

Eixo Temático: Inovação e Sustentabilidade

**COMPARAÇÃO DA CAPACIDADE PREDITIVA DE MODELOS
HETEROCEDÁSTICOS ATRAVÉS DA ESTIMAÇÃO DO VALUE-AT-RISK**

**PREDICTIVE ABILITY COMPARISON OF HETEROSKEDASTIC MODELS BY
ESTIMATING THE VALUE-AT-RISK**

Raphael Silveira Amaro, Paulo Sergio Ceretta, Michele Gendelsky De Oliveira e Glauco Oliveira
Rodrigues

RESUMO

Tanto as empresas quanto os gestores de carteiras de investimentos precisam de um modelo que possa ser capaz de quantificar os riscos inerentes aos seus investimentos da melhor maneira possível. Atualmente, existem inúmeros modelos heterogêneos que buscam quantificar riscos, tornando a escolha de um determinado modelo bastante complexa. Com o intuito de confrontar e encontrar modelos que possam servir, de forma eficiente, à quantificação de riscos, o objetivo desta pesquisa é o de comparar a capacidade preditiva de cinco modelos de heterocedasticidade condicional através da estimação do *Value-at-Risk*, levando em consideração oito distribuições de probabilidade estatística diferentes, para as séries de índices financeiros do mercado de capitais dos cinco maiores países emergentes: Brasil, Rússia, Índia, China e África do Sul, no período compreendido entre 26 de fevereiro de 2001 e 31 de dezembro de 2015. Este trabalho estimou previsões do *Value-at-Risk* para 50 passos à frente, para 240 modelos concorrentes, com reajuste dos parâmetros a cada passo. As evidências encontradas sugerem que o modelo heterocedástico *Component GARCH* é preferível, para realizar previsões do *Value-at-Risk*, a todos os outros modelos concorrentes, porém a distribuição de probabilidade estatística que este modelo utiliza interfere demasiadamente nos resultados das previsões obtidas por ele.

Palavras-chave: *value-at-risk*, modelos heterocedásticos, distribuições de probabilidade, capacidade preditiva.

ABSTRACT

Both companies and investment portfolio managers need to have a model that can be able to quantify the risks inherent in their investments in the best possible way. Currently, there are several heterogeneous models which seek to quantify risk, making the choice of a particular model very complex. In order to confront and find models that can serve, efficiently, to the quantification of risk, the objective of this research is to compare the predictive ability of five models of conditional heteroskedasticity by estimating the *Value-at-Risk*, assuming eight different statistical probability distributions, for the series of financial ratios of the capital market of the five largest emerging countries: Brazil, Russia, India, China and South Africa, in the period between February 26, 2001 and December 31, 2015. This paper estimated forecasts of *Value-at-Risk* for 50 steps ahead, for 240 competing models, with adjustment os parameters at every step. The evidences suggests that heterocedastic *Component GARCH* is preferable, to make predictions of *Value-at-Risk*, to all other competing models, however the distribution of statistical probability that this model uses interferes too much in the results of forecasts obtained by it.

Keywords: *value-at-risk*, heteroskedastic models, probability distributions, predictive ability.

1 INTRODUÇÃO

Em um ambiente econômico cada vez mais competitivo, a gestão de risco torna-se indispensável para a sobrevivência de empresas e de gestores de carteiras de investimento. As empresas precisam tomar decisões a respeito de quanto e onde devem investir seus recursos, de forma a maximizar a probabilidade da relação entre perda e ganho. Esta tarefa é bastante complexa, uma vez que existe uma correlação positiva entre risco e retorno, isto é, quanto maior for o risco inerente ao investimento, maior deverá ser o seu retorno (MARKOWITZ, 1952). Grande parte dos gestores não conseguem efetuar uma gestão de risco consistente e acabam obtendo resultados não muito satisfatórios. A recente crise financeira global, que teve início no ano de 2008, fez com que diversas instituições financeiras decretassem falência, constituindo bons exemplos de empresas que falharam na sua gestão de risco (ORHAN; KÖKSAL, 2012).

As empresas em geral estão expostas ao risco de mercado, que representa as possíveis perdas financeiras associadas às oscilações do mercado. Uma medida popular e intuitiva para mensurar este tipo de risco é a denominada de *Value-at-Risk* (VaR). O VaR pode ser utilizado para mensurar o risco em uma variedade de instrumentos financeiros, como, por exemplo, ações, títulos da dívida, *commodities*, câmbio, opções, futuros, *forwards* e *swaps* (SADORSKY, 2006). A popularidade do VaR resulta da sua maneira simples de agregar vários componentes de risco de mercado em um único número (GENÇAY; SELÇUK; ULUGÜLYAĞCI, 2003). O cálculo da medida do VaR pode ser considerado uma função determinística da volatilidade de um ativo financeiro, representando, em sua forma paramétrica, uma aplicação natural dos modelos de volatilidade (GIOT; LAURENT, 2004).

Embora existam inúmeras pesquisas e trabalhos acadêmicos a respeito da volatilidade, ainda não há um consenso sobre qual o modelo mais adequado para a sua estimação. Em 1982, Engle demonstrou que algumas séries temporais podem apresentar heterocedasticidade condicional na sua volatilidade e, para que a variância da série seja estimada de forma mais precisa, propôs o modelo *Autoregressive Conditional Heteroscedastic* (ARCH). Durante as últimas décadas, inúmeros modelos, variantes da família ARCH, têm sido desenvolvidos, estimados e validados tanto de uma perspectiva teórica quanto de uma perspectiva empírica (BERNARDI; CATANIA, 2015). Para a estimação dos modelos de heterocedasticidade condicional é comum a utilização da distribuição de probabilidade Gaussiana, entretanto esta distribuição nem sempre é a mais adequada. Geralmente, as séries financeiras apresentam uma distribuição com caudas mais pesadas do que a distribuição Gaussiana e são, normalmente, leptocúrticas (VALLE; AZZALINI, 2013). Esta constatação demonstra a importância de se utilizarem distribuições diferentes da distribuição Gaussiana para a modelagem de séries.

Devido à disponibilidade de um grande número de modelos e de distribuições de probabilidade estatística, prever o futuro nível de volatilidade de um ativo financeiro distancia-se muito da trivialidade e avaliar o desempenho de previsão de um determinado modelo apresenta um desafio ainda maior (WENNSTRÖM, 2014). Na tentativa de contribuir com novos estudos sobre este tema, o objetivo desta pesquisa é comparar a capacidade preditiva de cinco modelos de heterocedasticidade condicional através da estimação do VaR, levando em consideração oito distribuições de probabilidade estatística diferentes.

As séries utilizadas para alcançar o objetivo proposto referem-se a cinco séries de índices financeiros, que representam o desempenho médio das cotações dos ativos de maior negociabilidade e representatividade dos mercados de ações, dos cinco maiores países emergentes, conhecidos como BRICS. Os países que compõem o termo BRICS são, respectivamente, Brasil, Rússia, Índia, China e África do Sul. Os dados das séries temporais correspondem a informações diárias, no período compreendido entre 26 de fevereiro de 2001 e 31 de dezembro de 2015, do preço de fechamento, em pontos, de cada índice financeiro. Para que tal objetivo seja alcançado, realizam-se previsões do VaR para 50 passos à frente, para

todos os modelos concorrentes, com reajuste dos parâmetros a cada passo. Uma vez que todas as previsões foram computadas para todos os passos à frente, é possível realizar a comparação da capacidade preditiva dos modelos concorrentes estudados por meio de uma determinada função de perda específica.

Este estudo contribui para a literatura existente pois fornece uma importante ferramenta de auxílio para que se possa calcular uma medida mais precisa da exposição ao risco de mercado dos cinco maiores países emergentes do mundo. As evidências encontradas podem servir para que as empresas, ou os gestores de carteira de investimentos, tomem decisões a respeito de quanto e onde devem investir seus recursos, de forma a maximizar a probabilidade da relação entre perda e ganho de seus investimentos. Uma maior eficiência no planejamento sobre a viabilidade de investimentos pode proporcionar para as empresas um incremento nas suas margens de lucro, que pode viabilizar novos investimentos e elevar a otimização dos níveis de qualidade do serviço prestado, tornando-as mais competitivas no mercado em que se inserem. Ademais, os resultados encontrados podem ser úteis para novos estudos estatísticos que proporcionem métodos alternativos de mensuração de riscos de mercado, ou utilizem modelos concorrentes heterogêneos aos que foram utilizados neste estudo.

2 VALUE-AT-RISK

A especificação do VaR, válida para qualquer função de distribuição de probabilidade, discreta ou contínua, pode ser expressa por meio de termos probabilísticos. Uma vez que r_t representa a série de retornos de um ativo financeiro no tempo t , o $VaR_t(\alpha)$ no percentil $(1 - \alpha)$ é definido por $Pr(r_t \leq VaR_t(\alpha)) = \alpha$, que calcula a probabilidade do retorno no tempo t ser menor do que, ou igual a, $VaR_t(\alpha)$, dado um nível de significância $\alpha \in (0, 1)$. Em termos probabilísticos, o VaR é simplesmente um quantil da distribuição de perda (MCNEIL; FREY; EMBRECHTS, 2005, p. 38).

O cálculo do VaR pode ser consideravelmente simplificado na medida em que se utiliza uma abordagem paramétrica. Para calcular o VaR, por meio da abordagem paramétrica, precisa-se transformar uma função de distribuição geral $f(w)$ em uma função de distribuição padrão $\phi(\varepsilon_t)$, em que ε_t possui média zero e desvio padrão unitário. Para isso, associa-se w^* com o retorno de corte r_t^* e o r_t^* com um desvio padrão normal δ : $1 - \alpha = \int_{-\infty}^{w^*} f(w)dw = \int_{-\infty}^{r_t^*} f(r_t)dr = \int_{-\infty}^{-\delta} \phi(\varepsilon_t)d\varepsilon$. Sendo assim, o problema de encontrar o VaR torna-se equivalente a encontrar o desvio δ de tal modo que a área à esquerda da distribuição seja $1 - \alpha$. Na medida em que se supõe que r_t segue um processo $r_t = \mu + \varepsilon_t$, onde ε_t utiliza uma função de distribuição ϕ com média zero e variância σ_t^2 , a medida do VaR pode ser calculada através da seguinte expressão matemática: $VaR_t(\alpha) = \mu + \sigma_t \phi^{-1}(\alpha)$, em que $\phi^{-1}(\alpha)$ representa o valor do q -quantil ($q = 1 - \alpha$) da função de distribuição ϕ , μ expressa a média e σ_t expõe o desvio-padrão (volatilidade) do ativo financeiro que pode ser estimado através de um modelo estatístico (GENÇAY; SELÇUK; ULUGÜLYAGCI, 2003), como, por exemplo, os modelos heterocedásticos descritos na seção posterior.

3 MODELOS HETEROCEDÁSTICOS

O modelo *Autoregressive Conditional Heteroscedasticity* (ARCH), introduzido por Engle (1982), foi o primeiro a fornecer um caminho para modelar a heterocedasticidade condicional na volatilidade, dando início a uma nova classe de modelos. Durante as últimas décadas, inúmeros modelos, variantes da família ARCH, têm sido desenvolvidos, estimados e validados tanto de uma perspectiva teórica quanto de uma perspectiva empírica. Neste trabalho,

optou-se por utilizar os seguintes modelos variantes: *Generalized ARCH* (GARCH), *Exponential GARCH* (EGARCH), *GJR GARCH* (GJR-GARCH), *Asymmetric Power ARCH* (APARCH) e *Component GARCH* (CGARCH). Estes modelos, por apresentarem uma estrutura matemática mais complexa, possuem capacidade de exibir características mais peculiares da volatilidade de ativos financeiros e, conseqüentemente, devem apresentar uma capacidade preditiva melhor do que a dos modelos mais simplificados.

O modelo *Generalized Autoregressive Conditional Heteroscedasticity* (GARCH), desenvolvido por Bollerslev (1986), representa uma extensão do modelo clássico ARCH, uma vez que expressa a variância condicional como uma função linear do quadrado dos erros passados e dos próprios valores passados da série temporal. Seja v_t uma sequência de variáveis aleatórias independente e identicamente distribuídas (i.i.d.), com uma função de distribuição de probabilidade η , um processo ε_t é denominado de GARCH(p, q) se $\varepsilon_t = \sigma_t v_t$, em que σ_t^2 representa a variância condicional e é uma função de $\{\varepsilon_s, s < t\}$, que pode ser descrita pela seguinte equação matemática: $\sigma_t^2 = \alpha_0 + \sum_{i=1}^p \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2$, $t \in \mathbb{Z}$, em que $\varepsilon_{t-i} = r_{t-i} - \mu$, para $i = 1, 2, \dots, p$ no tempo $t = 1, 2, \dots, T$, $\mu \in \mathbb{R}$ e expressa a média de r_t , α_0 é uma constante maior que zero, $0 \leq \alpha_i < 1, \forall i = 1, 2, \dots, p$ e $0 \leq \beta_j < 1, \forall j = 1, 2, \dots, q$ com $\sum_{i=1}^p \alpha_i + \sum_{j=1}^q \beta_j < 1$ para preservar a estacionariedade e positividade na variância condicional (FRANCQ; ZAKOIAN, 2010, p. 20).

O processo GARCH(p, q) é eficiente para modelar a persistência da volatilidade, mas ele apresenta um ponto fraco: a sua estrutura assume que choques positivos e negativos possuem o mesmo impacto na volatilidade atual do ativo financeiro, desconsiderando que as inovações podem apresentar impactos assimétricos na volatilidade (WENNSTRÖM, 2014). Com o intuito de ser capaz de modelar o comportamento assimétrico que as inovações causam na volatilidade dos ativos e de superar outros pontos fracos do modelo GARCH, Nelson (1991) propôs a primeira extensão do modelo de Bollerslev (1986), denominada de modelo *Exponential GARCH* (EGARCH). Seja v_t uma sequência de variáveis aleatórias i.i.d., com $E(v_t) = 0$ e $Var(v_t) = 1$, um processo ε_t é denominado de EGARCH(p, q) se $\varepsilon_t = \sigma_t v_t$, em que σ_t^2 representa a variância condicional e é uma função de $\{\varepsilon_s, s < t\}$, que pode ser descrita pela seguinte equação matemática: $\log(\sigma_t^2) = \alpha_0 + \sum_{i=1}^p \left[\alpha_i \frac{\varepsilon_{t-i}}{\sigma_{t-i}} + \gamma_i \left(\left| \frac{\varepsilon_{t-i}}{\sigma_{t-i}} \right| - E \left| \frac{\varepsilon_{t-i}}{\sigma_{t-i}} \right| \right) \right] + \sum_{j=1}^q \beta_j \log(\sigma_{t-j}^2)$, em que $\alpha_0, \alpha_i, \gamma_i$ e $\beta_j \in \mathbb{R}$ e o parâmetro γ_i representa a resposta da assimetria dos choques. Caso $\gamma_i < 0$, isto é, se caso $\varepsilon_{t-1} < 0$, o impacto de choques negativos sobre a volatilidade futura será maior do que o de choques positivos de mesma magnitude (WENNSTRÖM, 2014). Note que, diferentemente do modelo GARCH clássico, os parâmetros α_i, β_j e γ_i não possuem restrições impostas para a positividade na variância condicional, pois o modelo utiliza uma especificação logarítmica para ser estimado, impedindo que a variância seja negativa (FRANCQ; ZAKOIAN, 2010, p. 247).

Um caminho alternativo para modelar o impacto de choques assimétricos foi proposto por Glosten, Jagannathan e Runkle (1993) através do modelo denominado GJR-GARCH, que pode ser interpretado como um modelo *Threshold* (CHEN, 2013). Seja v_t uma sequência de variáveis aleatórias i.i.d. com $E(v_t) = 0$ e $Var(v_t) = 1$, um processo ε_t é denominado de GJR-GARCH(p, q) se $\varepsilon_t = \sigma_t v_t$, em que σ_t^2 representa a variância condicional e é uma função de $\{\varepsilon_s, s < t\}$, que pode ser descrita pela seguinte equação matemática: $\sigma_t^2 = \alpha_0 + \sum_{i=1}^p [\alpha_i \varepsilon_{t-i}^2 + \gamma_i \varepsilon_{t-i}^2 I_{(\varepsilon_{t-i} > 0)}] + \sum_{j=1}^q \beta_j \sigma_{t-j}^2$, em que $I_{(\cdot)}$ representa uma função que assume um valor unitário se $\varepsilon_{t-i} < 0, \forall i = 1, 2, \dots, p$ e valor zero se $\varepsilon_{t-i} > 0, \forall i = 1, 2, \dots, p$. Observe que, devido a presença de $I_{(\cdot)}$, a persistência da especificação do modelo depende da assimetria da distribuição condicional utilizada para modelar ε_t : $P = \sum_{i=1}^p \alpha_i + \sum_{j=1}^q \beta_j + \sum_{i=1}^p \gamma_i \mathbb{P}(\varepsilon_{t-i} \leq 0)$, onde $\mathbb{P}(\varepsilon_{t-i} \leq 0)$ representa a probabilidade de observar choques negativos. Além disso, as

condições $\alpha_0 > 0$, $\alpha_i \geq 0, \forall i = 1, 2, \dots, p$, $\gamma_i \geq 0, \forall i = 1, 2, \dots, p$ e $\beta_j \geq 0, \forall j = 1, 2, \dots, q$ são necessárias para garantir que a variância condicional seja positiva (BERNARDI; CATANIA, 2015). Percebe-se que o modelo especifica a variância condicional como uma função das partes positivas e negativas das inovações passadas (FRANCO; ZAKOIAN, 2010, p. 250).

O modelo *Asymmetric Power ARCH* (APARCH), em sua forma generalizada, foi proposto por Ding, Granger e Engle (1993) e apresenta uma fórmula flexível, capaz de produzir, com algumas restrições, as principais características de alguns modelos sugeridos anteriormente. Seja v_t uma sequência de variáveis aleatórias i.i.d. com $E(v_t) = 0$ e $Var(v_t) = 1$, um processo ε_t é denominado de APARCH(p, q) se $\varepsilon_t = \sigma_t v_t$, em que σ_t^2 representa a variância condicional e é uma função de $\{\varepsilon_s, s < t\}$, que pode ser descrita pela seguinte equação matemática: $\sigma_t^\delta = \alpha_0 + \sum_{i=1}^p \alpha_i (|\varepsilon_{t-i}| - \gamma_i \varepsilon_{t-i})^\delta + \sum_{j=1}^q \beta_j \sigma_{t-j}^\delta$, em que $\delta \in \mathbb{R}^+$ e incrementa a flexibilidade do modelo, permitindo a seleção de um poder arbitrário. Além disso, $\alpha_0 > 0$, $\alpha_i \geq 0, \forall i = 1, 2, \dots, p$, $\gamma_i \geq 0, \forall i = 1, 2, \dots, p$ e $\beta_j \geq 0, \forall j = 1, 2, \dots, q$ são as restrições impostas nos parâmetros para garantir a positividade da variância condicional (FRANCO; ZAKOIAN, 2010, p. 256).

Em 1993, Engle e Lee propuseram o modelo denominado *Component GARCH* (CGARCH), que decompõe a variância condicional em dois componentes: um componente permanente e outro transitório, de modo a investigar os componentes de longo e de curto prazo da volatilidade da série. Seja v_t uma sequência de variáveis aleatórias i.i.d. com $E(v_t) = 0$ e $Var(v_t) = 1$, um processo ε_t é denominado de CGARCH(p, q) se $\varepsilon_t = \sigma_t v_t$, em que σ_t^2 representa a variância condicional e é uma função de $\{\varepsilon_s, s < t\}$, que pode ser descrita pela seguinte equação matemática: $\sigma_t^2 = \omega_t + \sum_{i=1}^p \alpha_i (\varepsilon_{t-i}^2 - \omega_{t-i}) + \sum_{j=1}^q \beta_j (\sigma_{t-j}^2 - \omega_{t-i})$, em que $\omega_t = \alpha_0 + \rho \omega_{t-1} + \eta (\varepsilon_{t-1}^2 - \sigma_{t-1}^2)$, $\sum_{i=1}^p \alpha_i + \sum_{j=1}^q \beta_j < 1$ e $\rho < 1$ são condições necessárias para garantir a estacionariedade do processo. Percebe-se que o intercepto do modelo varia ao longo do tempo e segue um tipo dinâmico de autorregressivo de primeira ordem (BERNARDI; CATANIA, 2015). As restrições dos parâmetros para que a variância condicional seja positiva são expostas por Engle e Lee (1993).

4 DISTRIBUIÇÕES DE PROBABILIDADE ESTATÍSTICA

Por serem as distribuições mais populares utilizadas pelos econométricos para modelar o termo de erro ε_t dos modelos de heterocedasticidade condicional (BERNARDI; CATANIA, 2015), este trabalho, optou-se por utilizar as seguintes distribuições: *Normal Distribution*, *Skew Normal Distribution*, *Student-t Distribution*, *Skew Student-t Distribution*, *Generalized Error Distribution*, *Skew Generalized Error Distribution*, *Johnson's Su Distribution* e *Generalized Hyperbolic Distribution*.

A distribuição condicional mais popular é a Gaussiana, conhecida como *Normal Distribution* (Norm), que pode ser descrita pelos seus dois primeiros momentos: pela sua média e pela sua variância (WALCK, 2007, p. 119). Formalmente, a função de densidade de probabilidade (f.d.p.) de uma variável aleatória x que segue uma Norm pode ser expressa pela

notação $x \sim N(\mu, \sigma^2)$ e definida como: $f(x|\mu, \sigma) = \frac{e^{-0,5 \frac{(x-\mu)^2}{\sigma^2}}}{\sigma \sqrt{2\pi}}$, $x \in (-\infty, \infty)$, $\mu \in (-\infty, \infty)$, $\sigma > 0$, em que μ representa a média, σ^2 é a variância e $x \in \mathbb{R}$. A Norm geralmente é representada por $N(\mu, \sigma^2)$ e a sua função de distribuição cumulativa é dada por $F(x|\mu, \sigma) = \int_{-\infty}^x f(t|\mu, \sigma) dt$.

A suposição de simetria para os dados pode fazer com que sejam realizadas inferências errôneas sobre os parâmetros de interesse. Neste sentido, a fim de minimizar problemas devidos à assimetria dos dados, Azzalini (1985) propôs a *Skew Normal Distribution* (SNorm) como uma

generalização do modelo normal. Formalmente, a f.d.p. de uma variável aleatória x que segue uma SNorm pode ser expressa pela notação $x \sim SN(0, 1, k)$ e definida como: $f(x|0, 1, k) = 2\phi(x)\Phi(kx)$, em que x e $k \in \mathbb{R}$, $\phi(\cdot)$ representa a função de densidade normal padrão, $\Phi(\cdot)$ expõe a função de distribuição e k expressa o parâmetro de assimetria. Caso $k = 0$, a equação se reduz a uma Norm. Já se $k \rightarrow \pm\infty$, a SNorm torna-se a metade de uma Norm (ELING, 2014).

Outra distribuição bastante utilizada no campo estatístico é a *Student-t distribution* (Std). As características da Std são semelhantes às da distribuição Gaussiana, porém possuem caudas mais pesadas, possibilitando a utilização de valores mais extremos. Seja Z e S variáveis aleatórias independentes, de tal modo que $Z \sim N(0,1)$ e $nS^2 \sim \chi_n^2$, a distribuição de $t = Z/S$ é denominada de Std com $df = n$ (KRISHNAMOORTHY, 2006, p. 171). Deste modo, a f.d.p. de uma variável aleatória x que segue uma Std com n graus de liberdade pode ser expressa pela notação $x \sim t_n$ e definida como: $f(x|n) = \frac{\Gamma[(n+1)/2]}{\Gamma(n/2)\sqrt{n\pi}} \frac{1}{(1+x^2/n)^{(n+1)/2}}$, $x \in (-\infty, \infty)$, $n \geq 1$, onde n representa um número inteiro positivo, $x \in \mathbb{R}$ e $\Gamma(\cdot)$ representa a função gama definida por $\Gamma(x) = \int_0^\infty z^{x-1}e^{-z} dz$.

Para minimizar os problemas devidos à assimetria dos dados, Branco e Dey (2001) desenvolveram a *Skew Student-t Distribution* (SStd). A utilização da SStd mostra-se mais robusta e mais adequada que a SNorm, pois permite regular tanto a assimetria quanto a curtose, que são essenciais para modelagem de determinadas séries temporais. Seja Z e S variáveis aleatórias independentes, de tal modo que $Z \sim N(0, 1, k)$ e $nS^2 \sim \chi_n^2$, a distribuição de $t = Z/\sqrt{S/n}$ é denominada de SStd com $df = n$ (ELING, 2014). Deste modo, a f.d.p. de uma variável aleatória x que segue uma SStd com n graus de liberdade pode ser expressa pela notação $x \sim SStd(k, \xi)$ e definida como: $f(x|k, \xi) = 2t_n(x)T_{n+1}\left(kx\sqrt{\frac{1+\xi}{\xi+x^2}}\right)$, $x \in (-\infty, \infty)$, $n \geq 1$, em que $t_n(\cdot)$ representa a f.d.p. de uma Std padrão com n graus de liberdade, T_{n+1} a função de distribuição acumulada de uma Std padrão com $n + 1$ graus de liberdade, k expressa o parâmetro de assimetria e ξ o parâmetro de curtose. Caso $k = 0$, a equação se reduz a uma Std padrão. Já se $k \rightarrow \pm\infty$, a SStd torna-se a metade de uma Std.

Uma outra possibilidade de distribuição é a *Generalized Error Distribution* (Ged), introduzida por Subbotin (1923), e que pertence a uma família exponencial. A f.d.p. de uma variável aleatória x que segue uma Ged pode ser expressa pela notação $x \sim G(\mu, \sigma^2, k)$ e definida como: $df(x|\mu, \sigma, k) = \frac{e^{-\frac{1}{2}|x-\mu|k}}{2^{k+1}\sigma\Gamma(k+1)} dx$. Se $k = 1/2$, a equação representa uma f.d.p. de uma Norm univariada, isto é, $G\left(\mu, \sigma^2, \frac{1}{2}\right) = N(\mu, \sigma^2)$. Caso $k = 1$, a equação refere-se a uma f.d.p. de uma distribuição Laplace, $L(\mu, 4\sigma^2)$. Já no limite $k \rightarrow 0$ a f.d.p. tende a uma distribuição uniforme, $U(\mu - \sigma, \mu + \sigma)$.

Do mesmo modo que a SNorm e a SStd, a *Skew Generalized Error Distribution* (SGed), introduzida por Theodossiou (1998), se refere a uma extensão da Ged que objetiva regular a assimetria e a curtose dos dados, tornando-se mais robusta para ser utilizada em dados de séries financeiras. Formalmente, a f.d.p. de uma variável aleatória x que segue uma SGed pode ser expressa pela notação $x \sim SG(\mu, \sigma^2, k, \xi)$ e definida como: $f(x|\mu, \sigma, k, \xi) = \frac{C}{\sigma} \exp\left(-\frac{1}{[1-\text{sign}(x-\mu+\psi\sigma)k]^\xi \theta^\xi \sigma^\xi} |x - \mu + \psi\sigma|^\xi\right)$. Em que $C = \frac{\xi}{2\theta} \Gamma\left(\frac{1}{\xi}\right)^{-1}$, $\theta = \Gamma\left(\frac{1}{\xi}\right)^{\frac{1}{2}} \Gamma\left(\frac{3}{\xi}\right)^{-\frac{1}{2}} S(k)^{-1}$, $\psi = 2kAS(k)^{-1}$, $S(k) = \sqrt{1 + 3k^2 - 4A^2k^2}$, $A = \Gamma\left(\frac{2}{\xi}\right) \Gamma\left(\frac{1}{\xi}\right)^{-\frac{1}{2}} \Gamma\left(\frac{3}{\xi}\right)^{-\frac{1}{2}}$, μ representa a média, σ expõe a variância, $k \in (0, \infty)$ controla o grau de assimetria, ξ controla a curtose, isto é, a altura e as caudas da função de densidade, sign é a

função sinal e $\Gamma(\cdot)$ representa a função gama definida por $\Gamma(x) = \int_0^{\infty} z^{x-1} e^{-z} dz$ (THEODOSSIOU, 1998).

A *Johnson's Su Distribution* (Jsu), desenvolvida por Johnson (1949), pode ser obtida a partir de uma transformação de uma variável aleatória normal padrão. Formalmente, a f.d.p. de uma variável aleatória x que segue uma Jsu pode ser expressa pela notação $x \sim Jsu(\mu, \sigma^2, k, \xi)$ e definida como: $f(x|\mu, \sigma, k, \xi) = \frac{\xi}{\sigma \sqrt{1 + \left(\frac{x-\mu}{\sigma}\right)^2}} \phi \left[k + \xi \sinh^{-1} \left(\frac{x-\mu}{\sigma} \right) \right]$, $\sigma > 0$, em que

\sinh^{-1} é a inversa da função seno hiperbólico, ϕ é a f.d.p. de uma Norm, $N(0,1)$, e ξ , como dito anteriormente, deve ser interpretado como um parâmetro de curtose e deve ser positivo.

A *Generalized Hyperbolic Distribution* (Ghyp), introduzida por Barndorff e Halgreen (1977), é bastante utilizada no tratamento de dados financeiros, uma vez que os seus parâmetros estimados são suficientemente flexíveis a ponto de se adequarem aos mais variados conjuntos de dados e contextos. Formalmente, a f.d.p. de uma variável aleatória x que segue uma Ghyp pode ser expressa pela notação $x \sim Ghyp(\mu, \sigma^2, k, \xi, \beta)$ e definida como: $f(x|\mu, \sigma, k, \xi, \beta) = a(\mu, \sigma, k, \xi, \beta) [\sigma^2 + (x - \mu)^2]^{1/2\xi - 1/4} \cdot B\left(\xi - 0.5, \beta \sqrt{\sigma^2 + x^2 - 2x\mu + \mu^2}\right) e^{k(x-\mu)}$, em que

$a(\mu, \sigma, k, \xi, \beta) = \frac{(\beta^2 - k^2)^{1/2\xi}}{\sqrt{2\pi\beta^{\xi-1/2}\sigma^{\xi}B(\xi, \sigma\sqrt{\beta^2 - k^2})}}$, $B(\xi, \cdot)$ representa a função de Bessel modificada de

terceira espécie com índice ξ e $x \in \mathbb{R}$. O domínio de variação dos parâmetros é representado por $\mu \in \mathbb{R}$, assim como: $\sigma \geq 0$ e $|k| < \beta$ se $\xi > 0$; $\sigma > 0$ e $|k| < \beta$ se $\xi = 0$, $\sigma > 0$ e $|k| \leq \beta$ se $\xi < 0$ (EBERLEIN; HAMMERSTEIN, 2004).

5 ESTUDOS EMPÍRICOS ANTERIORES

Os trabalhos publicados sobre modelos heterocedásticos que serviram de base para a elaboração deste trabalho, encontram-se resumidos no Quadro 1.

Quadro 1 – Síntese de trabalhos sobre capacidade preditiva de modelos heterocedásticos.

Autores	Séries financeiras	Período	Modelos	D.P.E.	Principais resultados
Angelidis, Benos e Degiannakis (2004)	CAC 40, DAX 30, FTSE 100, NIKKEI 225 e S&P 500	09/07/1987 a 18/10/2002	GARCH, TARARCH e EGARCH	Norm, Std e Ged	A especificação do melhor modelo heterocedástico foi divergente para cada índice de ações utilizado. Já as distribuições Std e Ged são preferíveis à Norm para realizar previsões do VaR.
Orhan e Köksal (2012)	Índices do mercado acionário dos seguintes países: Brasil, Turquia, Alemanha e Estados Unidos.	01/01/2006 a 31/08/2009	ARCH, GARCH, IGARCH, AVGARCH, SAGARCH, TGARCH, GJR-GARCH, GJR-PGARCH, EGARCH, PGARCH, NGARCH, AGARCH, NGARCHK, APGARCH, NPGARCH e NPGARCHK	Norm, Std	A especificação ARCH obteve a melhor performance, seguido pelo modelo GARCH(1, 1), e a distribuição Std mostrou-se ligeiramente superior à Norm.

Bucevska (2012)	MBI 10	04/01/2005 a 31/10/2011	GARCH, EGARCH, GJR- GARCH, APARCH e GARCH-M	Norm e Std	O modelo EGARCH, utilizando tanto a distribuição Std quanto a distribuição Norm, e o modelo GJR-GARCH, utilizando a distribuição normal, mostraram-se adequados e com capacidades preditivas semelhantes, sugerindo que existe mais de um modelo adequado para estimar e prever a volatilidade do mercado de ações da Macedônia.
Wennström (2014)	OMXS30, OMXC20 e OMXH25	02/01/2002 a 15/04/2014	MA, EWMA, ARCH, GARCH, EGARCH e GJR-GARCH	Norm e Std	Não existe um único modelo preferível para prever a volatilidade dos três índices financeiros, com base em todas as quatro funções de perda utilizadas.
Bentes (2015)	Preço do Ouro no mercado mundial	02/08/1976 a 06/02/2015	GARCH, IGARCH e FIGARCH	Norm	O melhor modelo, tanto para capturar as dependências lineares da variância condicional quanto para realizar previsões da volatilidade, foi o modelo FIGARCH com ordem de integração 0,52166.
Bernardi e Catania (2016)	SXP1E, SXA1E, SXXP e SXW1E	31/12/1991 a 24/06/2014	GARCH, EGARCH, APARCH, AVARCH, GJR- GARCH, TGARCH, NGARCH, CGARCH, GAS, SAV e AS	Norm e Std	Durante a crise da dívida soberana europeia os modelos de volatilidade altamente não-lineares foram preferíveis para os índices europeus. Já com relação aos índices da América do Norte e da Ásia, os resultados obtidos das previsões foram bastante homogêneos por ambos os modelos comparados

D.P.E.: Distribuições de Probabilidade Estatística.

Fonte: elaboração dos autores.

6 MÉTODO E BASE DE DADOS

As séries financeiras utilizadas para alcançar o objetivo proposto referem-se a cinco séries de índices financeiros, que representam o desempenho médio das cotações dos ativos de maior negociabilidade e representatividade dos mercados de ações, dos cinco maiores países emergentes do mundo, conhecidos como BRICS. Os países, cujas letras iniciais de seus nomes compõem o termo BRICS são, respectivamente, Brasil, Rússia, Índia, China e África do Sul (*South Africa*). Deste modo, para indicar o desempenho médio do mercado acionário dos países em questão, os índices financeiros utilizados são representados, respectivamente, por: índice *Bovespa* (Ibov); índice *RTS* (Irts); índice *BSE Sensex 30* (Ibse); índice *SSE* (Isse); e índice *South Africa 40* (Isa40). Os dados das séries temporais foram coletados no sítio do Yahoo¹ e do Investing.com² e correspondem a informações diárias, no período compreendido entre 26 de fevereiro de 2001 e 31 de dezembro de 2015, do preço de fechamento, em pontos, de cada índice financeiro. A escolha da data inicial deve-se à disponibilidade dos dados para todas as séries utilizadas. Já a data final se justifica pelo período de elaboração deste trabalho.

¹ Disponível em <<http://finance.yahoo.com/>>.

² Disponível em <<http://www.investing.com/>>.

Outro ponto a destacar é o de que as séries temporais financeiras podem possuir propriedades estatísticas comuns a outras séries temporais, como, por exemplo, tendências, sazonalidades, pontos influentes e não-linearidade (MORETTIN, 2008, p. 15). Em vista disto, com o intuito de minimizar ou até mesmo eliminar algumas destas características, utilizam-se dos retornos diários das séries temporais financeiras citadas, que podem ser definidos pela primeira diferença dos logaritmos neperianos dos preços dos índices financeiros.

Após a especificação das séries temporais financeiras, inicia-se o método aplicado para que se possa alcançar o objetivo proposto. Este processo é composto por quatro etapas. A primeira etapa consiste em analisar se as séries temporais possuem mudanças estruturais ao longo do tempo e, em caso afirmativo, precisa-se escolher um período em que os parâmetros estimados pelos modelos concorrentes sejam válidos para todas as observações da amostra. Para verificar a presença de mudanças estruturais nas séries, realizam-se os seguintes testes estatísticos: *Cumulative Sum Control Chart* (CUSUM), desenvolvido por Brown, Durbin e Evans (1975); *Moving Sums of Residuals* (MOSUM), proposto por Bauer e Hackl (1978); *supF*, desenvolvido por Andrews (1993); *aveF* e *expF*, propostos por Andrews e Ploberger (1994).

A segunda etapa consiste em verificar se as séries temporais são geradas por processos estocásticos estacionários, uma vez que a condição de estacionariedade nas séries é indispensável para que seja possível realizar inferências estatísticas para períodos de tempo fora do conjunto de dados observado. Para verificar se as séries apresentam desvios aparentes da condição de estacionariedade, este estudo utiliza-se dos seguintes testes estatísticos: *Augmented Dickey-Fuller* (ADF), desenvolvido por Dickey e Fuller (1981); e *Kwiatkowski-Phillips-Schmidt-Shin* (KPSS), desenvolvido por Kwiatkowski, Phillips, Schmidt e Shin (1992).

Cada série temporal completa é composta de um número T de observações e é representada por X_1, X_2, \dots, X_T . O número total de observações se divide em dois subconjuntos: $\{X_1, X_2, \dots, X_n\}$ e $\{X_{n+1}, X_{n+2}, \dots, X_T\}$, em que n representa o início da previsão. O primeiro subconjunto refere-se ao período denominado dentro da amostra, em que os parâmetros dos modelos concorrentes são estimados. O segundo subconjunto refere-se ao período denominado fora da amostra, em que as previsões dos modelos concorrentes são realizadas e avaliadas. Após a divisão da amostra, inicia-se a terceira etapa, que consiste na estimação dos parâmetros dos modelos e na previsão do VaR. Este processo se inicia com os seguintes passos: (a) define-se $z = n$ para ser o ponto de origem da previsão. Logo após, estimam-se os parâmetros dos modelos concorrentes usando o subconjunto $\{X_1, X_2, \dots, X_z\}$; (b) realiza-se a previsão para um passo à frente, para cada modelo concorrente, usando os seus parâmetros estimados no passo (a); e (c) aumenta-se a origem da previsão em uma observação ($z = z + 1$) e recomeça-se o processo no passo (a). Este procedimento é repetido até que se chegue na última observação disponível da série temporal, em que a origem da previsão z seja igual ao ponto T .

Conhecidos os passos do processo de previsão, realizam-se previsões para 50 passos à frente com reajuste dos parâmetros a cada passo para todos os modelos concorrentes. A escolha de 50 observações para o período para fora da amostra estabeleceu-se de forma arbitrária, justificando-se por apresentar um número suficiente, sem prejudicar o período de estimação dos parâmetros, para mensurar o desempenho de previsão dos modelos. Uma vez que os modelos heterocedásticos utilizados, para realizar previsões, referem-se a processos que podem ser representados por diferentes graus de polinômios (p, q) , neste trabalho, respeitam-se as seguintes limitações: $1 \geq p \geq 2$ e $1 \geq q \geq 2$. Estas limitações impostas fazem com que o número total de modelos concorrentes seja limitado em 240, o que resulta em uma matriz com um total de 12.000 previsões estimadas para cada série temporal utilizada. Totalizando, para todas as cinco séries temporais, 60.000 previsões.

Após todas as previsões serem computadas para todos os passos à frente, inicia-se a quarta etapa, que consiste na comparação da capacidade preditiva dos modelos a partir de medidas estatísticas estabelecidas. Neste trabalho, utilizam-se três funções de perda para

avaliação do desempenho de previsão dos modelos: *Mean Squared Error* ($MSE = \frac{1}{n} \sum_{t=1}^n (\hat{X}_t - X_t)^2$); *Root Mean Square Error* ($RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (\hat{X}_t - X_t)^2}$); e *Mean Absolute Deviation* ($MAD = \frac{1}{n} \sum_{t=1}^n |\hat{X}_t - X_t|$). Em que \hat{X}_t é o vetor de n previsões no tempo t e X_t é o vetor dos valores observados no tempo t que correspondem às entradas para a função que gerou as previsões. Percebe-se que as funções de perda utilizadas possuem como intuito avaliar a qualidade das previsões obtidas pelos modelos concorrentes por meio do cálculo do erro das previsões. Sendo assim, o modelo que fornecer os menores valores para as funções estabelecidas possuirá a melhor capacidade preditiva. Para que sejam realizadas as etapas 3 e 4, este trabalho faz uso do software livre R³, versão 3.2.4 *Revised*, e de um código-fonte de programação próprio, desenvolvido pelos autores, que pode ser solicitado por e-mail.

7 EVIDÊNCIAS EMPÍRICAS

Nesta seção, as evidências empíricas da performance de previsão do VaR são examinadas em detalhes. Inicia-se expondo, através da Figura 1, a evolução dos valores do preço de fechamento ao longo do tempo, em pontos e em primeira diferença dos logaritmos neperianos, para cada índice financeiro em estudo, no período compreendido entre 26 de fevereiro de 2001 e 31 de dezembro de 2015. Nota-se que as séries Ibov, Irts, Ibse, Isse e Isa40 apresentaram uma valorização nominal expressiva, totalizando, respectivamente, 172,80%, 353,81%, 535,05%, 81,06% e 419,90% no acumulado do período total. No entanto, a evolução destes crescimentos nominais mostrou-se bastante volátil ao longo do tempo, exibindo um comportamento com grandes oscilações de valores.

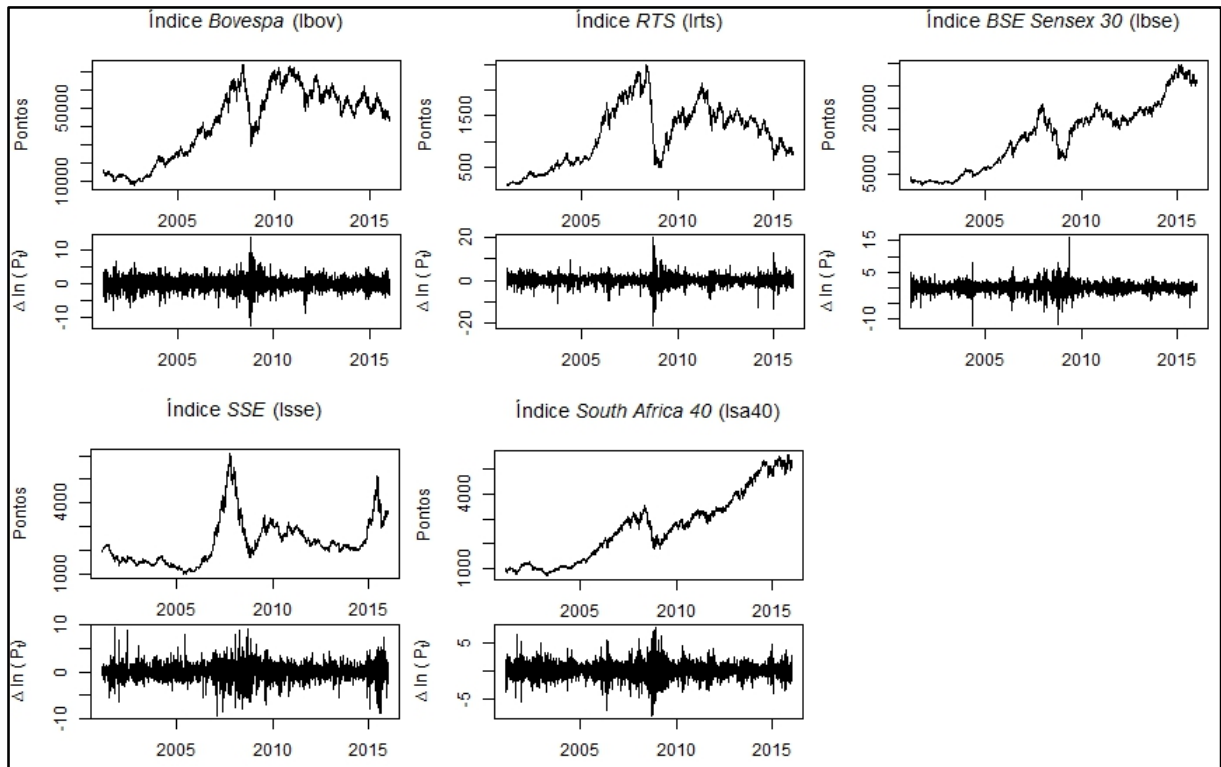
No ano de 2008, provavelmente pelo efeito relativo ao início da crise financeira mundial, percebe-se a existência de variações negativas expressivas nos preços de fechamento de todos os índices financeiros, levando-os a grandes desvalorizações em um período relativamente curto de tempo. As intensidades destas desvalorizações podem ser verificadas por meio dos vales deixados pelos gráficos dos índices no período em questão. Já a velocidade com que ocorreram as desvalorizações pode ser associada a um possível efeito de pânico nos mercados financeiros, que pode ser identificado pelo aumento da volatilidade no período. Embora tenham apresentado uma queda significativa dos seus valores no ano de 2008, no acumulado do período total estudado, os índices financeiros apresentaram um crescimento nominal acentuado, porém com comportamentos distintos entre eles ao longo do tempo.

Após a inspeção visual do comportamento da evolução dos valores dos preços de fechamento dos cinco índices financeiros, foi realizada a análise sobre a existência de mudanças estruturais ao longo do tempo em todas as séries financeiras. Como resultado, os cinco testes aplicados rejeitaram a hipótese nula da não existência de mudanças estruturais ao longo do tempo. Estes resultados evidenciam que os parâmetros dos modelos econométricos estimados são inapropriados para realizarem inferências estatísticas para todos os períodos das séries temporais estudadas. Deste modo, precisa-se identificar quantas mudanças estruturais ocorreram ao longo do tempo nas séries financeiras e definir em qual período devem ser realizadas as previsões do VaR utilizando os modelos concorrentes. A identificação do número de mudanças estruturais foi realizada através da estatística da *Residual Sum of Square* (RSS) e do critério de informação *Bayes Information Criterion* (BIC). Os resultados dos testes mostraram que a melhor escolha para as séries Ibov, Irts, Ibse, Isse e Isa40 se refere, respectivamente, a 4, 5, 4, 5 e 5 mudanças estruturais. Deste modo, utilizou-se neste trabalho o

³ Disponível em <<http://www.R-project.org/>>.

último período identificado após a última mudança estrutural de cada série financeira para estimar os modelos concorrentes e realizar as previsões para o VaR.

Figura 1 – Evolução dos valores do preço de fechamento de cada índice financeiro, no período compreendido entre 26 de fevereiro de 2001 e 31 de dezembro de 2015.



Fonte: elaboração dos autores.

As estatísticas descritivas das séries temporais encontram-se apresentadas na Tabela 2. Como esperado, percebem-se evidências de não-normalidade nas funções de distribuição de todos os índices financeiros, pois ambos os índices apresentaram excesso de curtose e um pouco de assimetria na sua distribuição. Estes resultados demonstraram que as séries possuem caudas mais pesadas do que a distribuição Gaussiana e são leptocúrticas. Além disso, o teste Jarque-Bera rejeita a hipótese nula de normalidade para quase todas as séries financeiras, com exceção de alguns períodos para fora da amostra. Entende-se que a aceitação da normalidade para os períodos para fora da amostra ocorre pelo fato de que eles possuem um número muito pequeno de observações. Em 2007, Thadewald e Büning demonstraram que, quando as amostras são pequenas, o teste de normalidade de Jarque-Bera possui uma baixa performance, apresentando um baixo percentual de acertos. Os resultados obtidos pelos testes de identificação de estacionariedade, ADF e KPSS, mostraram que todas as séries possuem uma ordem de integração unitária, isto é, todos os índices seguem um processo estocásticos estacionário apenas em sua primeira diferença dos logaritmos neperianos, com 1% de significância.

Uma vez que os comportamentos das séries foram analisados, foi realizado o processo descrito para obter as previsões das estimativas do VaR. Após todas as previsões serem computadas, iniciou-se o processo de comparação da capacidade preditiva dos modelos heterocedásticos por meio do cálculo das funções de perda. A Tabela 3 apresenta de forma resumida os cinco melhores resultados obtidos, classificados por ordem crescente dos valores estimados da estatística da função de perda RMSE, para todas as séries temporais. Devido a

limitação de espaço, os resultados encontrados para todas as funções de perda para todos os modelos concorrentes não serão apresentados, porém podem ser solicitados para os autores.

Tabela 1 – Estatísticas descritivas das séries temporais financeiras para o período identificado após a última mudança estrutural constatada.

Série	Obs.	Média	Mediana	Máximo	Mínimo	DP	Assimetria	Exc. Curtose	JB
<i>Dentro da amostra</i>									
Ibov	852	53.382,43	53.634,50	63.312,00	43.957,00	4.147,76	0,00	2,31	16,86
Irts	506	1.096,94	1.076,37	1.518,54	629,15	226,10	0,09	1,73	34,82
Ibse	500	25.320,18	26.354,72	29.681,77	19.379,77	2.965,15	-0,55	1,82	54,22
Isse	488	2.808,90	2.334,91	5.166,35	1.991,25	868,09	0,94	2,71	74,20
Isa40	545	4.948,28	4.996,50	5.510,80	4.261,40	308,67	-0,48	2,18	36,40
Série	Obs.	Média	Mediana	Máximo	Mínimo	DP	Assimetria	Exc. Curtose	JB
<i>Fora da amostra</i>									
Ibov	50	46.099,38	46.294,00	48.284,00	43.200,00	1.436,53	-0,39	2,09	2,98
Irts	50	826,69	839,20	897,56	756,07	42,92	-0,23	1,68	4,06
Ibse	50	26.066,40	25.877,55	27.470,81	25.036,05	638,92	0,71	2,78	4,32
Isse	50	3.520,41	3.535,35	3.651,77	3.316,70	100,17	-0,42	2,01	3,50
Isa40	50	5.264,51	5.275,05	5.550,00	4.912,20	171,53	-0,28	2,19	2,04

DP: desvio padrão; JB: Jarque-Bera.

Fonte: elaboração dos autores.

As evidências revelaram uma certa divergência entre os melhores modelos encontrados para os cinco índices financeiros estudados. Para o Ibov os modelos EGARCH e CGARCH apresentaram as melhores capacidades preditivas para todas as três funções de perda, ficando nas primeiras posições do ranking dos modelos com as melhores previsões realizadas, uma vez que estes modelos apresentaram os menores erros nas previsões. Estas evidências encontradas demonstram que os modelos com estruturas matemáticas mais complexas, que levam em conta características mais peculiares da volatilidade dos índices financeiros, possuem uma capacidade superior, de previsão do VaR, à dos modelos com estruturas matemáticas mais simplificadas. Com relação às distribuições de probabilidade estatística, nota-se que as distribuições que se sobressaíram no Ibov foram as assimétricas SStd e SNorm, demonstrando que a função de densidade da distribuição utilizada em tais modelos precisa considerar a existência de assimetria nos dados para que não sejam feitas inferências errôneas sobre os parâmetros de interesse.

Em referência ao Irts, os modelos que revelaram os melhores desempenhos nas previsões realizadas do VaR foram o GJR-GARCH, que permite modelar o impacto assimétrico dos choques através de uma função binária, e o CGARCH, que permite investigar os componentes de longo e de curto prazo da volatilidade do índice financeiro. Para ambos os modelos citados, a distribuição Norm mostrou-se superior a todas as outras distribuições, contradizendo as evidências preliminares obtidas por meio da análise das estatísticas descritivas da série em questão. Já nas evidências encontradas pelos índices Ibse, Isse e Isa40, percebe-se uma certa superioridade do modelo CGARCH para realizar previsões do VaR. Nestes índices citados, o modelo CGARCH ocupou as primeiras posições do ranking para todas as três funções de perda, divergindo apenas pelas distribuições de probabilidade estatística e pelos graus dos polinômios utilizados nos modelos concorrentes. Os resultados encontrados indicaram, mais uma vez, que os modelos que possuem estruturas mais complexas são superiores, para realizar previsões do VaR, aos modelos que possuem estruturas mais simplificadas, como o GARCH.

Tabela 2 – Resumo dos resultados obtidos pelas três funções de perda utilizadas, classificados por ordem crescente de valores da função de perda *Root Mean Square Error* (RMSE).

Série temporal/Modelos	RMSE		MSE		MAD	
	Estatística	Posição	Estatística	Posição	Estatística	Posição
Ibov						
EGARCH-SStd (1,1)	0,036407	1	0,001325	1	0,032804	1
CGARCH-SNorm (2,1)	0,036722	2	0,001348	2	0,033249	4
CGARCH-SNorm (2,2)	0,036810	3	0,001355	3	0,033098	2
EGARCH-SStd (1,2)	0,036839	4	0,001357	4	0,033321	5
GARCH-SNorm (2,1)	0,036860	5	0,001359	5	0,033376	7
Irts						
GJR-GARCH-Norm (1,2)	0,046936	1	0,002203	1	0,042928	2
GJR-GARCH-Norm (1,1)	0,047021	2	0,002211	2	0,043008	3
CGARCH-Norm (1,2)	0,047157	3	0,002224	3	0,043159	5
CGARCH-Norm (2,1)	0,047248	4	0,002232	4	0,042772	1
CGARCH-Norm (1,1)	0,047290	5	0,002236	5	0,043247	6
Ibse						
CGARCH-Norm (2,2)	0,020353	1	0,000414	1	0,018942	1
CGARCH-Norm (1,1)	0,020536	2	0,000422	2	0,019196	2
CGARCH-Norm (1,2)	0,020559	3	0,000423	3	0,019207	3
CGARCH-Norm (2,1)	0,020798	4	0,000433	4	0,019485	4
GARCH-Norm (2,1)	0,021218	5	0,000450	5	0,019776	5
Isse						
CGARCH-Jsu (2,2)	0,189343	1	0,035851	1	0,082784	1
CGARCH-Jsu (2,1)	0,189344	2	0,035851	2	0,082784	2
CGARCH-SStd (2,2)	0,206110	3	0,042481	3	0,083883	3
CGARCH-SStd (2,1)	0,206114	4	0,042483	4	0,083884	4
CGARCH-Ghyp (2,2)	0,208852	5	0,043619	5	0,085426	8
Isa40						
CGARCH-Ged (1,1)	0,028816	1	0,000830	1	0,026134	3
CGARCH-Ged (2,1)	0,028865	2	0,000833	2	0,026227	7
CGARCH-Ged (1,2)	0,028872	3	0,000834	3	0,026217	6
CGARCH-SNorm (1,2)	0,028895	4	0,000835	4	0,026197	4
CGARCH-SNorm (1,1)	0,028912	5	0,000836	5	0,026106	2

Fonte: elaboração dos autores.

As distribuições de probabilidade apresentaram um resultado significativamente diferente entre os índices Ibse, Isse e Isa. Para o Ibse a distribuição Norm mostrou-se mais relevante, ocupando as primeiras posições do ranking, o que contradiz, novamente, as evidências encontradas na análise das estatísticas descritivas da série em questão. Já para os índices Isse e Isa40 não existe uma distribuição predominantemente superior a todas as outras. Os modelos que ocuparam as primeiras posições do ranking do Isse utilizaram as distribuições Jsu, SStd e Ghyp, mostrando que a função de densidade da distribuição utilizada precisa empregar outros parâmetros que não estão presentes nas funções de densidade das distribuições com estruturas mais simplificadas. Já os modelos que ocuparam as primeiras posições do

ranking do Is40 utilizaram a distribuição Ged seguida pela distribuição assimétrica SNorm, evidenciando que o parâmetro que determina o grau de assimetria da distribuição é indispensável para que não sejam feitas inferências errôneas sobre os dados utilizados.

Na análise dos resultados divergentes encontrados, entre os índices financeiros, sobre os graus dos polinômios utilizados nos modelos heterocedásticos concorrentes, percebe-se que não existe uma predominância de qual grau é o ideal para empregar na estimação de um determinado modelo. Cada índice respondeu de maneira diferente aos graus dos polinômios utilizados nos modelos, tornando o resultado, de certa forma, inconclusivo.

8 CONSIDERAÇÕES FINAIS

Este trabalho teve como objetivo comparar a capacidade preditiva de cinco modelos de heterocedasticidade condicional através da estimação do VaR, levando em consideração oito distribuições de probabilidade estatística diferentes, para os índices financeiros dos mercados de ações dos cinco maiores países emergentes. A análise *a priori* indicou que os índices possuem diferentes comportamentos em seus preços e em suas volatilidades ao longo do tempo, sinalizando que, embora em alguns momentos eles possam caminhar para uma direção em comum, eles possuem, na maior parte do tempo, observações que seguem caminhos e intensidades heterógenas entre si.

A tarefa de fornecer um modelo ideal perfeito para que se possa quantificar todos os riscos inerentes ao processo de investimento nos países analisados mostrou-se utópica. Embora se possa dizer que o modelo heterocedástico CGARCH se sobressaiu, entre todos os modelos concorrentes, como o modelo que melhor prevê a estimação do VaR para os cinco índices financeiros, a distribuição de probabilidade estatística que este modelo utiliza interfere demasiadamente nos resultados das previsões obtidas por ele.

Os dados de cada índice financeiro estudado mostraram-se adequar-se a um determinado tipo de função de densidade de probabilidade diferente, não refletindo uma distribuição que possa ser considerada superior a todas as outras, o que impossibilita a indicação de uma determinada distribuição ideal a ser utilizada como padrão nos modelos de mensuração de risco dos investimentos realizados nos índices dos países emergentes. Ademais, também mostrou-se inconclusiva a utilização de diferentes graus de polinômios na estimação dos modelos heterocedásticos, pois em nenhum resultado encontrado existiu uma predominância de qual grau é o melhor para ser utilizado na estimação de um determinado modelo concorrente. Desta forma, em resumo, pode-se concluir que, entre os 240 modelos concorrentes analisados, não existe um único modelo heterocedástico que, utilizando uma determinada distribuição de probabilidade estatística e um número específico de graus de polinômios (p, q), possua uma capacidade preditiva superior a todos os outros modelos para estimar o VaR dos índices financeiros do mercado de ações dos principais países emergentes.

REFERÊNCIAS

- ANDREWS, D. W. K. Tests for parameter instability and structural change with unknown change point. **Econometrica**, v. 61, n. 4, p. 821-856, July 1993.
- ANDREWS, D. W. K.; PLOBERGER, W. Optimal tests when a nuisance parameter is present only under the alternative. **Econometrica**, v. 62, n. 6, p. 1383-1414, Nov. 1994.
- ANGELIDIS, T.; BENOS, A.; DEGIANNAKIS, S. The use of GARCH models in VaR estimation. **Statistical Methodology**, v. 1, n. 1-2, p. 105-128, Dec. 2004.
- AZZALINI, A. A class of distributions which includes the normal ones. **Scandinavian Journal of Statistics**, v. 12, n. 1, p. 171-178, 1985.

- BARNDORFF, N. O.; HALGREEN, C. Infinite divisibility of the hyperbolic and generalized inverse Gaussian distributions. **Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete**, v. 38, n. 4, p. 309-311, dec. 1977.
- BAUER, P.; HACKL, P. The use of Mosums for quality control. **Technometrics**, v. 20, n. 4, p. 431-436, nov. 1978.
- BENTES, S. R. Forecasting volatility in gold returns under the GARCH, IGARCH and FIGARCH frameworks: New evidence. **Physica A: Statistical Mechanics and its Applications**, v. 438, p. 355-364, nov. 2015.
- BERNARDI, M.; CATANIA, L. Comparison of Value-at-Risk models using the MCS approach. **Computational Statistics**, v. 31, n. 2, p. 579-608, feb. 2016.
- BERNARDI, M.; CATANIA, L. The Model Confidence Set package for R. **CEIS Tor Vergata No. 362**, v. 13, n. 11, nov. 2015.
- BOLLERSLEV, T. Generalized autoregressive conditional heteroskedasticity. **Journal of Econometrics**, v. 31, n. 3, p. 307-327, apr. 1986.
- BOLLERSLEV, T.; ENGLE, R.F.; NELSON, D. Arch models. **Handbook of Econometrics**. 1. ed. United States: Elsevier, 1994, v. 4, cap. 49, editado por Engle, R.F e McFadden D.L, p. 2959-3038.
- BRANCO, M. D.; DEY, D. K. A general class of multivariate skew-elliptical distributions. **Journal of Multivariate Analysis**, v. 79, n. 1, p. 99-113, oct. 2001.
- BROWN, R. L.; DURBIN, J.; EVANS, J. M. Techniques for testing the constancy of regression relationships over time. **Journal of the Royal Statistical Society B**, v. 37, n. 2, p. 149-163, 1975.
- BUCEVSKA, V. An Empirical Evaluation of GARCH Models in Value-at-Risk Estimation: Evidence from the Macedonian Stock Exchange. **Business System Research**, v. 4, n. 1, p. 49-64, dec. 2012.
- CHEN, M. Y. **Time Series Analysis: Conditional Volatility Models**. 1. ed. Taiwan: Department of Finance. National Chung Hsing University, feb. 2015. 47 p.
- DAVIS, R. **ARMA-GARCH models applied to Exchange-Traded Funds**. 2012. 75 p. Dissertação (Master of Science) - Department of Mathematical Sciences. The University of Texas at El Paso, Texas, United States, 2012.
- DICKEY, D. A.; W. A. FULLER. Likelihood ratio statistics for autoregressive time series with a unit root. **Econometrica**, v. 49, n. 4, p. 1057-1072, 1981.
- DING, Z.; GRANGER, C. W. J.; ENGLE, R. F. A long memory property of stock Market returns and a new model. **Journal of Empirical Finance**, v. 1, n. 1, p. 83-106, june 1993.
- EBERLEIN, E.; HAMMERSTEIN, E. A. V. Generalized Hyperbolic and Inverse Gaussian Distributions: Limiting Cases and Approximation of Processes. **Seminar on Stochastic Analysis, Random Fields and Applications IV**, Volume 58 da série Progress in Probability. Birkhäuser, Basel, 2004, p. 221-264.
- ELING, M. Fitting asset returns to skewed distributions: Are the skew-normal and skew-student good models?. **Insurance: Mathematics and Economics**, v. 59, p. 45-56, nov. 2014.
- ENGLE, R. F. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. **Econometrica**, v. 50, n. 4, p. 987-1007, july 1982.
- ENGLE, R. F.; Lee, G. G. **A permanent and transitory component model of stock return volatility**. University of California at San Diego, Economics Working Paper Series, 1993.
- FRANCQ, C.; ZAKOIAN, J. M. **GARCH Models: Structure, Statistical Inference and Financial Applications**. 1. ed. United Kingdom: Wiley, Chichester, 2010. 489 p.
- GENÇAY, R.; SELÇUK, F.; ULUGÜLYAĞCI, A. High volatility, thick tails and extreme value theory in Value-at-Risk estimation. **Insurance: Mathematics and Economics**, v. 33, n. 2, p. 337-356, oct. 2003.
- GILLER, G. L. **A Generalized Error Distribution**. Giller Investments Research Note, 2005.

- GIOT, P., LAURENT, S. Modelling daily Value-at-Risk using realized volatility and ARCH type models. **Journal of Empirical Finance**, v. 11, n. 3, p. 379-398, June 2004.
- GLOSTEN, L. R.; JAGANATHAN, R.; RUNKLE, D. E. On the relation between the expected value and the volatility of the nominal excess returns on stocks. **Journal of Finance**, v. 48, n. 5, p. 1779-1801, Dec. 1993.
- INVESTING.COM, 2016. Disponível em: <<http://www.investing.com/>>. Acesso em: 15 Jan. 2016.
- JOHNSON, N. L. Systems of frequency curves generated by methods of translation. **Biometrika**, v. 36, n. 1-2, p. 149-176, 1949.
- JORION, P. **Value at Risk: The New Benchmark for Managing Financial Risk**. 3. ed. United States: The McGraw-Hill Companies, 2007. 602 p.
- KRISHNAMOORTHY, K. **Handbook of statistical distributions with applications**. 1. ed. United States: Boca Raton: Chapman & Hall/CRC, 2006. 344 p.
- KWIATKOWSKI, D.; P. C. B. PHILLIPS; P. SCHMIDT; Y. SHIN. Testing the Null Hypothesis of Stationarity against the Alternative of a Unit Root. **Journal of Econometrics**, v. 54, n. 1-3, p. 159-178, Dec. 1992.
- MARKOWITZ, H. Portfolio selection. **Journal of Finance**, v. 7, n. 1, p. 77-99, 1952.
- MCNEIL, A. J.; FREY, R.; EMBRECHTS, P. **Quantitative Risk Management: Concepts, Techniques, and Tools**. 1st ed. United States: Princeton University Press, 2005. 538 p.
- MORETTIN, P. A. **Econometria financeira: um curso de séries temporais financeiras**. 1. ed. Brasil: São Paulo, Blucher, 2008. 319 p.
- NECULA, C. Modeling heavy-tailed stock index returns using the generalized hyperbolic distribution. **Romanian Journal of Economic Forecasting**, v. 6, n. 2, p. 118-131, June 2009.
- NELSON, D. B. Conditional heteroskedasticity in asset return: a new approach. **Econometrica**, v. 59, n. 2, p. 347-370, Mar. 1991.
- ORHAN, M.; KÖKSAL, B. A Comparison of GARCH models for VaR estimation. **Expert Systems with Applications**, v. 39, n. 3, p. 3582-3592, Feb. 2012.
- R Core Team (2016). **R: A language and environment for statistical computing**. R Foundation for Statistical Computing, Vienna, Austria. Available: <<https://www.R-project.org/>>.
- SADORSKY, P. Modeling and forecasting petroleum futures volatility. **Energy Economics**, v. 28, n. 4, p. 467-488, July 2006.
- SUBBOTIN, M. T. On the Law of Frequency of Error. **Matematicheskii Sbornik**, v. 31, n. 2, p. 296-301, 1923.
- THADEWALD, T.; BÜNING, H. Jarque-Bera test and its competitors for testing normality - A power comparison. **Journal of Applied Statistics**, v. 34, n. 1, p. 87-105, 2007.
- THEODOSSIOU, P. Financial Data and the Skewed Generalized t Distribution. **Management Science**, v. 44, n. 12, p. 1650-1661, Dec. 1998.
- VALLE, R. B. A.; AZZALINI, A. The centred parameterization and related quantities of the skew-t distribution. **Journal of Multivariate Analysis**, v. 113, p. 73-90, Jan. 2013.
- WALCK, C. **Hand-book on statistical distributions for experimentalists**. 2. ed. Sweden: University of Stockholm, Stockholm 2007. 190 p.
- WENNSTRÖM, A. **Volatility Forecasting Performance: Evaluation of GARCH type volatility models on Nordic equity indices**. 2014. 61 p. Dissertação (Master of Science) - Department of Mathematics. Royal Institute of Technology (KTH), Stockholm, Sweden, 2014.
- YAHOO FINANCE, 2016. Disponível em: <<http://finance.yahoo.com/>>. Acesso em: 15 Jan. 2016.