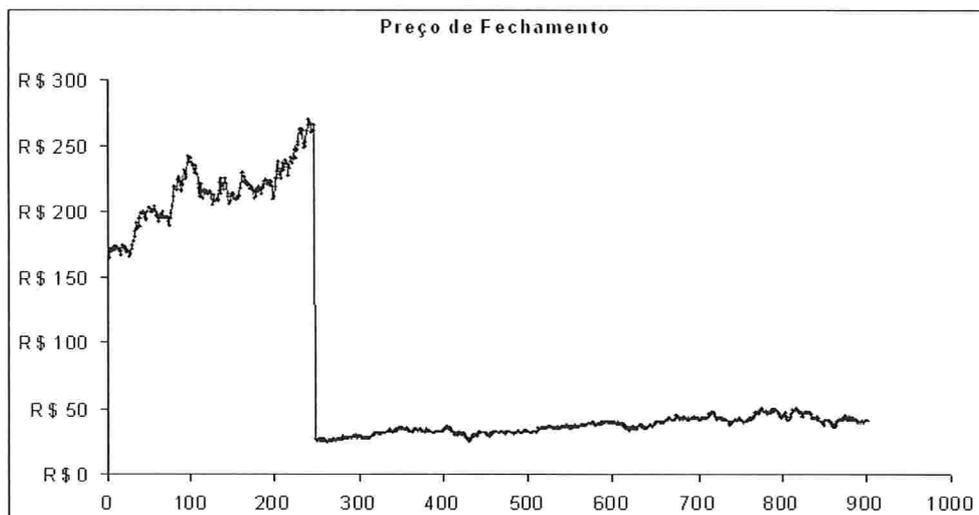


Figura 6.19: Preço de fechamento da Ação do Itaú

Na Figura 6.20, apresentamos (a) a série do itau4, (b) a série de retornos do itau4, (c) o histograma dos retornos do itau4 e (d) o gráfico QQ-plot dos retornos do itau4.

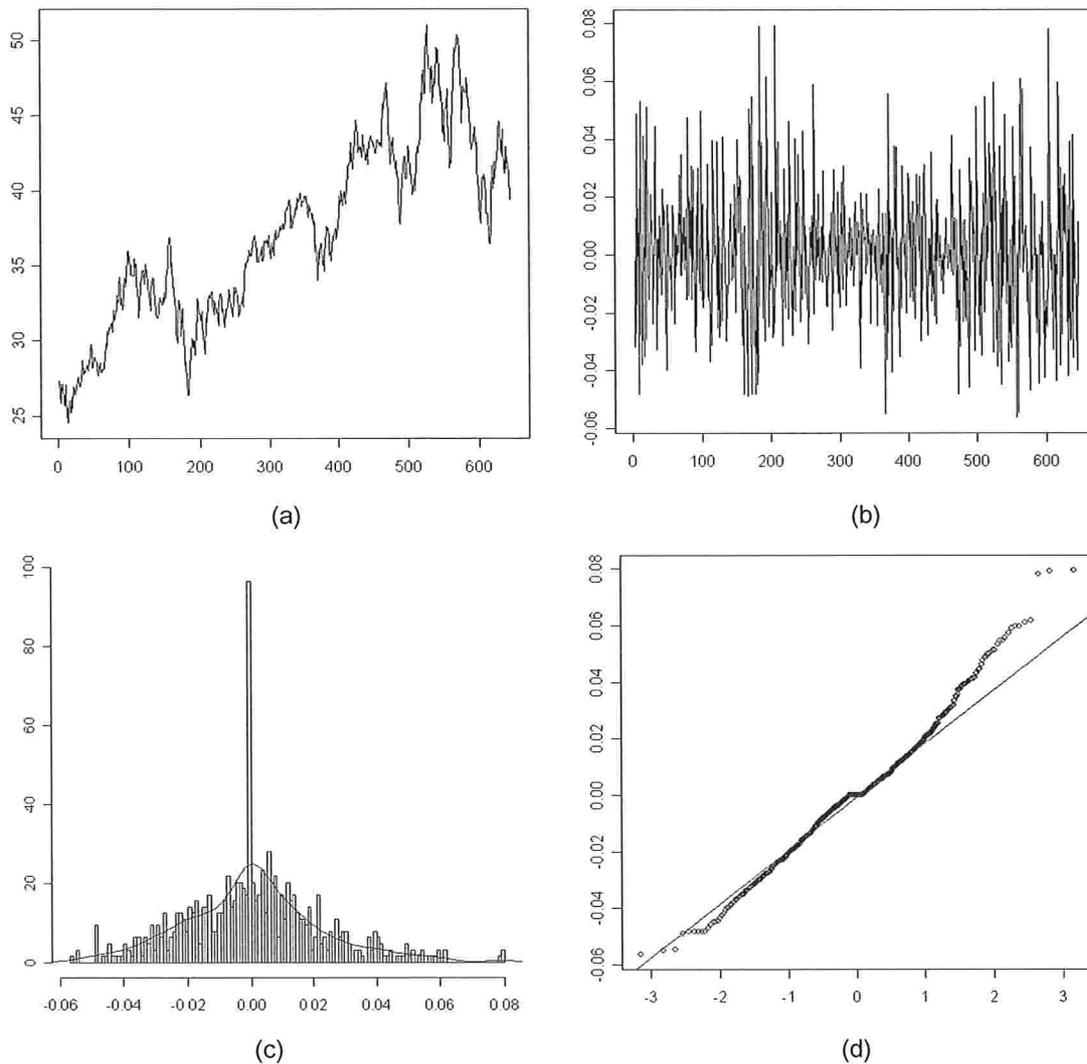


Figura 6.20: (a) Série do itau4, (b) Série de Retornos do itau4, (c) Histograma dos retornos do itau4 e (d) QQ-Plot dos Retornos do itau4.

Nota-se pela Figura 6.20 (b) que a série apresenta alguns grupos de volatilidade, indicando que esta não seja constante ao longo do tempo, o que sugere a utilização de um modelo GARCH, também pode-se observar em (d) que a distribuição da série parece destoar da distribuição Normal.

(a) Modelo GARCH

Assim como no caso da série do Ibovespa, inicialmente será ajustado um modelo GARCH para os dados e os resultados serão comparados com os resultados obtidos segundo o método proposto neste trabalho.

Na Figura 6.21, temos os gráficos das funções de autocorrelações (FAC) e autocorrelações parciais (FACP) da série de retornos do Itau4 e da série de retornos ao quadrado.

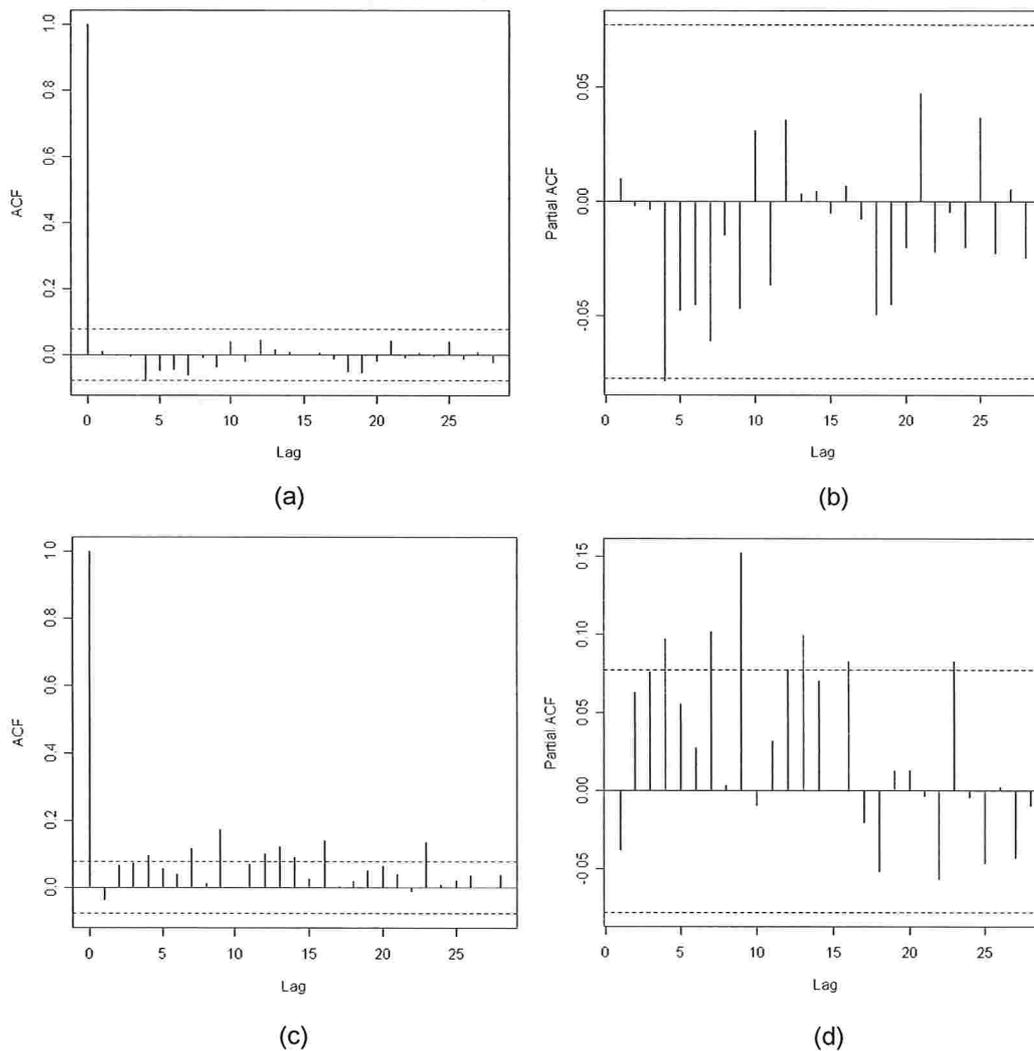


Figura 6.21: (a) FAC dos retornos, (b) FACP dos retornos, (c) FAC dos retornos ao quadrado, (d) FACP dos retornos ao quadrado.

Conforme observado na Figura 6.21, pode-se notar que não há evidências de que haja autocorrelação entre os retornos porém entre os retornos ao quadrado

os gráficos indicam fortes indícios da existência de autocorrelação, o que sugere o ajuste de um modelo GARCH.

Foi ajustado um modelo GARCH(1,1) assumindo inicialmente distribuição Normal para os erros.

Conforme observado no Apêndice A.1, todos os parâmetros são significantes ao nível de 5%, além disso, pode ser observado (Apêndice A.2) através do teste de Ljung Box tanto para os resíduos padronizados ε_t / σ_t , quanto para os resíduos padronizados ao quadrado $(\varepsilon_t / \sigma_t)^2$ que ao nível de 5% não rejeitamos a hipótese H_0 de que eles são independentes, ou seja, o modelo capturou com sucesso as estruturas de correlação serial na variância condicional.

Os testes de normalidade Jarque-Bera (p-valor <0,001) e Shapiro-Wilks (p-valor <0,001) para os resíduos padronizados apresentado no Apêndice A.2 indicam não normalidade.

A Figura 6.22 apresenta o gráfico QQ-plot dos resíduos padronizados e o gráfico das autocorrelações do quadrado dos resíduos padronizados. Observamos no gráfico QQ-plot que a suposição de normalidade não é apropriada, confirmando os resultados dos testes apresentados anteriormente, assim será ajustado a seguir um novo modelo GARCH(1,1) assumindo distribuição t de Student para os erros.

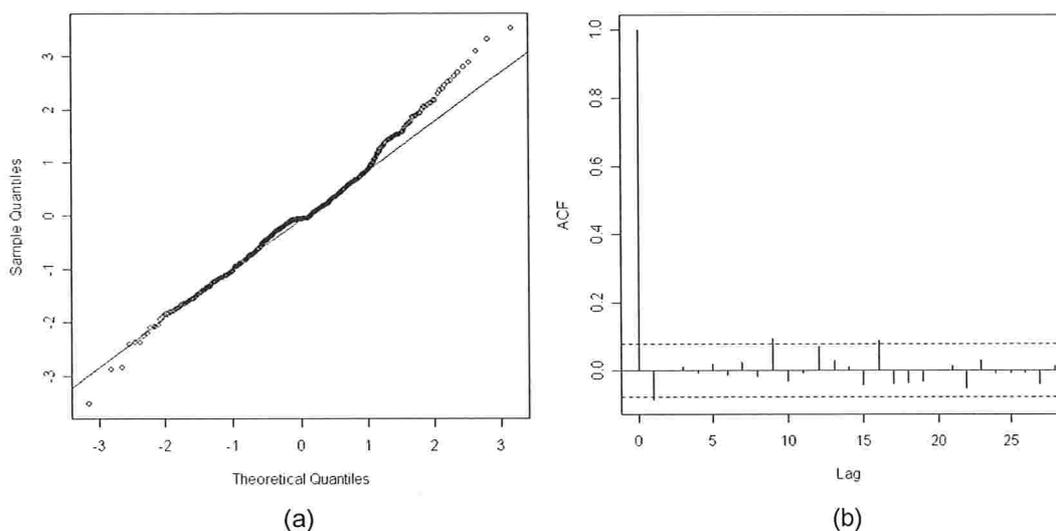


Figura 6.22: (a) QQ-Plot dos resíduos padronizados, (b) Autocorrelação do quadrado dos resíduos padronizados referentes ao modelo com distribuição Normal.

Agora vamos ajustar um modelo GARCH(1,1) supondo distribuição t de Student para os erros. Assim como no caso da série do Ibovespa, foram utilizados os graus de liberdade default do R-Plus. Os resultados podem ser visto no Apêndice A.3.

Observa-se através do Apêndice A.4, que o modelo ajustado considerando distribuição t de Student apresentou valores de BIC e AIC levemente menores que o modelo com a distribuição normal, indicando um ajuste melhor.

Através da Figura 6.23 (a), pode-se notar que a distribuição t-Student, assim como no caso da distribuição normal, parece não ajustar perfeitamente os dados, o que pode ser confirmado através dos testes de Jarque-Bera (p-valor<0,001) e Shapiro-Wilks (p-valor<0,001).

Assim como no caso do modelo com distribuição normal, utilizando a distribuição t de Student também foi possível capturar com sucesso as estruturas de correlação serial na variância condicional conforme pode ser observado na Figura 6.23 (b) e através dos testes de Ljung-Box.

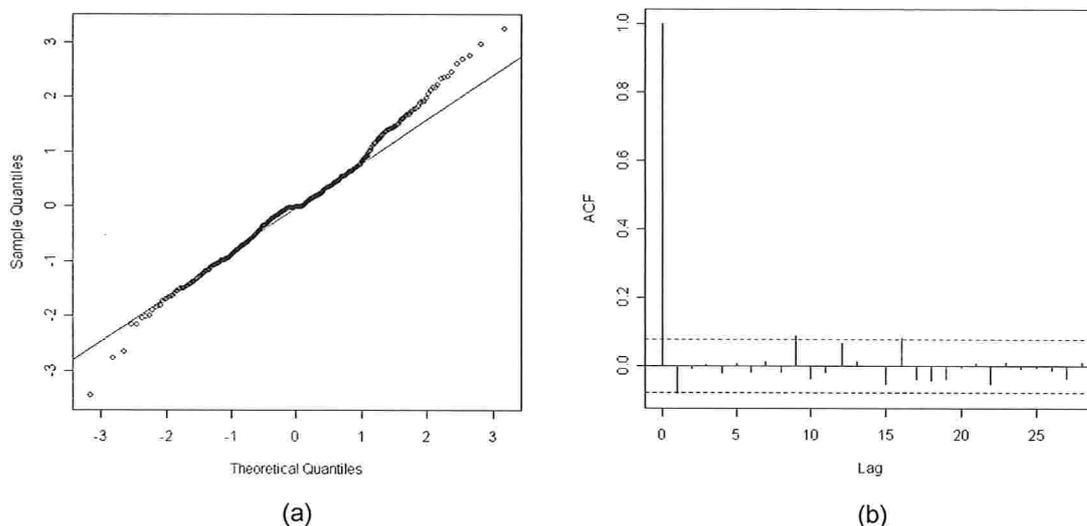


Figura 6.23: (a) QQ-Plot dos resíduos padronizados, (b) Autocorrelação do quadrado dos resíduos padronizados referentes ao modelo com distribuição t de Student.

Dado os resultados obtidos, foi selecionado o modelo assumindo distribuição t-Student, conforme segue:

$$X_t = \hat{\sigma}_t \varepsilon_t,$$

$$\hat{\sigma}_t^2 = 0,0815 * X_{t-1}^2 + 0,9252 * \hat{\sigma}_{t-1}^2. \quad (6.8)$$

(b) Modelo com coeficientes variando no tempo

Após o ajuste dos modelos GARCH apresentados anteriormente foi ajustado um modelo com coeficientes variando assim como no caso da série do Ibovespa.

A série de retorno do Itaú refere-se à 644 observações e neste caso também foi utilizada uma janela móvel de tamanho 300, dessa forma, foi possível rodar 345 modelos distintos. O modelo considerado foi o GARCH(1,1) com distribuição normal, o procedimento no ajuste dos modelos é exatamente o mesmo utilizado no caso da série do Ibovespa e por isso os detalhes serão desconsiderados. Ao término desse processo foi obtido um vetor de parâmetros com 4 colunas e 345 linhas e cada coluna foi tratada com uma série temporal independente. Para cada uma dessas séries foi ajustado um modelo conforme segue:

Ajuste do modelo para a série de parâmetros $\alpha(t)$

Na Figura 6.24, apresentamos (a) a série original de $\alpha(t)$, (b) a série da 1ª diferença de $\alpha(t)$, (c) o histograma da 1ª diferença de $\alpha(t)$ e (d) o gráfico QQ-plot da 1ª diferença de $\alpha(t)$.

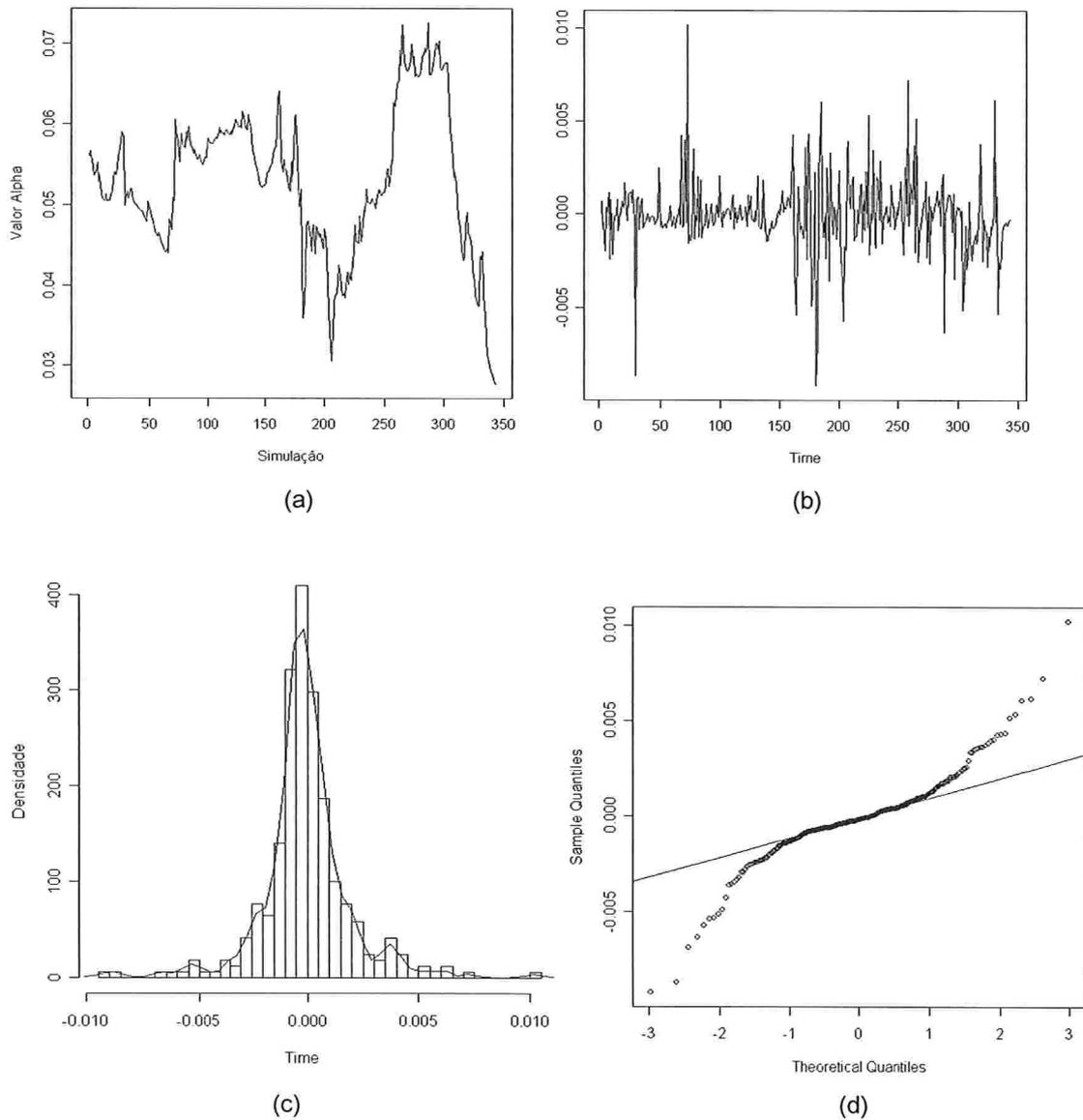


Figura 6.24: (a) Série original de $\alpha(t)$, (b) Série da 1ª diferença de $\alpha(t)$, (c) Histograma da 1ª diferença de $\alpha(t)$ e (d) QQ-Plot da 1ª diferença de $\alpha(t)$.

Na Figura 6.24 (b) pode-se ver que tomando a 1ª diferença da série de $\alpha(t)$, a mesma se torna estacionária, porém a série apresenta alguns valores muito altos ou baixos, referentes aos saltos que podem ser visto na série original em 6.24 (a). Estes saltos podem ser causados pelo fato da série considerada ser pequena ou devido à escala apresentar um range pequeno de variação. Estes aspectos são

importante pois podem prejudicar o ajuste do modelo. Em (d) pode-se observar que a distribuição da série parece destoar da distribuição normal.

Na Figura 6.25, são apresentados os gráficos das funções de autocorrelações (FAC) e autocorrelações parciais (FACP) da série da 1ª diferença dos $\alpha(t)$ e da 1ª diferença ao quadrado.

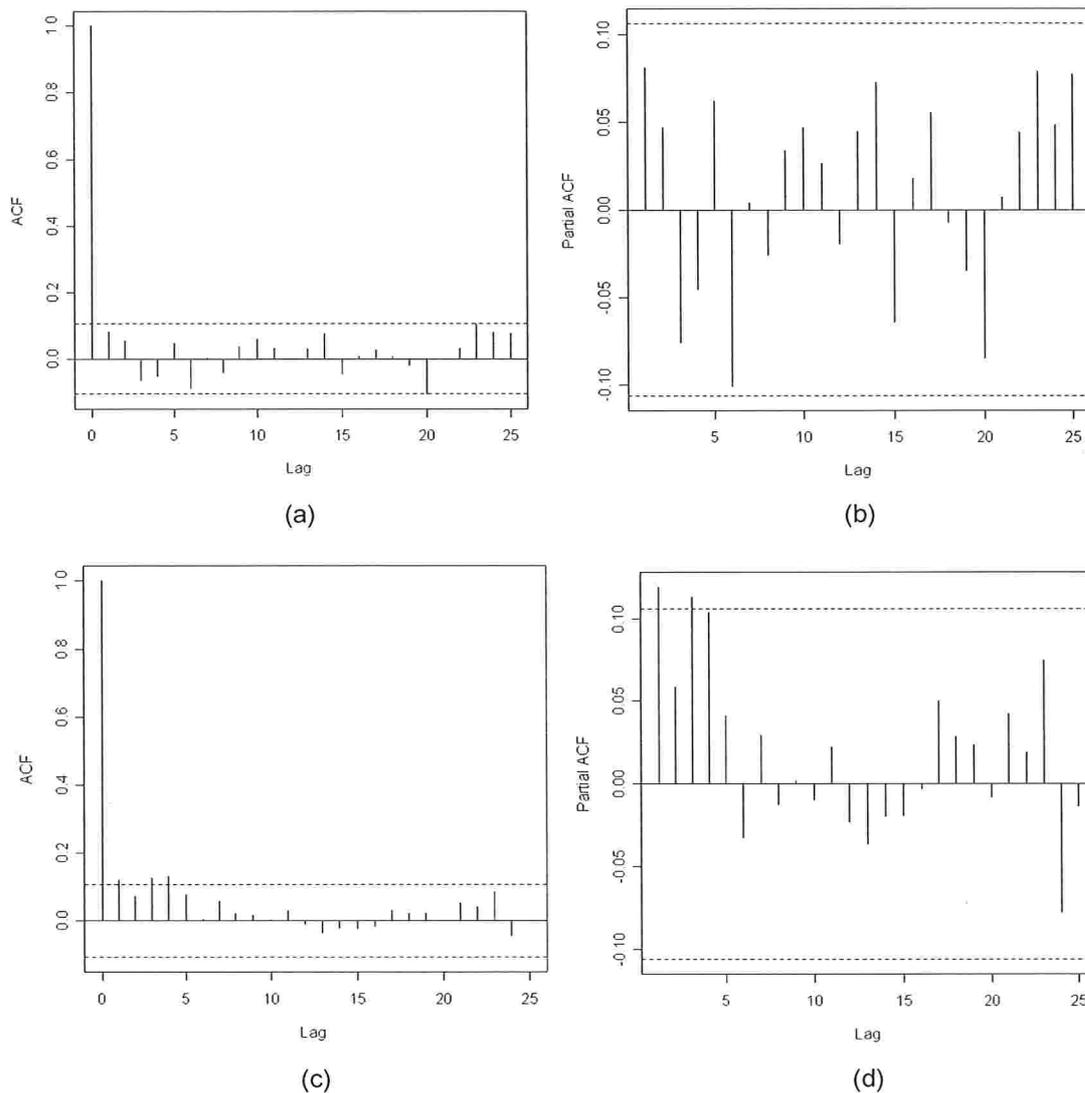


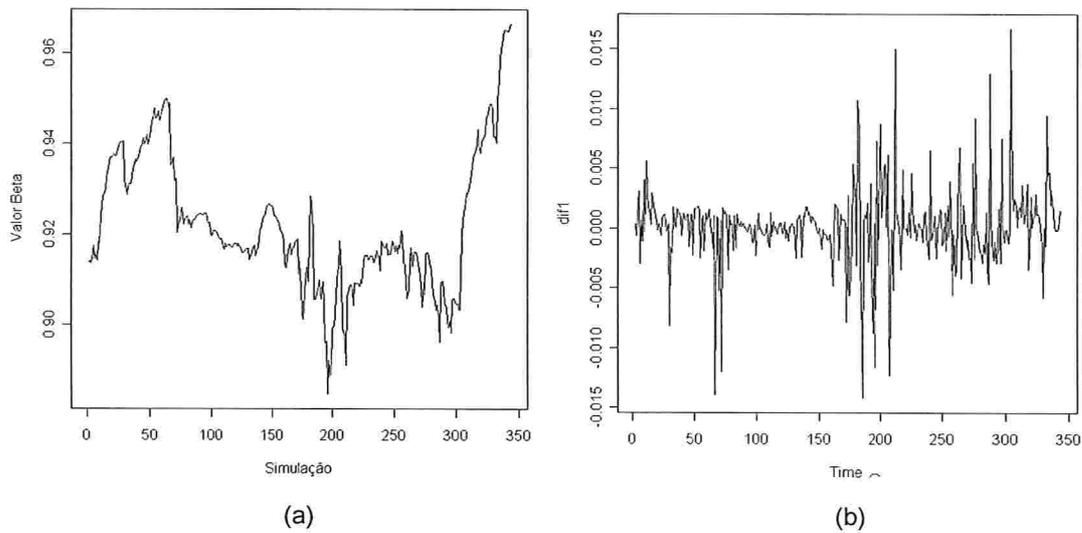
Figura 6.25: (a) FAC da 1ª diferença de $\alpha(t)$, (b) FACP da 1ª diferença de $\alpha(t)$, (c) FAC da 1ª diferença de $\alpha(t)$ ao quadrado, (d) FACP da 1ª diferença de $\alpha(t)$ ao quadrado.

Na Figura 6.25, pode-se notar que não há evidências de que exista autocorrelação na série referente à 1ª diferença de $\alpha(t)$, dessa forma, será considerado o seguinte modelo para a série $\alpha(t)$:

$$\alpha(t) = \alpha(t-1) + \varepsilon_t \quad (6.9)$$

Ajuste do modelo para a série de parâmetros $\beta(t)$

Na Figura 6.26, apresentamos (a) a série original de $\beta(t)$, (b) a série da 1ª diferença de $\beta(t)$, (c) o histograma da 1ª diferença de $\beta(t)$ e (d) o gráfico QQ-plot da 1ª diferença de $\beta(t)$.



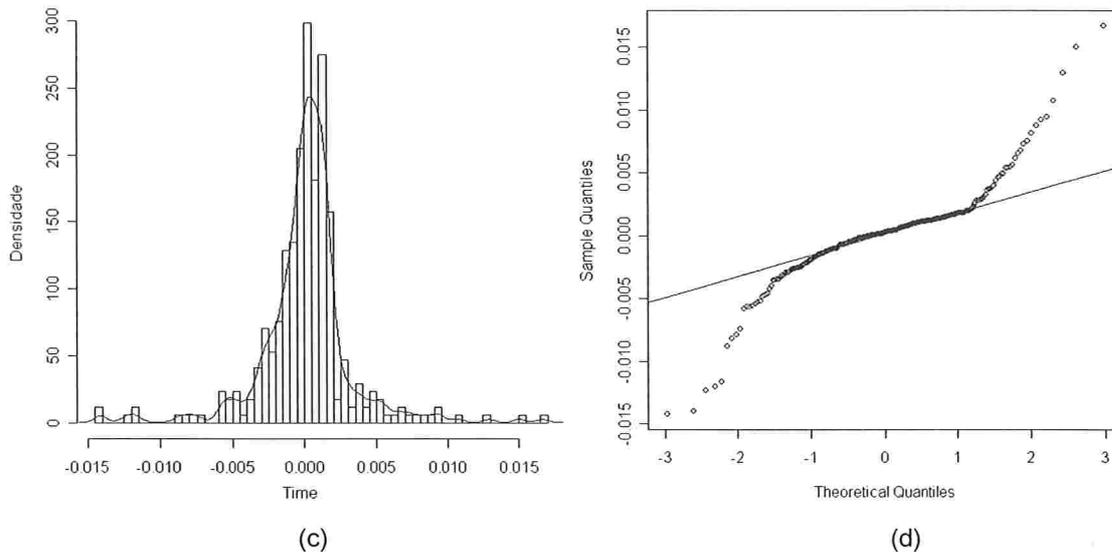


Figura 6.26: (a) Série original de $\beta(t)$, (b) Série da 1ª diferença de $\beta(t)$, (c) Histograma da 1ª diferença de $\beta(t)$ e (d) QQ-Plot da 1ª diferença de $\beta(t)$.

Na Figura 6.26 pode-se notar a presença de grupos de volatilidade e também que a série se tornou estacionário na média ao considerarmos a 1ª diferença da série de $\beta(t)$, além disso, em (d) pode-se observar que a distribuição da série parece destoar da distribuição normal por apresentar caudas mais pesadas que o esperado.

Na Figura 6.27, são apresentados os gráficos das funções de autocorrelações (FAC) e autocorrelações parciais (FACP) da série da 1ª diferença de $\beta(t)$ e da 1ª diferença ao quadrado.

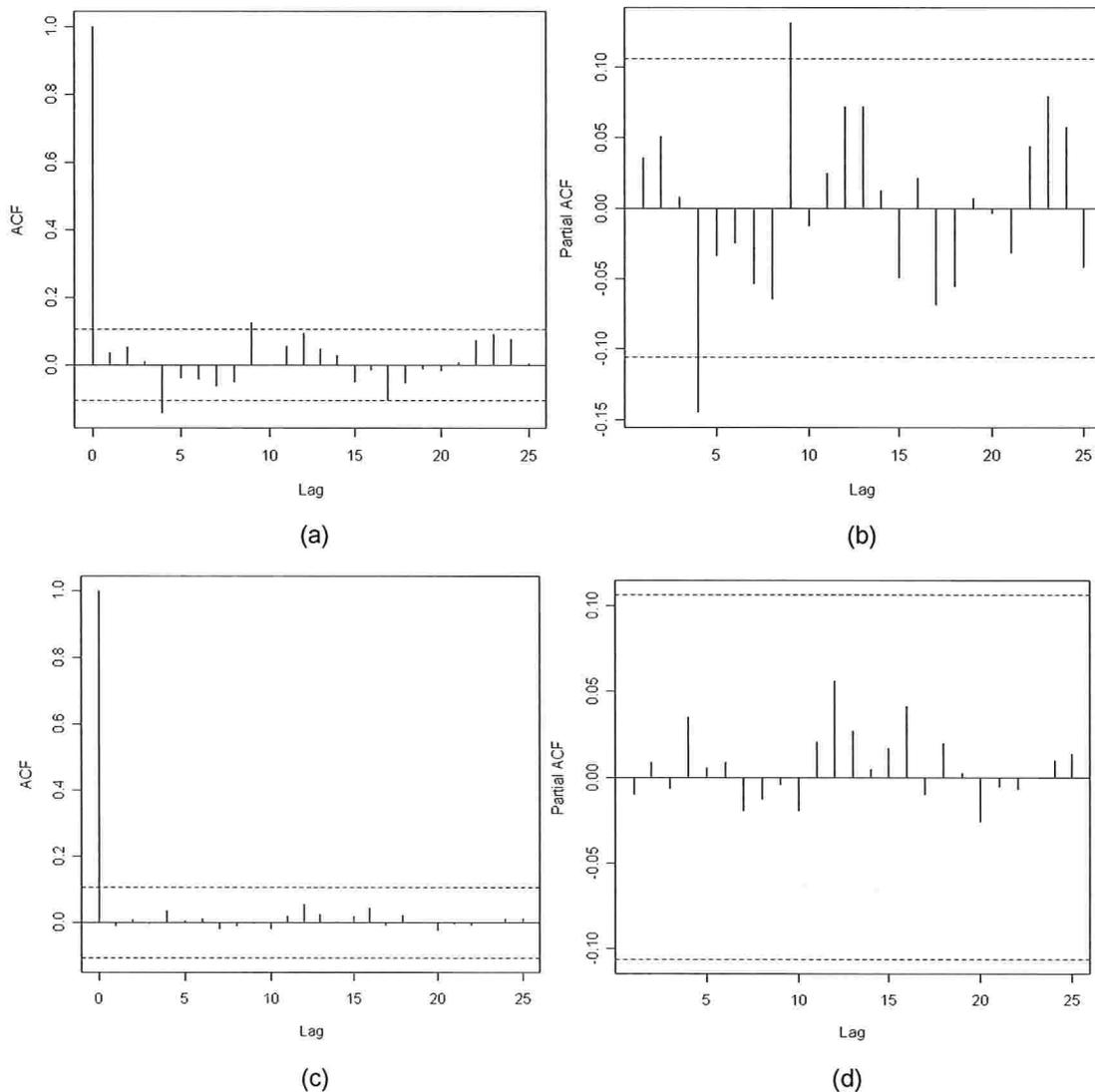


Figura 6.27: (a) FAC da 1ª diferença de $\beta(t)$, (b) FACP da 1ª diferença de $\beta(t)$, (c) FAC da 1ª diferença de $\beta(t)$ ao quadrado, (d) FACP da 1ª diferença de $\beta(t)$ ao quadrado.

Na Figura 6.27 (a) e (b), há evidências de que exista autocorrelação de lag 4 e lag 9 na série de $\beta(t)$. Observando a série referente à 1ª diferença dos $\beta(t)$ ao quadrado, Figura 6.10 (c) e (d), não há evidências da existência de autocorrelação, dessa forma, foi ajustado um modelo AR(9) sendo que todos os parâmetros diferentes de AR=4 e AR=9 foram desconsiderados.

Conforme observado Apêndice A.5, todos os parâmetros estimados são significantes ao nível de 5%, além disso, pode ser observado (Apêndice A.6)

através do teste de Box-Pierce (p-valor=0,3481) para os resíduos padronizados ε_t / σ_t , que ao nível de 5% não rejeitamos a hipótese H_0 de que eles são independentes, ou seja, o modelo capturou com sucesso as estruturas de correlação serial o que também pode ser observado pela Figura 6.28.

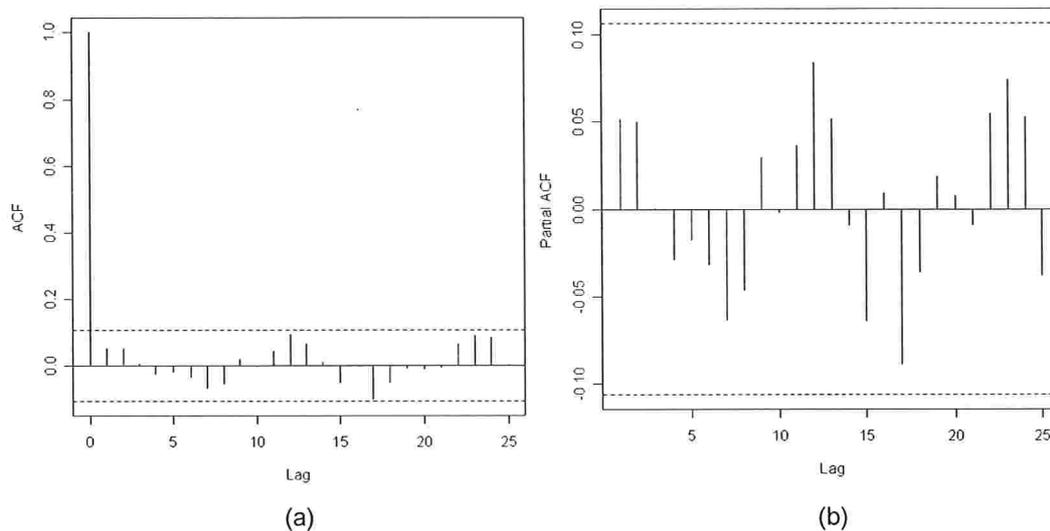


Figura 6.28: (a) FAC dos resíduos padronizados, (b) FACP dos resíduos padronizados

Os testes de normalidade Jarque-Bera (p-valor=0) e Shapiro-Wilks (p-valor=0) para os resíduos padronizados apresentado no Apêndice A.7 indicam não normalidade.

A Figura 6.29 apresenta o gráfico QQ-plot dos resíduos padronizados e o gráfico de autocorrelações dos resíduos padronizados. Observamos no gráfico QQ-plot que a suposição de normalidade não é apropriada, confirmando os resultados dos testes apresentados anteriormente.

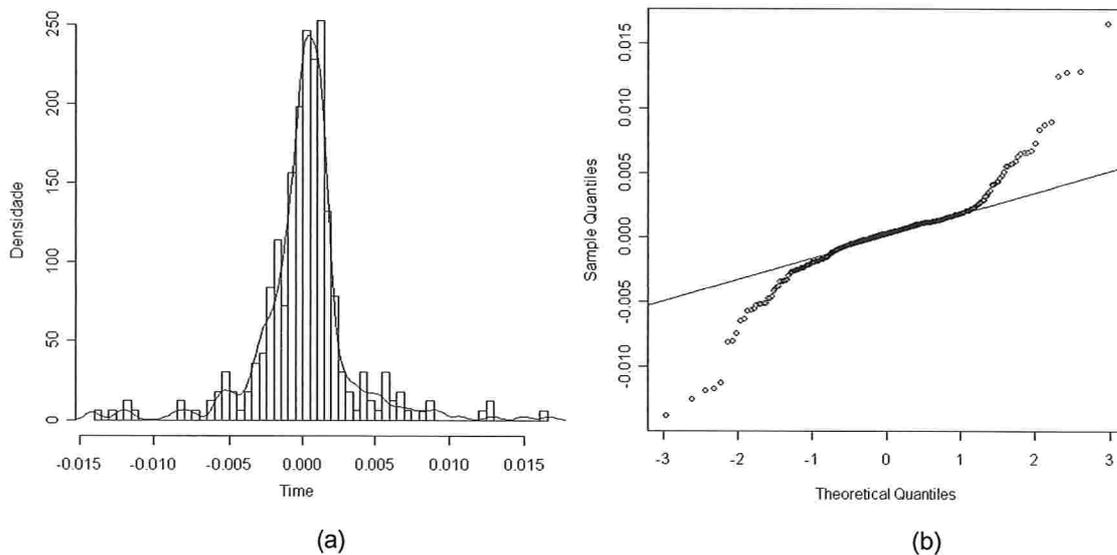


Figura 6.29: (a) Histograma dos resíduos padronizados, (b) QQ-Plot dos resíduos padronizados

Dados os resultados obtidos no ajuste, o modelo final é:

$$(\hat{\beta}(t) - \hat{\beta}(t-1)) = -0,13703 * (\hat{\beta}(t-4) - \hat{\beta}(t-5)) + 0,12445 * (\hat{\beta}(t-9) - \hat{\beta}(t-10)) \quad (6.10)$$

Ajuste do modelo para a série de parâmetros $\mu(t)$

Na figura 6.30, apresentamos (a) a série original de $\mu(t)$, (b) a série da 1ª diferença de $\mu(t)$, (c) o histograma da 1ª diferença de $\mu(t)$ e (d) e o gráfico QQ-plot da 1ª diferença de $\mu(t)$.

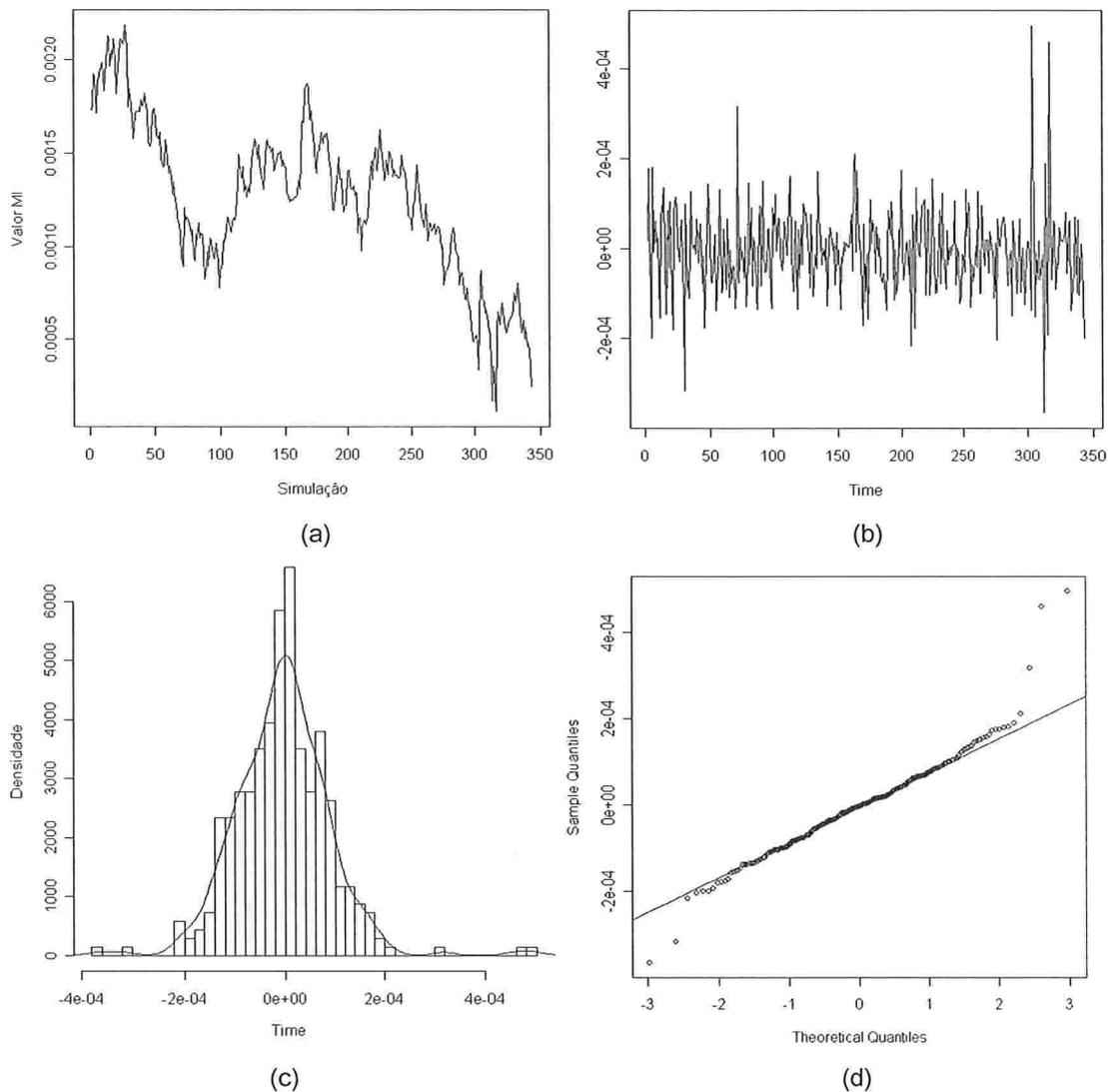


Figura 6.30: (a) Série original de $\mu(t)$, (b) Série da 1ª diferença de $\mu(t)$, (c) Histograma da 1ª diferença de $\mu(t)$ e (d) QQ-Plot da 1ª diferença de $\mu(t)$.

Na Figura 6.30 (b) pode-se notar a presença de grupos de volatilidade e também que a série se tornou estacionária na média ao considerarmos a 1ª diferença da série de $\mu(t)$, além disso, em (d) pode-se observar que a distribuição da série parece aproximar-se da distribuição Normal com exceção de alguns pontos na cauda.

Na Figura 6.31, são apresentados os gráficos das funções de autocorrelações (FAC) e autocorrelações parciais (FACP) da série da 1ª diferença dos $\mu(t)$ e da 1ª diferença ao quadrado.

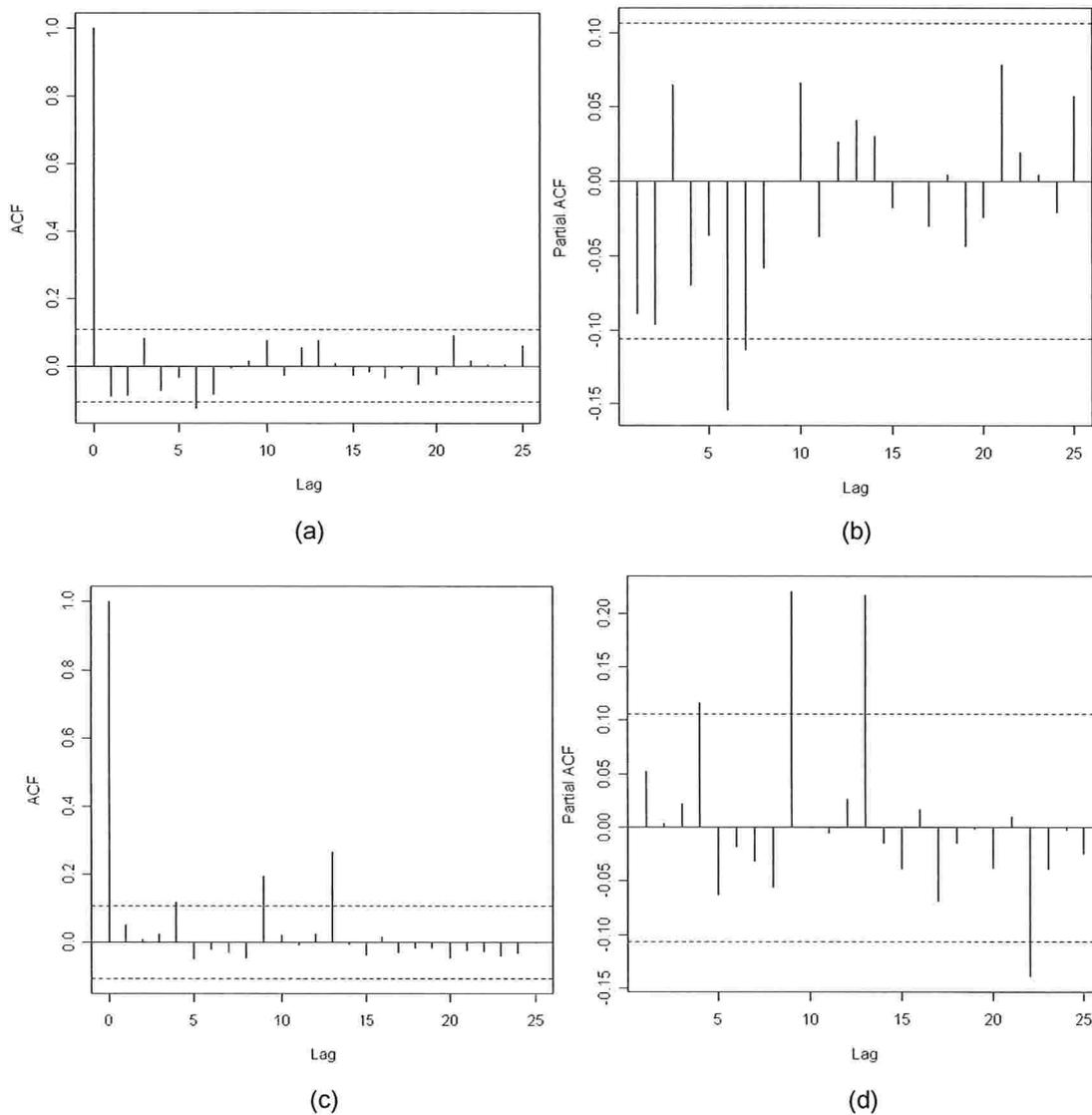


Figura 6.31: (a) FAC da 1ª diferença de $\mu(t)$, (b) FACP da 1ª diferença de $\mu(t)$, (c) FAC da 1ª diferença de $\mu(t)$ ao quadrado, (d) FACP da 1ª diferença de $\mu(t)$ ao quadrado.

Na Figura 6.31 (a) e (b), há evidências de que exista autocorrelação de lag 6 na série de $\mu(t)$. Observando os resultados referentes à 1ª diferença de $\mu(t)$ ao

quadrado, Figura 6.31 (c) e (d), também há evidências da existência de autocorrelação, porém como o objetivo neste momento é a modelagem da média e não da variabilidade a volatilidade será desconsiderada, dessa forma, foi ajustado um modelo AR(6) sendo que todos os parâmetros diferentes de AR=6 foram desconsiderados.

Conforme observado no Apêndice A.8, o parâmetro estimado é significativo ao nível de 5%, além disso, pode ser observado (Apêndice A.9) através do teste de Box-Ljung (p-valor=0,060) para os resíduos padronizados ε_t / σ_t , que ao nível de 5% não rejeitamos a hipótese H_0 de que eles são independentes, o que também pode ser observado pela Figura 6.32.

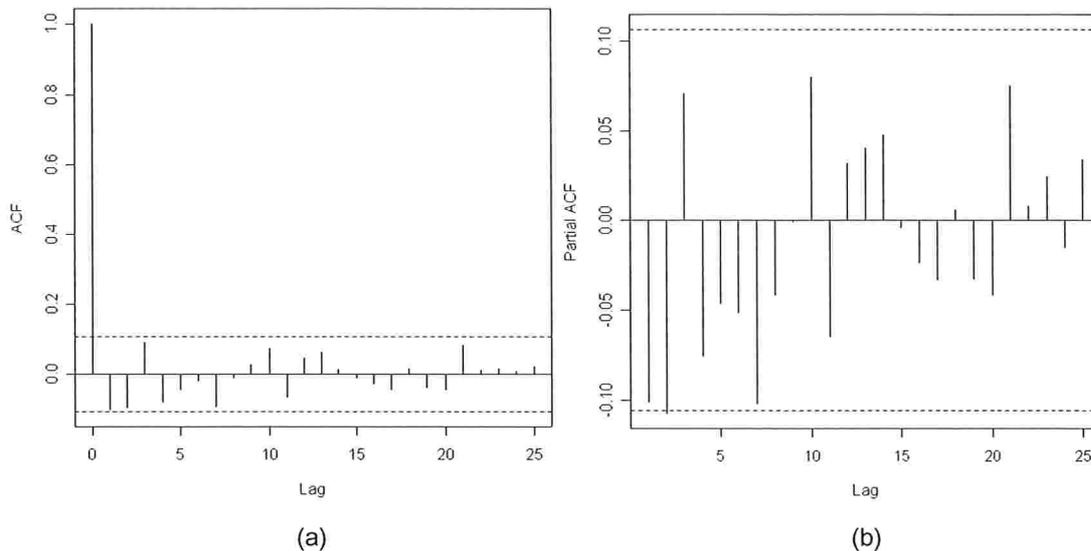


Figura 6.32: (a) FAC dos resíduos padronizados, (b) FACP dos resíduos padronizados

Os testes de normalidade Jarque-Bera (p-valor=0) e Shapiro-Wilks (p-valor=0) para os resíduos padronizados apresentados no Apêndice A.10 indicam não normalidade.

A Figura 6.33 apresenta o gráfico QQ-plot dos resíduos padronizados. Observamos no gráfico QQ-plot que de maneira geral a distribuição parece aproximar-se bem de uma normal porém alguns pontos na cauda destoam muito o que deve estar afetando os testes e fazendo com que a hipótese de normalidade

seja rejeitada conforme visto nos resultados dos testes apresentados anteriormente.

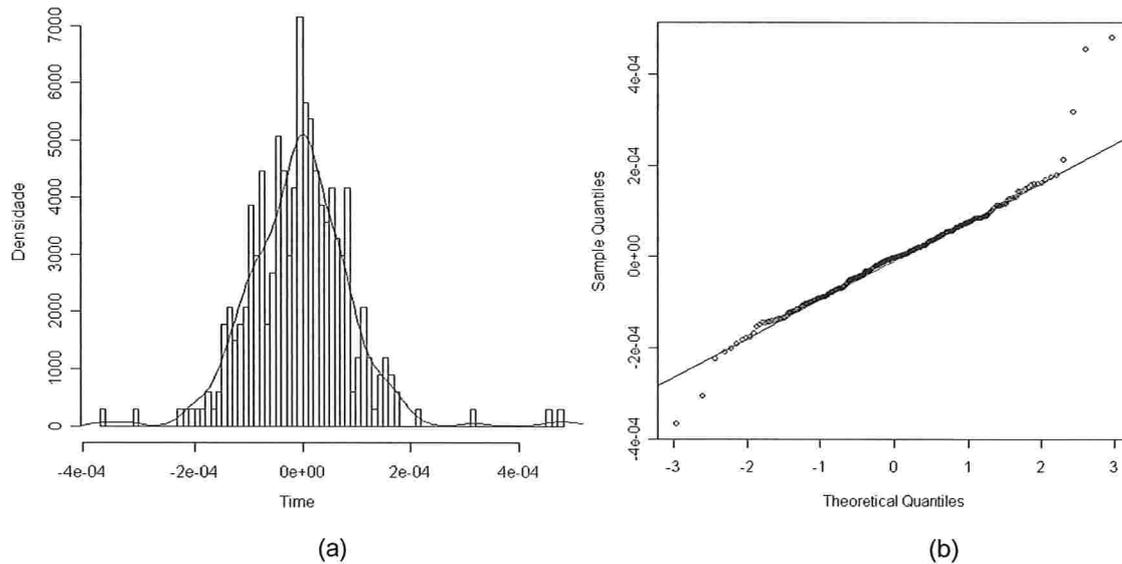


Figura 6.33: (a) Histograma dos resíduos padronizados, (b) QQ-Plot dos resíduos padronizados

Dados os resultados obtidos no ajuste, o modelo final é:

$$(\hat{\mu}(t) - \hat{\mu}(t-1)) = -0,12597 * (\hat{\mu}(t-6) - \hat{\mu}(t-7)) \quad (6.11)$$

Ajuste do modelo para a série de parâmetros $\alpha_0(t)$

Na Figura 6.34, apresentamos (a) a série original de $\alpha_0(t)$, (b) a série da 1ª diferença de $\alpha_0(t)$, (c) o histograma da 1ª diferença de $\alpha_0(t)$ e (d) e o gráfico QQ-plot da 1ª diferença de $\alpha_0(t)$.

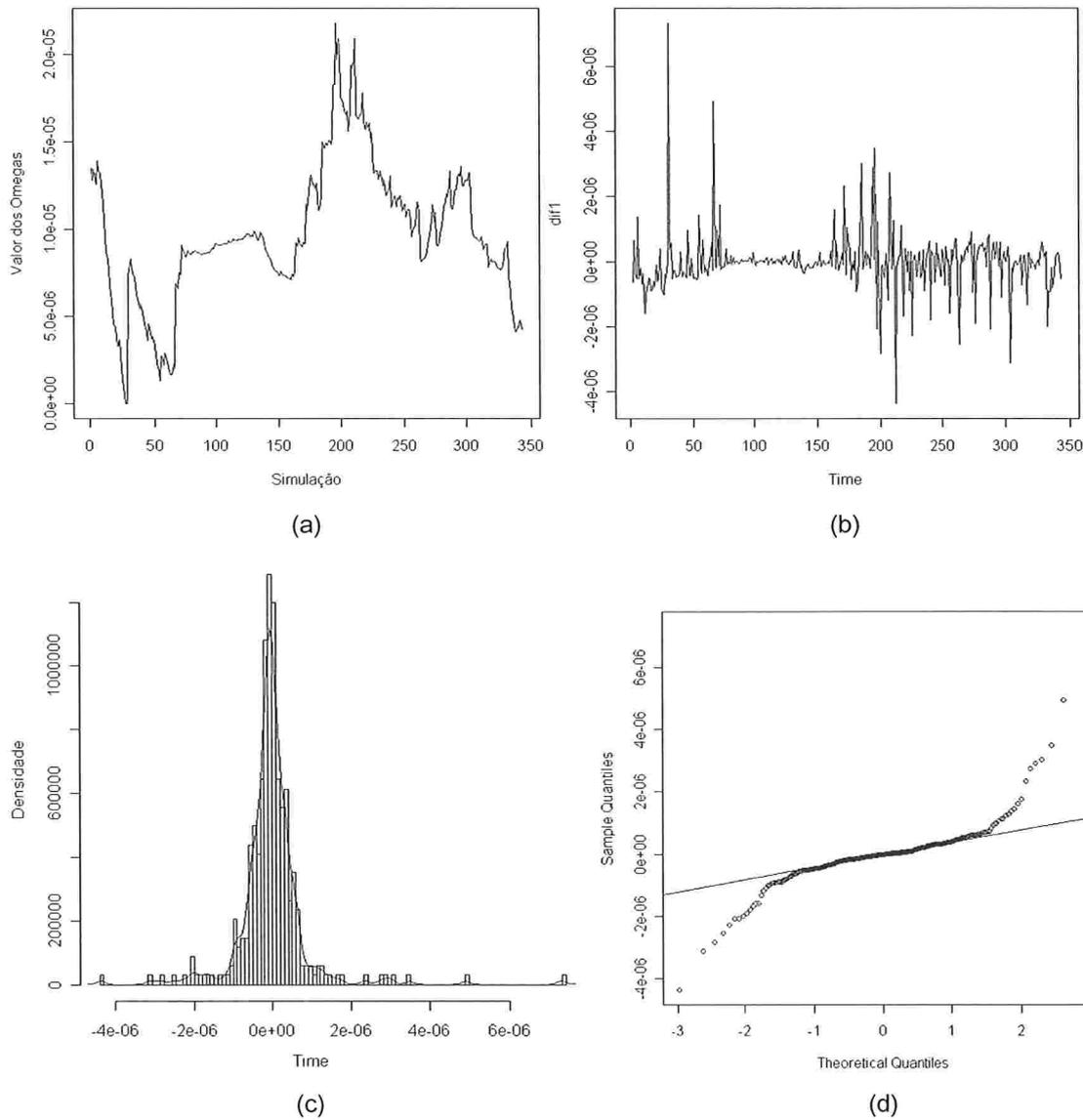


Figura 6.34: (a) Série original de $\alpha_0(t)$, (b) Série da 1ª diferença de $\alpha_0(t)$, (c) Histograma da 1ª diferença de $\alpha_0(t)$ e (d) QQ-Plot da 1ª diferença de $\alpha_0(t)$.

Na Figura 6.34 (b) pode-se ver que tomando a 1ª diferença da série de $\alpha_0(t)$, a mesma se torna estacionária e também é forte a presença de grupos de volatilidade, o que pode prejudicar o ajuste do modelo, além disso, em (d) pode-se observar que a distribuição da série parece destoar de uma normal.

Na Figura 6.35, são apresentados os gráficos das funções de autocorrelações (FAC) e autocorrelações parciais (FACP) da série da 1ª diferença de $\alpha_0(t)$ e da 1ª diferença ao quadrado.

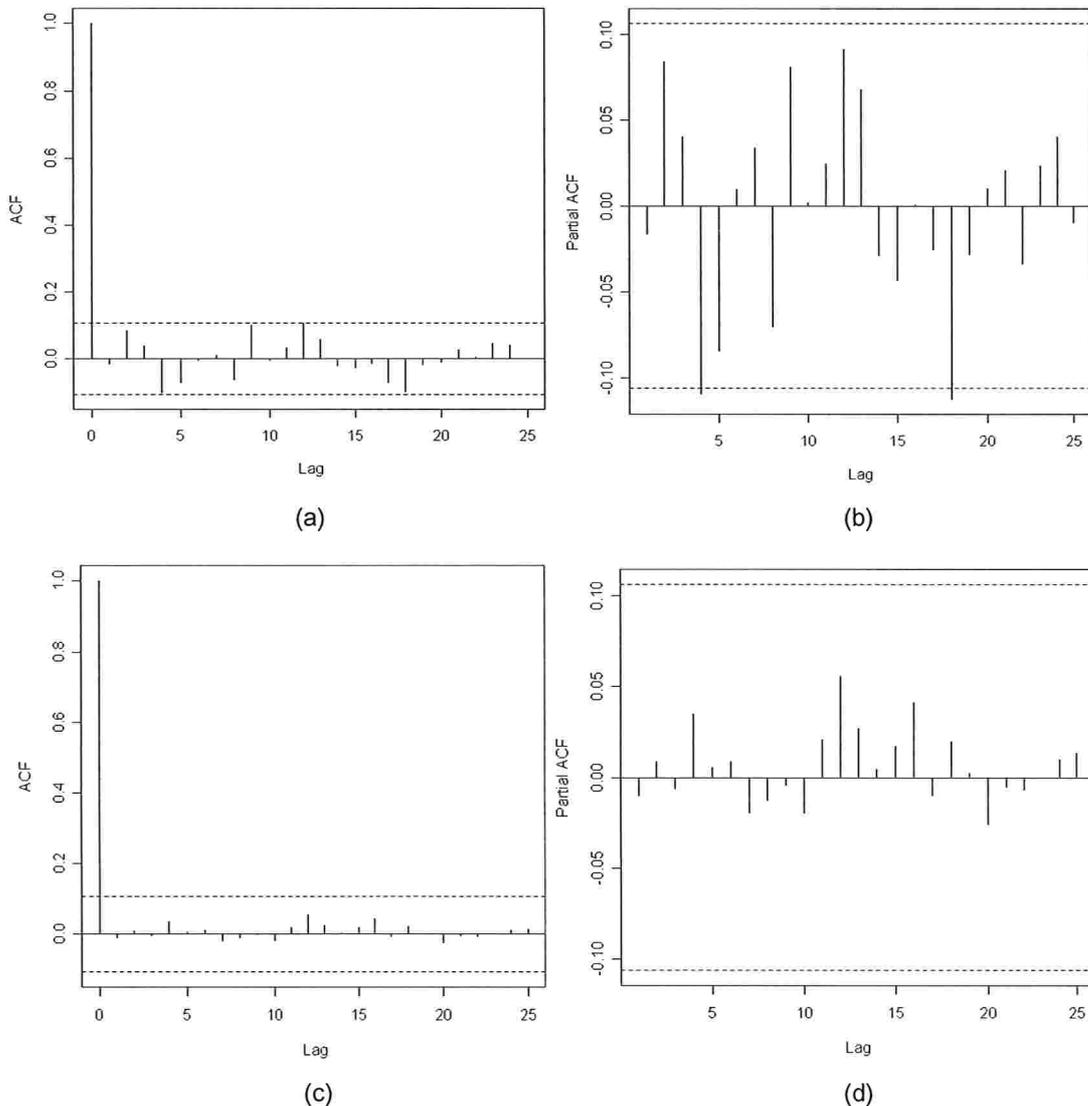


Figura 6.35: (a) FAC da 1ª diferença de $\alpha_0(t)$, (b) FACP da 1ª diferença de $\alpha_0(t)$, (c) FAC da 1ª diferença de $\alpha_0(t)$ ao quadrado, (d) FACP da 1ª diferença de $\alpha_0(t)$ ao quadrado.

Na Figura 6.35, pode-se notar que não há evidências de que exista autocorrelação na série referente à 1ª diferença de $\alpha_0(t)$, de qualquer forma, será

ajustado um modelo AR(18) onde todos os parâmetros diferentes de AR=4 e AR=18 serão desconsiderados de modo a confirmar esta hipótese.

Conforme observado no Apêndice A.11, adotando um nível de significância de 5% não rejeitamos a hipótese H_0 de que os coeficientes do modelo ajustado sejam 0, conforme já visto na Figura 6.35

Neste caso, como os dados não apresentaram correlação serial será considerado o seguinte modelo:

$$\alpha_0(t) = \alpha_0(t-1) + \varepsilon_t \quad (6.12)$$

Previsões

Após o desenvolvimento dos modelos citados anteriormente, eles foram utilizados para previsão. Os valores previstos foram comparados com os valores reais da série referentes ao período 21 de março de 2008 à 04 de abril de 2008. Assim como no caso da série do Ibovespa foi calculado o EQM para cada metodologia, abaixo seguem os resultados obtidos:

Cálculo do EQM – Modelo GARCH(1,1) com distribuição t de Student

Date	Preço Fechamento	Log(Pt / Pt-1)	Predito - Garch(1,1) Dist t Student	Erro Quadrático Médio
			Previsão σ_t^2	Garch(1,1)
21/3/2008	39,85	-		
24/3/2008	39,10	-0,0083	-	
25/3/2008	41,00	0,0206	0,0006418	4,71492E-08
26/3/2008	40,58	-0,0045	0,0006390	3,8321E-07
27/3/2008	39,50	-0,0117	0,0006363	2,49093E-07
28/3/2008	39,15	-0,0039	0,0006337	3,82834E-07
31/3/2008	40,01	0,0094	0,0006311	2,9378E-07
1/4/2008	41,25	0,0133	0,0006285	2,05028E-07
2/4/2008	41,32	0,0007	0,0006260	3,91185E-07
3/4/2008	41,50	0,0019	0,0006235	3,84345E-07
4/4/2008	41,25	-0,0026	0,0006211	3,7725E-07
7/4/2008	41,75	0,0052	0,0006187	3,49672E-07
				<u>3,06355E-06</u>

Tabela 6.3: Cálculo do EQM para o modelo GARCH(1,1)

A variância condicional σ_t^2 e o EQM (Erro quadrático médio) foram obtidos da mesma forma que no caso da série do Ibovespa. Conforme observado na tabela acima, o valor do EQM calculado foi $3,06355 e^{-6}$.

Cálculo do EQM – Modelo com coeficientes variando no tempo

Date	Preço Fecham.	Log (P _t / P _{t-1})	Metodologia Proposta				Previsão σ_t^2	Erro Quadrático Médio
			$\alpha_1(t)$	$\beta(t)$	$\mu(t)$	$\alpha_0(t)$		Metodologia Proposta
9/5/2008	39,85	-						
12/5/2008	39,10	-0,0083	-	-	-	-	0,0006749	
13/5/2008	41,00	0,0206	0,0277	0,9668	0,0003	0,0000042	0,0006587	5,47887E-08
14/5/2008	40,58	-0,0045	0,0277	0,9671	0,0002	0,0000042	0,0006528	4,00388E-07
15/5/2008	39,50	-0,0117	0,0277	0,9673	0,0003	0,0000042	0,0006363	2,49063E-07
16/5/2008	39,15	-0,0039	0,0277	0,9672	0,0003	0,0000042	0,0006236	3,7051E-07
19/5/2008	40,01	0,0094	0,0277	0,9671	0,0003	0,0000042	0,0006078	2,69131E-07
20/5/2008	41,25	0,0133	0,0277	0,9670	0,0003	0,0000042	0,0005944	1,7527E-07
21/5/2008	41,32	0,0007	0,0277	0,9670	0,0003	0,0000042	0,0005836	3,39987E-07
23/5/2008	41,50	0,0019	0,0277	0,9670	0,0003	0,0000042	0,0005686	3,19272E-07
26/5/2008	41,25	-0,0026	0,0277	0,9672	0,0003	0,0000042	0,0005542	2,99601E-07
27/5/2008	41,75	0,0052	0,0277	0,9672	0,0003	0,0000042	0,0005406	2,63349E-07
								2,74136E-06

Tabela 6.4: Cálculo do EQM para o modelo com coeficientes variando no tempo

As previsões 10 passos à frente para $\alpha_0(t)$, $\alpha_1(t)$, $\beta(t)$ e $\mu(t)$ mostradas na Tabela 6.4 foram obtidas com base nos modelos (6.9, 6.10, 6.11 e 6.12) ajustados para cada parâmetro apresentado anteriormente.

A previsão da variância condicional σ_t^2 foi obtida conforme fórmula (6.7) e o EQM foi calculado da mesma forma que no modelo GARCH(1,1).

Conforme observado na Tabela 6.4, o valor obtido foi $2,74136 e^{-6}$, indicando que a utilização da metodologia proposta ajudou a reduzir o EQM, porém a redução foi um pouco menor que no caso da série do Ibovespa.

Calculando a razão entre os EQMs, obtemos:

$$\text{Razão} = \frac{2,74136 e^{-6}}{3,06355 e^{-6}} = 0.895\%$$

indicando que foi possível reduzir o EQM em 10,5 pontos percentuais.

6.4 Preços da Ação da Vale (Vale5)

Os dados utilizados nesta aplicação referem-se aos valores diários de fechamento da ação da vale (VALE5), de 19 de agosto de 2004 à 23 de janeiro de 2009, totalizando 1147 observações. Em 19 de agosto de 2004 houve uma alteração nos valores da ação da Vale5, conforme pode ser observado na Figura 6.36 o que fez com que o preço da ação caísse bruscamente inviabilizando a utilização dos dados referentes ao período anterior à esta data. Foi utilizado o período de 26 de janeiro de 2009 à 06 de fevereiro de 2009 na validação do modelo.

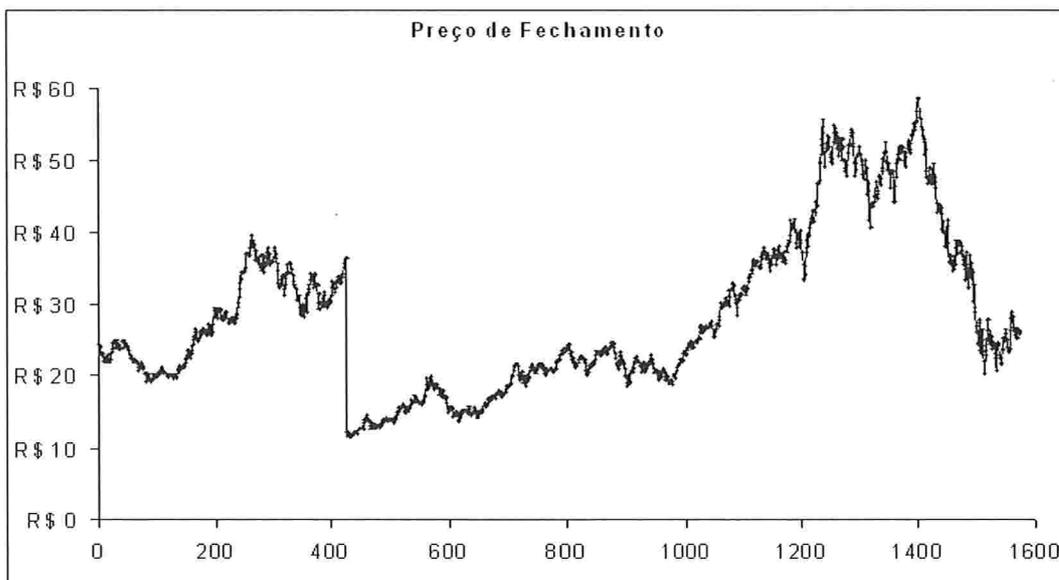


Figura 6.36: Preço de fechamento da Ação da Vale

Na Figura 6.37, apresentamos (a) a série original da vale5, (b) a série de retornos da vale5, (c) o histograma dos retornos da vale5 e (d) e o gráfico QQ-plot dos retornos da vale5.

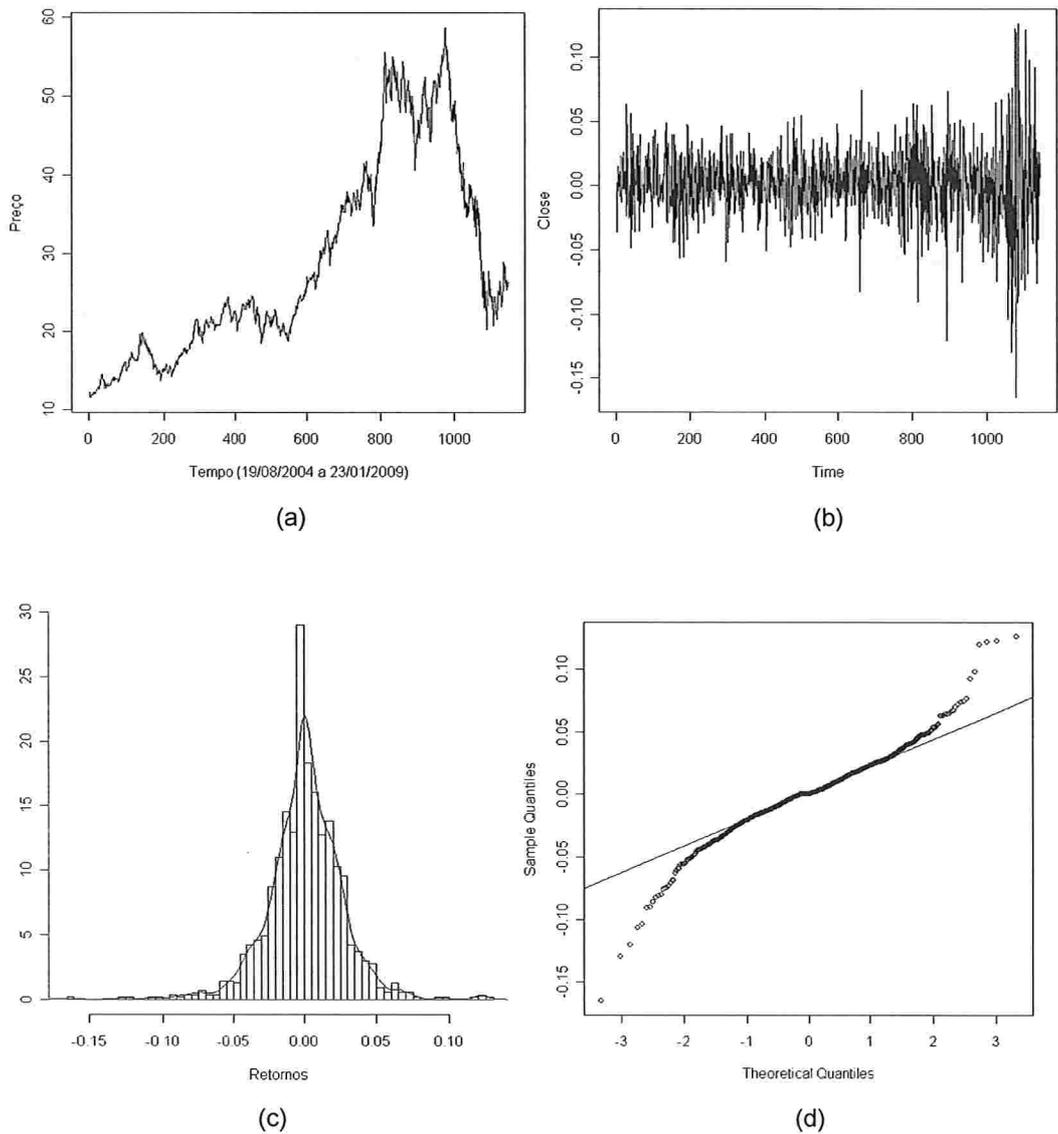


Figura 6.37: (a) Série da vale5, (b) Série de Retornos da vale5, (c) Histograma dos retornos da vale5 e (d) QQ-Plot dos Retornos da vale5.

Nota-se pela Figura 6.37 (b) que a série apresenta alguns grupos de volatilidade, indicando que variância não seja constante ao longo do tempo, o que sugere a utilização de um modelo GARCH, também pode-se observar em (d) que a distribuição da série parece destoar da distribuição Normal.

(a) Modelo GARCH

Assim como no caso das séries anteriores, inicialmente será ajustado um modelo GARCH para os dados e os resultados serão comparados com os resultados obtidos segundo o método proposto neste trabalho.

Na Figura 6.38, temos os gráficos das funções de autocorrelações (FAC) e autocorrelações parciais (FACP) da série de retornos da vale5.

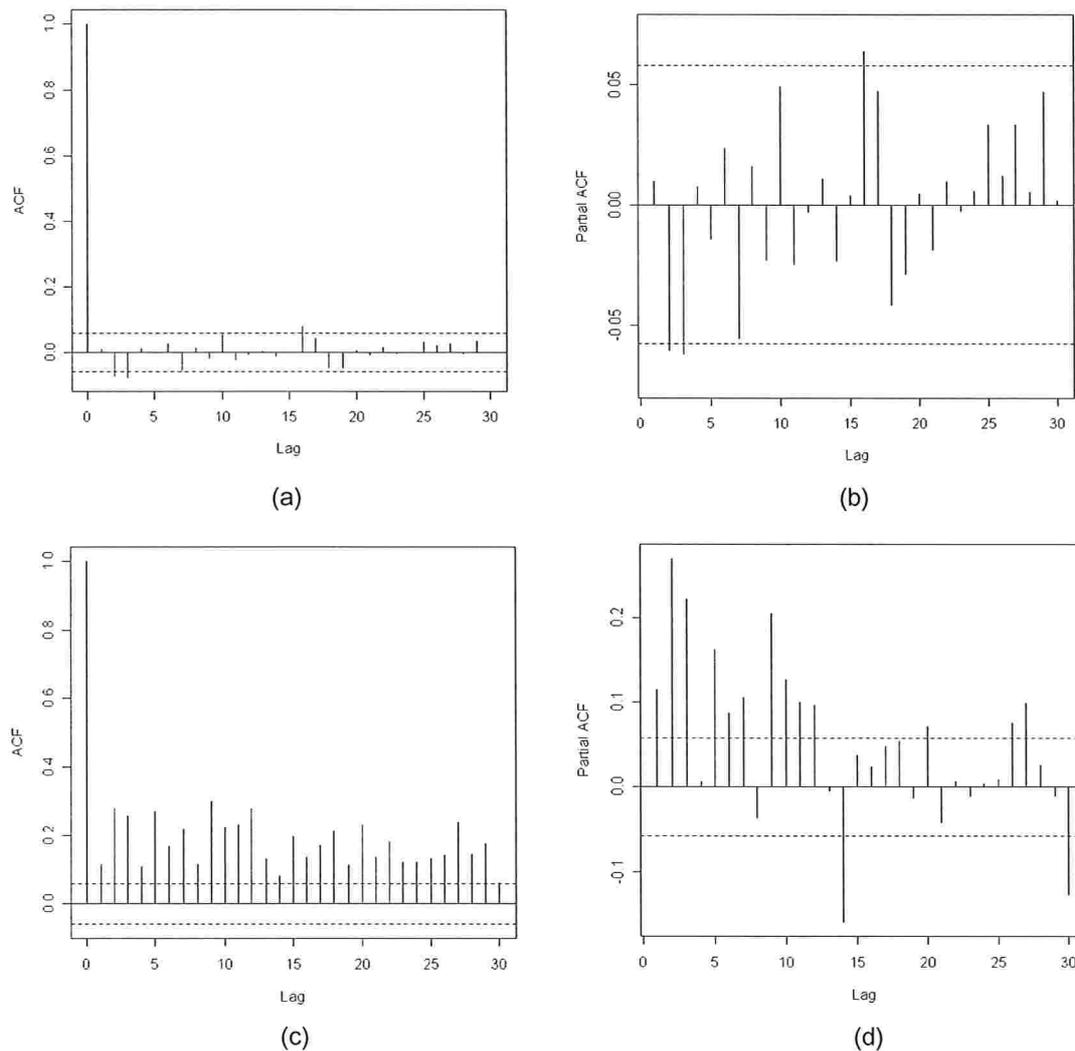


Figura 6.38: (a) FAC dos retornos, (b) FACP dos retornos, (c) FAC dos retornos ao quadrado, (d) FACP dos retornos ao quadrado.

Conforme observado pelos gráficos acima, pode-se notar que não há evidências de que haja autocorrelação entre os retornos. Com relação aos retornos

ao quadrado os gráficos indicam fortes indícios da existência de autocorrelação, o que sugere o ajuste de um modelo GARCH, dessa forma, foi ajustado inicialmente um modelo GARCH(1,1) assumindo distribuição Normal para os erros.

Conforme observado no Apêndice A.12, todos os parâmetros são significantes ao nível de 5%, além disso, pode ser observado através do teste de Ljung Box no Apêndice A.13 que tanto para os resíduos padronizados ε_t / σ_t quanto para os resíduos padronizados ao quadrado $(\varepsilon_t / \sigma_t)^2$ ao nível de 5% não rejeitamos a hipótese H_0 de que eles são independentes.

Os testes de normalidade (Apêndice A.13) Jarque-Bera (p-valor <0,001) e Shapiro-Wilks (p-valor <0,001) para os resíduos padronizados indicam não normalidade.

A Figura 6.39 apresenta o gráfico QQ-plot dos resíduos padronizados e o gráfico das autocorrelações do quadrado dos resíduos padronizados. Observa-se no gráfico QQ-plot que a distribuição parece destoar da distribuição normal, confirmando os resultados dos testes apresentados anteriormente, assim será ajustado a seguir um novo modelo GARCH(1,1) assumindo distribuição t de Student para os erros. Assim como no caso da série do Ibovespa, foram utilizados os graus de liberdade default do R-Plus

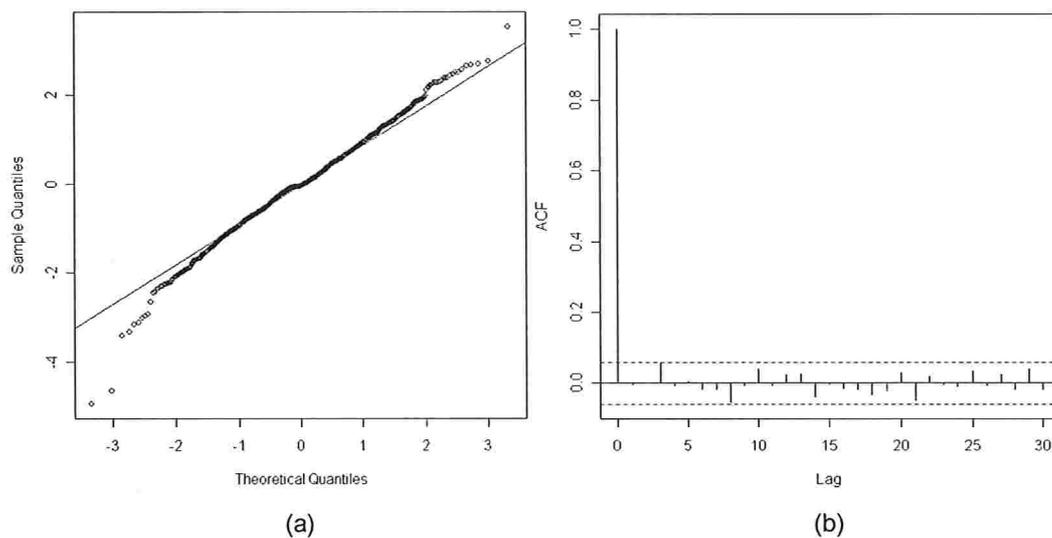


Figura 6.39: (a) QQ-Plot dos resíduos padronizados, (b) Autocorrelação do quadrado dos resíduos padronizados referentes ao modelo com distribuição Normal.

Agora será ajustado um modelo GARCH(1,1) supondo distribuição t de Student para os erros. No Apêndice A.14 são apresentados os resultados do ajuste que indicam ao nível de significância de 5% que todos os parâmetros são diferente de 0.

Observa-se através do valores do Apêndice A.15, que o modelo ajustado considerando a distribuição Normal apresentou valores de BIC e AIC levemente menores que o modelo com a distribuição t de Student, indicando um ajuste melhor.

Através da Figura 6.40 (a), pode-se notar que mesmo utilizando a distribuição t-Student, os dados não foram perfeitamente ajustados, o que pode ser confirmado através dos testes de Jarque-Bera ($p\text{-valor} < 0,001$) e Shapiro-Wilks ($p\text{-valor} < 0,001$).

Assim como no caso do modelo com distribuição normal, o modelo com distribuição t de Student também conseguiu capturar com sucesso a estrutura de correlação serial na variância condicional conforme pode ser observado na Figura 6.40 (b) e através dos testes de Ljung-Box.

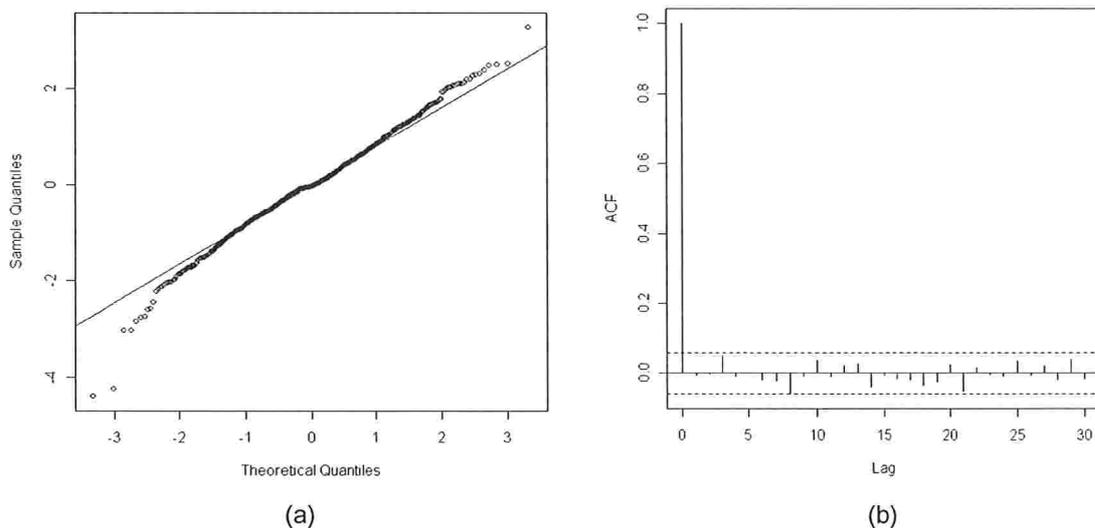


Figura 6.40: (a) QQ-Plot dos resíduos padronizados, (b) Autocorrelação do quadrado dos resíduos padronizados referentes ao modelo com distribuição t de Student.

Dado os resultados obtidos, foi selecionado o modelo assumindo distribuição normal, conforme segue:

$$X_t = \hat{\sigma}_t \varepsilon_t,$$
$$\hat{\sigma}_t^2 = 0,00001445 + 0,07754 * X_{t-1}^2 + 0,9006 * \hat{\sigma}_{t-1}^2 \quad (6.13)$$

(b) Modelo com coeficientes variando no tempo

Após o ajuste dos modelos GARCH apresentados anteriormente foi ajustado um modelo com coeficientes variando no tempo assim como no caso das séries anteriores.

A série de retorno da vale considerada refere-se à 1.147 observações e foi considerada uma janela móvel de tamanho 300, assim foram rodados 1.148 modelos distintos. O modelo considerado foi o GARCH(1,1) com distribuição normal e o procedimento utilizado no ajuste dos modelos é exatamente o mesmo utilizado no caso das séries anteriores. Ao término desse processo foi obtido um vetor de parâmetros com 4 colunas e 1.148 linhas, sendo que cada coluna refere-se à um parâmetro do modelo. Cada coluna foi tratada como uma série temporal independente e para cada uma dessas séries foi ajustado um modelo conforme segue:

Ajuste do modelo para a série de parâmetros $\alpha(t)$

Na figura 6.41, apresentamos (a) a série original de $\alpha(t)$, (b) a série da 1ª diferença de $\alpha(t)$, (c) o histograma da 1ª diferença de $\alpha(t)$ e (d) e o gráfico QQ-plot da 1ª diferença de $\alpha(t)$.

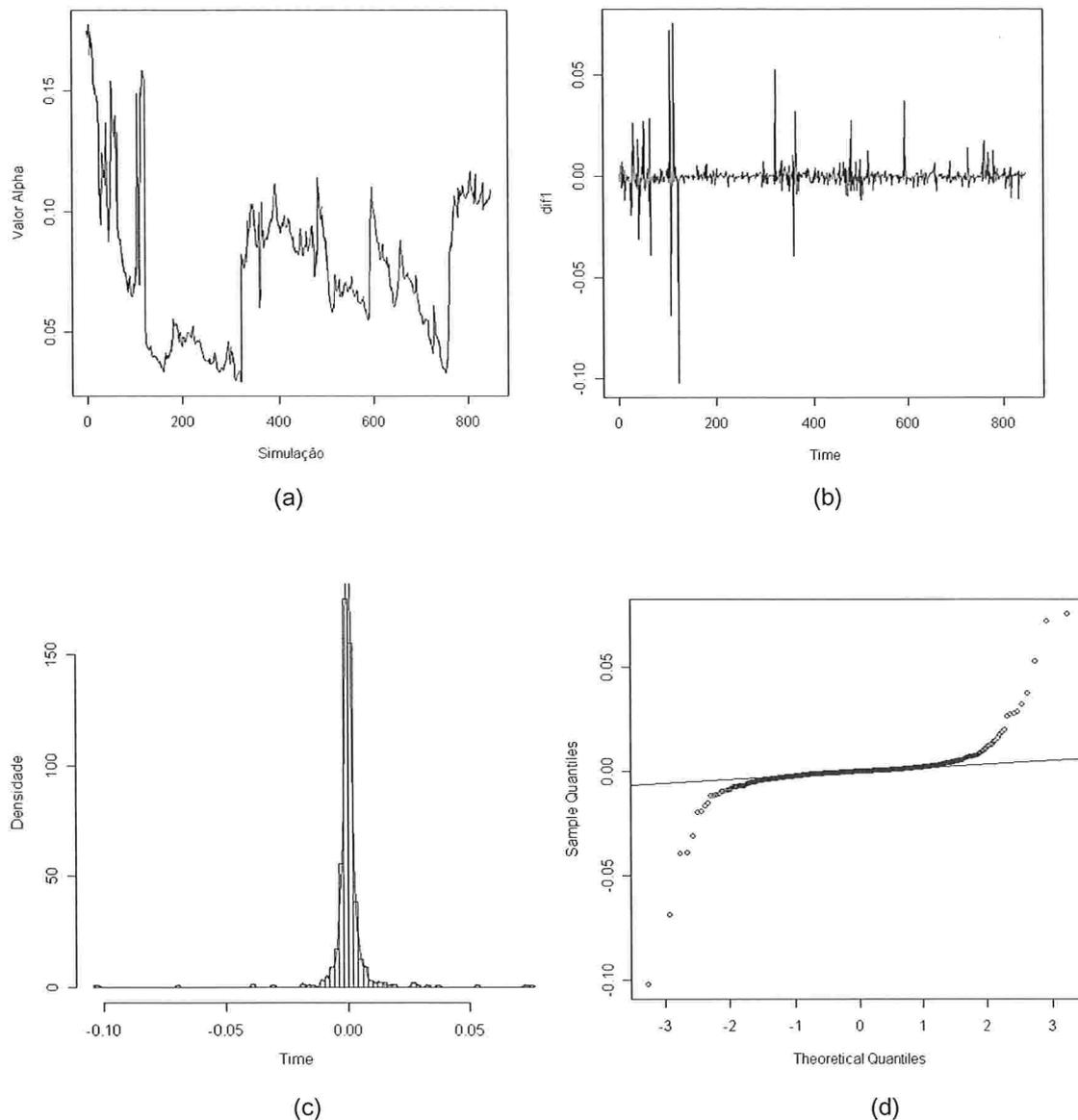


Figura 6.41: (a) Série original de $\alpha(t)$, (b) Série da 1ª diferença de $\alpha(t)$, (c) Histograma da 1ª diferença de $\alpha(t)$ e (d) QQ-Plot da 1ª diferença de $\alpha(t)$.

Pode-se ver pela Figura 6.41 (b), que tomando a 1ª diferença da série dos $\alpha(t)$, a mesma se torna estacionária, porém a série apresenta alguns valores muito altos ou baixos, referentes aos saltos que podem ser vistos na série original em 6.41 (a). Estes aspectos são importante pois podem prejudicar o ajuste do modelo. Em (d) pode-se observar que a distribuição da série parece se destoar da normal.

Na Figura 6.42, são apresentados os gráficos das funções de autocorrelações (FAC) e autocorrelações parciais (FACP) da série da 1ª diferença dos $\alpha(t)$.

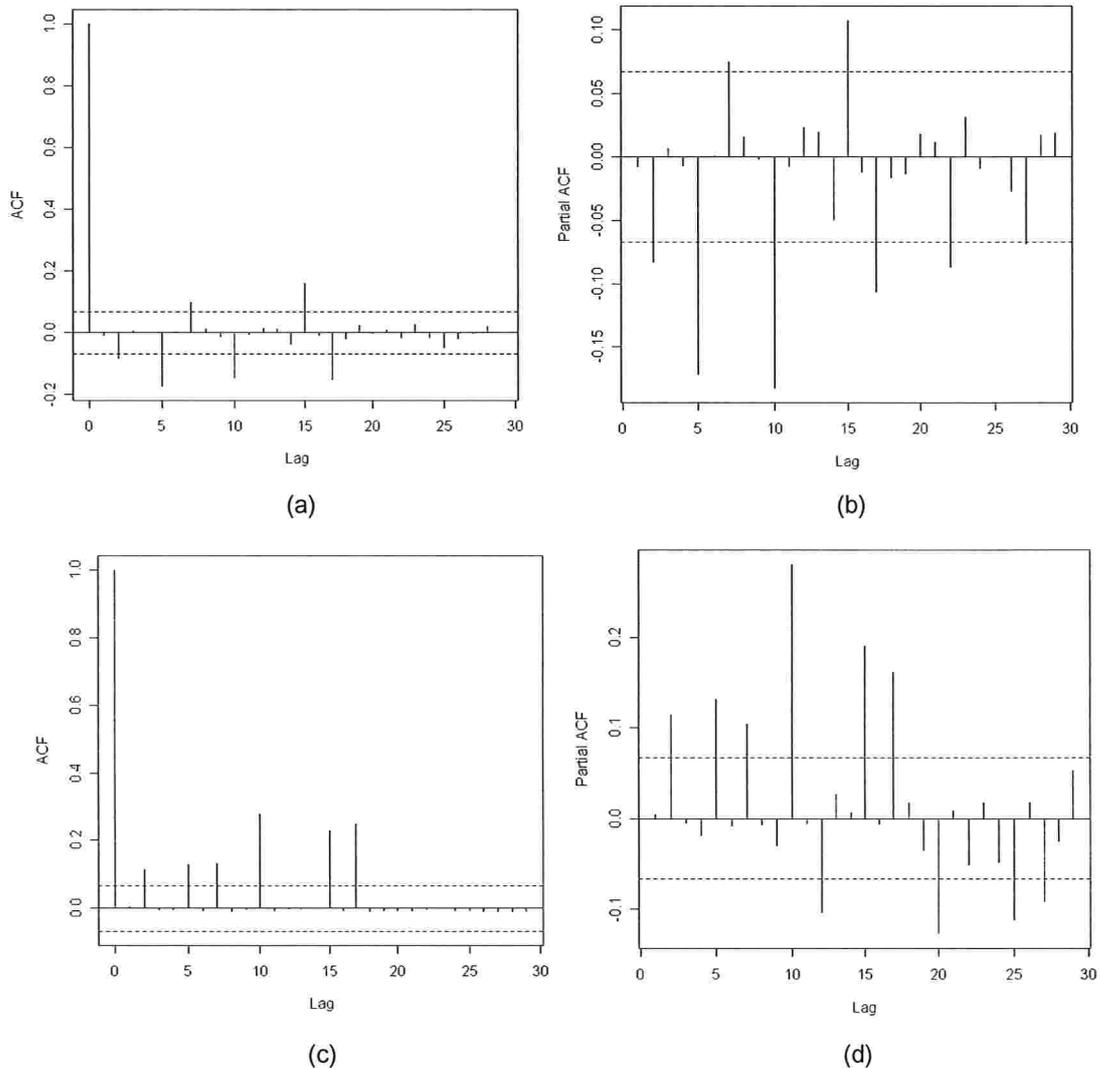


Figura 6.42: (a) FAC da 1ª diferença de $\alpha(t)$, (b) FACP da 1ª diferença de $\alpha(t)$, (c) FAC da 1ª diferença de $\alpha(t)$ ao quadrado, (d) FACP da 1ª diferença de $\alpha(t)$ ao quadrado.

Na Figura 6.42 (a) e (b), há evidências de que haja autocorrelação de lag 5, lag 10, lag 15 e lag 17 na série de $\alpha(t)$. Observando a série referente à 1ª diferença de $\alpha(t)$ ao quadrado, Figura 6.42 (c) e (d), também há evidências da

existência de autocorrelação porém como o interesse é na modelagem da média a volatilidade será desconsiderada. Dado o observado anteriormente foi ajustado um modelo AR(17) sendo que todos os parâmetros diferentes de AR=5, AR=10, AR=15 e AR=17 foram desconsiderados.

Conforme observado no Apêndice A.16, todos os parâmetros estimados são significantes ao nível de 5%, além disso, pode ser observado (Apêndice A.17) através do teste de Box-Pierce (p-valor=0,7613) para os resíduos padronizados ε_t / σ_t , que ao nível de 5% não rejeitamos a hipótese H_0 de que eles são independentes, ou seja, o modelo capturou com sucesso a estrutura de correlação serial o que também pode ser observado pela Figura 6.43.

Os testes de normalidade Jarque-Bera (p-valor=0) e Shapiro-Wilks (p-valor=0) para os resíduos padronizados apresentados no Apêndice A.18 indicam não normalidade.

A Figura 6.43 apresenta o gráfico QQ-plot dos resíduos padronizados e o gráfico das autocorrelações dos resíduos padronizados. Observamos no gráfico QQ-plot que a suposição de normalidade não é apropriada, confirmando os resultados dos testes apresentados anteriormente.

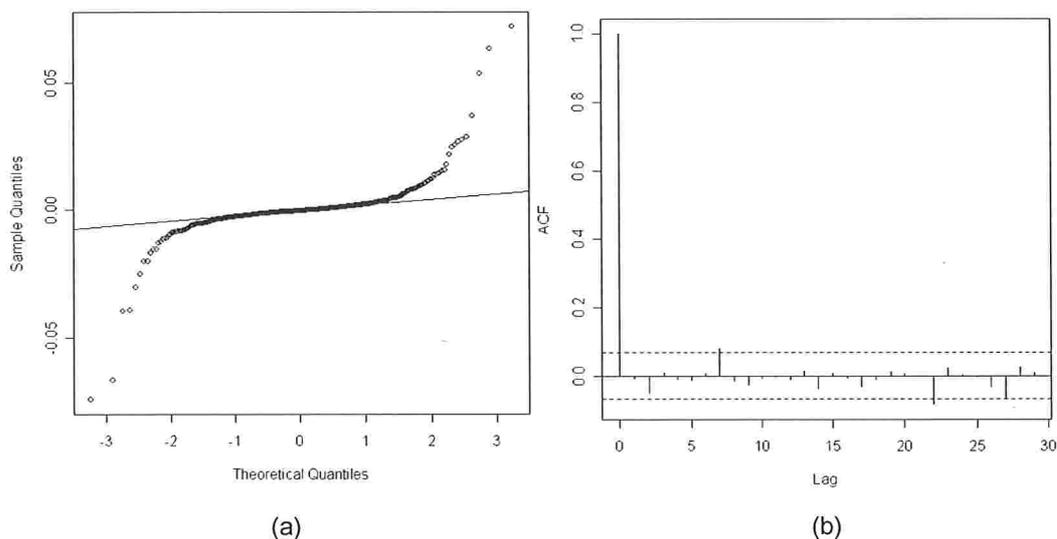


Figura 6.43: (a) QQ-Plot dos resíduos padronizados, (b) Autocorrelação dos resíduos padronizados.

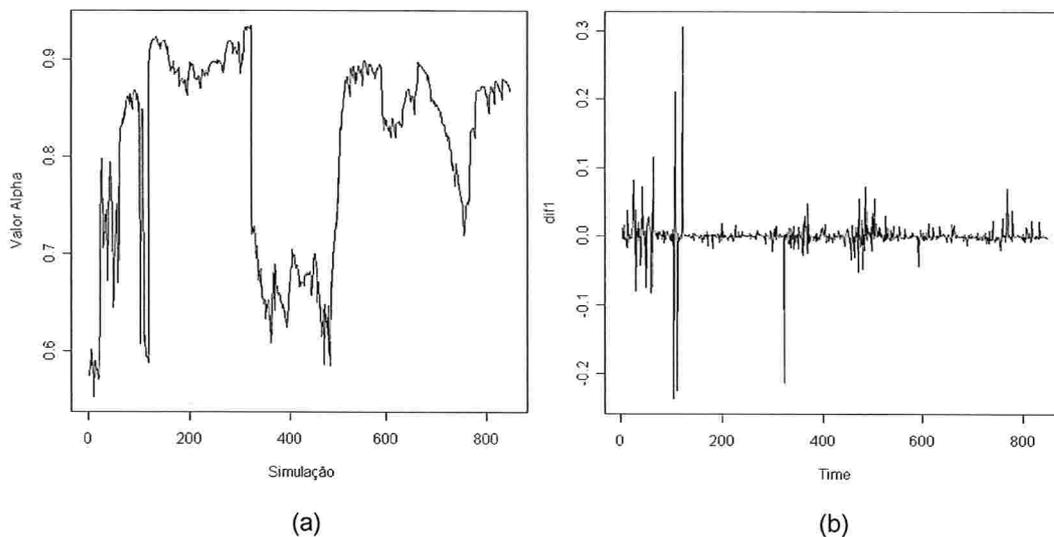
Dados os resultados obtidos no ajuste, o modelo final é:

$$\begin{aligned}
 (\hat{\alpha}(t) - \hat{\alpha}(t-1)) = & -0,18346 * (\hat{\alpha}(t-5) - \hat{\alpha}(t-6)) - 0,14949 * (\hat{\alpha}(t-10) - \hat{\alpha}(t-11)) + \\
 & + 0,09687 * (\hat{\alpha}(t-15) - \hat{\alpha}(t-16)) - 0,12622 * (\hat{\alpha}(t-17) - \hat{\alpha}(t-18))
 \end{aligned}
 \tag{6.14}$$

Um ponto importante neste modelo é que ele ficou com muitos parâmetros, inclusive assumindo lags de ordens bem altas. Isso pode ser efeito dos grandes saltos que a série apresenta conforme mostrado na Figura 6.41 (a).

Ajuste do modelo para a série de parâmetros $\beta(t)$

Na figura 6.44 apresentamos (a) a série original de $\beta(t)$, (b) a série da 1ª diferença de $\beta(t)$, (c) o histograma da 1ª diferença de $\beta(t)$ e (d) o gráfico QQ-plot da 1ª diferença de $\beta(t)$.



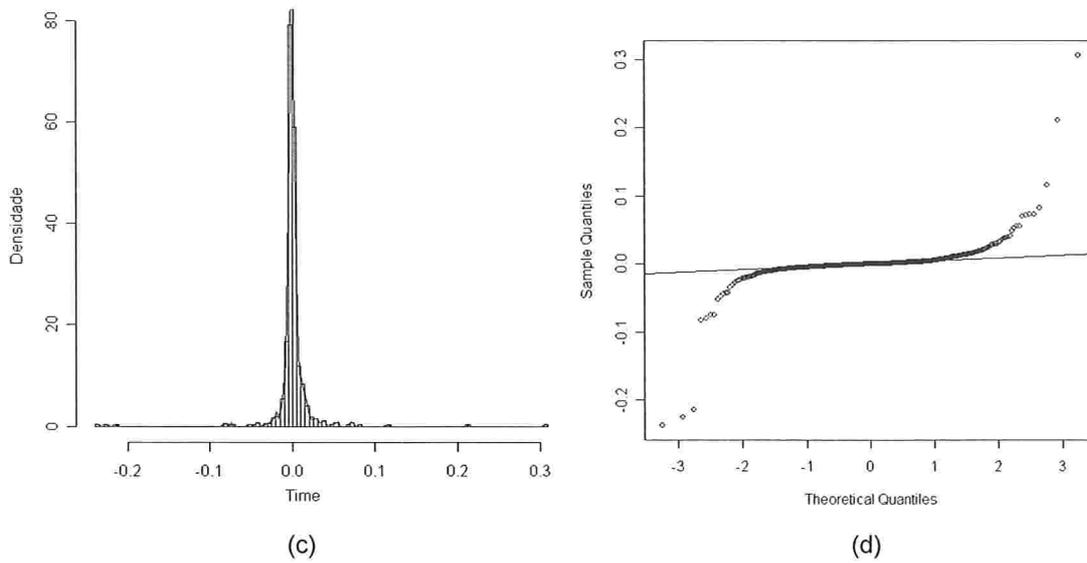


Figura 6.44: (a) Série original de $\beta(t)$, (b) Série da 1ª diferença de $\beta(t)$, (c) Histograma da 1ª diferença de $\beta(t)$ e (d) QQ-Plot da 1ª diferença de $\beta(t)$.

Na Figura 6.44 (b) pode-se notar a presença de grupos de volatilidade e também que a série se tornou estacionário na média ao considerarmos a 1ª diferença da série de $\beta(t)$, porém a série apresenta alguns valores muito altos ou baixos, referentes aos saltos que podem ser vistos na série original em 6.44 (a). Estes aspectos são importante pois podem prejudicar o ajuste do modelo. Em (d) pode-se observar que a distribuição da série parece destoar da distribuição Normal principalmente nas caudas.

Na Figura 6.45, são apresentados os gráficos das funções de autocorrelações (FAC) e autocorrelações parciais (FACP) da série da 1ª diferença de $\beta(t)$ e da 1ª diferença ao quadrado.

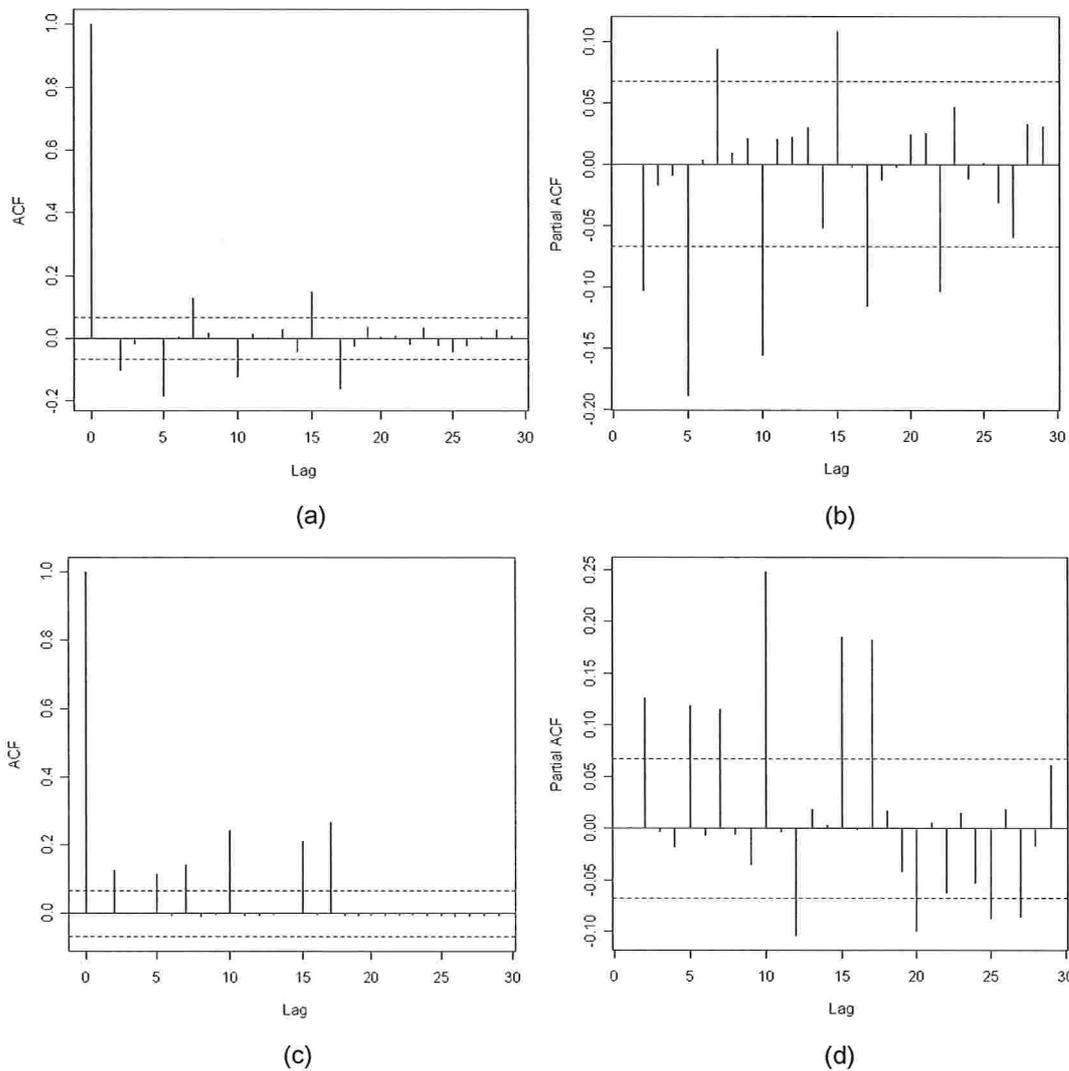


Figura 6.45: (a) FAC da 1ª diferença de $\beta(t)$, (b) FACP da 1ª diferença de $\beta(t)$, (c) FAC da 1ª diferença de $\beta(t)$ ao quadrado, (d) FACP da 1ª diferença de $\beta(t)$ ao quadrado.

Na Figura 6.45 (a) e (b), pode-se notar que há evidências de que exista autocorrelação de lag 5, lag 7, lag 10, lag 15, lag 17 e lag 22 na série de $\beta(t)$. Observando a série referente à 1ª diferença de $\beta(t)$ ao quadrado, Figura 6.45 (c) e (d), também há evidências da existência de autocorrelação, porém como o interesse é na modelagem da média a volatilidade será desconsiderada, assim foi ajustado um modelo AR(22) sendo que todos os parâmetros diferentes de AR=5, AR=7, AR=10, AR=15, AR=17 e AR=22 foram desconsiderados.

Conforme observado no Apêndice A.19, todos os parâmetros estimados são significantes ao nível de 5%, além disso, pode ser observado (Apêndice A.20) através do teste de Box-Pierce (p-valor=0,7588) para os resíduos padronizados ε_i / σ_i , que ao nível de 5% não rejeitamos a hipótese H_0 de que eles são independentes, o que também pode ser observado na Figura 6.46.

Os testes de normalidade Jarque-Bera (p-valor=0) e Shapiro-Wilks (p-valor=0) para os resíduos padronizados apresentados no Apêndice A.20 indicam não normalidade.

A Figura 6.46 apresenta o gráfico QQ-plot dos resíduos padronizados e o gráfico das autocorrelações dos resíduos padronizados. Observamos no gráfico QQ-plot que a suposição de normalidade não é apropriada, confirmando os resultados dos testes apresentados anteriormente.

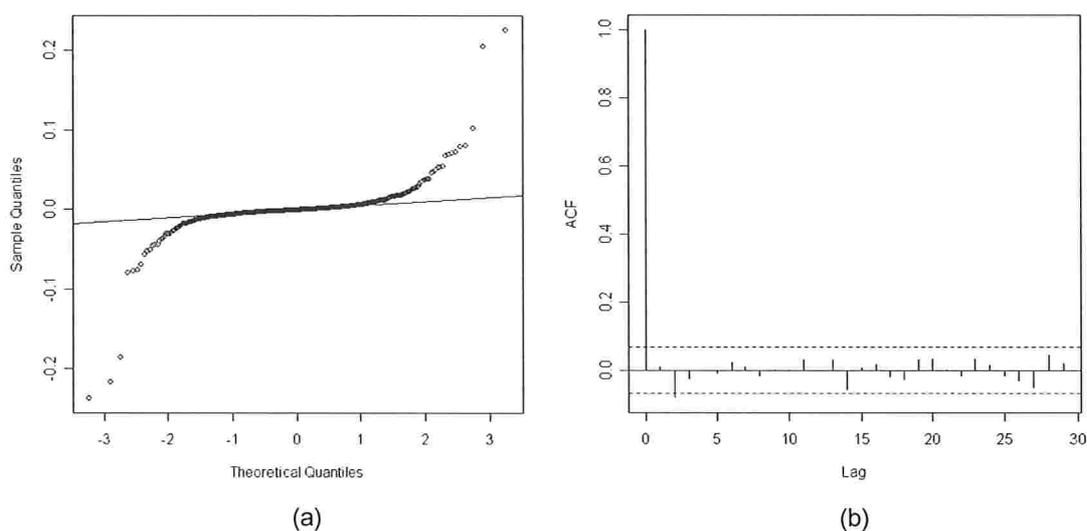


Figura 6.46: (a) QQ-Plot dos resíduos padronizados, (b) Autocorrelação dos resíduos padronizados.

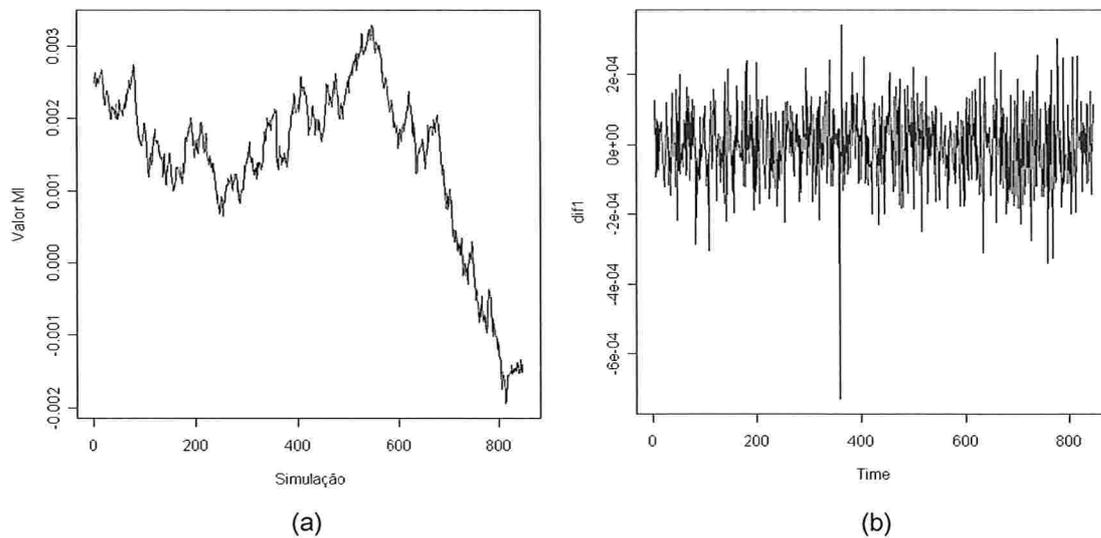
Dados os resultados obtidos no ajuste, o modelo final é:

$$\begin{aligned}
(\hat{\beta}(t) - \hat{\beta}(t-1)) = & -0,19859 * (\hat{\beta}(t-5) - \hat{\beta}(t-6)) + 0,10204 * (\hat{\beta}(t-7) - \hat{\beta}(t-8)) - \\
& - 0,12474 * (\hat{\beta}(t-10) - \hat{\beta}(t-11)) + 0,09938 * (\hat{\beta}(t-15) - \hat{\beta}(t-16)) - \\
& - 0,14118 * (\hat{\beta}(t-16) - \hat{\beta}(t-17)) + 0,10414 * (\hat{\beta}(t-22) - \hat{\beta}(t-23))
\end{aligned}
\tag{6.15}$$

Apesar do modelo ter conseguido capturar com sucesso a estrutura de correlação dos dados, ele ficou com muitos parâmetros além de lags de ordens altas o que prejudica sua interpretação.

Ajuste do modelo para a série de parâmetros $\mu(t)$

Na figura 6.47, apresentamos (a) a série original de $\mu(t)$, (b) a série da 1ª diferença de $\mu(t)$, (c) o histograma da 1ª diferença de $\mu(t)$ e (d) o gráfico QQ-plot da 1ª diferença de $\mu(t)$.



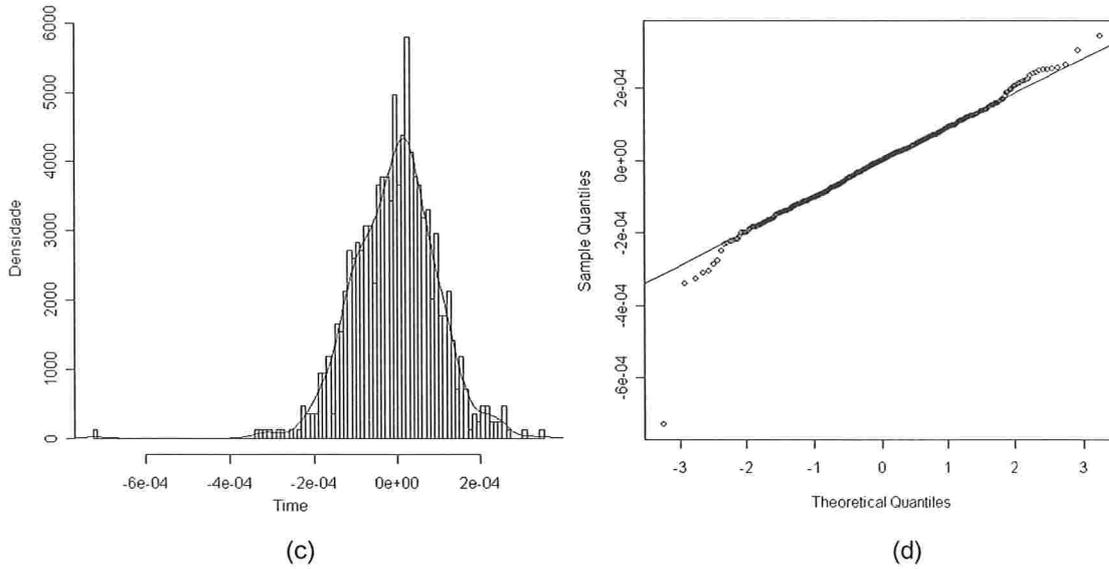
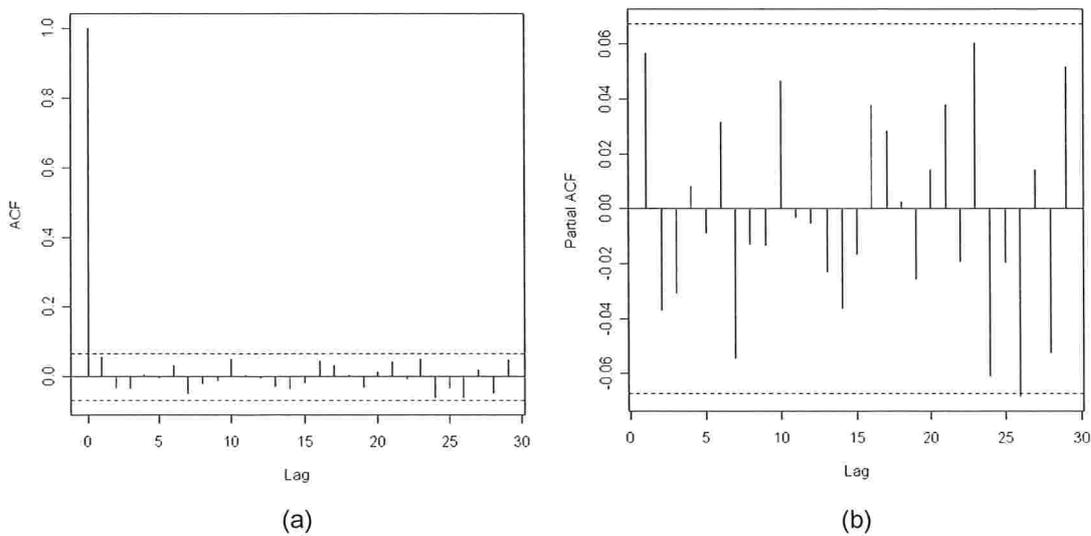


Figura 6.47: (a) Série original de $\mu(t)$, (b) Série da 1ª diferença de $\mu(t)$, (c) Histograma da 1ª diferença de $\mu(t)$ e (d) QQ-Plot da 1ª diferença de $\mu(t)$.

Na Figura 6.47 (b) pode-se ver que tomando a 1ª diferença da série de $\mu(t)$, a mesma se torna estacionária, além disso, em (d) pode-se observar que a distribuição da série parece aproximar-se de uma normal com exceção de alguns pontos na cauda.

Na Figura 6.48, são apresentados os gráficos das funções de autocorrelações (FAC) e autocorrelações parciais (FACP) da série da 1ª diferença de $\mu(t)$ e da 1ª diferença ao quadrado.



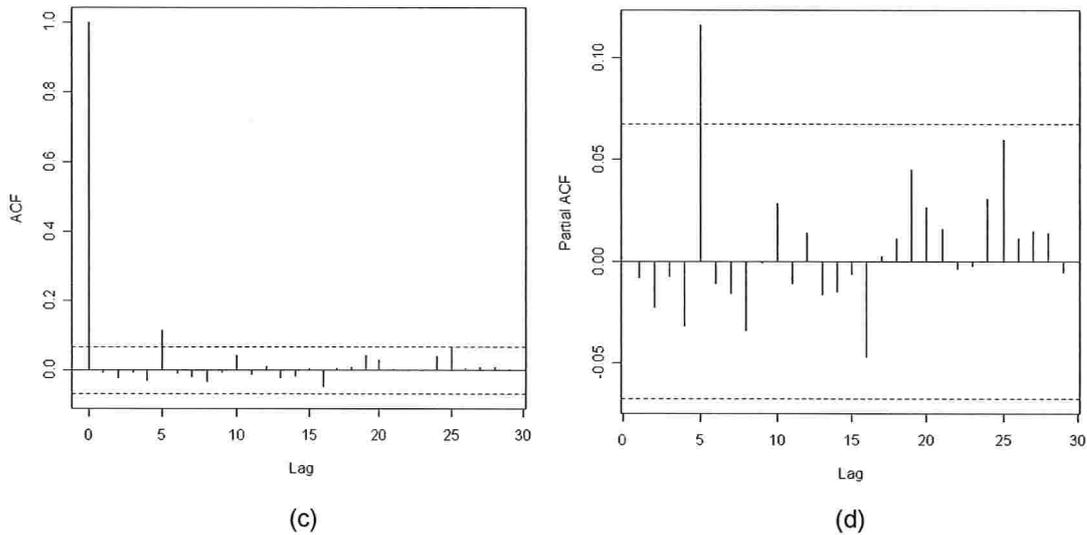


Figura 6.48: (a) FAC da 1ª diferença de $\mu(t)$, (b) FACP da 1ª diferença de $\mu(t)$, (c) FAC da 1ª diferença de $\mu(t)$ ao quadrado, (d) FACP da 1ª diferença de $\mu(t)$ ao quadrado.

Na Figura 6.48, pode-se notar que não há evidências de que exista autocorrelação na série referente à 1ª diferença de $\mu(t)$, de qualquer forma, será ajustado um modelo AR(1) de modo a confirmar esta hipótese.

Conforme observado no Apêndice A.22, adotando um nível de significância de 5% não rejeitamos a hipótese H_0 de que o coeficiente do modelo ajustado seja 0, conforme já visto na Figura 6.48.

Neste caso, como os dados não apresentaram correlação serial será considerado o seguinte modelo:

$$\mu(t) = \mu(t - 1) + \varepsilon_t \tag{6.16}$$

Ajuste do modelo para a série de parâmetros $\alpha_0(t)$

Na figura 6.49, apresentamos (a) a série original de $\alpha_0(t)$, (b) a série da 1ª diferença de $\alpha_0(t)$, (c) o histograma da 1ª diferença de $\alpha_0(t)$ e (d) e o gráfico QQ-plot da 1ª diferença de $\alpha_0(t)$.

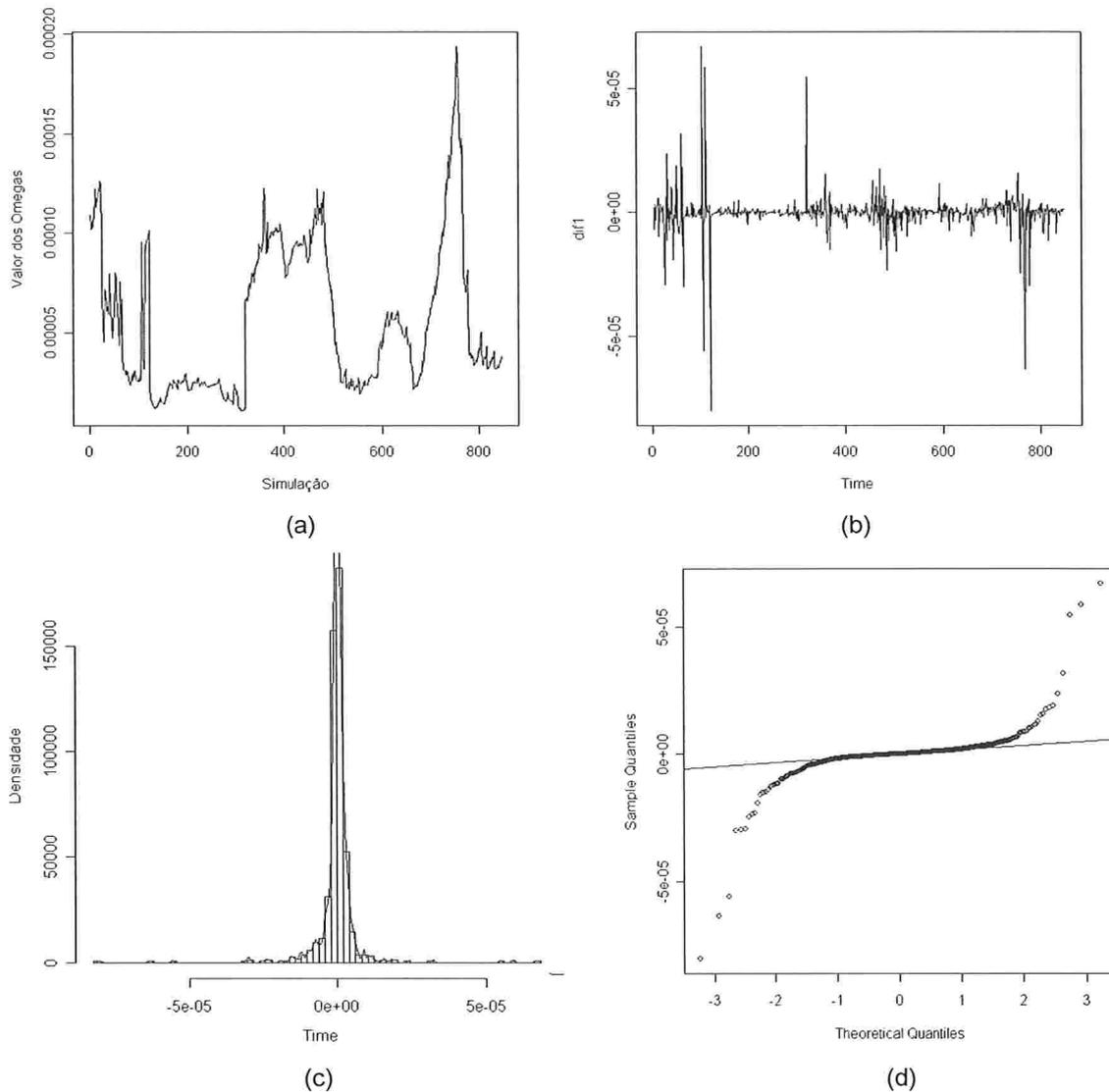


Figura 6.49: (a) Série original de $\alpha_0(t)$, (b) Série da 1ª diferença de $\alpha_0(t)$, (c) Histograma da 1ª diferença de $\alpha_0(t)$ e (d) QQ-Plot da 1ª diferença de $\alpha_0(t)$.

Na Figura 6.49 (b) pode-se ver que tomando a 1ª diferença da série de $\alpha_0(t)$, a mesma se torna estacionária e também é forte a presença de grupos de volatilidade, além disso, a série apresenta alguns valores muito altos ou baixos, referentes aos saltos que podem ser vistos na série original em 6.49 (a). Estes aspectos são importante pois podem prejudicar o ajuste do modelo. Em (d) pode-se observar que a distribuição da série parece destoar da normal.

Na Figura 6.50, são apresentados os gráficos das funções de autocorrelações (FAC) e autocorrelações parciais (FACP) da série da 1ª diferença dos $\alpha_0(t)$ e da 1ª diferença ao quadrado.

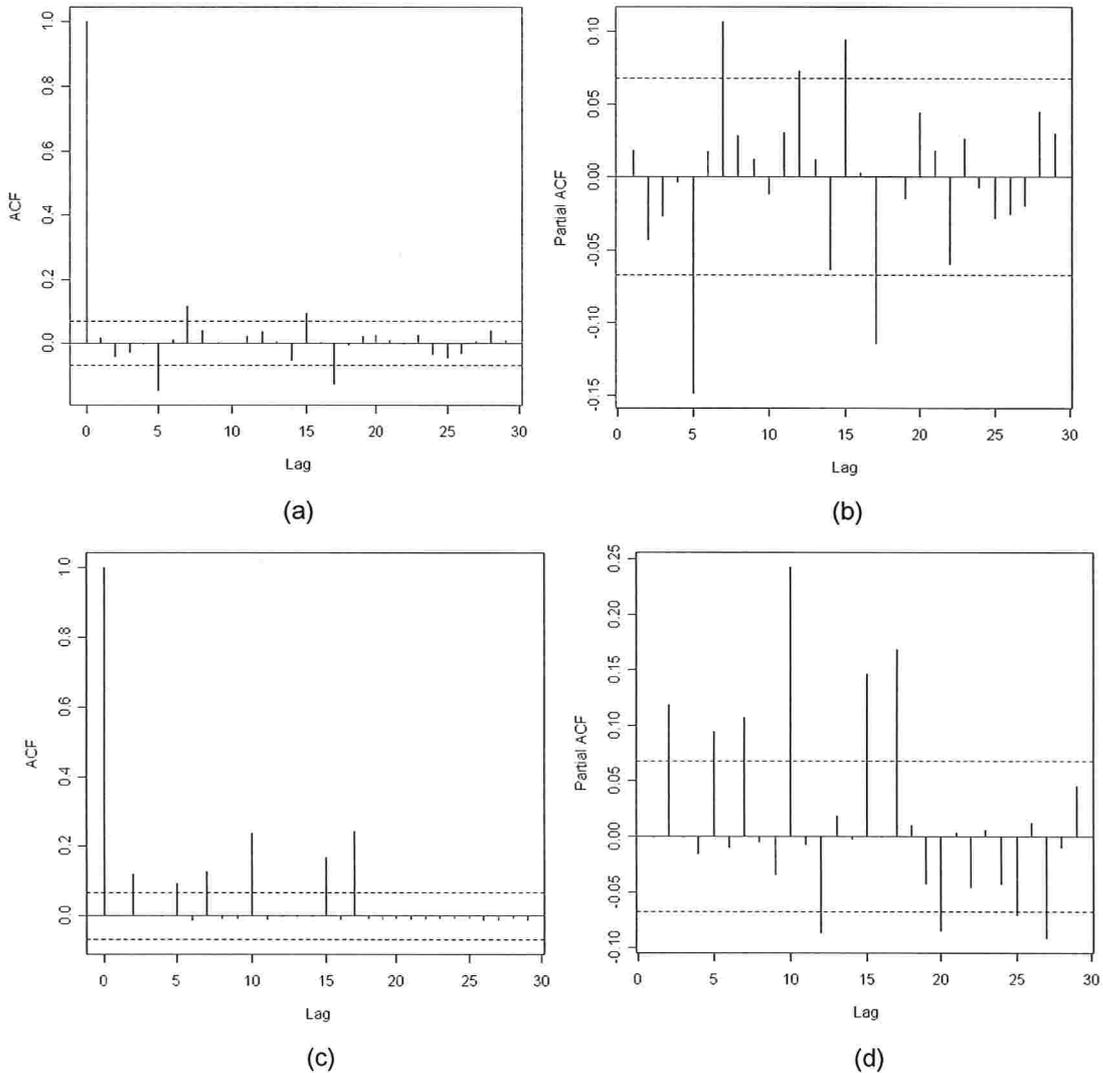


Figura 6.50: (a) FAC da 1ª diferença de $\alpha_0(t)$, (b) FACP da 1ª diferença de $\alpha_0(t)$, (c) FAC da 1ª diferença de $\alpha_0(t)$ ao quadrado, (d) FACP da 1ª diferença de $\alpha_0(t)$ ao quadrado.

Na Figura 6.50 (a) e (b), há evidências de que exista autocorrelação de lag 5, lag 7, lag 15 e lag 17 na série de $\alpha_0(t)$. Observando a série referente à 1ª diferença de $\alpha_0(t)$ ao quadrado, Figura 6.50 (c) e (d), também há evidências da existência de autocorrelação, porém como o interesse é na modelagem da média a

volatilidade será desconsiderada, assim foi ajustado um modelo AR(17) sendo que todos os parâmetros diferentes de AR=5, AR=7, AR=15 e AR=17 foram desconsiderados. Abaixo seguem os resultados do ajuste:

Conforme observado no Apêndice A.25, todos os parâmetros estimados são significantes ao nível de 5%, além disso, pode ser observado (Apêndice A.26) através do teste de Box-Pierce (p-valor=0,5255) para os resíduos padronizados ε_t / σ_t , que ao nível de 5% não rejeitamos a hipótese H_0 de que eles são independentes.

Os testes de normalidade Jarque-Bera (p-valor=0) e Shapiro-Wilks (p-valor=0) para os resíduos padronizados apresentado no Apêndice A.27 indicam não normalidade.

Observamos na Figura 6.51, através do gráfico QQ-plot dos resíduos padronizados que a suposição de normalidade não é apropriada, confirmando os resultados dos testes apresentados anteriormente.

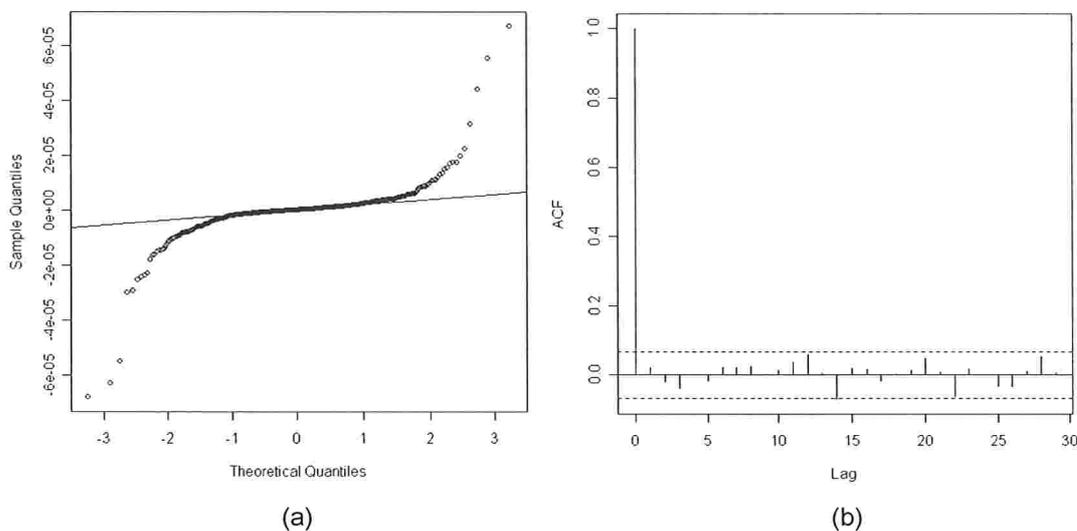


Figura 6.51: (a) QQ-Plot dos resíduos padronizados, (b) Autocorrelação dos resíduos padronizados.

Dados os resultados obtidos no ajuste, o modelo final é:

$$\begin{aligned}
 (\alpha_0(t) - \alpha_0(t-1)) = & -0,13655 * (\alpha_0(t-5) - \alpha_0(t-6)) + 0,10756 * (\alpha_0(t-7) - \alpha_0(t-8)) + \\
 & + 0,08501 * (\alpha_0(t-15) - \alpha_0(t-16)) - 0,11724 * (\alpha_0(t-17) - \alpha_0(t-18))
 \end{aligned}
 \tag{6.17}$$

Conforme observado na Figura 6.49 (b), existem alguns pontos aberrantes e estes podem ter feito com que o modelo ficasse com muitos parâmetros, além de lags de ordens altas, por isso, devemos ter cuidado ao analisá-lo.

Previsões

Após o desenvolvimento dos modelos citados anteriormente, eles foram utilizados para previsão. Os valores previstos foram comparados com os valores reais da série referentes ao período 26 de janeiro de 2009 à 06 de fevereiro de 2009. Assim como no caso das séries anteriores foi calculado o EQM para cada metodologia, abaixo seguem os resultados obtidos:

Cálculo do EQM – Modelo GARCH(1,1) com distribuição Normal

Date	Preço Fechamento	Log(P _t / P _{t-1})	Predito - Garch(1,1) Dist Normal	Erro Quadrático Médio
			Previsão σ_t^2	Garch(1,1)
22/1/2009	25,20	-		
23/1/2009	26,30	0,0186	-	
26/1/2009	26,95	0,0106	0,0011179	1,011E-06
27/1/2009	27,93	0,0155	0,0011079	7,522E-07
28/1/2009	29,37	0,0218	0,0010982	3,862E-07
29/1/2009	28,50	-0,0131	0,0010886	8,428E-07
30/1/2009	28,01	-0,0075	0,0010793	1,046E-06
2/2/2009	27,75	-0,0041	0,0010701	1,11E-06
3/2/2009	29,00	0,0191	0,0010612	4,83E-07
4/2/2009	30,05	0,0154	0,0010524	6,623E-07
5/2/2009	31,35	0,0184	0,0010438	4,978E-07
6/2/2009	32,48	0,0154	0,0010355	6,384E-07
				<u>7,43E-06</u>

Tabela 6.5: Cálculo do EQM para o modelo GARCH(1,1)

A variância condicional σ_t^2 e o EQM (Erro quadrático médio) foram obtidos da mesma forma que no caso das séries anteriores. Conforme observado na tabela acima, o valor do EQM calculado foi $7,43e^{-6}$.

Cálculo do EQM – Modelo com coeficientes variando no tempo

Date	Preço Fecham.	Log (P _t / P _{t-1})	Metodologia Proposta					Erro Quadrático Médio	
			$\alpha_1(t)$	$\beta(t)$	$\mu(t)$	$\alpha_0(t)$	Previsão σ_t^2	Metodologia Proposta	
22/1/2009	25,20	-	-	-	-	-	-	-	-
23/1/2009	26,30	0,0186	-	-	-	-	0,0012586	-	-
26/1/2009	26,95	0,0106	0,1086	0,8682	-0,0014	0,0000383	0,0011743		1,12756E-06
27/1/2009	27,93	0,0155	0,1101	0,8653	-0,0014	0,0000395	0,0010715		6,90358E-07
28/1/2009	29,37	0,0218	0,1096	0,8661	-0,0014	0,0000397	0,0009991		2,72875E-07
29/1/2009	28,50	-0,0131	0,1094	0,8673	-0,0014	0,0000397	0,0009652		6,31438E-07
30/1/2009	28,01	-0,0075	0,1088	0,8688	-0,0014	0,0000397	0,0008930		6,99339E-07
2/2/2009	27,75	-0,0041	0,1088	0,8694	-0,0014	0,0000398	0,0008202		6,46142E-07
3/2/2009	29,00	0,0191	0,1084	0,8682	-0,0014	0,0000397	0,0007526		1,49356E-07
4/2/2009	30,05	0,0154	0,1081	0,8687	-0,0014	0,0000400	0,0007393		2,50744E-07
5/2/2009	31,35	0,0184	0,1081	0,8686	-0,0014	0,0000403	0,0007132		1,40555E-07
6/2/2009	32,48	0,0154	0,1079	0,8692	-0,0014	0,0000405	0,0007027		2,17335E-07
									4,8257E-06

Tabela 6.6: Cálculo do EQM para o modelo com coeficientes variando no tempo

As previsões 10 passos à frente para $\alpha_0(t)$, $\alpha_1(t)$, $\beta(t)$ e $\mu(t)$ mostradas na Tabela 6.6 foram obtidas com base nos modelos (6.14, 6.15, 6.16 e 6.17) ajustados para cada parâmetro apresentados anteriormente.

A previsão da variância condicional σ_t^2 foi obtida da mesma forma como mostrado para as séries anteriores. O valor do EQM obtido foi $4,8257e^{-6}$, ou seja, a utilização da metodologia proposta ajudou a reduzir o EQM. Calculando a razão entre os EQMs, obtemos:

$$\text{Razão} = \frac{4,8257e^{-6}}{7,43e^{-6}} = 0.649\%$$

indicando que foi possível reduzir o EQM em 35 pontos percentuais.

6.4 Conclusão

Conforme observado em todos os exemplos mostrados neste capítulo o método proposto apresentou um desempenho melhor que o método normalmente utilizado (Modelos GARCH), sendo que o maior ganho foi obtido na série do Ibovespa e o menor ganho foi obtido com a série do Itaú.

Capítulo 7

Conclusão e Trabalhos Futuros

Neste trabalho apresentamos um método alternativo para estimação dos parâmetros do modelo GARCH, supondo que estes variam no tempo e comparamos os resultados obtidos na previsão 10 passos a frente com os resultados obtidos com o modelo GARCH tradicional em que os parâmetros são fixos ao longo do tempo.

O modelo GARCH com parâmetros fixos já foram muito estudados e hoje são bem conhecidos, além disso, eles conseguem capturar de maneira satisfatória muitos aspectos das séries temporais, como visto no capítulo de aplicação à dados reais em ele conseguiu capturar com êxito a estrutura de correlação existente nos 3 conjuntos de dados considerados Ibovespa, Vale e Itau. Apesar dos modelos GARCH serem muito usados e apresentarem boas propriedades, existem algumas séries temporais como no caso de ações no qual talvez alguns aspectos não sejam modelados de maneira satisfatória.

Em paralelo ao modelo GARCH tradicional, também foi ajustado um modelo com parâmetros variando no tempo segundo a metodologia proposta. Modelos com coeficiente variando no tempo já foram propostos anteriormente, o grande diferencial desse trabalho está na forma da função que esta sendo utilizada para definir o padrão de variação destes parâmetros, no qual seguem um modelo ARIMA. Neste caso, também foram realizadas previsões 10 passos à frente e conforme mostrado, para todas as séries (Ibovespa, Vale e Itau) tivemos ganho de previsão quando analisado o EQM (Erro Quadrático Médio) em comparação ao modelo GARCH tradicional. O maior ganho obtido, foi com relação a série do Ibovespa.

Todas as comparações foram feitas de maneira empírica devido ao grau de dificuldade envolvida na demonstração teórica dos resultados apresentados acima, o que fugia do escopo deste trabalho.

A sugestão do método proposto foi apenas o 1º passo de uma longa estrada que pode ser percorrida, assim se abrem muitas possibilidades de trabalhos futuros, como por exemplo, continuar testando experimentalmente este método em outras séries, mostrar teoricamente alguns resultados, comparar os resultados deste método com algumas variações do modelo GARCH, como por exemplo, o modelo TGARCH, pesquisar métodos que possam ser utilizados na definição do tamanho da janela móvel utilizada, entre outros. Ainda há muita coisa a ser feita e apesar da teoria de séries temporais ter evoluído muito nas últimas décadas existem alguns tópicos que ainda tem muito a evoluir e com certeza modelos com coeficientes variando no tempo é um deles.

Apêndice A

Outputs referentes ao modelos ajustados no capítulo 6

A.1 Ajuste do modelo GARCH(1,1) com dist. Normal para a série do Itau

```
> summary(obj)

Title:
  GARCH Modelling

Call:
  garchFit(formula = ~garch(1, 1), data = dif)

Mean and Variance Equation:
  ~arma(0, 0) + ~garch(1, 1)

Conditional Distribution:
  dnorm

Coefficient(s):
      mu      omega      alpha      beta1
1.18700e-03  8.75847e-06  5.20270e-02  9.30035e-01

Error Analysis:
      Estimate  Std. Error  t value Pr(>|t|)
omega  8.758e-06  4.808e-06    1.822 0.068524 .
alpha  5.203e-02  1.530e-02    3.400 0.000674 ***
beta1  9.300e-01  2.032e-02   45.763 < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Log Likelihood:
-1566.535    normalized: -2.436290
```

A.2 Teste de Normalidade e Ljung-Box referente ao modelo GARCH(1,1) com dist. Normal ajustado para a série do Itaú

```
Standadized Residuals Tests:

      Statistic p-Value
Jarque-Bera Test  R    Chi^2  15.74470  0.0003811371
Shapiro-Wilk Test R    W      0.9894315 0.0001386856
Ljung-Box Test   R    Q(10)  9.966657  0.4434235
Ljung-Box Test   R    Q(15)  12.97180  0.604476
Ljung-Box Test   R    Q(20)  16.31254  0.697055
Ljung-Box Test   R^2  Q(10)  13.42782  0.2007253
Ljung-Box Test   R^2  Q(15)  18.54376  0.2351573
Ljung-Box Test   R^2  Q(20)  26.44537  0.1516013
LM Arch Test     R    TR^2  15.69050  0.2058263

Information Criterion Statistics:
      AIC      BIC      SIC      HQIC
4.885022  4.912806  4.884946  4.895804

Description:
  Fri Jan 01 16:23:48 2010 by user: Acer
```

A.3 Ajuste do modelo GARCH(1,1) com dist. t-Student para a série do Itau

```

> summary(obj)

Title:
  GARCH Modelling

Call:
  garchFit(formula = ~garch(1, 1), data = dif, cond.dist = "dstd",
    include.shape = FALSE)

Mean and Variance Equation:
  ~arma(0, 0) + ~garch(1, 1)

Conditional Distribution:
  dstd

Coefficient(s):
      mu      omega      alpha      beta1
4.81823e-04  7.00660e-06  8.14967e-02  9.25247e-01

Error Analysis:
      Estimate Std. Error t value Pr(>|t|)
alpha1 8.150e-02  2.566e-02   3.177 0.00149 **
beta1  9.252e-01  2.345e-02  39.451 < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Log Likelihood:
-1566.330    normalized: -2.435973

```

A.4 Teste de Normalidade e Ljung-Box referente ao modelo GARCH(1,1) com dist. t-Student ajustado para a série do Itau

```

Standardized Residuals Tests:

      Jarque-Bera Test  R  Chi^2  18.76600  8.41423e-05
      Shapiro-Wilk Test  R  W      0.988734  7.504985e-05
      Ljung-Box Test    R  Q(10)  9.774954  0.4604544
      Ljung-Box Test    R  Q(15)  12.81099  0.6168941
      Ljung-Box Test    R  Q(20)  16.13362  0.7083006
      Ljung-Box Test    R^2 Q(10)  12.01671  0.2839398
      Ljung-Box Test    R^2 Q(15)  17.19301  0.3074580
      Ljung-Box Test    R^2 Q(20)  25.16482  0.1951982
      LM Arch Test      R  TR^2   15.00061  0.2414029

Information Criterion Statistics:
      AIC      BIC      SIC      HQIC
4.884387 4.912170 4.884310 4.895169

Description:
Fri Jan 01 16:19:35 2010 by user: Acer

```

A.5 Ajuste do modelo ARMA para a série de $\beta(t)$ referente a série do Itaú.

```
> summary(Beta)

Call:
arma(x = dif1, lag = list(ar = c(4, 9)), include.intercept = FALSE)

Model:
ARMA(9,0)

Residuals:
      Min       1Q   Median       3Q      Max
-0.0137654 -0.0009980  0.0002891  0.0012651  0.0164446

Coefficient(s):
      Estimate Std. Error t value Pr(>|t|)
ar4  -0.13703    0.05303   -2.584  0.00976 **
ar9   0.12445    0.05314    2.342  0.01919 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Fit:
sigma^2 estimated as 1.056e-05,  Conditional Sum-of-Squares = 0,  AIC = -2944.18
```

A.6 Teste de Box-Pierce para a série de $\beta(t)$ referente a série do Itaú.

```
> Box.test(res_beta)

      Box-Pierce test

data:  res_beta
X-squared = 0.8802, df = 1, p-value = 0.3481
```

A.7 Teste de Normalidade para a série de $\beta(t)$ referente a série do Itaú.

```
> jarque.bera.test(res_beta_norm)

      Jarque Bera Test

data:  res_beta_norm
X-squared = 483.8456, df = 2, p-value < 2.2e-16

> shapiro.test(res_beta_norm)

      Shapiro-Wilk normality test

data:  res_beta_norm
W = 0.8727, p-value = 5.962e-16
```

A.8 Ajuste do modelo ARMA para a série de $\mu(t)$ referente a série do Itaú.

```
> summary(MI)

Call:
arma(x = dif1, lag = list(ar = c(6)), include.intercept = FALSE)

Model:
ARMA(6,0)

Residuals:
    Min       1Q   Median       3Q      Max
-3.666e-04 -6.734e-05 -2.731e-06  4.791e-05  4.794e-04

Coefficient(s):
      Estimate Std. Error t value Pr(>|t|)
ar6  -0.12597    0.05328   -2.364  0.0181 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Fit:
sigma^2 estimated as 8.508e-09, Conditional Sum-of-Squares = 0, AIC = -5382.56
```

A.9 Teste de Box-Pierce para a série de $\mu(t)$ referente a série do Itaú.

```
> Box.test(res_mi, lag = 5, type = c("Ljung-Box"))

      Box-Ljung test

data:  res_mi
X-squared = 2.3502, df = 5, p-value = 0.06029
```

A.10 Teste de Normalidade para a série de $\mu(t)$ referente a série do Itaú.

```
> jarque.bera.test(res_mi_norm)

      Jarque Bera Test

data:  res_mi_norm
X-squared = 260.1642, df = 2, p-value < 2.2e-16

> shapiro.test(res_mi_norm)

      Shapiro-Wilk normality test

data:  res_mi_norm
W = 0.9537, p-value = 8.636e-09
```

A.11 Ajuste do modelo ARMA para a série de $\alpha_0(t)$ referente a série do Itaú.

```
> summary(Omega)

Call:
arma(x = dif1, lag = list(ar = c(4, 18)), include.intercept = FALSE)

Model:
ARMA(18,0)

Residuals:
      Min       1Q   Median       3Q      Max
-3.646e-04 -6.509e-05 -4.914e-06  4.626e-05  4.962e-04

Coefficient(s):
      Estimate Std. Error t value Pr(>|t|)
ar4  -0.078870   0.054010   -1.460   0.144
ar18 -0.004694   0.053075   -0.088   0.930

Fit:
sigma^2 estimated as 8.59e-09, Conditional Sum-of-Squares = 0, AIC = -5377.29
```

A.12 Ajuste do modelo GARCH(1,1) com dist. Normal para a série da Vale

```
> summary(obj)

Title:
  GARCH Modelling

Call:
  garchFit(formula = ~garch(1, 1), data = dif)

Mean and Variance Equation:
  ~arma(0, 0) + ~garch(1, 1)

Conditional Distribution:
  dnorm

Coefficient(s):
      mu      omega    alpha1    beta1
1.59571e-03  1.44498e-05  7.75369e-02  9.00597e-01

Error Analysis:
      Estimate Std. Error t value Pr(>|t|)
mu      1.596e-03   6.413e-04   2.488   0.0128 *
omega   1.445e-05   6.053e-06   2.387   0.0170 *
alpha1  7.754e-02   1.847e-02   4.199 2.68e-05 ***
beta1   9.006e-01   2.496e-02  36.089 < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Log Likelihood:
-2671.058      normalized: -2.330766
```

A.13 Teste de Normalidade e Ljung-Box referente ao modelo GARCH(1,1) com dist. Normal ajustado para a série da Vale

Standardized Residuals Tests:				
			Statistic	p-Value
Jarque-Bera Test	R	Chi ²	68.97517	9.992007e-16
Shapiro-Wilk Test	R	W	0.9898691	4.165714e-07
Ljung-Box Test	R	Q(10)	11.57224	0.3147043
Ljung-Box Test	R	Q(15)	13.62159	0.5544003
Ljung-Box Test	R	Q(20)	18.44491	0.5581252
Ljung-Box Test	R ²	Q(10)	10.03173	0.4377136
Ljung-Box Test	R ²	Q(15)	13.56053	0.5590858
Ljung-Box Test	R ²	Q(20)	17.12893	0.6445867
LM Arch Test	R	TR ²	10.91874	0.5358954

Information Criterion Statistics:				
AIC	BIC	SIC	HQIC	
4.668512	4.686118	4.668488	4.675159	

Description:
Sat Jan 02 19:09:48 2010 by user: Acer

A.14 Ajuste do modelo GARCH(1,1) com dist. t-Student para a série da Vale

```
> summary(obj)

Title:
GARCH Modelling

Call:
garchFit(formula = ~garch(1, 1), data = dif, cond.dist = "dstd",
include.shape = FALSE)

Mean and Variance Equation:
~arma(0, 0) + ~garch(1, 1)

Conditional Distribution:
dstd

Coefficient(s):
      mu      omega    alpha1    beta1
1.62711e-03  1.66038e-05  1.06109e-01  8.95747e-01

Error Analysis:
      Estimate Std. Error t value Pr(>|t|)
mu      1.627e-03  5.929e-04   2.744 0.006061 **
omega   1.660e-05  9.147e-06   1.815 0.069477 .
alpha1  1.061e-01  2.855e-02   3.717 0.000202 ***
beta1   8.957e-01  2.941e-02  30.453 < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Log Likelihood:
-2677.952    normalized: -2.336782
```

A.15 Teste de Normalidade e Ljung-Box referente ao modelo GARCH(1,1) com dist. t-Student ajustado para a série da Vale

```

Standadized Residuals Tests:

      Jarque-Bera Test   R   Chi^2   70.0871   5.551115e-16
      Shapiro-Wilk Test  R   W       0.9899055  4.361033e-07
      Ljung-Box Test     R   Q(10)   11.19732   0.3423534
      Ljung-Box Test     R   Q(15)   13.32340   0.5773344
      Ljung-Box Test     R   Q(20)   17.90588   0.5936084
      Ljung-Box Test     R^2  Q(10)   9.659225   0.4708819
      Ljung-Box Test     R^2  Q(15)   13.06316   0.5974194
      Ljung-Box Test     R^2  Q(20)   16.78987   0.6665729
      LM Arch Test       R   TR^2    10.21218   0.5973524

Information Criterion Statistics:
      AIC      BIC      SIC      HQIC
4.680545 4.698150 4.680520 4.687192

Description:
Sat Jan 02 20:43:46 2010 by user: Acer

```

A.16 Ajuste do modelo ARMA para a série de $\alpha(t)$ referente à série da Vale.

```

> summary(Alpha)

Call:
arma(x = dif1, lag = list(ar = c(5, 10, 15, 17)), include.intercept = FALSE)

Model:
ARMA(17,0)

Residuals:
      Min       1Q   Median       3Q      Max
-0.074320 -0.001621 -0.000303  0.001259  0.071969

Coefficient(s):
      Estimate Std. Error t value Pr(>|t|)
ar5   -0.18346   0.03388   -5.414 6.15e-08 ***
ar10  -0.14949   0.03409   -4.385 1.16e-05 ***
ar15   0.09687   0.03384    2.863 0.004199 **
ar17  -0.12622   0.03299   -3.826 0.000130 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Fit:
sigma^2 estimated as 5.285e-05, Conditional Sum-of-Squares = 0.04, AIC = -5922.55

```

A.17 Teste de Box-Pierce para a série de $\alpha(t)$ referente à série da Vale.

```

> Box.test(res_Alpha)

      Box-Pierce test

data:  res_Alpha
X-squared = 0.0923, df = 1, p-value = 0.7613

```

A.18 Teste de Normalidade para a série de $\alpha(t)$ referente à série da Vale.

```
> jarque.bera.test(res_Alpha_norm)

      Jarque Bera Test

data:  res_Alpha_norm
X-squared = 74576.92, df = 2, p-value < 2.2e-16

> shapiro.test(res_Alpha_norm)

      Shapiro-Wilk normality test

data:  res_Alpha_norm
W = 0.5571, p-value < 2.2e-16
```

A.19 Ajuste do modelo ARMA para a série de $\beta(t)$ referente à série da Vale.

```
> summary(Beta)

Call:
arma(x = dif1, lag = list(ar = c(5, 7, 10, 15, 17, 22)), include.intercept = FALSE)

Model:
ARMA(22,0)

Residuals:
      Min       1Q   Median       3Q      Max
-0.2358024 -0.0033924 -0.0001347  0.0034718  0.2232151

Coefficient(s):
      Estimate Std. Error t value Pr(>|t|)
ar5    -0.19859   0.03401   -5.840 5.23e-09 ***
ar7     0.10204   0.03317    3.076 0.002097 **
ar10   -0.12474   0.03405   -3.664 0.000248 ***
ar15    0.09938   0.03370    2.949 0.003189 **
ar17   -0.14118   0.03352   -4.212 2.53e-05 ***
ar22   -0.10414   0.03382   -3.079 0.002076 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Fit:
sigma^2 estimated as 0.0004634, Conditional Sum-of-Squares = 0.38, AIC = -4081.76
```

A.20 Teste de Box-Pierce para a série de $\beta(t)$ referente à série da Vale.

```
> Box.test(res_beta)

      Box-Pierce test

data:  res_beta
X-squared = 0.0942, df = 1, p-value = 0.7588
```

A.21 Teste de Normalidade para a série de $\beta(t)$ referente à série da Vale.

```
> jarque.bera.test(res_beta_norm)

      Jarque Bera Test

data:  res_beta_norm
X-squared = 114699.1, df = 2, p-value < 2.2e-16

> shapiro.test(res_beta_norm)

      Shapiro-Wilk normality test

data:  res_beta_norm
W = 0.5203, p-value < 2.2e-16
```

A.22 Ajuste do modelo ARMA para a série de $\mu(t)$ referente à série da Vale.

```
> summary(MI)

Call:
arma(x = dif1, lag = list(ar = c(1)))

Model:
ARMA(1,0)

Residuals:
      Min       1Q   Median       3Q      Max
-7.269e-04 -6.639e-05  4.147e-06  6.382e-05  3.478e-04

Coefficient(s):
              Estimate Std. Error t value Pr(>|t|)
ar1          5.676e-02  3.435e-02   1.652  0.0984 .
intercept -4.350e-06  3.450e-06  -1.261  0.2073
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Fit:
sigma^2 estimated as 1.006e-08, Conditional Sum-of-Squares = 0, AIC = -13174.18
```

A.23 Ajuste do modelo ARMA para a série de $\alpha_0(t)$ referente à série da Vale.

```
> summary(Omega)

Call:
arma(x = dif1, lag = list(ar = c(5, 7, 15, 17)), include.intercept = FALSE)

Model:
ARMA(17,0)

Residuals:
      Min          1Q      Median          3Q         Max
-6.781e-05 -1.104e-06  1.080e-07  1.401e-06  6.681e-05

Coefficient(s):
      Estimate Std. Error t value Pr(>|t|)
ar5   -0.13655   0.03346   -4.081 4.49e-05 ***
ar7    0.10756   0.03341    3.219 0.001286 **
ar15   0.08501   0.03339    2.546 0.010889 *
ar17  -0.11724   0.03339   -3.511 0.000447 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Fit:
sigma^2 estimated as 4.601e-11, Conditional Sum-of-Squares = 0, AIC = -17727.72
```

A.24 Teste de Box-Pierce para a série de $\alpha_0(t)$ referente à série da Vale.

```
> Box.test(res_Omega)

      Box-Pierce test

data:  res_Omega
X-squared = 0.4031, df = 1, p-value = 0.5255
```

A.25 Teste de Normalidade para a série de $\alpha_0(t)$ referente à série da Vale.

```
> jarque.bera.test(res_Omega_norm)

      Jarque Bera Test

data:  res_Omega_norm
X-squared = 70618.88, df = 2, p-value < 2.2e-16

> shapiro.test(res_Omega_norm)

      Shapiro-Wilk normality test

data:  res_Omega_norm
W = 0.565, p-value < 2.2e-16
```

Referências Bibliográficas

- [1] Black, F. (1976). Studies in stock price volatility changes, Proceedings of the 1976 Business Meeting of the Business and Economics Statistics Section, American Statistical Association, 177-181.
- [2] Bollerslev, T.(1986). Generalized Autoregressive Conditional Heteroskedasticity, *Journal of Econometrics*, 31, 307-327.
- [3] Bollerslev, T., Engle, R. F., Nelson, D. B., (1994). ARCH Models, in R. F. Engle and D.L. McFadden (eds.) *Handbook of econometrics*, Vol. 4, North Holland Press, Amsterdam.
- [4] Bollerslev, T.,Wooldridge, J. M. (1992). Quasi-maximum likelihood estimation and inference in dynamic models with timevarying covariances, *Econometric Reviews*, 11, 143-172.
- [5] Box, Jenkins, G.M. Reinsel, G.C. (1994). *Time Series Analysis, Forecasting and Control*, 3rd edn. Prentice Hall, New Jersey.
- [6] Box, G. E. P., Pierce, D. A. (1970). Distributions of autocorrelations in autoregressive moving average models. *Journal of The American Statistical Association*, 65, 1509-1526.
- [7] Brockwell, P. J., Davis, R.A. (1991). *Time Series: Theory and Methods*, 2nd. Springer-Verla, New York.
- [8] Cai, Z., Fan, J., Yao, Q. (2004). Functional Coefficient Regression Models for Nonlinear Time Series.
- [9] Chen, R., Tsay, R. S. (1993). Functional coefficient autoregressive models. *Journal of American Statistical Association*, 88, 298-308.
- [10] Diebold, F. X., Mariano, R. S. (1995). Comparing Predictive Accuracy. *Journal of Business and Economic Statistics* , 13, 253-263.
- [11] Engle, R. F. (1982). Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation, *Econometrica*, 50 (4), 987-1007.
- [12] Fama, E. F. (1965). The behavior of stock market prices, *Journal of Bussines*, 38, 34-105.
- [13] Fan, J., Yao, Q. (2005). *Nonlinear Time Series: Nonparametric and Parametric Methods*, Springer.
- [14] Fuller, W. A. (1996). *Introduction to Statistical Time Series*, Second Edition. John Wiley & Sons, New York.
- [15] Glosten, L. R., Jagannathan, R., Runkle, D. E. (1993). Relation Between the Expected Value and the Volatility of the Nominal Excess Return on Stocks, *Journal of Finance*, 48, 1779-1801.
- [16] Hagan, V., Ozaki, T. (1981). Modeling nonlinear vibrations using an amplitude dependent autoregressive time series model. *Biometrika*, 68, 189-196.
- [17] Hamilton, J.D. (1994). *Time Series Analysis*. Princeton University Press, New Jersey.
- [18] Harvill, J. L., Ray, B. K. (2005). Functional Coefficient Autoregressive Models for Vector Time

Series, Mississippi.

- [19] Hayashi, F. (2000). *Econometrics*. Princeton University Press, New Jersey.
- [20] He, C., Terasvirta, T. (1999a). Properties of Moments of a Family of GARCH Process, *Journal of Econometrics*, 92, 173-192.
- [21] He, C., Terasvirta, T. (1999b). Fourth Moment Structure of the GARCH(p,q). Process, *Econometrics Theory*, 15, 824-846.
- [22] Ljung, G., Box, G. E. P. (1978). On a measure of lack of fit in time series models. *Biometrika*, 65, 297-303.
- [23] Mandelbrot, B. B., (1963). The Variation of certain speculative prices, *Journal of Business*, 36, 394-419.
- [24] Mills, T. C., (1999). *The Econometric Modelling of Financial Time Series*. Second Edition. Cambridge University Press.
- [25] Mittnik, S., Rachev, S. T. (2000). *Stable Paretian Models in Finance: Some empirical and theoretical aspect*, in R. J. Adler, R. E. Feldman and M. S. Taqqu(eds.): *A Practical Guide to Heavy Tails: Statistical Techniques And Applications*, 79-110, Birkhäuser , Boston.
- [26] Mittnik, S., Rachev, S. T., Paoletta, M. S. (1998).
- [27] Morettin, P. A. (2006). *Econometria Financeira, Um Curso em Series Temporais*. Edgard Blucher, São Paulo.
- [28] Morettin, P. A., Tolo, C.M.C. (2004). *Análise de Séries Temporais*. Edgard Blucher, São Paulo.
- [29] Nelson, D. B. (1991). Conditional Heteroskedasticity in Asset Return: a New Approach, *Econometrica*, 59 (2), 347-370.
- [30] Newey, W. K., West, K. D. (1987). A Simple Positive Semidefinite Heteroskedasticity and Autocorrelations Consistent Covariance Matrix, *Econometrica*, 55, 703-708.
- [31] Pagan, A. R., Schwert, G. W. (1990). Alternative models for conditional stochastic volatility. *Journal of Econometrics*, 45, 267-290.
- [32] Peng, L., Yao, Q. (2003). Least Absolute deviation estimation for ARCH and GARCH models, *Biometrika*, 90 (4), 967-975.
- [33] Shen, H., Huang, J. Z. (2004). *Functional Coefficient Regression Models for Non-linear Time Series: A Polynomial Spline Approach*, Oxford.
- [34] Tjøstheim, D., Auestad, B. H. (1994). Nonparametric identification of nonlinear time series: projections. *Journal of American Statistical Association*, 89, 1398-1409.
- [35] Tong, H. (1990). *Non Linear Time Series Models*. Oxford: Oxford University Press.
- [36] Tsay, R. S. (2002). *Analysis of Financial Time Series*. Wiley, New York.
- [37] White, H. (1984). *Asymptotic Theory for Econometricians*. Academic Press, San Diego.
- [38] Zivot, E., Wang, J. (2003). *Modelling Financial Time Series with S-Plus*. Springer, New York.