



Revista Brasileira de Finanças

ISSN: 1679-0731

rbfin@fgv.br

Sociedade Brasileira de Finanças

Brasil

Gomes dos Santos, Douglas; Ziegelmann, Flávio Augusto  
Estimação e Previsão de Volatilidade em Períodos de Crise: Um Estudo Comparando Modelos  
GARCH e Modelos Aditivos Semi-Paramétricos  
Revista Brasileira de Finanças, vol. 10, núm. 1, 2012, pp. 49-70  
Sociedade Brasileira de Finanças  
Rio de Janeiro, Brasil

Disponível em: <http://www.redalyc.org/articulo.oa?id=305824774003>

- Como citar este artigo
- Número completo
- Mais artigos
- Home da revista no Redalyc

redalyc.org

Sistema de Informação Científica  
Rede de Revistas Científicas da América Latina, Caribe, Espanha e Portugal  
Projeto acadêmico sem fins lucrativos desenvolvido no âmbito da iniciativa Acesso Aberto

# Estimação e Previsão de Volatilidade em Períodos de Crise: Um Estudo Comparando Modelos GARCH e Modelos Aditivos Semi-Paramétricos

(Volatility Estimation and Forecasting During Crisis Periods: A Study Comparing GARCH Models with Semiparametric Additive Models)

Douglas Gomes dos Santos\*

Flávio Augusto Ziegelmann\*\*

## Resumo

Neste artigo, comparamos os modelos aditivos semi-paramétricos e os modelos GARCH quanto à capacidade de estimar e prever volatilidade em períodos de crise. Nossos estudos de Monte Carlo indicam uma melhor performance dos modelos GARCH quando suas formas funcionais não diferem da especificada no Processo Gerador de Dados (PGD). Contudo, na medida em que diferem do PGD, obtemos resultados que sugerem a superioridade dos modelos aditivos. Adicionalmente, efetuamos uma aplicação empírica em três períodos de alta volatilidade da série de retornos do IBOVESPA, onde ambas as classes de modelos obtêm resultados similares.

**Palavras-chave:** volatilidade; modelos aditivos semi-paramétricos; modelos GARCH; crise.

**Códigos JEL:** C14; C22; C52; C53.

## Abstract

In this paper, we compare semiparametric additive models with GARCH models in terms of their capability to estimate and forecast volatility during crisis periods. Our Monte Carlo studies indicate a better performance for GARCH models when their functional forms do not differ from that of the specified Data Generating Process (DGP). However, if they differ from the DGP, the results suggest the superiority of additive models. Additionally, we perform an empirical application in three selected periods of high volatility of IBOVESPA returns series, in which both families of models obtain similar results.

**Keywords:** volatility; semiparametric additive models; GARCH models; crisis.

---

Submetido em 2 de fevereiro de 2011. Reformulado em 22 de setembro de 2011. Aceito em 9 de novembro de 2011. Publicado on-line em 15 de abril de 2012. Os autores agradecem ao editor, bem como aos dois avaliadores por suas relevantes sugestões. O artigo foi avaliado segundo o processo de duplo anonimato além de ser avaliado pelo editor. Editor responsável: Márcio Laurini. Reprodução parcial ou total e trabalhos derivativos permitidos com a citação apropriada da fonte.

\*Programa de Pós-Graduação em Economia (UFRGS), Porto Alegre, RS, Brasil. E-mail: dgomess@hotmail.com

\*\*Departamento de Estatística e Programas de Pós-Graduação em Economia e Administração (UFRGS), Porto Alegre, RS, Brasil. E-mail: flavioaz@mat.ufrgs.br

## 1. Introdução

Em razão da relevância de temas como risco e incerteza na teoria econômica moderna, métodos capazes de modelar uma variância condicional dinâmica ao longo do tempo têm sido propostos na literatura. Neste contexto, o modelo ARCH (*Autoregressive Conditional Heteroskedastic*) de Engle (1982) e as extensões da família GARCH destacam-se por sua capacidade, variedade e ampla utilização. Inúmeros trabalhos envolvendo os modelos GARCH aparecem na literatura nas últimas décadas, levando a diversas variações nos mesmos. Estas variações visam, em geral, a descrever de forma mais apropriada as regularidades empíricas observadas em séries de tempo financeiras. Para revisões abrangentes, ver Bollerslev (2008), Bollerslev *et al.* (1992), Bollerslev *et al.* (1994), Rydberg (2000), entre outros.

Em contraste aos modelos GARCH paramétricos, a principal motivação para a modelagem não paramétrica surge quando não se possui conhecimento suficiente sobre o processo gerador dos dados (PGD), permitindo desta forma que os mesmos “falem por si” (para revisões sobre estes métodos, ver Fan & Gijbels (1996), Fan & Yao (2003), Wand & Jones (1995), entre outros). Assim, a análise dos dados é feita sob a suposição de modelos menos rígidos ou, em outras palavras, mais flexíveis. Pode-se dizer que a principal vantagem dos modelos não paramétricos está relacionada a esta flexibilidade e à consequente diminuição do risco de errar na especificação do modelo. Contudo, os custos desta maior flexibilidade se apresentam em taxas de convergência dos estimadores mais lentas, perda do poder de extrapolação e a dificuldade de estimação em alta dimensionalidade (*curse of dimensionality*).

Sob o enfoque multivariado, onde há a presença de mais de uma variável defasada na regressão, o modelo saturado (ou totalmente não paramétrico) destaca-se como a opção mais flexível. Todavia, ressalta-se que problemas de estimação relacionados a espaços de alta dimensão são freqüentes (*curse of dimensionality*). Nestes casos não se pode obter uma vizinhança local, essencial em modelagem não paramétrica. Assim, abre-se espaço para o uso de modelos semi-paramétricos, os quais impõem algum tipo de restrição mais forte nas funções a serem estimadas. Por essa razão, são mais rígidos que o modelo não paramétrico saturado, mas podem ser substancialmente mais flexíveis que os paramétricos. Discussões abrangentes sobre modelos semi-paramétricos podem ser encontradas em Fan & Yao (2003), Gao (2007), Hastie & Tibshirani (1990), Horowitz (1998), entre outros.

O modelo aditivo semi-paramétrico, amplamente discutido em Hastie & Tibshirani (1990), é uma alternativa para evitar o problema da alta dimensionalidade. Neste, a idéia é modelar aditivamente funções univariadas. A motivação principal do modelo aditivo é permitir que componentes do modelo de regressão assumam formas não pré-determinadas e possivelmente não lineares. Aqui o problema em espaços de alta dimensão é superado devido ao método de estimação ser construído a partir de suavizadores univariados. Carroll *et al.* (2002), Kim & Linton (2004),

Levine & Li (2007) e Linton & Mammen (2005) são algumas das referências que utilizam os modelos aditivos (ou suas versões generalizadas) em estimação de volatilidade.

Neste artigo, temos como objetivo trazer uma contribuição diferenciada, comparando os modelos aditivos e os modelos GARCH quanto à capacidade de estimar e prever a volatilidade de retornos em períodos de crise. Os momentos de turbulência no mercado financeiro certamente são aqueles que oferecem os maiores riscos, mas também as maiores oportunidades de alavancagem, motivando assim fortemente este estudo. As comparações baseiam-se em um estudo de Monte Carlo e em um estudo empírico do Índice BOVESPA. As simulações de Monte Carlo indicam uma melhor performance dos modelos GARCH quando suas formas funcionais não diferem da especificada no Processo Gerador de Dados (PGD). Contudo, na medida em que diferem, obtêm-se resultados que sugerem um melhor desempenho dos modelos aditivos. No estudo empírico, os resultados indicam uma equivalência entre os modelos em relação à performance preditiva tanto dentro quanto fora da amostra.

O restante do artigo é estruturado da seguinte forma: a seção 2 descreve sucintamente os modelos e métodos de estimação utilizados, focando em modelos aditivos; a seção 3 descreve o estudo de Monte Carlo e a análise empírica do Índice BOVESPA; e, por fim, as considerações finais são apresentadas na seção 4.

## 2. Modelos e sua Estimação

Nesta seção, primeiramente descreveremos de forma breve os já bem conhecidos modelos da família GARCH. A seguir, devido ao uso menos popular, descreveremos um pouco mais detalhadamente modelos semi-paramétricos aditivos e estimação polinomial local.

### 2.1 Modelos GARCH

O modelo  $ARCH(q)$ , introduzido por Engle (1982), pode ser definido como

$$\begin{aligned}\varepsilon_t &= \sigma_t z_t \\ \sigma_t^2 &= \omega + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2\end{aligned}\tag{1}$$

onde  $z_t \sim i.i.d. D(0, 1)$ ,  $t = 1, \dots, T$ ,  $D(\cdot)$  representa uma função densidade de probabilidade (fdp) qualquer,  $\omega > 0$  e  $\alpha_i \geq 0$ ,  $i = 1, \dots, q$ . A condição necessária e suficiente para (1) definir um processo estritamente estacionário único  $\{\varepsilon_t\}$ , com  $E\varepsilon_t^2 < \infty$ , é  $\sum_{i=1}^q \alpha_i < 1$  (ver Fan & Yao (2003)).

Evidências empíricas mostram que modelos ARCH de ordens elevadas devem ser selecionados para poder captar a dinâmica da variância condicional, desta forma, envolvendo a estimação de diversos parâmetros. O modelo ARCH Generalizado (GARCH) de Bollerslev (1986) surge como resposta ao problema.

O modelo  $GARCH(p, q)$  pode ser expresso como:

$$\sigma_t^2 = \omega + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2 \quad (2)$$

onde  $\omega > 0$ ,  $\alpha_i \geq 0$  para  $i = 1, \dots, q$ , e  $\beta_j \geq 0$  para  $j = 1, \dots, p$ . A condição necessária e suficiente para (2) definir um processo estritamente estacionário único  $\{\varepsilon_t\}$ , com  $E\varepsilon_t^2 < \infty$ , é  $\sum_{i=1}^q \alpha_i + \sum_{j=1}^p \beta_j < 1$  (ver Fan & Yao (2003)).

Nos modelos ARCH e GARCH retornos de mesma magnitude, em termos absolutos, resultam em uma mesma volatilidade futura. Entretanto, pode existir uma assimetria nesta resposta, visto que várias vezes retornos negativos levam a uma maior volatilidade futura do que retornos positivos de mesmo módulo. O modelo  $GJR(p, q)$ , proposto por Glosten *et al.* (1993), incorpora esta característica de assimetria. Sua versão generalizada é dada por:

$$\sigma_t^2 = \omega + \sum_{i=1}^q (\alpha_i \varepsilon_{t-i}^2 + \gamma_i S_{t-i}^- \varepsilon_{t-i}^2) + \sum_{j=1}^p \beta_j \sigma_{t-j}^2 \quad (3)$$

onde  $S_t^-$  é uma variável *dummy*. Neste modelo, é assumido que o impacto de  $\varepsilon_t^2$  em  $\sigma_t^2$  é distinto, dependendo se  $\varepsilon_t$  é positivo ou negativo.

O modelo  $TS-GARCH(p, q)$  de Taylor (1986) e Schwert (1989) parametriza o desvio padrão condicional como uma defasagem distribuída dos resíduos absolutos e dos desvios padrão condicionais defasados:

$$\sigma_t = \omega + \sum_{i=1}^q \alpha_i |\varepsilon_{t-i}| + \sum_{j=1}^p \beta_j \sigma_{t-j} \quad (4)$$

Esta especificação mitiga a influência de grandes (em termos absolutos) observações quando se compara ao modelo  $GARCH(p, q)$  tradicional.

O principal método de estimação dos modelos GARCH é o de máxima verossimilhança, normalmente condicional às observações iniciais.<sup>1</sup> No estudo empírico, iremos selecioná-los<sup>2</sup> por meio dos critérios de informação de Akaike (AIC) e de Schwarz (BIC) e a partir do valor máximo da função de log-verossimilhança. Ainda, discussões a respeito das propriedades assintóticas dos estimadores de máxima verossimilhança condicional para modelos GARCH podem ser encontradas em Hall & Yao (2003), e de forma compacta em Fan & Yao (2003).

## 2.2 Métodos não paramétricos

Há vários métodos não paramétricos utilizados para modelagem estocástica. Dentre eles, podem ser citados com destaque aqueles baseados em suavização via *kernels* e aqueles baseados em *splines*. Descrições de métodos de suavização via

<sup>1</sup>O pacote fGarch, do *software* livre R, possui estas implementações.

<sup>2</sup>Para seleção em modelos GARCH, ver Fan & Yao (2003), Tsay (2010), entre outros.

*kernels* podem ser encontradas em Bowman & Azzalini (1997), Fan & Gijbels (1996), Fan & Yao (2003), entre outros.

Em regressão não paramétrica, o estimador polinomial local de suavização *kernel* é amplamente utilizado. Na apresentação do método, considere a sequência de variáveis aleatórias bidimensionais<sup>3</sup>  $(Y_t, X_t), t = 1, \dots, T$ . A esperança condicional e a variância condicional são definidas respectivamente como  $m(x) = E(Y_t|X_t = x)$  e  $\sigma^2(x) = Var(Y_t|X_t = x) > 0, t = 1, \dots, T$ . Supõe-se um modelo que relaciona as variáveis  $(Y_t, X_t)$ , com a seguinte forma:

$$Y_t = m(X_t) + \sigma(X_t)\varepsilon_t \quad (5)$$

onde  $E(\varepsilon_t|X_t) = 0$  e  $Var(\varepsilon_t|X_t) = 1$ . Utilizando expansão de Taylor, pode-se aproximar a função de regressão desconhecida  $m(X_t)$  por um polinômio de ordem  $p$ . Desta forma, considerando um ponto de interesse  $x$  no domínio da variável aleatória  $X$ , pode-se definir o estimador polinomial local de  $m(x)$  como  $\hat{m}(x) = \hat{\beta}_0$ , com  $\hat{\beta}_0$  dado pela solução do problema de mínimos quadrados ponderados

$$(\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_p) = \arg \min_{\beta_0, \beta_1, \dots, \beta_p} \sum_{t=1}^T \left\{ Y_t - \sum_{i=0}^p \beta_i (X_t - x)^i \right\}^2 K_h(X_t - x) \quad (6)$$

onde  $K(\cdot)$  é uma função densidade de probabilidade simétrica em  $\mathbb{R}$ ,  $K_h(u) = (1/h)K(u/h)$ , e  $h > 0$  é o parâmetro de suavização ou janela (*bandwidth*). Ainda, quando  $p = 1$ , obtém-se o estimador linear local.

Utiliza-se neste artigo, com  $p = 1$ , o suavizador loess<sup>4</sup> de Cleveland & Devlin (1988). Em loess, as vizinhanças locais são determinadas por uma janela baseada em vizinhos próximos. Complementarmente,  $0 < f \leq 1$  expressa uma fração dos dados ( $T$ ) a serem incluídos em cada vizinhança, de tal forma que  $f$  representa o parâmetro de suavização (*span*). Adotamos como método de seleção do *span* a minimização do critério de informação aproximado de Akaike (AIC), como sugerido em Hastie (1993) e Hastie & Tibshirani (1990).

Em regressão múltipla, foca-se a atenção na relação estrutural entre a variável resposta  $Y$  e o vetor de covariáveis  $(X_1, X_2, \dots, X_d)'$ . Desta maneira, pode-se estender o modelo (5) ao caso multivariado, onde  $\mathbf{X}$  passa a ser um vetor<sup>5</sup> de dimensão- $d$  em  $\mathbb{R}^d$  e  $Y$  é um escalar em  $\mathbb{R}$ . A esperança e a variância condicionais são dadas por  $m(\mathbf{x}) = E(Y|\mathbf{X} = \mathbf{x})$  e  $\sigma^2(\mathbf{x}) = Var(Y|\mathbf{X} = \mathbf{x}) > 0$ , respectivamente. O modelo (5) assume a forma

$$Y_t = m(\mathbf{X}_t) + \sigma(\mathbf{X}_t)\varepsilon_t \quad (7)$$

onde  $E(\varepsilon_t|\mathbf{X}_t) = 0$  e  $Var(\varepsilon_t|\mathbf{X}_t) = 1$ .

<sup>3</sup>No contexto de séries temporais,  $X_t$  representa  $Y_{t-1}$ .

<sup>4</sup>Versão disponível no Pacote gam (ver Hastie (1993)) do software livre R.

<sup>5</sup>No contexto de séries de tempo,  $\mathbf{X}_t = (Y_{t-1}, Y_{t-2}, \dots, Y_{t-d})'$ .

Para modelar a variância condicional, utilizamos o estimador em dois passos baseado nos resíduos, o qual chamamos de estimador residual da volatilidade. Alguns autores como Fan & Yao (1998), Ruppert *et al.* (1997), Ziegelmann (2002, 2008, 2011) apresentam implementações e resultados tanto teóricos como numéricos sobre a técnica. Resumidamente, o método consiste em reescrever o modelo (7) como  $Y_t - m(\mathbf{X}_t) = \sigma(\mathbf{X}_t)\varepsilon_t$ , tal que ao tomar a esperança condicional dos resíduos ao quadrado,  $r_t^2 = [Y_t - m(\mathbf{X}_t)]^2 = \sigma^2(\mathbf{X}_t)\varepsilon_t^2$ , obtém-se  $E(r_t^2|\mathbf{X}_t) = \sigma^2(\mathbf{X}_t)$ . Assim, pode-se estimar  $\sigma^2(\mathbf{X}_t)$  a partir de uma regressão de  $\hat{r}_t^2$  contra  $\mathbf{X}_t$ , onde  $\hat{r}_t^2 = [Y_t - \hat{m}(\mathbf{X}_t)]^2$ . Nesta situação,  $\hat{r}_t^2$  representa a nova variável resposta, substituindo  $Y_t$  nos modelos anteriores.

Entretanto, como mencionado anteriormente, o modelo não paramétrico (7) não é tão aplicável devido ao problema de estimação em alta dimensionalidade. Então, possíveis soluções costumam estar relacionadas a algum mecanismo de redução de dimensão, como os modelos aditivos semi-paramétricos, por exemplo.

O modelo aditivo semi-paramétrico é uma generalização do modelo de regressão linear tradicional, sendo sugerido por Friedman & Stuetzle (1981) e substancialmente desenvolvido por Hastie & Tibshirani (1990). No modelo de regressão linear múltipla, a função de regressão  $m(\cdot)$  é assumida como linear, e portanto aditiva nas variáveis explicativas. Nos modelos aditivos, o pressuposto de linearidade é abandonado, mas a forma aditiva é mantida.

Na modelagem da variância condicional,<sup>6</sup> o modelo aditivo impõe ao modelo não paramétrico (7) a seguinte estrutura, o modelo aditivo impõe ao modelo não paramétrico (7) a seguinte estrutura

$$\sigma^2(\mathbf{X}_t) = \sum_{i=0}^{d-1} f_i(X_{t-i}) \quad (8)$$

onde  $\mathbf{X}_t = (X_t, X_{t-1}, \dots, X_{t-d+1})'$  é o vetor de covariáveis. As funções desconhecidas  $f_0, \dots, f_{d-1}$  são univariadas, o que permite evitar o problema de estimação em alta dimensão sem a imposição de restrições paramétricas aos termos aditivos. Ainda, note que (8) inclui como caso específico os modelos ARCH paramétricos.

Para estimar a variância condicional em (8), pode-se utilizar qualquer suavizador não paramétrico (linear local, loess, etc.) na estimação de cada função univariada  $f_i(\cdot)$ ,  $i = 0, \dots, d-1$ . Contudo, ao invés de realizar  $d$  regressões univariadas, o procedimento de estimação deve levar em consideração a influência das demais variáveis no momento do ajuste de cada função. Adicionalmente, de forma a evitar constantes livres nas funções<sup>7</sup>  $f_i(\cdot)$ , usualmente impõe-se que

$$E\{f_i(X_{t-i})\} = 0, \quad i = 0, \dots, d-1 \quad (9)$$

resultando na incorporação de um intercepto ao modelo

<sup>6</sup>Esta especificação pode ser assumida simultaneamente, ou não, para a função de esperança condicional.

<sup>7</sup>Para prevenir ambigüidade e garantir a identificabilidade do modelo (8).

$$\sigma^2(\mathbf{X}_t) = \alpha + \sum_{i=0}^{d-1} f_i(X_{t-i}) \quad (10)$$

tal que  $\alpha = E(r_t^2)$ , onde  $r_t^2 = [Y_t - m(\mathbf{X}_t)]^2$ .

Se o modelo aditivo (10) é válido, tem-se que

$$E \left\{ r_t^2 - \alpha - \sum_{i \neq k} f_i(X_{t-i}) | X_{t-k} \right\} = f_k(X_{t-k}), \quad k = 0, \dots, d-1 \quad (11)$$

o que sugere um algoritmo iterativo para a estimação das funções univariadas  $f_0, \dots, f_{d-1}$ . De fato, dados  $\alpha$  e as funções  $f_i, i \neq k$ , a função  $f_k$  pode ser estimada por meio de uma simples regressão univariada. Ainda, denotando o suavizador univariado de  $f_k$  por  $S_k$ , e as estimativas iniciais de  $f_k$  por  $f_k^0, k = 0, \dots, d-1$ , o algoritmo utilizado, dito *backfitting*,<sup>8</sup> segue as etapas:

(i) Início:  $\hat{\alpha} = (T - d + 1)^{-1} \sum_{t=d}^T \hat{r}_t^2$ ,  $\hat{f}_k = f_k^0, \quad k = 0, \dots, d-1$ .

(ii) Para cada  $k = 0, \dots, d-1$ , calcule  $\hat{f}_k(\cdot) = S_k \{ \hat{r}_t^2 - \hat{\alpha} - \sum_{i \neq k} \hat{f}_i(\cdot) | X_{t-k} \}$  e após centralize<sup>9</sup> o estimador para obter  $\hat{f}_k^*(\cdot) = \hat{f}_k(\cdot) - (T - d + 1)^{-1} \sum_{t=d}^T \hat{f}_k(X_{t-k})$ .

(iii) Repita a etapa (ii) até convergir.

Para selecionar o número de funções a serem incorporadas no modelo, utilizamos o teste F aproximado e o critério de informação aproximado de Akaike (AIC). Os métodos propostos podem ser encontrados em Hastie (1993), Hastie & Tibshirani (1990) e Bowman & Azzalini (1997).

### 3. Implementações Numéricas

Nesta seção descrevemos o estudo de Monte Carlo e a análise empírica de dados do IBOVESPA. Todas as análises são feitas no *software* R. Nas simulações de Monte Carlo, são especificados dois PGDs distintos, com duas distribuições para os erros em cada um. Ainda, seis modelos têm suas estimativas e previsões comparadas em momentos selecionados de alta volatilidade. No estudo empírico, são utilizadas três séries de retornos diários do IBOVESPA. As comparações entre os modelos ocorrem em três períodos de crise escolhidos *ad hoc* para cada série.

Em ambos os estudos (Monte Carlo e empírico), duas medidas de erro são utilizadas para avaliar o desempenho dos modelos dentro e fora da amostra, a saber:

<sup>8</sup>Ver Hastie & Tibshirani (1990) para o estimador *Backfitting* e Linton & Nielsen (1995) para o método alternativo de Integração Marginal. Para informações referentes à convergência do algoritmo *Backfitting*, ver Hastie & Tibshirani (1990).

<sup>9</sup>Condição necessária para respeitar a restrição (9).



## i. Raiz do Erro Quadrático Médio (REQM)

$$REQM = \left[ \frac{1}{T} \sum_{t=1}^T (\hat{\sigma}_t^2 - \sigma_t^2)^2 \right]^{1/2} \quad (12)$$

## ii. Erro Absoluto Médio (EAM)

$$EAM = \frac{1}{T} \sum_{t=1}^T |\hat{\sigma}_t^2 - \sigma_t^2| \quad (13)$$

Nas expressões (12) e (13),  $\{\sigma_t^2\}_{t=1}^T$  representa a série de variâncias condicionais verdadeiras,<sup>10</sup>  $T$  é o número de dias estimados ou previstos e  $\{\hat{\sigma}_t^2\}_{t=1}^T$  denota a série de estimativas ou previsões um passo à frente da variância condicional de um determinado modelo.

No estudo empírico, utiliza-se adicionalmente o teste de comparação de capacidade de previsão de Diebold & Mariano (1995). Resumidamente, considere a série de erros de previsão para dois modelos concorrentes,  $\{\varepsilon_{i,t}\}_{t=1}^T = \{\hat{\sigma}_{i,t}^2 - \sigma_t^2\}_{t=1}^T$ ,  $i = 1, 2$ . A performance de cada previsão é mensurada por uma função perda particular,  $L(\varepsilon_{i,t})$ ,  $i = 1, 2$ . Utilizamos as funções quadrática  $L(\varepsilon_{i,t}) = (\varepsilon_{i,t})^2$  e absoluta  $L(\varepsilon_{i,t}) = |\varepsilon_{i,t}|$ . A hipótese nula de igual capacidade preditiva é  $H_0 : E[d_t] = 0$ ,  $t = 1, \dots, T$ , onde  $d_t = L(\varepsilon_{1,t}) - L(\varepsilon_{2,t})$ . Sob a hipótese nula, a estatística de Diebold-Mariano  $DM = \bar{d} / \sqrt{\hat{\omega}/T}$  é assintoticamente  $N(0, 1)$ , onde  $\bar{d}$  é a média amostral da série  $\{d_t\}_{t=1}^T$  e  $\hat{\omega}$  é um estimador consistente da variância assintótica de  $\sqrt{T}\bar{d}$ .

## 3.1 Estudo de Monte Carlo

Nesta seção busca-se comparar por meio de simulações de Monte Carlo os modelos da família GARCH e os Modelos aditivos. Nos experimentos, replicamos 250 séries de tempo<sup>11</sup> com 2.500 observações cada. Primeiramente, para cada série replicada, procura-se a partir da observação 1.500, ou seja, nas últimas 1.000 observações, o período (composto de 50 observações) mais volátil. Após a identificação (localização) do período com alta volatilidade, calculamos as estimativas e previsões um passo à frente atualizadas neste trecho da série que reúne 50 observações.

Os Processos Geradores de Dados (PGD) são os seguintes:

<sup>10</sup>No estudo empírico, ao invés das volatilidades verdadeiras (desconhecidas), são utilizados os quadrados dos resíduos.

<sup>11</sup>Para procedimentos computacionalmente caros em séries temporais, especialmente modelos (não/semi) paramétricos, o número de réplicas adotado é um usualmente empregado na literatura (ver Chen *et al.* (2010), Fan & Gu (2003), Linton *et al.* (2009), Liu *et al.* (2010)).

i. PGD<sub>1</sub>

$$\begin{aligned} y_t &= 0,2y_{t-1} + 0,1y_{t-2} + \sigma_t z_t \\ \sigma_t^2 &= 0,1 + 0,5\varepsilon_{t-1}^2 + 0,2\varepsilon_{t-2}^2 \end{aligned} \quad (14)$$

ii. PGD<sub>2</sub>

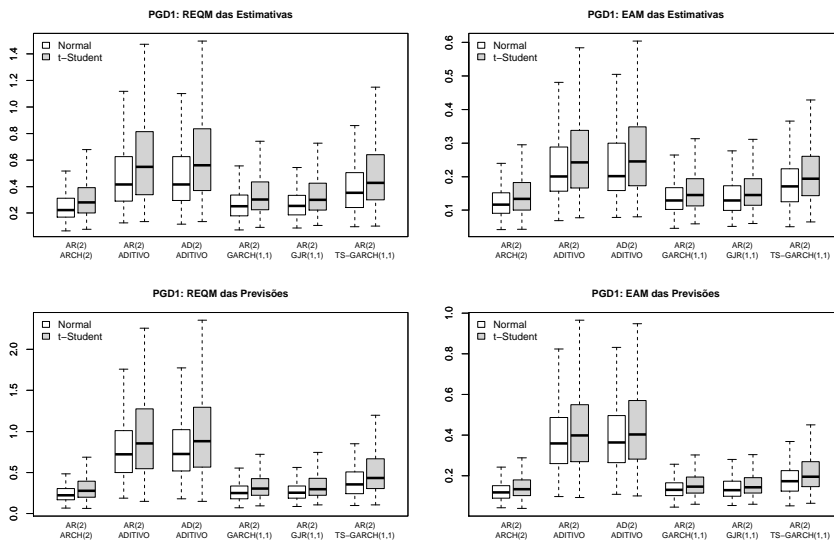
$$\begin{aligned} y_t &= 0,2y_{t-1} + 0,1y_{t-2} + \sigma_t z_t \\ \sigma_t^2 &= 0,1 + 0,5|\varepsilon_{t-1}| + 0,2|\varepsilon_{t-2}| \end{aligned} \quad (15)$$

onde  $\varepsilon_t = \sigma_t z_t$ , sendo especificadas duas distribuições para  $z_t$ ,  $z_t \sim N(0, 1)$  e  $z_t \sim t(6)$ .

Para cada série de tempo replicada, por ambos os PGDs, calculam-se as estimativas e previsões a partir de seis modelos distintos. Conhecendo o número de defasagens dos processos geradores, são analisados os modelos  $AR(2) - ARCH(2)$ ,  $AR(2) - \text{Aditivo}$ ,  $AD(2) - \text{Aditivo}$ ,<sup>12</sup>  $AR(2) - GARCH(1, 1)$ ,  $AR(2) - GJR(1, 1)$  e  $AR(2) - TS - GARCH(1, 1)$ . As estimativas e previsões de volatilidade, oriundas destes modelos, são utilizadas em conjunto com as volatilidades verdadeiras para calcular, para cada uma das 250 séries de tempo, as medidas de erro. A partir das estatísticas  $REQM_i$  e  $EAM_i$ , para  $i = 1, \dots, 250$ , pode-se apresentar na Figura 1 e Figura 2 os box plots que comparam os seis modelos citados (dentro e fora da amostra) para o PGD<sub>1</sub> e PGD<sub>2</sub>, respectivamente. Adicionalmente, a Tabela A.1 do Apêndice reúne as estatísticas dos experimentos de Monte Carlo.

A Figura 1 indica um melhor desempenho dentro e fora da amostra, de acordo com ambas as medidas, do modelo  $AR(2) - ARCH(2)$ , que possui a mesma especificação do processo gerador (PGD<sub>1</sub>). Este é seguido de perto pelos outros dois modelos cuja forma funcional depende de quadrados de resíduos, o  $GARCH(1, 1)$  e o  $GJR(1, 1)$ . O modelo  $TS - GARCH(1, 1)$  apresenta comportamento inferior em relação aos últimos. Para o PGD<sub>1</sub>, o pior desempenho é atribuído aos modelos aditivos, que indicam performance similar entre si.

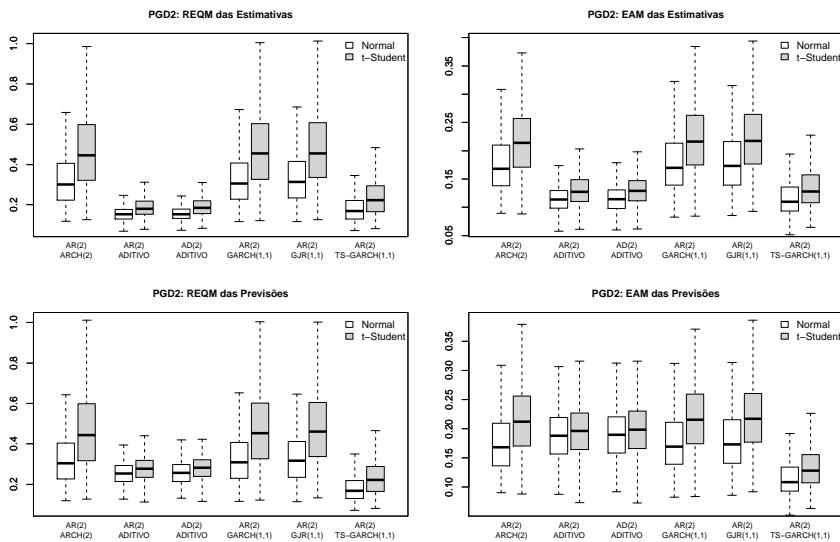
<sup>12</sup>Modelo Aditivo com duas funções univariadas tanto para a esperança condicional como para a variância condicional. Ainda, destaca-se que em ambos os modelos aditivos o *span* foi mantido constante, *span*=0,8, devido à performance ter sido equivalente a dos modelos com *span* variável (por série de tempo) e devido à redução do tempo computacional.



**Figura 1**

Medidas de erro de estimação e previsão para o PGD<sub>1</sub>. A Figura 1 exclui o pequeno número existente de *outliers*.

A Figura 2 reúne as comparações entre os seis estimadores para o PGD<sub>2</sub>. Nesta, fica clara a superioridade dos modelos aditivos para as medidas dentro da amostra. Cabe dizer que, para distribuição normal e métrica EAM, o modelo  $TS - GARCH(1,1)$  apresenta desempenho equivalente aos modelos aditivos. Ainda, os demais modelos da família GARCH, cuja forma funcional difere da especificada no PGD<sub>2</sub>, destacam-se pelo pior desempenho dentro da amostra. Para a análise fora da amostra, o modelo  $TS - GARCH(1,1)$  apresenta os melhores resultados. Adicionalmente, os modelos aditivos apresentam a segunda melhor performance. Excetuando-se a medida EAM, para distribuição normal, os modelos aditivos oferecem resultados superiores em previsão em relação aos modelos GARCH baseados em quadrados de resíduos.


**Figura 2**

Medidas de erro de estimação e previsão para o PGD<sub>2</sub>. A Figura 2 exclui o pequeno número existente de *outliers*.

Os resultados gerais do estudo de Monte Carlo sugerem um melhor desempenho dos modelos da família GARCH, notadamente aqueles especificados via quadrados de resíduos, para o primeiro processo (PGD<sub>1</sub>). Contudo, tendo em vista sua dependência à forma funcional, previamente especificada, estes apresentam dificuldade em se adaptar a processos que fujam do padrão de quadrados de resíduos, como o PGD<sub>2</sub>. Os modelos aditivos, por sua vez, baseados em funções livres, são capazes de se adaptar aos diferentes processos. Embora com performance inferior no PGD<sub>1</sub>, apresentam resultados robustos no PGD<sub>2</sub>.

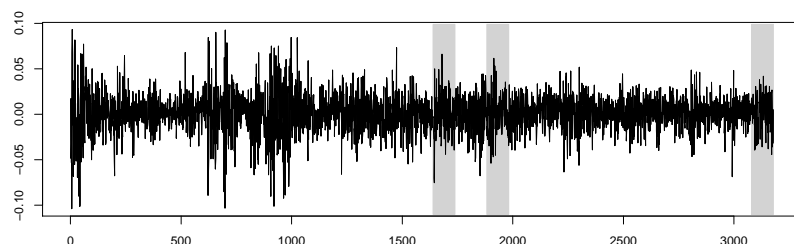
### 3.2 Estudo empírico

Utilizando dados do IBOVESPA, procura-se comparar mais uma vez os modelos GARCH e os Modelos aditivos em momentos de crise. Para isto, são utilizadas três séries de tempo compostas dos preços de fechamento diários do índice e seus respectivos retornos diários, definidos como  $y_t = \ln X_t - \ln X_{t-1}$  (onde  $X_t$  é o preço do ativo no tempo  $t$ ). Definimos três anos recentes com crise no mercado financeiro internacional, sendo a comparação delimitada aos momentos de alta volatilidade de 2001, 2002 e 2007. Primeiramente, baseando-se nas três amostras de retornos, os modelos mais adequados são identificados e selecionados (por amostra). Após, os modelos selecionados têm suas estimativas e previsões comparadas nos períodos denominados de crise (detalhes serão explicados adiante).



## Características dos dados

A amostra mais extensa dos retornos do IBOVESPA contém 3.178 observações, de 03/01/1995 até 26/11/2007. A evolução dos retornos para este período segue na Figura 3. A série compreende um período posterior à crise do México (fim de 1994) e atravessa momentos de alta volatilidade, como as crises da Ásia (1997) e da Rússia (1998). Em 2001, notam-se turbulências decorrentes do atentado de 11 de setembro nos Estados Unidos, e mais à frente, em 2002, os meses de setembro e outubro refletem os desdobramentos da bolha das empresas de alta tecnologia. Para completar o período amostral, nos meses de julho e agosto de 2007, o índice sofre influência da crise no mercado hipotecário *subprime* norte-americano.



**Figura 3**

Retornos do IBOVESPA (1995–2007). Áreas sombreadas denotam os períodos de crise de 2001, 2002 e 2007.

A identificação e escolha dos modelos aditivos, bem como dos modelos GARCH, ocorre para cada uma das três amostras dos retornos do IBOVESPA, cujas estatísticas descritivas estão na Tabela A.2 do Apêndice. A primeira amostra compreende os retornos de 03/01/1995 até 04/02/2002, ou seja, englobando o período em que ocorreu o ataque terrorista nos Estados Unidos. A segunda amostra reúne os dados de 03/01/1995 até 24/01/2003. Nesta série, o objetivo é mensurar o comportamento dos modelos a partir do mês de setembro de 2002, cuja volatilidade nos mercados financeiros é elevada. Por fim, a terceira série dos dados inicia em 03/01/1995 e se estende até 26/11/2007. Este período contempla os desdobramentos da crise no mercado hipotecário *subprime* nos EUA, que se intensificam ao final de julho e durante as primeiras semanas de agosto.

A partir da Tabela A.2 do apêndice, observa-se uma autocorrelação nos retornos das três amostras do IBOVESPA. Em relação aos quadrados dos retornos, as três amostras revelam a existência de autocorrelação.

## Modelos

Ao iniciar a modelagem pela série menos extensa do IBOVESPA, observando-se uma autocorrelação significativa nos retornos, ajustou-se uma estrutura para a esperança condicional. Devido à comparação das estimativas e previsões de volatilidade dos modelos ser realizada por meio dos quadrados dos resíduos,

optou-se por utilizar uma estrutura ARMA na equação da esperança condicional, para ambos os modelos,<sup>13</sup> preservando a base de comparação.

Através dos procedimentos usuais de identificação em séries temporais, selecionou-se um modelo<sup>14</sup> AR(2) para filtrar a série dos retornos. A Tabela A.3 do apêndice apresenta os resultados da seleção, de forma resumida, e os testes de correlação serial para os resíduos do modelo AR(2). Sendo assim, os resíduos do modelo AR(2) são utilizados para se obter as estimativas de volatilidade.

Na estimação da volatilidade via modelos aditivos, utilizamos os testes  $F$  e AIC aproximados como método de seleção dos modelos. Desta forma, buscou-se de forma simultânea a definição do número de funções univariadas a serem incluídas no modelo aditivo bem como o parâmetro de suavização (*span*) que minimizasse o critério AIC. Modelos compostos de uma até nove funções foram testados. Adicionalmente, a amplitude testada para o *span* variou de 0,5 até 0,8, com incrementos<sup>15</sup> de 0,05. Os modelos com seis funções ou mais apresentaram problemas de convergência ou estimativas negativas para volatilidade, sendo portanto excluídos da análise. A Tabela A.4 do Apêndice descreve os resultados dos testes  $F$  bem como os critérios AIC aproximados para os melhores modelos identificados (estimados) para a primeira amostra do IBOVESPA.<sup>16</sup>

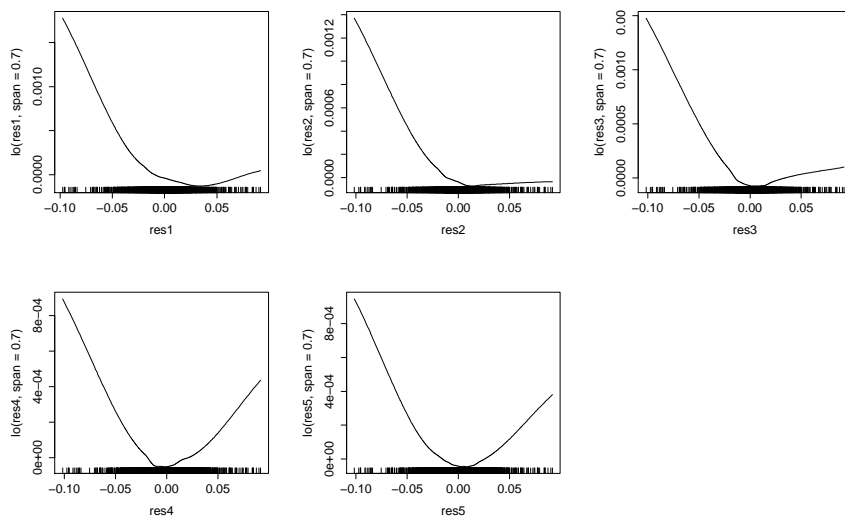
Da Tabela A.4 do Apêndice percebe-se que a inclusão sucessiva de componentes funcionais se mostra significativa até a quinta defasagem. Os critérios AIC aproximados reforçam os resultados do teste  $F$ , que indicam a superioridade do modelo com cinco funções. A Figura 4 apresenta a descrição visual das contribuições de cada função univariada para a explicação da variável resposta. Nesta, pode-se notar uma assimetria nas funções em relação à ocorrência de resíduos positivos ou negativos de mesma magnitude. As funções indicam uma maior inclinação na região de resíduos negativos e, conseqüentemente, sugerem uma maior resposta da volatilidade perante a ocorrência dos mesmos.

<sup>13</sup>Como visto nas simulações de Monte Carlo, as estimativas de volatilidade dos modelos aditivos indicam robustez, não dependendo do estimador utilizado na equação da esperança condicional.

<sup>14</sup>A identificação se deu via função de autocorrelação (fac) e função de autocorrelação parcial (fapc). Para seleção, comparamos critérios de informação (AIC), o valor máximo da função de log-verossimilhança, a significância dos parâmetros e a eliminação da autocorrelação serial (ver Tsay (2010)). Ainda, o modelo AR(2) se mostrou o mais adequado para as três amostras analisadas.

<sup>15</sup>Nota-se que a amplitude definida gera funções estimadas mais suaves e resultados superiores.

<sup>16</sup>O modelo aditivo selecionado na primeira amostra segue sendo o melhor nas duas amostras subsequentes. Sendo assim, são apresentados somente os resultados de identificação para a primeira amostra.



**Figura 4**

Funções univariadas do Modelo aditivo (5) – (Amostra 1). Os termos  $res1, res2, \dots, res5$  indicam as defasagens da variável explicativa.

Para os modelos GARCH, buscou-se seleccioná-los através dos critérios de informação (AIC, BIC) e a partir do valor máximo do log-verossimilhança. Na estimação via função de verossimilhança condicional, foram supostas duas distribuições condicionais para os erros padronizados, a normal e a t-Student. A Tabela A.5 do Apêndice fornece informação sobre os melhores modelos comparados para a primeira amostra do IBOVESPA. Aqui, optou-se por seleccionar um modelo  $ARCH(q)$  e outra parametrização entre os GARCH. Desta forma, foram seleccionados os modelos  $ARCH(7)$  e  $GJR(1,1)$ , ambos com distribuição t-Student<sup>17</sup> para os erros.

### Comparações nos períodos de crise

As comparações entre os modelos aditivos e GARCH são realizadas a partir das três séries de retornos do IBOVESPA (Tabela A.2 do Apêndice). Para avaliar o desempenho dos modelos em previsão dentro da amostra, os mesmos são estimados incorporando as observações contidas nos momentos de crise (cem últimos dias de negociação de cada amostra). Este período é definido em 2001 por (03.09.2001-04.02.2002), em 2002 por (02.09.2002-24.01.2003) e em 2007 por (02.07.2007-26.11.2007).

Para a análise fora da amostra, as previsões um passo à frente atualizadas são efetuadas para os mesmos momentos referenciados de cem dias. Sendo assim,

<sup>17</sup>Salienta-se que os dois modelos seleccionados para a primeira amostra seguem sendo os melhores nas amostras subsequentes.

os modelos são estimados até o período  $t$  (imediatamente anterior as cem últimas observações), e então são obtidas as previsões para  $t + 1$ . Em  $t + 1$ , os modelos são reestimados e as previsões são obtidas para  $t + 2$ , e assim sucessivamente até  $t + 100$ . As medidas de erro, REQM e EAM, são calculadas somente sobre as previsões (dentro e fora da amostra) nos intervalos de cem dias de negociação. Ainda, a capacidade preditiva dos modelos é comparada via teste de Diebold-Mariano.<sup>18</sup>

Apresenta-se na Tabela 1 o desempenho, dentro da amostra, dos modelos selecionados para os três momentos de crise. Segundo as métricas REQM e EAM, o modelo aditivo tem performance superior em 2001, ao passo que o modelo  $GJR(1, 1)$  é o de melhor desempenho nas crises de 2002 e 2007. Contudo, a capacidade preditiva dos modelos não é estatisticamente diferente em nenhum dos períodos. Os resultados comparativos para as previsões fora da amostra, contidos na Tabela 2, indicam o mesmo comportamento apresentado dentro da amostra.

**Tabela 1**

Desempenho dos modelos (previsão dentro da amostra) para o IBOVESPA

	M. Aditivo (5)	ARCH (7)	GJR (1,1)
Crise de 2001			
REQM	<b>8,460</b> (0,999)	9,176 (0,999)	8,745
EAM	<b>5,246</b> (0,760)	5,910 (0,721)	5,552
Crise de 2002			
REQM	7,084 (0,999)	6,889 (0,999)	<b>6,831</b>
EAM	4,988 (0,764)	4,870 (0,856)	<b>4,688</b>
Crise de 2007			
REQM	5,049 (0,999)	4,990 (0,999)	<b>4,873</b>
EAM	3,812 (0,856)	3,894 (0,793)	<b>3,631</b>

Nota: As REQMs e os EAMs foram multiplicados por  $10^4$  (os menores valores estão em negrito). Entre parênteses, os p-valores do teste de Diebold-Mariano (todos os modelos versus  $GJR(1, 1)$ ).

**Tabela 2**

Desempenho dos modelos (previsão fora da amostra) para o IBOVESPA

	M. Aditivo (5)	ARCH (7)	GJR (1,1)
Crise de 2001			
REQM	<b>8,565</b> (0,999)	9,250 (0,999)	8,778
EAM	<b>5,300</b> (0,789)	5,971 (0,686)	5,567
Crise de 2002			
REQM	7,401 (0,999)	6,943 (0,999)	<b>6,838</b>
EAM	5,487 (0,430)	4,910 (0,833)	<b>4,699</b>
Crise de 2007			
REQM	5,305 (0,999)	5,005 (0,999)	<b>4,880</b>
EAM	4,109 (0,634)	3,904 (0,786)	<b>3,632</b>

Nota: As REQMs e os EAMs foram multiplicados por  $10^4$  (os menores valores estão em negrito). Entre parênteses, os p-valores do teste de Diebold-Mariano (todos os modelos versus  $GJR(1, 1)$ ).

<sup>18</sup>Utilizamos a rotina implementada no Pacote *forecast* do *software* R.



#### 4. Considerações Finais

O objetivo deste artigo é o de comparar modelos aditivos semi-paramétricos com modelos GARCH paramétricos quanto à capacidade de estimar e prever alta volatilidade em períodos turbulentos. Os momentos de crise financeira são caracterizados por elevada incerteza no mercado, aumento da aversão ao risco, alta volatilidade nos retornos dos ativos, sendo que a exposição ao risco e as perdas potenciais neste cenário são substancialmente aumentadas (para uma abordagem histórica e empírica em crises financeiras, ver Kindleberger (2005), Krugman (2009), Reinhart & Rogoff (2009); entre outros). Ainda, embora com alguns padrões históricos semelhantes, cada crise costuma apresentar novos desafios (peculiaridades) ao gerenciamento de risco. Nesta situação, o uso conjunto de métodos distintos para a mensuração do risco se mostra fundamental (para uma abordagem em gerenciamento de risco em períodos de turbulência, ver Bénéplanc & Rochet (2011) e Stulz (2008)).

Inicialmente, realizou-se um experimento de Monte Carlo em que foram supostos dois processos geradores de dados. No primeiro ( $PGD_1$ ), baseado em quadrados de resíduos, os modelos GARCH com a mesma especificação do processo foram superiores aos modelos aditivos. No segundo ( $PGD_2$ ), baseado em resíduos absolutos, os modelos aditivos apresentaram melhor desempenho que os modelos GARCH com formas funcionais baseadas em quadrados de resíduos. Embora tenhamos trabalhado com alta volatilidade em ambos os PGDs, ao criar uma não linearidade no  $PGD_2$ , pôde-se analisar o comportamento, em particular, dos modelos GARCH tradicionais em um contexto similar ao de períodos de alta turbulência, qual seja, de elevado desconhecimento do processo gerador dos dados. Desta forma, os modelos aditivos semi-paramétricos justificam sua aplicabilidade por serem mais flexíveis que os GARCH paramétricos, adaptando-se aos diferentes PGDs.

Adicionalmente, realizou-se um estudo empírico avaliando a performance dos modelos em três séries de retornos do IBOVESPA contendo momentos de crise. Neste estudo, os modelos comparados tanto dentro quanto fora da amostra não apresentaram uma capacidade preditiva estatisticamente distinta de acordo com o teste de Diebold-Mariano.

Portanto, tais resultados nos levam a concluir que em momentos de grande turbulência no mercado (crise), associados a um alto grau de desconhecimento do verdadeiro PGD, é recomendável o uso de modelos aditivos semi-paramétricos em conjunto aos modelos GARCH paramétricos. Cabe salientar que os modelos aditivos podem ser utilizados de forma complementar aos modelos paramétricos, fornecendo sugestões de assimetrias (como ocorre neste estudo) e não linearidades a serem incorporadas nestes modelos.

## Apêndice

**Tabela A.1**

Estatísticas das medidas de erro

Modelos		PGD <sub>1</sub>				PGD <sub>2</sub>			
		Estimativas		Previsões		Estimativas		Previsões	
		N(.)	t(.)	N(.)	t(.)	N(.)	t(.)	N(.)	t(.)
AR(2)-ARCH(2)									
REQM	M	0,256	0,323	0,258	0,324	0,324	0,496	0,323	0,493
	DP	0,132	0,171	0,133	0,172	0,130	0,257	0,129	0,253
EAM	M	0,130	0,148	0,131	0,148	0,177	0,229	0,176	0,227
	DP	0,062	0,068	0,062	0,067	0,053	0,087	0,052	0,085
AR(2)-ADITIVO									
REQM	M	0,497	0,635	0,813	0,975	0,157	0,199	0,258	0,278
	DP	0,289	0,387	0,451	0,544	0,038	0,093	0,059	0,064
EAM	M	0,237	0,272	0,406	0,436	0,116	0,133	0,192	0,199
	DP	0,123	0,138	0,222	0,222	0,027	0,037	0,047	0,047
AD(2)-ADITIVO									
REQM	M	0,505	0,664	0,824	0,998	0,159	0,197	0,261	0,282
	DP	0,300	0,404	0,452	0,546	0,039	0,078	0,059	0,065
EAM	M	0,240	0,280	0,411	0,444	0,117	0,133	0,193	0,201
	DP	0,123	0,139	0,222	0,222	0,027	0,033	0,047	0,047
AR(2)-GARCH(1,1)									
REQM	M	0,279	0,350	0,281	0,351	0,329	0,503	0,328	0,501
	DP	0,141	0,178	0,142	0,179	0,130	0,261	0,129	0,258
EAM	M	0,143	0,163	0,143	0,163	0,180	0,231	0,179	0,229
	DP	0,065	0,071	0,065	0,071	0,053	0,087	0,052	0,085
AR(2)-GJR(1,1)									
REQM	M	0,283	0,354	0,285	0,355	0,334	0,511	0,333	0,509
	DP	0,144	0,190	0,145	0,191	0,131	0,264	0,130	0,258
EAM	M	0,145	0,165	0,146	0,165	0,182	0,234	0,181	0,232
	DP	0,067	0,075	0,068	0,075	0,053	0,087	0,052	0,084
AR(2)-TS-GARCH(1,1)									
REQM	M	0,394	0,502	0,398	0,508	0,182	0,246	0,180	0,243
	DP	0,211	0,284	0,214	0,288	0,067	0,117	0,066	0,113
EAM	M	0,185	0,215	0,187	0,217	0,115	0,139	0,114	0,137
	DP	0,083	0,098	0,084	0,100	0,031	0,048	0,030	0,046

Nota: *M* e *DP* indicam média e desvio padrão, respectivamente.

**Tabela A.2**

Estatísticas descritivas dos retornos diários do IBOVESPA

	03.01.95–04.02.02	03.01.95–24.01.03	03.01.95–26.11.07
Média	0,000326	0,000211	0,000667
Mediana	0,001232	0,000962	0,001419
Desvio padrão	0,024455	0,024047	0,021452
Assimetria	-0,245186	-0,231737	-0,282123
Curtose	5,058688	4,972813	5,361856
Máximo	0,093123	0,093123	0,093123
Mínimo	-0,103836	-0,103836	-0,103836
Observações	1.738	1.981	3.178
Jarque-Bera	324,33 (0,000)	338,98 (0,000)	780,83 (0,000)
$Q_1(1)$	17,67 (0,000)	17,61 (0,000)	19,18 (0,000)
$Q_1(6)$	26,18 (0,000)	25,43 (0,000)	31,09 (0,000)
$Q_1(25)$	56,07 (0,000)	52,24 (0,001)	63,23 (0,000)
$Q_2(1)$	177,44 (0,000)	186,01 (0,000)	292,48 (0,000)
$Q_2(6)$	622,73 (0,000)	675,20 (0,000)	1.138,01 (0,000)
$Q_2(25)$	1.175,79 (0,000)	1.285,50 (0,000)	2.334,95 (0,000)
ADF	-37,69 (0,000)	-40,49 (0,000)	-52,11 (0,000)

Nota: Números em parênteses indicam p-valores.  $Q_1(k)$  e  $Q_2(k)$  referem-se às estatísticas de Ljung-Box nas séries dos retornos e retornos ao quadrado com  $k$  defasagens, respectivamente.

**Tabela A.3**

Modelos para a esperança condicional do IBOVESPA

	03.01.95–04.02.02	03.01.95–24.01.03	03.01.95–26.11.07
AR(1)	-7.981,39 (3.992,69)	-9.162,13 (4.583,06)	-15.414,24 (7.709,12)
AR(2)	<b>-7.982,04 (3.994,02)</b>	<b>-9.163,00 (4.584,37)</b>	<b>-15.416,96 (7.711,48)</b>
$Q_1(6)$	7,58 (0,271)	7,13 (0,309)	9,42 (0,151)
$Q_1(12)$	19,33 (0,081)	16,91 (0,153)	23,00 (0,035)
$Q_1(25)$	35,66 (0,077)	32,49 (0,144)	39,74 (0,031)
$Q_2(6)$	629,55 (0,000)	678,51 (0,000)	1.136,70 (0,000)
$Q_2(12)$	902,73 (0,000)	981,34 (0,000)	1.694,56 (0,000)
$Q_2(25)$	1.212,71 (0,000)	1.316,39 (0,000)	2.368,02 (0,000)

Nota (1): Valores em negrito denotam as melhores medidas para AIC (log-verossimilhança).

Nota (2):  $Q_1(k)$  e  $Q_2(k)$  referem-se às estatísticas de Ljung-Box ( $k$  defasagens) para os resíduos e quadrados dos resíduos do modelo AR(2), respectivamente. Os p-valores estão entre parênteses.

**Tabela A.4**

Seleção dos modelos aditivos para o IBOVESPA (Amostra 1)

	$SQR$	$df^{res}$	Teste $F$	AIC
Modelo aditivo (1; 0,55)	0,0019650	1.725,22		-18.769,22
Modelo aditivo (2; 0,60)	0,0018740	1.721,11	20,32 (0,000)	-18.843,06
Modelo aditivo (3; 0,60)	0,0017952	1.716,65	16,93 (0,000)	-18.908,58
Modelo aditivo (4; 0,60)	0,0017696	1.712,20	5,56 (0,000)	-18.924,51
Modelo aditivo (5; 0,70)	0,0017606	1.711,44	11,52 (0,002)	-18.931,79

Nota (1): Nos modelos aditivos: (número de funções;  $span$ ).

Nota (2): No teste  $F$  os números em parênteses indicam os p-valores.

**Tabela A.5**

Seleção dos modelos GARCH para o IBOVESPA (Amostra 1)

Modelo	Normal			t-Student		
	AIC	BIC	Max [Log L]	AIC	BIC	Max [Log L]
ARCH(5)	-4,791641	-4,766507	4.171,94	-4,810022	-4,771746	4.188,91
ARCH(6)	-4,790010	-4,761734	4.171,52	-4,810598	-4,779180	4.190,41
ARCH(7)	-4,796750	-4,765332	4.178,38	<b>-4,813980</b>	<b>-4,779420</b>	<b>4.194,35</b>
GARCH(1,1)	-4,817056	-4,801347	4.191,02	-4,835027	-4,816176	4.207,64
GARCH(1,2)	-4,817729	-4,798878	4.192,61	-4,835385	-4,813392	4.208,95
TS-GARCH(1,1)	-4,712940	-4,697231	4.100,55	-4,799981	-4,781130	4.177,18
GJR(1,1)	-4,845881	-4,827030	4.217,07	<b>-4,858101</b>	<b>-4,836108</b>	<b>4.228,69</b>

Nota: Em negrito, os melhores modelos.

## Referências

- Bénéplanc, Gilles, & Rochet, Jean-Charles. 2011. *Risk Management in Turbulent Times*. Oxford: Oxford University Press.
- Bollerslev, Tim. 1986. Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, **31**, 307–327.
- Bollerslev, Tim. 2008. *Glossary to ARCH (GARCH)*. Disponível em <http://ssrn.com/abstract=1263250>. Acesso em: 20/09/2008.
- Bollerslev, Tim, Chou, Ray Y., & Kroner, Kenneth F. 1992. ARCH Modeling in Finance: A Review of the Theory and Empirical Evidence. *Journal of Econometrics*, **52**, 5–59.
- Bollerslev, Tim, Engle, Robert F., & Nelson, Daniel B. 1994. ARCH Models. Pages 2959–3038 of: Engle, R. F., & McFadden, D. L. (eds), *Handbook of Econometrics*. Amsterdam: Elsevier Sciences B. V.
- Bowman, Adrian W., & Azzalini, Adelchi. 1997. *Applied Smoothing Techniques for Data Analysis: The Kernel Approach with S-Plus Illustrations*. Oxford: Oxford University Press.
- Carroll, Raymond J., Härdle, Wolfgang, & Mammen, Enno. 2002. Estimation in an Additive Model When the Components are Linked Parametrically. *Econometric Theory*, **18**, 886–912.
- Chen, Song X., Delaigle, Aurore, & Hall, Peter. 2010. Nonparametric Estimation for a Class of Levy Processes. *Journal of Econometrics*, **157**, 257–271.
- Cleveland, William S., & Devlin, Susan J. 1988. Locally Weighted Regression: An Approach to Regression Analysis by Local Fitting. *Journal of the American Statistical Association*, **83**, 596–610.
- Diebold, Francis X., & Mariano, Roberto S. 1995. Comparing Predictive Accuracy. *Journal of Business and Economic Statistics*, **13**, 253–263.
- Engle, Robert F. 1982. Autoregressive Conditional Heteroskedasticity with Estimates of the Variances of UK Inflation. *Econometrica*, **50**, 987–1008.
- Fan, Jianqing, & Gijbels, Irene. 1996. *Local Polynomial Modeling and Its Applications*. London: Chapman and Hall.
- Fan, Jianqing, & Gu, Juan. 2003. Semiparametric Estimation of Value-at-Risk. *Econometrics Journal*, **6**, 261–290.
- Fan, Jianqing, & Yao, Qiwei. 1998. Efficient Estimation of Conditional Variance Functions in Stochastic Regression. *Biometrika*, **85**, 645–660.



- Fan, Jianqing, & Yao, Qiwei. 2003. *Nonlinear Time Series: Nonparametric and Parametric Methods*. New York: Springer-Verlag.
- Friedman, Jerome H., & Stuetzle, Werner. 1981. Projection Pursuit Regression. *Journal of the American Statistical Association*, **76**, 817–823.
- Gao, Jiti. 2007. *Nonlinear Time Series: Semiparametric and Nonparametric Methods*. Boca Raton: Chapman and Hall.
- Glosten, Lawrence R., Jagannathan, Ravi, & Runkle, David. 1993. On the Relation Between the Expected Value and the Volatility of the Nominal Excess Return on Stocks. *Journal of Finance*, **48**, 1779–1801.
- Hall, Peter, & Yao, Qiwei. 2003. Inference in ARCH and GARCH Models with Heavy-Tailed Errors. *Econometrica*, **71**, 285–317.
- Hastie, Trevor J. 1993. Generalized Additive Models. *Pages 249–307 of: Chambers, J. M., & Hastie, T. J. (eds), Statistical Models in S*. London: Chapman and Hall.
- Hastie, Trevor J., & Tibshirani, Robert J. 1990. *Generalized Additive Models*. London: Chapman and Hall.
- Horowitz, Joel L. 1998. *Semiparametric Methods in Econometrics*. New York: Springer-Verlag.
- Kim, Woocheol, & Linton, Oliver. 2004. The Live Method for Generalized Additive Volatility Models. *Econometric Theory*, **20**, 1094–1139.
- Kindleberger, Charles P. 2005. *Manias, Panics, and Crashes: A History of Financial Crises*. New York: Wiley.
- Krugman, Paul. 2009. *The Return of Depression Economics and the Crisis of 2008*. New York: Norton.
- Levine, Michael, & Li, Jinguang. 2007. *Local Instrumental Variable (LIVE) Method for the Generalized Additive-Interactive Nonlinear Volatility Model*. Disponível em <http://www.stat.purdue.edu/research/profiles/levine-ch1234ref.pdf>. Acesso em: 21/09/2008.
- Linton, Oliver, & Mammen, Enno. 2005. Estimating Semiparametric ARCH ( $\infty$ ) Models by Kernel Smoothing Methods. *Econometrica*, **73**, 771–836.
- Linton, Oliver, & Nielsen, Jens P. 1995. A Kernel Method of Estimating Structured Nonparametric Regression Based on Marginal Integration. *Biometrika*, **82**, 93–100.
- Linton, Oliver, Nielsen, Jens P., & Nielsen, Soren F. 2009. Non-Parametric Regression with a Latent Time Series. *The Econometrics Journal*, **12**, 187–207.

- Liu, Jun M., Chen, Rong, & Yao, Qiwei. 2010. Nonparametric Transfer Function Models. *Journal of Econometrics*, **157**, 151–164.
- Reinhart, Carmen M., & Rogoff, Kenneth S. 2009. *This Time is Different: Eight Centuries of Financial Folly*. New Jersey: Princeton University Press.
- Ruppert, David, Wand, Matthew P., Holst, Ulla, & Hössjer, Ola. 1997. Local Polynomial Variance Function Estimation. *Technometrics*, **39**, 262–273.
- Rydberg, Tina H. 2000. Realistic Statistical Modelling of Financial Data. *International Statistics Review*, **68**, 233–258.
- Schwert, William. 1989. Why Does Stock Market Volatility Change Over Time? *Journal of Finance*, **44**, 1115–1153.
- Stulz, René M. 2008. Risk Management Failures: What Are They and When Do They Happen? *Journal of Applied Corporate Finance*, **20**, 39–48.
- Taylor, Stephen. 1986. *Modelling Financial Time Series*. New York: Wiley.
- Tsay, Ruey S. 2010. *Analysis of Financial Time Series*. New York: Wiley.
- Wand, Matthew P., & Jones, Chris. 1995. *Kernel Smoothing*. London: Chapman and Hall.
- Ziegelmann, Flávio A. 2002. Estimation of Volatility Functions: The Local Exponential Estimator. *Econometric Theory*, **18**, 985–992.
- Ziegelmann, Flávio A. 2008. A Local Linear Least-Absolute-Deviations Estimator of Volatility. *Communications in Statistics: Simulation and Computation*, **37**, 1543–1564.
- Ziegelmann, Flávio A. 2011. Semiparametric Estimation of Volatility: Some Models and Complexity Choice in the Adaptive Functional-Coefficient Class. *Journal of Statistical Computation and Simulation*, **81**, 707–728.

