

COMBINAÇÃO DA PROJEÇÃO DA VOLATILIDADE PERCEBIDA POR REDES NEURAIIS E HAR

ALCIDES CARLOS DE ARAUJO

FACULDADE DE ECONOMIA, ADMINISTRAÇÃO E CONTABILIDADE DA UNIVERSIDADE DE SÃO PAULO -
FEA

alcides.carlos@yahoo.com.br

ALESSANDRA DE AVILA MONTINI

CONRE - CONSELHO REGIONAL DE ESTATÍSTICA

amontini@usp.br

Introdução

A volatilidade possui impacto em diversas aplicações financeiras como gestão de carteiras, alocação de recursos, apreçamento de derivativos e gestão de riscos. A volatilidade é uma variável latente; diante disto, diversos métodos foram propostos na literatura para mensurá-la. Uma das formas de mensuração que recebeu relevância nos últimos anos foi a medida de Volatilidade Percebida (Realized Volatility - RV), métrica que utiliza as variações intradiárias. Conforme Zivot (2005) e Andersen et al. (2007) projetar a volatilidade utilizando esta métrica apresentou melhor desempenho de projeção.

Problema de Pesquisa e Objetivo

Um problema da área financeira é a análise de bancos de dados de negociações realizadas em tempo real e em altíssima frequência (High Frequency Trading - HFT). Um grande desafio do HFT é a gestão dos riscos, dado que o algoritmo pode apresentar retornos excelentes em um dia e revertê-los em perdas em um único momento. Deste modo, projetar a volatilidade de forma eficiente é fundamental para um melhor controle dos riscos. Diante disto, o artigo propõe a combinação dos métodos Heterogeneous Autoregressive Model of Realized Volatility - HAR e Redes Neurais para melhor projetar a volatilidade.

Fundamentação Teórica

O modelo HAR foi sugerido em Corsi (2003). Este método de projeção é uma adaptação dos modelos HARARCH propostos em Müller et al. (1997). Este método recebeu grande relevância nos últimos anos. Kuan e White (1994) apresentam a técnica de Redes Neurais como um modelo de entrada-saída de massivo paralelismo, dependência não linear e processamento de informações em multicamadas. Conforme Timmermann (2006) combinar as projeções possibilita a utilização de vários modelos em busca de minimizar os erros de previsão. A combinação de projeções demonstrou sucesso empírico em diversas aplicações econômicas

Metodologia

A pesquisa possuiu 4 etapas: obtenção da série de volatilidade percebida (RV), estimação dos modelos HAR e Redes Neurais, combinação das projeções e avaliação de performance. Para realizar as análises foi utilizada a série de Volatilidade Percebida do índice Bovespa no período entre 03/01/2000 a 01/10/2014, total de 2692 observações. Para estimar os modelos foi utilizada 70% da amostra total e os demais 30% para testes. Para realizar as combinações foram utilizados os métodos Pesos Iguais e Mean Squared Forecast Error. Os desempenhos foram comparados por meio das medidas RMSE e MAE.

Análise dos Resultados

Na comparação entre ambos os modelos dentro do período de estimação, observa-se não existir diferenças na capacidade de previsão dos métodos. Sendo que os modelos de Redes Neurais apresentaram os resultados levemente melhores quanto a minimização do erro dentro da amostra de treino. Os desempenhos observados na amostra de teste ocorreram conforme esperado pela teoria. Ou seja, combinar as projeções possibilitou minimizar erros de previsão melhor de que a escolha de um modelo individual. Os dois métodos de combinação utilizados apresentaram os menores valores de erros.

Conclusão

A combinação de ambas as técnicas para projeção da Volatilidade Percebida é inovativo, dado que duas áreas de conhecimento diferentes são utilizadas para melhorar o desempenho de projeções. Nos resultados fora da amostra, observou-se que a combinação de ambos os modelos resultaram numa melhor previsibilidade da Volatilidade Percebida. Consequentemente, mais eficiência para gestão de riscos. Em trabalhos futuros, mais modelos e mais métodos de combinação de projeções devem ser estudados. As análises também deverão utilizar uma quantidade maior de ativos e diferentes períodos de previsão.

Referências Bibliográficas

- CORSI, F. A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics*, v. 7, n. 2, p. 174-196, 2009.
- KUAN, C.; WHITE, H. Artificial neural networks: an econometric perspective. *Econometric Reviews*, v. 13, n. 1, p. 1-91, 1994.
- FERNANDES, M.; MEDEIROS, M. C.; SCHARTH, M. Modeling and predicting the CBOE market volatility index. *Journal of Banking & Finance*, v. 40, p. 1-10, 2014.
- TIMMERMANN, A. Forecast combinations. *Handbook of economic forecasting*, v. 1, p. 135-196, 2006.

COMBINAÇÃO DA PROJEÇÃO DA VOLATILIDADE PERCEBIDA POR REDES NEURAI E HAR

INTRODUÇÃO

A volatilidade possui impacto em diversas aplicações financeiras. As aplicações vão desde gestão de carteiras, alocação de recursos, apreçamento de derivativos, gestão de riscos até tomadas de decisão das instituições a respeito da estabilidade financeira, conforme apresenta Danielsson (2011).

Em essência, a volatilidade é uma variável latente; diante disto, diversos métodos foram propostos na literatura para mensurá-la. Uma das formas de mensuração que recebeu relevância nos últimos anos foi a medida de Volatilidade Percebida/Volatilidade Realizada (*Realized Volatility* - RV), proposta por Andersen e Bollerslev (1998).

De acordo com Danielsson (2011), esta medida utiliza dados passados com frequência intradiária; deste modo, torna-se uma medida puramente direcionada pelos dados. Conforme apresentam Zivot (2005) e Andersen et al. (2007), modelos de projeção que utilizam a métrica da Volatilidade Percebida demonstraram melhor desempenho de projeção.

Um problema da área financeira é a análise de grandes bancos de dados de negociações na bolsa de valores. Estas negociações são realizadas em tempo real e, para diversos ativos, em altíssima frequência.

Uma das atividades de finanças que fomenta e prioriza a análise de dados em alta frequência é o HFT (*High Frequency Trading*), tipo de negociação associada a rápida realocação de capital, em que as transações podem ocorrer em milésimos de segundos.

Um grande desafio do HFT é a gestão dos riscos, dado que o algoritmo pode apresentar retornos excelentes em um dia e em seguida revertê-los em perdas em um único momento. Deste modo, projetar a volatilidade de forma eficiente é fundamental para um melhor controle dos riscos.

No presente artigo, dois métodos de projeção da Volatilidade Percebida foram estudados. O modelo denominado *Heterogeneous Autoregressive Model of Realized Volatility* (HAR), sugerido em Corsi (2003) e o método de Redes Neurais.

Combinar as projeções dos modelos HAR e Rede Neural possui justificativa interessante, dado que o primeiro modelo é baseado em pressupostos e suposições da teoria econômica, enquanto que o segundo está baseado em modelos determinísticos não lineares advindos da área de aprendizagem de máquina. Diante disto, combinar o conhecimento de duas áreas importantes da ciência em finanças pode melhorar o desempenho das projeções.

Estes dois modelos são exemplos da amplitude de métodos de projeção desenvolvidos para dados em alta frequência, porém por meio desta amplitude surge um novo problema. Torna-se complicado para um gestor de algoritmos de negociação selecionar uma medida de risco para controlar as perdas de suas operações ou mensurar o quanto poderá perder caso ocorra o pior cenário.

Diante disto, como objetivo geral o artigo propõe a combinação dos métodos HAR e Redes Neurais para melhor projetar a Volatilidade Percebida e, conseqüentemente, trazer maior eficiência na gestão de riscos de operações HFT.

Os estudos a respeito de combinação de projeções possuem significativa aplicação prática. Conforme Timmermann (2006) a combinação de projeções demonstrou sucesso empírico em diversas aplicações econômicas, como projeção do PIB, preço do petróleo e retorno de ações.

2 REVISÃO DA LITERATURA

2.1 Medidas de volatilidade com dados em alta frequência

Um tema de pesquisa relativamente recente, conforme Zivot (2005) é o estudo das medidas de volatilidade para dados em alta frequência. Esta área de estudos possui importância dado que produz estimadores que apresentam melhor exatidão para a volatilidade e correlação condicional quando comparados com estimadores que utilizam dados diários ou mensais.

Os estimadores de volatilidade para dados em alta frequência são derivados a partir do conceito da Variância Integrada (*Integrated Variance – IV*), conforme descrevem Andersen et al. (2001), Barndorff-Nielsen e Shephard (2004) e Bauwens (2012). Conforme os autores, dado a velocidade e liquidez nos mercados financeiros, as operações tendem a ocorrer de forma muito próxima do tempo contínuo.

Seja Y_t um vetor de preços $N \times 1$, em que $Y_t = \ln p_t$, sendo p_t o preço negociado, a difusão dos preços é expresso na equação 1,

$$dY_t = \mu_t dt + \sigma_t dW_t, t \geq 0, \quad (1)$$

em que, dY_t é o incremento do preço em logaritmo, μ_t é a direção (*drift*) do processo contínuo localmente delimitado, σ_t é a volatilidade instantânea estritamente positiva e estacionária e W_t é o Movimento Browniano Padrão.

Pode-se assumir que t é um dia de operações, sendo $i-s$ o espaçamento entre duas observações, com $s = 1$ e $i = 1 \dots M$ períodos intradiários, o retorno $r_{t,i}$ é expresso na equação 2.

$$r_{t,i} \equiv Y_{t,i} - Y_{t,i-1} = \int_{t,i-1}^{t,M} \mu_{t,s} ds + \int_{t,i-1}^{t,M} \sigma_{t,s} dW_{t,s}, i = 1 \dots M. \quad (2)$$

Por meio da expressão 2, a distribuição de probabilidade dos retornos é definida na expressão 3,

$$r_{t,i} \sim N \left(\int_{t,i-1}^{t,M} \mu_{t,s} ds, IV_{t,M} \right), \quad (3)$$

em que, $IV_{t,s}$ é dada pela expressão 4,

$$IV_{t,s} \equiv \int_{t,i-1}^{t,M} \sigma_{t,s}^2 ds, \quad (4)$$

em que, IV é a Variância Integrada (*Integrated Variance* – IV).

Como se observa, os valores da variável IV precisam ser estimados, uma vez que é uma variável latente. No caso, IV pode ser estimada utilizando dados intradiários, Andersen e Bollerslev (1998) propõe o estimador da Variância Percebida (*Realized Variance* - RV), expressão 5,

$$RV_t = \sum_{i=1}^M r_{t,i} r'_{t,i}. \quad (5)$$

De acordo com Andersen *et al.* (2003), o estimador foi apresentado com o objetivo de integrar dados de negociação em alta frequência e medidas de estimação de volatilidades. Conforme os autores, o estimador é fácil de ser implementado e melhora o desempenho dos modelos preditivos.

Por meio da estimação do RV_t para cada dia, obtém-se a série de Volatilidade Percebida. Conforme apresenta Corsi (2009), esta série é altamente assimétrica a direita e leptocúrtica, consequentemente, desvia-se da normalidade. Porém, este comportamento é padrão para as séries de Volatilidades, sendo um resultado observado desde Bollerslev et al. (1994).

2.2 Projeção com Modelo HAR

O modelo HAR (*Heterogeneous Autoregressive Model of Realized Volatility*) foi sugerido em Corsi (2003), sendo apresentada uma versão revisada em Corsi (2009). Este método de projeção é uma adaptação dos modelos *Heterogeneous Autorregressive Conditional Heteroskedasticity* (HARCH) propostos em Müller et al. (1997) e Dacorogna et al. (1998).

Uma vez que séries de Volatilidades Percebidas são altamente assimétricas, tendo comportamento que se desvia da normalidade, Corsi (2009) e Fernandez et al. (2014) apresentam que o modelo HAR é adequado a tais características. Dado que o modelo possibilita a observação do impacto da volatilidade tanto no curto quanto no longo prazo.

Para estimação do modelo, Boudt et al. (2013) descreve primeiramente sobre a necessidade de obter as volatilidades percebidas (RV), uma medida foi apresentada na expressão 5. Durante esta etapa estima-se o RV como agregação de vários dias por meio da expressão 6,

$$RV_t^{(\cdot)} = \frac{1}{h} \left[\sum_{i=1}^h RV_{t-i} \right], \quad (6)$$

em que, h é o número de dias que se busca obter o RV agregado.

Na expressão 7 é apresentado o modelo HAR para k parâmetros,

$$RV_{t+h} = \sum_{i=0}^k \beta^{(i)} RV_t^{(i)} + \epsilon_{t+h}, \quad (7)$$

em que, os parâmetros $\beta^{(i)}$ são os coeficientes para os horizontes de tempo selecionados para a estimação. Normalmente as escolhas para “ i ” são 1, 5 e 22 dias, conforme é sugerido em Corsi (2009).

2.3 Projeção com Modelo de Redes Neurais

Devido ao comportamento complexo de uma série de volatilidade, diversas pesquisas estão empregando a utilização da técnica de Redes Neurais para estimá-la e projetá-la. Alguns trabalhos que apresentaram resultados significativos foram os de Donaldson e Kamstra (1997), Hu e Tsoukalas (1999) e Hamid e Iqbal (2004). Para o caso específico da Volatilidade Percebida, podem ser citados os artigos de Hillebrand e Medeiros (2010) e Fernandez et al. (2014).

Conforme apresentam Oliveira et al. (2011), uma Rede Neural é uma técnica de análise de dados que mapeia algum tipo de entrada de informação, processa esta informação em diversas camadas e sugere uma determinada saída após um processo de aprendizagem. As principais vantagens para utilização desta técnica são generalização, flexibilidade e modelagem não linear. Na Figura 1 pode ser visualizado o funcionamento da técnica.

Nas variáveis X1 a X5 é observada a camada de entrada (variáveis preditoras); H1 a H3 são denominados neurônios, estes compõem a camada oculta e Y é a camada de saída (variável de previsão). Os pesos β_i e γ_i são pesos utilizados para ajustar os valores observados na camada de entrada de forma a melhor prever os resultados observados na camada de saída.

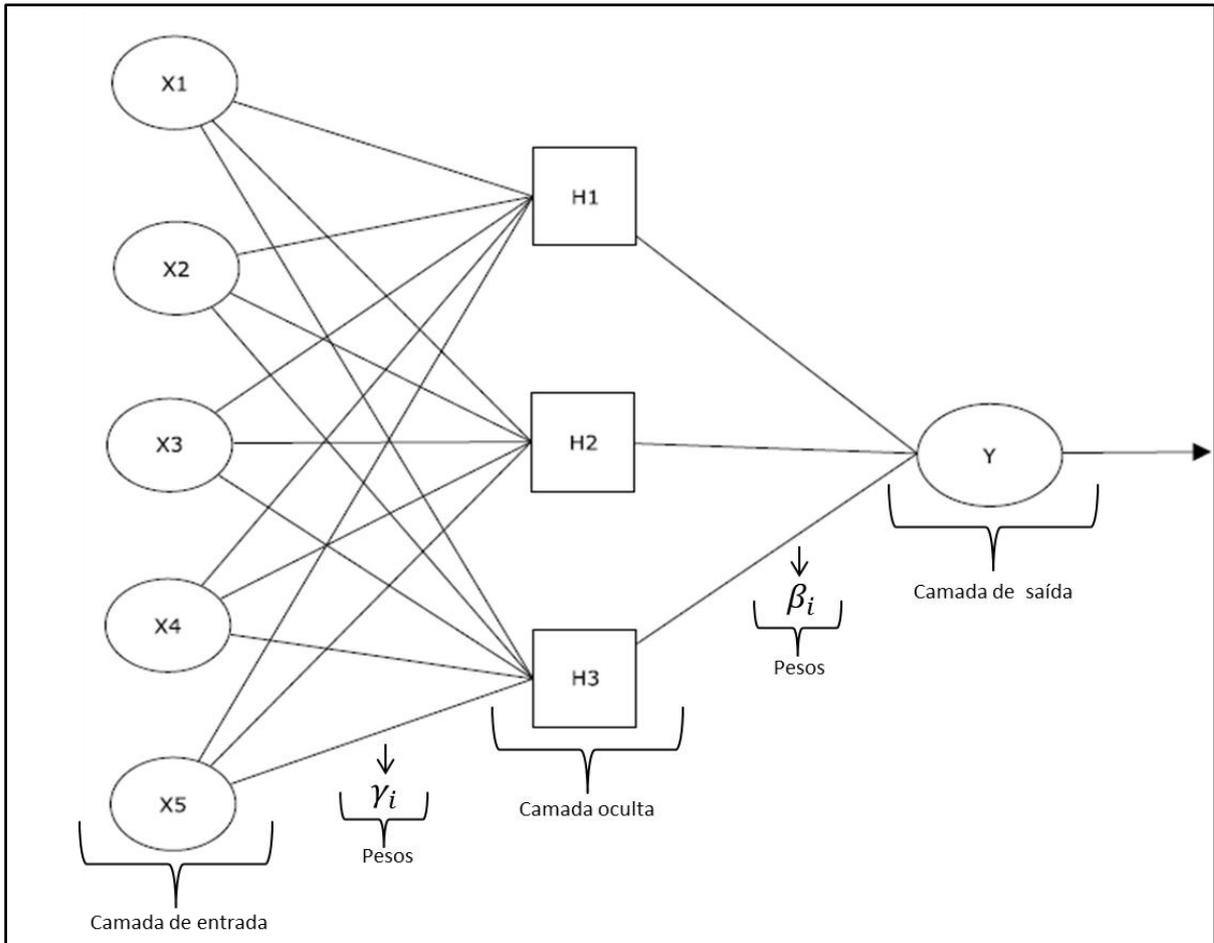


Figura 1 – Modelo de Redes Neurais

Fonte: elaborado pelos autores.

Do ponto de vista econométrico, Kuan e White (1994) apresentam a técnica de Redes Neurais como um modelo de entrada-saída, tendo características como: massivo paralelismo, capacidade de capturar dependência não linear entre variáveis e processamento de informações em multicamadas. Conforme citam Fernandez et al (2014), estas características explicam o sucesso na utilização da técnica para modelagem da Volatilidade.

No trabalho de Hillebrand e Medeiros (2010), os autores adaptam a técnica de Redes Neurais como uma extensão não linear do Modelo HAR, conforme é apresentado na expressão 8,

$$RV_{t+h} = T(\mathbf{RV}_t) + \epsilon_{t+h}, \quad (8)$$

em que, $T(\mathbf{RV}_t)$ é uma função não linear e $\mathbf{RV}_t = (RV_t^{(1)}, RV_t^{(2)}, RV_t^{(3)}, \dots, RV_t^{(h)})$.

O objetivo da modelagem por Redes Neurais é aproximar a função $T(\mathbf{RV}_t)$ pela expressão 9,

$$T(\mathbf{RV}_t) \approx G(\mathbf{RV}_t; \boldsymbol{\psi}) := \beta_0' \mathbf{RV}_t + \sum_{i=1}^H \beta_i f(\mathbf{RV}_t; \boldsymbol{\gamma}_i), \quad (9)$$

em que, H é o número neurônios na camada oculta, $G(\mathbf{RV}_t; \boldsymbol{\psi})$ é uma função não linear das variáveis \mathbf{RV}_t , ajustada pelo vetor de parâmetros $\boldsymbol{\psi} = (\beta_i, \gamma_i)$, sendo que $f(\mathbf{RV}_t; \gamma_i)$ pode ser ajustada por função logística (expressão 10),

$$f(\mathbf{RV}_t; \gamma_i) = \frac{1}{1 + e^{-\gamma_i \mathbf{RV}_t}}, \quad (10)$$

ou uma função tangente hiperbólica (expressão 11),

$$f(\mathbf{RV}_t; \gamma_i) = \frac{e^{2\gamma_i \mathbf{RV}_t} - 1}{e^{2\gamma_i \mathbf{RV}_t} + 1}. \quad (11)$$

2.4 Combinação de projeções

No presente tópico é apresentada a abordagem da combinação de projeções. Esta metodologia foi inspirada no problema dos tomadores de decisão possuírem múltiplas projeções para a mesma variável.

A seleção de uma única projeção possui a vantagem do modelo utilizado ter demonstrado o menor erro de previsão. Entretanto, Timmermann (2006) demonstra que combinar as projeções possibilita a utilização do "melhor" que cada modelo pode oferecer, minimizando os erros de previsão.

A combinação pela média simples (pesos iguais) apresenta resultados difíceis de serem superados, conforme apresentam Makridakis e Winkler (1983), Makridakis et al. (1982), Clemen (1989), Chan, Stock e Watson (1999), Timmerman (2006), Smith e Wallis (2009), Genre et al. (2013) e Graefe et al. (2014). Portanto, frequentemente é utilizado como *benchmark*.

Por este motivo, utilizou-se neste trabalho a combinação a ser realizada com pesos iguais. Na expressão 12 é apresentada a ponderação por pesos iguais,

$$\omega_{i,t}^{ig} = \frac{1}{K}, \quad (12)$$

em que, K é o número de modelos a serem combinados.

Outro método de combinação muito comentado na literatura é a utilização da Média do Erro de Previsão (*Mean Squared Forecast Error - MSFE*). Conforme é apresentado em Stock e Watson (2001) e Stock e Watson (2004) o método associa maiores ponderações as previsões individuais que demonstrem menor MSFE na amostra utilizada para estimação do valor de m . O método pela MSFE é apresentado nas expressões 13 e 14,

$$\omega_{i,t}^{msfe} = \frac{m_{i,t}^{-1}}{\sum_{i=1}^n m_{i,t}^{-1}}, \quad (13)$$

$$m_{i,t} = \sum_{i=T_0}^t (y_{s+h}^h - \hat{y}_{i,s+h}^h)^2. \quad (14)$$

No Quadro 1 são apresentados alguns trabalhos que utilizaram a combinação de projeções.

Quadro 1 –Resumoutilização de combinação de projeções

Autor	Dados	Combinação	Resultado
Makridakis e Winkler (1983)	Séries macro e micro econômicas	Pesos iguais e MSFE	Pesos iguais melhor performance
Chan, Stock e Watson (1999)	Séries macro e micro econômicas do mercado americano	Pesos iguais, MSFE, Regressão, PCA, Mediana, PLS	Combinação melhor performance, pesos iguais melhor combinação
Stock e Watson (2004)	73 projeções de dados econômicos de 7 países da OECD	MSFE, <i>shrinkage</i> , pesos iguais, PCA, Filtro de Kalman	Combinação pesos iguais melhor performance
Cavaleri (2008)	Ibovespa, Dow Jones e IGP-M	Pesos iguais, <i>shrinkage</i> e Filtro de Kalman	Combinação por <i>shrinkage</i> melhor performance
Smith e Wallis (2009)	GDP e Índice de Produção Industrial	Pesos Iguais e MSFE	Resultados similares entre os métodos
Genre et al. (2013)	Inflação, GDP e Taxa de desemprego na Europa	Pesos iguais, PCA, Regressão, <i>shrinkage</i> , MSFE	Resultados similares entre os métodos
Huang e Lee (2013)	Série em alta frequência do índice S&P500	Pesos iguais, mediana das projeções e PCA	Combinação melhor performance
Andrada-Felix et al. (2013)	Índice S&P 100	Pesos iguais	Resultados similares com projeções individuais

Fonte: elaborado pelos autores.

3 METODOLOGIA

3.1 Dados e métodos

O objetivo do presente artigo é propor a combinação das projeções de Volatilidade Percebida obtidas pelos modelos HAR e Redes Neurais. Para realizar as projeções, combinações e testes foi utilizada a série de Volatilidade Percebida do índice Bovespa no período entre 03/01/2000 a 01/10/2014, total de 2692 observações.

A pesquisa possuiu 4 etapas: obtenção da série de volatilidade percebida (RV), estimação dos modelos HAR e Redes Neurais, combinação das projeções e avaliação de performance.

Para obtenção da série de volatilidade percebida foram utilizados os retornos continuamente compostos em intervalos de cinco minutos. No artigo de Wink Júnior e Pereira (2011) os autores testaram a utilização das frequências de 1, 5 e 15 minutos; os resultados demonstraram que a frequência de 5 minutos apresentou o menor intervalo de confiança.

Ainda como procedimento de limpeza dos dados, os retornos obtidos nos períodos de abertura e fechamento foram retirados, conforme foi proposto em Rydberg e Shephard (2003). O objetivo deste procedimento é retirar os efeitos dos leilões de abertura e fechamento das séries.

Na segunda etapa foram estimados os modelos HAR e Redes Neurais, foi utilizada 70% da amostra total, sendo ao todo 1884 observações, compreendendo o período entre 03/01/2000 a 01/09/2009.

Nesta etapa foram analisadas duas formas de estimar e prever a Volatilidade Percebida por meio do modelo HAR. A primeira forma foi utilizar o modelo HAR clássico apresentado em

Corsi (2003), em que os componentes diário (1 dia), semanal (5 dias) e mensal (22 dias) são utilizados como variáveis explicativas. No presente artigo, este modelo será denominado de HAR(1,5,22), sendo apresentado na expressão 15,

$$RV_{t+h}^{(d)} = \beta_0 + \beta^{(d)}RV_t^{(d)} + \beta^{(w)}RV_t^{(w)} + \beta^{(m)}RV_t^{(m)} + \epsilon_{t+h}. \quad (15)$$

Em que, os parâmetros $\beta^{(d)}$, $\beta^{(w)}$ e $\beta^{(m)}$ são os coeficientes para os horizontes de tempo diário, semanal e mensal respectivamente e ϵ_{t+h} é o resíduo. As séries $RV_t^{(w)}$ e $RV_t^{(m)}$ foram construídas conforme a expressão 6.

A segunda forma para estimar o HAR foi utilizar todos os componentes de agregações da Volatilidade Percebida dentro da amplitude entre 1 a 22 dias. A seleção sobre quais destas agregações se manteriam no modelo foi realizado pelo método *Stepwise*, buscando-se selecionar o conjunto de variáveis que maximizassem o BIC (*Bayesian information criterion*).

Neste artigo, este modelo será denominado HAR *Stepwise*, sendo apresentado na expressão 16,

$$RV_{t+h}^{(d)} = \sum_{i=0}^{22} \beta^{(i)}RV_t^{(i)} + \epsilon_{t+h}, \quad (16)$$

em que, os parâmetros $\beta^{(i)}$ são os coeficientes para os horizontes de tempo a serem selecionados pelo método *Stepwise* durante a estimação.

Dado que o modelo HAR foi analisado de duas formas, buscou-se analisar a técnica de redes Neurais também de duas formas. Na primeira forma, estendeu-se o modelo HAR clássico para uma função não linear, apresentada na expressão 17, no artigo este modelo será denominado de NNHAR(1,5,22),

$$RV_{t+h} = \sum_{i=1}^H \beta_i \left(\frac{e^{2(\gamma^{(d)}RV_t^{(d)} + \gamma^{(w)}RV_t^{(w)} + \gamma^{(m)}RV_t^{(m)})} - 1}{e^{2(\gamma^{(d)}RV_t^{(d)} + \gamma^{(w)}RV_t^{(w)} + \gamma^{(m)}RV_t^{(m)})} + 1} \right) + \epsilon_{t+h}. \quad (17)$$

Em que, H é o número neurônios na camada oculta e os parâmetros $\gamma^{(d)}$, $\gamma^{(w)}$ e $\gamma^{(m)}$ são os coeficientes para os horizontes de tempo diário, semanal e mensal respectivamente e β_i são os coeficientes dos neurônios de saída.

A segunda forma para projetar a Volatilidade Percebida por Redes Neurais foi utilizar todos os componentes de agregações da Volatilidade Percebida dentro da amplitude entre 1 a 22 dias. Neste artigo, este modelo será denominado NNHAR(1:22), sendo apresentado na expressão 18,

$$RV_{t+h} = \sum_{i=1}^H \beta_i \left(\frac{e^{2(\sum_{i=1}^{22} \gamma^{(i)}RV_t^{(i)})} - 1}{e^{2(\sum_{i=1}^{22} \gamma^{(i)}RV_t^{(i)})} + 1} \right) + \epsilon_{t+h}. \quad (18)$$

Em que, H é o número neurônios na camada oculta, β_i são os coeficientes dos neurônios de saída e $\gamma^{(i)}$ são os coeficientes para os horizontes de tempo que variam entre 1 a 22 dias.

No processo de estimar NNHAR(1,5,22) e NNHAR(1:22) é necessário definir a quantidade de neurônios na camada oculta H . Porém, selecionar um número pequeno de neurônios pode-se realizar uma estimação pouco eficaz. Por outro lado, escolher um número grande para H pode-se gerar um ajuste demasiado no modelo (*overfitting*), fazendo com que a técnica apresente fraco desempenho na amostra de teste.

Deste modo, buscou-se analisar o número ótimo de H numa janela entre 3 a 10 neurônios, procedimento similar utilizado em Fernandes et al. (2014). Como forma de controlar o ajuste demasiado foram utilizados dois critérios: fator de decaimento do erro e regularização Bayesiana.

No critério de fator de decaimento do erro, o processo de estimação é parado quando numa determinada iteração o decaimento do erro de previsão se torna aproximadamente estável. No critério de regularização Bayesiana, busca-se balancear o número de parâmetros estimados e o melhor ajuste do modelo. Deste modo, o algoritmo controla o *overfitting* tornando alguns dos coeficientes β_i e $\gamma^{(i)}$ praticamente insignificantes.

Ainda na etapa 2 foi possível realizar a comparação dos modelos dentro da amostra. As medidas utilizadas para avaliação foram Erro Médio Absoluto (*Mean Absolute Error* - MAE) e Raiz do Erro Médio Quadrático (*Root Mean Squared Error* - RMSE), expressões 19 e 20 respectivamente,

$$MAE = \frac{1}{n} \sum_{i=1}^n |\hat{Y}_i - Y_i|, \quad (19)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{Y}_i - Y_i)^2}, \quad (20)$$

desta forma, quanto menor os valores destas métricas, menor será o erro obtido pelo modelo utilizado para realizar a previsão.

Na terceira etapa foram obtidas as projeções dos modelos HAR e Redes Neurais, as previsões da volatilidade foram realizadas seguindo a sequência de projeções um passo a frente. Deste modo, utilizaram-se os valores atuais para representar a variável dependente defasada, conforme procedimento seguido por Val et al. (2014). Estes valores foram obtidos da amostra restante, sendo ao todo 808 observações; compreendendo o período de 10/07/2008 a 01/10/2014.

Considerando os métodos de estimação e os controles para evitar *overfitting*, foram realizadas previsões para 6 modelos: HAR(1,5,22), HAR *Stepwise*, NNHAR(1,5,22) com fator de decaimento de erro, NNHAR(1:22) com fator de decaimento de erro, NNHAR(1,5,22) com regularização Bayesiana e NNHAR(1:22) regularização Bayesiana.

Dado a existência de previsões para 6 modelos, buscou-se realizar as combinações com 2 a 6 modelos diferentes. Deste modo, foram realizadas 15 combinações com 2 modelos, 20 com 3 modelos, 15 combinações com 4 modelos, 6 combinações com 5 modelos e 1 combinação utilizando todos os modelos.

Para estimar os pesos das combinações, foi utilizada a expressão 12 para o método de pesos iguais e a expressão 13 para o método MSFE. Para estimar o valor de m da expressão 14, utilizou-se a própria amostra de estimação, compreendendo o período entre 03/01/2000 a 08/07/2008.

Na quarta etapa foi realizada a avaliação de desempenho de previsão, as medidas de comparação foram o MAE (expressão 19) e RMSE (expressão 20). As análises foram viabilizadas pelo software RStudio.

4 ANÁLISE DOS DADOS

4.1 Estimação e projeção dos modelos

Os resultados dos modelos HAR estimados são apresentados na Tabela 1, o período de análise ficou compreendido entre 03/01/2000 a 08/07/2008. Foram estimados os modelos HAR(1,5,22) e HAR *Stepwise*. No modelo HAR(1,5,22) os parâmetros $\beta^{(d)}$, $\beta^{(w)}$ e $\beta^{(m)}$ correspondem aos coeficientes para os horizontes de tempo diário, semanal (5 dias) emensal (22 dias).

Por meio da análise dos coeficientes do modelo HAR(1,5,22) observa-se a presença de alta persistência na volatilidade para o componente diário e semanal. Esta maior importância dos componentes diário e semanal para explicar a volatilidade percebida diária também foi encontrada nos artigos de Andersen et al.(2007), Corsi (2009) e Val et al.(2014).

Quanto ao modelo HAR *Stepwise*, observou-se valores altos de VIF, o que impossibilita a interpretação dos parâmetros dado a presença de multicolineariedade das variáveis explicativas (WOOLDRIDGE, 2005). Deste modo, este modelo é utilizado especificamente para objetivos de previsão, uma vez que a multicolineariedade não afeta a capacidade de previsão do modelo, conforme apresentam Makridakis et al. (1998), Berry e Vaughan (2005) e Shmueli (2010).

Na comparação entre ambos os modelos dentro do período de estimação, observa-se não existir diferenças na capacidade de previsão dos métodos. Sendo que o HAR(1,5,22) apresentou melhor resultado de MAE e o HAR *Stepwise* melhor desempenho quanto ao RMSE.

Tabela 1 - Resultados estimacões modelos HAR

HAR(1,5,22)	Coef.	VIF	HAR Stepwise	Coef.	VIF ¹
β_0	0,0072**	-	β_0	0,0077***	
$\beta^{(d)}$	0,2186***	1,8926	$\beta^{(2d)}$	0,3643***	3,1337
$\beta^{(w)}$	0,5640***	2,6467	$\beta^{(4d)}$	0,3108***	4,0726
$\beta^{(m)}$	0,1273***	1,9941	$\beta^{(13d)}$	-1,2018***	18,0688
			$\beta^{(14d)}$	1,8518***	21,4929
			$\beta^{(18d)}$	-2,3059***	26,9398
			$\beta^{(19d)}$	2,6489***	29,2493
			$\beta^{(22d)}$	-0,7637***	11,0907
RMSE	0,07720710			0,07560311	
MAE	0,03277686			0,03333526	

Significante a 5%. *Significante a 1%.

¹VIF: *Variance Inflation Factor*.

Fonte: dados da pesquisa.

Na Tabela 2 são apresentados os desempenhos dos modelos de Redes Neurais na amostra de estimacão. Os valores selecionados para H foram: 9 - NNHAR(1,5,22) decaimento de erro, 8 - NNHAR(1:22) decaimento de erro, 10 - NNHAR(1,5,22) regularizacão Bayesiana e 9 - NNHAR(1:22) regularizacão Bayesiana.

Comparando-se os resultados de RMSE e MAE, percebem-se os melhores resultados nas estimacões realizadas pelas Redes com regularizacão Bayesiana, destacando-se a utilizacão das volatilidades de 1 a 22 dias para a camada de entrada.

Tabela 2 - Resultados estimacões modelos Rede Neural

Modelo	Rede	RMSE dentro amostra	MAE dentro da amostra
NNHAR(1,5,22) decaimento de erro	3-9-1	0,07213572	0,03190166
NNHAR(1:22) decaimento de erro	22-8-1	0,06325006	0,03074167
NNHAR(1,5,22) regularizacão Bayesiana	3-10-1	0,05020683	0,02693103
NNHAR(1:22) regularizacão Bayesiana	22-9-1	0,04157260	0,02303130

Fonte: dados da pesquisa.

4.2 Combinaçao e Comparacão de performance

Após as estimacões dos modelos, foram realizadas as projeçoes individuais e as combinaçoes das projeçoes utilizando os métodos de pesos iguais e MSFE. O período para análise da amostra de teste ficou compreendido entre 10/07/2008 a 01/10/2014, sendo ao todo 808 observaçoes. O resumo dos principais resultados é apresentado na Tabela 3.

Tabela 3 – Comparaçoes dos modelos

Modelo	RMSE forada amostra	MAE fora da amostra
HAR(1, 5, 22)	0,0371	0,0209^b
HAR Stepwise	0,0370	0,0213
NNHAR(1,5,22) decaimento de erro	0,0367^a	0,0212
NNHAR(1:22) decaimento de erro	0,0417	0,0211
NNHAR(1,5,22) regularizacão Bayesiana	0,0384	0,0209^b
NNHAR(1:22) regularizacão Bayesiana	0,0379	0,0214
Combinaçao Pesos Iguais menor RMSE	0,0364^c	0,0208
Combinaçao método MSFE menor RMSE	0,0365^c	0,0210
Combinaçao Pesos Iguais menor MAE	0,0378	0,0205^d
Combinaçao método MSFE menor MAE	0,0379	0,0205^d

^amenor RMSE modelo individual, ^bmenor MAE modelo individual, ^cmenor RMSE combinaçao, ^dmenor MAE combinaçao.

Por meio da Tabela 3, podem-se comparar os resultados dos modelos individuais e das combinações. Quanto aos modelos individuais, destacou-se o modelo NNHAR(1,5,22) com fator de decaimento de erro por ter apresentado o menor RMSE na amostra de teste. Em relação ao menor MAE, destacaram-se os modelos HAR(1,5,22) e NNHAR(1,5,22) regularização Bayesiana.

Em relação a combinação pelo método de Pesos Iguais, a combinação que demonstrou o menor RMSE ocorreu quando as previsões dos modelos HAR *Stepwise*, NNHAR(1,5,22) decaimento de erro e NNHAR(1:22) regularização Bayesiana foram combinadas. A combinação que apresentou o menor MAE ocorreu quando as previsões dos modelos NNHAR(1,5,22) decaimento de erro, NNHAR(1,5,22) regularização Bayesiana e NNHAR(1:22) regularização Bayesiana foram combinadas.

Quanto ao desempenho pelo método MSFE, a combinação que demonstrou o menor RMSE foi observada quando as previsões dos modelos HAR(1:22) e NNHAR(1,5,22) decaimento de erro foram combinadas, sendo os pesos de 47,65% para HAR(1:22) e 52,35% para NNHAR(1,5,22).

A combinação que demonstrou menor MAE pode ser observada quando os modelos HAR *Stepwise*, NNHAR(1,5,22) decaimento de erro e NNHAR(1,5,22) regularização Bayesiana foram utilizados. Os pesos estimados foram 22,90% para HAR *Stepwise*, 25,16% para NNHAR(1,5,22) decaimento de erro e 51,94% para NNHAR(1,5,22) regularização Bayesiana.

Estes desempenhos observados na amostra de teste ocorreram conforme esperado pela teoria. Ou seja, combinar as projeções possibilitou minimizar erros de previsão melhor de que a escolha de um modelo individual. Os dois métodos de combinação utilizados apresentaram os menores valores de RMSE e MAE, destacando-se o método de combinação com Pesos Iguais, apresentado na literatura como difícil de ser superado.

Como resultado específico do presente artigo, observa-se a eficiência em combinar as projeções do modelo HAR e do método de Redes Neurais. Sendo que as combinações que apresentaram melhor desempenho tenderam a utilizar as previsões de ambos os modelos para obter os menores valores de RMSE e MAE.

5 CONSIDERAÇÕES FINAIS

O objetivo do presente artigo foi propor a combinação das projeções de Volatilidade Percebida obtidas pelos modelos HAR e Redes Neurais Bayesianas. Conforme Timmermann (2006) e diversos outros autores, a combinação de projeções demonstrou sucesso empírico em diversas aplicações econômicas, como projeção do PIB, preço do petróleo e retorno de ações.

Combinar ambas as técnicas para projeção da Volatilidade Percebida é inovativo, dado que duas áreas de conhecimento diferentes são utilizadas para melhorar o desempenho de projeções. A primeira é baseada em pressupostos da economia, enquanto que a segunda teve origem na área de física e engenharia.

Para analisar os resultados foi utilizada uma amostra da série de Volatilidade Percebida diária do Ibovespa, no período entre 03/01/2000 a 01/10/2014, com um total de 2714 observações. Nos resultados fora da amostra, observou-se que a combinação de ambos os modelos

resultaram numa melhor previsibilidade da Volatilidade Percebida. Consequentemente, mais eficiência para gestão de riscos.

Como limitações, o presente artigo analisou somente a série de Volatilidade Percebida do Ibovespa, sendo necessário verificar os resultados em mais séries. O estimador da Volatilidade Percebida utilizado foi o proposto por Andersen e Bollerslev (1998), sendo que medidas mais robustas a saltos e ruídos de microestrutura foram propostas nos últimos anos.

Em trabalhos futuros, mais métodos de combinação de projeções devem ser estudados, sendo que no presente trabalho limitou-se ao estudo dos métodos de pesos iguais e MSFE. Ainda em relação a combinação de projeções, existe uma necessidade de trabalhos na área, dado que ainda não existe um consenso entre os pesquisadores sobre qual peso ótimo deveria ser utilizado.

REFERÊNCIAS

ANDERSEN, T. G.; BOLLERSLEV, T. *Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. International Economic Review*, v. 39, n. 4, p. 885-905, 1998.

ANDERSEN, T. G. et al. *The distribution of realized stock return volatility. Journal of Financial Economics*, v. 61, p. 43-76, 2001.

ANDERSEN, T. G. et al. *Modeling and forecasting realized volatility. Econometrica*, v. 71, n. 2, p. 579-625, 2003.

ANDERSEN, T. G.; BOLLERSLEV, T.; DIEBOLD, F. X. *Roughing it up: including jump components in the measurement, modelling and forecasting of return volatility. The Review of Economics and Statistics*, v. 89, p. 701-720, 2007.

ANDRADA-FELIX, J.; FERNANDEZ-RODRIGUEZ, F.; FUERTES, A. *Combining nearest neighbor predictions and model-based predictions of realized variance: Does it pay? Working Paper*, 2013. Disponível em: <http://ssrn.com/abstract=2258457>. Acesso em: 20/01/2014.

BARNDORFF-NIELSEN, O. E.; SHEPHARD, N. *Power and bipower variation with stochastic volatility and jumps. Journal of Financial Econometrics*, Oxford Univ. Press, v. 2, n. 1, p. 1-37, 2004.

BAUWENS, L.; HAFNER, C.; LAURENT, S. *Volatility models and their applications*. New Jersey: John Wiley & Sons, 2012.

BERRY, K. E.; VAUGHAN, T. S. *Using Monte Carlo techniques to demonstrate the meaning and implications of multicollinearity. Journal of Statistics Education*, v. 13, n. 1, p. 1, 2005.

BOUDT, K.; CORNELISSEN, J.; PAYSEUR, S. *Highfrequency: Toolkit for the analysis of high frequency financial data in R. Working Paper*, 2013.

CAVALERI, R. **Combinações de previsões aplicada à volatilidade**. Dissertação (Mestrado). Programa de Pós-Graduação da Faculdade de Ciências Econômicas da Universidade Federal do Rio Grande do Sul - UFRGS, 2008.

CHAN, Y.; STOCK, J.; WATSON, M.A *Dynamic Factor Model Framework for Forecast Combination*. **Spanish Economic Review**, v. 1, p. 91–121, 1999.

CLEMEN, R. T. *Combining forecasts: a review and annotated bibliography*. **International Journal of Forecasting**, v. 5, p. 559-583, 1989.

CORSI, F. *A simple long memory model of realized volatility*. **Working Paper**, 2003. Disponível em: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=626064. Acesso em 17/03/2014.

CORSI, F. *A simple approximate long-memory model of realized volatility*. **Journal of Financial Econometrics**, v. 7, n. 2, p. 174-196, 2009.

DACOROGNA, M. et al. *Modeling short term volatility with GARCH and HARMA models*. In: DUNIS, C.; ZHOU, B. **Nonlinear Modeling of High Frequency Financial Time Series**. Chichester, UK: Wiley, 1998.

DANIÉLSSON, J. **Financial risk forecasting: The theory and practice of forecasting market risk with implementation in R and Matlab**. Chichester: John Wiley & Sons, 2011.

DONALDSON, R. G.; KAMSTRA, M. *An artificial neural network-GARCH model for international stock return volatility*. **Journal of Empirical Finance**, v. 4, n. 1, p. 17-46, 1997.

FERNANDES, M.; MEDEIROS, M. C.; SCHARTH, M. *Modeling and predicting the CBOE market volatility index*. **Journal of Banking & Finance**, v. 40, p. 1-10, 2014.

GENRE, V. et al. *Combining expert forecasts: Can anything beat the simple average?* **International Journal of Forecasting**, v. 29, n. 1, p. 108-121, 2013.

GRAEFE, A. et al. *Combining forecasts: An application to elections*. **International Journal of Forecasting**, v. 30, n. 1, p. 43-54, 2014.

HAMID, S. A.; IQBAL, Z. *Using neural networks for forecasting volatility of S&P 500 Index futures prices*. **Journal of Business Research**, v. 57, n. 10, p. 1116-1125, 2004.

HILLEBRAND, E.; MEDEIROS, M. C. *The benefits of bagging for forecast models of realized volatility*. **Econometric Reviews**, v. 29, n. 5-6, p. 571-593, 2010.

HUANG, H.; LEE, T. *Forecasting Value-at-Risk using high-frequency information*. **Econometrics**, v. 1, p. 127-140, 2013.

KUAN, C.; WHITE, H. *Artificial neural networks: an econometric perspective*. **Econometric Reviews**, v. 13, n. 1, p. 1-91, 1994.

MAKRIDAKIS, S. et al. *The accuracy of extrapolation (time series) methods: Results of a forecasting competition*. **Journal of forecasting**, v. 1, n. 2, p. 111-153, 1982.

MAKRIDAKIS, S.; WINKLER, R. L. *Averages of forecasts: Some empirical results. Management Science*, v. 29, n. 9, p. 987-996, 1983.

MAKRIDAKIS, S.; WHEELWRIGHT, S. C.; HYNDMAN, R. J. *Forecasting methods and applications*. John Wiley & Sons, 2008.

MÜLLER, U. et al. *Volatilities of different time resolutions: Analysing the dynamics of market components. Journal of Empirical Finance*, v. 4, p. 213-239, 1997.

OLIVEIRA, M. A.; MONTINI, A. A.; MENDES-DA-SILVA, W.; BERGMANN, D. R. Estabilidade de preços de ações no mercado de capitais brasileiro: um estudo aplicando redes neurais e expoentes de Lyapunov. **Revista de Administração - RAUSP**, v. 46, n. 2, p. 161-177, 2011.

RYDBERG, T. H.; SHEPHARD, N. *Dynamics of trade-by-trade price movements: decomposition and models. Journal of Financial Econometrics*, v. 1, n. 1, p. 2-25, 2003.

SHMUELI, G. *To explain or to predict? . Statistical science*, p. 289-310, 2010.

SMITH, J.; WALLIS, K. F. *A simple explanation of the forecast combination puzzle. Oxford Bulletin of Economics and Statistics*, v. 71, n. 3, p. 331-355, 2009.

STOCK, J. H.; WATSON, M. W. *A comparison of linear and nonlinear univariate models for forecasting macroeconomic time series. Working Paper*, 2001.

STOCK, J. H.; WATSON, M. W. *Combination forecasts of output growth in a seven country data set. Journal of Forecasting*, v. 23, p. 405-430, 2004.

TIMMERMANN, A. *Forecast combinations. Handbook of economic forecasting*, v. 1, p. 135-196, 2006.

VAL, F. F.; PINTO, A. C. F.; KLOTZLE, M. C. Volatilidade e previsão de retorno com modelos de alta frequência e GARCH: evidências para o mercado brasileiro. **Revista Contabilidade & Finanças**, v. 25, n. 65, p. 189-201, 2014.

WOOLDRIDGE, J. M. *Introductory econometrics: A modern approach*. Nelson Education, 2015.

ZIVOT, E. *Analysis of high frequency financial data: models, methods and software. Part II: Modeling and forecasting realized variance measures. Technical Report UWEC-2005-03*, 2005.