

Aplicação de Redes Neurais Polinomiais na Previsão do Ibovespa e Merval

Everton Anger Cavalheiro¹

Kelmara Mendes Vieira²

Paulo Sérgio Ceretta³

José Carlos Severo Correa⁴

Carlos Frederico de Oliveira Cunha⁵

Resumo

Este artigo analisa a eficiência das redes neurais polinomiais *Group Method of Data Handling* (GMDH) na previsão dos retornos, em bases mensais, dos principais indicadores do mercado de capitais do brasileiro (Ibovespa) e argentino (Merval). Inicialmente, para a determinação da variável exógena, calculou-se o retorno logarítimo de cada um dos índices. Em seguida, para a determinação das variáveis endógenas, realizaram-se defasagens em $t-1$, $t-2$ e $t-3$ da variável exógena. Calcularam-se até nove camadas alimentadas para frente. Os resultados sugerem relativa previsibilidade de ambos os mercados, denotando certa ineficiência. A ineficiência, especialmente do mercado argentino, é corroborada por testes de causalidade de Granger adicionais, que demonstram a influência da Bolsa de Valores de São Paulo em cima da Bolsa de Valores de Buenos Aires e não havendo esta influência no sentido inverso.

Palavras-chave: Redes neurais. Cointegração. Causalidade de Granger.

¹ Mestre em Administração pela Universidade Federal de Santa Maria (UFSM). Professor da Universidade de Cruz Alta (Unicruz). ecavalheiro@unicruz.edu.br

² Doutora em Administração pela Universidade Federal do Rio Grande do Sul (UFRGS). Professora adjunta da Universidade Federal de Santa Maria (UFSM). kelmar@terra.com.br

³ Doutor em Engenharia de Produção pela Universidade Federal de Santa Catarina (UFSC). Professor Adjunto da Universidade Federal de Santa Maria (UFSM). Bolsista do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). ceretta@smail.ufsm.br

⁴ Mestre em Engenharia de Produção pela Universidade Federal de Santa Maria (UFSM). Professor da Universidade de Cruz Alta (Unicruz). jcorrea@unicruz.edu.br

⁵ Mestre em Desenvolvimento, Agricultura e Sociedade pela Universidade Federal Rural do Rio de Janeiro (UFRRJ). Professor da Universidade de Cruz Alta (Unicruz). Professor colaborador da Universidade Regional do Noroeste do Estado do Rio Grande do Sul (Unijui). ccunha@unicruz.edu.br

Abstract

This article analyses the efficiency of Group Method of Data Handling (GMDH) polynomial neural networks when anticipating return, on a monthly basis, on the return of the main Brazilian (Ibovespa) and Argentinean (Merval) market indicators. Initially, in order to determine the exogenous variable, we calculated the logarithmical return on each index. Afterwards, in order to determine the endogenous variables, we have performed $t-1$, $t-2$ and $t-3$ lags on the exogenous variable. We computed up to nine front fed layers. Results suggest some predictability on both markets, denoting some inefficiency. Inefficiency, especially on the Argentinean market, is validated by the additional causality Granger tests that demonstrate the influence of the São Paulo Stock Market over the Buenos Aires Stock Market and no such influence the other way round.

Keywords: Neural networks. Co-integration. Granger causality.

Consideradas caixas de ressonância das finanças, da economia e das políticas regionais, as bolsas de valores dos países pertencentes ao Mercado Comum do Sul (Mercosul) não representam com fidelidade a dimensão dos mercados de capitais nos países do bloco comercial. A exemplo, a Bolsa de Valores de Montevideú que negocia um volume significativo de recursos aportados em títulos de renda fixa e um pequeno volume de investimentos atrelado ao mercado acionário. Por outro lado, no Mercosul duas bolsas se destacam: a brasileira Bovespa e a argentina Merval.

Mariano Grondona, um dos mais influentes colonistas de jornal e televisão de Buenos Aires, intitulou a Argentina como uma província do Brasil, pois o país irmão, conforme dados da balança comercial de 2010, demonstra realmente depende do parceiro do Mercosul para sobreviver economicamente. Naquele ano, a Argentina exportou aproximadamente 14% dos produtos ao Brasil, enquanto este exportou somente 9% de seus produtos à Argentina. Desde 2004 a Argentina é deficitária perante o Brasil. Em números absolutos, no ano de 2010 os portenhos importaram US\$ 18 bilhões e exportaram ao Brasil US\$ 14 bilhões.

Esta relação de dependência, contudo, poderia extrapolar as relações de comércio exterior porque, apesar de a Argentina não dispor de empresas listadas na Bovespa, dentre os 14 papéis do índice Merval duas empresas são controladas pela Petrobras. A Petrobras argentina está na quinta posição, com 7,01% em volume de negócios, em que se destaca a siderúrgica Tenaris, com uma representatividade de 24,75% no índice Merval, enquanto a Petrobras Energia aparece em 12º lugar, com 3,01% do giro total.

Fundado na década de 20, o Mercado de Valores de Buenos Aires (Merval) chegou a movimentar em torno de US\$ 100 milhões por dia nos anos 90, durante o período da paridade cambial da moeda argentina com o dólar. Constantes crises políticas e econômicas, bem como a escassez crônica de crédito, atingiram todos os segmentos da sociedade portenha, acarretando o encolhimento da Bolsa de Buenos Aires. Por outro lado, a Bolsa de Valores de São Paulo movimenta diariamente cerca de US\$ 3,3 bilhões, sendo, aparentemente, o principal parâmetro de preços da Bolsa de Buenos Aires.

Os índices Merval e Ibovespa são considerados os mais importantes indicadores do mercado acionário de seus países, o que torna relevante as pesquisas relacionadas à previsibilidade destes mercados. Ferson (2007) destaca que o interesse pela previsão do comportamento dos preços das ações é provavelmente tão antigo quanto os próprios mercados, bem como a literatura sobre o assunto é vasta e significativa. Neste sentido, o uso de técnicas pertencentes à grande área de inteligência artificial (IA) para a previsão de séries temporais financeiras tem se mostrado cada vez mais presente nos testes empíricos, havendo um especial destaque para o uso da técnica de Redes Neurais Artificiais (RNAs).

De acordo com Haykin (1999), as pesquisas relacionadas às RNAs buscam incessantemente modelar o comportamento do cérebro, isto porque a obtenção de um modelo matemático que determine o entendimento completo do seu funcionamento daria condições para que todo o processamento de sinais pudesse ser reproduzido por alguma máquina. Apesar da complexidade do cérebro, algumas características já são conhecidas e podem ser modeladas por Redes Neurais Artificiais (RNA). A aplicação de RNA em modelagens matemáticas tem se mostrado cada vez mais presente na vida do homem. Nesse sentido, o uso de RNA na previsão de séries temporais financeiras tem gerado constante demanda. Sua utilização é representativa, pois, conforme Boose (1994), a área de inteligência artificial é um campo de estudo multidisciplinar e interdisciplinar, que se apoia no conhecimento e evolução de outras áreas deste.

O *Group Method of Data Handling* (GMDH) é uma derivação do campo das Redes Neurais Artificiais. Esse modelo é caracterizado pela sua habilidade de selecionar a combinação de variáveis exógenas mais significativas para predizer uma variável endógena, a partir da manipulação dos dados para encontrar o melhor ajuste. Raros estudos, contudo, têm utilizado as modelagens polinomiais, especialmente as do tipo GMDH, em séries temporais financeiras. Desta forma, a pergunta de pesquisa deste artigo é: “*O uso de redes polinomiais GMDH é suficientemente capaz de demonstrar não aleatoriedade no retorno mensal dos índices Merval e Ibovespa?*”

O objetivo geral do trabalho é comparar a previsibilidade dos mercados de capitais do Brasil e da Argentina, mediante o uso de redes neurais GMDH para séries de dados do retorno mensal dos índices destes países. Secundariamente, procurou-se identificar possíveis tendências comuns entre as duas séries de dados que poderiam indicar certa cointegração dos mercados, os quais poderiam ser influenciados pela existência de empresas de controladoras brasileiras listadas no índice Merval.

Inteligência artificial e redes neurais artificiais

A palavra “inteligência” vem do latim *inter* (entre) e *legere* (escolher). De acordo com Legg e Hutter (2007), inteligência significa aquilo que permite ao ser humano escolher entre uma coisa e outra. Além disso, inteligência é a habilidade de realizar de forma eficiente uma determinada tarefa. A palavra “artificial” vem do latim *artificiale*; significa algo não natural, isto é, produzido pelo homem. Inteligência Artificial (IA), portanto, é uma espécie de inteligência produzida pelo homem para dotar as máquinas de algum tipo de habilidade que simula a inteligência humana.

As Redes Neurais Artificiais (RNAs) são um conjunto de técnicas que tentam simular, em meio computacional, o funcionamento do cérebro humano de uma maneira simplificada. Elas são capazes de reconhecer padrões, extrair regularidades e detectar relações subjacentes em um conjunto de dados aparentemente desconexos. Além disso, elas apresentam capacidade de lidar com dados ruidosos, incompletos ou imprecisos. Sua capacidade de prever sistemas não lineares torna a sua aplicação no mercado financeiro muito objetiva. Para Haykin (1999), rede neural pode ser definida como um processador maciço e paralelamente distribuído, constituído de unidades de processamento simples, que têm a propensão natural para armazenar conhecimento experimental e torná-lo disponível para uso. A rede neural assemelha-se ao cérebro humano em dois aspectos: (1) o conhecimento é

adquirido pela rede a partir de seu ambiente por intermédio de um processo de aprendizagem; (2) forças de conexão entre neurônios (pesos sinápticos) são utilizadas para armazenar o conhecimento adquirido.

Já para Lippmann (1987), as RNAs são sistemas físicos que podem adquirir, armazenar e utilizar conhecimentos experimentais, adquirindo uma boa performance devido à sua interconexão entre os nós da rede. Basicamente há dois tipos de Redes Neurais Artificiais: *feedforward* e *feedback*. Nas redes *feedforward*, os sinais se propagam em apenas uma direção a partir da unidade de entrada, passando pelas camadas intermediárias até a saída. Nas redes *feedback*, os sinais de entrada podem propagar da saída de qualquer neurônio para a entrada em um outro neurônio.

As RNAs apresentam duas fases distintas em sua metodologia: a da aprendizagem e a do teste. Na fase de aprendizagem são apresentados estímulos de entrada, denominados padrões de treinamento, que farão com que a rede aprenda com os dados. Na fase de teste é demonstrada a capacidade de generalização da rede, pois seus resultados deverão ser significativos após seus pesos terem sido ajustados na fase anterior.

A propriedade mais importante das Redes Neurais Artificiais é a habilidade de aprender por meio de seu ambiente e, com isso, melhorar o seu desempenho. Segundo Ackley, Hinton e Sejnowski (1985), a habilidade de aprender ocorre mediante um processo iterativo de ajustes aplicado aos pesos: o treinamento. O aprendizado ocorre quando a rede neural atinge uma solução generalizada para uma classe de problemas.

Redes polinomiais GMDH

O método GMDH (*Group Method of Data Handling*) é composto por um algoritmo proposto pelo russo Alexey Ivakhnenko no final da década de 60. Ele consiste em um método matemático para estimar estados de um sistema, saídas de controladores e funções de atuadores (Ivakhnenko, 1969). O algoritmo proposto inicialmente pode ser considerado auto-organizável e

de propagação indutiva para a solução de problemas práticos e complexos. Além disso, é possível obter um modelo matemático do processo a partir de observações de amostras de dados, que será utilizado na identificação e no reconhecimento de padrões ou, até mesmo, para descrever o próprio processo.

De acordo com Ahmadi, Mottaghtalab e Nariman-Zadeh (2007), o uso de redes de auto-organizáveis do tipo GMDH tem sido bem-sucedido em uma ampla gama de áreas de estudo. Complementam Mottaghtalab et al. (2010) com os bons resultados deste tipo de rede em áreas mais específicas, como a Engenharia e a Economia. A maioria dos algoritmos GMDH emprega funções de referência polinomiais. Uma conexão genérica entre variáveis de entrada e de saída pode ser expressa pela série funcional de Volterra, análogo do polinômio de Kolmogorov-Gabor:

$$y = \beta_0 + \sum_{i=1}^n \beta_i x_i + \sum_{i=1}^n \sum_{j=1}^n \beta_{ij} x_i x_j + \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n \beta_{ijk} x_i x_j x_k + \xi. \quad [1]$$

Em [1], x_i, x_j, x_k são as variáveis endógenas, β_0, β_{ij} e β_{ijk} os coeficientes do polinômio e ξ o erro.

O conteúdo do algoritmo de Ivakhnenko (1969) desenvolveu-se como veículo para identificar relações lineares e não lineares entre *inputs* e *outputs*, gerando uma estrutura que tende a ótima a partir de um processo sucessivo de várias manipulações dos dados, mediante a incorporação de novas camadas.

De acordo com Schneider e Steiner (2006), os algoritmos da rede GMDH são mais poderosos que os modelos estatísticos tradicionais devido aos seus procedimentos auto-organizáveis. Estes algoritmos extraem informação dos dados existentes utilizando-se da complexidade obtida por meio de modelos matemáticos flexíveis, e, além disso, quebram o conceito de caixa preta, característico das redes neurais.

O modelo GMDH pode ser analisado como uma combinação de redes neurais e conceitos estocásticos, conforme cita Valença (2005). As redes GMDH são criadas com funções de ativação nos neurônios das camadas escondidas, além de apresentarem um critério de seleção para decidir quantas camadas serão construídas. Na formulação original, cada neurônio da camada escondida, a ser construída, recebe duas entradas e tem como função de ativação um polinômio de segundo grau. Mediante a combinação de cada par desses neurônios de entrada, portanto, será gerada uma função de saída polinomial cuja complexidade será função do número de camadas, isto é, se existir duas camadas, tem-se uma função polinomial de quarto grau; para três camadas, uma função polinomial de oitavo grau, e assim por diante. Desta forma, tais redes são chamadas de redes polinomiais, pois o modelo resultante é uma função polinomial.

Algoritmo GMDH

Algoritmo é um processo de cálculo em que se estipulam, com generalidade e sem restrições, regras formais para a obtenção do resultado ou da solução do problema. Nas redes GMDH são desenvolvidos, entre outros, o proposto por Ivakhnenko (1971), em que os neurônios gerados nas camadas intermediárias são funções polinomiais de segundo grau compostos de duas variáveis. A equação [2] exemplifica o caso de uma saída (y) gerada a partir de três entradas (x_1 , x_2 e x_3).

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_1^2 + \beta_3 x_1 x_2 + \beta_4 x_1 x_3 + \beta_5 x_2 x_3 + \beta_6 x_2 + \beta_7 x_2^2 + \beta_8 x_3 + \beta_9 x_3^2 + \xi. \quad [2]$$

Em [2], y é a saída estimada, x_1 , x_2 e x_3 as variáveis de entrada, β_i seus pesos e ξ o erro. Valença (2005) destaca que cada camada construída terá uma quantidade de neurônios que depende do número de variáveis de entrada. Supondo que o número de variáveis de entrada é n , o número de neurônios na primeira camada será de C_2^n . Esses neurônios terão seu desempenho avaliado por uma função objetivo com limiar que decidirá

quais neurônios deverão continuar e quais neurônios deverão parar. Este processo é então repetido até que reste apenas um neurônio, ou então quando as camadas adicionais não provocarem nenhuma melhoria, ou até mesmo degeneração.

Valença (2005) salienta que um importante aspecto a ser considerado no algoritmo GMDH é o uso de um critério externo de seleção dos melhores neurônios em cada camada construída. Vários são os critérios de seleção que podem ser empregados para decidir sobre a qualidade de cada modelo testado, de tal forma que o autor detalha alguns deles: i) regularidade, que utiliza o erro quadrático médio sobre o conjunto de validação; ii) critério de mínima tendência, que tem por objetivo minimizar a diferença entre os valores calculados para os conjuntos de dados A e B; iii) critério de predição, sendo indicado quando o conjunto de dados é suficientemente grande para ser dividido em três conjuntos (A, B e C) e; iv) critério de balanço de variáveis, sendo este mais indicado para previsões de longo prazo.

Algoritmo Combinatorial

O aspecto básico do algoritmo combinatorial é a sua estrutura em uma única camada e, de acordo com Valença (2005), as funções de soma são geradas para todas as combinações de variáveis de entrada de forma similar a uma busca polinomial. Esse algoritmo se torna eficiente pela utilização de uma técnica recursiva dos mínimos quadrados, uma vez que os modelos gerados são lineares em relação aos parâmetros.

Para um melhor entendimento, supondo-se uma função com três variáveis de entrada (x_1 , x_2 e x_3) e uma variável de saída (y), haverá uma busca polinomial do melhor modelo até o segundo grau, devendo se começar pelo modelo mais simples ($y = \beta_0 + \xi$) até o mais complexo, semelhante ao demonstrado na equação [2]. A seguir, demonstra-se o incremento gradual da complexidade para este algoritmo contendo uma variável:

$$y = \beta_0 + \xi, \quad [3]$$

$$y = \beta_1 x_1 + \xi, \quad [4]$$

$$y = \beta_2 x_1^2 + \xi, \quad [5]$$

...

$$y = \beta_9 x_3^2 + \xi. \quad [6]$$

Como demonstrado nas equações [3, 4, 5 e 6], existirão $C_{10}^1 = 10$ neurônios, sendo os pesos determinados pelo método dos mínimos quadrados, com um conjunto de treino. Na sequência, os erros obtidos por equação são comparados a um limiar estabelecido para cada função objetivo (complemento externo), utilizando-se um conjunto de dados teste. Finalmente, as unidades selecionadas, ditas sobreviventes, serão consideradas aptas para terem sua complexidade evoluída. Num segundo momento, todos os neurônios com duas variáveis são determinados, totalizando $C_{10}^2 = 45$ equações parciais.

$$y = \beta_0 + \beta_1 x_1 + \xi, \quad [7]$$

$$y = \beta_0 + \beta_2 x_1^2 + \xi, \quad [8]$$

...

$$y = \beta_8 x_3 + \beta_9 x_3^2 + \xi. \quad [9]$$

Os neurônios serão construídos até que no último passo obtenha $C_{10}^{10} = 1$ combinações, que é o polinômio completo demonstrado na equação [2]. É importante notar que, dependendo se o número de entradas for grande, irá ocorrer uma explosão de combinações, pois o número de modelos que devem ser testados aumenta rapidamente com o número de variáveis de entrada, como demonstra: $M=2^n-1$. No exemplo de três variáveis de entrada (x_1, x_2 e x_3) acrescida de uma constante, que se desdobraram em dez variáveis.

Rede Multilayer

De acordo com Valença (2005), a rede *multilayer* é uma estrutura paralela, constituída por um polinômio de segundo grau como função de ativação. O número de neurônios gerados, em cada camada, depende do número de variáveis de entrada. Neste modelo, cada duas variáveis são combinadas para gerar um neurônio escondido, por exemplo: x_i e x_j são passados para a unidade k que pode ser construída por uma função como demonstrado:

$$y = \alpha_0 + \alpha_1 x_i + \alpha_2 x_j + \alpha_3 x_i x_j + \alpha_4 x_i^2 + \xi \quad [10]$$

Os pesos são estimados utilizando-se um conjunto de treinamento e um dado valor de limiar, além disso, um critério de erro é usado para avaliar o ajuste destes neurônios gerados. Quando se tem n variáveis de entrada, a primeira camada gerará C_2^n , isto é, $n(n-1)/2$ neurônios parciais. Desses, escolhem-se os que são superiores a um dado valor de limiar estabelecido, que passarão a compor a próxima camada. Este processo é repetido para sucessivas camadas até que o mínimo global para um dado critério de erro seja encontrado. Os critérios adotados para finalizar o treinamento, citados por Valença (2005), são: i) a camada de saída possui apenas um neurônio, ii) a camada adicionada não fornece ganhos significativos em relação à anterior.

Aspectos metodológicos

Neste estudo foram utilizadas como variáveis o retorno mensal, na forma logarítmica, dos índices Ibovespa (Brasil) e Merval (Argentina), na forma da primeira diferença em $t-1$, $t-2$ e $t-3$, isto é, r_{t-1} , r_{t-2} e r_{t-3} que formarão x_1 , x_2 e x_3 , respectivamente.

A escolha dos índices deu-se pelo fato de serem considerados os mais importantes indicadores do desempenho médio das cotações do mercado de ações destes dois países. Foram empregados, na fase de treinamento, os

retornos mensais de janeiro de 2000 até fevereiro de 2007, perfazendo 86 observações. Na fase de teste foram simuladas 47 novas observações, entre março de 2007 até janeiro de 2011.

Tsay (2002) cita que a maioria dos estudos de séries temporais financeiras utiliza os retornos em vez dos preços dos ativos em questão; neste contexto, o autor comenta que existem duas razões principais para o uso dos retornos em estudos financeiros: em primeiro lugar, para a média dos investidores o retorno dos ativos é uma medida adequada para a comparação entre oportunidade de investimentos e, em segundo lugar, as séries de retorno são mais fáceis de lidar do que uma série de preços, porque os primeiros apresentam propriedades estatísticas mais atraentes. Dentre tais propriedades pode-se citar a ausência de tendenciosidade comum em séries de dados não estacionados. Tsay (2002) comenta que a utilização de log-retornos em estudos financeiros é indicada pela hipótese de que os retornos dos ativos são independente e identicamente distribuídos (i.i.d.) com média μ e variância σ^2 .

Inicialmente, as variáveis puras, x_1 , x_2 e x_3 foram acrescidas de uma constante. Na sequência, estas variáveis formarão uma nova combinação multiplicando-se duas a duas, formando as novas variáveis: $1, x_1, x_1^2, x_1x_2, x_1x_3, x_2x_3, x_2, x_2^2, x_3, x_3^2$.

Cálculo da primeira camada da rede

Para iniciar o cálculo da primeira camada, e para definir o número de neurônios a serem testados, foi utilizado o conceito do algoritmo combinatorial (Combi). Houve o desenvolvimento de combinações entre as dez variáveis, citadas anteriormente. Na equação [11] é apresentado o modelo de cálculo do total de combinações que foram testadas em cada camada, onde N é o número de neurônios.

$$N = C_{10}^1 + C_{10}^2 + C_{10}^3 + C_{10}^4 + C_{10}^5 + C_{10}^6 + C_{10}^7 + C_{10}^8 + C_{10}^9 + C_{10}^{10} \quad [11]$$

Posteriormente, houve uma busca polinomial, semelhante ao método *stepwise*. A diferença básica está no critério de decisão que será adotado no modelo. O critério de decisão adotado foi o *Predict Squared Error (PSE)*, pois este critério, além de levar em consideração o erro médio quadrático (MSE), leva em consideração a variância do erro e uma penalidade para os modelos mais complexos, o que poderia evitar o problema de *overfitting*. O método de cálculo do *PSE* foi definido na equação [12].

$$PSE = \left[\frac{\sum_1^N (y_i - \hat{y}_i)^2}{n - 2} \right] + 2\sigma_\varepsilon^2 (p / n) \rightarrow \min. \quad [12]$$

Em [12], σ_ε^2 é a variância do erro; p é o número de parâmetros; n o número de observações no arquivo de treinamento da rede; y_i são os valores tabelados da variável endógena e \hat{y}_i são os valores calculados de acordo com o modelo.

Foram selecionados os três melhores neurônios nesta primeira camada segundo o critério *PSE*. Como o *PSE* apresenta uma relação direta entre o erro médio quadrático e a variância do erro, quanto menor o valor calculado em cada neurônio, maior será o ajustamento da variável estimada (\hat{y}_i) com a variável observada (y_i).

Cálculo nas camadas subsequentes

Os três neurônios selecionados formaram as novas variáveis de entrada para a próxima camada, isto é, w_1 , w_2 e w_3 . Houve um novo desdobramento das variáveis e após ocorreu uma busca polinomial, entre as próximas combinações de variáveis, selecionando os três melhores neurônios a fim de diminuir o critério de decisão *PSE*. Esse processo de treinamento e seleção foi então repetido para cada nova camada.

Critério para a finalização do treinamento e análise da previsão

Conforme destaca Valença (2005), os critérios utilizados para finalizar o treinamento são: a camada adicionada não oferece ganhos significativos em relação à anterior ou o valor do MSE para o conjunto de treinamento é superior ao MSE da camada anterior.

Nesse caso, o modelo ótimo estaria na camada anterior. Para avaliar o sucesso das previsões, conforme orientam Ivakhnenko, Ivakhnenko e Müller (1993), foi utilizada a equação [12]. Os resultados inferiores, ou iguais, a 0,5 seriam considerados adequados, os que estiverem entre $0,5 < \delta^2 < 0,8$ seriam considerados satisfatórios, os maiores que 1 seriam considerados desinformação e as modelagens seriam consideradas como ineficientes.

$$\delta_i^2 = \frac{\sum_1^N (y_i - \hat{y}_i)^2}{\sum_1^N (y_i - \bar{y})^2} \rightarrow \min. \quad [13]$$

Conforme destacam Ivakhnenko e Ivakhnenko (1995), com essa abordagem, para cada modelagem a ser realizada em redes GMDH usou-se dois critérios: em um primeiro momento, houve uma exaustiva pesquisa entre todos os neurônios candidatos, conforme as combinações citadas no subitem “Cálculo nas camadas subsequentes”, e um pequeno número de modelos cuja estrutura esteve perto de ótima foram selecionados. Neste caso foi adotado o PSE , conforme foi demonstrado na equação [12]. Então, apenas o neurônio selecionado como ótimo na última camada foi testado em sua conformidade com um especial critério discriminatório, sendo, neste artigo, demonstrado na equação [13].

Análise dos erros e comparação entre mercados

A fim de se comparar a previsibilidade dos mercados, foi utilizado o coeficiente de determinação amostral R^2 . Este critério procura medir a proporção ou percentual de variação de y previsto pelos modelos, conforme demonstrado na equação [14].

$$R^2 = 1 - \frac{\sum_1^N (\hat{y}_i)^2}{\sum_1^N (y_i - \bar{y})^2}. \quad [14]$$

Outros dois indicadores foram utilizados: o erro quadrático médio (MSE) e o erro absoluto médio (MAE), que são demonstrados nas equações [15] e [16].

$$MSE = \sqrt{\frac{1}{N} \sum_1^N (y_i - \hat{y}_i)^2}. \quad [15]$$

$$MAE = \frac{1}{N} \sum_1^N \left| \sqrt{y_i^2} - \sqrt{\hat{y}_i^2} \right|. \quad [16]$$

Adicionalmente foram analisados os coeficientes de desigualdade de Theil, também chamados de U . O numerador de U é o MSE , mas a escala do denominador é tal que U existe no intervalo de 0 a 1. Em que $U=0$ seria um ajustamento perfeito da previsão com o valor observado e que $U=1$, o desempenho da previsão do modelo seria o pior possível. O coeficiente de desigualdade de Theil foi demonstrado na equação [17].

$$U = \frac{\sqrt{\frac{1}{N} \sum_1^N (y_i - \hat{y}_i)^2}}{\sqrt{\frac{1}{N} \sum_1^N (y_i)^2} + \sqrt{\frac{1}{N} \sum_1^N (\hat{y}_i)^2}}. \quad [17]$$

Além do coeficiente de desigualdade de Theil, analisou-se as proporções U^M e U^S (proporção da tendenciosidade e proporção da variância), as quais permitem decompor o erro em suas fontes características.

Conforme Pindyck e Rubinfeld (1991), a proporção de tendenciosidade (U^M) analisa um possível erro sistemático, pois mede quanto os valores médios das séries simulada e efetiva se desviam uns dos outros. Qualquer que seja o valor do coeficiente de desigualdade (U), espera-se que U^M seja próximo de 0. Um valor elevado de U^M (acima de 0,1 ou 0,2) seria preocupante porque indicaria a presença de tendenciosidade sistemática, de modo que seria necessário rever os modelos. Nas equações [18] e [19] foram demonstradas a proporção de tendenciosidade e proporção de variância, respectivamente.

$$U^M = \frac{\left(\bar{y}^S - \bar{y}^A\right)^2}{(1/T) \sum \left(y_t^S - y_t^A\right)^2} \quad [18]$$

$$U^S = \frac{\left(\sigma_S - \sigma_A\right)^2}{(1/T) \sum \left(y_t^S - y_t^A\right)^2} \quad [19]$$

Em [17 e 18] \bar{y}^S , \bar{y}^A , σ_S e σ_A são as médias e os desvios-padrão dos valores estimados e observados, respectivamente. A proporção de variância U^S , conforme citam Pindyck e Rubinfeld (1991), indica a capacidade de replicar o grau de variabilidade na variável que interessa. Se U^S fosse alto, significaria que a série efetiva flutuou muito enquanto a série simulada mostrou pouca flutuação, ou vice-versa. Isso também seria preocupante, e poderia levar à revisão dos modelos.

Volatilidade dos mercados

A fim de se verificar possíveis diferenças de volatilidade entre os retornos mensais de cada um dos mercados, utilizou-se a equação [20].

$$\lambda = \sqrt{\frac{1}{(N-1) \times \sum_1^N \left(X - \bar{X}\right)^2}} \quad [20]$$

Em [20] λ é a volatilidade mensal dos mercados, N o número de meses, X o retorno mensal de cada ativo e \bar{X} a média dos retornos de cada índice.

Resultados

Foram calculadas até 9 camadas na rede para cada uma das previsões dos índices Ibovespa e Merval.

Resultados da fase de teste

Na Tabela 1 são demonstrados os resultados das previsões do índice argentino Merval.

Tabela 1 – Resultados das previsões para o índice argentino Merval

Camada	R^2	Correlação	Sinais	EMQ	MAE	U	U^M	U^S	Ivakhnenko
Camada 1	0,0008	0,0274	0,5745	0,0118	0,0754	0,0660	0,0459	0,0013	1,2308
Camada 2	0,0000	0,0019	0,4681	0,0114	0,0819	0,0644	0,0259	0,0015	1,1861
Camada 3	0,0479	0,2190	0,5106	0,0095	0,0723	0,0530	0,0441	0,0013	0,9847
Camada 4	0,1736	0,4167	0,5745	0,0083	0,0648	0,0451	0,0903	0,0011	0,8651
Camada 5	0,2146	0,4632	0,5745	0,0081	0,0622	0,0437	0,1089	0,0011	0,8382
Camada 6	0,2146	0,4632	0,5745	0,0081	0,0622	0,0437	0,1089	0,0011	0,8382
Camada 7	0,2146	0,4632	0,5745	0,0081	0,0622	0,0437	0,1089	0,0011	0,8382
Camada 8	0,2146	0,4632	0,5745	0,0081	0,0622	0,0437	0,1089	0,0011	0,8382
Camada 9	0,2103	0,4586	0,5532	0,0081	0,0637	0,0442	0,1185	0,0011	0,8468

Fonte: Elaborada pelos autores.

Ao se analisar a Tabela 1 evidencia-se, por meio do critério de Ivakhnenko demonstrado na equação [15], que houve uma evolução significativa na capacidade de previsão do índice argentino, à medida que foram adicionadas novas camadas à rede. Observa-se que, na primeira camada, os resultados eram considerados desinformação e, à medida que foi aumentada a complexidade dos modelos, pelo acréscimo de novas camadas, os resultados foram considerados satisfatórios. Todos os demais indicadores (R^2 ajustado,

correlação, acerto de sinais, Erro Médio Quadrático, Erro Médio Absoluto e de tendenciosidade) melhoraram significativamente, havendo um especial destaque para o coeficiente de determinação amostral que as modelagens explicaram 21,03% da variação do retorno mensal do índice da Bolsa de Valores de Buenos Aires.

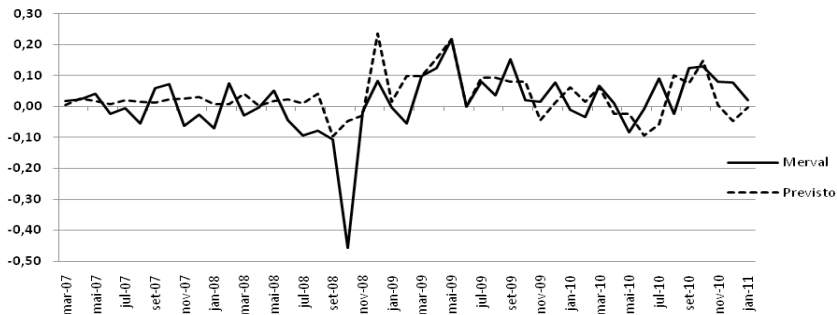


Figura 1 – Retornos observados e estimados para o índice Merval

Fonte: Elaborado pelos autores.

Na Figura 1 evidencia-se que a modelagem foi eficiente para a previsão do argentino Merval, o que corrobora com a hipótese de que os resultados foram satisfatórios para as previsões em $t+1$, fato que pode ser explicado pela maturidade do mercado de capitais argentino.

Na Tabela 2 são demonstrados os resultados das previsões do mercado brasileiro, representado pelo Ibovespa. Nas modelagens, foram testados 324.291 neurônios, sendo que 188.232 neurônios foram testados na previsão do Ibovespa.

Tabela 2 – Resultados das previsões para o Ibovespa

Camada	R^2	Correlação	Sinais	EMQ	MAE	U	U^M	U^S	Ivakhnenko
Camada 1	0,0734	0,2710	0,6595	0,0053	0,0593	0,0470	0,0490	0,0023	0,9348
Camada 2	0,0951	0,3084	0,6808	0,0080	0,0662	0,0697	0,0071	0,0021	1,4006
Camada 3	0,0765	0,2766	0,7021	0,0055	0,0612	0,0480	0,0137	0,0021	0,9703
Camada 4	0,0159	0,1262	0,6595	0,0056	0,0663	0,0489	0,0020	0,0020	0,9941
Camada 5	0,0508	0,2253	0,6595	0,0051	0,0631	0,0437	0,0025	0,0017	0,9067
Camada 6	0,2006	0,4479	0,6808	0,0043	0,0593	0,0329	0,0010	0,0008	0,7656
Camada 7	0,2021	0,4495	0,6808	0,0043	0,0591	0,0328	0,0010	0,0008	0,7638
Camada 8	0,2016	0,4490	0,6808	0,0043	0,0592	0,0328	0,0010	0,0008	0,7645
Camada 9	0,2016	0,4490	0,6808	0,0043	0,0592	0,0328	0,0010	0,0008	0,7645

Fonte: Elaborada pelos autores.

Ao se analisar os resultados das previsões em $t+1$ para o índice Ibovespa, no teste proposto por Ivakhnenko e Ivakhnenko (1995) percebe-se que houve uma evolução significativa nos resultados à medida em que foram sendo incrementadas novas camadas e todos os resultados foram considerados satisfatórios. Esse fato pode ser evidenciado na evolução de todos os demais indicadores e, em especial, é possível se destacar o R^2 que evoluiu de 0,07 (da primeira camada), para 0,20 (na última camada da rede), demonstrando a eficiência da modelagem polinomial na previsão desse índice. A correlação entre os valores observados e os valores previstos evoluiu durante a criação da rede, corroborando com a eficiência desta em prever a rentabilidade do índice brasileiro.

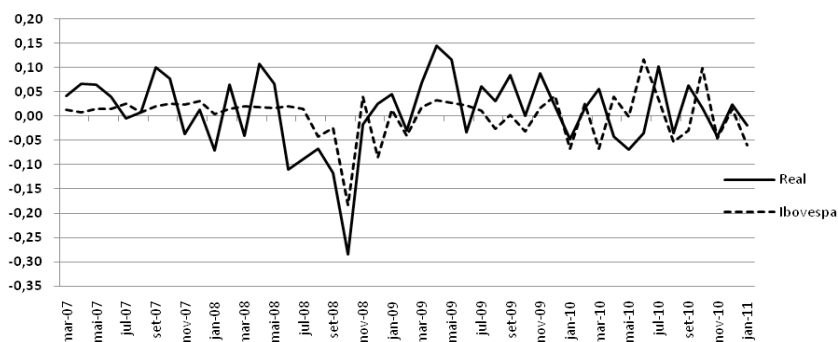


Figura 2 – Retornos observados e estimados para o índice Ibovespa

Fonte: Elaborado pelos autores.

Na Figura 2 evidencia-se que o sucesso da previsão do Ibovespa pode ser parcialmente explicado pelo acerto dos sinais (indicando queda ou aumento do retorno em $t+1$), visto que os retornos previstos variaram próximos de zero.

Tabela 3 – Resumo dos resultados obtidos na previsão dos índices Merval e Ibovespa

País	Volatilidade	R^2	Correlação	Sinais	EMQ	MAE	U'	U''	U'''	Ivakhnenko
Brasil	0,0785	0,2016	0,4490	0,6809	0,0044	0,0592	0,0329	0,0010	0,0008	0,7645
Argentina	0,1069	0,2103	0,4585	0,5532	0,0082	0,0637	0,0442	0,1185	0,0011	0,8468

Fonte: Elaborada pelos autores.

Na Tabela 3, evidencia-se que o método foi mais eficaz para encontrar não aleatoriedade para o índice brasileiro (Ibovespa), o que pode ser explicado pela menor volatilidade do retorno mensal deste índice quando comparado com o argentino Merval. A maturidade dos mercados brasileiro e argentino pode explicar, parcialmente, o grau de explicação das modelagens realizadas por meio das redes neurais polinomiais GMDH. A Figura 3 evidencia a evolução, em pontos, dos índices analisados.

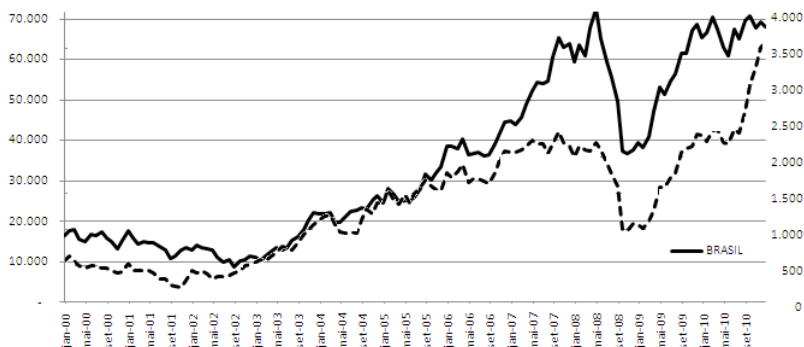


Figura 3 – Evolução, em pontos, dos índices brasileiro e argentino

Fonte: Elaborada pelos autores.

Observa-se à esquerda, da Figura 3, a pontuação do Ibovespa e à direita a pontuação do índice Merval. Nesta mesma figura é possível evidenciar que o índice brasileiro apresentou maior crescimento em pontuação, quando comparado com o argentino Merval. Apesar disso, o comportamento dos índices mostrou-se semelhante, de modo que se destaca o movimento altista de ambos. Esta tendência altista é destacada a partir de janeiro de 2003, período que coincide com a posse do novo governo federal brasileiro, com o decréscimo do risco país, assim como uma maior abertura do mercado brasileiro e de um maior esforço, por parte da Bolsa de Valores de São Paulo, em popularizá-la, gerando um expressivo crescimento na participação de pessoas físicas na Bovespa.

Nesse mesmo ano de 2003, um novo governante foi eleito na Argentina (Néstor Kirchner). A partir desse período houve uma significativa retomada no crescimento da economia do país portenho que, dois anos antes, havia passado pela sua pior crise, com o decreto de moratório da dívida externa e o chamado *corralito*: os depósitos bancários foram suspensos e os saques não podiam ser superiores a 1.000 pesos por mês. Como efeito desta crise, que se instalou na Argentina a partir de 2001, o PIB castelhano chegou ao menor valor (102,04 bilhões de dólares) em 2002, retomando um constante crescimento a partir de 2003. Esta retomada no crescimento do PIB argentino, possivelmente, influenciou o retorno do mercado de capitais deste país, levando a este movimento altista.

Este cenário, de tendência altista na pontuação dos índices analisados, especialmente após o ano de 2003, e de menor volatilidade dos retornos do índice brasileiro, poderia ajudar a explicar, parcialmente, a eficiência do método em encontrar não aleatoriedade nos dados das séries temporais.

De acordo com Harris (1995), a interpretação econômica da cointegração é que um conjunto de variáveis possuem uma relação de equilíbrio de longo prazo, então mesmo que as séries possam conter tendências estocásticas (isto é, serem não estacionárias), elas irão mover-se juntas no tempo e a diferença entre elas será estável (isto é, estacionária). Em suma, o conceito de cointegração indica a existência de um equilíbrio de longo prazo, para o qual o sistema econômico converge no tempo.

Ao observarmos certa semelhança na evolução dos pontos do índice Merval e Ibovespa (Figura 3), e considerando-se que ambos os índices apresentaram tendência altista mais proeminente a partir de 2003, período que coincide com importantes mudanças políticas e econômicas de ambos os países, procurou-se analisar a presença de raiz unitária e realizar a análise de cointegração entre as duas séries temporais.

Nesse trabalho foi utilizada a estatística denominada Dickey-Fuller Aumentado (ADF) proposto por Dickey e Fuller (1979), para testar a presença de raiz unitária na série. O teste Dickey-Fuller baseia-se no seguinte modelo na equação [21]:

$$\Delta y_t = \alpha + \beta_t + \eta y_{t-1} + \xi_t \tag{21}$$

Sendo que:

$$\eta = \sum_{j=1}^p \rho_j - 1 \tag{22}$$

Em que y denota a variável dependente e Δ denota o operador de diferença ($\Delta y_t = y_t - y_{t-1}$). Os parâmetros a serem estimados são α , β e η . As estatísticas τ_τ e τ_μ e τ apresentadas por Dickey e Fuller (1981) correspondem ao teste t para a estimativa do coeficiente da variável y_{t-1} da equação [21]. Essas estatísticas são especificadas para um modelo que inclui uma constante, uma tendência e defasagem (τ_τ), um modelo incluindo defasagem e uma constante (τ_μ) e um modelo sem constante e sem tendência (τ). As hipóteses testadas nesses modelos correspondem a uma hipótese nula de que a série não é estacionária ($H_0 : y_t$ não é $I(0)$ ou $\eta = 0$); contra a hipótese alternativa de que a série não é integrada, ou seja, trata-se de uma série estacionária ($H_1 : y_t$ é $I(0)$).

Tabela 4 – Resultados de raiz unitária do teste Dickey-Fuller Aumentado para a hipótese nula de que os índices Merval e Ibovespa não são estacionários

Total de Lags	τ_τ	τ_μ	T	Total de Lags	τ_τ	τ_μ	T
	Índice Merval				Ibovespa		
1	-2,525**	-2,194**	-2,186**	1	-2,401**	-2,095**	-2,057**
2	-2,199**	-1,926*	-1,931*	2	-2,286**	-1,987**	-1,959*
3	-1,879*	-1,561**	-1,543**	3	-2,376**	-2,037**	-1,995**
4	-1,494**	-1,240**	-1,221**	4	-2,363**	-1,968*	-1,937*
5	-1,090**	-0,843**	-0,815**	5	-2,326**	-2,011**	-1,994**

*indica que a hipótese nula é rejeitada no nível de significância de 10%.

** indica que a hipótese nula é rejeitada no nível de significância de 5%.

Fonte: Elaborada pelos autores.

Inicialmente, aplicou-se o teste de raiz unitária nos modelos do Ibovespa e do índice Merval. Verifica-se na Tabela 4 que a hipótese nula de que os índices não são estacionários não pode ser rejeitada. Desta forma realizou-se o procedimento de Engle e Granger.

Tabela 5 – Teste de Engle-Granger para a cointegração do retorno logaritmo do índice Merval com o retorno logaritmo do Ibovespa

	Estimativa	Erro Padrão	t valor	Pr(> t)
(Intercepto)	0.0065	0.008	0.806	0.422
Ibovespa	0.6902	0.103	6.727	0.000***

*** indica que a hipótese nula é rejeitada no nível de significância de 1%.

Fonte: Elaborada pelos autores.

O teste de Engle-Granger apresentou R^2 ajustado igual a 0,2511, em que o p valor do teste f apresentou-se significativo a 99% de confiança. Evidencia-se na Tabela 5 que o retorno logaritmo do índice Merval está cointegrado com o retorno logaritmo do Ibovespa. Esta informação é particularmente importante em estratégias de *hedge*. A alternativa de utilização desta cointegração entre os mercados como estratégia de *hedge* parte do pressuposto de que, apesar dos retornos dos índices apresentarem alta correlação (0,5067 e sig. 0,000 para o período analisado), este fato não é suficiente para se garantir o desempenho de longo prazo do *hedge*. Existe a necessidade de que as metodologias usuais de modelagem de risco-retorno sejam ampliadas para se levar em conta as tendências comum de longo prazo dos ativos. Isso é exatamente o que a cointegração proporciona. De acordo com Alexander (2001), a cointegração mede os movimentos conjuntos de longo prazo dos preços que podem ocorrer mesmo no decorrer de períodos em que as correlações estatísticas parecem ser baixas, as metodologias de *hedge* baseadas em ativos financeiros cointegrados devem ser mais eficientes no longo prazo. Além disso, a metodologia da cointegração não desperdiça nada da análise tradicional. Ela simplesmente amplia o modelo de correlação básico com a inclusão de um estágio preliminar, no qual os dados dos

preços multivariados são analisados e, então, estende o modelo de correlação de forma a incluir a análise dinâmica do comportamento de antecedência-defasagem dos retornos.

Alexander (2001) comenta que os índices de mercado de diferentes países podem ser cointegrados se a paridade do poder de compra se verifica. O autor demonstra que nos mercados de ações da Holanda, Alemanha e da França têm sido altamente cointegrados desde 1996, com a maior parte da causalidade, em bases diárias, vinda do mercado alemão. Neste sentido, Taylor e Tonks (1989) mostraram um elevado nível de cointegração existente entre os mercados de ações do Reino Unido, Holanda, Alemanha e Japão entre 1979 e 1986, com causalidade de Granger vinda do mercado do Reino Unido para esses outros mercados, mas não vice-versa.

De acordo com Alexander (2001), o conceito de “causalidade de Granger” independe da existência de cointegração, embora seja suficiente. Segundo esta autora, quando as séries de tempo são cointegradas, deve haver algum fluxo causal do tipo de Granger no sistema. A cointegração não é essencial para que as relações de antecedência-defasagem existam: pode ser que os fluxos causais existam entre as séries de tempo porque elas têm algumas outras características em comum. A fim de se verificar uma possível existência de um fluxo causal no retorno logaritmo dos índices Merval e Ibovespa, realizou-se o Teste de Causalidade de Granger, em bases mensais.

Tabela 6 – Teste de Causalidade de Granger entre os retornos logaritmo-mensais do índice Merval e o retorno logaritmo do Ibovespa com até três defasagens

	Uma defasagem		Duas defasagens		Três defasagens	
	Teste <i>f</i>	<i>R</i> ²	Teste <i>f</i>	<i>R</i> ²	Teste <i>f</i>	<i>R</i> ²
Ibovespa causa Granger Merval	10.12081***	0.087	5.879856***	0.100	3.93718***	0.104
Merval causa Granger Ibovespa	0.13007	0.027	0.411645	0.032	0.5675544	0.039

*** Nível de significância de 1%. ** Nível de significância de 5%.

Fonte: Elaborada pelos autores.

Na Tabela 6 evidencia-se a existência de uma relação de causalidade do retorno do Ibovespa no índice Merval, inexistindo no sentido inverso. Embora a relação de causalidade seja fraca (R^2 próximo a 10%), este fluxo causal permanece nas três defasagens realizadas. O mecanismo que vincula séries cointegradas é a “causalidade”, não no sentido de que ao ocorrer uma mudança estrutural em uma das séries, a outra também irá mudar, mas no sentido de que pontos de reversão de uma das séries precedem pontos de reversão de outra. Esta informação é particularmente importante para a alocação de ativos, na busca de alternativas para a arbitragem, bem como no gerenciamento de *hedge*, pois supõe certa ineficiência nos mercados, especialmente do mercado argentino.

Esta relação de causalidade do retorno do Ibovespa no retorno do índice Merval ocorre não necessariamente pelas relações de comércio ou dependência econômica entre os países. Ela pode ser explicada por dois argumentos vinculados à própria estrutura dos mercados de capitais dos países.

A influência da Bolsa de São Paulo sobre a Bolsa de Buenos Aires pode ser explicada, parcialmente, pelo volume de negócios diários de cada país. Em 2010, o mercado brasileiro movimentou diariamente em média US\$ 3,3 bilhões, volume anual que representa cerca de 41% do PIB. Na Argentina este volume total do ano de 2010 beirou 1% do PIB castelhano. Além disso, a Causalidade de Granger pode ser explicada pela efetiva negociação, no mercado argentino, de papéis de empresas controladas por empresas brasileiras.

A participação de dois papéis da Petrobras na composição do Ibovespa gira, atualmente, em torno de 13%, participação similar, ao peso no Merval (10%), das ações das duas controladas da estatal brasileira no mercado argentino. Além da existência destes dois papéis listados no índice Merval ocorreram, recentemente, negociações com vistas a uma maior participação do Banco do Brasil no Banco da Patagônia, crescendo a importância das empresas brasileiras no mercado acionário argentino.

Esta relação de causalidade entre os mercados que, para o período analisado mostraram-se relativamente previsíveis, apresenta melhores oportunidades de investimento no mercado acionário argentino, bem como dá fomento a discussões das estratégias de *hedge* entre os dois países e a possibilidade de arbitragem entre os mercados.

Considerações finais

O Mercado Comum do Sul (Mercosul) em sua formação original era composto por quatro países: Argentina, Brasil, Paraguai e Uruguai. Nessa região duas economias se destacam: a argentina e a brasileira. Considerados caixas de ressonância das finanças, da economia e das políticas locais, os mercados de capitais são partes fundamentais para a economia da região. No Mercosul destacam-se duas grandes bolsas de valores: a argentina Merval e a brasileira Bovespa.

O interesse em pesquisar não aleatoriedade nos retornos dos índices de países emergentes tem se mostrado cada vez mais atual, conforme cita Lim (2009). Neste trabalho procurou-se testar a eficiência das redes neurais GMDH na previsão dos índices mais representativos das bolsas de valores argentina e brasileira. Nesse sentido, a tentativa de utilização do método *Group Method of Data Handling* (GMDH) apresentou resultados significativos na previsão do índice brasileiro e do índice argentino, denotando certa ineficiência para o período analisado. Esta ineficiência é especialmente destacada no mercado argentino, em que os testes adicionais de causalidade de Granger demonstram a influência da Bolsa de Valores de São Paulo em cima da Bolsa de Valores de Buenos Aires e não havendo esta influência no sentido inverso

A principal força da modelagem GMDH pode ser expressa pela exceção na previsão do Ibovespa em $t+1$, sendo ineficaz em modelagens para períodos posteriores a $t+1$. Uma das fragilidades pode ser evidenciada na dificuldade do método em encontrar padrões em dados altamente voláteis.

A qualidade do método em encontrar padrões para a previsão do índice brasileiro pode ser explicado, fundamentalmente, pela menor volatilidade dos dados, no período analisado, quando comparado ao índice Merval. É de se esperar que, em havendo menor volatilidade nos dados, maior será a capacidade de uma rede neural encontrar padrões entre dados que aparentemente são desconexos e, desta forma, de demonstrar possíveis resultados não aleatórios nas séries.

A satisfatória previsibilidade dos índices brasileiro e argentino também pode ser especialmente explicada pela tendência altista do mercado destes países. O deslocamento da pontuação do mercado acionário brasileiro e argentino, a partir de janeiro de 2003, é contemporâneo das mudanças políticas e macroeconômicas destes dois países, destacando-se a posse de novos governantes, do decréscimo vertiginoso do risco-país brasileiro e argentino, e de uma maior abertura econômica, bem como uma nova postura da Bolsa brasileira, que promoveu ações que incentivavam uma maior participação do capital oriundo de pessoas físicas na Bolsa. Cabe, portanto, sugerir o uso destas variáveis em futuros estudos associados a esta temática.

Apesar de o índice Merval apresentar maior volatilidade que o brasileiro, mostrou-se cointegrado, em bases mensais, com o mercado brasileiro, em que existe uma relação de causalidade (embora fraca) do Ibovespa no índice Merval, revelando maior ineficiência do mercado acionário argentino. Tal fato poderia explicar, parcialmente, o sucesso da técnica em encontrar não aleatoriedade nos retornos do índice argentino.

Sugere-se ainda o uso de modelos com aprendizagem por correção de erros ou aprendizagem baseada em memória, ou ainda modelos de redes recorrentes, uma vez que essa modelagem foi feita por meio de redes alimentadas diretamente com múltiplas camadas. Para a criação de modelagens com aprendizagem por correção recomenda-se ainda a regra delta, ou regra de Widrow-Hoff. Para a modelagem com aprendizado baseado em memória, indica-se o uso de conceitos da lógica difusa na determinação de padrões.

Referências

- ACKLEY, D. H.; HINTON, G. E.; SEJNOWSKI, T. J. A learning algorithm for Boltzmann machines. *Cognitive Science*, vol. 9, p. 147-169, 1985.
- AHMADI, H.; MOTTAGHITALAB, M.; NARIMAN-ZADEH, N. Group method of data handling-type neural network prediction of broiler performance based on dietary metabolizable energy, methionine, and lysine. *Journal of Applied Poultry Research*, v. 16, p. 494-501, 2007.
- ALEXANDER, C. *Market Models: A Guide to Financial Data Analysis*. New York: John Wiley & Sons, 2001.
- BOOSE, J. H. *Personal construct theory and the transfer of human expertise*. Proceedings of AAAI-84. California: American Association for Artificial intelligence, 1984. p. 27-33.
- DICKEY, D. A; FULLER, W. A. Distribution of the Estimators for Autoregressive Time Series with a Unit Root. *Journal of the American Statistical Association*, 74, p. 427-431, 1979.
- DICKEY, D. A; FULLER, W. A. Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root. *Econometrica*, Econometric Society, vol. 49(4), p. 1.057-1.072, june, 1981.
- FERSON, W. Market Efficiency and Forecasting. In: SATCHELL, Stephen (Ed.). *Forecasting Expected Returns in the Financial Markets*. Elsevier: Academic Press, 2007. p. 1-16.
- HARRIS, R. I. D. *Using Cointegration Analysis in Econometric Modelling*. London: Prentice Hall, 1995.
- HAYKIN, S. *Neural Networks: A Comprehensive Foundation*. Second edition. Londres: Prentice Hall, 1999.
- IVAKHNENKO. A. G. *Self-teaching systems of recognition and automatic control*. Moscou: Tekhnika, 1969. p. 392.
- IVAKHNENKO. A. G. Polynomial theory of complex systems, IEEE Transaction on Systems, Man, *Cybernetics*, vol, 1, p. 364-378, october, 1971.
- IVAKHNENKO, A. G.; IVAKHNENKO, G. A.; MULLER, J. A. Self-Organization of Optimum Physical Clustering of the Data Sample for Weakened Description and Forecasting of Fuzzy Objects. *Pattern Recognition and Image Analysis*, vol, 3, no, 4, p. 415-422, 1993.

IVAKHNENKO, A. G.; IVAKHNENKO, G. A. The Review of Problems Solvable by Algorithms of the Group Method of Data Handling (GMDH). *Pattern Recognition and Image Analysis*, vol, 5, n. 4, p. 527-535, 1995.

LEGG, S.; HUTTER, M. *A Collection of Definitions of Intelligence*. Proceeding of the 2007 conference on Advances in Artificial General Intelligence: Concepts, Architectures and Algorithms: Proceedings of the AGI Workshop 2006, p. 17-24, June 7, 2007.

LIM, K. P. *An Empirical Analysis of the Weak-form Efficiency of Stock Markets*. Tese Phd Department of Econometrics and Business Statistics of Monash University, Australia, June de 2009.

LIPPMAN, R. P. *An introduction to computing neural networks*. IEEE ASSP Magazine., vol. 3-4, pp. 4-22, 1987.

MOTTAGHITALAB, M. et al. Ahmadi Predicting caloric and feed efficiency in turkeys using the group method of data handling-type neural networks. *Poult. Sci.*, June 1, 2010, 89(6): 1.325-1.331.

PINDYCK, R.; RUBINFELD, D. *Econometric Models and Economic Forecasts*. 3rd. ed. New York: McGraw-Hill, 1991.

SCHNEIDER, S.; STEINER, M. *Conditional Asset Pricing*: Predicting time varying Beta-Factors with Group Method of Data Handling Methods, SSRN-ID667468, 2006.

SILVA FILHO, J. I.; ABE, J. M. *Fundamentos das redes neurais artificiais paraconsistentes*. São Paulo: Arte & Ciência, 2000.

TSAY, R. S. *Analysis of financial time series*. New York: Ed, Wiley, 2002.

VALENÇA, M. *Aplicando redes neurais: um guia completo*. Olinda: Editora Livro Rápido: 2005.

TAYLOR, M. P.; TONKS, M. The internationalisation of stock markets and the abolition of UK Exchange control. *Review of Economics and Statistics*, v. 7, p. 332-336, 1989.

Recebido em: 8/6/2011

Accito em: 1/8/2011