

UNIVERSIDADE ESTADUAL DE CAMPINAS
INSTITUTO DE MATEMÁTICA, ESTATÍSTICA E COMPUTAÇÃO CIENTÍFICA
DEPARTAMENTO DE ESTATÍSTICA

**Estimação por Máxima Quase-Verossimilhança no Domínio do
Tempo de Modelos de Volatilidade Estocástica
com Memória Longa**

Rosemeire de Olanda Ferraz

Prof. Dr. Luiz Koodi Hotta

Orientador

Dissertação apresentada junto ao Departamento de Estatística do Instituto de Matemática Estatística e Computação Científica da Universidade Estadual de Campinas, para obtenção do título de Mestre em Estatística.

Campinas

2003

Resumo

O objetivo deste trabalho é a estimação por máxima quase-verossimilhança de modelos de volatilidade estocástica com memória longa. A estimação é feita no domínio do tempo usando uma representação aproximada do modelo na forma de espaço de estados. Os processos foram aproximados nas formas auto-regressivas e médias móveis para diferentes truncamentos na representação do operador de diferença fracionária. Os resultados são comparados com os estimadores usuais, no domínio da frequência, GPH (Geweke e Porter-Hudak) e GSE (*Gaussian Semiparametric Estimator*) aplicados a modelos ARFIMA.

Os métodos são aplicados às séries simuladas com tamanhos iguais a 1000 e às séries de índices da Bolsa de Valores do Estado de São Paulo, IBOVESPA, observadas em intervalos de 1 minuto e de 5 minutos, no período de 18/Dezembro/2001 à 02/Janeiro/2002.

Abstract

The aim of this work is to use of the Quasi-Maximum Likelihood estimator of Long Memory Stochastic Volatility models. The estimation by quasi-maximum likelihood is done in the time domain using an approximate state-space representation of the model. We analyse the autoregressive and moving average approximation using different maximum lag approximation. The results are compared to usual frequency domain estimators proposed by Geweke and Porter-Hudak and Robinson applied to ARFIMA models.

The methods are applied to simulated series of sample size 1000 and to the São Paulo Stock Exchange Index (IBOVESPA) series. The series was collected at one-minute and five-minute intervals, from december 18th, 2001 to January 2^{sd}, 2002.

Conteúdo

1	Introdução	1
2	Modelos com Memória Longa	5
2.1	Introdução	5
2.2	Modelos ARIMA	6
2.2.1	Modelos ARMA(p,q)	6
2.2.2	Modelos ARIMA(p,d,q)	10
2.3	Modelos ARFIMA	11
3	Modelos de Volatilidade Estocástica	19
3.1	Introdução	19
3.2	Retornos e Fatos Estilizados	20
3.3	Modelos de Volatilidade Estocástica	22
3.3.1	Propriedades do Modelo de Volatilidade Estocástica	24
3.4	Modelos de Volatilidade Estocástica com Memória Longa	27
3.4.1	Representação na Forma de Espaço de Estados	28
3.4.2	Comportamentos dos Coeficientes das Formas AR e MA	34
4	Métodos de Estimação	38
4.1	Introdução	38
4.2	Métodos Semiparamétricos	40
4.2.1	Método GPH - Geweke e Porter-Hudak	41
4.2.2	Método GSE - Estimador Semiparamétrico Gaussiano	44
4.3	Método de Estimação por Máxima Quase-Verossimilhança	46
4.3.1	Introdução	46
4.3.2	Função de Quase-Verossimilhança	47

5	Simulação	49
5.1	Introdução	49
5.2	Modelos VE-ARFIMA(0,d,0)	51
5.3	Modelos VE-ARFIMA(1,d,0)	60
5.4	Conclusão	73
6	Aplicação dos Métodos de Estimação	75
6.1	Introdução	75
6.2	Análise dos Índices IBOVESPA	76
6.3	Conclusão	90
A	Método de Estimação por Máxima Verossimilhança	92
	Bibliografia	96

Lista de Figuras

3.1	Comportamentos dos coeficientes das componentes AR e MA	34
3.2	Somas Acumuladas dos Coeficientes da Componente MA	35
3.3	Somas Acumuladas dos coeficientes das componentes AR e MA-Mod	36
5.1	Dispersão das Estimativas de d obtidas com as formas AR e MA VE-ARFIMA(0,d,0)	56
5.2	Dispersão das estimativas de d obtidas com as formas AR30, MA30, GPH e GSE - VE-ARFIMA(0,d,0)	57
5.3	<i>Box Plot</i> das Estimativas de d - VE-ARFIMA(0,d,0)	58
5.4	Dispersão das Estimativas de d obtidas com as formas AR e MA VE-ARFIMA(1,d,0)	68
5.5	Dispersão das Estimativas de d obtidas com as formas AR30, MA30, GPH e GSE - VE-ARFIMA(1,d,0)	70
5.6	<i>Box Plot</i> das Estimativas de d - VE-ARFIMA(1,d,0)	72
6.1	Séries de Índices IBOVESPA	76
6.2	Retornos absolutos das séries de índices IBOVESPA	77
6.3	Histogramas dos retornos das séries de índices IBOVESPA	78
6.4	Correlograma dos retornos de cada série	79
6.5	Correlograma do log dos retornos ao quadrado para cada série	79
6.6	Resíduos dos ajustes por ARMA(1,1), para a Série 1min e Série 5min	83
6.7	Histogramas dos resíduos dos ajustes por ARMA(1,1), para a Série 1min e Série 5min	84
6.8	Auto-correlações dos resíduos dos ajustes por ARMA(1,1), para a Série 1min e Série 5min	85
6.9	ACF-Empírica dos resíduos do ARMA(1,1) que Ajusta a série de 5min e Teórica do processo VE-ARFIMA(0,d,0) com a média das estimativas para d por QMV	88

6.10	ACF-Empírica dos resíduos do ARMA(1,1) que Ajusta a série de 5min e Teórica do processo VE-ARFIMA(0,d,0) com a média das estimativas para d por GPH e GSE .	88
6.11	Resíduos absolutos, volatilidade filtrada e suavizada - VE-ARFIMA(0,d,0)	89

Lista de Tabelas

5.1	Comparação das estimativas QMV para d - VE-ARFIMA(0,d,0)	52
5.2	Comparação das estimativas para d - VE-ARFIMA(0,d,0)	52
5.3	Comparação das estimativas de σ_{η}^2 e σ_{ν}^2 - VE-ARFIMA(0,d,0)	53
5.4	Correlações, variâncias e covariâncias das estimativas de d :VE-ARFIMA(0,d,0)	55
5.5	Comparação das estimativas por MQV de d e ϕ - VE-ARFIMA(1,d,0)	60
5.6	Comparação das estimativas de d - VE-ARFIMA(1,d,0)	61
5.7	Comparação das estimativas de σ_{η}^2 e σ_{ν}^2 - VE-ARFIMA(1,d,0)	62
5.8	Correlações, variâncias e covariâncias das estimativas de d :VE-ARFIMA(1,d,0)	65
6.1	Estatísticas descritivas dos retornos- IBOVESPA	77
6.2	Estimativas dos parâmetros para a Série 1min - IBOVESPA	80
6.3	Resultados do Teste Ljung-Box Modificado para a Série 1min - IBOVESPA	81
6.4	Estimativas dos parâmetros para a Série 5min - IBOVESPA	82
6.5	Resultados do Teste LBM para a Série 5min - IBOVESPA	82
6.6	Estimativas para os parâmetros do VE-ARFIMA(0,d,0)-Série 1min	86
6.7	Estimativas para os parâmetros do VE-ARFIMA(1,d,0)-Série 1min	86
6.8	Estimativas para os parâmetros do VE-ARFIMA(0,d,0)-Série de 5min	87
6.9	Estimativas para os parâmetros do VE-ARFIMA(1,d,0)-Série de 5min	87

Capítulo 1

Introdução

Nos estudos das séries financeiras é comum analisar o comportamento dos retornos, mais do que da própria série em si. Os retornos podem ser vistos como ganho ou perda relativa por parte dos investidores.

Durante muito tempo estudou-se o comportamento dos retornos com base na suposição de que a variabilidade neles existente seria constante, ou seja, invariante no tempo. Mas observou-se, em recentes estudos, que a variabilidade condicional (condicionada a valores passados) dos retornos, mais conhecida como volatilidade, era variante no tempo e, portanto, seria interessante modelá-la.

Muitos modelos foram propostos e os que obtiveram maiores destaques foram o modelo ARCH, proposto por Engle (1982) (ver também o trabalho de Shephard (1996)) e o modelo de volatilidade estocástica (ver por exemplo, Taylor (1994) e Ghysels, Harvey e Renault (1996)).

O modelo ARCH sofreu várias modificações, dando origem à família dos modelos XARCH, a qual pertencem os modelos GARCH, IGARCH, EGARCH, QGARCH, SWGARCH e etc.

O modelo de volatilidade estocástica também tem sido extensivamente estudado, mas menos utilizado, pelo fato de a função de verossimilhança ser mais facilmente calculada nos modelos da família XARCH. A possibilidade de colocar os modelos de volatilidade estocástica na forma de espaço de estados, aumentou o interesse em estudar a volatilidade baseando-se nestes modelos. No modelo de volatilidade estocástica, como em Harvey (1993), a log-volatilidade segue um processo AR(1); ver também Breidt *et al.* (1998).

A modelagem da volatilidade pelos modelos de volatilidade estocástica, com o uso dos processos AR(1), foi extensivamente utilizada até, recentes estudos, constatarem que em, algumas séries, a volatilidade apresentava comportamento de persistência longa.

O comportamento de memória longa em séries temporais tem despertado o interesse de muitos estudiosos desde o trabalho de Harold Edwin Hurst (1951). Hurst foi um dos primeiros a estudar esse comportamento, utilizando-se, em seus estudos, dados dos níveis anuais das águas de alguns rios, como por exemplo, o rio Nilo.

Motivados pelos estudos de Hurst, estudiosos passaram a estudar o comportamento de memória longa em séries financeiras.

A memória longa pode ser melhor visualizada através do comportamento das auto-correlações, as quais decaem hiperbolicamente, ao invés do decaimento exponencial, tradicionalmente, encontrado nos modelos estacionários da família XARCH.

Os processos mais utilizados para capturarem a memória longa nas séries são os processos ARFIMA, uma extensão natural dos processos ARIMA da metodologia de Box e Jenkins (1976). Os processos ARFIMA permitem que o parâmetro de diferenciação também assuma valores não-inteiros, isto é, fracionários.

Assim, o comportamento de memória longa na volatilidade passou a ser estudado utilizando-se os modelos de volatilidade estocástica, nos quais a log-volatilidade é modelada pelos processos ARFIMA, tais processos são denominados de VE-ARFIMA (ver Harvey (1993) e Taylor (2000)) .

Os modelos da família XARCH, mais especificamente os modelos GARCH e EGARCH, tiveram a incorporação de raízes fracionárias, originando os processos FIGARCH e FIEGARCH (ver, por exemplo, Baillie, Bollerslev e Mikkelsen (1996) e Baillie (1996)).

Alguns autores preferem explicar a dependência longa através de modelos de mudança de regime, (ver Liu (2000)), nos quais a incorporação de dependência longa é feita pela utilização de uma distribuição de cauda longa no tempo de permanência de cada regime.

Diferentemente dos processos da família XARCH, o modelo de volatilidade estocástica não possui função de verossimilhança analítica, o que dificulta a estimação dos parâmetros. Esta dificuldade foi em parte contornada pelo fato de ser possível, por meio de linearização, colocá-lo na forma de espaço de estados (FEE).

A formulação de espaço de estados permite modelar, separadamente, as várias componentes de um processo (como sazonalidade, tendência, volatilidade e etc). Com essa formulação, a função de verossimilhança, para o caso de inovações gaussianas, pode ser calculada facilmente com as saídas do Filtro de Kalman¹(FK).

No entanto, para a utilização da recursão de Kalman, é necessário satisfazer a suposição de normalidade, suposição não satisfeita pelos processos de volatilidade estocástica, pois, uma vez linearizados, estes processos passam a ter perturbações com distribuição não-gaussiana; em particular, uma distribuição log-quadrado, quando o processo de volatilidade estocástica tem perturbações com distribuição Normal. Assim, a função de verossimilhança obtida com as saídas do Filtro de Kalman é uma função de verossimilhança aproximada, ou função de quase-verossimilhança (ver Durbin e Koopman (1997) e Pérez e Ruiz (2001)).

Com a correção da função de quase-verossimilhança, utilizando-se técnicas de amostragem por importância e de simulação, pode-se obter a função de verossimilhança exata; ver, por exemplo, de Jong e Shephard (1995) e Durbin e Koopman (2001a).

Assim, é possível estimar todos os parâmetros dos modelos de volatilidade estocástica com memória longa, sem grandes dificuldades. Uma vez colocados na forma de espaço de estados, estes modelos podem ser estendidos para casos mais complexos, como os multivariados.

A estimação dos parâmetros pode ser feita utilizando-se métodos paramétricos, como o mencionado anteriormente, que utilizam a função de verossimilhança, ou por métodos semiparamétricos, como os de GPH, de Geweke e Porter-Hudak (1983) e GSE (*Gaussian Semiparametric Estimator*) de Robinson (1995b). Os métodos GPH e GSE são bastante utilizados por estarem baseados na regressão linear simples do log-periodograma e assim ser possível estimar os parâmetros, semelhante a regressão linear, pelo método de mínimos quadrados ordinários. Em Crato e Ray (2002) pode ser encontrado uma comparação do método GPH com outros métodos semiparamétricos.

¹O Filtro de Kalman é um procedimento recursivo bastante utilizado e, assim, não será detalhado neste trabalho, para isto veja Durbin e Koopman (2001a).

O enfoque desta dissertação é a estimação dos modelos de volatilidade estocástica com memória longa, mais especificamente dos processos VE-ARFIMA, utilizando-se o método da máxima quase-verossimilhança. As estimativas obtidas, para o parâmetro de diferença, são comparadas com as obtidas pelos métodos GPH e GSE.

Este trabalho encontra-se organizado da seguinte forma. No Capítulo 2, introduzimos um dos processos mais utilizados nos estudos das séries temporais, os processos ARIMA, com enfoque em algumas de suas características, apresentando também a sua extensão para incorporar a característica de memória longa, os processos ARFIMA. No Capítulo 3, iniciamos alguns conceitos de retornos e seus fatos estilizados, e em seguida apresentamos o uso dos processos ARFIMA na modelagem dos processos de volatilidade estocástica com memória longa. No Capítulo 4, são apresentados apenas os métodos de estimação que utilizamos. No Capítulo 5, temos as simulações dos processos de volatilidade estocástica com memória longa, nos quais as log-volatilidades são modeladas por processos ARFIMA, aproximados nas formas auto-regressivas (AR) e médias móveis (MA). Finalmente, no Capítulo 6, aplicamos os métodos de estimação a uma série real de índices IBOVESPA (índices da Bolsa de Valores do Estado de São Paulo) observada em intervalos de um minuto e de cinco minutos, totalizando 3471 e 696 observações, respectivamente, no período de 18/12/01 à 02/01/02.

Capítulo 2

Modelos com Memória Longa

2.1 Introdução

Um dos primeiros a se interessar pelo fenômeno da memória longa em séries temporais foi Harold Edwin Hurst (1951), em seus estudos dos níveis anuais das águas de alguns rios, como por exemplo, do rio Nilo. Hurst utilizou-se de um coeficiente para identificar o comportamento de memória longa em séries temporais, que passou a ser chamado de coeficiente de Hurst.

Mandelbrot e Taqqu (1979), motivados pelos trabalhos de Hurst, passaram a estudar o comportamento de memória longa em séries financeiras.

Os modelos com memória longa são caracterizados pela existência de dependência significativa mesmo entre as observações distanciadas por um longo período de tempo. Esta dependência pode ser melhor visualizada através da função de auto-correlação, a qual apresenta decaimento hiperbólico e da função espectral, a qual é ilimitada em frequências próximas de zero.

Constatando-se que o fenômeno da memória longa não podia ser capturado pelos modelos da classe ARIMA (utilizando-se parâmetro de diferença assumindo valores inteiros), Granger e Joyeux (1980) e Hosking (1981) propuseram os modelos ARFIMA (Auto-regressivos Fracionariamente Integrados e de Médias Móveis), uma generalização da classe ARIMA, que permite que o parâmetro de diferença assumam também valores fracionários.

Devido à importância dos modelos das classes ARIMA e ARFIMA nos estudos das séries temporais, apresentamos, inicialmente, na Seção 2.2, os modelos ARIMA, com algumas de suas principais características, e, na Seção 2.3, estendemos a classe ARIMA para incorporar o comportamento de memória longa com os modelos ARFIMA, também enfatizando suas principais características.

2.2 Modelos ARIMA

A metodologia dos modelos ARIMA (Auto-regressivos integrados e de médias móveis) foi desenvolvida por Box e Jenkins (1976) e é a mais utilizada para a análise das séries temporais. Esta metodologia abrange os modelos que necessitam de diferenciação inteira para serem estacionarizados, sendo um dos mais simples o modelo ARMA(p,q), para o qual nenhuma diferenciação é necessária.

2.2.1 Modelos ARMA(p,q)

Para o melhor entendimento da definição dos processos estacionários ARMA (Auto-regressivos e de médias móveis), segue, primeiramente, a definição de um processo ruído branco, a partir do qual os processos ARMA são definidos.

Definição 2.2.1. O processo $\{\epsilon_t\}$ é um processo ruído branco com média 0 e variância σ_ϵ^2 , ou seja, $\{\epsilon_t\} \sim RB(0, \sigma_\epsilon^2)$ se, e somente se, $\{\epsilon_t\}$ tem função de auto-covariância dada por

$$\gamma_\epsilon(k) = \text{Cov}(\epsilon_t, \epsilon_{t-k}) = \begin{cases} \sigma_\epsilon^2 & , \quad k = 0 \\ 0 & , \quad k \neq 0 \end{cases} ,$$

ou seja, ϵ_t são não-correlacionados e, quando satisfeita a suposição de normalidade, também tem-se a independência entre dois tempos quaisquer.

A densidade espectral do processo ruído branco $\{\epsilon_t\}$ é dada por

$$\begin{aligned} f_\epsilon(\omega) &= \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \gamma(k) e^{-i\omega k} \\ &= \frac{\sigma_\epsilon^2}{2\pi}, \quad -\pi \leq \omega \leq \pi. \end{aligned} \quad (2.1)$$

Definição 2.2.2. O processo $\{y_t\}$ para $\{t = 0, \pm 1, \dots\}$ é um processo auto-regressivo e de médias móveis de ordens p e q , ARMA(p, q), se puder ser escrito como

$$y_t - \mu = \phi_1(y_{t-1} - \mu) + \dots + \phi_p(y_{t-p} - \mu) + \epsilon_t + \theta_1\epsilon_{t-1} + \dots + \theta_q\epsilon_{t-q},$$

em que $\{\epsilon_t\} \sim RB(0, \sigma_\epsilon^2)$.

Sem perda de generalidade, a partir daqui, utilizaremos $\mu = 0$ e assim, de forma abreviada, temos o processo ARMA(p, q) dado por

$$\Phi(B)y_t = \Theta(B)\epsilon_t, \quad (2.2)$$

em que $\Phi(B) = 1 - \sum_{j=1}^p \phi_j B^j$ é o polinômio auto-regressivo de ordem p , $\Theta(B) = 1 + \sum_{j=1}^q \theta_j B^j$ é o polinômio de médias móveis de ordem q , B é o operador de defasagem, ou seja, $By_t = y_{t-1}$, assim, $B^j y_t = y_{t-j}$ e $\{\epsilon_t\}$ é um processo ruído branco com média 0 e variância σ_ϵ^2 . Os parâmetros podem ser representados por um único vetor $\beta = (\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q, \sigma_y^2, \sigma_\epsilon^2)$.

Algumas restrições sobre os parâmetros do processo ARMA são necessárias para garantir propriedades de estacionariedade e invertibilidade. O processo ARMA(p, q) é estacionário e invertível se as raízes de $\Phi(B) = 0$ e de $\Theta(B) = 0$, respectivamente, estiverem fora do círculo unitário. Uma suposição é que os polinômios $\Phi(B)$ e $\Theta(B)$ não tenham raízes em comum, para que não haja problema de identificabilidade.

O processo em (2.2) tem uma representação MA (médias móveis) da seguinte forma

$$y_t = \frac{\Theta(B)}{\Phi(B)} \epsilon_t = \Psi(B) \epsilon_t = \sum_{j=0}^{\infty} \psi_j \epsilon_{t-j} \quad , \quad (2.3)$$

em que $\Psi(B) = \sum_{j=0}^{\infty} \psi_j B^j = (\psi_0 + \psi_1 B + \psi_2 B^2 + \dots)$, para $\psi_0 = 1$.

Dada a estacionariedade, tem-se que

$$\sum_{j=0}^{\infty} |\psi_j| < \infty$$

e

$$\sum_{j=0}^{\infty} \psi_j^2 < \infty$$

A segunda condição é necessária para garantir que a variância do processo seja finita.

De forma análoga, o processo também tem uma representação AR (auto-regressiva) dada por¹

$$\frac{\Phi(B)}{\Theta(B)} y_t = \Pi(B) y_t = \sum_{j=0}^{\infty} \pi_j y_{t-j} = \epsilon_t \quad ,$$

com a condição de estacionariedade e que $\sum_{j=0}^{\infty} |\pi_j| < \infty$.

Os comportamentos das séries temporais podem ser visualizados através dos gráficos das funções de auto-correlação, auto-covariância e densidade espectral. Seguem as formas destas funções para os processos ARMA(p,q).

¹Para facilidade de notação, sempre consideraremos $\Pi(B)$ como sendo um polinômio auto-regressivo e $\Psi(B)$ como sendo um polinômio de médias móveis, para quaisquer representações AR e MA, respectivamente.

Função de Auto-covariância

O processo ARMA(p,q) tem função de auto-covariância dada por

$$\gamma(k) = \phi_1 \gamma(k-1) + \dots + \phi_p \gamma(k-p) + g_k(\beta') \quad , \quad (2.4)$$

em que

$$g_k(\beta') = \begin{cases} \theta_1 \gamma_{y\epsilon}(k-1) + \dots + \theta_q \gamma_{y\epsilon}(k-q) & , \quad k < (q+1) \\ 0 & , \quad k \geq (q+1) \end{cases} \quad ,$$

para $\gamma_{y\epsilon}(k) = E(y_{t-k}\epsilon_t)^2$ e $\beta' = (\theta_1, \dots, \theta_q)$, um vetor com os coeficientes da parte médias móveis.

Função de Auto-correlação

A função de auto-correlação de um processo ARMA(p,q), definida por $\rho(k) = \gamma(k)/\gamma(0)$, é geometricamente limitada, isto é,

$$|\rho(k)| \leq C r^{|k|} \quad , \quad k = 1, 2, \dots \quad , \quad (2.5)$$

para $C > 0$ e $0 < r < 1$. Por isso, o processo ARMA é considerado, por alguns autores, como um processo de memória curta³.

²Lembrar que ϵ_t e, conseqüentemente, y_t , têm médias iguais a zero.

³Não existe uma definição única de memória curta (ou longa) mas, de forma não rigorosa, pode-se dizer que em um processo de memória curta não existe dependência significativa entre as observações distantes, enquanto que no processo de memória longa existe.

Uma característica da função de auto-correlação é que ela satisfaz uma equação de diferenças homogênea de ordem $|p - q|$.

Função de Densidade Espectral

A função de densidade espectral de um processo ARMA(p,q) é dada por

$$f_y(\omega) = \frac{\sigma_\epsilon^2 |\Theta(e^{-i\omega})|^2}{2\pi |\Phi(e^{-i\omega})|^2}, \quad -\pi \leq \omega \leq \pi, \quad (2.6)$$

em que $\Theta(e^{-i\omega}) = 1 - \sum_{j=1}^q \theta_j e^{-ij\omega}$ e $\Phi(e^{-i\omega}) = 1 - \sum_{j=1}^p \phi_j e^{-ij\omega}$.

2.2.2 Modelos ARIMA(p,d,q)

Os processos ARIMA são aqueles que, diferenciados d vezes (sendo d um número inteiro), reduzem-se a processos ARMA. Os processos ARIMA são também chamados de processos integrados de ordem d e abreviados por $I(d)$. Diferentemente dos processos ARMA(p,q), os processos ARIMA(p,d,q) não são estacionários, pois d raízes da equação característica, $\Phi(B) = 0$, caem sobre o círculo unitário, ou seja, são iguais a um, por isto a necessidade de d diferenças para torná-los estacionários. Assim, os processos ARIMA são processos não-estacionários homogêneos.

Definição 2.2.3. Um processo $\{y_t\}$ auto-regressivo integrado ⁴ e de médias móveis ARIMA(p,d,q), com o parâmetro d assumindo somente valores inteiros, é dado por

$$\Phi(B)(1 - B)^d y_t = \Theta(B)\epsilon_t, \quad (2.7)$$

em que $\Phi(B)$ e $\Theta(B)$ são os polinômios auto-regressivos e de médias móveis, definidos na expressão (2.2).

⁴O nome integrado vem do fato de que, na análise dos processos ARIMA, é comum utilizar a forma inversa da diferenciação, ou seja, $(1 - B)^{-d}$, referida como integração.

Similarmente aos processos ARMA, os processos ARIMA também têm representações nas formas AR e MA. Na forma AR tem-se que

$$\Pi(B) = \frac{\Phi(B)(1-B)^d}{\Theta(B)} \quad ,$$

e na forma MA,

$$\Psi(B) = \frac{\Theta(B)(1-B)^{-d}}{\Phi(B)} \quad .$$

2.3 Modelos ARFIMA

Nesta seção, estenderemos a classe de processos ARIMA para incorporar o comportamento de memória longa (ver Hosking (1981) e Granger e Joyeux (1980)). Para tanto, propõe-se que o parâmetro de diferença d dos processos ARIMA também assuma valores fracionários. Tais processos são denominados de ARFIMA - Auto-regressivos fracionariamente integrados e de médias móveis e servem para modelar séries, com características de memória longa, de diversas áreas, como: hidrologia e economia (ver, por exemplo, Hurst (1951), Lawrance e Kottegoda (1977), Hipel e McLeod (1978b) e Granger (1980)).

Os processos da classe ARFIMA têm comportamentos de memória longa, ou persistência, para $d \in (0, 0.5)$, comportamentos de memória intermediária, ou anti-persistência, para $d \in (-0.5, 0)$ e podem ser reduzidos a processos ARMA, para $d = 0$, (ver os trabalhos de Baillie (1996) e Ding e Granger (1996)).

Como a característica de memória longa, ou curta, é definida pelo valor do parâmetro d , d recebe o nome de parâmetro de memória. Granger (1980), Granger e Joyeux (1980) e Hosking (1981) demonstraram que os processos ARFIMA(p,d,q), para $d \in (-0.5, 0.5)$, são estacionários e invertíveis.

Definição 2.3.1. Um processo $\{y_t\}$ é um processo ARFIMA(p,d,q) , auto-regressivo fracionariamente integrado e de médias móveis de ordens p,d e q, com $d \in (-0.5, 0.5)$, se satisfaz à equação

$$\Phi(B)(1-B)^d y_t = \Theta(B)\epsilon_t \quad , \quad (2.8)$$

em que $\{\epsilon_t\}$, $\Phi(B)$ e $\Theta(B)$ são definidos como em (2.2). A restrição sobre o parâmetro d , para assumir valores no intervalo $(-0.5, 0.5)$, é necessária para manter a estacionariedade ($d < 0.5$) e invertibilidade ($d > -0.5$).

O operador de diferença fracionária $(1-B)^d$, ou seja, para o caso em que d assume valores não-inteiros, é dado por⁵

$$(1-B)^d = \sum_{j=0}^{\infty} \frac{\Gamma(j-d)}{\Gamma(j+1)\Gamma(-d)} B^j = \sum_{j=0}^{\infty} \pi_j B^j \quad , \quad (2.9)$$

em que $\Gamma(\cdot)$ é a função Gama. O coeficiente π_j pode ser, recursivamente, obtido por $\pi_j = \pi_{j-1} (j-d-1)/j$, considerando-se $\pi_0 = 1$.

O operador de diferença fracionária também pode ser escrito como a seguinte expansão binomial

$$(1-B)^d = \sum_{k=0}^{\infty} \binom{d}{k} (-B)^k = 1 - dB + \frac{d(d-1)}{2!} B^2 - \frac{d(d-1)(d-2)}{3!} B^3 + \dots \quad .$$

⁵A forma $(1-B)^{-d}$ é obtida substituindo-se no numerador $\Gamma(j-d)$ por $\Gamma(j+d)$ e no denominador $\Gamma(-d)$ por $\Gamma(d)$.

Função de Auto-Correlação e Densidade Espectral

A identificação dos processos como de memória longa, ou curta, pode ser feita com a análise do comportamento da função de auto-correlação e da densidade espectral.

De forma geral, um processo fracamente estacionário⁶ é dito ser de memória longa se tiver função de auto-correlação com decaimento hiperbólico, isto é,

$$\rho(k) \sim C_1 k^{2d-1} \quad , \quad \text{quando } k \rightarrow \infty \quad , \quad (2.10)$$

para algum $C_1 \neq 0$ e $d \neq 0$.

O parâmetro d controla a taxa de decaimento das auto-correlações, as quais não são absolutamente somáveis para $d > 0$. Os processos com memória curta, como os ARMA, têm auto-correlações com decaimento exponencial.

Alternativamente, pode-se dizer que um processo tem memória longa se tiver espectro , $f(\omega)$, com comportamento assintótico da forma

$$f(\omega) \sim C_2 |\omega|^{-2d} \quad , \quad \text{quando } \omega \rightarrow 0 \quad , \quad (2.11)$$

em que $d \neq 0$ e $C_2 > 0$.

Para $d > 0$, o espectro diverge em zero, ou seja, $f(\omega) \uparrow \infty$, quando $\omega \rightarrow 0$. E, assim, a função de densidade espectral é ilimitada para frequências próximas de zero.

⁶Ver definição de estacionariedade fraca em Harvey(1993).

Considerando-se o processo ARFIMA(p,d,q) dado pela expressão (2.8), observamos que a mesma equação pode ser interpretada de várias maneiras. Por exemplo, considerando-se que

$$(1 - B)^d y_t = x_t \quad ,$$

então x_t é um processo ARMA definido por

$$x_t = \Phi^{-1}(B)\Theta(B)\epsilon_t \quad .$$

Ou seja, ao passar o processo $\{y_t\}$ por um operador de diferença fracionária (ou filtro linear infinito), obtém-se um processo ARMA.

Mas também podemos escrever a expressão (2.8) considerando-se que

$$y_t = \Phi^{-1}(B)\Theta(B)\tilde{x}_t \quad ,$$

em que $\tilde{x}_t$ é um processo ARFIMA(0,d,0) dado por

$$\tilde{x}_t = (1 - B)^{-d}\epsilon_t \quad .$$

Assim, a densidade espectral do processo ARFIMA(p,d,q) pode ser escrita em termos da densidade espectral do processo ARMA(p,q), por

$$f_y(\omega) = |1 - e^{i\omega}|^{-2d} f_x(\omega) \quad , \quad -\pi \leq \omega \leq \pi, \quad (2.12)$$

em que $f_x(\omega)$ é a densidade espectral do processo ARMA(p,q) dada pela expressão (2.6). E, observando-se que $|1 - e^{i\omega}| = 2\sin(\omega/2)$ e que $\lim_{\omega \rightarrow 0} \omega(2\sin(\omega/2))^{-1} = 1$, então a densidade espectral pode ser aproximada por

$$f_y(\omega) \sim |\omega|^{-2d} f_x(0) \quad , \quad -\pi \leq \omega \leq \pi.$$

Exemplo 1: Processo ARFIMA(0,d,0)

O caso mais simples de processo com memória longa é o processo ARFIMA(0,d,0), o qual tem propriedades análogas às do processo ruído branco fracionalmente integrado (*Fractionally Integrated White Noise*, FIWN), escrito por

$$(1 - B)^d y_t = \epsilon_t \quad , \quad \{\epsilon_t\} \sim RB(0, \sigma_\epsilon^2) \quad , \quad (2.13)$$

em que $(1 - B)^d$ é dado pela expressão (2.9).

O processo FIWN pode ser escrito, na forma MA, por

$$y_t = (1 - B)^{-d} \epsilon_t = \sum_{j=0}^{\infty} \psi_j \epsilon_{t-j} \quad ,$$

em que $\psi_j = \Gamma(j + d)/\Gamma(j + 1)\Gamma(d)$ ou, na forma recursiva,

$$\psi_j = \frac{j + d - 1}{j} \psi_{j-1} \quad , \quad j > 0,$$

considerando-se que $\psi_0 = 1$.

O processo FIWN possui variância dada por

$$\gamma_y(0) = \text{Var}(y_t) = \sigma_\epsilon^2 \frac{\Gamma(1 - 2d)}{[\Gamma(1 - d)]^2}$$

Quando $d \geq 0.5$, a variância de y_t é infinita e o processo é não-estacionário; por isto, para o caso que estudamos, é importante restringir os valores de d a $(-0.5, 0.5)$.

O processo FIWN é covariância estacionário e os seguintes resultados valem⁷:

As auto-correlações do processo FIWN são dadas por

$$\rho_y(1) = \frac{d}{1-d}, \quad \rho_y(2) = \frac{d(d+1)}{(1-d)(2-d)}, \quad \rho_y(3) = \frac{d(d+1)(d+2)}{(1-d)(2-d)(3-d)}, \dots, \quad (2.14)$$

ou em termos da função Gama,

$$\rho_y(k) = \frac{\Gamma(1-d)\Gamma(k+d)}{\Gamma(d)\Gamma(k+1-d)}. \quad (2.15)$$

Pode-se obter $\rho_y(k)$, recursivamente, por

$$\rho_y(k) = \frac{d+k-1}{k-d} \rho_y(k-1),$$

com $\rho_y(0) = 1$. Ou, assintoticamente, por

$$\frac{\rho_y(k)}{k^{2d-1}} \rightarrow \frac{\Gamma(1-d)}{\Gamma(d)}, \quad \text{quando } k \rightarrow \infty. \quad (2.16)$$

A função de densidade espectral do processo FIWN é dada por

$$f_y(\omega) = \frac{\sigma_\epsilon^2}{2\pi} |1 - e^{-i\omega}|^{-2d} = \frac{\sigma_\epsilon^2}{2\pi} \left[2 \sin\left(\frac{\omega}{2}\right) \right]^{-2d}, \quad \omega > 0. \quad (2.17)$$

⁷Algumas propriedades matemáticas de FIWN são sumarizadas em Baillie (1996) e Taylor(2000).

Visto que $\sin \omega \sim \omega$, quando $\omega \rightarrow 0$, a função de densidade espectral pode ser aproximada por

$$f_y(\omega) \sim \frac{\sigma_\epsilon^2}{2\pi} \omega^{-2d} \quad , \quad \omega \rightarrow 0 \quad (2.18)$$

e $f_y(0)$ é finito se, e somente se, $d \leq 0$.

Aplicando-se a fórmula de Stirling

$$\Gamma(x) \sim \sqrt{2\pi} e^{-x+1} (x-1)^{x-1/2} \quad , \quad x \rightarrow \infty \quad ,$$

os coeficientes π_j , ψ_j e $\rho(j)$ podem ser aproximados, quando $j \rightarrow \infty$, por

$$\begin{aligned} \pi_j &\sim \frac{j^{-d-1}}{\Gamma(-d)} \quad , \\ \psi_j &\sim \frac{j^{d-1}}{\Gamma(d)} \quad , \\ \rho(j) &\sim j^{2d-1} \frac{\Gamma(1-d)}{\Gamma(d)} \quad . \end{aligned}$$

Exemplo 2: Processo ARFIMA(1,d,0)

O processo ARFIMA(1,d,0) é dado por

$$(1 - \phi B)(1 - B)^d y_t = \epsilon_t \quad , \quad (2.19)$$

em que $|\phi| < 1$, $\epsilon_t \sim RB(0, \sigma_\epsilon^2)$ e $(1 - B)^d$ é dado pela expressão (2.9).

O processo em (2.19) pode ser escrito na forma MA por

$$y_t = (1 - \phi B)^{-1}(1 - B)^{-d} \epsilon_t \quad ,$$

em que $(1 - \phi B)^{-1} = 1 + \phi B + (\phi B)^2 + (\phi B)^3 + \dots$.

As propriedades dinâmicas dos processos ARFIMA(1,d,0) podem ser obtidas em termos dos processos ARFIMA(0,d,0), pois considerando-se que $\tilde{x}_t$ é um processo ARFIMA(0,d,0), então $y_t = (1 - \phi B)^{-1} \tilde{x}_t$ é um processo ARFIMA(1,d,0).

Capítulo 3

Modelos de Volatilidade Estocástica

3.1 Introdução

Muitas séries temporais, em especial as séries financeiras, apresentam variâncias condicionais que mudam ao longo do tempo. A variância condicional dos retornos, mais conhecida como volatilidade, tornou-se alvo de recentes estudos e, com isto, uma grande variedade de modelos tem sido proposta para sua modelagem.

Um dos primeiros modelos para tal intuito foi o modelo ARCH, sugerido por Engle(1982), no qual modela-se a volatilidade em uma forma auto-regressiva, e a partir do qual muitos outros modelos foram sugeridos, gerando assim a família XARCH (ARCH, FIGARCH, FIEGARCH e etc). Estas abordagens também podem ser encontradas nos trabalhos de Baillie *et al.*(1996), Robinson (1991), Nelson(1991), entre outros.

Uma proposta alternativa ao modelo ARCH é o modelo de volatilidade estocástica VE (ver, por exemplo, trabalhos de Shephard (1996) e Harvey (1993)). Uma importante característica deste modelo é que pode ser colocado na forma de espaço de estados, uma abordagem que facilita a estimação dos parâmetros e que explicaremos na Seção 3.4.1 .

Nos processos VE, a volatilidade segue algum processo estocástico latente, pois ela é uma variável não-observável, sendo quase sempre utilizado o processo auto-regressivo de ordem 1, AR(1), estacionário.

No caso especial da volatilidade ter comportamento de memória longa, uma sugestão para o processo estocástico latente é a utilização de processos da classe ARFIMA, sendo essa abordagem o enfoque desta dissertação¹.

Para melhor entendimento do que apresentaremos, temos, na Seção 3.2, a descrição de alguns comportamentos tipicamente encontrados em séries de retornos, conhecidos como fatos estilizados. Na Seção 3.3, enfocamos os modelos de volatilidade estocástica com algumas de suas principais propriedades. Finalmente, na Seção 3.4, apresentamos os processos de volatilidade estocástica, com comportamentos de memória longa, assim como sua estrutura na forma de espaço de estados.

3.2 Retornos e Fatos Estilizados

Em geral, nos estudos das séries financeiras, modelam-se os retornos mais do que as próprias séries em si, isto porque os investidores estão mais interessados em suas perdas e ganhos relativos.

Existem vários tipos de retornos, entre os quais temos: o retorno simples e o retorno composto. O mais utilizado é o retorno composto dado por

$$y_t = \log \frac{p_t}{p_{t-1}} = \log p_t - \log p_{t-1} \quad ,$$

em que $\log(\cdot)$ é o logaritmo neperiano e $\{p_t\}$ é o preço de algum ativo financeiro.

Os retornos são livres de escalas e possuem propriedades estatísticas interessantes, como a estacionariedade e ergodicidade.

¹Nelson(1991) propõe o modelo FIEGARCH para a modelagem da volatilidade com comportamento de memória longa. Em Baillie (1996), a modelagem da volatilidade é feita utilizando-se modelos FIGARCH.

Nos estudos dos retornos, alguns comportamentos, por serem bastante encontrados, passaram a ser referidos como fatos estilizados, tais como:

- A distribuição dos retornos é leptocúrtica (pois apresenta, usualmente, curtose maior do que a de uma distribuição Normal, ou seja, maior do que 3). No entanto, as caudas da distribuição dos retornos, geralmente, não são tão pesadas, de forma que pode-se garantir a existência de, pelo menos, até o quarto momento;
- A função de auto-correlação dos retornos ao quadrado apresenta, em geral, comportamento similar a de um modelo ARMA(1,1): primeira auto-correlação pequena (mas estatisticamente significativa) e as demais decaindo lentamente;
- A existência de conglomerados de volatilidade. O conglomerado de volatilidade (*Volatility Clustering*) foi primeiro observado por Mandelbrot (1963) e caracterizou-se em séries financeiras por períodos de baixas frequências seguidos por períodos de baixas frequências e o mesmo ocorrendo para períodos de altas frequências. Como Mandelbrot afirmou: "...*large changes tend to be followed by large changes - of either sign - and small changes by small changes...*".
- Os retornos em si não são correlacionados, ou apresentam correlações desprezíveis, enquanto que os retornos ao quadrado e retornos absolutos apresentam correlações não desprezíveis.
- O comportamento de memória longa, ou persistência, na volatilidade. Apresenta-se decaimento lento das auto-correlações dos retornos ao quadrado e dos retornos absolutos.

Dos fatos estilizados apresentados, o comportamento de memória longa na volatilidade foi o que mais ganhou interesse pelos estudiosos no sentido de explicarem esse comportamento por algum processo estocástico.

Para esse intuito, os modelos de volatilidade estocástica são bastante utilizados na literatura e, por ser o enfoque desta dissertação, será detalhado na próxima seção.

3.3 Modelos de Volatilidade Estocástica

A forma mais simples de modelar a volatilidade é considerar que o retorno pode ser escrito como o produto de um processo ruído branco ϵ_t com variância igual a 1, por um processo independente σ_t , isto é, como

$$y_t = \sigma_t \epsilon_t \quad , \quad \epsilon_t \sim RB(0, 1) \quad , \quad (3.1)$$

como a variância de y_t condicionada a σ_t é igual a σ_t^2 , σ_t é a volatilidade (condicional) dos retornos. Em geral, assume-se que as perturbações ϵ_t têm distribuição gaussiana ou t de Student padronizada.

No modelo de volatilidade estocástica, σ_t pode ser escrito da forma

$$\sigma_t = \sigma \exp(h_t/2) \quad , \quad (3.2)$$

em que $\{h_t\}$ segue algum processo estocástico e $\sigma > 0$.

Um exemplo simples de modelo de volatilidade estocástica é obtido quando considera-se que $\{h_t\}$ segue um processo auto-regressivo de ordem 1, com perturbações gaussianas,

$$h_t = \phi h_{t-1} + \eta_t \quad , \quad |\phi| < 1 \quad , \quad (3.3)$$

em que $\eta_t \sim N(0, \sigma_\eta^2)$ e $\epsilon_t \sim N(0, 1)$. Sendo que η_t e ϵ_t são serialmente não-correlacionados e, portanto, independentes.

A análise do modelo de volatilidade estocástica torna-se bastante simplificada quando transformado, como descrevemos a seguir.

Se aplicarmos o quadrado e logaritmo neperiano em (3.1), temos

$$\log y_t^2 = \log \sigma_t^2 + \log \epsilon_t^2 \quad , \quad (3.4)$$

e com a aplicação feita em (3.1) e, conseqüentemente, em (3.2), temos

$$x_t = \gamma + h_t + \nu_t, \quad (3.5)$$

considerando-se que $x_t = \log y_t^2$, $\gamma = \log \sigma^2 + E(\log \epsilon_t^2)$ e $\nu_t = \log \epsilon_t^2 - E(\log \epsilon_t^2)$, e, para $\epsilon_t \sim N(0, 1)$, $E(\log \epsilon_t^2) = -1.27$ e $\sigma_\nu^2 = \pi^2/2$. Assim, as expressões (3.3) e (3.5) estão na forma linear de espaço de estados², sendo que o erro da equação (3.5) tem como distribuição a resultante do logaritmo de uma qui-quadrado com um grau de liberdade centrada em zero, denominada, por simplicidade, de log-qui-quadrado.

A expressão em (3.3) é a equação de estados (ou de transição) e a expressão em (3.5) é a equação de observação (ou de medida). Devemos ressaltar que, na equação de observação, a perturbação ν_t tem distribuição não-gaussiana.

A função de auto-covariância de $\{x_t\}$ é dada por

$$\gamma_x(k) = \text{Cov}(x_t, x_{t+k}) = \gamma_h(k) + \sigma_\nu^2 \mathbf{I}_{k=0} \quad , \quad (3.6)$$

em que

$$\mathbf{I}_{k=0} = \begin{cases} 1, & k = 0 \\ 0, & k \neq 0 \end{cases} \quad ,$$

²Mais detalhes sobre modelos na forma de espaço de estados, ver Durbin e Koopman (2001a).

e $\gamma_h(k)$ é a função de auto-covariância do processo AR(1), isto é,

$$\gamma_h(k) = \phi^k \frac{\sigma_\eta^2}{1 - \phi^2} \quad , \quad k=0,1,2,\dots \quad .$$

A função de densidade espectral de $\{x_t\}$, denotada por $f_x(\cdot)$, é dada por

$$f_x(\omega) = \frac{\sigma_\eta^2}{2\pi} \frac{1}{(1 - \phi e^{-i\omega})(1 - \phi e^{i\omega})} + \frac{\sigma_\nu^2}{2\pi} \quad , \quad -\pi \leq \omega \leq \pi \quad . \quad (3.7)$$

Assim, considerando-se, respectivamente, $f_h(\omega)$ e $f_\epsilon(\omega)$, como sendo as funções de densidade espectral dos processos AR(1) em (3.3) e ruído branco, como em (2.1), temos que $f_x(\omega) = f_h(\omega) + f_\epsilon(\omega)$ é a densidade espectral do processo $\{x_t\}$.

Propriedade análoga a dada pela expressão (3.7), para processos ARFIMA(p,d,q) com ruídos adicionados, pode ser encontrada em Crato e Ray (2002).

3.3.1 Propriedades do Modelo de Volatilidade Estocástica

Uma das vantagens do modelo de volatilidade estocástica é que suas propriedades dinâmicas podem ser facilmente obtidas. Esta facilidade vem do fato de que ϵ_t e η_t são independentes e de que σ_t^2 , em (3.1), tem distribuição lognormal. Assim, as propriedades dos modelos VE seguem das de uma distribuição lognormal. A seguir enunciamos algumas das principais propriedades do modelo de volatilidade estocástica (ver Harvey (1993)).

Segue das propriedades da distribuição lognormal que a variância de y_t é dada por

$$\text{Var}(y_t) = \sigma^2 \exp(\sigma_h^2/2) \quad , \quad (3.8)$$

em que σ_h^2 é a variância do processo latente h_t .

Assumindo-se que as perturbações ϵ_t e η_t são mutuamente independentes, a função de auto-correlação dos valores absolutos das observações elevados a uma potência c é dada por

$$\rho_c(k) = \frac{\exp(\frac{c^2}{4}\sigma_h^2\rho_h) - 1}{k_c \exp(\frac{c^2}{4}\sigma_h^2) - 1}, \quad k \geq 1, \quad (3.9)$$

em que k_c é

$$k_c = \frac{E(|y_t|^{2c})}{\{E(|y_t|^c)\}^2}, \quad (3.10)$$

e ρ_h , $k=0,1,2,\dots$, é a função de auto-correlação de h_t . Quando $c = 2$, k_c é a curtose da distribuição normal, a qual é igual a 3³.

Também pode-se escrever k_c como dado por

$$k_c = \frac{\Gamma(c+1/2)\Gamma(1/2)}{\{\Gamma(c/2+1/2)\}^2}. \quad (3.11)$$

A expressão (3.11) pode ser avaliada para outras distribuições como, por exemplo, para a distribuição t de Student e neste caso

$$k_c = \frac{\Gamma(c+1/2)\Gamma(-c+\nu/2)\Gamma(1/2)\Gamma(\nu/2)}{\{\Gamma(c/2+1/2)\Gamma(-c/2+\nu/2)\}^2}, \quad c < \nu/2. \quad (3.12)$$

Observe que ν precisa ser pelo menos igual a 5, se c for igual a 2.

³Taylor (1986,p.75) utiliza $c = 1$ e $c = 2$ e ϵ_t normalmente distribuído.

Características da Função de Auto-correlação

A função de auto-correlação $\rho_c(k)$ tem as seguintes características:

1. Se σ_h^2 for pequeno e $\rho(k)$ próximo a 1, então

$$\rho_c(k) \cong \rho_h \frac{\exp(\frac{c^2}{4}\sigma_h^2) - 1}{(k_c \exp(\frac{c^2}{4}\sigma_h^2) - 1)}, \quad k \geq 1 \quad . \quad (3.13)$$

Assim, a forma da função de auto-correlação de h_t é, aproximadamente, determinada por $\rho_c(k)$, exceto que ela é multiplicada por um fator de proporcionalidade, o qual precisa ser menor do que 1, quando k_c for maior do que 1.

2. Para a distribuição t de Student, k_c declina quando $\nu \rightarrow \infty$. Assim, $\rho_c(k)$ é máximo para a distribuição normal. Por outro lado, para uma distribuição com menor curtose que a da normal surgirão valores mais altos de $\rho_c(k)$.

Embora (3.9) dê um explícito relacionamento entre $\rho_c(k)$ e c , não parece ser possível afirmar que $\rho_c(k)$ atinja seu máximo para certos valores de c . De fato, diferentes valores de σ_h^2 levam a diferentes valores de c que maximizam $\rho_c(k)$.

A função de auto-correlação de $x_t = \log y_t^2$ é dada por

$$\rho_x(k) = \frac{\rho_h(k)}{1 + \sigma_\nu^2 / \sigma_h^2} \quad , \quad (3.14)$$

sendo $\nu_t = \log \epsilon_t^2 - E(\log \epsilon_t^2)$, e assim σ_ν^2 é a variância de $\log \epsilon_t^2$ ⁴.

⁴A expressão (3.14) é mantida mesmo para o caso em que ϵ_t e η_t são correlacionados, visto que $\log \epsilon_t^2$ e η_t são não-correlacionados se a distribuição conjunta de ϵ_t e η_t for simétrica.

O processo y_t^2 é também covariância e estritamente estacionário. Os momentos de y_t^2 são obtidos das propriedades da distribuição lognormal,

$$\begin{aligned} E[y_t^2] &= \exp(\sigma_h^2/2)\sigma^2 \\ \text{Var}(y_t^2) &= \sigma^4[1 + \text{Var}(\epsilon_t^2)\exp\{2\sigma_h^2\} - \exp\{\sigma_h^2\}] \\ \text{Cov}(y_t^2, y_{t+k}^2) &= \sigma^4[\exp\{\sigma_h^2 + \gamma_h(k)\} - \exp\{\sigma_h^2\}] \quad , \quad k \neq 0 \quad . \end{aligned} \tag{3.15}$$

O processo AR(1) é o mais simples para a modelagem da log-volatilidade. Para o caso em que a volatilidade tem comportamento de memória longa, processos como os da classe ARFIMA devem ser considerados para sua modelagem e estes processos, diferentemente do AR(1), não possuem uma forma direta de representação na forma de espaço de estados. Para os processos ARFIMA, a forma de espaço de estados é aproximada utilizando-se o truncamento do operador de diferença fracionária. Esta abordagem pode ser encontrada em Chan e Palma (1998) e será explorada nesta dissertação para a estimação dos modelos de volatilidade estocástica com memória longa, apresentados na próxima seção.

3.4 Modelos de Volatilidade Estocástica com Memória Longa

Em muitos casos torna-se mais adequado utilizar $\{h_t\}$ assumindo processos que capturem o comportamento de memória longa, ou persistência, da volatilidade, originando os processos de volatilidade estocástica com memória longa LMSV, *Long Memory Stochastic Volatility*, (ver Breidt, Crato e de Lima (1998) e Harvey (1993)). Aplicações de modelos com memória longa na volatilidade podem ser encontradas em Pérez e Ruiz (2001).

O comportamento de memória longa pode ser capturado por processos como os ARFIMA(p,d,q), apresentados no Capítulo 2. Em particular, os processos de volatilidade estocástica com comportamento de memória longa podem ter a log-volatilidade modelada por processos ARFIMA(p,d,q) e são denominados de processos VE-ARFIMA.

Os processos ARFIMA(p,d,q) são escritos em termos de um operador de diferença fracionária com soma infinita e, portanto, não possuem representação exata na forma de espaço de estados. A representação dos processos na forma de espaço de estados é adequada para a estimação dos parâmetros, apresentada na próxima subseção.

3.4.1 Representação na Forma de Espaço de Estados

Um modelo na forma de espaço de estados é escrito de forma simplificada por

$$y_t = Z_t \alpha_t + \epsilon_t \quad , \quad \epsilon_t \sim N(0, H_t) \quad (3.16)$$

$$\alpha_{t+1} = T_t \alpha_t + R_t \eta_t \quad , \quad \eta_t \sim N(0, Q_t) \quad , \quad (3.17)$$

em que Z_t , T_t , R_t são matrizes de seleção (para o caso que estudamos são invariantes no tempo e portanto suprimimos o índice de tempo); α_t é o vetor de estados do processo, ϵ_t e η_t são as perturbações das respectivas equações de observação (ou medida) e de estados (ou de transição).

O processo de volatilidade estocástica, no qual $\{h_t\}$ segue um processo ARFIMA(p,d,q), tem a seguinte representação na forma de espaço de estados,

$$x_t = \gamma + h_t + \nu_t \quad (3.18)$$

$$\Phi(B)(1-B)^d h_t = \Theta(B) \eta_t \quad , \quad (3.19)$$

cujas componentes γ , ν_t são como dadas em (3.5).

A forma reduzida de x_t é dada por

$$\Phi(B)(1-B)^d x_t = \gamma' + \Theta^*(B) \xi_t \quad , \quad (3.20)$$

em que

$$\begin{aligned} \gamma' &= \Phi(B)(1-B)^d \gamma \\ \Theta^*(B) \xi_t &= \Theta(B) \eta_t + \Phi(B)(1-B)^d \nu_t \end{aligned}$$

e $\Theta^*(B)$ é um polinômio de ordem infinita com $(p+q+1)$ parâmetros, sendo que $p+1$ parâmetros aparecem no polinômio auto-regressivo.

Pela expressão (3.20), temos o processo x_t similar a um processo ARMA(p', q'). As ordens p' e q' são infinitas, o que impossibilita escrevê-los na forma de espaço de estados. Uma alternativa é utilizar uma abordagem similar a de Chan e Palma (1998), na qual trunca-se o operador de diferença fracionária em m componentes. Para o processo escrito na forma AR, consideramos o operador $(1 - B)^d$ com truncamento em m e assim o processo h_t pode ser aproximado por um processo ARMA($p+m, q$). E, considerando o processo na forma MA, podemos aproximar o processo h_t por um processo ARMA($p, q+m$). Em ambos os casos, têm-se apenas $(p + q + 3)$ parâmetros ($d, \sigma_\eta^2, \sigma_\nu^2$ e os parâmetros dos polinômios $\Phi(B)$ e $\Theta(B)$).

Com o truncamento do operador, podemos colocar o processo de volatilidade estocástica na forma linear aproximada de espaço de estados, mas agora tem-se o problema de se determinar o valor mais adequado para m .

Apenas por simplicidade, apresentamos, com mais detalhes, a representação do processo de volatilidade estocástica na forma de espaço de estados, para o caso em que h_t segue um processo ARFIMA(0,d,0),

$$\begin{aligned} x_t &= \gamma + h_t + \nu_t \\ (1 - B)^d h_t &= \eta_t \quad , \end{aligned} \tag{3.21}$$

para $\eta_t \sim N(0, \sigma_\eta^2)$, ν_t com distribuição log-qui-quadrado e $(\gamma + h_t)$ representando o sinal do processo. As expressões, em (3.21), são as equações de observação e de estados, respectivamente.

O operador de diferença fracionária, $(1 - B)^d$, com truncamento em m componentes, é dado por

$$(1 - B)^d \cong \sum_{j=0}^m \pi_j B^j \tag{3.22}$$

em que os coeficientes π_j são obtidos, recursivamente, como em (2.9).

Assim, o processo $h_t \sim ARFIMA(0, d, 0)$ pode ser escrito na forma AR por

$$h_{t+1} \cong \sum_{j=1}^m \pi_j h_{t-j+1} + \eta_t \quad ,$$

e, portanto, temos a seguinte representação matricial para o modelo VE-ARFIMA(0,d,0) na forma de espaço de estados⁵

$$\begin{aligned} \mathbf{h}_{t+1} &= T \cdot \mathbf{h}_t + H \cdot \eta_t \\ x_t &= Z \cdot \mathbf{h}_t + G \cdot \nu_t \quad , \end{aligned}$$

em que $\mathbf{h}_{t+1}$ é o vetor de estados com dimensão truncada em m , no operador de diferença fracionária $(1 - B)^d$, dado por

$$\mathbf{h}_{t+1} = \begin{pmatrix} \gamma \\ h_{t+1} \\ \pi_2 h_t + \pi_3 h_{t-1} + \dots + \pi_m h_{t-m} \\ \vdots \\ \pi_m h_{t-m} \end{pmatrix}$$

e T , Z , H e G são matrizes de seleção. T e Z são dadas por

$$T = \begin{pmatrix} 1 & 0 & 0 & 0 & \dots & 0 \\ 0 & \pi_1 & 1 & 0 & \dots & 0 \\ 0 & \pi_2 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \pi_{m-1} & 0 & 0 & \dots & 1 \\ 0 & \pi_m & 0 & 0 & \dots & 0 \end{pmatrix}_{(m+1) \times (m+1)} \quad Z = (1 \quad 1 \quad 0 \quad 0 \quad \dots \quad 0)_{1 \times (m+1)} \quad .$$

⁵A forma matricial utilizada é adequada para a linguagem Ox, ver Koopman, Shephard e Doornik (1999).

Juntar as matrizes T e Z na forma

$$\Phi = \begin{pmatrix} T \\ Z \end{pmatrix} ,$$

representa a matriz de transição.

A seleção das perturbações é feita através das matrizes H e G , dadas por

$$H = \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \end{pmatrix}_{(m+1) \times 2} \quad \text{e} \quad G = (0 \quad 1)_{1 \times 2} .$$

De forma abreviada, podemos escrever

$$\mathbf{h}_{t+1}^* = \Phi \cdot \mathbf{h}_t + \begin{pmatrix} H & 0 \\ 0 & G \end{pmatrix} \cdot \mathbf{u}_t , \quad (3.23)$$

em que $\mathbf{u}_t = \begin{pmatrix} \eta_t \\ \nu_t \end{pmatrix}_{2 \times 1}$ é o vetor de perturbações e $\mathbf{h}_{t+1}^* = \begin{pmatrix} \mathbf{h}_{t+1} \\ x_t \end{pmatrix}$. A forma em (3.23) é apropriada em termos computacionais. Devemos salientar que, para um mesmo modelo, existem diferentes representações na forma de espaço de estados⁶.

Também utilizamos a representação MA para o processo ARFIMA(0,d,0). A representação MA foi modificada pela aplicação do operador $(1 - B)$, pois, desta forma, os coeficientes decaem mais rapidamente e, conseqüentemente, será necessário utilizar menor valor para m .

⁶Existe uma representação denominada de forma canônica, ver Kim *et al.* (1998).

Para a representação MA modificada, que denotamos por MA-Mod ⁷, o processo de volatilidade estocástica tem a seguinte representação na forma de espaço de estados

$$\begin{aligned} x_t &= \gamma + h_t + \nu_t \\ (1 - B) h_t &= (1 - B)^{1-d} \eta_t \quad . \end{aligned}$$

Considerando-se que o processo h_t é escrito como

$$h_{t+1} = h_t + \sum_{j=0}^m \pi_j^* \eta_{t-j} \quad ,$$

em que $\pi_j^* = \pi_{j-1}^*(j-2-d)/j$. Então, o vetor de estados é dado por

$$\mathbf{h}_{t+1} = \begin{pmatrix} \gamma \\ h_{t+1} \\ h_t + \pi_1^* \eta_{t-1} + \pi_2^* \eta_{t-2} + \dots + \pi_m^* \eta_{t-m} \\ \vdots \\ h_t + \pi_m^* \eta_{t-m} \end{pmatrix} \quad ,$$

as matrizes T e Z dadas por

$$T = \begin{pmatrix} 1 & 0 & 0 & 0 & \dots & 0 \\ 0 & 1 & 1 & 0 & \dots & 0 \\ 0 & 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & 0 & \dots & 0 \end{pmatrix}_{(m+2) \times (m+2)} \quad \quad Z = (1 \quad 1 \quad 0 \quad 0 \quad \dots \quad 0)_{1 \times (m+2)} \quad .$$

⁷Chan e Palma (1998) escrevem os processos ARFIMA com outra forma de representação MA, mas não para os processos VE-ARFIMA.

As matrizes H e G são dadas por

$$H = \begin{pmatrix} 0 & 0 \\ \pi_0^* & 0 \\ \pi_1^* & 0 \\ \vdots & \\ \pi_m^* & 0 \end{pmatrix}_{(m+2) \times 2} \quad G = \begin{pmatrix} 0 & 1 \end{pmatrix}_{1 \times 2} .$$

Para o modelo VE-ARFIMA(1,d,0), a representação é feita de forma análoga, considerando-se que o processo $\{h_t\} \sim ARFIMA(1, d, 0)$ é escrito na forma MA-Mod por

$$(1 - \phi B)(1 - B)h_t = (1 - B)^{1-d}\eta_t . \quad (3.24)$$

Com a representação em (3.24), temos que o parâmetro d aparece apenas na expansão truncada da componente MA e o coeficiente ϕ somente aparece na componente AR.

A forma de espaço de estados é bastante atraente, em especial para o procedimento de estimação, por sua estrutura recursiva e semelhança com os modelos estruturais. Isto facilita a inclusão de componentes como a sazonal, a volatilidade, entre outras, possibilitando assim estudos adicionais do processo, além de ser possível estudar processos mais complexos, como os multivariados (Harvey, Ruiz e Shephard (1994)).

A forma de espaço de estados foi explorada, inicialmente, para propor a estimação do modelo de volatilidade estocástica pelo método da Máxima Quase-Verossimilhança (ver, por exemplo, Broto e Ruiz (2002)) e, em Durbin e Koopmann (2001a), pela Máxima Verossimilhança. Shephard (1996) também explora a forma de espaço de estados para propor um estimador bayesiano. Um resumo desta abordagem pode ser encontrado em Motta e Hotta (2001), em Chib, Nardari e Shephard (2002) e em Liesenfeld e Richard (2003).

3.4.2 Comportamentos dos Coeficientes das Formas AR e MA

A representação dos processos nas formas auto-regressivas AR e médias móveis MA facilita o estudo dos mesmos. A forma MA que utilizamos é uma modificação da forma comum utilizada. A modificação foi feita com a aplicação do operador $(1 - B)$, pois assim os coeficientes decaem mais rapidamente e menos observações passadas são necessárias para a explicação de um processo. Chan e Palma (1998), com o mesmo intuito, diferenciam um processo ARFIMA(p,d,q).

Seguem alguns gráficos para a visualização dos comportamentos dos coeficientes, para cada representação considerada e para vários valores do parâmetro d .

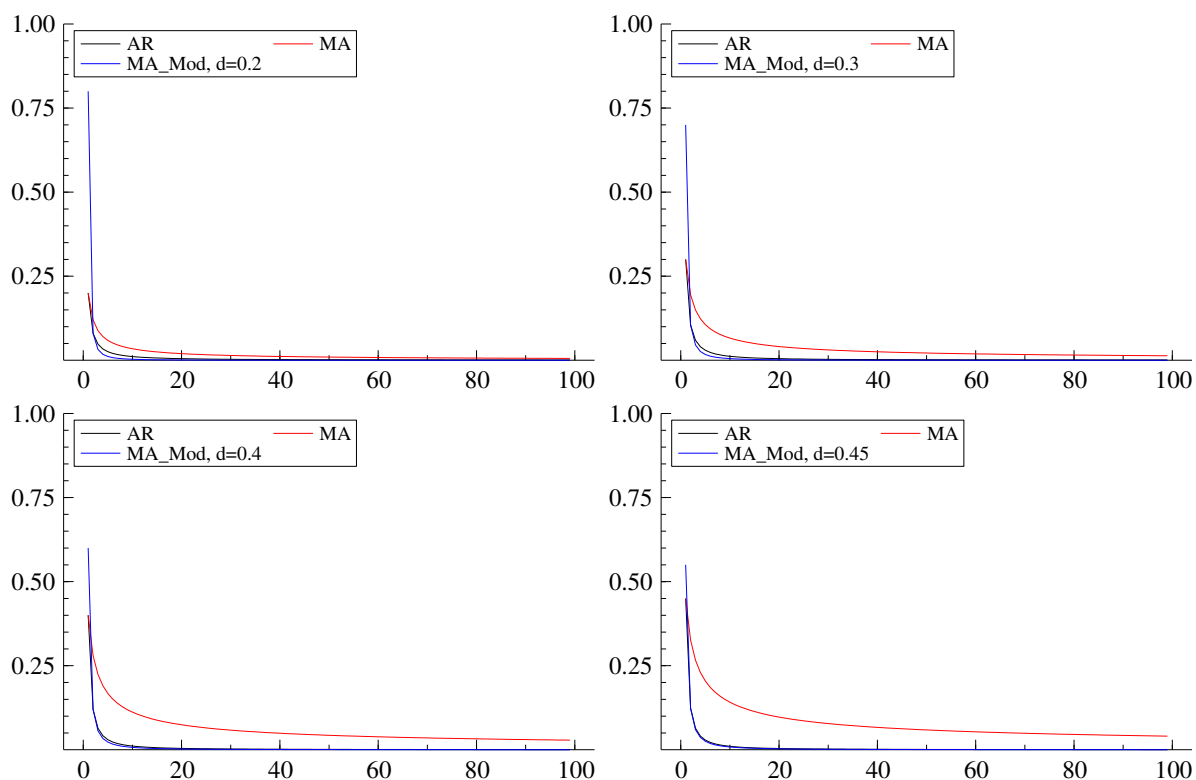


Figura 3.1: Comportamentos dos coeficientes das componentes AR e MA

Conforme a Figura 3.1, constatamos que os coeficientes que decaem mais rapidamente são os correspondentes às representações AR e MA-Mod. Com isto, menos informações são necessárias para explicar o processo estudado e, portanto, a convergência será mais rápida.

Abaixo seguem os gráficos das somas acumuladas dos coeficientes (em valores absolutos) das representações AR, MA e MA-Mod.

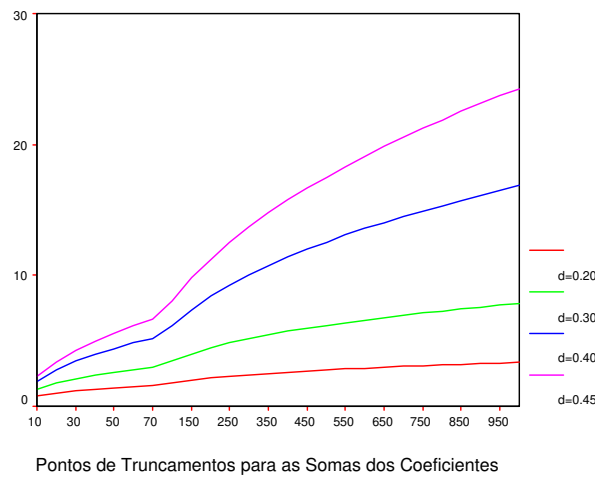


Figura 3.2: Somas Acumuladas dos Coeficientes da Componente MA

Através das Figuras (3.2) e (3.3), vemos que, à medida que aumenta o número de observações consideradas, a soma dos coeficientes da componente MA também aumenta o seu valor, enquanto que para a soma dos coeficientes das componentes AR e MA-Mod, a soma torna-se, aproximadamente, igual a 1.

Com isto, parece mais adequado o uso das representações AR e MA-Mod, pois assim utiliza-se um menor número de observações para a explicação de um processo.

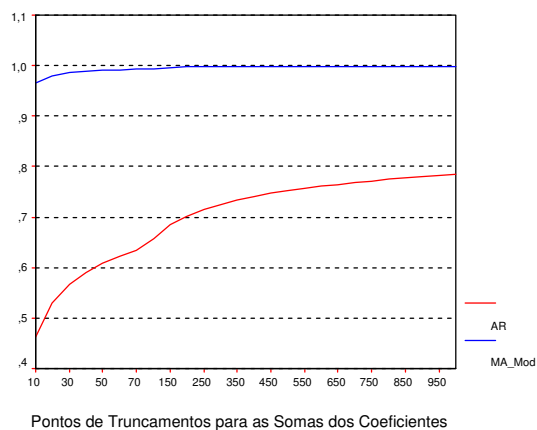
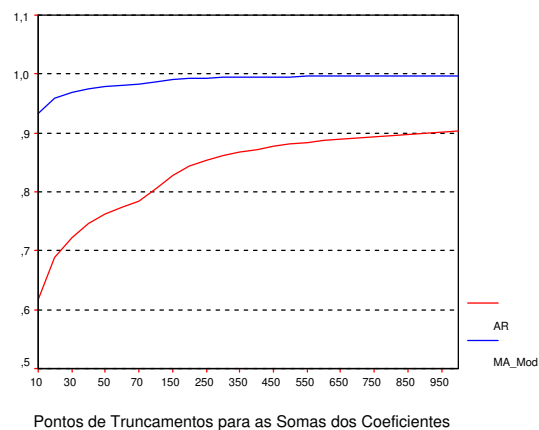
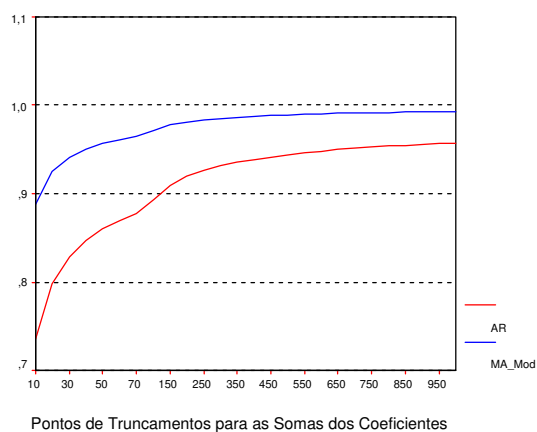
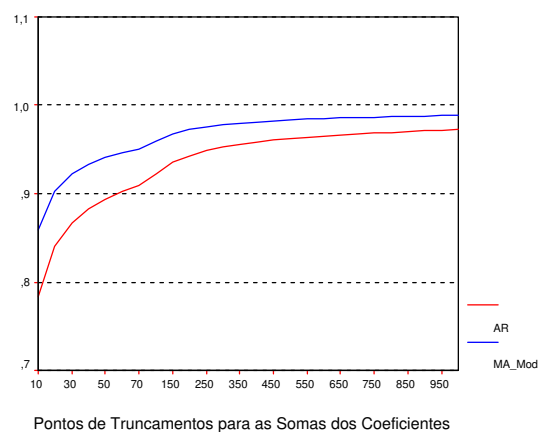
(a) $d = 0.2$ (b) $d = 0.3$ (c) $d = 0.4$ (d) $d = 0.45$

Figura 3.3: Somas Acumuladas dos coeficientes das componentes AR e MA-Mod

Como a soma, para as representações AR e MA-Mod, converge para 1, é razoável imaginar que muitas observações possuam coeficientes próximos a zero, tornando-as desnecessárias para a explicação do processo. Assim, é preciso determinar um ponto de truncamento para a soma das componentes.

Em muitos trabalhos, utiliza-se truncamentos com valores para m iguais a 30, 40 ou 50. Mas, como podemos constatar pelas expressões (3.22), os coeficientes da soma, assim como o ponto de truncamento, dependem do valor assumido para o parâmetro d e, portanto, os valores apropriados para o truncamento são obtidos por meio de simulações, estudados no Capítulo 5.

Capítulo 4

Métodos de Estimação

4.1 Introdução

Existem vários métodos, não apresentados aqui, de estimação do parâmetro de diferença, os quais podem ser classificados como semiparamétricos e paramétricos. Entre os semiparamétricos, um dos mais utilizados é o método GPH, proposto por Geweke e Porter-Hudak (1983). A popularidade deste método está na simplicidade de sua aplicação, pois está baseado na idéia de uma regressão linear simples do log-periodograma, no qual a estimação também pode ser feita utilizando-se o método de mínimos quadrados ordinários. Aplicações do método GPH, no contexto da volatilidade, podem ser encontradas em Andersen e Bollerslev (1997), Ray e Tsay (2000), Wright (2000), entre outros.

Deo e Hurvich (2001) consideram que a série de volatilidade é gaussiana e desenvolvem a teoria assintótica para o estimador GPH, baseando-se no logaritmo dos retornos ao quadrado. Isso fornece justificativa para o uso do estimador GPH para estimar a memória longa nos modelos de volatilidade estocástica, além de motivar um estimador semiparamétrico do parâmetro d , com base no que foi proposto originalmente por Geweke e Porter-Hudak (1983) para modelos ARFIMA, estudando seus vícios e variância assintótica.

Outro método semiparamétrico, que utilizamos, é o GSE (*Gaussian Semiparametric Estimator*), proposto por Robinson (1995b) e Robinson e Henry (1998), que aproxima a função de log-verossimilhança de um processo gaussiano ARFIMA(p,d,q).

O método paramétrico que estudamos é o da estimação pela Máxima Quase-Verossimilhança (MQV), no domínio do tempo. Justifica-se seu uso pois para modelos com memória longa não existe uma forma fechada para a função de verossimilhança e uma abordagem, recentemente usada, é a utilização de uma função de verossimilhança aproximada para a estimação dos parâmetros.

Em muitos trabalhos, como os de Durbin e Koopman (1997, 2001a) a função de verossimilhança aproximada é corrigida utilizando-se as técnicas de simulação (de Jong e Shephard (1995)) e amostragem por importância.

Deve-se salientar que o primeiro algoritmo para a aplicação da técnica de amostragem por importância foi de Geweke (1994), o qual motivou, além do trabalho de Durbin e Koopman (2001a), já mencionado, outros trabalhos como os de Shephard e Pitt (1997), Sandmann e Koopman (1998), entre outros.

O método da máxima quase-verossimilhança, no domínio da frequência, foi utilizado por Harvey (1998), em seus estudos dos modelos de volatilidade estocástica com memória longa (LMSV, *Long Memory Stochastic Volatility*), utilizando-se o processo ARFIMA (0,d,0), assim como por Breidt *et al.* (1998), que estuda o modelo LMSV no qual a log-volatilidade segue um processo ARFIMA(p,d,q) e demonstra que o estimador de máxima quase-verossimilhança é fortemente consistente. Deo (2002) prova a normalidade assintótica do estimador MQV, no domínio do tempo.

Embora o método MQV não seja o mais eficiente, ele tem a vantagem de ser facilmente implementado para a análise de séries reais e poder ser generalizado para processos mais complexos, como os multivariados.

Chan e Petris (2000) propõem uma abordagem bayesiana para desenvolver inferências no domínio do tempo, com base no método da verossimilhança truncada, proposto por Chan e Palma (1998), no qual o modelo LMSV está escrito na forma linear de espaço de estados. Com o modelo na forma de espaço de estados é possível obter a função de quase-verossimilhança e de verossimilhança exata, para inovações gaussianas, com as saídas do Filtro de Kalman.

Neste capítulo, apresentamos os métodos da seguinte forma: na Seção 4.2, colocamos os métodos semiparamétricos GPH e GSE e, na Seção 4.3, está o método da máxima quase-verossimilhança, escrito na formulação de espaço de estados.

4.2 Métodos Semiparamétricos

A simplicidade dos métodos que utilizam a regressão linear tornou estes métodos bastante utilizados na literatura. Apenas os estimadores de GPH e GSE, baseados na regressão do log-periodograma, serão utilizados para comparação, sendo brevemente apresentados nesta seção. Os métodos de regressão são bastante utilizados pois consistem na idéia simples de transformar a função espectral, com a aplicação do logaritmo, em uma equação similar à regressão linear simples. Assim, o parâmetro d pode ser estimado, de forma análoga aos parâmetros de uma regressão, pelo método de mínimos quadrados ordinários.

Aplicando-se o logaritmo neperiano na função de densidade espectral de um processo $\{y_t\}$ ARFIMA(p,d,q), dada pela expressão (2.12), obtemos

$$\ln f_y(\omega) = \ln f_x(\omega) - d \ln(2 \sin(\omega/2))^2, \quad (4.1)$$

ou ainda, adicionando-se e subtraindo-se o termo $\ln f_x(0)$ na expressão (4.1), temos

$$\ln f_y(\omega) = \ln f_x(0) - d \ln(2 \sin(\omega/2))^2 + \ln(f_x(\omega)/f_x(0)) \quad (4.2)$$

Assim, as várias formas propostas para estimar o espectro diferenciam os métodos que apresentaremos.

4.2.1 Método GPH - Geweke e Porter-Hudak

O método da regressão usando a função periodograma foi apresentado por Geweke e Porter-Hudak (1983), e tem sido utilizado em várias pesquisas, tais como: Sowell(1990), Brockwell e Davis (1987), Janacek (1982), Ray (1991), entre outros.

Considerando-se uma amostra $y_1, \dots, y_n$, podemos estimar a função densidade espectral $f(\omega)$ pela função periodograma, dada por

$$I_y(\omega) = \frac{1}{2\pi} [R(0) + 2 \sum_{s=1}^{n-1} R(s) \cos(s\omega)], \quad \omega \in [-\pi, \pi] \quad , \quad (4.3)$$

em que $R(s)$ é a função de auto-covariância amostral, isto é,

$$R(s) = \frac{1}{n} \sum_{i=1}^{n-s} (y_i - \bar{y})(y_{i+s} - \bar{y}), \quad s = 0, \pm 1, \dots, \pm(n-1)$$

e $\bar{y}$ é a média amostral.

Substituindo-se, em (4.3), ω por $\omega_j = \frac{2\pi j}{n}$, as frequências de Fourier e adicionando-se, $\ln(I_y(\omega_j))$ em (4.2), temos

$$\ln I_y(\omega_j) = \ln f_x(0) - d \ln(2 \sin(\omega_j/2))^2 + \ln(f_x(\omega_j)/f_x(0)) + \ln(I_y(\omega_j)/f_y(\omega_j)) \quad . \quad (4.4)$$

Considera-se o limite superior de j igual a $g(n)$, o qual é escolhido satisfazendo que

$$\frac{g(n)}{n} \longrightarrow 0 \quad , \quad \text{quando } n \longrightarrow \infty \quad ,$$

para ω_j próximo de zero, $\omega_j \leq \omega_{g(n)}$ e $\omega_{g(n)}$ pequeno.

E, considerando-se que o termo $\ln(f_x(\omega_j)/f_x(0))$ é desprezível, em comparação aos demais termos, para ω_j próximo de zero, podemos aproximar $\ln I_y(\omega_j)$ por

$$\ln I_y(\omega_j) \cong \ln f_x(0) - d \ln(2 \operatorname{sen}(\omega_j/2))^2 + \ln(I_y(\omega_j)/f_y(\omega_j)) \quad , \quad (4.5)$$

que tem a forma de uma equação de regressão linear simples

$$y_j = a + b x_j + e_j, \quad j = 1, \dots, g(n), \quad (4.6)$$

se considerarmos $y_j = \ln I_y(\omega_j)$, $x_j = \ln(2 \operatorname{sen}(\omega_j/2))^2$ e $e_j = \ln(I_y(\omega_j)/f_y(\omega_j)) + c$, $b = -d$, $a = \ln f_x(0) - c$ e $c = E[-\ln(I_y(\omega_j)/f_y(\omega_j))]$. O número de observações utilizadas na equação de regressão é definido por $g(n)$, o qual é uma função do tamanho amostral n , isto é, $g(n) = n^\alpha$, $0 < \alpha < 1$.

Lema 4.2.1. (*Reisen (1994)*) *Seja $\{y_t\}$ um processo ARFIMA(p, d, q) escrito por*

$$y_t = (1 - B)^{-d} \frac{\Theta(B)}{\Phi(B)} \epsilon_t, \quad (4.7)$$

em que $d < 0$ e $\{\epsilon_t\}$ é um processo ruído branco de distribuição normal.

Quando $n \rightarrow \infty$, a sequência $\{-\ln(I_y(\omega_j)/f_y(\omega_j))\}$, para $\omega_j = 2\pi j/n$ e $j = 1, \dots, [n/2] - 1$, é de termos independentes, de distribuição Gumbel, com média 0.577216 (constante de Euler) e variância $\pi^2/6$.

Pelo Lema 4.2.1, quando $d \in (-0.5, 0.0)$, as variáveis $\{\ln(I_y(\omega_j)/f_y(\omega_j))\}$, para $j = 1, \dots, g(n)$, são aproximadamente independentes de distribuição Gumbel, com média -0.577216 e variância $\pi^2/6$. Portanto, as variáveis $\{\epsilon_j\}$ têm a mesma distribuição, mas com médias iguais a zero. O resultado anterior sugere o estimador de d pelo método de mínimos quadrados da regressão de $y_1, \dots, y_{g(n)}$ em termos de $x_1, \dots, x_{g(n)}$.

Assim, temos

$$\hat{b} = \frac{\sum_{i=1}^{g(n)} (x_i - \bar{x}) y_i}{\sum_{i=1}^{g(n)} (x_i - \bar{x})^2} , \quad (4.8)$$

em que $\bar{x} = g(n)^{-1} \sum_{i=1}^{g(n)} x_i$.

O estimador de d usando a função periodograma no método de regressão é dado por

$$\hat{d} = -\hat{b} , \quad (4.9)$$

com as seguintes propriedades

$$E(\hat{d}) = d \quad (4.10)$$

$$Var(\hat{d}) = \frac{\pi^2}{6 \sum_{i=1}^{g(n)} (x_i - \bar{x})^2} \quad (4.11)$$

A distribuição assintótica de $\hat{d}$ é dada pelo seguinte teorema que sugere $g(n) = n^\alpha$, ($0 < \alpha < 1$).

Teorema 4.2.1. *(Geweke e Porter-Hudak (1983)) Seja $\{y_t\}$ um processo ARFIMA(p, d, q), com $d < 0$ e $I(w_j)$ sendo a função periodograma de $\{y_t\}$ nas frequências de Fourier $w_j = 2\pi j/n$. Seja $\hat{d}$ o estimador de d na forma (4.9) e supondo que $g(n) \rightarrow \infty$ e $g(n)/n \rightarrow 0$, quando $n \rightarrow \infty$. Então, $\hat{d}$ converge em probabilidade para d .*

Se $\lim\{(\ln(n))^2/g(n)\} = 0$, quando $n \rightarrow \infty$, então

$$\frac{\hat{d} - d}{\sqrt{Var(\hat{d})}} \xrightarrow{\mathcal{D}} N(0, 1) . \quad (4.12)$$

Uma importante questão é a escolha de $g(n)$ para o estimador GPH. Neste trabalho, utilizamos $g(n) = n^{1/2}$, como sugerido por Geweke e Porter-Hudak (1983) e investigado por Reisen (1994).

4.2.2 Método GSE - Estimador Semiparamétrico Gaussiano

O método de estimação GSE, utilizado por Robinson (1995b) e Robinson e Henry (1998), consiste na idéia de modelar a memória longa de uma série covariância estacionária y_t para $t = 0, \pm 1, \dots$, semiparametricamente, por

$$f_y(\omega) \sim G \omega^{1-2H} \quad \text{quando } \omega \rightarrow 0^+, \quad (4.13)$$

em que $H = d + 0.5$, $0.5 < H < 1$, $0 < G < \infty$ e $f_y(\omega)$ é a densidade espectral de y_t , satisfazendo

$$\gamma(k) = \text{Cov}(y_t, y_{t+k}) = \int_{-\pi}^{\pi} f_y(\omega) \cos(k\omega) d\omega, \quad k = 0, \pm 1, \dots$$

Pela expressão (4.13), $f_y(\omega)$ tem um pólo em $\omega = 0$, para $0.5 < H < 1$ (caso em que há memória longa em y_t), $f_y(\omega)$ é finita e positiva para $H = 0.5$ (caso em que há memória curta em y_t) e $f_y(0) = 0$, para $0 < H < 0.5$ (quando há memória intermediária, ou anti-persistência, em y_t).

Com base nas observações de y_t , para $t = 1, \dots, n$, define-se a função periodograma como em (4.3) e estima-se H por

$$\hat{H} = \text{argmin}_{\Delta_1 \leq h \leq \Delta_2} R(h), \quad (4.14)$$

para $0 < \Delta_1 < \Delta_2 < 1$. Ou seja, $\hat{H}$ é o argumento que minimiza a função $R(h)$, dada por

$$R(h) = \log \left\{ \frac{1}{g(n)} \sum_{j=1}^{g(n)} \frac{I_y(\omega_j)}{\omega_j^{1-2h}} \right\} - (2h-1) \frac{1}{g(n)} \sum_{j=1}^{g(n)} \log \omega_j, \quad (4.15)$$

para $g(n) \in (0, [n/2])$ e $\omega_j = 2\pi j/n$.

Como explicado em Robinson(1995b), para $g(n) = [n/2]$, $\hat{H}$ é uma forma gaussiana ou estimador de Whittle, sob o modelo paramétrico $f(\omega) = G|\omega|^{1-2H}$, para todo $\omega \in (-\pi, \pi]$, e suas propriedades assintóticas são discutidas em Fox e Taqu (1986), Giraitis e Surgailis (1990), entre outros, sob a normalidade ou, mais genericamente, sob a suposição de que y_t é linear, com as inovações independentes e identicamente distribuídas. Entretanto, essas propriedades dependem da correta especificação de $f(\omega)$, no intervalo $(-\pi, \pi]$, pois, em caso de má especificação, as estimativas serão inconsistentes.

Para $g(n) < [n/2]$ e $n \rightarrow \infty$,

$$\frac{1}{g(n)} + \frac{g(n)}{n} \longrightarrow 0 \quad , \quad (4.16)$$

$\hat{H}$ pode ser visto como um estimador semiparamétrico baseado em (4.13).

Suposição 4.2.1.

$$y_t = E(y_t) + \sum_{j=0}^{\infty} \alpha_j \epsilon_{t-j} \quad , \quad \sum_{j=0}^{\infty} \alpha_j^2 < \infty \quad ,$$

sendo que ϵ_t satisfaz pelo menos que

$$\begin{aligned} E(\epsilon_t | F_{t-1}) &= 0 \quad \text{quase certamente} \quad (q.c.), \\ \sigma_t^2 &\doteq V(\epsilon_t | F_{t-1}) = \sigma^2 \quad q.c. \quad , \end{aligned}$$

para todo t . $F_t = \sigma\{\epsilon_s, s \leq t\}$, ou seja, é a σ -álgebra de eventos gerados por $\{\epsilon_s, s \leq t\}$ e σ^2 é uma constante positiva.

Sob (4.13), pela Suposição 4.2.1 e (4.16) e outras condições de regularidade, Robinson (1995b) mostra que $\hat{H}$ é consistente para H e, com adicionais condições, tem-se

$$g(n)^{1/2}(\hat{H} - H) \xrightarrow{\mathcal{D}} N(0, 1/4) \quad , \quad \text{quando } n \longrightarrow \infty \quad .$$

4.3 Método de Estimação por Máxima Quase-Verossimilhança

4.3.1 Introdução

O modelo de volatilidade estocástica não tem função de verossimilhança exata, no domínio do tempo, definida de forma direta. Para a estimação do modelo, é mais apropriado colocá-lo na forma de espaço de estados, obtida pela linearização do modelo, com a transformação logaritmica das observações ao quadrado.

A forma de espaço de estados é recursiva e a função de verossimilhança é calculada via Filtro de Kalman (FK), um procedimento recursivo que fornece estimativas com propriedades ótimas. No entanto, para a utilização de FK, é necessário satisfazer a suposição de normalidade, uma suposição não satisfeita pelos processos de volatilidade estocástica que, após a linearização, possuem perturbações com distribuição não-gaussiana.

Assim, para os processos de volatilidade estocástica, uma proposta é supor normalidade e obter, com as saídas do FK, uma função de verossimilhança aproximada, denominada de função de quase-verossimilhança. Com isto, os parâmetros podem ser estimados maximizando-se essa função.

O estimador por máxima quase-verossimilhança (MQV) foi proposto, independentemente, por Nelson (1988) e Harvey *et al.* (1994). Ruiz (1994) mostrou que o estimador MQV é consistente e assintoticamente normal. Durbin e Koopman (2001a) calculam a função de verossimilhança exata, pela correção da função de quase-verossimilhança, através de amostragem por importância.

4.3.2 Função de Quase–Verossimilhança

A função de quase-verossimilhança é calculada com as saídas do FK e, para tanto, é necessário que o modelo esteja na forma de espaço de estados, assumindo-se que modelo seja condicionalmente gaussiano.

Considerando-se que o modelo de volatilidade estocástica, após a linearização, possa ser escrito na forma de espaço de estados, como na expressão (3.23), a função de log-verossimilhança gaussiana é obtida, através da decomposição do erro de predição, por

$$\begin{aligned} l &= \log p(x_1, \dots, x_n; \varphi) = \sum_{t=1}^n \log p(x_t | x_1, \dots, x_{t-1}; \varphi) \\ &= -\frac{nN}{2} \log(2\pi) - \frac{1}{2} \sum_{t=1}^n (\log |F_t| + \nu_t' F_t^{-1} \nu_t) \quad , \end{aligned} \quad (4.17)$$

em que $x_t = \log y_t^2$, φ é um vetor de parâmetros para um modelo especificado, ν_t é o erro de predição de x_t , um passo à frente e F_t é o correspondente erro quadrático médio de ν_t .

Os vetores ν_t e F_t são obtidos pelas equações do Filtro de Kalman,

$$\begin{aligned} \nu_t &= y_t - c_t - Z_t \alpha_t & F_t &= Z_t P_t Z_t' + H_t \\ K_t &= T_t P_t Z_t' F_t^{-1} & L_t &= T_t - K_t Z_t \\ a_{t+1} &= d_t + T_t \alpha_t + K_t \nu_t & P_{t+1} &= T_t P_t L_t' + R_t Q_t R_t' \quad , \end{aligned}$$

sendo K_t chamada de matriz ganho.

Para inicializar o FK, utiliza-se $a_1 = a$ e $P_1 = P_* + k P_\infty$, com $k = 10^7$ e $t = 1, \dots, n$. Em geral, quando é desconhecida a distribuição dos estados, inicializa-se o FK com $a_1 = 0$ e $P_1 = k P_\infty$, com $k = 10^7$, sendo P_∞ uma matriz identidade.

As equações de atualização a_{t+1} e P_{t+1} escritas, recursivamente, em termos de a_t e P_t , expressam, respectivamente, as médias e as variâncias dos estados, dadas por

$$\begin{aligned} a_{t+1} &= E(T_t \alpha_t + R_t \eta_t | Y_t) = T_t E(\alpha_t | Y_t) \\ P_{t+1} &= \text{Var}(T_t \alpha_t + R_t \eta_t | Y_t) = T_t \text{Var}(\alpha_t | Y_t) T_t' + R_t Q_t R_t' \quad , \end{aligned}$$

para todo $t = 1, \dots, n$.

A necessidade de assumir a normalidade faz o método de QML ter limitações em sua utilização, mas apesar disto, ainda é considerado um método bastante flexível, pois é possível sua generalização. Uma interessante generalização é feita para capturar o efeito de alavancagem, no qual a resposta da volatilidade é maior quando o retorno é negativo do que quando é positivo (ver Black (1986), Harvey e Shephard (1996)).

Capítulo 5

Simulação

5.1 Introdução

Neste capítulo, aplicamos os métodos de estimação, apresentados no Capítulo 4, em modelos de volatilidade estocástica com comportamentos de memória longa, nos quais as log-volatilidades seguem processos ARFIMA(p,d,q), tais modelos são denotados por VE-ARFIMA(p,d,q).

Consideramos séries que simulam os processos VE-ARFIMA(0,d,0) e VE-ARFIMA(1,d,0), com tamanhos iguais a 1000, para 300 e 160¹ replicações, respectivamente. Os parâmetros considerados foram: $\phi \in \{0.7, 0.9\}$, $d \in \{0.2, 0.3, 0.4, 0.45\}$ e $\sigma_\eta^2 = 1$, por serem os mais encontrados nos recentes estudos (ver, por exemplo, Hurvich, Deo e Brodsky (1998)). O valor $d = 0.45$ foi considerado para avaliar os comportamentos dos estimadores quando d tem valor próximo ao limite da não-estacionariedade dos modelos VE-ARFIMA(p,d,q), ou seja, d próximo a 0.5.

Os processos ARFIMA(0,d,0) e ARFIMA(1,d,0) foram aproximados pelas representações auto-regressivas (AR) e médias móveis (MA), com truncamentos, no operador de diferença fracionária, iguais a 20 e 30. O programa utilizado, para gerar as séries de processos ARFIMA, foi o `arima.frac.diff`, encontrado no aplicativo S-Plus 2000 .

¹O número de replicações para os processos VE-ARFIMA(1,d,0) é menor, devido ao maior tempo necessário para a realização das simulações.

Sabemos que, na prática, a volatilidade é estacionária, pois não pode ter comportamento explosivo. Não colocamos, nos programas, a restrição $|d| < 0.5$, pois com ela o desempenho dos estimadores melhoraria bastante, principalmente para valores pequenos de d .

Para obter as estimativas pelo método da máxima quase-verossimilhança MQV aproximada, com a função de quase-verossimilhança aproximada calculada via Filtro de Kalman e formulação de espaço de estados, utilizamos a função de maximização maxBFGS (função de maximização pelo método de Quase-Newton, desenvolvido por Broyden-Fletcher-Goldfarb-Shanno, BFGS, para a linguagem Ox), ver Doornik e Ooms(1999)².

Com relação aos métodos GPH e GSE, utilizamos a forma da linguagem Ox ³ e consideramos $g(n) = n^{1/2}$ pontos na regressão do log-periodograma, sendo os pontos avaliados nas frequências de Fourier, $2\pi j/n$, para $j=1, \dots, g(n)$.

Para comparar os métodos de estimação, calculamos a média das estimativas, assim como dos valores obtidos para os seus erros quadráticos médios (EQM), vícios e desvios-padrão.

O Capítulo está organizado da seguinte forma. Na Seção 5.2, apresentamos os resultados obtidos com as séries simuladas dos processos VE-ARFIMA(0,d,0). Na Seção 5.3, temos os resultados das séries simuladas para os processos VE-ARFIMA(1,d,0). Na Seção 5.4, temos a conclusão para os resultados fornecidos pelos métodos de estimação.

²As séries não foram geradas utilizando-se a linguagem Ox, devido ao alto custo computacional.

³O programa da linguagem Ox implementa o método GSE para estimar d , como discutido em Robinson e Henry (1998).

5.2 Modelos VE-ARFIMA(0,d,0)

Nesta seção, apresentamos os resultados obtidos com as séries simuladas dos processos de volatilidade estocástica com memória longa VE-ARFIMA(0,d,0), para $d \in \{0.2, 0.3, 0.4, 0.45\}$. Temos também o interesse de verificar o comportamento dos estimadores MQV, GPH e GSE quando d tem valor próximo ao limite da não-estacionariedade desses processos, ou seja, d próximo a 0.5.

Inicialmente, comparamos os quatro estimadores por máxima quase-verossimilhança obtidos com as representações AR e MA e truncamentos iguais a 20 e 30. Em seguida, comparamos os estimadores MQV com as formas AR e MA, com truncamentos iguais a 30, pois apresentam melhores resultados (em termos dos valores dos EQM, vícios e desvios-padrão) com os métodos GPH e GSE. Os sinais dos vícios são mantidos, mas a comparação é feita em termos das magnitudes dos mesmos.

As tabelas possuem algumas abreviações, como:

- ▷ AR20, AR30, MA20 e MA30 - Representações AR e MA com truncamentos iguais a 20 e 30;
- ▷ dp - Desvio-padrão das estimativas.

Tabela 5.1: Comparação das estimativas QMV para d - VE-ARFIMA(0,d,0)

d	Dp				EQM				Viés			
	AR20	AR30	MA20	MA30	AR20	AR30	MA20	MA30	AR20	AR30	MA20	MA30
0.20	0.188	0.187	0.121	0.139	0.041	0.043	0.021	0.021	-0.079	-0.088	-0.082	-0.043
0.30	0.189	0.159	0.141	0.148	0.037	0.030	0.025	0.024	-0.037	-0.064	-0.068	-0.045
0.40	0.180	0.155	0.143	0.149	0.033	0.025	0.025	0.024	0.013	-0.037	-0.066	-0.038
0.45	0.168	0.147	0.142	0.146	0.029	0.022	0.023	0.022	0.022	-0.024	0.052	-0.023

Os valores em negrito são os menores obtidos para cada medida considerada.

Tabela 5.2: Comparação das estimativas para d - VE-ARFIMA(0,d,0)

d	Dp				EQM				Viés			
	AR30	MA30	GPH	GSE	AR30	MA30	GPH	GSE	AR30	MA30	GPH	GSE
0.20	0.187	0.139	0.143	0.117	0.043	0.021	0.032	0.029	-0.088	-0.043	-0.110	-0.123
0.30	0.159	0.148	0.137	0.116	0.030	0.024	0.035	0.033	-0.064	-0.045	-0.125	-0.140
0.40	0.155	0.149	0.138	0.116	0.025	0.024	0.031	0.030	-0.037	-0.038	-0.109	-0.127
0.45	0.147	0.146	0.154	0.125	0.022	0.022	0.038	0.031	-0.024	-0.023	-0.119	-0.124

Os valores em negrito são os menores obtidos para cada medida considerada.

Tabela 5.3: Comparação das estimativas de σ_η^2 e σ_ν^2 - VE-ARFIMA(0,d,0)

d	esti ma dor	$\sigma_\eta^2 = 1$			$\sigma_\nu^2 \approx 4.9348$		
		$\hat{\sigma}_\eta^2$			$\hat{\sigma}_\nu^2$		
		dp	EQM	Viés	dp	EQM	Viés
0.20	AR20	2.239	9.290	2.068	2.210	9.016	-2.033
	AR30	3.051	13.585	2.068	3.042	13.362	-2.027
	MA20	7.997	84.837	4.570	7.954	83.631	-4.513
	MA30	2.332	8.564	1.768	2.291	8.244	-1.731
0.30	AR20	2.145	6.818	1.489	2.076	6.294	-1.409
	AR30	3.275	13.282	1.599	3.252	12.873	-1.516
	MA20	2.544	10.408	1.984	2.437	9.443	-1.872
	MA30	4.550	22.660	1.399	4.488	21.881	-1.319
0.40	AR20	1.674	3.126	0.569	1.578	2.791	-0.549
	AR30	1.800	3.992	0.867	1.685	3.520	-0.825
	MA20	2.021	5.489	1.185	1.918	4.926	-1.117
	MA30	1.824	4.116	0.888	1.730	3.705	-0.844
0.45	AR20	1.402	2.091	0.354	1.354	1.929	-0.310
	AR30	1.458	2.467	0.584	1.398	2.221	-0.517
	MA20	1.550	3.095	0.832	1.481	2.741	-0.740
	MA30	1.326	2.078	0.565	1.270	1.860	-0.497

Os valores em negrito são os menores obtidos, para cada medida considerada.

Conforme resultados da Tabela 5.1, vemos que as melhores estimativas para o parâmetro d , em termos do EQM, são obtidas com o método MQV, utilizando-se a representação MA com truncamento igual a 30, além de terem menores vícios do que as obtidas pela representação AR com igual truncamento.

Para truncamento igual a 20, a representação AR produz estimativas com menores vícios, porém com maiores variabilidades; enquanto que a representação MA fornece estimativas com maiores vícios, mas com menores variabilidades.

Assim, obtemos estimativas com menores vícios, porém com maiores variabilidades pela representação AR e com maiores vícios, mas com menores variabilidades pela representação MA.

A vantagem da representação MA é que, ao aumentar o truncamento de 20 para 30, embora as variabilidades aumentem um pouco, os vícios das estimativas diminuem, consideravelmente; enquanto que para a representação AR, ocorre o contrário: com o aumento do truncamento, os vícios aumentam e as variabilidades diminuem.

Assim, a escolha entre as representações AR e MA deve ser feita com base nos valores dos EQM das estimativas obtidas. Como já mencionamos, a representação MA, com truncamento igual a 30, fornece estimativas com menores valores para EQM e, portanto, ela é preferível.

As representações AR e MA sub-estimam o parâmetro d , exceto a representação MA, com truncamento igual a 20 e $d = 0.45$ e a representação AR, com igual truncamento e $d \in \{0.4, 0.45\}$.

Conforme Tabela 5.2, comparando-se os resultados obtidos pelas representações AR e MA, com truncamentos iguais a 30, com os obtidos pelos métodos GPH e GSE, notamos que as estimativas obtidas pela representação MA têm menores valores para os EQM e vícios, exceto para $d = 0.4$, em que as de menores vícios são obtidas pela representação AR. As estimativas obtidas pelo métodos GPH e GSE têm variabilidades menores, mas possuem grandes vícios negativos, sub-estimando o parâmetro d .

Em termos do EQM, a representação MA, com truncamento igual a 30, fornece melhores estimativas que os métodos GPH e GSE.

É interessante observar que à medida que o valor do parâmetro d é aumentado, ou seja, aproximado do limite de não-estacionariedade dos modelos VE-ARFIMA(0,d,0), d próximo a 0.5, os vícios das estimativas, obtidas pelas representações AR e MA, são diminuídos, e aumentados para as obtidas pelos métodos GPH e GSE. Constatamos, assim, que o método MQV fornece estimativas mais próximas ao verdadeiro valor do parâmetro, para modelos não-estacionários, que os métodos GPH e GSE.

Pelos resultados da Tabela 5.3, vemos que, as estimativas, obtidas para os parâmetros σ_η^2 e σ_ν^2 , tornam-se melhores (em termos do EQM, viés e d.p.) quando aumenta-se o valor do parâmetro d , para qualquer uma das representações, AR e MA. Notamos também que, para $d \in \{0.2, 0.45\}$, a representação MA, com truncamento igual a 30, fornece estimativas com menores EQM. E, para $d \in \{0.3, 0.4\}$, os menores são obtidos pela representação AR, com truncamento igual a 20. Com os mesmos truncamentos, obtemos menores vícios pela representação MA, para $d \in \{0.2, 0.3\}$ e pela representação AR, para $d \in \{0.4, 0.45\}$.

Tabela 5.4: Correlações, variâncias e covariâncias das estimativas de d :VE-ARFIMA(0,d,0)

	VE-ARFIMA(0,0.2,0)						VE-ARFIMA(0,0.3,0)					
	AR20	AR30	MA20	MA30	GPH	GSE	AR20	AR30	MA20	MA30	GPH	GSE
NC	2	4	10	10	-	-	7	4	9	5	-	-
AR20	0.035	0.823	0.550	0.523	0.271	0.379	0.036	0.921	0.795	0.841	0.441	0.559
AR30	0.029	0.035	0.551	0.594	0.333	0.442	0.028	0.025	0.800	0.825	0.444	0.553
MA20	0.012	0.012	0.015	0.903	0.355	0.452	0.019	0.017	0.019	0.944	0.399	0.555
MA30	0.013	0.015	0.014	0.019	0.306	0.401	0.022	0.019	0.019	0.022	0.442	0.578
GPH	0.007	0.009	0.006	0.006	0.020	0.844	0.011	0.009	0.008	0.009	0.019	0.784
GSE	0.008	0.010	0.006	0.007	0.014	0.014	0.012	0.010	0.009	0.009	0.012	0.013
	VE-ARFIMA(0,0.4,0)						VE-ARFIMA(0,0.45,0)					
	AR20	AR30	MA20	MA30	GPH	GSE	AR20	AR30	MA20	MA30	GPH	GSE
NC	2	1	2	1	-	-	1	1	1	1	-	-
AR20	0.032	0.965	0.758	0.802	0.644	0.741	0.028	0.978	0.789	0.821	0.669	0.788
AR30	0.027	0.024	0.779	0.835	0.650	0.743	0.024	0.022	0.806	0.839	0.668	0.781
MA20	0.019	0.017	0.020	0.976	0.606	0.677	0.019	0.017	0.020	0.979	0.576	0.694
MA30	0.021	0.019	0.020	0.022	0.632	0.697	0.020	0.018	0.019	0.021	0.606	0.712
GPH	0.016	0.014	0.012	0.013	0.019	0.841	0.017	0.015	0.012	0.013	0.023	0.840
GSE	0.015	0.013	0.011	0.012	0.013	0.013	0.016	0.014	0.012	0.013	0.016	0.015

Os valores marcados em negrito são as variâncias, acima destes estão as correlações e abaixo as covariâncias.

NC : Número de não-convergências em cada uma das representações.

Conforme Tabela 5.4, notamos que as correlações entre as estimativas obtidas por todos métodos aumentam com o aumento do valor do parâmetro d . As estimativas, obtidas pelas representações AR e MA, são mais correlacionadas com as obtidas com o método GSE e mais correlacionadas entre si quando aumenta-se o truncamento. Vale salientar que, quando $d \in \{0.4, 0.45\}$, as estimativas, obtidas pela representação AR, são mais correlacionadas com as obtidas pelos métodos GPH e GSE, que as obtidas pela representação MA.

Pelos gráficos da Figura 5.1, vemos claramente, que a dependência linear entre as estimativas torna-se mais evidente com o aumento do parâmetro d , para cada uma das representações e, também, entre as representações.

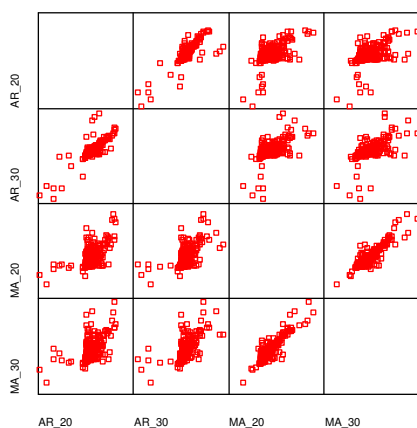
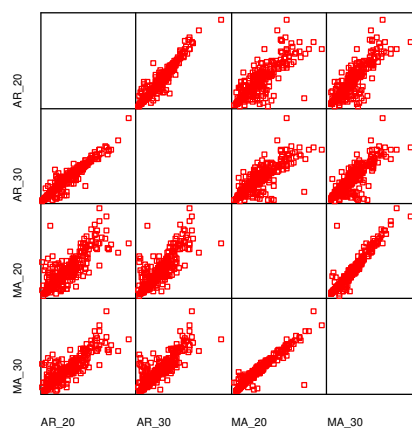
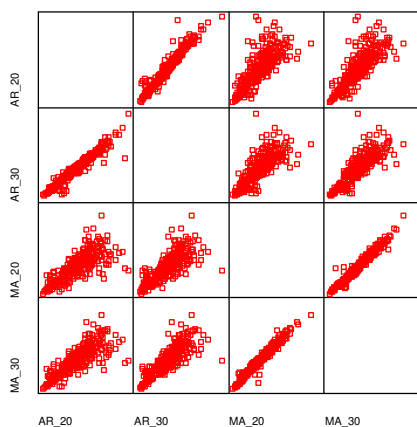
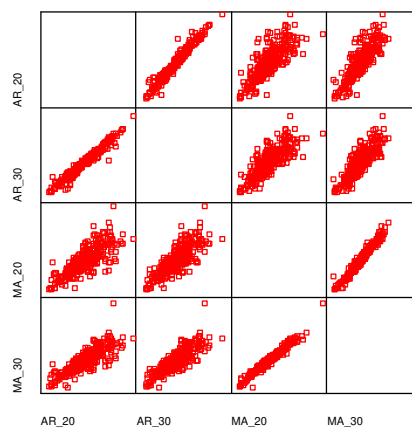
(a) $d = 0.2$ (b) $d = 0.3$ (c) $d = 0.4$ (d) $d = 0.45$

Figura 5.1: Dispersão das Estimativas de d obtidas com as formas AR e MA - VE-ARFIMA(0,d,0)

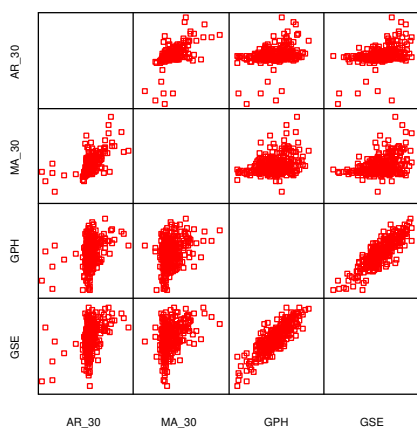
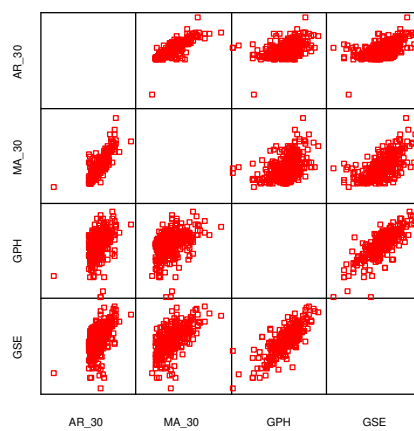
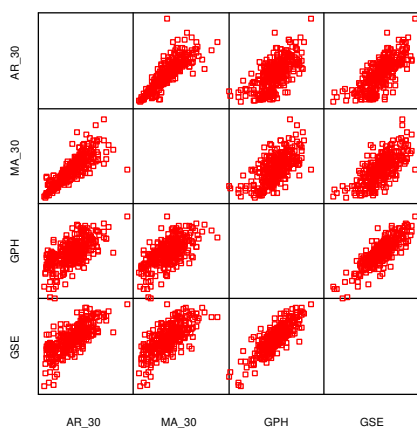
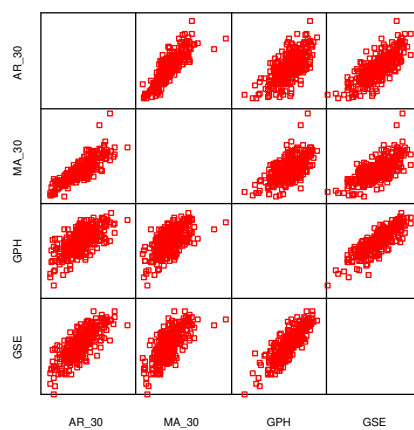
(a) $d = 0.2$ (b) $d = 0.3$ (c) $d = 0.4$ (d) $d = 0.45$

Figura 5.2: Dispersão das estimativas de d obtidas com as formas AR30, MA30, GPH e GSE - VE-ARFIMA(0,d,0)

Através dos gráficos da Figura 5.2, constatamos que a dependência linear entre as estimativas, obtidas por meio das formas AR30 e MA30 e pelos métodos GPH e GSE, também aumenta com o valor do parâmetro d , sendo mais forte entre estes dois últimos.

O comportamento da distribuição das estimativas obtidas por cada método de estimação pode ser verificado com uso dos gráficos de *Box Plot*.

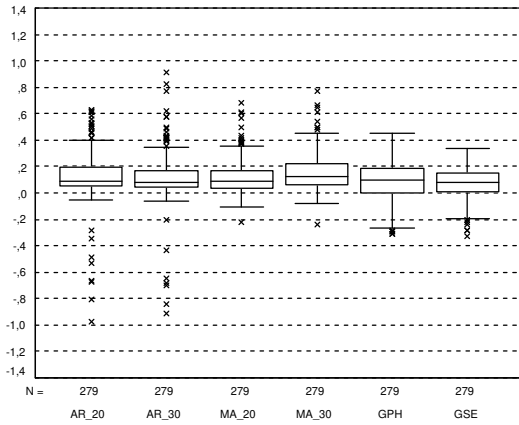
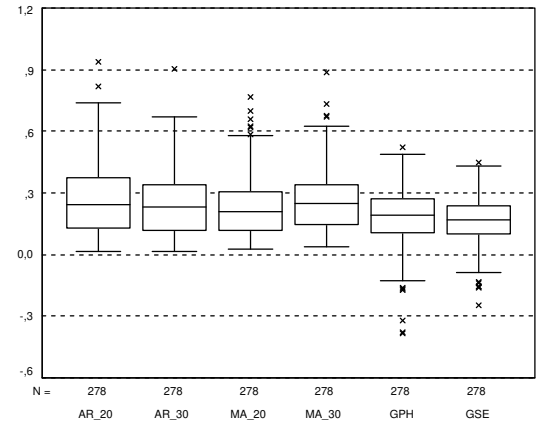
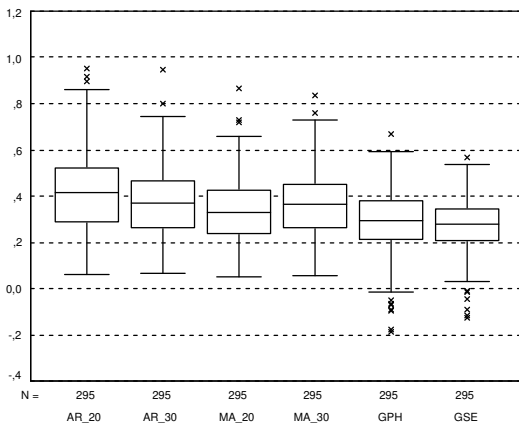
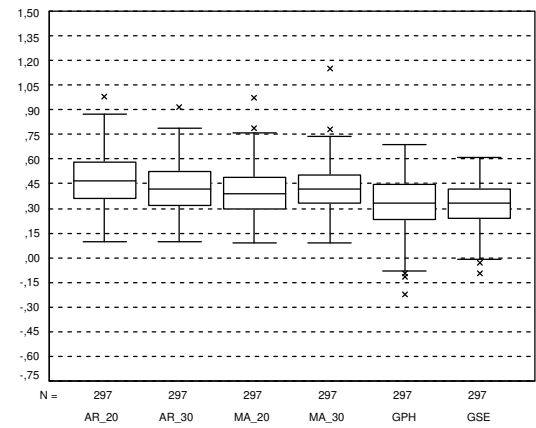
(a) $d = 0.2$ (b) $d = 0.3$ (c) $d = 0.4$ (d) $d = 0.45$

Figura 5.3: *Box Plot* das Estimativas de d - VE-ARFIMA(0,d,0)

Observamos pela Figura 5.3, que as medianas das estimativas obtidas pelos métodos GPH e GSE tornam-se cada vez mais distantes dos verdadeiros valores dos parâmetros d , à medida em que estes aumentam. Também temos, para quase todos métodos, que as estimativas obtidas para d têm medianas abaixo do valor do parâmetro, exceto as da representação AR com truncamento em 20 e parâmetro $d \in \{0.4, 0.45\}$, as quais têm valores muito próximos aos dos verdadeiros valores dos parâmetros.

Vemos que os métodos GPH e GSE têm terceiros quartis com valores abaixo dos verdadeiros valores dos parâmetros, ou seja, a maioria das estimativas produzidas por esses métodos tem valores menores que os dos parâmetros.

Com o auxílio dos gráficos de *Box Plot*, podemos verificar a presença de valores discrepantes, o que não pode ser visto apenas com os valores das estimativas. Pode-se notar que a representação AR produz muitos valores discrepantes e que este número diminui com o aumento do valor de d .

5.3 Modelos VE-ARFIMA(1,d,0)

Nesta seção, apresentamos os resultados obtidos com as séries simuladas dos processos de volatilidade estocástica com memória longa VE-ARFIMA(1,d,0), para $d \in \{0.2, 0.3, 0.4, 0.45\}$ e $\phi \in \{0.7, 0.9\}$. De forma análoga à seção anterior, comparamos, primeiramente, as estimativas obtidas pelas representações AR e MA, com truncamentos iguais a 20 e 30. Em seguida, comparamos os estimadores AR e MA, com truncamentos iguais a 30, com os métodos GPH e GSE, com base nas médias dos erros quadráticos médios EQM, desvios-padrão e vícios das estimativas.

Conforme resultados da Tabela 5.5, comparando-se as estimativas obtidas pelas representações AR e MA, para o parâmetro d , considerando-se $\phi = 0.7$, observamos que, em geral, a representação MA, com truncamento, no operador de diferença fracionária, igual a 30, fornece melhores estimativas (em termos do dp, EQM e viés), exceto para $d = 0.45$, em que a representação AR, com truncamento igual a 30, produz estimativas com menores vícios.

Tabela 5.5: Comparação das estimativas por MQV de d e ϕ - VE-ARFIMA(1,d,0)

d	esti ma dor	$\phi = 0.7$						$\phi = 0.9$					
		$\hat{d}$			$\hat{\phi}$			$\hat{d}$			$\hat{\phi}$		
		dp	EQM	viés	dp	EQM	viés	dp	EQM	viés	dp	EQM	viés
0.20	AR20	0.205	0.045	-0.055	0.261	0.075	-0.083	0.224	0.051	-0.016	0.173	0.032	-0.048
	AR30	0.222	0.057	-0.088	0.237	0.058	-0.046	0.209	0.044	-0.005	0.134	0.020	-0.041
	MA20	0.123	0.017	-0.046	0.213	0.047	-0.038	0.175	0.032	-0.036	0.170	0.030	-0.029
	MA30	0.118	0.014	-0.010	0.248	0.070	-0.092	0.183	0.034	0.009	0.174	0.032	-0.046
0.30	AR20	0.200	0.041	-0.025	0.225	0.056	-0.072	0.231	0.054	-0.006	0.104	0.012	-0.029
	AR30	0.191	0.039	-0.052	0.207	0.044	-0.039	0.209	0.044	0.000	0.091	0.009	-0.026
	MA20	0.167	0.031	-0.053	0.221	0.050	-0.025	0.180	0.037	-0.065	0.074	0.006	0.004
	MA30	0.161	0.027	-0.021	0.230	0.056	-0.059	0.189	0.036	-0.007	0.081	0.007	-0.015
0.40	AR20	0.179	0.033	0.016	0.268	0.097	-0.159	0.278	0.078	0.024	0.131	0.021	-0.063
	AR30	0.196	0.040	-0.032	0.222	0.056	-0.083	0.244	0.062	0.049	0.112	0.016	-0.062
	MA20	0.184	0.037	-0.052	0.235	0.058	-0.055	0.252	0.064	0.013	0.151	0.026	-0.053
	MA30	0.163	0.027	-0.012	0.248	0.071	-0.099	0.229	0.054	0.038	0.142	0.023	-0.057
0.45	AR20	0.195	0.040	0.043	0.262	0.096	-0.164	0.306	0.095	-0.036	0.136	0.020	-0.037
	AR30	0.202	0.041	-0.002	0.236	0.067	-0.106	0.280	0.080	0.035	0.146	0.025	-0.062
	MA20	0.198	0.041	-0.042	0.264	0.075	-0.075	0.265	0.074	0.058	0.214	0.053	-0.086
	MA30	0.180	0.032	-0.009	0.254	0.074	-0.098	0.228	0.058	0.074	0.161	0.031	-0.073

Os valores em negrito, são os menores, para cada medida considerada.

Tabela 5.6: Comparação das estimativas de d - VE-ARFIMA(1,d,0)

d	estimator	$\phi = 0.7$			$\phi = 0.9$		
		dp	EQM	viés	dp	EQM	viés
0.20	AR30	0.222	0.057	-0.088	0.209	0.044	-0.005
	MA30	0.118	0.014	-0.010	0.183	0.034	0.009
	GPH	0.151	0.023	0.001	0.154	0.076	0.229
	GSE	0.128	0.016	-0.002	0.134	0.075	0.239
0.30	AR30	0.191	0.039	-0.052	0.209	0.044	0.000
	MA30	0.161	0.027	-0.021	0.189	0.036	-0.007
	GPH	0.148	0.022	0.001	0.150	0.075	0.228
	GSE	0.135	0.018	-0.004	0.149	0.077	0.235
0.40	AR30	0.196	0.040	-0.032	0.244	0.062	0.049
	MA30	0.163	0.027	-0.012	0.229	0.054	0.038
	GPH	0.137	0.019	0.023	0.138	0.083	0.252
	GSE	0.123	0.016	0.018	0.127	0.082	0.256
0.45	AR30	0.202	0.041	-0.002	0.280	0.080	0.035
	MA30	0.180	0.032	-0.009	0.228	0.058	0.074
	GPH	0.152	0.024	0.035	0.143	0.086	0.256
	GSE	0.132	0.018	0.026	0.131	0.083	0.257

Os valores em negrito, são os menores, para cada medida considerada.

Observamos, conforme resultados na Tabela 5.5, que os vícios das estimativas do parâmetro d , para $\phi = 0.7$, são todos negativos e, portanto, os métodos sub-estimam o parâmetro, exceto para o caso da representação AR, com truncamento igual a 20 e $d \in \{0.4, 0.45\}$.

As estimativas obtidas, para o parâmetro $\phi = 0.7$, têm menores vícios quando utiliza-se a representação MA e truncamento igual a 20 e têm menores valores para EQM e dp, quando obtidas pela representação AR e truncamento igual a 30, exceto para $d = 0.2$, em que a representação MA, com truncamento igual a 20, tem melhores resultados.

Comparando-se as estimativas obtidas, para o parâmetro d , considerando-se $\phi = 0.9$, notamos que as obtidas pela representação AR, truncamento igual a 30, têm menores vícios, exceto para $d = 0.4$ (obtem-se menores vícios para as obtidas pela representação MA, com truncamento igual a 20). As estimativas com menores valores para EQM e d.p., são as obtidas pelas representações MA, com truncamento igual a 20 (para $d = 0.2$) e MA, com truncamento igual a 30, (para $d \in \{0.4, 0.45\}$). Notamos que o método QMV super-estima o valor de d , quando $d \in \{0.4, 0.45\}$, exceto para o caso da representação AR, com truncamento igual a 20 e $d = 0.45$.

Tabela 5.7: Comparação das estimativas de σ_η^2 e σ_ν^2 - VE-ARFIMA(1,d,0)

d	esti ma dor	$\phi = 0.7, \sigma_\eta^2 = 1 \quad e \quad \sigma_\nu^2 \approx 4.9348$						$\phi = 0.9, \sigma_\eta^2 = 1 \quad e \quad \sigma_\nu^2 \approx 4.9348$					
		$\hat{\sigma}_\eta^2$			$\hat{\sigma}_\nu^2$			$\hat{\sigma}_\eta^2$			$\hat{\sigma}_\nu^2$		
		dp	EQM	Viés	dp	EQM	Viés	dp	EQM	Viés	dp	EQM	Viés
0.20	AR20	1.490	2.841	0.788	1.237	1.924	-0.628	0.742	0.676	0.354	0.643	0.485	-0.267
	AR30	1.393	2.484	0.737	1.175	1.732	-0.593	0.577	0.407	0.272	0.561	0.361	-0.216
	MA20	1.827	3.721	0.619	1.589	2.795	-0.519	0.667	0.551	0.325	0.586	0.402	-0.242
	MA30	2.120	5.200	0.840	1.823	3.817	-0.703	0.673	0.511	0.241	0.593	0.390	-0.195
0.30	AR20	1.433	2.525	0.687	1.081	1.386	-0.466	0.762	0.711	0.361	0.534	0.330	-0.213
	AR30	1.367	2.285	0.645	1.008	1.201	-0.430	0.770	0.703	0.332	0.531	0.321	-0.198
	MA20	1.561	2.827	0.625	1.227	1.690	-0.430	0.697	0.662	0.420	0.517	0.325	-0.241
	MA30	1.584	2.943	0.659	1.250	1.772	-0.458	0.701	0.577	0.293	0.519	0.299	-0.173
0.40	AR20	1.569	3.134	0.820	1.162	1.692	-0.585	0.532	0.362	0.281	0.470	0.262	-0.204
	AR30	1.036	1.371	0.546	0.813	0.810	-0.386	0.493	0.277	0.184	0.447	0.223	-0.154
	MA20	0.883	0.995	0.464	0.676	0.563	-0.326	0.475	0.293	0.259	0.440	0.229	-0.188
	MA30	1.382	2.260	0.592	1.064	1.317	-0.430	0.441	0.227	0.180	0.448	0.222	-0.146
0.45	AR20	0.920	1.130	0.533	0.692	0.576	-0.312	0.632	0.568	0.410	0.497	0.292	-0.212
	AR30	0.877	0.988	0.468	0.641	0.486	-0.274	0.634	0.471	0.263	0.499	0.269	-0.143
	MA20	1.151	1.612	0.536	0.868	0.860	-0.327	0.650	0.482	0.243	0.475	0.244	-0.135
	MA30	1.029	1.305	0.496	0.763	0.673	-0.302	0.503	0.268	0.124	0.430	0.190	-0.073

Os valores em negrito, são os menores, para cada medida considerada.

Para o parâmetro $\phi = 0.9$, as estimativas têm menores vícios quando obtidas pela representação MA e truncamento igual a 20, exceto as obtidas pela representação AR, com truncamento igual a 20, para $d = 0.45$. Em geral, os menores valores para EQM e d.p. são das estimativas obtidas pela representação AR, exceto os das estimativas obtidas pela representação MA, com truncamento igual a 20, para $d = 0.3$.

As estimativas obtidas para $\phi = 0.9$ têm, em geral, maiores vícios quando aumentamos o valor de d .

Observamos que, para $d = 0.3$, as estimativas obtidas para $\phi = 0.9$, pela representação MA e truncamento igual a 20 são as melhores, em termos de dp, EQM e viés. Para $d = 0.45$, as melhores são as obtidas pela representação AR e truncamento igual a 20.

Com o aumento do parâmetro ϕ , as representações AR e MA produzem estimativas, para o parâmetro d , com maiores valores para dp e EQM.

Quando utilizamos a representação AR para estimar o parâmetro d , para $\phi = 0.7$, notamos que o aumento do truncamento também aumenta a variabilidade das estimativas; enquanto que para as estimativas obtidas para ϕ , as variabilidades são diminuídas. Para a representação MA ocorre o contrário: o aumento no truncamento diminui a variabilidade das estimativas de d e aumenta das estimativas de ϕ .

Para $\phi = 0.9$, o aumento no truncamento para a representação AR, provoca a diminuição dos valores de EQM e dp das estimativas de d e ϕ , exceto para as estimativas de ϕ , quando $d = 0.45$. Para a representação MA temos que, para $d \in \{0.2, 0.3\}$, o aumento no truncamento aumenta a variabilidade das estimativas de d e ϕ ; enquanto que para $d \in \{0.4, 0.45\}$, essas variabilidades diminuem.

Com relação às estimativas obtidas para $\phi = 0.7$, com o aumento do truncamento, os valores para dp, EQM e vícios diminuem quando obtidas pela representação AR e, aumentam para as obtidas pela representação MA, exceto os valores de dp e EQM das estimativas obtidas, quando $d = 0.45$.

As estimativas obtidas, para o parâmetro $\phi = 0.9$, pela representação AR possuem valores para dp, EQM e vícios que diminuem com o aumento do truncamento, exceto para $d = 0.45$. Já as obtidas pela representação MA, têm vícios, dp e EQM que aumentam com o aumento do truncamento, exceto os vícios das estimativas obtidas quando $d = 0.45$ e, dp e EQM, quando $d \in \{0.4, 0.45\}$. Em geral, o parâmetro $\phi = 0.9$ é sub-estimado pelo método MQV, para quaisquer representações, pois produz vícios negativos.

Conforme Tabela 5.6, comparando-se as estimativas obtidas pelas representações AR e MA, com truncamentos iguais a 30, com as obtidas pelos métodos GPH e GSE, observamos que, para $\phi = 0.7$, as estimativas têm menores dp e EQM quando obtidas pelo método GSE (exceto para $d = 0.2$, em que os menores valores são produzidos pela representação MA) e, menores vícios quando obtidas pelo método GPH, exceto para $d \in \{0.4, 0.45\}$, em que o método MQV produz estimativas com menores vícios.

Notamos que os métodos GPH e GSE são sensíveis ao aumento do parâmetro d , ou seja, à medida que d tem valor próximo ao limite de não-estacionariedade dos processos VE-ARFIMA (d próximo a 0.5), esses métodos falham para estimar adequadamente esse parâmetro.

Para $\phi = 0.9$, as estimativas do parâmetro d com menores valores para EQM e vícios são as obtidas pelo método MQV (para quaisquer representações) e com menores variabilidades as obtidas pelo método GSE.

De forma geral, os métodos GPH e GSE têm melhores resultados quando $\phi = 0.7$ e, as representações AR e MA, com truncamentos iguais a 30, quando $\phi = 0.9$. Em termos do EQM, a representação MA, com truncamento igual a 30, produz melhores estimativas.

Embora o aumento no valor do parâmetro d não provoque muitas mudanças nas variabilidades das estimativas obtidas pelos métodos GPH e GSE, ele aumenta consideravelmente os vícios das mesmas, para quaisquer valores do parâmetro ϕ . Esses métodos super-estimam o parâmetro d .

Pelos resultados da Tabela 5.7, as estimativas obtidas para os parâmetros σ_η^2 e σ_ν^2 , para $\phi = 0.7$, possuem menores valores de dp e EQM quando obtidas pela representação AR, com truncamento igual a 30, exceto para $d = 0.4$. E possuem menores vícios quando obtidas pela representação MA, com truncamento igual a 20, exceto as obtidas pela representação AR, para $d = 0.45$ e truncamento igual a 30. Para $\phi = 0.9$, os menores valores para EQM e vícios são das estimativas obtidas pela representação MA, truncamento igual a 30, exceto para $d = 0.2$, em que a representação AR, com truncamento igual a 30, produz menores EQM e dp.

Em suma, as melhores estimativas para os parâmetros σ_η^2 e σ_ν^2 , para $\phi = 0.7$, são obtidas pela representação AR e, para $\phi = 0.9$, obtidas pela representação MA, ambas com truncamentos iguais a 30.

Pela Tabela 5.8, vemos que, para um mesmo valor do parâmetro d , aumentar o valor do parâmetro ϕ provoca a diminuição das correlações entre os métodos de estimação. E, para um mesmo valor do parâmetro ϕ , com o aumento do parâmetro d , em geral, as correlações entre os métodos aumentam, principalmente para $\phi = 0.7$. E, ainda podemos dizer, que o com aumento simultâneo dos valores de d e ϕ as correlações entre os métodos diminuem.

Constatamos que as estimativas obtidas pelas representações AR e MA são mais correlacionadas entre si e com as obtidas pelos métodos GPH e GSE, quando $\phi = 0.7$. Em particular, as obtidas pela representação AR são mais correlacionadas com as obtidas pelo método GSE, que as obtidas pela representação MA.

Tabela 5.8: Correlações, variâncias e covariâncias das estimativas de d :VE-ARFIMA(1,d,0)

	VE-ARFIMA(1,0.20,0) e $\phi = 0.7$						VE-ARFIMA(1,0.20,0) e $\phi = 0.9$					
	AR20	AR30	MA20	MA30	GPH	GSE	AR20	AR30	MA20	MA30	GPH	GSE
NC	-	-	-	2	-	-	-	-	-	-	-	-
AR20	0.042	0.716	0.502	0.612	0.432	0.639	0.050	0.500	0.279	0.285	-0.009	0.007
AR30	0.032	0.049	0.572	0.668	0.566	0.759	0.023	0.044	0.516	0.483	0.295	0.289
MA20	0.012	0.015	0.015	0.894	0.562	0.601	0.011	0.019	0.031	0.852	0.329	0.184
MA30	0.014	0.017	0.013	0.014	0.660	0.705	0.011	0.018	0.027	0.033	0.218	0.125
GPH	0.013	0.019	0.010	0.012	0.023	0.800	-0.001	0.009	0.009	0.006	0.024	0.780
GSE	0.017	0.021	0.009	0.011	0.015	0.016	0.000	0.008	0.004	0.003	0.016	0.018
	VE-ARFIMA(1,0.30,0) e $\phi = 0.7$						VE-ARFIMA(1,0.30,0) e $\phi = 0.9$					
	AR20	AR30	MA20	MA30	GPH	GSE	AR20	AR30	MA20	MA30	GPH	GSE
AR20	0.040	0.871	0.635	0.702	0.542	0.618	0.054	0.688	0.401	0.445	-0.087*	-0.133*
AR30	0.033	0.036	0.708	0.741	0.584	0.641	0.033	0.044	0.500	0.519	-0.079*	-0.062*
MA20	0.021	0.023	0.028	0.855	0.535	0.602	0.017	0.019	0.032	0.807	-0.036*	-0.122*
MA30	0.022	0.023	0.023	0.026	0.563	0.591	0.019	0.021	0.027	0.036	-0.038*	-0.126*
GPH	0.016	0.016	0.013	0.013	0.022	0.851	-0.003	-0.002	-0.001	-0.001	0.022	0.840
GSE	0.016	0.016	0.013	0.013	0.017	0.018	-0.004	-0.002	-0.003	-0.003	0.019	0.022
	VE-ARFIMA(1,0.40,0) e $\phi = 0.7$						VE-ARFIMA(1,0.40,0) e $\phi = 0.9$					
	AR20	AR30	MA20	MA30	GPH	GSE	AR20	AR30	MA20	MA30	GPH	GSE
NC	-	-	-	2	-	-	-	-	-	-	-	-
AR20	0.032	0.857	0.725	0.811	0.745	0.821	0.077	0.482	0.187	0.235	-0.013*	-0.044*
AR30	0.030	0.038	0.654	0.727	0.728	0.801	0.033	0.059	0.357	0.448	0.132*	0.036*
MA20	0.024	0.024	0.034	0.924	0.632	0.701	0.013	0.022	0.063	0.712	0.226	0.193*
MA30	0.024	0.023	0.028	0.027	0.692	0.747	0.015	0.025	0.041	0.053	0.247	0.187*
GPH	0.018	0.019	0.016	0.015	0.019	0.825	-0.001	0.004	0.008	0.008	0.019	0.804
GSE	0.018	0.019	0.016	0.015	0.014	0.015	-0.001	0.001	0.006	0.005	0.014	0.016
	VE-ARFIMA(1,0.45,0) e $\phi = 0.7$						VE-ARFIMA(1,0.45,0) e $\phi = 0.9$					
	AR20	AR30	MA20	MA30	GPH	GSE	AR20	AR30	MA20	MA30	GPH	GSE
AR20	0.038	0.930	0.693	0.736	0.692	0.810	0.094	0.506	0.216	0.280	-0.128*	-0.164*
AR30	0.037	0.041	0.685	0.780	0.682	0.792	0.043	0.078	0.226	0.404	0.077*	-0.085*
MA20	0.027	0.027	0.039	0.885	0.647	0.736	0.017	0.017	0.071	0.724	0.280	0.244
MA30	0.026	0.028	0.031	0.032	0.647	0.735	0.019	0.026	0.044	0.052	0.335	0.209
GPH	0.021	0.021	0.019	0.017	0.023	0.834	-0.005	0.003	0.011	0.011	0.020	0.796
GSE	0.021	0.021	0.019	0.017	0.016	0.017	-0.006	-0.003	0.008	0.006	0.015	0.017

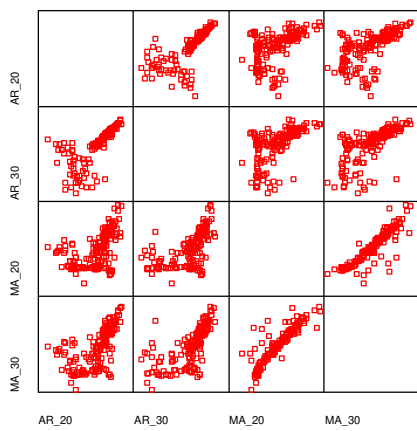
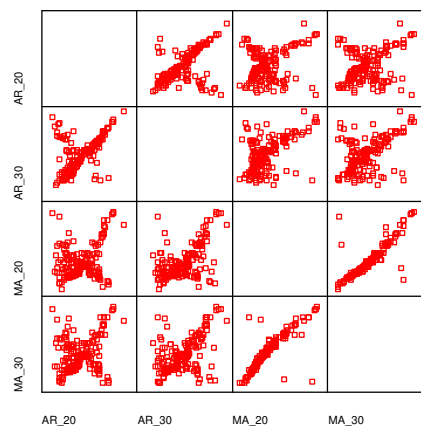
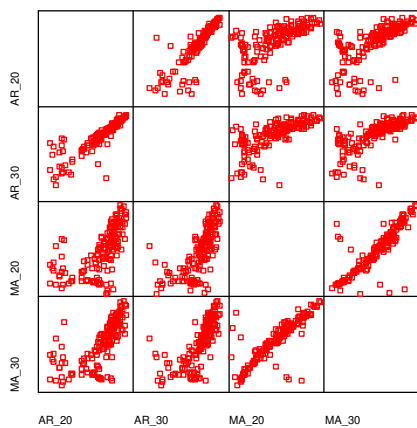
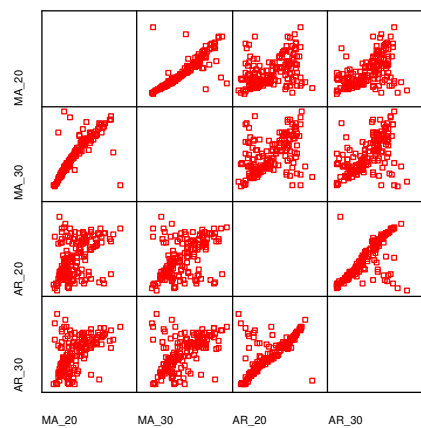
Os valores com asteriscos são as correlações não significativas a 1%, segundo Teste de Pearson. Os valores marcados em negrito são as variâncias, acima destes estão as correlações e abaixo as covariâncias.

NC: Número de não-convergências, em cada uma das representações.

Com o aumento no truncamento, aumenta-se a correlação entre as estimativas obtidas pelas representações AR e MA.

Os métodos GPH e GSE são fortemente correlacionados entre si e devemos lembrar que o GSE é uma modificação do método GPH, assim a alta correlação entre eles é razoável.

Pelos gráficos de dispersão, a seguir, podemos ver a existência de relação linear entre os métodos. Como já vimos pelos valores das correlações e pode-se ver claramente pelos gráficos, da Figura 5.4, que os métodos são mais correlacionados entre si para valor de $\phi = 0.7$. Inclusive, entre as estimativas obtidas pelas representações AR e MA as correlações são maiores para valor de $\phi = 0.7$ e truncamentos iguais a 30.

(a) $d = 0.2$ e $\phi = 0.7$ (b) $d = 0.2$ e $\phi = 0.9$ (c) $d = 0.3$ e $\phi = 0.7$ (d) $d = 0.3$ e $\phi = 0.9$

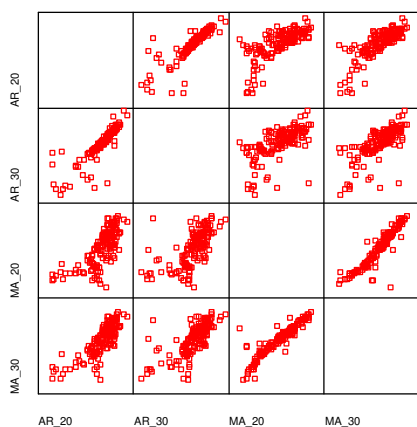
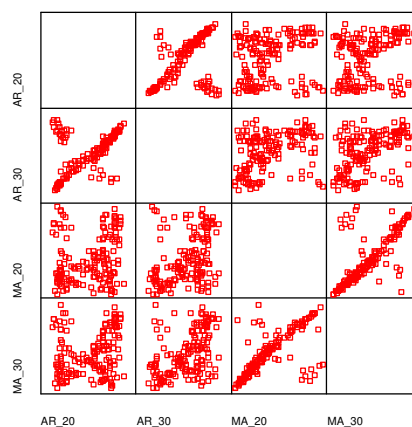
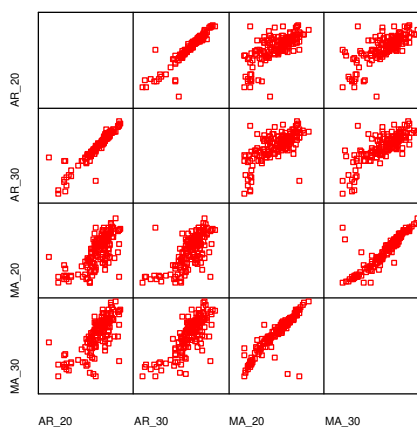
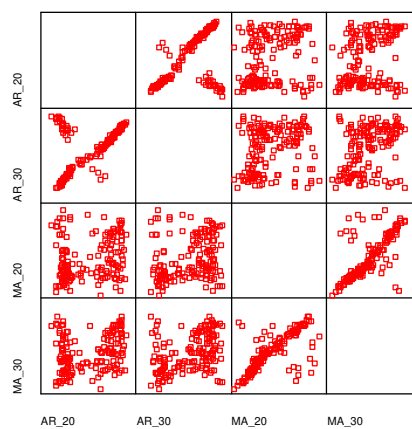
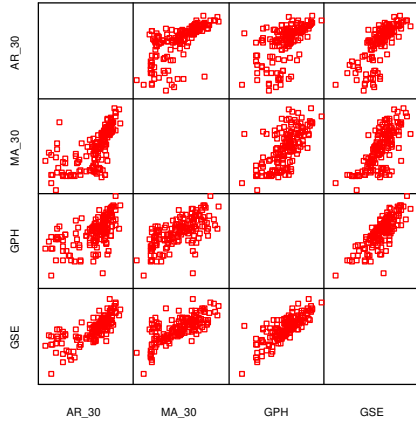
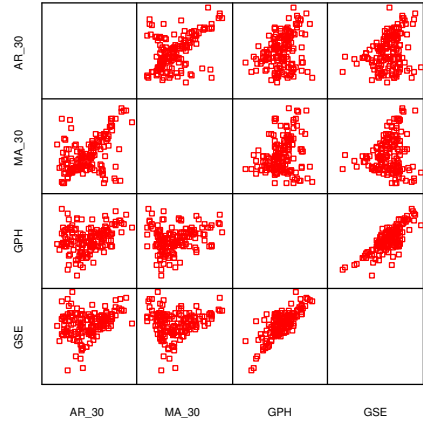
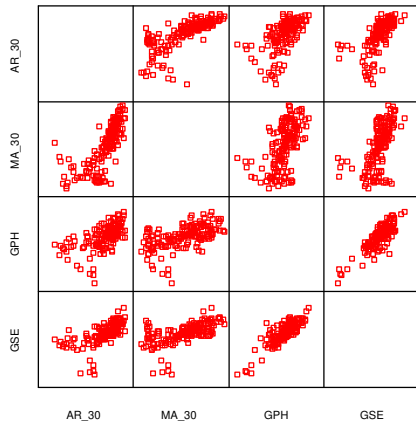
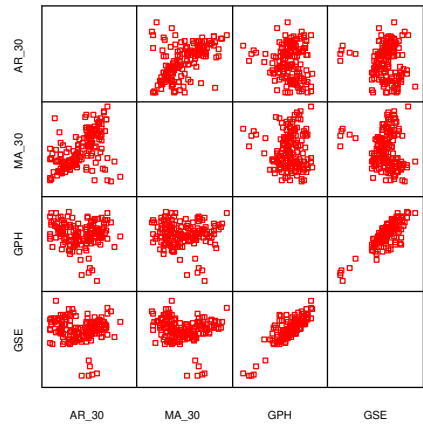
(e) $d = 0.4$ e $\phi = 0.7$ (f) $d = 0.4$ e $\phi = 0.9$ (g) $d = 0.45$ e $\phi = 0.7$ (h) $d = 0.45$ e $\phi = 0.9$

Figura 5.4: Dispersão das Estimativas de d obtidas com as formas AR e MA - VE-ARFIMA(1,d,0)

(a) $d = 0.2$ e $\phi = 0.7$ (b) $d = 0.2$ e $\phi = 0.9$ (c) $d = 0.3$ e $\phi = 0.7$ (d) $d = 0.3$ e $\phi = 0.9$

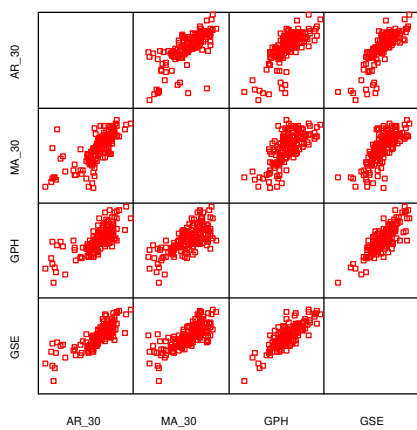
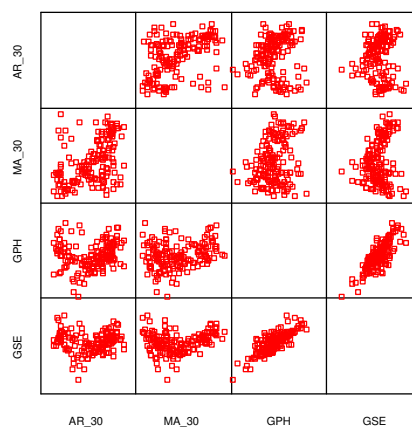
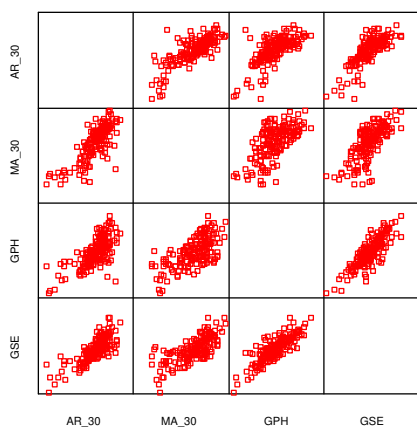
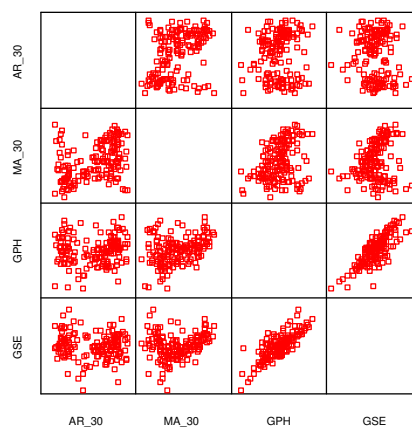
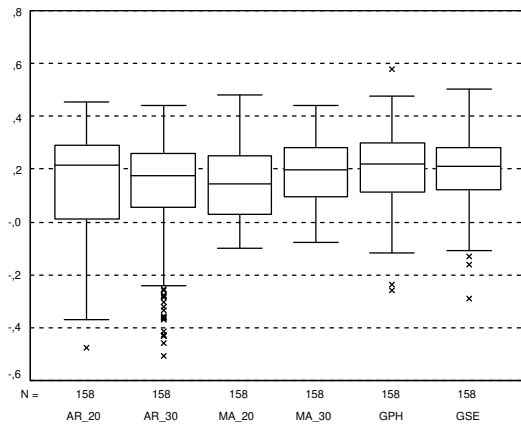
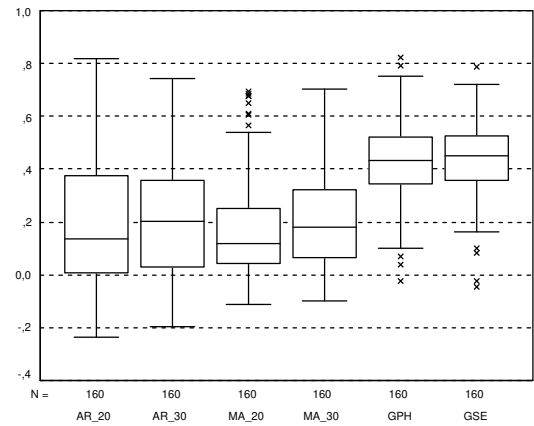
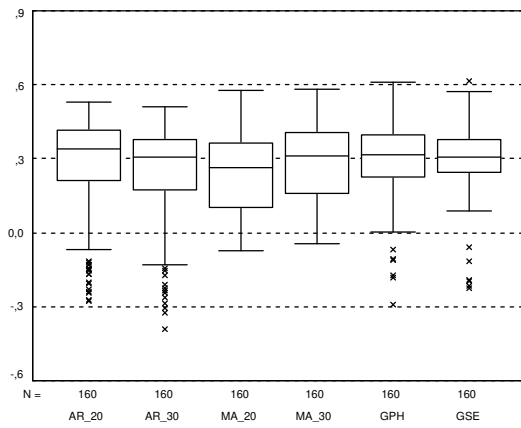
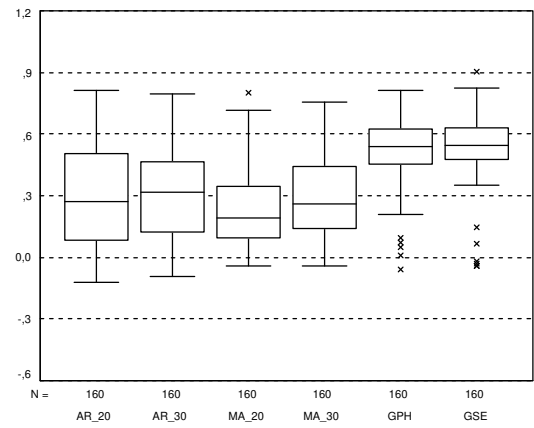
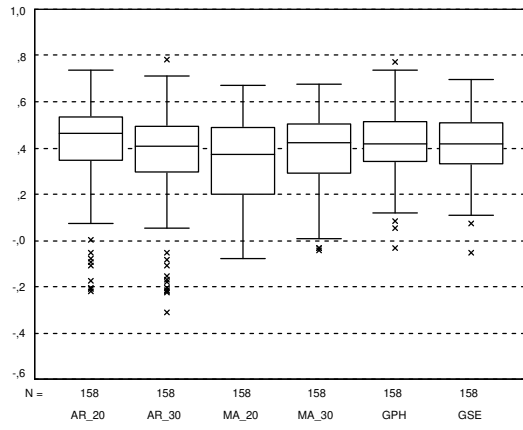
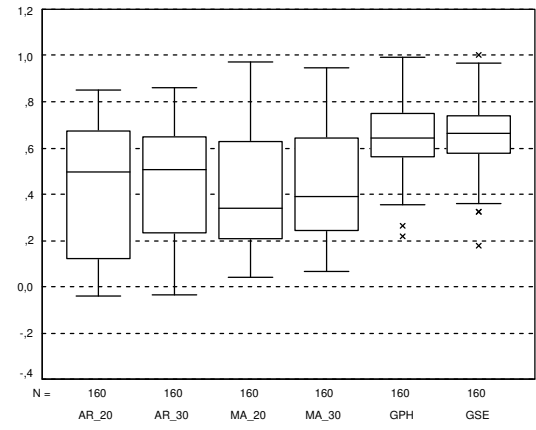
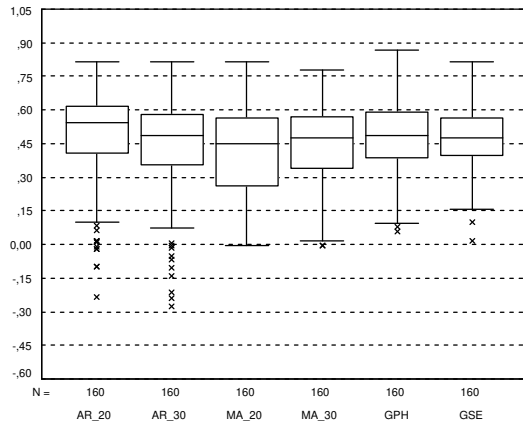
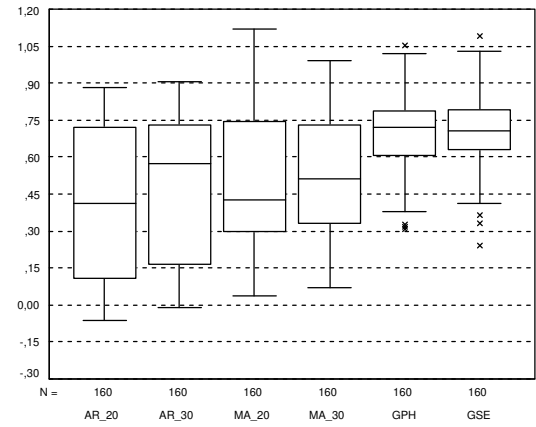
(e) $d = 0.4$ e $\phi = 0.7$ (f) $d = 0.4$ e $\phi = 0.9$ (g) $d = 0.45$ e $\phi = 0.7$ (h) $d = 0.45$ e $\phi = 0.9$

Figura 5.5: Dispersão das Estimativas de d obtidas com as formas AR30, MA30, GPH e GSE - VE-ARFIMA(1,d,0)

(a) $d = 0.2$ e $\phi = 0.7$ (b) $d = 0.2$ e $\phi = 0.9$ (c) $d = 0.3$ e $\phi = 0.7$ (d) $d = 0.3$ e $\phi = 0.9$

(e) $d = 0.4$ e $\phi = 0.7$ (f) $d = 0.4$ e $\phi = 0.9$ (g) $d = 0.45$ e $\phi = 0.7$ (h) $d = 0.45$ e $\phi = 0.9$ Figura 5.6: *Box Plot* das Estimativas de d - VE-ARFIMA(1,d,0)

Observe pelos gráficos dos *Box Plot*, na Figura 5.6, que o aumento do parâmetro ϕ faz as medianas das estimativas obtidas pelos métodos GPH e GSE ficarem bem mais distantes dos verdadeiros valores do parâmetro d , indicando que esses métodos super-estimam o parâmetro. As estimativas obtidas pelo método MQV, com as representações AR e MA, têm medianas próximas aos verdadeiros valores dos parâmetros. Com o aumento do parâmetro ϕ , as estimativas de todos os métodos estão bem mais dispersas.

Diferentemente do que foi observado para os processos VE-ARFIMA(0,d,0), a representação AR produz muitos valores discrepantes que não estão diminuindo com o aumento do valor de d .

5.4 Conclusão

Para o caso dos processos VE-ARFIMA(0,d,0) em qualquer uma das representações, AR ou MA, o método QMV tem melhores estimativas (em termos do EQM) do que os métodos GPH e GSE. Também podemos concluir que tanto pela representação AR quanto pela MA, à medida em que o valor do parâmetro d está próximo ao limite da não-estacionariedade desses processos, obtém-se melhores resultados.

Os métodos GPH e GSE sub-estimam o valor do parâmetro d e suas estimativas tornam-se ainda mais viesadas quando aumenta-se o valor desse parâmetro.

A representação AR produz estimativas com menores vícios, porém com maiores variabilidades, devido à presença de muitos valores discrepantes e a representação MA fornece estimativas com maiores vícios, mas com menores variabilidades, assim, a escolha entre elas pode ser feita em termos do EQM. E, como a representação MA fornece menores valores de EQM e também nos dá a possibilidade de contornar o problema do vício com o aumento do truncamento, ela é preferível.

As correlações entre os métodos aumentam com parâmetro d , sendo as estimativas obtidas pela representação AR mais correlacionadas com as obtidas pelo método GSE.

Para o caso dos processos VE-ARFIMA(1,d,0), entre os métodos utilizados, o que produziu melhores estimativas para o parâmetro de memória d também foi o método de máxima quase-verossimilhança, MQV, utilizando-se a representação MA e truncamento $m = 30$ no operador de diferença fracionária e parâmetro $\phi = 0.9$. Ou seja, o método MQV tem melhores resultados próximos ao limite de não-estacionariedade do que os métodos de GPH e GSE.

Os métodos GPH e GSE super-estimam o parâmetro d , pois produz estimativas com vícios positivos.

As estimativas para o parâmetro d , quando $\phi = 0.9$, têm menores vícios quando obtidas pela representação AR e menores valores de EQM pela representação MA.

Com relação ao parâmetro ϕ , as estimativas com menores variabilidades são obtidas pela representação AR e, com menores vícios, pela representação MA. Para o parâmetro d , obtém-se menores variabilidades, quando obtidas pela representação MA, melhoradas com o aumento do truncamento.

Concluimos que os métodos GPH e GSE são muito sensíveis à presença de não-estacionariedade nas séries e, portanto, não são adequados para estimar esses tipos de processos.

Em suma, para os processos VE-ARFIMA com parâmetro d e ϕ próximos ao limite de não-estacionariedade, ou seja, d próximo a 0.5 e ϕ próximo a 1, a representação MA, com truncamento igual a 30, fornece melhores estimativas.

Durante a estimação não foram utilizadas restrições sobre os valores de d e σ_ν^2 , ($\sigma_\nu^2 = \pi^2/2$, sob a hipótese de normalidade). A primeira restrição, pelo fato de saber que a volatilidade, para a série empírica do tipo analisado, não pode ser explosiva. Estas restrições poderiam melhorar bastante as estimativas obtidas. Os estimadores GPH e GSE podem ser melhorados utilizando-se a modificação sugerida por Deo e Hurvich (2001).

Capítulo 6

Aplicação dos Métodos de Estimação

6.1 Introdução

Neste capítulo, aplicamos os métodos de estimação, apresentados no Capítulo 4, a uma série real de índices de fechamento da BOVESPA (Bolsa de Valores do Estado de São Paulo) observada em intervalos de um minuto e de cinco minutos (das 11 horas do dia 18/12/01 até 18 horas do dia 02/01/02), totalizando 3471 e 696 observações, respectivamente ¹. Para analisar os índices, utilizamos os seus retornos corrigidos (esta correção tem por finalidade evitar problemas de valores indefinidos)².

Ajustamos as séries de retornos por processos ARMA(p,q) com ordens p e q , para valores inferiores a 2. A seleção dos modelos foi feita com base nos valores do AIC (*Akaike Information Criterion*) de cada modelo, do teste de correlações de Ljung-Box Modificado (LBM) e significância dos parâmetros ³.

¹Agradecemos ao *Finance Lab* do *Ibmec Business School* pelo fornecimento dos dados.

²Mais detalhes ver Kim *et al.* (1998).

³Para mais informações sobre a medida AIC e teste de Ljung-Box Modificado, ver Wei (1990).

Ajustamos o logaritmo dos resíduos ao quadrado, do modelo selecionado, por processos VE-ARFIMA(0,d,0) e VE-ARFIMA(1,d,0), com truncamentos, no operador de diferença fracionária $(1 - B)^d$, iguais a 20 e 30.

O capítulo está organizado da seguinte forma. Na Seção 6.2, está a análise das séries dos índices IBOVESPA e na Seção 6.3, apresentamos a conclusão para os resultados obtidos.

6.2 Análise dos Índices IBOVESPA

Os comportamentos, das séries de índices IBOVESPA estudados, podem ser vistos pela Figura 6.1. Observamos nas séries a presença de pontos baixos e altos, bastante evidentes. Pelos dados, constatamos que o menor índice foi observado no dia 20/12/01, com o fim do funcionamento do mercado financeiro e o maior índice observado no dia 27/12/01, com o início das atividades financeiras, após o período do feriado.



(a) Série 1min



(b) Série 5min

Figura 6.1: Séries de Índices IBOVESPA

Para analisar a série de índices IBOVESPA, utilizamos os retornos corrigidos, dados por

$$r_t = 100 * \left\{ y_t - \frac{1}{n} \sum_{i=1}^n y_i \right\} ,$$

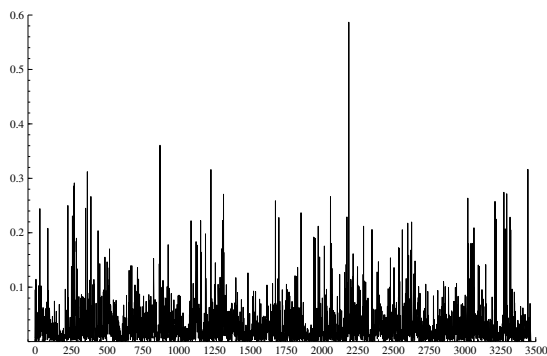
sendo y_t os retornos compostos, como apresentados no Capítulo 3, e n o número de observações.

Na Tabela 6.1, estão algumas estatísticas descritivas para os retornos de cada série considerada, ou seja, série observada em intervalos de um minuto e de cinco minutos, que chamaremos de Série 1min e Série 5min.

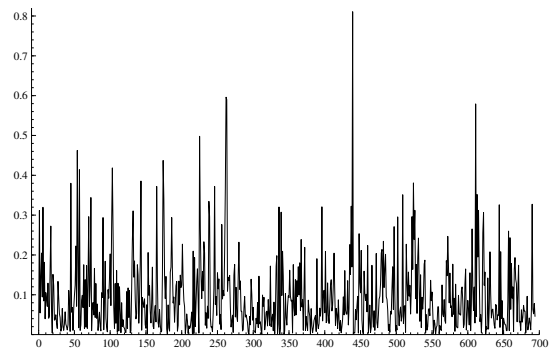
Tabela 6.1: Estatísticas descritivas dos retornos- IBOVESPA

Retornos	Mínimo	Quartil1	Mediana	Média	Quartil3	Máximo	DP	Assimetria	E.Curtose
1MIN	-0.3608	-0.0209	-0.0020	-0.0000	0.0163	0.5868	0.0498	1.1174	13.0345
5MIN	-0.4979	-0.0708	-0.0103	0.0000	0.0572	0.8107	0.1340	0.8899	4.0037

Os retornos, de ambas as séries, têm excessos de curtoses, em relação a distribuição Normal, bastante grandes e portanto, possuem distribuição com caudas mais pesadas que as da distribuição Normal, um dos fatos estilizados apresentados no Capítulo 3.



(a) Série 1min



(b) Série 5min

Figura 6.2: Retornos absolutos das séries de índices IBOVESPA

Pela Figura 6.2, dos retornos absolutos, podemos ver mais claramente os pontos de alta, observados na Figura 6.1. E, pela Figura 6.3, podemos ver que as caudas da distribuição dos retornos são mais pesadas que as da distribuição normal, sendo maior a dispersão dos retornos da Série 5min, como é esperado pelo fato de serem observados em intervalos maiores de tempo que os da Série 1min.

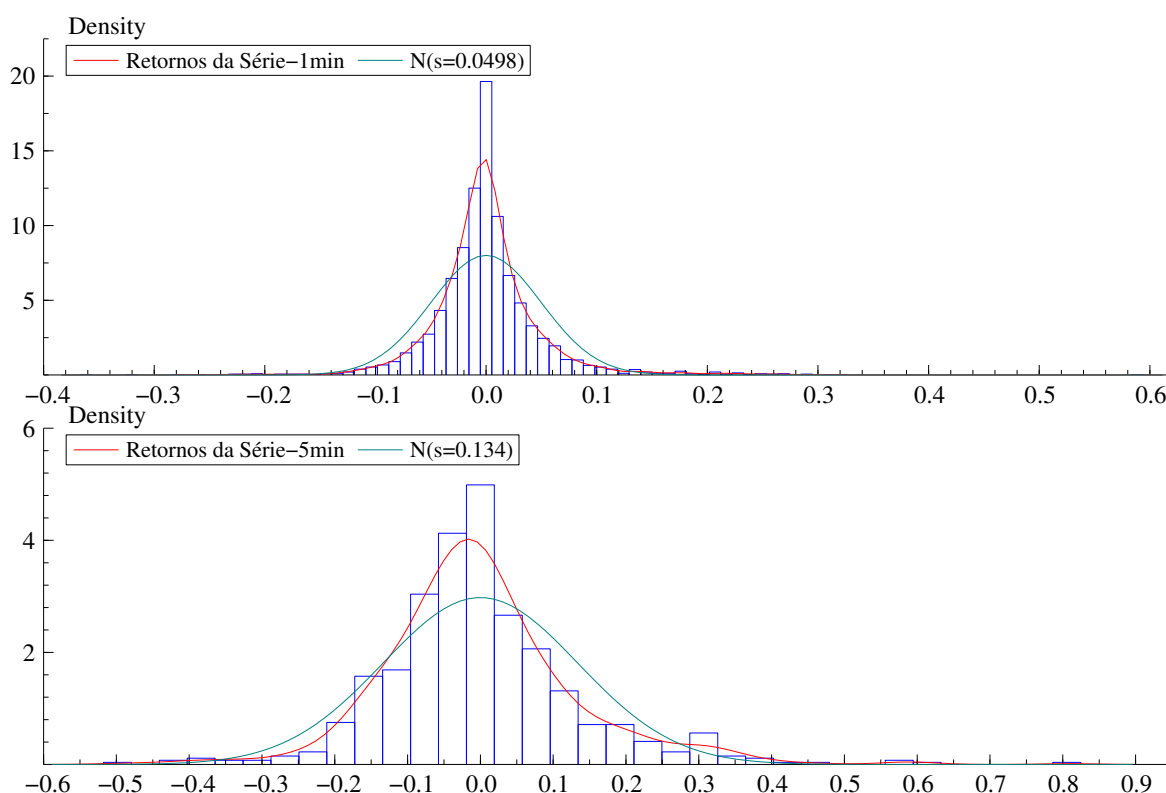
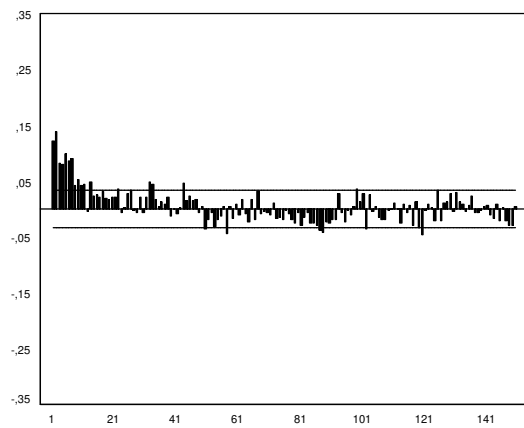
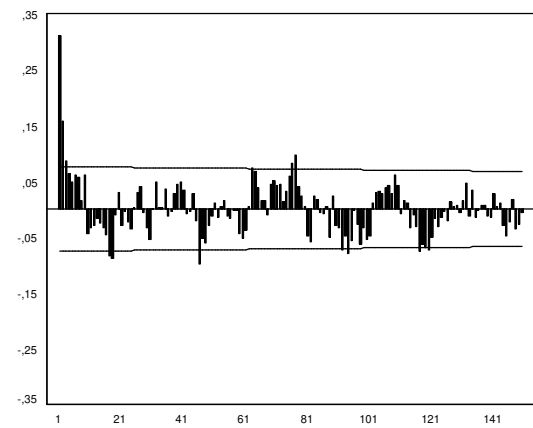


Figura 6.3: Histogramas dos retornos das séries de índices IBOVESPA

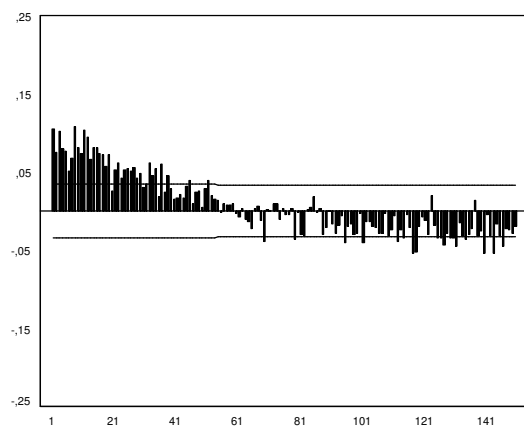


(a) Série 1min

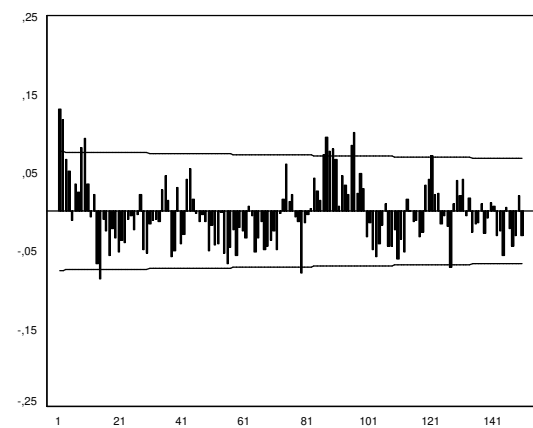


(b) Série 5min

Figura 6.4: Correlograma dos retornos de cada série



(a) Série 1min



(b) Série 5min

Figura 6.5: Correlograma do log dos retornos ao quadrado para cada série

Conforme Figura 6.4 e Figura 6.5, podemos notar que a Série 1min tem comportamento da função de auto-correlação amostral similar a dos processos com memória longa, ou seja, decaimento lento das auto-correlações, o que não observamos, claramente, na Série 5min. Notamos a existência de correlações fora do intervalo de confiança, mesmo para observações distantes, principalmente para os gráficos da Figura 6.5. Devemos lembrar que um dos fatos estilizados dos retornos é que os mesmos não são correlacionados, mas que os retornos ao quadrado (ou absolutos) possuem correlações não desprezíveis. Notamos, pela Figura 6.4, que a primeira correlação da Série 5min é bastante evidente.

Fizemos a modelagem dos retornos das séries IBOVESPA (Série 1min e Série 5min) utilizando os modelos ARMA(p,q), para ordens p e q com valores inferiores a 2. E o modelo mais adequado foi escolhido com base nos valores do AIC, resultados do teste Ljung-Box Modificado e significância dos parâmetros.

Conforme esses valores, selecionamos o modelo ARMA(1,1), como o de melhor ajuste para a Série 1min. Os ajustes pelos processos AR(1) e AR(2) tiveram correlações significativamente diferentes de zero. E, com o aumento das ordens p e q do processo ARMA(1,1), não obtivemos coeficientes adicionais significativos.

Tabela 6.2: Estimativas dos parâmetros para a Série 1min - IBOVESPA

	ARMA(1, 1)			
Parâmetros	Coef	d.p	T	Valor - p
Φ_1	0.8947	0.0243	36.83	0.000
Φ_2	—	—	—	—
Θ_1	0.7997	0.0324	24.69	0.000
Θ_2	—	—	—	—
Constante	0.0000	0.0002	0.00	0.998
Média	0.0000	0.0016	—	—

Coef abrevia coeficientes.

Valor-p é o valor discriminante do teste.

Na Tabela 6.2, apresentamos, por simplicidade, apenas as estimativas obtidas para o modelo selecionado. A Série 1min tem comportamento que pode ser confundido com não-estacionariedade, pois os valores estimados, para os coeficientes do ARMA(1,1), estão próximos a 1. Mas, a não-estacionariedade somente pode ser verificada por meio de testes de raízes unitárias.

O processo ARMA(1,1) ajustado, para a Série 1min, tem média zero (como esperado pois os retornos foram corrigidos) e constante não significativamente diferente de zero. Notamos, pela Tabela 6.3, que, com o aumento das ordens p e q , ainda existem correlações significativamente diferentes de zero ao nível de significância de 5%, mas não a 1%.

Tabela 6.3: Resultados do Teste Ljung-Box Modificado para a Série 1min - IBOVESPA

	AR(1)			AR(2)			ARMA(1,1)		
Lag	χ^2	g.l	Valor-p	χ^2	g.l	Valor-p	χ^2	g.l	Valor-p
12	149.9	10	0.000	63.4	9	0.000	19.1	9	0.024
24	170.5	22	0.000	79.2	21	0.000	25.1	21	0.241
36	191.6	34	0.000	101.4	33	0.000	43.6	33	0.102
48	205.5	46	0.000	114.4	45	0.000	59.0	45	0.078
	ARMA(1,2)			ARMA(2,1)			ARMA(2,2)		
Lag	χ^2	g.l	Valor-p	χ^2	g.l	Valor-p	χ^2	g.l	Valor-p
12	18.3	8	0.019	18.3	8	0.019	16.6	7	0.020
24	24.4	20	0.225	24.4	20	0.227	22.7	19	0.249
36	43.4	32	0.086	43.4	32	0.086	42.1	31	0.089
48	58.9	44	0.065	58.9	44	0.065	57.6	43	0.068

A abreviação g.l. é o número de graus de liberdade para a distribuição qui-quadrado, χ^2 .

Valor-p é o valor discriminante do Teste LBM.

De forma análoga, ajustamos a Série 5min por processos ARMA e, conforme os valores do AIC, do Teste LBM e significância dos parâmetros, também selecionamos o modelo ARMA(1,1) como o de melhor ajuste. O modelo AR(1) poderia ser selecionado mas, para manter uma comparação também entre as séries, utilizamos apenas o modelo ARMA(1,1). O aumento das ordens p e q resultou em coeficientes adicionais não-significativos.

Conforme resultados da Tabela 6.4, temos que os coeficientes do modelo ARMA(1,1), ajustado para a Série 5min, não são tão próximos a 1, como os obtidos para o ARMA(1,1) no ajuste da Série 1min e assim, a Série 5min tem comportamento de uma série estacionária.

Tabela 6.4: Estimativas dos parâmetros para a Série 5min - IBOVESPA

$ARMA(1, 1)$				
Parâmetros	Coef	d.p	T	Valor-p
Φ_1	0.5295	0.0986	5.37	0.000
Φ_2	—	—	—	—
Θ_1	0.2447	0.1127	2.17	0.030
Θ_2	—	—	—	—
Constante	0.0000	0.0036	0.00	0.998
Média	0.0000	0.0077	—	—

Coef abrevia coeficientes.

Valor-p é o valor discriminante do teste.

Tabela 6.5: Resultados do Teste LBM para a Série 5min - IBOVESPA

AR(1)				AR(2)			ARMA(1,1)		
Lag	χ^2	g.l	Valor-p	χ^2	g.l	Valor-p	χ^2	g.l	Valor-p
12	14.5	10	0.150	10.2	9	0.331	9.8	9	0.369
24	25.7	22	0.265	22.0	21	0.401	21.5	21	0.430
36	34.6	34	0.440	32.1	33	0.510	31.8	33	0.527
48	46.7	46	0.443	44.7	45	0.483	44.4	45	0.495
ARMA(1,2)				ARMA(2,1)			ARMA(2,2)		
Lag	χ^2	g.l	Valor-p	χ^2	g.l	Valor-p	χ^2	g.l	Valor-p
12	9.7	8	0.290	9.4	8	0.313	9.2	7	0.239
24	21.0	20	0.395	20.7	20	0.416	20.6	19	0.358
36	31.1	32	0.510	30.7	32	0.533	30.7	31	0.481
48	43.8	44	0.479	43.5	44	0.494	43.5	43	0.449

A abreviação g.l. é o número de graus de liberdade para a distribuição qui-quadrado, χ^2 .

Valor-p é o valor discriminante do Teste LBM.

Na Figura 6.6, apresentamos o comportamento dos resíduos para os modelos ARMA(1,1) ajustados para as Série 1min e Série 5min. Existem pontos de picos correspondentes aos períodos que mencionamos no início desta seção.

Pelos histogramas da Figura 6.7, vemos que os resíduos têm distribuição com caudas mais pesadas que as da distribuição Normal, mais evidentes para a Série 1min. Obviamente, a Série 1min tem valores mais concentrados em torno de zero que a Série 5min, pois é observada em intervalos de tempo menores, portanto, com pouca variabilidade.

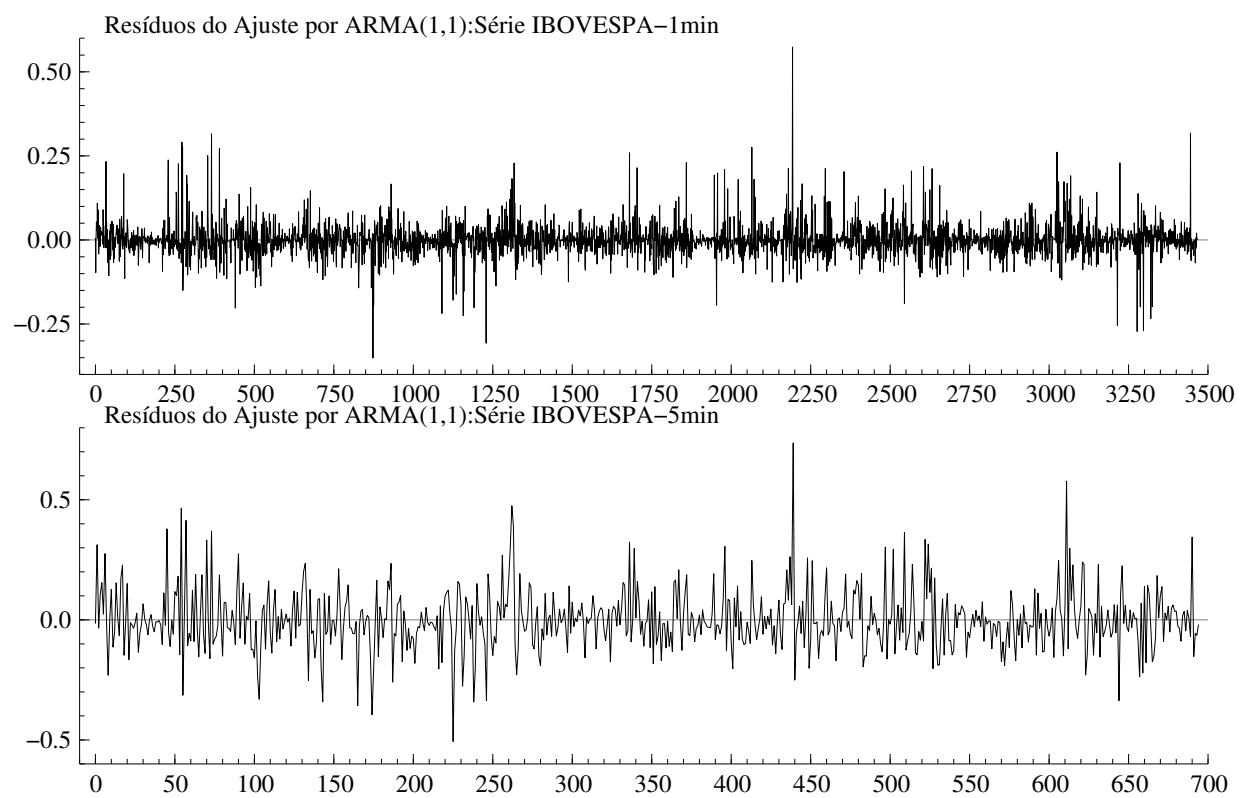


Figura 6.6: Resíduos dos ajustes por ARMA(1,1), para a Série 1min e Série 5min

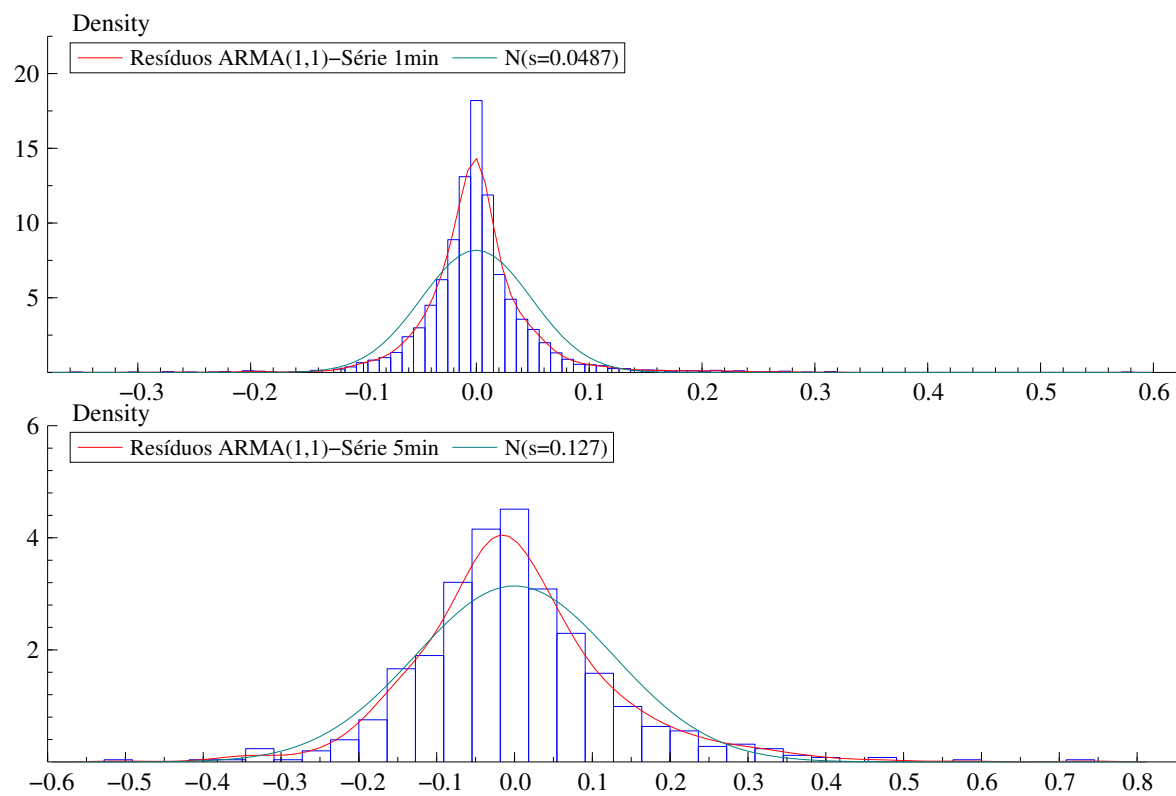


Figura 6.7: Histogramas dos resíduos dos ajustes por ARMA(1,1), para a Série 1min e Série 5min

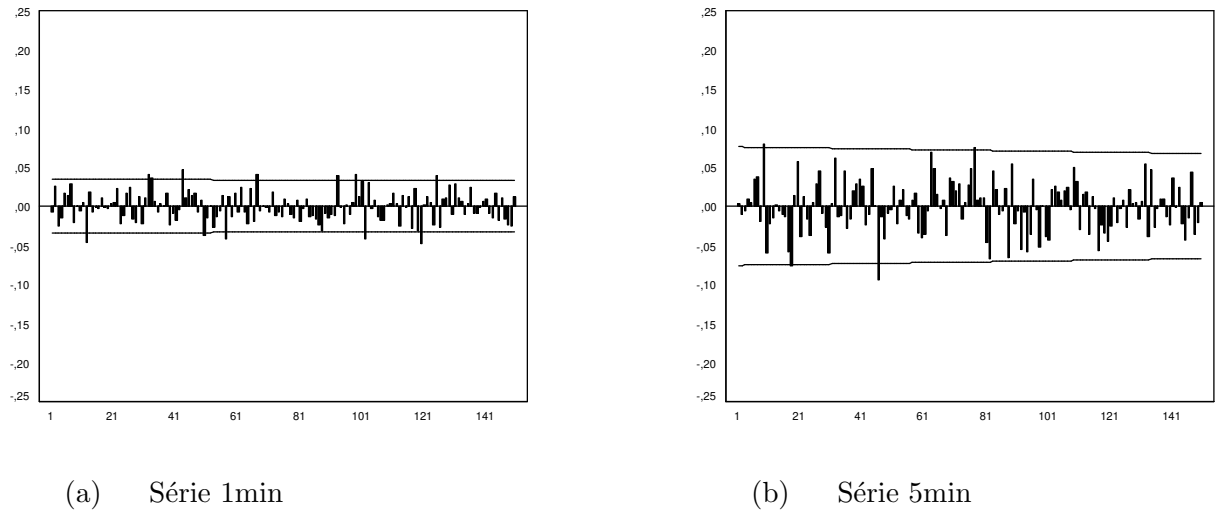


Figura 6.8: Auto-correlações dos resíduos dos ajustes por $ARMA(1,1)$, para a Série 1min e Série 5min

Conforme Figura 6.8, vemos que as correlações são muito pequenas, praticamente todas estão dentro dos intervalos de confiança, indicando que os resíduos são não-correlacionados, como suposto para os processos ARMA.

Ajustamos o logaritmo dos resíduos ao quadrado dos modelos $ARMA(1,1)$ (ajustados para ambas as séries, Série 1min e Série 5min) por processos $VE-ARFIMA(0,d,0)$ e $VE-ARFIMA(1,d,0)$.

A Tabela 6.6 tem as estimativas obtidas para os parâmetros do ajuste da Série 1min e podemos notar que o parâmetro d é estimado pelo método MQV pelas representações AR e MA (truncamentos em 20 e 30) com valores próximos ao valor limite da não-estacionariedade dos processos $VE-ARFIMA$, ou seja, próximo a 0.5, enquanto que os métodos GPH e GSE fornecem estimativas com valores muito maiores que este limite, com desvios-padrão muito pequenos. Conforme resultados das simulações, sabemos que estes dois últimos métodos não conseguem estimar adequadamente o parâmetro d para processos com comportamento de não-estacionariedade e que, nestes casos, o método MQV fornece melhores resultados.

Pela Tabela 6.7, as estimativas de d são próximas a 0.5 e coeficiente ϕ estimado com valores negativos pequenos.

Além de estimar o parâmetro de memória d , o método MQV também estima os demais parâmetros do modelo e, pelos resultados das simulações, os valores estimados estariam bem próximos aos verdadeiros valores dos parâmetros do processo, em particular, melhores obtidos pela representação MA, com truncamento, no operador de diferença fracionária, igual a 30.

Tabela 6.6: Estimativas para os parâmetros do VE-ARFIMA(0,d,0)-Série 1min

		Máxima Quase-Verossimilhança						GPH		GSE	
Repres	m	d	dp	σ_η^2	dp	σ_ν^2	dp	d	dp	d	dp
AR	20	0.548	0.060	0.564	0.215	5.832	0.252	0.835	0.000	0.816	0.000
	30	0.501	0.059	0.732	0.269	5.689	0.287	—	—	—	—
MA	20	0.621	0.084	0.449	0.223	5.933	0.256	—	—	—	—
	30	0.616	0.083	0.467	0.228	5.912	0.258	—	—	—	—

Tabela 6.7: Estimativas para os parâmetros do VE-ARFIMA(1,d,0)-Série 1min

		Máxima Quase-Verossimilhança							
Repres	m	d	dp	ϕ	dp	σ_η^2	dp	σ_ν^2	dp
AR	20	0.540	0.064	-0.120	0.197	0.755	0.455	5.689	0.407
	30	0.487	0.066	-0.129	0.146	1.041	0.565	5.446	0.507
MA	20	0.613	0.089	-0.125	0.241	0.603	0.445	5.825	0.389
	30	0.606	0.089	-0.142	0.218	0.657	0.459	5.776	0.403

Para o caso da Série 5min, como verificamos em análise anterior, ela apresenta comportamento de presença de estacionariedade e, assim, os métodos GPH e GSE forneceriam estimativas melhores que as obtidas pelo método MQV. Pela Tabela 6.8, as estimativas obtidas, para o parâmetro d , pelos métodos MQV, GPH e GSE possuem valores opostos. Os métodos GPH e GSE têm estimativas com maiores desvios-padrão do que as do método MQV, com as representações AR e MA.

Conforme Tabela 6.9 e resultados das simulações, do Capítulo 5, podemos dizer que as estimativas para d obtidas pelo método MQV, com a representação MA, truncamento igual a 30, próximas a 0.1 e de ϕ próximas a 0.5, indicam estacionariedade.

Tabela 6.8: Estimativas para os parâmetros do VE-ARFIMA(0,d,0)-Série de 5min

		Máxima Quase-Verossimilhança						GPH		GSE	
Repres	m	d	dp	σ_η^2	dp	σ_ν^2	dp	d	dp	d	dp
AR	20	0.1464	0.0779	3.5189	2.1783	2.6631	2.1107	-0.1590	0.2992	-0.0984	0.3155
	30	0.1436	0.0811	3.4527	2.2991	2.7341	2.2308	—	—	—	—
MA	20	0.1220	0.0602	4.5035	2.2597	1.7307	2.1972	—	—	—	—
	30	0.1368	0.0708	4.0052	2.2595	2.2080	2.1916	—	—	—	—

Tabela 6.9: Estimativas para os parâmetros do VE-ARFIMA(1,d,0)-Série de 5min

		Máxima Quase-Verossimilhança							
Repres	m	d	dp	ϕ	dp	σ_η^2	dp	σ_ν^2	dp
AR	20	0.156	0.182	0.229	0.566	1.361	1.376	4.673	1.121
	30	-0.460	0.214	0.925	0.049	0.849	0.698	5.098	0.711
MA	20	0.046	0.116	0.532	0.278	0.793	0.705	5.133	0.678
	30	0.087	0.154	0.477	0.355	0.825	0.807	5.109	0.743

Apresentamos abaixo o gráfico da função de auto-correlação empírica (correlograma) para os resíduos do ARMA(1,1), que ajusta a Série 5min, juntamente com o gráfico da função de auto-correlação teórica do VE-ARFIMA(0,d,0). Para escrever a função de auto-correlação teórica utilizamos o valor médio das estimativas, para o parâmetro d , obtidas pelas representações AR e MA. No caso dos métodos GPH e GSE, foi considerada a média das estimativas obtidas pelos dois.

Os valores estimados pelos métodos MQV, GPH e GSE são iguais em magnitudes, mas têm sinais opostos, e observamos, nos Gráficos 6.9 e 6.10, que as funções de auto-correlações têm sentidos opostos.

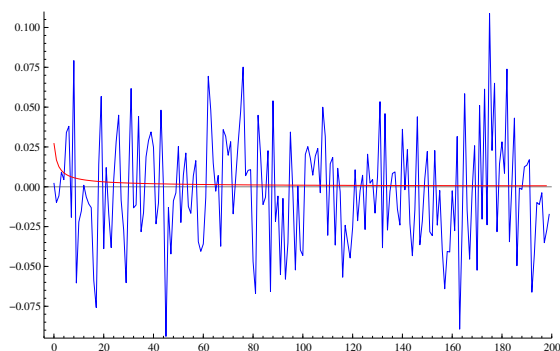


Figura 6.9: ACF-Empírica dos resíduos do ARMA(1,1) que Ajusta a série de 5min e Teórica do processo VE-ARFIMA(0,d,0) com a média das estimativas para d por QMV

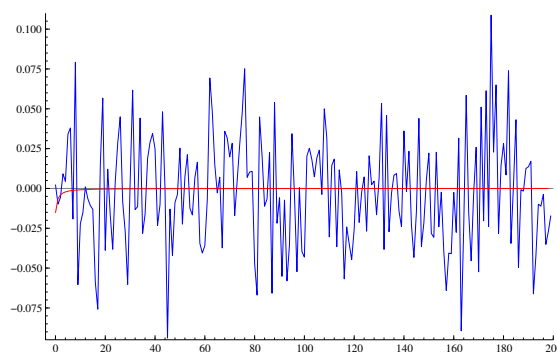
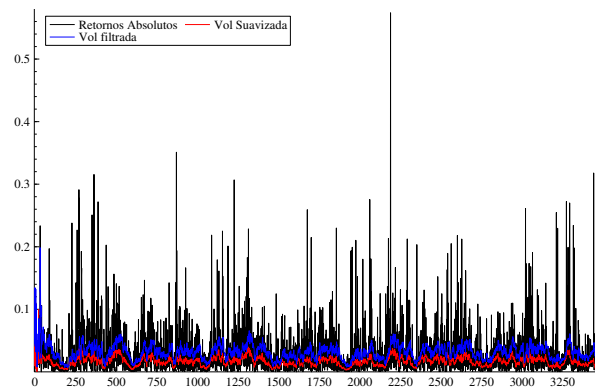
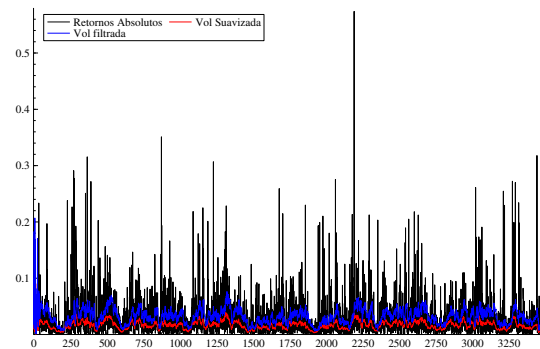


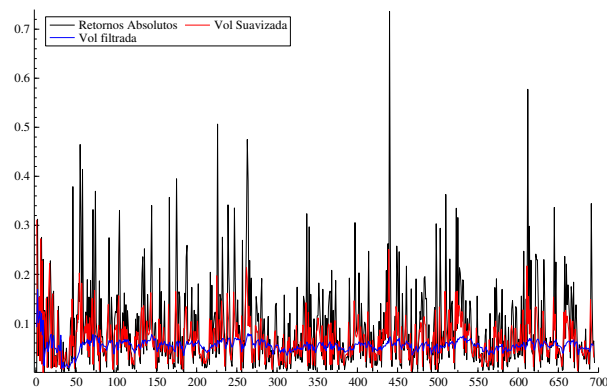
Figura 6.10: ACF-Empírica dos resíduos do ARMA(1,1) que Ajusta a série de 5min e Teórica do processo VE-ARFIMA(0,d,0) com a média das estimativas para d por GPH e GSE



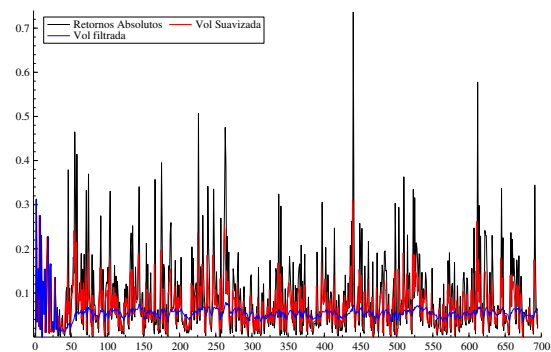
(a) Repres. AR para Série 1min



(b) Repres. MA para Série 1min



(c) Repres. AR para Série 5min



(d) Repres. MA para Série 5min

Figura 6.11: Resíduos absolutos, volatilidade filtrada e suavizada - VE-ARFIMA(0,d,0)

Os gráficos, na Figura 6.11, são para avaliarmos os comportamentos das volatilidades filtradas e suavizadas, com relação aos resíduos absolutos, para o caso do VE-ARFIMA(0,d,0), em cada uma das representações, AR e MA.

A alta frequência dos dados estudados torna um pouco difícil a visualização, mas ainda assim percebemos que as volatilidades acompanham os comportamentos dos resíduos absolutos.

6.3 Conclusão

Nesta dissertação, utilizamos para a estimação dos modelos de volatilidade estocástica com memória longa, em particular dos modelos VE-ARFIMA, o método paramétrico baseado na maximização da função de quase-verossimilhança aproximada, calculada via Filtro de Kalman, utilizando-se a formulação de espaço de estados. Comparamos as estimativas, assim obtidas, para o parâmetro d , com as obtidas pelos métodos semiparamétricos GPH e GSE, baseados na regressão do log-periodograma.

Os processos VE-ARFIMA(0,d,0) e ARFIMA(1,d,0) utilizados foram aproximados nas formas AR e MA, com truncamentos no operador de diferença fracionária iguais a 20 e 30.

Constatamos, por meio de simulações e aplicação a uma série real de índices IBOVESPA, que os métodos GPH e GSE são bastante sensíveis na presença de não-estacionariedade nas séries, produzindo estimativas com muitos vícios. Sendo que o método GSE fornece estimativas com menores vícios que o GPH.

Assim, processos com comportamentos não-estacionariedade não seriam, adequadamente, estimados pelos métodos GPH e GSE, sendo mais aconselhável o uso do método MQV (Máxima Quase-Verossimilhança).

O método MQV produz, em geral, melhores estimativas que os métodos GPH e GSE, quando utiliza-se a representação MA, com truncamento igual a 30 e valores para os parâmetros d e ϕ próximos ao limite de não-estacionariedade dos processos VE-ARFIMA, ou seja, d próximo a 0.5 e ϕ próximo a 1.

Para a estimação do parâmetro ϕ , o método MQV, com a representação AR, geralmente, é melhor (em termos do EQM, desvio-padrão e viés).

Na prática, sabemos que as séries são, em sua maioria, não-estacionárias e portanto o método MQV é mais adequado. Outra vantagem desse método é que ele pode ser generalizado para processos mais complexos, como os multivariados, além de também ser possível estimar todos os parâmetros de processos que possam ser escritos na forma de espaço de estados.

O método MQV, que utilizamos, apesar de suas limitações, pois assumimos perturbações gaussianas e ainda utilizamos função de quase-verossimilhança aproximada no truncamento do operador de diferença fracionária, forneceu resultados pelas séries simuladas bastante satisfatórios, que motivam o estudo dos processos VE-ARFIMA com a adição de outras componentes como a sazonalidade, a tendência e etc. Além de motivar, também, a correção da função de quase-verossimilhança com o uso da amostragem por importância.

Apêndice A

Método de Estimação por Máxima Verossimilhança

Neste apêndice, apresentamos, resumidamente, a metodologia descrita em Durbin e Koopman (2001a), que utiliza amostragem por importância para a correção da função de quase-verossimilhança e variáveis antitéticas e de controle, para melhorar a simulação.

Considere o modelo na forma de espaço de estados dada em (3.23), concentrando-se em θ_t (também chamado de sinal), em vez do vetor completo de estados α_t , para obter ganhos computacionais.

Considere $p(y/\psi)$, $p(\theta/\psi)$, $p(y, \theta/\psi)$ e $p(y/\theta, \psi)$ as densidades marginais de y e θ , a conjunta de y e θ , a condicional de y dado θ , respectivamente, todas condicionadas a ψ (vetor de parâmetros desconhecidos) e $\theta = (\theta'_1, \dots, \theta'_n)'$.

A função de verossimilhança pode ser escrita como

$$L(\psi) = p(y/\psi) = \int p(y, \theta/\psi) d\theta = \int p(y/\theta, \psi) p(\theta/\psi) d\theta \quad . \quad (1.1)$$

Para obter ganhos computacionais na estimação de $L(\psi)$, utilizam-se técnicas de amostragem por importância, variáveis antitéticas e variáveis de controle.

Para usar a técnica de amostragem por importância, considera-se que $\tilde{p}(\theta/y, \psi)$ seja uma densidade condicional arbitrária, a qual é definida positiva para todo θ condicionado a y e ψ .

Então, pode-se reescrever (1.1) como sendo

$$L(\psi) = \int p(y/\theta, \psi) \frac{p(\theta/\psi)}{\tilde{p}(\theta/y, \psi)} \tilde{p}(\theta/y, \psi) d\theta = \tilde{E} \left\{ p(y/\theta, \psi) \frac{p(\theta/\psi)}{\tilde{p}(\theta/y, \psi)} \right\} \quad , \quad (1.2)$$

em que $\tilde{E}$ é a esperança utilizando-se a densidade de importância $\tilde{p}(\theta/y, \psi)$.

Considerando-se N escolhas independentes $\theta^{(i)}$ da densidade de $\tilde{p}(\theta/y, \psi)$, estima-se $L(\psi)$ por

$$\hat{L}(\psi)_1 = \frac{1}{N} \sum p(y/\theta^{(i)}, \psi) \frac{p(\theta^{(i)}/\psi)}{\tilde{p}(\theta^{(i)}/y, \psi)} \quad . \quad (1.3)$$

A eficiência melhora à medida em que se tem $\tilde{p}(\theta/y, \psi)$ mais próxima de $p(\theta/y, \psi)$. Uma forma elegante é desenvolvida calculando-se a verossimilhança $L_g(\psi)$ para um aproximado modelo linear gaussiano pela rotina de Filtro de Kalman e, depois, calcula-se a verdadeira verossimilhança $L(\psi)$ com uma correção, obtida por simulação.

O modelo aproximado é obtido assumindo-se que as observações y_t são geradas pela equação linear de espaço de estados, em termos de ϵ_t , $t=1, \dots, n$. Para permitir mais flexibilidade do que a da equação linear, considera-se que $\epsilon_t \sim N(\mu_t, H_t)$, μ_t e H_t são encontrados por simulação.

Considera-se que $p(\cdot)$, $p(\cdot, \cdot)$ e $p(\cdot/\cdot)$ sejam, respectivamente, as densidade marginais, conjunta e condicional de um modelo não-gaussiano e que $g(\cdot)$, $g(\cdot, \cdot)$ e $g(\cdot/\cdot)$, para o caso de um modelo aproximado gaussiano. Assim, a verossimilhança para o modelo não-gaussiano é dada por

$$L(\psi) = p(y) = \int p(y, \theta) d\theta = \int p(y/\theta) p(\theta) d\theta \quad , \quad (1.4)$$

e, para o caso do modelo aproximado gaussiano, dada por

$$L_g(\psi) = g(y) = \frac{g(y, \theta)}{g(\theta/y)} = \frac{g(y/\theta) p(\theta)}{g(\theta/y)} \quad . \quad (1.5)$$

Assim, substituindo-se em (1.4) a $p(\theta)$ obtida em (1.5), tem-se

$$L(\psi)L_g(\psi) = \int \frac{p(y/\theta)}{g(y/\theta)} g(\theta/y) d\theta = L_g(\psi) E_g \left\{ \frac{p(y/\theta)}{g(y/\theta)} \right\} , \quad (1.6)$$

em que E_g é a esperança utilizando-se a densidade gaussiana $g(\theta/y)$.

A expressão (1.6) é bem mais simples do que a dada em (1.2) e tem a vantagem de somente ser necessária a estimação do desvio das verossimilhanças, da verossimilhança Gaussiana, e não a estimação da verossimilhança completa.¹

A eficiência depende da escolha do modelo aproximado utilizado para fazer a correção da verossimilhança. Esta correção é obtida utilizando-se o suavizador de simulação (*Simulation Smoother*) de de Jong e Shephard (1995).

Seleção de Modelo Aproximado

Apresentamos a seguir uma maneira de escolher μ_t e H_t para o caso do modelo aproximado linear gaussiano.

Considere-se que $\hat{\theta} = E_g(\theta)$ é o valor suavizado de θ , obtido com a aplicação do Filtro e Suavizamento de Kalman, sob a suposição de que o modelo aproximado está correto. Com a simulação suavizada, os valores simulados de θ_t são normalmente distribuídos com média $\hat{\theta}$. Assim, escolhe-se μ_t e H_t tais que $p(y/\theta)$ e $g(y/\theta)$ sejam próximas na vizinhança de $\hat{\theta}$.

Considere-se que

$$\ell(\theta) = \log p(y/\theta) - \log g(y/\theta) . \quad (1.7)$$

Escolhe-se, segundo Durbin e Koopman (1997), μ_t e H_t tais que $\partial \ell(\theta)/\partial \theta = 0$ e $\partial^2 \ell(\theta)/\partial \theta \partial \theta' = 0$ em $\theta = \hat{\theta}$.

¹Similares eficiências não são obtidas no caso de a densidade de importância não ser gaussiana.

Visto que

$$\ell(\theta) = \sum_{t=1}^n \{\log(p(y_t/\theta_t)) - \log(g(y_t/\theta_t))\} \quad , \quad (1.8)$$

em que

$$\log g(y_t/\theta_t) = a + 0.5(y_t - \theta_t - \mu_t)' H_t^{-1} (y_t - \theta_t - \mu_t) \quad , \quad (1.9)$$

as equações para μ_t e H_t são obtidas por

$$\left. \frac{\partial \log p(y_t/\theta_t)}{\partial \theta_t} \right|_{\theta_t = \hat{\theta}_t} - H_t^{-1} (y_t - \hat{\theta}_t - \mu_t) = 0 \quad (1.10)$$

$$\left. \frac{\partial^2 \log p(y_t/\theta_t)}{\partial \theta \partial \theta'} \right|_{\theta_t = \hat{\theta}_t} + H_t^{-1} = 0 \quad . \quad (1.11)$$

As equações (1.10) e (1.11) são resolvidas, iterativamente, com valores iniciais $\mu_t = 0$ e H_t arbitrário. Para cada par de μ_t e H_t , um novo valor de $\hat{\theta}_t$ é obtido por Filtro e Suavizador de Kalman.

É importante notar que a solução para as expressões (1.10) e (1.11), com H_t definida positiva, sempre existe no caso de observações da família exponencial, embora obtenham-se válidas soluções para observações com erros que tenham, por exemplo, distribuição t de Student.

É possível melhorar a estimação utilizando-se variáveis antitéticas e de controle, para mais detalhes sobre essa abordagem, veja trabalhos de Motta e Hotta (2001a), Durbin e Koopman(2001).

Bibliografia

- [1] Agiakloglou, C.; Newbold, P. e Wohar, M. (1993). Bias in an estimator of the fractional difference parameter. *Journal of Time Series Analysis*, **14**(3): 235-246.
- [2] Alizadeh, S.; Brandt, M. W. e Diebold, F. X. (2002). Range-based estimation of stochastic volatility models. *The Journal of Finance*, **LVII**(3): 1047-1091.
- [3] Andersen, T. G. e Bollerslev, T. (1997). Intraday periodicity and volatility persistence in financial markets. *Journal of Empirical Finance*, **4**: 115-158.
- [4] Andersen, T. G.; Chung, H-J. e Sorensen, B. E. (1999). Efficient method of moments estimation of a stochastic volatility model: a Monte Carlo study. *Journal of Econometrics*, **91**: 61-87.
- [5] Andersen, T. G.; Bollerslev, T. e Das, A. (2001). Variance-ratio statistics and high-frequency data: testing for changes in intraday volatility patterns. *The Journal of Finance*, **LVI**(1): 305-327.
- [6] Anderson, J. (2002). An improvement of the GPH estimator. *Economics Letters*, **77**: 137-146.
- [7] Baillie, R. T. (1996). Long memory processes and fractional integration in econometrics. *Journal of Econometrics*, **73**: 5-59.
- [8] Baillie, R. T.; Bollerslev, T. e Mikkelsen, H. O. (1996). Fractionally integrated generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, **74**: 3-30.
- [9] Beltrao, K. I. (1987). Determining the bandwidth of a kernel spectrum estimate. *Journal of Time Series Analysis*, **18**(1): 21-38.
- [10] Beran, J. (1994). *Statistics for Long-Memory Processes*. New York: Chapman & Hall.

- [11] Beran, J. (1995). Maximum likelihood estimation of the differencing parameter for invertible short and long memory autoregressive integrated moving average models. *Journal Royal Statistics Society, B*, **57**(4): 659-672.
- [12] Bisaglia, L. e Guégan, D. (1998). A comparison of techniques of estimation in long-memory processes. *Computational Statistics & Data Analysis*, **27**: 61-81.
- [13] Black, R. (1986). Noise. *Journal of Finance*, **3**: 529-543.
- [14] Blair, B. J.; Poon, S-H. e Taylor, S. J. (2001). Forecasting *S&P*100 volatility: the incremental information content of implied volatilities and high-frequency index returns. *Journal of Econometrics*, **105** : 5-26.
- [15] Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, **31**: 307-327.
- [16] Bollerslev, T. e Mikkelsen, H. O. (1996). Modelling and pricing long-memory in stock market volatility. *Journal of Econometrics*, this issue.
- [17] Bollerslev, T. e Mikkelsen, H. O. (1999). Long-term equity anticipation securities and stock market volatility dynamics. *Journal of Econometrics*, **92**: 75-99.
- [18] Bollerslev, T. e Wright, J. H. (2000). Semiparametric estimation of long memory volatility dependencies: the role of high-frequency data. *Journal of Econometrics*, **98**: 81-106.
- [19] Box, G. E. P. e Jenkins, G. M. (1976). *Time Series Analysis: Forecasting and Control*. Revised Edition. Holden Day, San Francisco, CA.
- [20] Breidt, F. J.; Crato, N. e Lima, Pedro de (1998). The detection and estimation of long memory in stochastic volatility. *Journal of Econometrics*, **83**: 325-348.
- [21] Brockwell, P. J. e Davis, R. A. (1987). *Time Series Analysis: Theory and Methods*. New York: Springer.
- [22] Brockwell, P. J. e Davis, R. A. (1991). *Time Series: Theory and Methods*. New York: Springer.
- [23] Brockwell, P. J. e Davis, R. A. (1996). *Introduction to Time Series and Forecasting*. New York: Springer-Verlag.
- [24] Broto, C. e Ruiz, E. (2002). Estimation methods for stochastic volatility models: a survey. *Preprint*: 1-49.

- [25] Chan, N. H. e Palma, W. (1998). State space modelling of long-memory processes. *Annals of Statistics*, **26**(2): 719-740.
- [26] Chan, N. H. e Petris, G. (2000). Long memory stochastic volatility: a bayesian approach. *Communications in Statistics. Theory and Methods*, **29**: 1367-1378.
- [27] Chen, G.; Abraham, B. e Peiris, S. (1994). Lag window estimation of the degree of differencing in fractionally integrated time series models. *Journal of Time Series Analysis*, **15**(5): 473-487.
- [28] Chib, S., Nardari, F. e Shephard, N. (2002). Markov chain Monte Carlo methods for stochastic volatility models. *Journal of Econometrics*, **108**: 281-316.
- [29] Crato, N. e Ray, B. K. (2002). Semi-parametric smoothing estimators for long-memory processes with added noise. *Journal of Statistical Planning and Inference*, **105**: 283-297.
- [30] de Jong, P. (1988). The likelihood for a state space model. *Biometrika*, **75**(1):165-169.
- [31] de Jong, P. e Shephard, N. (1995). The simulation smoother for time series models. *Biometrika*, **82**: 339-350.
- [32] Deo, R. S. (2002). On EMM and QML estimation for the long memory stochastic volatility model. *Manuscript*.
- [33] Deo, R. S. e Hurvich, C. M. (2001). On the log periodogram regression estimator of the memory parameter in long memory stochastic volatility models. *Econometric Theory*, **17**: 686-710.
- [34] Diebold, F. X. e Inoue, A. (2001). Long memory and regime switching. *Journal of Econometrics*, **105**: 131-159.
- [35] Ding, Z. e Granger, C. W. J. (1996). Modelling volatility persistence of speculative returns: a new approach. *Journal of Econometrics*, **73**: 185-215.
- [36] Doornik, J. A. (1999). *Object-oriented Matrix Programming using Ox 2.1*. London: Timberlake Consultants Press.
- [37] Doornik, J. A. e Ooms, M. (1999). A package for estimating, forecasting and simulating ARFIMA models: arfima package 1.0 for Ox. *Preprint*: 1-31.
- [38] Doornik, J. A. e Ooms, M. (2001). Computational aspects of maximum likelihood estimation of autoregressive fractionally integrated moving average models. *Preprint*: 1-14.

- [39] Durbin, J. e Koopman, S. J. (1997). Monte Carlo maximum likelihood estimation for non-gaussian state space models. *Biometrika*, **84**(3): 669-684.
- [40] Durbin, J. e Koopman, S. J. (2000). Time series analysis of non-gaussian observations based on state space models from both classical and bayesian perspectives (with discussion). *Journal of the Royal Statistical Society, Series B*, **62**: 3-56.
- [41] Durbin, J. e Koopman, S. J. (2001a). *Time Series Analysis by State Space Methods*. New York: Oxford University Press Inc.
- [42] Durbin, J. e Koopman, S. J. (2001b). A simple and efficient simulation smoother for state space time series analysis. *Preprint. London School of Economics and Political Science*. (forthcoming in: *Biometrika*).
- [43] Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, **50**(4): 987-1007.
- [44] Fox, R. e Taqqu, M.S. (1986). Large sample properties of parameter estimates for strongly dependent stationary gaussian time series. *The Annals of Statistics*, **14**(2), 517-532.
- [45] Gallant, A. R.; Hsieh, D. e Tauchen, G. (1997). Estimation of stochastic volatility models with diagnostics. *Journal of Econometrics*, **81**: 159-192.
- [46] Geweke, J. e Porter-Hudak, S. (1983). The estimation and application of long memory time series models. *Journal of Time Series Analysis*, **4**(4): 221-238.
- [47] Geweke, J. (1994). Bayesian comparison of econometric models. *Manuscript, Federal Reserve Bank*. Minneapolis, Minnesota.
- [48] Giraitis, L. e Surgailis, D. (1990). A central limit theorem for quadratic form in strongly dependent linear variables and its application to Whittle's estimate. *Probability Theory and Related Fields*, **86**:87-104.
- [49] Ghysels, E.; Harvey, A. C. e Renault, E. (1996). Stochastic Volatility, In G. S. Maddala and C.R. Rao (Eds.), *Handbook of Statistics, Statistical Methods in Finance*, **14**: 119-191.
- [50] Granger, C. W. G. e Joyeux, R. (1980). An introduction to long memory time series models and fractional differencing. *Journal of Time Series Analysis*, **1**(1): 15-29.
- [51] Granger, C. W. G. (1980). Long memory relationships and the aggregation of dynamic models. *Journal of Econometrics*, **14**: 227-238.

- [52] Harvey, A. C. (1993). Long memory in stochastic volatility. *London School of Economics, Preprint*.
- [53] Harvey, A. C. (1994). *Time Series Models*, 2ed. Massachusetts: Cambridge.
- [54] Harvey, A. C. (1998). Long memory in stochastic volatility. In J. Knight and S. Satchell (eds.) *Forecasting Volatility in Financial Markets*, Butterworth - Haineman, Oxford: 307-320.
- [55] Harvey, A. C., Ruiz, E. e Shephard, N. G. (1994). Multivariate stochastic variance models. *Review of Economic Studies*, **61**: 247-264.
- [56] Harvey, A. C. e Shephard, N. (1996). Estimation of an asymmetric stochastic volatility model for asset returns. *Journal of Business and Economic Statistics*, **14**: 429-434.
- [57] Haslett, J. e Raftery, A. E. (1989). Space-time modelling with long-memory dependence: assessing Ireland's wind power resource. *Applied Statistical*, **38**(1): 1-50.
- [58] Hassler, U. (1993). Regression of spectral estimators with fractionally integrated time series. *Journal of Time Series Analysis*, **14**(4): 369-380.
- [59] Hipel, K. W. e McLeod, A. I. (1978a). Preservation of the rescaled adjusted range, 2: simulation studies using Box-Jenkins models. *Water Resources Research*, **14**: 509-516.
- [60] Hipel, K. W. e McLeod, A. I. (1978b). Preservation of the rescaled adjusted range, 3: fractional gaussian noise algorithms. *Water Resources Research*, **14**: 517-518.
- [61] Hol, E. e Koopman, S. J. (2002). Stock index volatility forecasting with high frequency data. *Timbergen Institute Discussion Paper, TI2002-068/4. Preprint*: 1-22.
- [62] Hosking, J. R. M. (1981). Fractional differencing. *Biometrika*, **68**(1): 165-176.
- [63] Hosking, J. R. M. (1984). Modelling persistence in hydrological time series using fractional differencing. *Water Resources Research*, **20**(12): 1898-1908.
- [64] Hsu, N-J. e Breidt, F. J. (2002). Bayesian analysis of fractionally integrated ARMA with additive noise. *Preprint*: 1-37.
- [65] Hurst, H. E. (1951). Long term storage capacity of reservoirs. *Transactions of the American Society of Civil Engineers*, **116**: 770-799.
- [66] Hurvich, C. M. e Soulier, P. (2002). Testing for long memory in volatility. *Preprint*: 1-17.

- [67] Hurvich, C. M., Deo, R. e Brodsky, J. (1998). The mean squared error of Geweke and Porter-Hudak's estimator of the memory parameter of a long memory time series. *Journal of Time Series Analysis*, **19**(1): 19-46.
- [68] Hurvich, C. M. e Beltrao, K. I. (1994). Automatic semiparametric estimation of the memory parameter of a long-memory time series. *Journal of Time Series Analysis*, **15**(3): 285-302.
- [69] Hwang, S. (2000). Structural breaks, persistence and long memory. *Preprint*: 1-31.
- [70] Hwang, S. e Pereira, P. L. V. (2002). Property of volatility persistence: revisited with GARCH models. *Preprint*.
- [71] Janacek, G. J. (1982). Determining the degree of differencing for time series via the log spectrum. *Journal of Time Series Analysis*, **3**: 177-283.
- [72] Kim, S.; Shephard, N. e Chib, S. (1998). Stochastic volatility: likelihood inference and comparison with ARCH models. *Review of Economic Studies*, **65**: 361-393.
- [73] Koopman, S. J. (1993). Disturbance smoother for state space models. *Biometrika*, **80**(1): 117-126.
- [74] Koopman, S. J.; Shephard, N. e Doornik, J. A. (1999). Statistical algorithms for models in state space using SsfPack 2.2. *Econometrics Journal*, **2**: 113-166.
- [75] Künsch, H. R. (1987). Statistical aspects of self-similar processes. In *Proceedings of the First World Congress of the Bernoulli Society* (eds. Yu. Prohorov and V. V. Sazanov), **1**: 67-74.
- [76] Lawrance, A. J. e Kottegoda, N. T. (1977). Stochastic modelling of riverflow time series. *Journal of the Royal Statistical Society, A*, **140**: 1-47.
- [77] Li, W. K. e McLeod, A. I. (1986). Fractional time series modelling. *Biometrika*, **73**(1): 217-221.
- [78] Liesenfeld, R. e Richard, J-F. (2003). Univariate and multivariate stochastic volatility models: estimation and diagnostics. *Journal of Empirical Finance*, **207**: 1-27.
- [79] Liu, M. (2000). Modeling long memory in stock market volatility. *Journal of Econometrics*, **99**: 139-171.
- [80] Mandelbrot, B. B. (1963). The variation of certain speculative prices. *Journal of Business*, **36**: 394-419.

- [81] Mandelbrot, B. B. (1975). Limit theorems of the self-normalized range for weakly and strongly dependent processes. *Z. Wahr. Verw. Geb.*, **31**: 271-285.
- [82] Mandelbrot, B. B. (1983). *The Fractal Geometry of Nature*. New York: Freeman.
- [83] Mandelbrot, B. B. e Taqqu, M. S. (1979). Robust analysis of long-run serial correlation. *Proc. 42nd Session of the Internat. Statist. Institute*, Mamila, Bull. Int. Statist. Institute, **48**(2):69-104.
- [84] McLeod, A. I. e Hipel, K. W. (1999). Exploratory spectral analysis of hydrological time series. *Preprint*: 1-17, (Updated version of: Exploratory spectral analysis of hydrological time series, stochastic hydrology and hydraulics (1995), **9**: 171-205).
- [85] Montanari, A.; Taqqu, M. S. e Teverovsky, V. (1999). Estimating long-range dependence in the presence of periodicity: an empirical study. *Mathematical and Computer Modelling*, **29**: 217-228.
- [86] Motta, A. C. O. e Hotta, L. K. (2001). Modelos de Espaço de Estado Não-Gaussianos e o Modelo de Volatilidade Estocástica. Tese de Mestrado. Departamento de Estatística, UNICAMP.
- [87] Nelson, D. B. (1988). The Time Series Behavior of Stock Market Volatility and Returns, *Unpublished Ph D dissertation, Massachusetts, Institute of Technology, Cambridge, MA*.
- [88] Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: a new approach. *Econometrica*, **59**: 347-370.
- [89] Parke, W. R. (1999). What is fractional integration?. *Working Paper*, 99-01.
- [90] Pérez, A. e Ruiz, E. (2001). Finite sample properties of a QML estimator of stochastic volatility models with long memory. *Economics Letters*, **70**: 157-164.
- [91] Perron, P. (1989). The great crash, the oil price shock, and the unit root hypothesis. *Econometrica*, **57**(6): 1361-1401.
- [92] Porter-Hudak, S.(1990). An application of the seasonal fractionally differenced model to the monetary aggregates. *J. Amer. Statist. Assoc.*, **85**: 338-344.
- [93] Ray, B. K. (1991). Fractionally Differenced ARMA Models: Seasonality and Forecasting Issues. Ph.D. Thesis. Columbia University.

- [94] Ray, B. K. (1993). Modelling long-memory processes for optimal long-range prediction. *Journal of Time Series Analysis*, **14**: 511-525.
- [95] Ray, B. K. e Tsay, R. S. (2000). Size effects on common long-range dependence in stock volatilities. *J. Business Econom. Statist.*, **18**: 254-262.
- [96] Reisen, V. A. (1994). Estimation of the fractional difference parameter in the ARIMA(p,d,q) model using the smoothed periodogram. *Journal of Time Series Analysis*, **15**(3): 335-350.
- [97] Reisen, V. A.; Abraham, B. e Toscano, E. M.(2000). Parametric and semiparametric estimations of stationary univariate ARFIMA models. *Brazilian Journal of Probability and Statistics*, **14**: 185-206.
- [98] Reisen, V. A.; Sena Jr, M. R. e Lopes, R. C. (2002). Estimating the long-range dependence parameter in the presence of non-gaussian errors. *Preprint*: 1-31.
- [99] Robinson, P. M. (1991). Testing for strong serial correlation and dynamic conditional heteroskedasticity in multiple regression. *Journal of Econometrics*, **47**: 67-84.
- [100] Robinson, P. M. (1992). Semiparametric analysis of long-memory time series. *Annals of Statistics*, **22**: 515-539.
- [101] Robinson, P. M. (1994a). Rates of convergence and optimal spectral bandwidth for long range dependence. *Probab. Theory Related Fields*, **99**: 443-473.
- [102] Robinson, P. M. (1994b). Time series with strong dependence. *In Advances in Econometrics. Sixth World Congress (C.A. Sims,ed)*. Cambridge Univ. Press, **1**: 47-96.
- [103] Robinson, P. M. (1995a). Log-periodogram regression of time series with long range dependence. *The Annals of Statistics*, **23**(3): 1048-1072.
- [104] Robinson, P. M. (1995b). Gaussian semiparametric estimation of long range dependence. *The Annals of Statistics*, **23**(5): 1630-1661.
- [105] Robinson, P. M. e Henry, M.(1998). Long and short memory conditional heteroscedasticity in estimating the memory parameter of levels. *Discussion Paper STICERD Econometrics Discussion Paper EM/98/357*. London School of Economics and Political Science: 1-33.
- [106] Ruiz, E. (1994). Quasi-maximum likelihood estimation of stochastic volatility models. *Journal of Econometrics*, **63**: 289-306.

- [107] Sandmann, G. e Koopman, S. J. (1998). Estimation of stochastic volatility models via Monte Carlo maximum likelihood. *Journal of Econometrics*, **87**: 271-301.
- [108] Shephard, N. (1996). *Statistical aspects of ARCH and stochastic volatility*. In O. E. Barndorff-Nielsen, D. R. Cox e D. V. Hinkley (eds.). *Statistical Models in Econometrics, Finance and other Fields*. London: Chapman & Hall: 1-67.
- [109] Shephard, N. e Pitt, M. K. (1997). Likelihood analysis of non-gaussian measurement time series. *Biometrika*, **84**(3): 653-667.
- [110] Sibbertsen, P. (2001). Long-memory versus structural breaks: an overview. *Preprint*: 1-28.
- [111] Sowell, F. (1990). The fractional unit root distribution. *Econometrica*, **58**: 495-508.
- [112] Taylor, S. J. (1986). *Modelling Financial Time Series*. New York: Wiley.
- [113] Taylor, S. J. (1994). Modelling stochastic volatility. *Mathematical Finance*, **4**: 183-204.
- [114] Taylor, S. J. (2000). Consequences for option pricing of a long memory in volatility. *Preprint*: 1-55.
- [115] Wei, W. W. S. (1990). *Time Series Analysis - Univariate and Multivariate Methods*. New York: Addison-Wesley Publishing Company, Inc..
- [116] Whittle, P. (1951). *Hypothesis Testing in Time Series Analysis*. New York: Hafner.
- [117] Wright, J. (1999). A new estimator of the fractionally integrated stochastic volatility model. *Economics Letters*, **63**: 295-303.
- [118] Wright, J. (2000). Log periodogram estimation of long memory volatility dependencies with conditionally heavy tailed returns. *Board of Governors of the Federal Reserve System, International Finance Discussion Paper Number 685*.