



Valor em risco de longo prazo: uma abordagem para modelos da Família Ach e redes neurais

Long term value-at-risk: an Arch Models and neural networks approach

Leandro Santos Maciel -

leandro_maciell@hotmail.com

Graduando em Ciências Econômicas, Departamento de Teoria Econômica
Instituto de Economia da Universidade Estadual de Campinas – UNICAMP/Brasil

Editor Científico Responsável Prof. Dr. Humberto Elias Garcia Lopes
Programa de Pós-Graduação em Administração da Pontifícia Universidade Católica de Minas
Gerais/Brasil

Submissão: 04/08/2009

Aprovação: 29/11/2010

Resumo

Este artigo comparou diferentes estimativas para o Valor em Risco (VaR) de longo prazo no mercado de ações brasileiro. Foram avaliados três modelos: o modelo convencional sugerido pelo *RiskMetrics*, baseado no produto do VaR de curto prazo pela raiz quadrada do *holding period*; o modelo proposto por Dowd (2001); e um modelo proposto neste trabalho, baseado no VaR de curto prazo, em termos de parâmetros previstos. A metodologia de computo do VaR proposta fundamentou-se na previsão da volatilidade e dos retornos das ações por meio de um processo do tipo GARCH (1,1) e por um modelo de Rede Neuronal multicamadas, respectivamente. Os resultados mostraram que o modelo sugerido pelo *RiskMetrics* superestima o valor real do VaR. Além disso, o modelo com base em previsão proposto apresentou superioridade aos demais, visto que suas estimativas mais se aproximaram dos valores observados no mercado.

Palavras-chave: *Value-at-Risk* de longo prazo. Redes Neurais Artificiais. Modelos GARCH.

Abstract

This paper compared different Long Term Value-at-Risk estimative in Brazilian Stock Markets. Three models was tested: the conventional model suggested by RiskMetrics, based on short term VaR multiplied by the square root of holding period; the model proposed by Dowd (2001); and, the model proposed in this work, based on short term VaR in terms of forecasting parameters. The VaR valuated methodology proposed was founded in assets volatility and returns forecasting by GARCH (1,1) process and a multilayer neural network model, respectively. The results showed that the model suggested by RiskMetrics overestimates the real VaR. Nevertheless, the suggested forecasting model performed superiority in comparison to the others methods, since their estimative most approached to the values observed in the stock market.

Keywords: Long term Value-at-Risk. Artificial Neural Networks. GARCH Models.



1 INTRODUÇÃO

Um dos desenvolvimentos mais significativos na área de gestão de riscos foi a emergência da ferramenta do *Value-at-Risk*. Utilizado para a mensuração dos riscos de mercado, o Valor em Risco, ou VaR, surgiu como uma resposta aos desastres financeiros do começo dos anos de 1990. O VaR de uma carteira ou portfólio mensura a perda esperada ao longo de determinado intervalo de tempo, sob condições normais de mercado e dentro de um nível de confiança (JORION, 2003). Trata-se de uma mensuração muito útil para representação dos riscos de mercado, e é amplamente utilizada para os mais diversos instrumentos financeiros¹. Além disso, o VaR pode ser utilizado para avaliação de riscos de fluxos de caixa e eventuais riscos operacionais (ver DANIELSSON et al., 1998; DOWD, 1998, 2000). Entretanto as aplicações dessa metodologia estão associadas principalmente com riscos relativos a horizontes curtos, em que os riscos de longo prazo deixam de ter importância central, por motivos de sua dificuldade de extrapolação.

Como a estimação do VaR inclui as observações históricas dos retornos dos ativos ou carteira em questão, a previsão da perda potencial em prazos mais longos torna-se cada vez mais complexa. O problema maior está centrado na previsão da volatilidade, argumento fundamental para o computo do VaR, que se torna mais difícil para horizontes mais longos. Dowd (2001) propôs uma nova abordagem para a estimação do VaR com posições de longo prazo, em contraposição à fórmula apresentada na emenda sobre risco de mercado do acordo de Basileia, ou seja, o produto do VaR de curto prazo com a raiz quadrada do horizonte temporal². Para o autor, a metodologia convencional é irreal e superestima o valor do VaR. Em seu trabalho, apresenta uma fórmula analítica para o VaR de longo prazo com base nos parâmetros usuais do VaR, porém não requer a estimação dia-a-dia para a volatilidade em longos horizontes de tempo.

Este trabalho oferece uma comparação, para o mercado de ações brasileiro, do modelo do VaR de longo prazo proposto por Dowd com a estimação indicada pelo Comitê da Basileia e, também, com um modelo proposto de previsão dia-a-dia para as variáveis que compõem o

¹ Um bom exemplo dessas aplicações é o sistema de riscos do J. P. Morgan, o *RiskMetrics*TM.

² Especificamente, a emenda sobre risco de Mercado sugere que os bancos devem estimar o VaR para um horizonte de 10 dias úteis, e que essas estimativas devem se basear no VaR diário, ou seja, de curto prazo, multiplicando-o pela raiz quadrada do horizonte temporal, ou *holding period* (Comitê da Basileia, 1996, Seção B-4, parágrafo c, página 44).



VaR, baseado num modelo GARCH (1,1) para previsão da volatilidade e num modelo de Redes Neurais para a previsão do preço do ativo em questão. Foram escolhidas umas das ações mais líquidas da BM&FBOVESPA: as ações preferenciais da Petrobrás, Vale do Rio Doce, Grupo Gerdau, Banco ItaúUnibanco e BM&FBOVESPA. O conjunto de dados é composto pelos preços de fechamento diários das ações no período de 2000 a 2006, utilizados para o cálculo do VaR de longo prazo, ou seja, com *holding period* de 5, 10, 21 e 31 dias úteis. O modelo de previsão dia-a-dia corresponde à abordagem tradicional do VaR em termos de previsão dos parâmetros, i.e., a volatilidade será prevista por um processo do tipo GARCH (1,1) e o valor de mercado da ação será previsto por um modelo de rede neuronal multicamadas do tipo *feedforward*, com base no algoritmo de retropropagação do erro (*backpropagation*). Os valores do VaR de longo prazo foram comparados com os valores do VaR de curto prazo, baseados nos dados observados no mercado, ou seja, as previsões foram comparadas com seus valores reais, com a utilização da métrica do erro quadrático³.

O presente artigo está estruturado como segue. A Seção 2 apresenta o modelo de VaR de longo prazo convencional e modelo proposto por Dowd (2001). O modelo de previsão dia-a-dia é descrito na Seção 3, juntamente com uma síntese de processos de tipo GARCH e o modelo de redes neurais. Segue-se, então, a metodologia empregada. A análise e a discussão dos resultados compõem a Seção 5. Finalmente, a conclusão encerra este trabalho, seguida das notas e das referências bibliográficas.

2 VALOR EM RISCO DE LONGO PRAZO

2.1 Modelo convencional: Acordo de Basileia e *RiskMetrics*

O Value at Risk (Valor em Risco), VaR, sintetiza a maior (ou pior) perda esperada dentro de determinados períodos de tempo e nível de confiança (JORION, 2003). A maior vantagem do cálculo do VaR consiste em resumir, num único número de fácil compreensão, a

³ Este trabalho corresponde a uma abordagem inédita do VaR de longo prazo para o mercado brasileiro e um teste empírico para o modelo de Dowd (2001). Além disso, apresenta uma nova proposta para previsão do VaR em prazos mais longos.

exposição total ao risco de mercado de uma instituição. Trata-se, contudo, de uma medida probabilística e, portanto, nada diz a respeito da magnitude das perdas. Além disso, assume que as relações pretéritas entre as variáveis de interesse continuarão sendo verificadas no futuro (DOWD, 1998).

Para o cálculo do VaR, fixa-se um intervalo de tempo t e um nível de significância estatística de $\alpha\%$ ($0 < \alpha < 1$). O Value-at-risk (VaR_t) de uma carteira de ativo ou portfólio de valor P_t , no período t é definido como:

$$P_t \{ \Delta P_t \leq VaR_t \} = \alpha\% \quad (1)$$

em que P_t representa a variação do valor da carteira ou ativo de preço P_t .

Existem dois métodos para a estimação do VaR: o paramétrico e o não-paramétrico. O método paramétrico consiste na pressuposição de se conhecer a distribuição de probabilidades da variável em estudo. Dessa forma, o valor do VaR pode ser derivado diretamente com a estimação da variância, por isso paramétrica, envolvendo a estimação de parâmetros. Esse método é simples, convincente e produz medidas de VaR mais precisas (JORION, 2003). Dentre os modelos paramétricos tem-se o VaR Delta-Normal. O método não-paramétrico consiste na análise de amostras dos retornos, no qual os dados são usados diretamente como cenário para o cálculo do VaR. Como exemplo de método não-paramétrico tem-se a Simulação Histórica.

Neste artigo, estima-se o VaR por meio do método não-paramétrico. Para isso supõe-se a normalidade para a distribuição dos retornos das séries estudadas, como na maioria dos trabalhos empíricos da área. Assim, o VaR pode ser calculado da seguinte forma:

$$VaR_t = -P_{t-1} \cdot Z_{\alpha\%} \cdot \bar{\sigma}_t \quad (2)$$

em que $Z_{\alpha\%}$ representa o quantil de uma distribuição Normal de probabilidade com um nível de confiança de $\alpha\%$, e $\bar{\sigma}_t$ representa a volatilidade estimada⁴.

⁴ O sinal negativo corresponde ao VaR associado às perdas, uma vez que os riscos estão associados em casos de ocorrências de perdas não previstas.

Se o VaR (h) representa o VaR para um horizonte temporal de h dias (ou seja, para um determinado *holding period*), ele pode ser obtido pela equação em (2) multiplicada pela raiz quadrada de h , e isso significa:

$$VaR(h) = VaR_t \cdot \sqrt{h} \quad (3)$$

Essa abordagem para o VaR de longo prazo é baseada numa sólida fundamentação teórica, sugerida pelo *RiskMetrics* e pelo Comitê da Basileia. Pode ser derivada de um movimento geométrico browniano – em que a variável segue um processo aleatório com distribuição Normal – com média zero; porém, não pode ser generalizada quando não satisfaz essas condições. Com base nessas questões, Dowd (2001) propôs uma nova abordagem para o computo do VaR de longo prazo, que será apresentada na subseção seguinte.

2.2 Modelo de Dowd

Dowd (2001) afirma que a metodologia convencional para o cálculo do VaR de longo prazo, além de ser irreal, superestima o valor do VaR; e, sobretudo, a estimação da volatilidade dia-a-dia se torna cada vez mais difícil com o aumento do período a ser avaliado. Para superar esses problemas, apresentou uma abordagem com base nos parâmetros convencionais do VaR de curto prazo, porém não requer previsão dia-a-dia para a volatilidade e pode gerar estimativas do VaR de longo prazo subjetivo, baseado em hipóteses probabilísticas para cenários futuros. Esse modelo será descrito a seguir.

Suponha um portfólio que gere um retorno logarítmico aleatório diário com média $\mu > 0$ e desvio-padrão (ou volatilidade) σ . Os retornos positivos correspondem a lucros, enquanto os negativos a perdas. Considere, por conveniência, que a composição do portfólio não mude ao longo do período correspondente ao investimento. O nível de confiança do VaR é $\alpha\%$, e considere-o em um horizonte de h dias.

O autor supõe três distribuições alternativas para os retornos logarítmicos diários para avaliar as diferentes assunções de distribuição para os resultados do VaR: distribuição Normal, t de *Student* com 5 graus de liberdade, e distribuição do Valor Extremo (ou

distribuição de Gumbel). Essas distribuições apresentam vantagens e limitações, que estão descritas no Quadro 1.

Distribuições	Vantagens	Limitações
Normal	Conveniente e se torna útil pela propriedade assimétrica	Evidências empíricas mostram que as séries financeiras apresentam leptocurtose (ou caldas pesadas)
t de <i>Student</i>	Permite acomodar a característica das caudas pesadas	Com o aumento dos graus de liberdade converge a uma Normal, o que gera problemas dependendo do horizonte temporal e do grau de leptocurtose
Valor Extremo	Literatura apresenta como a melhor forma para se estimar o VaR com elevado nível de confiança (EMBRECHTS; KLUPPELBERG; MIKOSCH, 1997; LONGIN, 1997).	Ainda apresenta poucas evidências empíricas de sua efetividade.

Quadro 1: Vantagens e limitações das distribuições de probabilidades
Fonte: DOWD, 2001.

Dowd (2001) apresenta uma fórmula diferente para o cálculo do VaR de longo prazo de acordo com cada distribuição de probabilidade. O VaR associado à distribuição Normal para os retornos logarítmicos é:

$$VaR(h) = -(P - P)^* = -\left(P - \exp\left[\mu \cdot h + Z_{\alpha\%} \cdot \sigma \cdot \sqrt{h} + \ln P\right]\right) \quad (4)$$

em que P é o valor corrente do portfólio, P^* corresponde ao $(1 - \alpha)$ percentil (ou percentil crítico) do valor final do portfólio depois de um *holding period* de h dias, e $Z_{\alpha\%}$ é a variável de uma Normal associada ao nível de confiança $\alpha\%$ escolhido (i.e., se $\alpha = 95\%$, $Z_{\alpha\%} = -1,645$).

Com base na distribuição *t-Student*, o VaR se apresenta como:

$$VaR(h) = -(P - P^*) = -\left(P - \exp\left[\mu \cdot h + t_{gl} \cdot \sigma \cdot \sqrt{h} + \ln P\right]\right) \quad (5)$$

em que t_{gl} é uma variável que corresponde à uma variável *t-Student* com um certo nível de confiança, e um certo número de graus de liberdade (gl).

Por fim, o VaR associado à distribuição de valores extremos é:

$$VaR(h) = -(P - P^*) = -(P - \exp[a(h) - b(h) \cdot \log[\log(1/(1 - \alpha\%))]] + \ln P) \quad (6)$$

em que $a(h)$ e $b(h)$ são conhecidos como parâmetros de escala e local (ou meio) da distribuição. $a(h)$ e $b(h)$ podem ser escritos em termos de μ e σ :

$$a(h) = \mu \cdot h - 0,57722 \cdot \sigma \cdot \frac{\sqrt{6 \cdot h}}{\pi} \quad (7)$$

$$b(h) = \sigma \cdot \frac{\sqrt{6 \cdot h}}{\pi} \quad (8)$$

3 Modelo de previsão dia-a-dia: GARCH e Redes Neurais

Para uma comparação mais consistente entre os modelos de VaR de longo prazo, propôs-se neste trabalho um modelo de previsão dia-a-dia, representado por:

$$VaR(h) = -\hat{P}(h) \cdot Z_{\alpha\%} \cdot \hat{\sigma}(h) \quad (9)$$

em que $\hat{P}(h)$ representa o valor previsto da carteira para h passos no futuro, $Z_{\alpha\%}$ o valor correspondente a uma distribuição Normal de probabilidade com um nível de confiança de $\alpha\%$ e $\hat{\sigma}(h)$ a previsão da volatilidade para h passos no futuro.

Diferentemente dos demais modelos, a proposta deste estudo consiste numa abordagem de previsão do VaR para períodos mais longos. Entretanto o que caracteriza a metodologia são os instrumentos para previsão das variáveis fundamentais: Redes Neurais (previsão para o preço) e um processo GARCH (previsão para a volatilidade). Essas metodologias têm apresentado destaque na literatura de finanças, devido a seus resultados satisfatórios e relativa facilidade de implementação⁵, e são apresentadas com mais detalhes a seguir.

⁵ Ver Gately (1996), Zang, Patatuwo e Hu (1998), Wong e Selvi (1998) e Chatterjee et al. (2000) Revista Economia & Gestão - v. 10, n. 24, set/dez. 2010.

3.1 Processos GARCH

As séries financeiras apresentam diversas características peculiares, dentre elas: a heteroscedasticidade, na qual a volatilidade ou variância muda ao longo do tempo; grandes (baixos) valores em certo instante de tempo são seguidos por valores também altos (baixos) nos próximos períodos (agrupamentos de volatilidade); os retornos não são independentes ao longo do tempo; assim como o quadrado dos retornos apresenta autocorrelação estatisticamente significativa. Com isso, essa autocorrelação no quadrado dos retornos implica autocorrelação na variância, pois a esperança do quadrado dos retornos corresponde à variância em diferentes instantes de tempo, decorrente da hipótese de que a média dos retornos assume o valor zero. Ao se determinar essa estrutura de correlação, podem-se prever valores futuros da variância. Entretanto essas previsões serão as condicionais, ou seja, que levam em conta as informações do presente como também as observadas no passado, e, além disso, as não-condicionais que apresentam maior variabilidade.

A variância condicional é um processo autorregressivo (AR), o que resulta em erros condicionalmente heteroscedásticos. Dada a importância das previsões de variância condicional, métodos estatísticos começaram a ser propostos com essa finalidade. O pioneiro nesse campo foi Engle (1982), que, com o objetivo de estimar a variância da inflação, propôs o modelo ARCH (Autoregressive Conditional Heteroscedasticity). Esse modelo é baseado num processo autorregressivo na variância com o objetivo de prever a variância condicional no próximo período. Um processo autorregressivo pode ser definido como um processo estocástico em que uma variável aleatória estabelece uma relação de dependência com seus valores passados e com erros aleatórios. Assim, um modelo ARCH (p) pode ser definido como:

$$\begin{aligned} r_t &= \sqrt{\sigma_t^2} \varepsilon_t, \\ \sigma_t^2 &= \alpha_0 + \alpha_1 r_{t-1}^2 + \dots + \alpha_p r_{t-p}^2 \end{aligned} \quad (10)$$

sendo ε_t uma sequência de variáveis aleatórias independentes e identicamente distribuídas (i.i.d.) com média zero e variância unitária, $\alpha_0 > 0$, $\alpha_i \geq 0$, $i > 0$; σ_t^2 representa a variância e r_t representa o retorno de um ativo financeiro qualquer no instante t , no qual é calculado por uma simples relação com o preço do ativo, P_t , em questão em determinado instante t :

$$r_t = \ln \frac{P_t}{P_{t-1}} = \ln P_t - \ln P_{t-1} \quad (11)$$

As séries financeiras se enquadram a esses modelos, pois apresentam a variância condicional evoluindo com o tempo.

Uma generalização do modelo ARCH foi proposta por (BOLLERSLEV, 1986), denominado GARCH (Generalized Autoregressive Conditional Heteroscedasticity). Trata-se de uma técnica de séries temporais, na qual as variâncias observadas no passado servem para explicar a variância no futuro, ou uma técnica de séries temporais que permite modelar a dependência serial da variância. O modelo GARCH pode ser usado para descrever a volatilidade com menos parâmetros que um modelo ARCH. Esses tipos de modelos, entretanto, são não-lineares no que se refere à variância. Um modelo GARCH (p, q) pode ser representado como:

$$\begin{aligned} r_t &= \sqrt{\sigma_t^2} \varepsilon_t, \\ \sigma_t^2 &= \alpha_0 + \sum_{i=1}^p \alpha_i r_{t-i}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2 \end{aligned} \quad (12)$$

no qual ε_t i.i.d. com média 0 e variância 1.0, $\alpha_0 > 0$, $\alpha_i \geq 0$, $\beta_j \geq 0$ (condições de não-negatividade da variância condicional), e $\sum_{i=1}^m (\alpha_i + \beta_i) < 1$, $m = \max(p, q)$.

A expressão que relaciona σ_t^2 expressa a volatilidade, ou seja, como a variância σ_t^2 se comporta de acordo com o retorno dos períodos passados, r_{t-i}^2 , e a própria variância passada σ_{t-i}^2 (BOLLERSLEV et al., 1992). Assim, no modelo ARCH, a variância do retorno em determinado período depende dos retornos dos períodos passados. No modelo GARCH, essa variância depende dos retornos mas também das variâncias passadas (BOLLERSLEV, 1986; ENGLE, 1982). Assim como nos modelos ARCH, quando se têm volatilidades altas, estas são precedidas de retornos ou volatilidades grandes (MORETTIN; TOLOI, 2004).

Um modelo GARCH (p, q) apresenta a seguinte variância não-condicional:

$$Var(r_t) = \frac{\alpha_0}{1 - \sum_{i=1}^q \alpha_i - \sum_{j=1}^p \beta_j} \quad (13)$$

Se essa variância não for constante, trata-se de um processo não estacionário. Contudo não se pode dizer a mesma coisa da variância condicional, pois a observação de heterocedasticidade condicional não é motivo para não estacionaridade.

Ao se trabalhar com séries financeiras, o modelo usado com maior frequência é o modelo GARCH (1,1), sendo o mais utilizado pelos agentes de mercado (SOUZA, 1999). Esse modelo, comumente usado na prática, pode ser representado como:

$$\begin{aligned} r_t &= \sqrt{\sigma_t^2} \varepsilon_t, \\ \sigma_t^2 &= \alpha_0 + \alpha_1 r_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \end{aligned} \quad (14)$$

com $\alpha_1 \geq 0$, $\beta_1 < 1$, $\alpha_1 + \beta_1 < 1$ (condição para a estacionaridade do modelo).

A utilização de modelos GARCH, contudo, implica algumas ressalvas. Apesar da eficiência do modelo, ao se avaliarem os riscos de ativos negociados no mercado financeiro, os resultados obtidos funcionam apenas como parte para um conjunto de soluções com relação à alocação financeira, dado que as decisões nesse campo são tomadas com base em inúmeros outros fatores, devido à incerteza desses mercados (NELSON, 1991). Além disso, o modelo GARCH apresenta melhor desempenho em condições de mercado com relativa estabilidade. Embora modele a variância condicional, esse modelo não consegue capturar flutuações elevadas nos mercados, como *crashes*, que são capazes de levar a resultados qualitativamente distintos, como mudanças estruturais (ENDERS, 1995; MELLAGI FILHO; ISHIKAWA, 2000). Outra característica das séries financeiras, a qual não é capturada pelos modelos GARCH, são as “caudas pesadas”. Entretanto a heterocedasticia consegue explicar essa ocorrência em partes, por isso distribuições do tipo *t-Student* são frequentemente utilizadas na modelagem GARCH (ANDERS, 2006).

3.2 Redes Neurais Artificiais

Redes Neurais Artificiais (RNAs) podem ser definidas, basicamente, como sendo sistemas de processamento paralelo e distribuído, baseados no sistema nervoso biológico (HAYKIN, 2001). São compostas por elementos computacionais, chamados neurônios. Os neurônios artificiais captam os dados de entrada, ponderam de acordo com determinados pesos sinápticos, que, após passar por uma função de transferência ou ativação, restringem a saída do neurônio (FIG. 1).

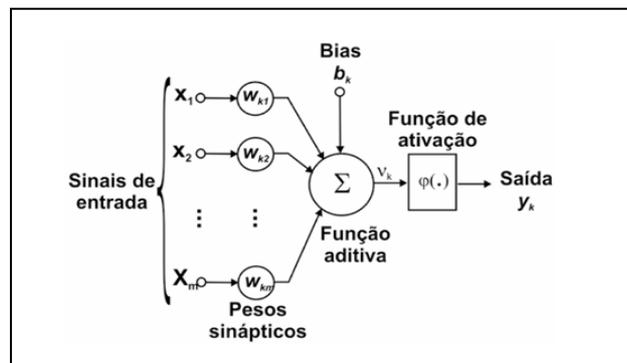


Figura 1: Neurônio Artificial
 Fonte: HAYKIN, 2001

Uma característica fundamental das redes neurais é sua arquitetura ou topologia. A rede *perceptron* multicamadas (MultiLayer Perceptron – MLP) é uma das mais conhecidas e aplicadas arquiteturas de redes neurais, ao qual apresenta uma generalização do *perceptron* proposto por Roseblatt (1958)⁶. A topologia desse modelo consiste em uma camada de entrada, uma ou mais camadas intermediárias e uma camada de saída, como pode ser visto na Figura 2.

O treinamento da topologia MLP é supervisionado, ou seja, é necessário um “professor” para indicar a resposta desejada para o padrão de entrada apresentado à rede durante a fase de aprendizagem. Um sinal de erro é definido como a diferença entre a resposta desejada e a resposta observada. Os parâmetros da rede (pesos e limiares) são ajustados de acordo com esse sinal. O método de aprendizado mais utilizado é o algoritmo de retropropagação do erro. Compõe-se por duas fases. Na primeira fase, conhecida como fase

⁶ As RNAs podem ser do tipo Single-Layer Feed Forward, também conhecido como *perceptrons*, que possui apenas uma camada de nós de entrada ligada diretamente à camada de neurônios de saída, ou Multilayer Feed Forward, que, assim como a arquitetura anterior, possui camadas de entrada e saída e também uma ou mais camadas intermediárias (HAYKIN, 2001).

forward, as entradas são apresentadas e propagadas através da rede, camada por camada, calculando a saída de cada neurônio. Nessa fase, os pesos são fixos e a saída calculada é comparada com a saída desejada, resultando em um erro para cada unidade. Na segunda fase, o erro calculado é propagado da camada de saída para a camada de entrada, fase *backward*, e os pesos são ajustados de acordo com a regra de correção do erro, originando o termo “retropropagação do erro”.

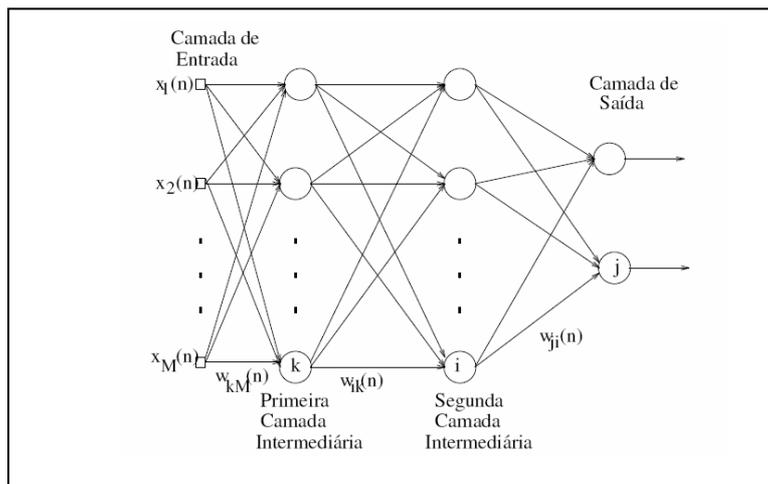


Figura 2 - Rede Neural Perceptron Multicamadas
Fonte: BALLINI, 2000

O algoritmo de aprendizagem de retropropagação (*back-propagation*) consiste em alterar os pesos da rede a fim de minimizar a diferença entre as saídas desejadas e as saídas fornecidas pelo modelo. Se o erro na camada de saída não é menor que uma dada tolerância, então o erro é retropropagado através da rede, tendo como base para a atualização dos pesos a Regra Delta, a qual implementa o Método do Gradiente Descendente (HAYKIN, 2001).

Seja $y_{d_j}(n)$ a saída desejada do neurônio j e $y_j(n)$ a saída calculada pela rede para uma entrada n . O erro $e_j(n)$ na camada de saída é:

$$e_j(n) = y_{d_j}(n) - y_j(n) \tag{15}$$

O valor do erro quadrático para o neurônio j é definido como sendo $\frac{1}{2} \cdot (e_j(n))^2$. A soma dos erros quadráticos é obtida para todos os neurônios da camada de saída, isto é:

$$\varepsilon(n) = \frac{1}{2} \sum_j (e_j(n))^2 \tag{16}$$

Com N representando o número total de pares entrada/saída-desejada, contidos no conjunto de dados de treinamento, o erro quadrático médio é obtido pela soma do erro $\varepsilon(n)$ sobre todo n , normalizado com relação ao número de padrões N :

$$E = \frac{1}{N} \sum_{n=1}^N \varepsilon(n) \quad (17)$$

O processo de treinamento objetiva ajustar os parâmetros livres (pesos) da rede, minimizando uma função objetivo representada pelo erro quadrático médio⁷. Ou seja,

$$\min E = \min \frac{1}{N} \sum_{n=1}^N \varepsilon(n) \quad (18)$$

O nível de ativação interna do neurônio j , $v_j(n)$, é uma função linear das saídas y_i dos neurônios que estão conectados ao neurônio j através dos pesos w_{ji} dado por:

$$v_j(n) = \sum_{i=0}^p w_{ji}(n) y_i(n) \quad (19)$$

em que p é o número de neurônios da camada imediatamente anterior ao neurônio j . O peso sináptico w_{j0} é igual ao limiar θ_0 e corresponde à entrada $y_0 = -1$.

A saída $y_j(n)$ do neurônio j é um valor real, dada por uma função de ativação não-linear:

$$y_j(n) = f_j(v_j(n)) \quad (20)$$

Para minimizar (18) usando o método do gradiente descendente, é necessário calcular a derivada parcial de \mathcal{E} em relação a cada peso da rede⁸.

⁷ Para a derivação do algoritmo de retropropagação, é considerado, inicialmente, um método de treinamento em que os pesos são ajustados entrada a entrada, ou seja, o ajuste dos pesos é realizado de acordo com o erro calculado para cada entrada apresentada à rede.

4. METODOLOGIA

Esta seção descreve a metodologia empregada para a construção dos três modelos avaliados para a mensuração do Valor em Risco de Longo Prazo, desde a composição da amostra de dados à estruturação e estimação dos modelos.

4.1 Dados

Os dados utilizados correspondem aos valores de fechamento diários das ações preferenciais da Petrobrás (PETR4), Vale do Rio Doce (VALE5), Grupo Gerdau (GGBR4), Banco ItaúUnibanco (ITAU4) e BM&FBOVESPA (BMEF3), comercializadas na bolsa de valores brasileira⁹. A amostra se inicia em 3 de janeiro de 2000 e vai até o dia 29 de dezembro de 2006. Como o objetivo é obter o Valor em Risco para prazos de 5, 10, 21 e 31 dias úteis (du), as estimações do VaR avaliaram a máxima perda desses ativos para os dias: 8 de janeiro de 2007 (prazo 5 du), 15 de janeiro de 2007 (prazo 15 du), 30 de janeiro de 2007 (prazo 21 du) e 13 de fevereiro de 2007 (prazo 31 du). Portanto, com os dados históricos até dezembro de 2006, previu-se o VaR desses ativos para o início de 2007.

4.2 Modelo Convencional: *Acordo de Basileia e RiskMetrics*

Para estruturação do modelo de VaR de Longo Prazo, proposto pelo acordo de Basileia e sugerido pelo *RiskMetrics*, construiu-se um processo GARCH (1,1) para a estimação da volatilidade das séries avaliadas. A escolha desse modelo se explica por sua facilidade de implementação, por ser uma estrutura que apresenta bons resultados para séries

⁸ A derivação do modelo do gradiente descendente pode ser vista em Haykin (2001).

⁹ Os dados foram obtidos em <<http://www.bovespa.com.br>>. Acesso em: 26/11/2008.
Revista Economia & Gestão - v. 10, n. 24, set/dez. 2010.

financeiras e por ser amplamente utilizado por agentes de mercado e instituições financeiras (JORION, 2003)¹⁰. Para a estimação do modelo utilizou-se o logaritmo dos retornos das séries (equação 11)¹¹.

Assumiu-se distribuição Normal para a os retornos logarítmicos das séries. Essa assunção, apesar de forte e restritiva, permite uma adequação significativa dos dados, gerando resultados de estimação acurados. Na verdade, não se buscou satisfazer a realidade de todas as hipóteses, mas tentar chegar a um modelo de estimação que mais se aproxime dos valores reais observados no mercado¹².

Portanto, com o valor do logaritmo dos retornos, assunção de Normalidade para as séries e estimação da volatilidade dos ativos por um processo GARCH (1,1), o VaR de longo prazo dos ativos realizou-se por meio da equação (3), ou seja, o VaR para os prazos estipulados se deu com base no VaR dos ativos calculado para dezembro de 2006; assim, esse valor foi extrapolado para os vértices por meio da multiplicação pela raiz quadrada do prazo (*holding period*)¹³. A Tabela 1 apresenta o valor obtido para o VaR nos diferentes prazos estabelecidos.

Tabela 1:
VaR de Longo Prazo: Modelo Convencional (Acordo de Basileia)

Papéis	VaR (5 du)	VaR (10 du)	VaR (21 du)	VaR (31 du)
PETR4	-0,6424	-0,9085	-1,3165	-1,5996
GGBR4	-1,0892	-1,5404	-2,2323	-2,7122
VALE5	-0,7292	-1,0312	-1,4944	-1,8157
ITAU4	-0,8805	-1,2453	-1,8046	-2,1925
BMEF3	-0,9422	-1,3324	-1,9309	-2,3461

Fonte: Resultados da pesquisa.

4.3 Modelo de Dowd

¹⁰ A modelagem foi realizada no *software* Matlab®.

¹¹ A estatística descritiva das séries avaliadas mostrou que elas apresentam excesso de curtose, característico das séries financeiras, pela presença das chamadas *caudas pesadas* (ou *fat tails*). Alguns valores, que por alguns podem ser vistos como *outliers*, foram mantidos na amostra, uma vez que o objetivo da modelagem é prevenir perdas, retirá-los significaria perder grande parte do propósito do trabalho de estimação.

¹² Para avaliação da distribuição das séries aspectos de assimetria e achatamento foram levados em consideração, em que se comparou os percentis da calda esquerda das séries empíricas com os de uma Normal (aspecto de assimetria), e realizou-se o teste de Chebyshev (aspecto de achatamento). Os resultados mostraram que pode ser considerada uma boa *proxy* a distribuição com inovações gaussianas para as séries estudadas.

¹³ Denominam-se vértices os períodos avaliados neste trabalho, 5, 10, 21 e 31 dias úteis, uma vez que se trata de um termo muito utilizado no mercado financeiro para vencimentos de títulos ou negócios.

O modelo de Dowd (2001), utilizado neste trabalho, corresponde ao associado à distribuição Normal dos logaritmos dos retornos (como apresentado na Equação 4), uma vez que o próprio autor apresenta em seu trabalho resultados que comprovem a boa adequação desse modelo. Além disso, como este artigo apresenta uma comparação entre diferentes abordagens, a assunção da mesma hipótese com relação à distribuição de probabilidades das séries avaliadas torna-se necessária para verificação dos resultados.

As variáveis para aplicação do modelo de Dowd foram as mesmas utilizadas no modelo convencional. Ou seja, o preço do ativo é representado pelo retorno logarítmico dos preços, a volatilidade foi obtida por meio de um modelo GARCH (1,1), assumiu-se a hipótese de distribuição Normal para as séries e a média dos retornos foi obtida pela estimativa convencional não-viesada da média¹⁴. A Tabela 2 apresenta os valores do VaR de Longo Prazo pelo modelo de Dowd.

Tabela 2
VaR de Longo Prazo: Modelo de Dowd

Papéis	VaR (5 du)	VaR (10 du)	VaR (21 du)	VaR (31 du)
PETR4	-0,4112	-0,5874	-0,8433	-1,1183
GGBR4	-0,8221	-1,1725	-1,7760	-2,2670
VALE5	-0,4870	-0,6629	-1,0280	-1,3374
ITAU4	-0,5879	-0,9025	-1,3222	-1,6998
BMEF3	-0,6750	-0,9701	-1,4857	-1,9039

Fonte: Resultados da pesquisa.

Ao observar a diferença entre os VaR obtidos pelos modelos até agora apresentados, pode-se perceber que os valores para os diferentes prazos obtidos pelo modelo de Dowd são menores se comparados ao modelo convencional, principalmente para os prazos mais longos, o que pode confirmar a hipótese de superestimação do VaR pelo modelo convencional, conforme Dowd (2001).

4.4 Modelo de Previsão dia-a-dia

¹⁴ Vale ressaltar que o coeficiente de confiança utilizado nos modelos foi de 95%, e, além disso, foi realizada uma análise da correlação e correlação parcial do logaritmo dos retornos para a confirmação da utilização da estrutura (1,1) para o processo GARCH.

Neste modelo, calculou-se o VaR de Longo Prazo em termos de valores previstos, ou seja, aplicou-se a equação (2), porém os parâmetros, o preço e a volatilidade, serão previstos pelo método de Redes Neurais e por um processo GARCH (1,1), respectivamente.

Na modelagem da volatilidade, a metodologia empregada não difere dos demais modelos para a estruturação. Isso significa que foi construído um modelo de tipo GARCH (1,1) com base nos dados históricos que compõem a amostra. Entretanto, em vez de se considerar a volatilidade constante, como nos demais modelos, aqui se fez sua previsão com o número de passos de acordo com o prazo, em dias úteis, avaliado. Ao realizar esse procedimento, trabalhou-se com previsão em cima de previsão, ou seja, previsão com n passos no futuro. Assumem-se as limitações desse tipo de previsão, uma vez que o erro aumenta quando se preveem valores com base em valores já previstos, entretanto buscou-se um modelo que mais se aproximasse da realidade. Com isso, a volatilidade foi prevista para os períodos de 5, 10, 21 e 31 dias úteis, com base num modelo GARCH (1,1)¹⁵.

Para o preço, representado pelo logaritmo dos retornos, utilizou-se um modelo de previsão com base no método de Redes Neurais, técnica de inteligência artificial amplamente utilizada na área de finanças. Na literatura, não existe trabalho que especifique os parâmetros ideais para a construção de uma modelagem de Rede Neural que leve aos melhores resultados; portanto, envolve escolhas empíricas que variam de acordo com a especificidade dos dados e o objetivo de previsão (KAASTRA; BOYD, 1995). Dessa forma, a estrutura apresentada neste trabalho corresponde a um resultado obtido por diversas tentativas de alteração dos valores que estruturam a rede em busca aos melhores resultados.

Os dados de entrada (*input data*) utilizados para se prever o logaritmo dos retornos no futuro das séries foram os últimos 35 retornos logarítmicos observados no passado. Ou seja, estruturou-se um modelo de Rede Neural de tipo multicamadas, com o objetivo de se prever o logaritmo dos retornos das ações no futuro e, para isso, os dados apresentados à rede foram os retornos logarítmicos observados nos últimos 35 dias úteis. Os dados foram divididos em três amostras distintas: treinamento, validação e teste. Os conjuntos treinamento e validação compõem os dados históricos que formam a amostra deste trabalho, em que cada um contém 90% e 10% do total dos dados, respectivamente. O conjunto teste contém os valores observados no mês de janeiro e fevereiro de 2007, para que se possa avaliar o quanto a rede foi capaz de se aproximar aos valores observados no mercado. Portanto, a estrutura da rede

¹⁵ Também neste caso, assumiu-se distribuição Normal para o logaritmo dos retornos das séries em questão.
Revista Economia & Gestão - v. 10, n. 24, set/dez. 2010.

corresponde a um modelo do tipo *Multilayer Perceptrons*, com 35 neurônios na camada de entrada, uma camada intermediária composta por 15 neurônios e uma camada de saída, que representa o valor previsto para o logaritmo do retorno do papel avaliado.

O treinamento do modelo de Rede Neural foi realizado com o algoritmo *backpropagation*, com um número de épocas estipulado em 1 000¹⁶. Este número alto de épocas foi escolhido propositalmente para obrigar a rede a convergir quando atingisse o erro mínimo desejável e não o número máximo de épocas. Valores do erro do treinamento usado na simulação foram da ordem de (10^{-3}) ¹⁷. Dessa forma, com os valores previstos para os retornos e as volatilidades, apurou-se o valor do VaR para os prazos especificados, conforme a Tabela 3 apresenta.

Tabela 3
VaR de Longo Prazo: Modelo de Previsão dia-a-dia (GARCH e RNA)

Papéis	VaR (5 du)	VaR (10 du)	VaR (21 du)	VaR (31 du)
PETR4	-0,3505	-0,5029	-0,8465	-1,0397
GGBR4	-0,8042	-1,1334	-1,7542	-2,1043
VALE5	-0,4281	-0,6658	-0,9054	-1,2121
ITAU4	-0,5459	-0,8153	-1,3077	-1,8016
BMEF3	-0,6793	-0,9440	-1,4660	-1,7647

Fonte: Resultados da pesquisa

Com o valor do VaR de longo prazo obtido pelas metodologias avaliadas é possível realizar a comparação entre os modelos com os valores de mercado, como apresenta a seção seguinte.

5 Análise e discussão dos resultados

Na seção anterior apresentou-se a metodologia empregada para a estimação do VaR de longo prazo com base nos três modelos avaliados. Entretanto, para comparar qual foi o modelo que melhor estimou a máxima perda dos ativos para os prazos selecionados, fez-se uma comparação de seus valores com os observados no mercado de ações brasileiro. Ou seja,

¹⁶ Uma época é definida como toda apresentação de um conjunto entrada-saída para o treinamento da rede, na busca de minimizar a diferença entre o valor desejado e o valor estipulado pela rede.

¹⁷ O valor adotado da taxa de aprendizagem foi 0,04. Este é apresentado como *default* pelo algoritmo e apresentou os melhores resultados

calculou-se o VaR de curto prazo para as datas estudadas com base no modelo simples, como na Equação em 2, utilizando os valores para retorno e volatilidade observados no mercado; assim, verificou-se qual modelo se aproximou mais da realidade.

Com isso, o valor do VaR para as datas 8 de janeiro de 2007 (prazo 5 du), 15 de janeiro de 2007 (prazo 15 du), 30 de janeiro de 2007 (prazo 21 du) e 13 de fevereiro de 2007 (prazo 31 du) foi obtido com a utilização do retorno observado em cada dia e com a volatilidade estimada para cada data específica, i. e., a máxima perda esperada em cada dia, de acordo com os dados observados (reais). A Tabela 4 apresenta o valor do VaR para os vértices avaliados.

Tabela 4
VaR de Curto Prazo para os Holding Periods avaliados

Papéis	VaR (8 jan 2007)	VaR (15 jan 2007)	VaR (30 jan 2007)	VaR (13 fev 2007)
PETR4	-0,2291	-0,3356	-0,5492	-0,7464
GGBR4	-0,6675	-0,9582	-1,4810	-1,8368
VALE5	-0,2460	-0,4686	-0,7181	-0,9394
ITAU4	-0,4243	-0,6731	-1,0314	-1,5254
BMEF3	-0,4969	-0,7672	-1,2237	-1,5195

Fonte: Resultados da pesquisa

O próximo passo é avaliar, dentre os modelos empregados, o que mais se aproximou do valor do VaR de curto prazo. Para isso, utilizou-se a métrica do erro quadrático, obtido por:

$$EQ = (VaR.CP - VaR.LP)^2 \quad (21)$$

em que $VaR.CP$ representa o VaR de curto prazo para as datas estimadas (valor real) e $VaR.LP$ o VaR de longo prazo estimado pelos modelos (valor estimado).

A Tabela 5 apresenta o valor do erro quadrático, em termos percentuais, para os modelos de VaR de Longo Prazo utilizados neste trabalho em todos os *holding periods* escolhidos. Ela nos permite inferir uma série de conclusões. Em primeiro lugar, pode-se confirmar a crítica de Dowd (2001) ao modelo convencional de estimação do VaR de longo prazo, como proposto no documento do Acordo da Basileia e pelo *RiskMetrics*, uma vez que é notória a superestimação do valor do VaR, principalmente para os vértices mais longos (21 e 31 dias úteis, como na TAB. 5). A superioridade do modelo de Dowd (2001) sobre o modelo convencional foi confirmada pela análise empírica deste trabalho. O modelo proposto pelo Revista Economia & Gestão - v. 10, n. 24, set/dez. 2010.

autor confere uma estimaco mais precisa, e, alm disso, principalmente para os prazos mais longos, a superioridade  ainda mais clara¹⁸. Entretanto, vale acrescentar que o erro aumenta com a elevaco do prazo, o que  coerente com qualquer modelo que visa prever valores no futuro, visto que a volatilidade aumenta em decorrncia das incertezas relativas ao futuro.

Tabela 5
Erro quadrtico dos modelos de VaR de longo prazo

Erro Quadrtico						
VaR (5 dias teis)				VaR (10 dias teis)		
Papis	Basileia	Dowd	Previso	Basileia	Dowd	Previso
PETR4	17,8%	3,31%	1,47%	32,81%	6,34%	2,80%
GGBR4	17,79%	2,39%	1,87%	33,89%	4,59%	3,07%
VALE5	23,35%	5,81%	3,32%	31,66%	3,78%	3,89%
ITAU4	20,81%	2,68%	1,48%	32,74%	5,26%	2,02%
BMEF3	19,83%	3,17%	3,33%	31,95%	4,12%	3,12%
VaR (21 dias teis)				VaR (31 dias teis)		
Papis	Basileia	Dowd	Previso	Basileia	Dowd	Previso
PETR4	58,88%	8,65%	8,84%	72,80%	13,83%	8,60%
GGBR4	56,44%	8,70%	7,47%	76,62%	18,59%	7,15%
VALE5	60,26%	9,60%	3,51%	76,79%	15,84%	7,44%
ITAU4	59,79%	8,46%	7,63%	44,51%	3,04%	7,63%
BMEF3	50,02%	8,87%	5,87%	68,32%	14,78%	6,01%

Fonte: Resultados da pesquisa

O modelo de previso proposto neste trabalho apresentou os melhores resultados, em geral. Vale a pena acrescentar, ainda, que seus resultados se aproximaram muito do modelo de Dowd, porm a maioria dos resultados do modelo de previso dia-a-dia foram superiores. Esses resultados se explicam, em grande parte, pela capacidade elevada de previso dos retornos por meio das Redes Neurais. O modelo de rede proposto conseguiu prever acuradamente os retornos das aoes, implicando melhor estimaco do VaR de longo prazo. Mesmo com um desvio de acurcia por meio do modelo GARCH, por se tratar de um modelo tradicional de sries temporais, sua capacidade de errar nos prazos mais longos pode ser compensada pelos bons resultados obtidos da Rede Neural. Alm disso, o bom desempenho dos modelos, modelo de Dowd e de previso dia-a-dia, pode ser confirmado mesmo com a assuno de distribuio Normal para o logaritmo dos retornos das sries em questo¹⁹.

¹⁸ Dowd (2001) afirma que seu modelo  mais eficaz para a estimaco do VaR de acordo com o aumento do prazo. Essa proposio foi confirmada por este teste emprico.

¹⁹ Vale ressaltar, ainda, que a no utilizao do teste de Kupiec se d pelo fato de se estimar o VaR para perodos especficos no futuro e no em se obterem estimativas dirias, o que resultaria em uma amostra significativa para avaliao por meio das taxas de exceoes.

6 CONCLUSÕES

O presente artigo avaliou três modelos para computo do *Value-at-Risk* de longo prazo para as ações preferenciais da Petrobrás, Vale do Rio Doce, Grupo Gerdau, Banco ItaúUnibanco e BM&FBOVESPA. O conceito de VaR de longo prazo se refere à estimação da máxima perda, com determinado nível de confiança e em situações normais de mercado, para as ações em intervalos de tempo que podem ser considerados de longo prazo neste mercado: 5, 10, 21 e 31 dias úteis. Os três modelos utilizados foram: modelo convencional, adotado pelo *RiskMetrics* e sugerido pelo Acordo de Basileia, com base no cálculo do VaR de curto prazo multiplicado pela raiz quadrada do horizonte temporal (ou *holding period*); o modelo proposto por Dowd (2001), embasado numa fórmula obtida pela manipulação algébrica do modelo convencional do VaR de curto prazo; e um modelo proposto neste trabalho, que calcula os parâmetros do VaR em termos de valores previstos, com a utilização do processo GARCH para a previsão da volatilidade, e do modelo de Redes Neurais Artificiais para a previsão do logaritmo dos retornos das ações.

Para realizar essa análise, a amostra dos dados continham os preços de fechamentos das ações avaliadas no período de 2000 a 2006, para a previsão do VaR dessas ações no mês de janeiro e fevereiro de 2007. A volatilidade dos modelos foi estimada com um processo de tipo GARCH (1,1), supondo-se que os retornos logarítmicos seguem uma distribuição com inovações gaussianas, i.e., distribuição Normal. Para o modelo de previsão dia-a-dia proposto, foi construído um modelo de rede neural multicamadas, com base no algoritmo de retropropagação do erro, para prever o valor dos retornos logarítmicos das ações para os períodos estipulados. Finalmente, os resultados foram comparados com o valor do VaR de curto prazo auferidos com dados reais dos mercados nas datas sugeridas, na tentativa de conhecer o modelo que mais se aproximou da realidade.

Os resultados confirmaram a proposição de Dowd (2001), em que o modelo convencional do computo do VaR de longo prazo utilizado pelo *RiskMetrics* superestima o valor verdadeiro do VaR, e, para todos os vértices avaliados, apresentou os piores resultados. Entretanto o modelo que apresentou estimativas mais próximas da realidade foi o modelo de previsão dia-a-dia, proposto neste trabalho. As estimativas obtidas foram superiores, em geral,



se comparadas às obtidas pelo modelo de Dowd, além disso, foram significativamente acuradas.

Trabalhos futuros podem se concentrar na tentativa de se obter o processo gerador da volatilidade por um modelo de rede neuronal, visto que essa técnica tem apresentado melhores resultados em comparação com modelos tradicionais de séries temporais e não requer tanto esforço computacional em vista de suas boas estimativas. Além disso, estudos podem se concentrar na tentativa de utilizar outros algoritmos de aprendizagem ou estrutura de rede neural na tentativa de incrementar os resultados.

Referências

ANDERS, W. Garch forecasting performance under different distribution assumptions. **Journal of Forecasting**, v. 25, n. 5, p. 561-578, 2006.

BALLINI, R. **Análise e Previsão de Vazões Utilizando Modelos de Séries Temporais, Redes Neurais e Redes Neurais Nebulosas**. 2000.169 f. Tese (Doutorado em Engenharia Elétrica) – Universidade Estadual de Campinas, Faculdade de Engenharia Elétrica e Computação, Campinas.

BASILÉIA (BASLE), Committee on Banking Supervision. **Amendment to the Capital Accord to Incorporate Market Risk**. Bank for International Settlements, January, 1996.

BOLLERSLEV, T. R. Generalized Autoregressive Conditional Heteroskedasticity. **Journal of Econometrics**, v. 31, n.51, p. 307-327, 1986.

BOLLERSLEV, T. R.; CHOU, R. Y.; KRONER, K. F. ARCH Modeling in Finance: A Review of the Theory and Empirical Evidence. **Journal of Econometrics**, n.52, p.5-59, 1992.

CHATTERJEE, A.; AYADI, O. F.; BOONE, B. E. Artificial neural networks and the financial markets: A survey. **Managerial Finance**, v. 26, n. 4, p.32-45, 2000.

DANIELSSON, J.; HARTMANN, P; DE VRIES, C. G. The cost of the conservatism: Extreme Returns, Value-at-Risk, and the Basle ‘Multiplication Factor’. **Risk**, v. 11, n. 1, p. 101-120, 1998.

DOWD, K. **Beyond Value at Risk: The New Science of Risk Management**. Nova Iorque: Wiley & Sons, 1998. 234p.

DOWD, K. Estimating Value at Risk: a subjective approach. **Journal of Risk Finance**, v. 1, n. 4, p. 43-56, 2000.

DOWD, K. **Long-Term Value at Risk**. The Pensions Institute, Birkbeck College, University of London Discussion Paper, June, 2001.



EMBRECHTS, P.; KLUPPELBERG, C.; MIKOSCH, T. **Modelling Exponential Events for Insurance and Finance**. Berlin: Springer, 1997. 648p.

ENDERS, W. **Applied Econometric Time Series**. 2. ed. Nova Iorque: Wiley & Sons, 1995, 149p.

ENGLE, R. F. Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingston Inflation. **Econometrica**, v. 50, n. 50, p. 987-1 008, 1982.

GATELY, E. J. **Neural Networks for Financial Forecasting**. Nova Iorque: John Wiley & Sons, 1996. 242p.

HAYKIN, S. **Neural Networks: A Comprehensive Foundation**. Nova Iorque: IEEE Press, 2001. 657p.

JORION, P. **Value at Risk: A nova fonte de referência para a gestão do risco financeiro**. São Paulo: Bolsa de Mercadorias e Futuros, 2003. 488p.

KAASTRA, I; Boyd, M. Designing a neural network for forecasting financial and economic time series. **Neurocomputing**, v. 10, n. 3, p. 215-236, 1995.

LONGIN, F. M. Beyond the VaR. **The Journal of Derivatives**, v. 8, n. 4, p. 36-48, 2001.

MELLLAGI FILHO, A.; ISHIKAWA, S. **Mercado Financeiro e de Capitais**. São Paulo: Atlas, 2000. 199p.

MORETTIN, P. A.; TOLOI, C. M. C. **Análise de Séries Temporais**. São Paulo: Edgar Blücher, 2004. 567p.

NELSON, D. B. Conditional Heteroskedasticity in Asset Pricing: A New Approach. **Econometrica**, v. 59, n. 2, p. 347-370, 1991.

ROSEMBLATT, F. The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain. **Psychological Review**, v. 65, n. 6, p. 386-408, 1958.

SOUZA, L. A. R. **Valor em Risco em Épocas de Crise**. 1999. 121 f. Dissertação (Mestrado em Economia) – Universidade de São Paulo, Departamento de Economia da Faculdade de Economia, Administração e Contabilidade, São Paulo.

WONG, B. K.; SELVI, Y. Neural network applications in business: A review and analysis of the literature. **Information and Management**, v. 34, n. 2, p. 129-139, 1998.

ZANG, G.; PATATUWO, B. E.; HU, M. Y. Forecasting with artificial neural networks: The state of the art. **International Journal of Forecasting**, v. 14, n. 1, p. 35-62, 1998.