



CEFAGE-UE Working Paper  
2007/05

---

## **Aplicação das Redes Neurais Artificiais à Detecção dos Mercados Euronext Mais Rentáveis**

---

*Paulo Horta<sup>a</sup>, Carlos Mendes<sup>b</sup>*

<sup>a</sup>CMVM – Comissão do Mercado de Valores Mobiliários

<sup>b</sup>UNINOVA – DEE, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa

# **Aplicação das Redes Neurais Artificiais à Detecção dos Mercados Euronext Mais Rentáveis**

**Paulo Horta\***

CMVM – Comissão do Mercado de Valores Mobiliários

**Carlos Mendes\***

UNINOVA – DEE, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa

---

\* As opiniões expressas neste artigo não reflectem necessariamente as opiniões das instituições com as quais os autores colaboram. Os autores podem ser contactados através dos seguintes endereços: [paulohorta@cmvm.pt](mailto:paulohorta@cmvm.pt) e [cmvmendes@gmail.com](mailto:cmvmendes@gmail.com).

# **Aplicação das Redes Neurais Artificiais à Detecção dos Mercados Euronext Mais Rentáveis**

## **RESUMO**

Com este estudo investiga-se a possibilidade de utilização de uma rede neuronal artificial na detecção dos mercados de acções da Euronext que proporcionam a melhor rendibilidade diária. A rede, treinada com o algoritmo de Levenberg-Marquardt, recomenda a um investidor hipotético a escolha do índice de preços representativo do mercado que se prevê que maior rendibilidade oferece no dia de negociação seguinte. Verifica-se que as recomendações da rede superam as rendibilidades dos benchmarks nos primeiros meses de 2007. O teste de Pesaran-Timmermann permite concluir que os resultados obtidos na previsão da direcção dos movimentos dos mercados Euronext (subida ou descida) não são devidos ao acaso. De igual modo, a matriz das classificações permite concluir que o desempenho do modelo na classificação diária de cada mercado em “1º melhor”, “2º melhor”, “3º melhor” ou “4º melhor”, também não é fruto do “acaso máximo” ou do “acaso proporcional”. Finalmente, como se utilizou, por um lado, uma rede do tipo feedforward com quatro neurónios na camada de output e, por outro lado, como se efectuou uma experiência de bootstrap que mostra que as elasticidades dos mercados são, em geral, estatisticamente significativas, conclui-se que os mercados Euronext estão significativamente relacionados entre si, deixando assim em aberto a possibilidade dos agentes do mercado (v.g. gestores de carteiras) potenciarem a rentabilização dos seus investimentos recorrendo a este tipo de modelos neuronais.

**Palavras-chave:** Redes neuronais, feedforward, classificação, mercados Euronext, bootstrap, avaliação de performance, estratégias de investimento

**Classificação JEL:** G11, G15

## 1. INTRODUÇÃO

As redes neuronais artificiais são modelos que surgiram originalmente na década de 1940 como tentativa de reprodução do funcionamento do cérebro humano, sendo o complexo sistema de neurónios biológicos a sua principal fonte de inspiração.

Os modelos neuronais têm tido inúmeras aplicações nas mais diversas áreas, desde as telecomunicações ao mercado imobiliário, das despesas militares ao turismo (Shachmurove, 2002; Law e Pine, 2004), desde a robótica à visão (Kröse e Smagt, 1996), das relações internacionais (Beck *et al.*, 2000) às questões de política interna (Eisinga *et al.*, 1997). Na área financeira vários problemas têm sido abordados recorrendo às redes neuronais, como a análise do risco de crédito (Nargundkar e Priestley, 2004), a previsão da insolvência de empresas (Neves e Vieira, 2004), a modelização da inflação (McNelís, 2005), a modelização das taxas de câmbio (Li *et al.*, 2004), o *rating* de obrigações, a previsão da volatilidade das opções (McNelís, 2005), a previsão das rendibilidades de acções (Thawornwong e Enke, 2004; Lawrence, 1997; Zhang *et al.*, 2002; Saad *et al.*, 1998), a previsão de índices e tendências de mercados accionistas (Walczak, 2004; Shachmurove e Witkowska, 2000; McCluskey, 1993; Pan *et al.*, 2005; Roman e Jameel, 1996), *inter alia*.

A capacidade preditiva das redes neuronais não tem passado despercebida aos agentes do mercado, prendendo a especial atenção dos gestores de carteiras, dos corretores e dos bancos de investimento. A *Goldman Sachs* e a *Morgan Stanley* têm inclusivamente departamentos dedicados à implementação destes modelos. A *Fidelity Investments* criou um fundo de investimento, em que a escolha da respectiva carteira é baseada exclusivamente nas recomendações produzidas por redes neuronais (Shachmurove, 2000, 2002). De igual forma, a *Standard & Poor's* utiliza as redes neuronais para tomar decisões relativamente à composição de uma carteira denominada "*Neural Fair Value 20*".

Com este estudo pretende-se investigar a possibilidade de utilização destes modelos no apoio à selecção dos mercados Euronext que proporcionam a melhor rendibilidade diária, no horizonte de um dia. Assim, entre os 4 principais mercados *spot*, Euronext Lisbon, Euronext Paris, Euronext Amsterdam e Euronext Brussels, uma rede neuronal treinada com o algoritmo de Levenberg-Marquardt apoiará um investidor hipotético na escolha do índice de mercado que a própria rede prevê que maior rendibilidade oferece no

dia de negociação seguinte. Uma das medidas utilizadas para aferir a qualidade das escolhas sugeridas pela rede é a comparação entre um investimento hipotético que segue o *output* da rede, com outros investimentos alternativos que seguem uma estratégia de *buy and hold* nos vários mercados da Euronext N.V.. A rede será considerada útil se permitir ao investidor a obtenção de mais-valias superiores às dos investimentos alternativos (*benchmarks*).

A estratégia do investidor hipotético é a seguinte: no dia 3/Jan/2007, após o fecho do Nasdaq, a rede faz uma previsão relativamente ao índice de mercado que maior rentabilidade deverá proporcionar no dia 4/Jan/2007. Supondo que o investidor consegue tomar uma posição longa no dia 3/Jan/2007, ao preço de fecho desse índice, ele aplica a totalidade do dinheiro disponível para investimento exclusivamente na compra desse índice. Esta posição é mantida durante o dia de negociação seguinte, e o investidor auferirá a rentabilidade diária proporcionada pelo mercado que a rede sugeriu, independentemente deste mercado ser ou não de facto o mais rentável. No final da sessão do Nasdaq do dia seguinte (4/Jan/2007), a rede efectua uma nova previsão para o mercado Euronext mais rentável do dia de negociação seguinte (5/Jan/2007). Se essa previsão coincidir com a previsão anterior, o investidor mantém a sua carteira inalterada, caso contrário vende toda a sua carteira (ao preço de fecho do dia 4/Jan/2007) e investe de imediato todo o dinheiro disponível na compra do novo índice sugerido pela rede, auferindo a rentabilidade proporcionada por esse índice durante o dia 5/Jan/2007. Esta estratégia é mantida sistematicamente até ao dia 16/Abril/2007, altura em que o investidor vende a sua carteira e avalia os resultados obtidos.

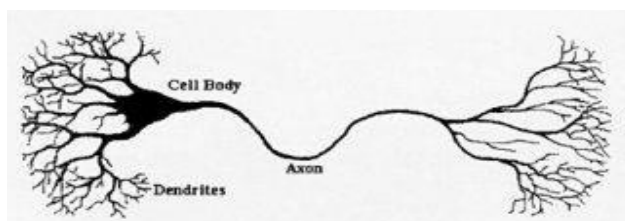
O estudo está organizado do seguinte modo. Na secção 2 apresenta-se um breve enquadramento teórico das redes neuronais, mais concretamente das redes que são utilizadas neste estudo empírico: *feedforward neural networks*. Na secção 3 discute-se a arquitectura da rede e os dados utilizados. Na secção 4 apresentam-se os resultados e avalia-se a qualidade do modelo. Finalmente, na secção 5, apresentam-se as principais conclusões.

## 2. REDES NEURONAIS ARTIFICIAIS

### 2.1. Conceito

Uma rede neuronal artificial consiste essencialmente num conjunto de unidades de processamento simples (neurónios) que comunicam entre si enviando sinais através de um número elevado de conexões (Kröse e Smagt, 1996). Em termos biológicos, se a informação acumulada no núcleo central de um determinado neurónio atingir um certo limite, o neurónio “dispara”, transmitindo um sinal electroquímico a um neurónio adjacente, através de um canal emissor denominado de axónio. A extremidade do axónio é composta por ramificações (as sinapses) que por sua vez estão ligadas à estrutura do neurónio receptor através de outras ramificações denominadas de dentrites. O cérebro humano comporta aproximadamente 10 mil milhões de neurónios, sendo que cada um está ligado a cerca de 10 000 neurónios vizinhos<sup>1</sup>.

Gráfico 1. Representação do neurónio biológico (Reed *et al.*, 1999)



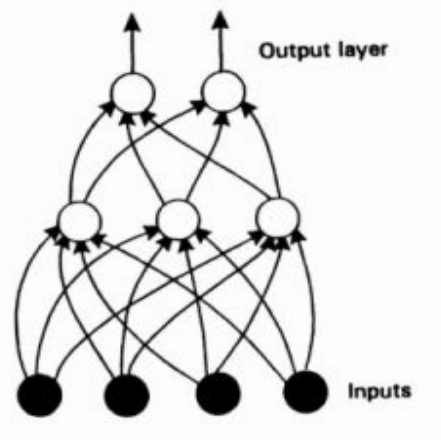
No gráfico, as ramificações do lado direito representam as sinapses. As redes neuronais artificiais não são compostas por tantos neurónios, mas aqueles que as compõem seguem uma estrutura que se assemelha de alguma forma à dos homólogos biológicos.

Existem diversos tipos de redes, sendo que cada tipo é especializado numa tarefa concreta. McNelis (2005), por exemplo, utiliza os modelos neuronais para a resolução de tarefas de previsão ou ajustamento de funções, tarefas de classificação, e tarefas de redução de variáveis. Para as tarefas de previsão, as redes mais utilizadas designam-se na terminologia anglo-saxónica por *feedforward neural networks*. Estas redes são treinadas para produzirem um *output* com o menor desvio possível em relação ao verdadeiro valor das séries a prever. É este tipo de rede que utilizamos neste estudo.

---

<sup>1</sup> Stanford University's Computer Science Education, em <http://www-cse.stanford.edu/classes/sophomore-college/projects-00/neural-networks/Biology/index.html>.

Gráfico 2. Exemplo de uma rede do tipo feedforward neural network (Gurney, 1997)



As redes do tipo *feedforward* são constituídas por camadas (*layers*) de neurónios<sup>2</sup>: uma camada de *input*, uma ou várias camadas ocultas<sup>3</sup> (*hidden*) e uma camada de *output*. Como não existe processamento de informação na camada de *input*, devido à ausência de uma função de activação, alguns autores optam por não considerar os elementos desta camada como neurónios. Kröse e Smagt (1996) chamam às células de *input*, unidades “fan-out”.

Dentro de cada camada existem vários neurónios trabalhando em paralelo que recebem informação da camada anterior e enviam informação apenas para os neurónios da camada subsequente. Nestas redes não há ligação entre os neurónios pertencentes à mesma camada.

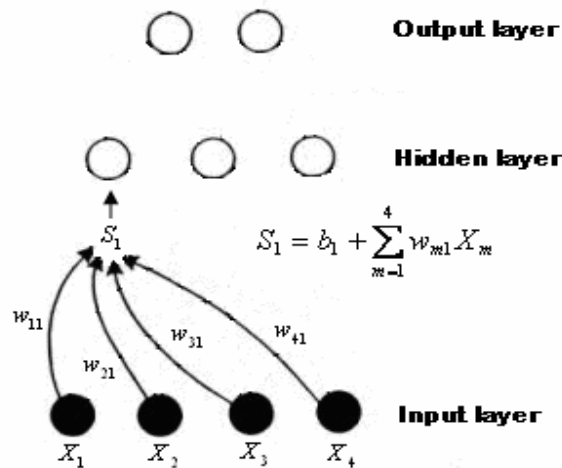
A figura anterior exemplifica isso mesmo: a camada de *input* corresponde aos 4 círculos pretos (cada círculo preto representa uma variável de *input*,  $X_m, m = 1, \dots, 4$ ). A camada intermédia é a única camada oculta desta rede, e é composta por 3 neurónios; a última camada é a camada de *output*, composta por 2 neurónios. Como se pode verificar, as células de uma determinada camada estão ligadas apenas às células da camada subsequente. As setas, para além de ilustrarem as ligações entre os neurónios, mostram ainda que a informação é transmitida apenas no sentido ascendente (*feedforward*).

<sup>2</sup> Os neurónios são representados na figura pelos círculos brancos. Os círculos pretos representam as células de *input*.

<sup>3</sup> Apesar da terminologia, convém referir que nada está oculto: as formas funcionais das células das camadas ocultas são conhecidas, tal como as das restantes camadas. Na figura, a camada intermédia é a única camada oculta desta rede.

A informação contida na camada de *input* é agrupada (ou somada) de acordo com uma função de propagação específica, e enviada (propagada) como *input* para as células da camada subsequente. A função de propagação mais utilizada é a aditiva<sup>4</sup> e tem a forma  $S_h = b_h + \sum_m w_{mh} X_m$ , em que  $X_m$  é a variável de *input*  $m$ ,  $w_{mh}$  é o ponderador ou força de conexão sináptica entre a variável de *input*  $m$  e o neurónio  $h$  da camada subsequente<sup>5</sup>,  $b_h$  é um parâmetro de enviesamento (*bias* ou *threshold*) associado ao neurónio  $h$ , e  $S_h$  é a informação propagada da camada de *input* para o neurónio  $h$  da camada subsequente.

Gráfico 3. Propagação da informação da camada de input para a camada oculta



É a partir dos neurónios das camadas ocultas que a informação começa a ser verdadeiramente processada. Estes neurónios são dos mais importantes na rede na medida em que são eles que em primeira instância extraem a informação da camada de *input* e identificam padrões nessa informação, que subsequentemente são transmitidos às camadas superiores. Após receber a informação  $S_h$ , cada neurónio da primeira camada oculta é activado através de uma função de activação, produzindo um determinado sinal. O conjunto dos sinais de todos os neurónios desta camada é depois enviado para a camada subsequente

<sup>4</sup> Feldman e Ballard (1982) utilizaram uma outra função de propagação conhecida como “unidade sigma-pi”.

<sup>5</sup> As forças de conexão sinápticas são caracterizadas como “estimuladoras” se assumirem valores positivos, e “inibidoras” se assumirem valores negativos (Kröse e Smagt, 1996).



através da função de propagação. Este processo repete-se de camada em camada até que a camada de *output* seja atingida.

A função de activação encontra paralelo no mundo biológico da seguinte forma: cada neurónio biológico vai acumulando informação no núcleo (*cell body*); quando a quantidade de informação acumulada atinge um determinado limite (*bias* ou *threshold*)<sup>6</sup>, o neurónio é activado e emite um impulso electroquímico para os neurónios vizinhos. Assim, o estado de cada neurónio pode ser representado por uma variável binária: 0 se não é activado, ou 1 se é activado (não há estados intermédios no neurónio biológico)<sup>7</sup>.

O homólogo artificial pode assumir estados intermédios, nomeadamente por uma questão de facilidade no tratamento matemático. Se for possível considerar uma função de activação que não assuma apenas 0 ou 1, mas que possa assumir, por exemplo, qualquer valor entre 0 e 1, e que seja diferenciável em todo o domínio, o tratamento matemático sai beneficiado.

A literatura faz referência às mais diversas funções de activação, mas a mais utilizada para os neurónios ocultos, para tarefas de ajustamento de funções, é a função logística (ou *sigmoid*), que tem a forma:  $F(x) = \frac{1}{1 + e^{-x}}$ . Como se pode verificar, esta função é diferenciável em  $R$ , e além disso produz valores entre 0 e 1.

Deste modo, se um determinado neurónio artificial for associado a uma função de activação logística, podemos facilmente medir o grau de intensidade de activação do neurónio. Se os valores produzidos pela função de activação forem próximos de 0 dizemos que o neurónio é activado com pouco vigor, se forem próximos de 1 dizemos que é activado de forma vigorosa. Quando um neurónio é activado de forma vigorosa pode significar que detectou um padrão importante na informação de *input*, sendo esse padrão enviado de seguida para a camada subsequente através da função de propagação.

As redes do tipo *feedforward* são ainda caracterizadas pelo facto de os neurónios de uma determinada camada serem representados pela mesma função de activação. Entre

---

<sup>6</sup> Este limite é representado por "*b*" no caso dos neurónios artificiais. O neurónio "dispara" se a informação somada for superior a  $b$ :  $\sum w_{mh} X_m \geq b_h$ . Esta situação é válida apenas para as funções de activação binárias. Caso contrário, o neurónio "dispara" sempre, com maior ou menor intensidade.

<sup>7</sup> Stanford University's Computer Science Education, em <http://www-cse.stanford.edu/classes/sophomore-college/projects-00/neural-networks/Biology/index.html>.

neurónios de camadas distintas podem existir funções de activação diferentes. No caso de previsão de variáveis financeiras, é comum utilizar-se uma função de activação linear (ou função identidade:  $F(x) = x$ ) na camada de *output*, precisamente para não restringir os valores previstos pela rede ao intervalo (0,1), tal como aconteceria se fosse utilizada uma função de activação logística. Se quisermos prever uma variável dicotómica (por exemplo, as subidas ou as descidas do preço de uma acção) então uma função de activação logística poderá ser adequada à modelização da camada de *output*, indicando grosso modo a probabilidade de subida do preço da acção.

Em relação à rede representada no gráfico 2, se utilizarmos o índice “(2)” para fazer referência às variáveis que ligam a camada 2 à camada seguinte, se considerarmos que os 3 neurónios da camada oculta têm função de activação logística, se chamarmos  $Y_1$  e  $Y_2$  aos neurónios de *output*, e se considerarmos que estes são representados por uma função de activação linear, então a forma reduzida da rede pode ser matematicamente representada por um sistema de tantas equações quantos os neurónios de *output* existentes:

$$\begin{cases} Y_1 = b_1^{(2)} + w_{11}^{(2)} \frac{1}{1 + e^{-\frac{1}{m}(b_1 + \sum w_{m1} X_m)}} + w_{21}^{(2)} \frac{1}{1 + e^{-\frac{1}{m}(b_2 + \sum w_{m2} X_m)}} + w_{31}^{(2)} \frac{1}{1 + e^{-\frac{1}{m}(b_3 + \sum w_{m3} X_m)}} \\ Y_2 = b_2^{(2)} + w_{12}^{(2)} \frac{1}{1 + e^{-\frac{1}{m}(b_1 + \sum w_{m1} X_m)}} + w_{22}^{(2)} \frac{1}{1 + e^{-\frac{1}{m}(b_2 + \sum w_{m2} X_m)}} + w_{32}^{(2)} \frac{1}{1 + e^{-\frac{1}{m}(b_3 + \sum w_{m3} X_m)}} \end{cases}$$

Neste sistema visualiza-se facilmente, para as equações de  $Y_1$  e  $Y_2$ , os 3 neurónios ocultos dispostos em paralelo. Repare-se ainda que os neurónios de *output* utilizam exactamente os mesmos neurónios das camadas precedentes. O que distingue os neurónios  $Y_1$  e  $Y_2$  são as forças de conexão sinápticas e os parâmetros de enviesamento referenciadas com o índice “(2)”, o que significa que estes neurónios de *output* reagem da mesma forma a grande parte da informação de *input*, seguindo caminhos distintos apenas na fase terminal da rede, isto é, quando se passa da última camada oculta para a camada de *output*. Isto quer dizer que os neurónios de *output* de uma rede deste tipo estão fortemente interligados.

A forma de representação analítica permite verificar que uma rede não é mais do que uma função composta por várias funções de propagação e de activação intercaladas:

$$Y = \text{activação}(\text{propagação}(\text{activação}(\text{propagação}(X))))).$$

## 2.2. Treino

Após definida a arquitectura da rede, a aplicação a um problema concreto passa a centrar-se na estimação dos diversos parâmetros (os “ $w$ ” e os “ $b$ ”). O processo de determinação destes parâmetros designa-se por treino, sendo o algoritmo de *backpropagation* um dos mais utilizados nas redes de tipo *feedforward*. Walkzak (2004), referenciando diversos autores escreve o seguinte: “*The backpropagation learning algorithm has been shown to be able to model arbitrary associations (White, 1990) and to be a universal approximator (Hornik, 1991; Hornik et al., 1989). While various other neural network training algorithms exist, the backpropagation algorithm is the most widely used (Fu, 1994; Walczak e Cerpa, 1999; Widrow et al., 1994).*”

A estratégia subjacente ao algoritmo consiste em encontrar os valores para os parâmetros que permitam a minimização de uma determinada função de erro<sup>8</sup>, através de um método denominado de *gradient descent*. A ideia passa por efectuar alterações aos parâmetros de forma proporcional ao simétrico da derivada da função de erro  $E_t$ , em ordem

a cada parâmetro:  $\Delta_t w_m = -\theta \frac{\partial E_t}{\partial w_m}$ , em que  $\theta$  é uma constante de proporcionalidade,

conhecida como taxa de aprendizagem. A taxa de aprendizagem define a magnitude das alterações a efectuar aos parâmetros. Se  $\theta$  for pequeno, a aprendizagem da rede é lenta, e a convergência para uma solução óptima é demorada. Se  $\theta$  for demasiado grande, pode correr-se o risco do algoritmo oscilar, produzindo eventualmente erros ainda maiores. Hagan *et al.* (1996) discutem a escolha da taxa de aprendizagem adequada.

Existem diversas variantes ao método de *backpropagation* “puro”. O algoritmo de Levenberg-Marquardt, por exemplo, utiliza um parâmetro  $\mu$ , conhecido como *damping parameter*, que permite conjugar a *backpropagation* com o método de Newton. Quando  $\mu$  tende para zero, o algoritmo tende para o método de Newton; quando  $\mu$  é elevado, o

---

<sup>8</sup> Utilizámos a média do quadrado dos erros:  $E = \frac{1}{N} \sum_t (T_t - Y_t)^2$ , em que  $E$  representa o erro global produzido pela rede,  $T_t$  representa o *target* ou valor real para a observação  $t$ , relativamente à qual a rede pretende aproximar-se o mais possível,  $Y_t$  representa o *output* produzido pela rede, e  $N$  o número de observações.

algoritmo utiliza a *backpropagation*. Como o método de Newton é mais rápido e preciso a convergir na vizinhança de um erro mínimo, o objectivo do algoritmo de Levenberg-Marquardt passa por aproximar-se progressivamente em direcção ao método de Newton (*ex vi* diminuição de  $\mu$ ) sempre que o valor da função de erro diminua em cada iteração. Caso se preveja que a função de erro tende a aumentar, então aumenta-se  $\mu$ , fazendo com que o algoritmo de Levenberg-Marquardt se transforme progressivamente na *backpropagation*. Desta forma, o valor da função de erro diminui sempre após cada iteração (Demuth *et al.*, 2006). O resultado prático é uma velocidade de convergência mais elevada. Uma aplicação deste algoritmo ao treino dos modelos neuronais pode ser encontrada em Hagan e Menhaj (1994). Demuth *et al.* (2006) referem que este algoritmo aparenta ser o método de treino mais rápido para redes neuronais *feedforward* de dimensão moderada, isto é, redes cuja arquitectura pode incluir até várias centenas de forças de conexão sinápticas.

O número de parâmetros ( $P$ ) a estimar depende do número variáveis de *input*, do número de camadas e do número de neurónios de cada camada da rede. Se  $w_{mh}$  for o termo geral de uma matriz com  $m'$  linhas e  $h'$  colunas, em que  $m'$  é o número de variáveis de *input* e  $h'$  o número de neurónios ocultos (no gráfico 2,  $m \times h' = 4 \times 3$ );  $b_h$  o termo geral de um vector  $h \times 1$ ;  $w_{hn}^{(2)}$  o termo geral de uma matriz  $h \times n'$ , em que  $n'$  é o número de neurónios de *output* (no gráfico 2,  $n' = 2$ ); e finalmente  $b_n^{(2)}$  um vector  $n \times 1$ , então o número de parâmetros a estimar será:  $P = m \times h' + h \times 1 + h \times n' + n \times 1$ . No caso da rede do gráfico 2, seriam 23 os parâmetros a estimar.

### 2.3. Algumas vantagens e desvantagens dos modelos neuronais

A utilização das redes neuronais para previsão de séries financeiras pode justificar-se pelas vantagens que estes modelos trazem relativamente a outros métodos de previsão. Shachmurove (2002), por exemplo, enumera algumas vantagens face aos modelos económicos tradicionais. Uma das vantagens referidas é que estes modelos conseguem captar com precisão padrões complexos existentes na informação. A este propósito, Maasoumi *et al.* (1994) mencionam que as séries temporais são dinâmicas por natureza, sendo necessário a existência de ferramentas não lineares para captar padrões

comportamentais existentes no interior das séries. Uma outra vantagem, talvez das mais relevantes, é que não é necessário elaborar hipóteses relativamente à natureza da distribuição dos dados. Em vez disso, estes modelos utilizam os próprios dados para produzir representações internas das relações entre as variáveis. Como consequência, é de esperar a obtenção de melhores resultados com a aplicação das redes neuronais quando as relações entre as variáveis não seguem um comportamento pré-definido.

Relativamente às desvantagens, Shachmurove (2002) refere que não existe uma metodologia estruturada que permita efectuar a melhor escolha relativamente à arquitectura da rede, ao treino da rede ou à verificação da qualidade da mesma. Por exemplo, o número de camadas a utilizar na rede ou o número de neurónios que cada camada deve ter são normalmente aspectos *inter alia* com opiniões divergentes. Thawornwong e Enke (2004) para concluírem que as redes neuronais permitem obter melhores resultados que outros modelos na previsão das rendibilidades de acções, compilaram numa tabela as técnicas de modelização neuronais encontradas em 45 estudos diferentes. Verifica-se que raramente há consenso em relação à escolha dos diversos aspectos, o que permite concluir que a escolha da melhor arquitectura para a rede continua a depender em grande medida da sensibilidade e das experiências realizadas por cada investigador.

Uma outra crítica que se faz amiúde aos modelos neuronais é a crítica da “caixa negra” (*black box*), segundo a qual não é possível perceber como são estimadas as relações nos neurónios ocultos. Eisinga *et al.* (1997) mostram que esta crítica é um pouco exagerada ao conseguirem, por um lado, desenhar uma rede que lhes permitiu controlar o *timing* de activação dos neurónios ocultos e, por outro lado, efectuar inferência relativamente ao impacto que as variáveis independentes provocam nas dependentes. Não obstante, enquanto que nos modelos econométricos lineares a avaliação da relevância das variáveis independentes, e do impacto que estas provocam na dependente, é trivialmente avaliada recorrendo às estatísticas *t*, nos modelos neuronais essa tarefa afigura-se mais complexa.

### 3. ARQUITECTURA DA REDE E DADOS

#### 3.1. Primeira experiência

Uma vez que a utilização dos modelos neuronais não requer necessariamente que se efectuem hipóteses relativamente à distribuição dos dados (Gorr, 1994; Shachmurove, 2002), a nossa primeira experiência foi no sentido de tentar tirar partido dessa característica, utilizando as variáveis em nível (valores de fecho) e não em rendibilidades ou diferenças, como é habitual proceder com variáveis financeiras. A rede neuronal foi treinada com o algoritmo de Levenberg-Marquardt, com um parâmetro de *damping*  $\mu = 0,001$ , com um parâmetro de incremento de  $\mu$  igual a  $inc^+ \mu = 10$ , um parâmetro de decremento igual a  $inc^- \mu = 0,1$ , e uma taxa de aprendizagem  $\theta = 0,01$ . A rede é do tipo *feedforward*, com uma tipologia 7-15-4, com função de propagação aditiva, função de activação logística nos neurónios da camada oculta, e função de activação linear nos neurónios de *output*.<sup>9</sup>

A opção pela escolha do algoritmo de Levenberg-Marquardt justifica-se pelo facto deste método de treino convergir rapidamente para uma solução óptima, quando aplicado a redes de dimensão moderada, como é o caso da que utilizamos neste estudo. A escolha de 7 células de *input* e 4 neurónios de *output* está relacionada com a existência de 7 variáveis independentes e 4 dependentes neste estudo. Em relação à camada oculta, a escolha de 15 neurónios resulta das experiências efectuadas. Com efeito, verificou-se que o valor da função de erro resultante do treino da rede assumia em média valores menores quando eram utilizados 15 neurónios ocultos, *ceteris paribus*. A mesma justificação sustenta a escolha das funções de activação. Quanto aos restantes parâmetros, optou-se por considerar os valores por defeito existentes no *software* de treino, visto que as experiências efectuadas não permitiram melhorar, em média, o valor da função de erro.

As variáveis foram previamente normalizadas para média nula e desvio-padrão unitário, para reduzir o espaço de pesquisa do algoritmo de treino, tornando a rede mais eficiente.

---

<sup>9</sup> Uma tipologia 7-15-4 significa que a rede tem 3 camadas, e que na 1ª camada existem 7 células, na 2ª existem 15 neurónios, e na 3ª existem 4 neurónios.

Como o objectivo é prever qual dos 4 mercados Euronext vai subir mais no dia de negociação seguinte, as variáveis dependentes utilizadas foram os valores de fecho dos índices PSI20 (Lisboa), CAC40 (Paris), AEX (Amesterdão) e BEL20 (Bruxelas).

As variáveis de *input* utilizadas foram 7: os valores de fecho dos índices desfasados 1 dia: PSI20(-1), CAC40(-1), AEX(-1) e BEL20(-1)<sup>10</sup>; o valor da taxa de juro Euribor a 3 meses desfasado 1 dia, EUR(-1); o valor de fecho do índice petrolífero XOI, convertido em euros, desfasado 1 dia, XOI(-1); e o valor de fecho do índice Nasdaq100 desfasado 1 dia, NAS(-1).

A justificação para a utilização destas variáveis é a seguinte: os 4 índices desfasados são utilizados para tentar captar a autocorrelação existente nas séries. Os preços de fecho dos índices num determinado dia dependem em grande medida do fecho do dia anterior, não se afastando muito entre si em dias consecutivos. Os investidores tendem a utilizar a informação de fecho dos índices para elaborar as suas expectativas relativamente à evolução dos mercados no dia seguinte. Espera-se que a rede reflecta esta situação.

No mesmo sentido, encontra-se justificação para a utilização do índice Nasdaq desfasado. Como o mercado norte-americano fecha às 21H00 (hora de Lisboa), existe um período de cerca de 4,5 horas, entre o fecho dos mercados Euronext e o fecho do Nasdaq, em que a informação relevante para os investidores não se encontra incorporada nos mercados Euronext.

A utilização da taxa de juro Euribor e do índice de preços do petróleo tem um propósito mais de natureza “fundamental”. Em termos teóricos é de esperar que uma diminuição nas taxas de juro provoque um aumento nos preços das acções. Por um lado, se as taxas de juros diminuem, os instrumentos de dívida podem tornar-se menos atractivos para os investidores, fazendo eventualmente com que a procura de acções aumente, aumentando assim o seu preço. Paralelamente, se as taxas de juro diminuem os custos das empresas também diminuem, o que pode levar a um aumento do lucro e à consequente valorização das suas acções, principalmente se as empresas recorrerem em boa medida ao financiamento por capital alheio.

---

<sup>10</sup> Apesar de estas séries não serem estacionárias, as redes neuronais são capazes de lidar com esta situação sem necessidade prévia de transformação das variáveis.

Em relação ao preço do petróleo a justificação está igualmente relacionada com os custos das empresas. Uma vez que esta matéria-prima é indispensável à actividade das empresas, se o seu preço aumentar, os custos destas também aumentam, fazendo com que as expectativas de lucros futuros diminuam, diminuindo consequentemente o interesse dos investidores pelas acções, o que levará a uma eventual redução dos seus preços.

Os dados utilizados neste estudo foram retirados da *Bloomberg*. O *software* utilizado para o treino da rede foi o MATLAB. A amostra de treino começa no início do ano 1999, data em que dispomos de valores para todas as variáveis, e termina no final do ano 2006 (2017 observações diárias<sup>11</sup>). A amostra para previsão, que nunca foi apresentada ao *software* durante a fase de treino, começa no início de 2007 e termina no dia 16/Abr/2007 (72 observações).

Um procedimento que adoptámos, que não é o mais usual neste tipo de estudos empíricos é a utilização de uma rede neuronal com mais que um neurónio de *output*. Este procedimento tem a vantagem de evidenciar a interligação dos mercados, se o ajustamento do modelo for satisfatório.

### 3.2. Segunda experiência

A 2ª experiência realizada foi no sentido de aplicar a mesma arquitectura às variáveis referidas anteriormente, mas desta vez convertidas em rendibilidades. Assim, a rendibilidade do PSI20, por exemplo, foi definida como  $\frac{PSI20 - PSI20(-1)}{PSI20(-1)}$ . As restantes variáveis tiveram um tratamento semelhante. Com estas transformações as variáveis ficaram com média próxima de zero e desvio-padrão reduzido, pelo que não foi necessário normalizá-las para melhorar a eficiência da rede.

Apesar da 1ª experiência não ter produzido resultados tão bons como os da 2ª, foi no entanto muito útil, nomeadamente para interpretar os resultados do modelo recorrendo à análise gráfica.

---

<sup>11</sup> A amostra de treino foi dividida de modo a obter-se um vector de dados para validação, com o objectivo de evitar o *overfitting*, permitindo que o modelo adquirisse uma boa capacidade de generalização (*i.e.*, bons resultados *out of sample*).



#### 4. RESULTADOS

A avaliação da qualidade das redes, a sua capacidade preditiva, é normalmente efectuada recorrendo a indicadores como a raiz do erro quadrático médio (RMSE), o erro percentual médio em módulo (MAPE), ou o erro absoluto médio (MAE)<sup>12</sup>, por exemplo. Vários autores notaram que quando se cometem erros, mesmo que pequenos, que vão na direcção oposta à do movimento dos mercados, a capacidade preditiva da rede pode sair prejudicada caso se utilizem estes indicadores. Por esta razão, sugerem que a proporção de previsões correctas da direcção do movimento dos índices deve ser o critério primordial para avaliar a qualidade previsional dos modelos (Walczak, 2001, 2004; Thawornwong e Enke, 2004; Tahai *et al.*, 1998; Green e Pearson, 1994).

Pesaran e Timmermann (1992) construíram uma estatística que permite testar se a direcção dos movimentos efectivos dos mercados e as previsões desses movimentos dadas pelo modelo estão relacionadas. A hipótese nula do teste sugere que os movimentos reais e os movimentos dados pelo modelo são independentes, o que significa que o modelo não é adequado para previsão. Vejamos sinteticamente como é construída a estatística. As variáveis  $T_t$  e  $Y_t$  representam, respectivamente, o *target* e do *output* (em rendibilidades) no momento  $t$ . A variável  $P$  representa aqui a percentagem de previsões correctas na direcção do mercado:

$$P = \frac{1}{N} \sum_{t=1}^N D_t, \text{ em que } D_t = \begin{cases} 1, \text{ se } T_t \cdot Y_t > 0 \\ 0, \text{ caso contrário} \end{cases}, \text{ e } N \text{ representa a dimensão da amostra}$$

utilizada para previsão *out of sample*. Quando  $P = 0$  significa que não é efectuada qualquer previsão correcta na direcção do mercado;  $P = 1$  significa que todas as previsões efectuadas acertam na direcção em que o mercado se move. É ainda necessário construir as variáveis seguintes:

$$A_t = \begin{cases} 1, \text{ se } T_t > 0 \\ 0, \text{ caso contrário} \end{cases}$$

$$B_t = \begin{cases} 1, \text{ se } Y_t > 0 \\ 0, \text{ caso contrário} \end{cases}$$

---

<sup>12</sup> Na terminologia anglo-saxónica: *Root Mean Square Error*, *Mean Absolute Percentage Error* e *Mean Absolute Error*, respectivamente.

$$P_A = \frac{1}{N} \sum_{t=1}^N A_t ; P_B = \frac{1}{N} \sum_{t=1}^N B_t ; P_* = P_A P_B + (1 - P_A)(1 - P_B)$$

A estatística de Pesaran-Timmermann,  $PT$ , definida como:  $PT = \frac{P - P_*}{[V(P) - V(P_*)]^{1/2}} \stackrel{a}{\sim} N(0,1)$

segue assintoticamente uma distribuição normal *standardizada* sob a hipótese nula de que  $T_t$  e  $Y_t$  são variáveis aleatórias independentes.

As variáveis  $V(P)$  e  $V(P_*)$  representam as variâncias de  $P$  e  $P_*$ , respectivamente, e definem-se do seguinte modo:  $V(P) = N^{-1} P_*(1 - P_*)$ ,

$$V(P_*) = N^{-1} (2P_B - 1)^2 P_A (1 - P_A) + N^{-1} (2P_A - 1)^2 P_B (1 - P_B) + 4N^{-2} P_B P_A (1 - P_B)(1 - P_A).$$

Outro indicador importante para avaliar a qualidade da rede é o resultado da comparação entre as rendibilidades que um investidor hipotético consegue obter se seguir os conselhos da rede, e a rendibilidade de um investimento de referência ou *benchmark*.

Considerámos 5 *benchmarks*, que representam os investimentos efectuados entre 2/Jan/2007 e 16/Abr/2007 em 5 dos principais índices dos mercados Euronext. Além da rendibilidade absoluta apresenta-se ainda o índice de *Sharpe*, que leva em conta não só a rendibilidade dos investimentos - ou melhor, o excesso de rendibilidade relativamente ao investimento num activo sem risco -, como ainda o risco ou volatilidade associado a esses investimentos<sup>13</sup>. O índice de *Sharpe* permite hierarquizar por ordem de preferência os diversos investimentos: quanto maior for o índice melhor é o investimento.

Quadro 1. Benchmarks

<b>Benchmark</b>	<b>PSI20</b>	<b>CAC40</b>	<b>AEX</b>	<b>BEL20</b>	<b>NXT100</b>
<b>Rendibilidade</b>	7,37%	4,35%	6,21%	3,61%	5,69%
<b>Índice de Sharpe</b>	2,00	0,79	1,24	0,59	1,19

Faz-se ainda referência a um outro indicador: a matriz de classificações, prevista em Reis (2001). Esta matriz permite comparar a classificação original, isto é, o desempenho (relativo) real dos 4 mercados Euronext (qual foi efectivamente o 1º melhor mercado, o 2º, o 3º e o 4º), com a classificação prevista pela rede. A matriz, de termo geral  $n_{ij}$ , exhibe em linha a classificação prevista pela rede, e em coluna a classificação real. Por exemplo, o

<sup>13</sup> Considerámos a Euribor a 3 meses como *proxy* para a rendibilidade do activo sem risco. O risco foi medido com o desvio-padrão das rendibilidades.

elemento  $n_{11}$  da matriz das classificações do quadro 2 mostra que em 38% das vezes a rede acertou no mercado Euronext com a maior rendibilidade diária; o elemento  $n_{41}$  significa que em 19% das vezes que a rede previu que determinado mercado seria o 4º melhor (i.e., o mercado com menor rendibilidade), na realidade esse mercado foi o 1º melhor. Trata-se de um erro de previsão da rede. Assim, os elementos da matriz fora da diagonal principal mostram os insucessos da rede, enquanto que os elementos da diagonal principal (a sombreado) mostram a percentagem de previsões correctas efectuadas pela rede.

Reis (2001) sugere que para considerar aceitável a percentagem de casos correctamente classificados (PCC), deve fazer-se uma comparação com os resultados de uma classificação totalmente feita ao acaso. Como existem 4 grupos (4 mercados) com igual dimensão (72 observações<sup>14</sup>), a percentagem de casos correctamente classificados de forma aleatória seria de 25%, de acordo com os critérios do acaso máximo e do acaso proporcional (Reis, 2001). Assim, se a média dos elementos da diagonal principal (PCC) for superior a 25% podemos considerar aceitável a qualidade da rede, de acordo com este indicador.

Quadro 2. Resultados da 1ª experiência<sup>15</sup>

<i>in sample</i>	<b>PSI20</b>	<b>CAC40</b>	<b>AEX</b>	<b>BEL20</b>
<b><math>R^2</math></b>	99,84%	99,70%	99,78%	99,76%
<b>RMSE (raiz do erro quadrático médio)</b>	84,87	54,74	5,73	28,88
<b>MAPE (média do erro percentual em módulo)</b>	0,65%	0,94%	0,94%	0,78%
<i>out sample</i>	<b>PSI20</b>	<b>CAC40</b>	<b>AEX</b>	<b>BEL20</b>
<b>RMSE (raiz do erro quadrático médio)</b>	97,49	76,73	6,00	51,25
<b>MAPE (média do erro percentual em módulo)</b>	0,63%	1,08%	0,94%	0,93%
<b>Acertos na direcção do mercado</b>	57,5%	60,3%	58,9%	49,3%
<b>Estatística de Pesaran-Timmermann</b>	0,50	1,62 ***	1,33 ***	-0,28
<b>Rendibilidade seguindo os conselhos da rede</b>	11,63%			
<b>Índice de Sharpe</b>	2,67			

<sup>14</sup> 73 no caso da rede da 2ª experiência. Como na 1ª experiência se obtêm previsões dos preços de fecho, perdeu-se uma observação para permitir o apuramento posterior das rendibilidades.

<sup>15</sup> \* significa que a hipótese nula é rejeitada com 1% de significância; \*\* significa que a hipótese nula é rejeitada com uma significância de 5%; \*\*\* significa que a hipótese nula é rejeitada com 10% de significância.

		Mercado real			
		1º	2º	3º	4º
Previsão da rede	1º	38%	21%	21%	21%
	2º	21%	24%	26%	29%
	3º	22%	25%	32%	21%
	4º	19%	31%	21%	29%

**Matriz das classificações**  
(PCC = 31%)

Como podemos observar no quadro, apesar da estatística de Pesaran-Timmermann indicar que apenas os mercados francês e holandês apresentam resultados significativos (com uma probabilidade de 90%), o índice de *Sharpe* sugere que a estratégia proposta pela rede é preferível a qualquer um dos *benchmarks*. Com efeito, um investidor que apostasse passivamente no PSI20, que foi o melhor mercado no período em análise (com uma rentabilidade de 7,37%), obteria ainda assim uma mais-valia inferior à da estratégia proposta pela rede (rentabilidade de 11,63%). Verifica-se assim que esta rede supera qualquer um dos investimentos de referência.

Como se pode verificar pela matriz das classificações, a rede acertou 38% das vezes no melhor mercado Euronext, percentagem substancialmente superior aos 25% que resultariam de uma escolha puramente aleatória. Durante o período de previsão, que engloba 72 dias de negociação, o mercado português foi o que ofereceu maior rentabilidade diária em 29 ocasiões, tendo a rede previsto correctamente 15 dessas ocasiões. O mercado holandês foi o melhor mercado por 17 vezes (a rede acertou em 6). O mercado francês foi o melhor em 12 ocasiões (a rede acertou em 3), e o mercado belga foi o melhor em 14 (a rede acertou em 3).

Convém referir que estes resultados foram obtidos ignorando os custos de transacção em mercado. Do ponto de vista dos intermediários financeiros que são membros dos mercados Euronext, esta hipótese não retira validade à análise, na medida em que estes custos se traduzem num valor fixo por negócio, de acordo com os vários pacotes de “trading fees” disponibilizados pelo grupo Euronext aos membros dos seus mercados. Por esta razão, quanto maior forem os volumes transaccionados pelos intermediários financeiros, mais diluídos se tornam os custos de transacção, tornando-se mesmo negligenciáveis para montantes de negociação elevados. Do ponto de vista do pequeno investidor, que não tem acesso directo ao mercado, esta situação deixa de ter validade. Nestes casos, como os intermediários financeiros cobram normalmente aos seus clientes

quantias variáveis, que dependem dos volumes negociados, os custos de transacção passam a ser relevantes e anulam mesmo as mais-valias identificadas.

Para avaliar o impacto das variáveis independentes sobre as dependentes aplicamos o método das diferenças finitas, abordado por McNelis (2005). Em geral, como a forma funcional das redes pode ser relativamente complexa, dada a sua não linearidade, o cálculo das derivadas analíticas pode revelar-se também uma tarefa complexa. O método das diferenças finitas, por ser de aplicação simples e universal, facilita este tipo de análise.

Representando a forma funcional da rede por  $Y_j = f_j(X_1, \dots, X_i, \dots, X_m)$ , em que  $j = 1, \dots, 4$ ;  $i = 1, \dots, m$ ; e  $m = 7$ , o impacto da variável independente  $X_i$  no *output*  $Y_j$  pode ser aproximado através da expressão seguinte:

$$\frac{\partial Y_j}{\partial X_i} = \frac{f_j(X_1, \dots, X_i + \varepsilon, \dots, X_m) - f_j(X_1, \dots, X_i, \dots, X_m)}{\varepsilon},$$

em que  $\varepsilon = 10^{-6}$  representa o acréscimo infinitesimal. Como as variáveis  $X_i$  assumem escalas diferentes, a comparação dos seus impactos nas variáveis dependentes pode não ser de fácil percepção. A normalização desses impactos, que facilita a comparação e a visualização gráfica, pode ser conseguida através do cálculo das elasticidades, de acordo com a expressão seguinte:

$$Elasticidade_{Y_j/X_i} = \frac{[f_j(X_1, \dots, X_i + \varepsilon, \dots, X_m) - f_j(X_1, \dots, X_i, \dots, X_m)]X_i}{\varepsilon \cdot f_j(X_1, \dots, X_i, \dots, X_m)}$$

Para medir a significância estatística das elasticidades recorreremos à técnica de *bootstrap* utilizada por McNelis (2005), e detalhada por Efron e Tibshirani (1993).

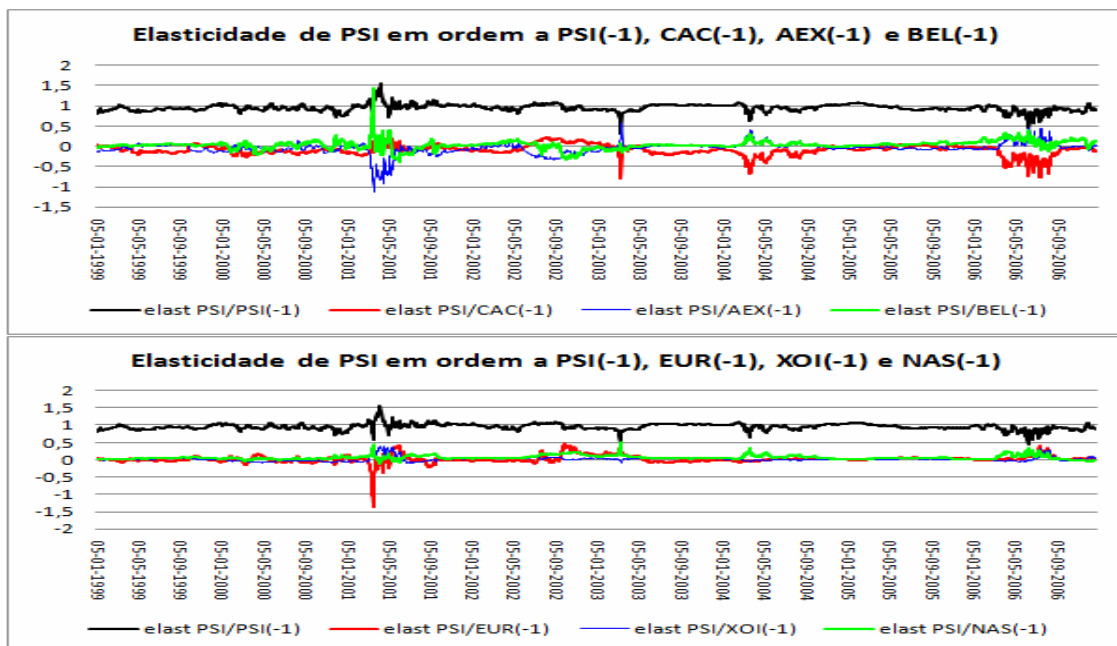
A partir dos dados originais dos *inputs* e dos *targets*,  $[T, X]$ , obtém-se um valor  $Y$  para o *output* através do treino da rede neuronal, e apura-se o vector de erros,  $\delta = T - Y$ . A partir deste vector efectua-se aleatoriamente uma tiragem com reposição dos seus elementos, construindo-se assim o primeiro vector de choques,  $\delta^{(1)}$ , para utilizar na primeira experiência de *bootstrap*. Este vector permite construir um novo vector de *targets*,  $T^{b1} = Y + \delta^{b1}$ . De seguida, as variáveis  $[T^{b1}, X]$  são usadas para voltar a estimar a rede e obter as elasticidades através do método das diferenças finitas. Este procedimento é

repetido 500 vezes (até  $T^{b500}$  e  $\delta^{b500}$ ), obtendo-se assim 500 valores para as elasticidades<sup>16</sup>. De seguida os 500 valores são ordenados por ordem crescente de forma a construir a função de probabilidade das elasticidades. A partir desta função podem calcular-se os *p-values* para cada uma das elasticidades, considerando como hipótese nula que as elasticidades são iguais a zero, isto é, *H0*: as variáveis independentes não são relevantes para explicar as dependentes.

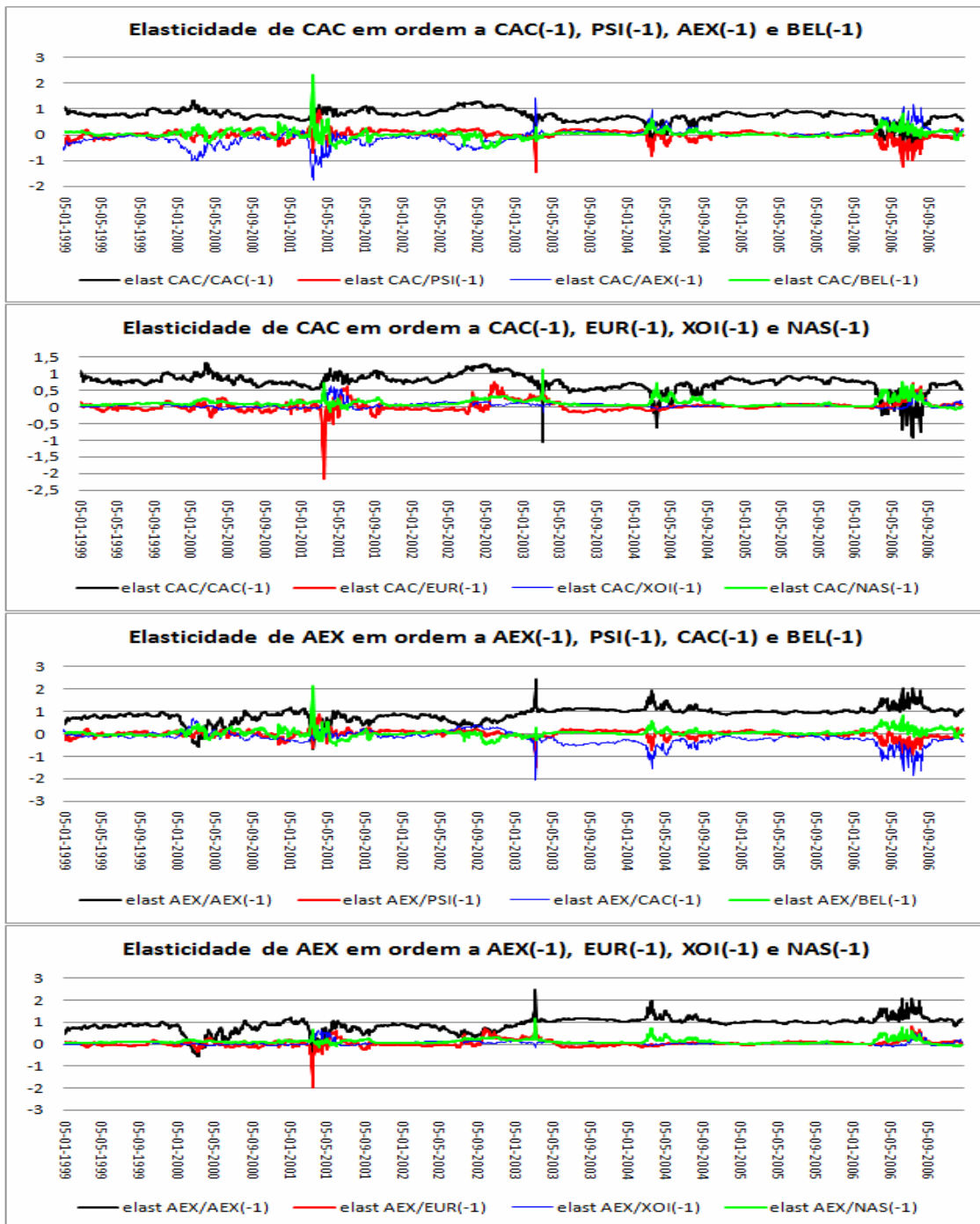
Em relação à fiabilidade deste método de *bootstrap*, McNelis (2005) refere que apesar de moroso e pesado em termos computacionais, é um método fiável, dando em geral a garantia de obtenção de resultados significativos.

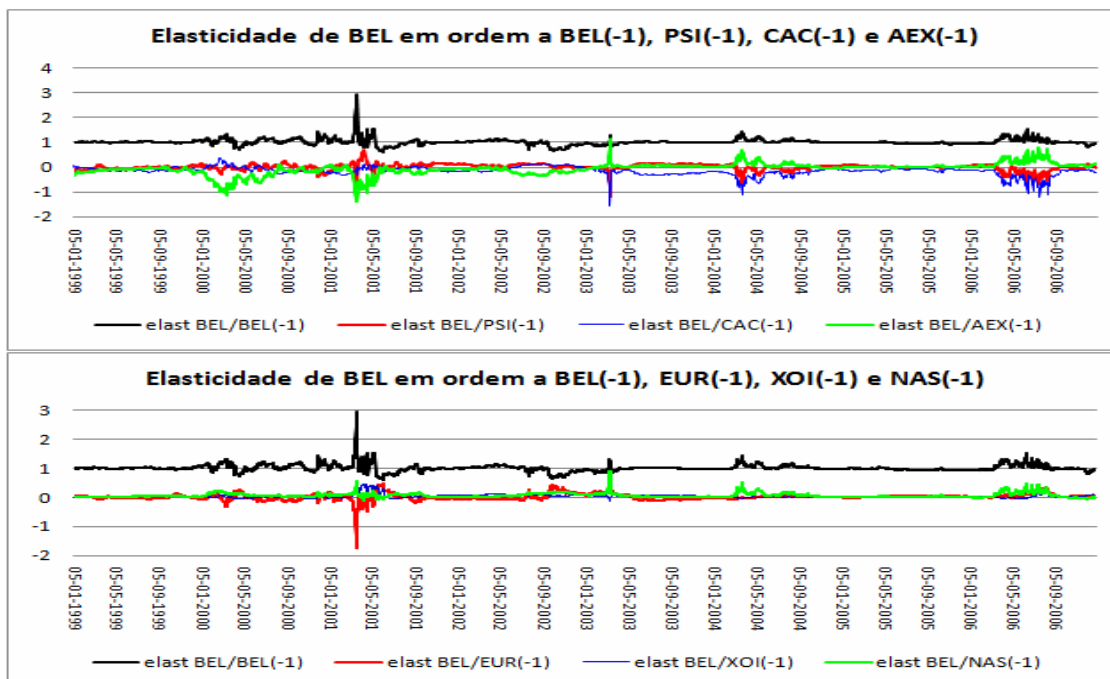
O gráfico seguinte mostra os valores das elasticidades das 4 variáveis dependentes em ordem a cada uma das 7 variáveis independentes. Optou-se por construir 2 gráficos para cada variável dependente para não prejudicar excessivamente o efeito visual, uma vez que a maior parte das elasticidades assume valores relativamente próximos.

Gráfico 4. Elasticidades



<sup>16</sup> 500 valores para cada observação amostral de cada variável de *input*, e para cada *target*, o que perfaz mais de 28 milhões de observações (500 x 2016 x 7 x 4).





Da análise dos gráficos três situações principais podem ser identificadas. Em primeiro lugar, tal como se esperava, o principal factor explicativo para um determinado índice de mercado é o próprio índice desfasado um dia. Esta situação pode ser confirmada em cada um dos diversos gráficos: a elasticidade de um índice em ordem ao índice desfasado é a que assume o maior valor, aproximando-se de “1”, ao passo que as elasticidades das restantes variáveis não se afastam normalmente muito de “0”. Em segundo lugar, existem determinados momentos em que as outras variáveis independentes contribuem para explicar o índice, sobrepondo-se inclusivamente à capacidade explicativa do índice desfasado. Nos gráficos estas situações podem ser visualizadas nos pontos em que, simultaneamente, a elasticidade do índice em ordem ao próprio índice desfasado se afasta de “1” e tende para “0”, e a elasticidades do índice em ordem às outras variáveis se afasta de “0”. Em terceiro lugar, e ao contrário do inicialmente esperado, as variáveis XOI(-1) e EUR(-1), o preço do petróleo e a taxa de juro, nem sempre influenciam de forma negativa o valor dos índices, numa base diária. Resta saber se os valores das elasticidades têm ou não significado estatístico. A figura seguinte mostra os *p-values* obtidos na experiência de *bootstrap*. Os *p-values* correspondem a testes unilaterais: quando os valores das elasticidades dados pela rede são positivos, o *p-value* representa a probabilidade no lado esquerdo da distribuição; de forma análoga, quando os valores das



elasticidades são negativos, o *p-value* reflecte a massa de probabilidade existente no lado direito da distribuição.

Quadro 3. Elasticidades do PSI20

Data	Variáveis independentes						
	PSI(-1)	CAC(-1)	AEX(-1)	BEL(-1)	EUR(-1)	XOI(-1)	NAS(-1)
05-01-1999	<b>0,889</b>	-0,015	<b>-0,102</b>	-0,015	0,005	-0,017	0,013
19-10-1999	<b>0,878</b>	0,065	-0,125	0,014	<b>0,108</b>	0,013	<b>0,070</b>
09-08-2000	<b>0,905</b>	-0,199	0,103	0,001	0,057	-0,042	<b>0,059</b>
04-06-2001	<b>1,057</b>	0,072	<b>-0,209</b>	-0,329	0,391	0,170	-0,044
25-03-2002	<b>0,962</b>	-0,153	0,021	<b>0,103</b>	-0,053	0,044	<b>0,040</b>
16-01-2003	<b>0,919</b>	0,023	-0,085	-0,004	<b>0,112</b>	0,043	<b>0,088</b>
28-10-2003	<b>1,036</b>	-0,131	-0,010	0,031	-0,042	0,024	<b>0,036</b>
10-08-2004	<b>0,859</b>	-0,269	0,115	<b>0,118</b>	0,021	-0,022	<b>0,138</b>
19-05-2005	<b>0,990</b>	-0,005	-0,064	0,042	0,033	0,002	0,042
24-02-2006	<b>0,990</b>	<b>-0,026</b>	-0,074	<b>0,162</b>	-0,017	-0,015	0,018
06-12-2006	<b>1,008</b>	<b>-0,035</b>	-0,069	0,014	0,054	0,067	-0,048
29-12-2006	<b>0,893</b>	<b>-0,128</b>	0,034	0,149	0,034	0,054	-0,013

Data	p-values (Hipótese nula: elasticidade = 0)						
	PSI(-1)	CAC(-1)	AEX(-1)	BEL(-1)	EUR(-1)	XOI(-1)	NAS(-1)
05-01-1999	<b>0,000</b>	0,542	<b>0,070</b>	0,626	0,440	0,470	0,252
19-10-1999	<b>0,000</b>	0,534	0,108	0,518	<b>0,030</b>	0,492	<b>0,060</b>
09-08-2000	<b>0,000</b>	0,114	0,888	0,140	0,460	0,130	<b>0,058</b>
04-06-2001	<b>0,000</b>	0,410	<b>0,060</b>	0,370	0,280	0,712	0,850
25-03-2002	<b>0,000</b>	0,218	0,626	<b>0,068</b>	0,372	0,284	<b>0,078</b>
16-01-2003	<b>0,000</b>	0,574	0,282	0,374	<b>0,020</b>	0,372	<b>0,004</b>
28-10-2003	<b>0,000</b>	0,136	0,490	0,452	0,452	0,418	<b>0,036</b>
10-08-2004	<b>0,000</b>	0,156	0,442	<b>0,078</b>	0,454	0,348	<b>0,020</b>
19-05-2005	<b>0,000</b>	0,208	0,332	0,166	0,438	0,410	0,106
24-02-2006	<b>0,000</b>	<b>0,072</b>	0,698	<b>0,036</b>	0,546	0,330	0,292
06-12-2006	<b>0,000</b>	<b>0,070</b>	0,776	0,138	0,496	0,478	0,752
29-12-2006	<b>0,000</b>	<b>0,088</b>	0,240	0,164	0,542	0,534	0,706

Quadro 4. Elasticidades do CAC40

Data	Variáveis independentes						
	PSI(-1)	CAC(-1)	AEX(-1)	BEL(-1)	EUR(-1)	XOI(-1)	NAS(-1)
05-01-1999	-0,105	<b>0,927</b>	<b>-0,493</b>	0,107	0,027	0,012	0,046
19-10-1999	-0,131	<b>1,028</b>	<b>-0,222</b>	-0,062	<b>0,209</b>	0,072	<b>0,124</b>
09-08-2000	-0,007	<b>0,611</b>	0,121	-0,012	0,086	-0,008	<b>0,112</b>
04-06-2001	<b>0,084</b>	<b>0,909</b>	<b>-0,158</b>	-0,457	0,591	0,278	-0,049
25-03-2002	0,086	<b>0,758</b>	-0,054	0,101	-0,099	<b>0,134</b>	<b>0,054</b>
16-01-2003	-0,012	<b>0,886</b>	-0,093	-0,033	<b>0,170</b>	<b>0,129</b>	<b>0,137</b>
28-10-2003	<b>0,175</b>	<b>0,594</b>	0,066	0,043	-0,081	<b>0,089</b>	<b>0,039</b>
10-08-2004	-0,274	<b>0,336</b>	0,297	0,226	0,056	0,010	<b>0,274</b>
19-05-2005	0,017	<b>0,874</b>	-0,047	0,006	0,064	<b>0,056</b>	<b>0,057</b>
24-02-2006	0,140	<b>0,732</b>	0,046	0,063	-0,012	0,021	0,015
06-12-2006	0,157	<b>0,746</b>	0,006	-0,120	0,116	0,152	-0,074
29-12-2006	-0,043	<b>0,515</b>	0,241	0,130	0,091	0,130	-0,004

Data	p-values (Hipótese nula: elasticidade = 0)						
	PSI(-1)	CAC(-1)	AEX(-1)	BEL(-1)	EUR(-1)	XOI(-1)	NAS(-1)
05-01-1999	0,164	<b>0,000</b>	<b>0,026</b>	0,250	0,394	0,234	0,150
19-10-1999	0,578	<b>0,000</b>	<b>0,082</b>	0,314	<b>0,008</b>	0,196	<b>0,030</b>
09-08-2000	0,422	<b>0,000</b>	0,974	0,850	0,380	0,604	<b>0,006</b>
04-06-2001	<b>0,034</b>	<b>0,000</b>	<b>0,070</b>	0,402	0,222	0,442	0,928
25-03-2002	0,166	<b>0,000</b>	0,228	0,146	0,482	<b>0,026</b>	<b>0,030</b>
16-01-2003	0,776	<b>0,000</b>	0,386	0,336	<b>0,010</b>	<b>0,076</b>	<b>0,000</b>
28-10-2003	<b>0,092</b>	<b>0,000</b>	0,240	0,524	0,460	<b>0,068</b>	<b>0,012</b>
10-08-2004	0,468	<b>0,004</b>	0,228	0,136	0,398	0,170	<b>0,010</b>
19-05-2005	0,262	<b>0,000</b>	0,588	0,494	0,378	<b>0,026</b>	<b>0,056</b>
24-02-2006	0,128	<b>0,000</b>	0,138	0,146	0,702	0,314	0,174
06-12-2006	0,328	<b>0,006</b>	0,144	0,742	0,346	0,242	0,820
29-12-2006	0,618	<b>0,008</b>	0,166	0,248	0,410	0,320	0,788

Quadro 5. Elasticidades do AEX

Data	Variáveis independentes						
	PSI(-1)	CAC(-1)	AEX(-1)	BEL(-1)	EUR(-1)	XOI(-1)	NAS(-1)
05-01-1999	-0,048	0,084	<b>0,468</b>	0,037	0,011	0,007	0,038
19-10-1999	-0,130	0,058	<b>0,801</b>	-0,057	<b>0,168</b>	0,063	<b>0,114</b>
09-08-2000	-0,013	-0,267	<b>1,008</b>	-0,001	0,061	-0,006	<b>0,114</b>
04-06-2001	0,060	-0,027	<b>0,816</b>	-0,442	0,555	0,272	-0,051
25-03-2002	0,102	-0,106	<b>0,841</b>	0,069	-0,098	<b>0,116</b>	<b>0,049</b>
16-01-2003	-0,055	-0,068	<b>0,898</b>	-0,002	<b>0,163</b>	0,119	<b>0,128</b>
28-10-2003	0,159	<b>-0,411</b>	<b>1,085</b>	0,051	-0,092	0,086	<b>0,033</b>
10-08-2004	-0,269	<b>-0,666</b>	<b>1,284</b>	0,231	0,045	0,002	<b>0,278</b>
19-05-2005	0,005	<b>-0,109</b>	<b>0,924</b>	0,058	0,062	<b>0,045</b>	<b>0,043</b>
24-02-2006	0,100	<b>-0,256</b>	<b>1,008</b>	0,192	-0,028	0,000	-0,008
06-12-2006	0,159	<b>-0,143</b>	<b>0,868</b>	-0,066	0,122	0,154	-0,111
29-12-2006	-0,072	<b>-0,403</b>	<b>1,128</b>	0,239	0,086	0,126	-0,038

Data	p-values (Hipótese nula: elasticidade = 0)						
	PSI(-1)	CAC(-1)	AEX(-1)	BEL(-1)	EUR(-1)	XOI(-1)	NAS(-1)
05-01-1999	0,164	0,402	<b>0,004</b>	0,390	0,456	0,256	0,184
19-10-1999	0,498	0,626	<b>0,000</b>	0,358	<b>0,030</b>	0,226	<b>0,056</b>
09-08-2000	0,302	0,218	<b>0,000</b>	0,850	0,470	0,564	<b>0,016</b>
04-06-2001	0,114	0,468	<b>0,000</b>	0,398	0,252	0,472	0,922
25-03-2002	0,144	0,224	<b>0,000</b>	0,190	0,438	<b>0,070</b>	<b>0,042</b>
16-01-2003	0,682	0,136	<b>0,000</b>	0,336	<b>0,012</b>	0,130	<b>0,000</b>
28-10-2003	0,144	<b>0,018</b>	<b>0,000</b>	0,572	0,446	0,110	<b>0,034</b>
10-08-2004	0,448	<b>0,060</b>	<b>0,000</b>	0,184	0,444	0,272	<b>0,030</b>
19-05-2005	0,278	<b>0,080</b>	<b>0,000</b>	0,440	0,430	<b>0,074</b>	<b>0,092</b>
24-02-2006	0,158	<b>0,024</b>	<b>0,000</b>	0,124	0,572	0,550	0,738
06-12-2006	0,350	<b>0,028</b>	<b>0,000</b>	0,768	0,394	0,306	0,762
29-12-2006	0,652	<b>0,040</b>	<b>0,000</b>	0,236	0,492	0,414	0,720

Quadro 6. Elasticidades do BEL20

Data	Variáveis independentes						
	PSI(-1)	CAC(-1)	AEX(-1)	BEL(-1)	EUR(-1)	XOI(-1)	NAS(-1)
05-01-1999	-0,064	0,027	<b>-0,253</b>	<b>0,974</b>	0,021	0,009	0,016
19-10-1999	-0,069	0,026	<b>-0,153</b>	<b>0,966</b>	<b>0,134</b>	0,049	<b>0,064</b>
09-08-2000	0,019	-0,257	0,067	<b>0,973</b>	0,059	0,005	<b>0,078</b>
04-06-2001	<b>0,052</b>	-0,013	<b>-0,148</b>	<b>0,649</b>	0,447	0,224	-0,042
25-03-2002	0,082	-0,177	-0,024	<b>1,065</b>	-0,077	<b>0,100</b>	<b>0,029</b>
16-01-2003	0,019	-0,075	-0,053	<b>0,961</b>	<b>0,092</b>	<b>0,090</b>	<b>0,079</b>
28-10-2003	0,137	<b>-0,251</b>	0,044	<b>1,002</b>	-0,058	<b>0,062</b>	<b>0,015</b>
10-08-2004	-0,188	<b>-0,434</b>	0,212	<b>1,150</b>	0,043	0,006	<b>0,178</b>
19-05-2005	0,024	<b>-0,081</b>	-0,016	<b>0,991</b>	0,039	<b>0,034</b>	<b>0,029</b>
24-02-2006	0,100	<b>-0,126</b>	0,036	<b>0,959</b>	0,000	0,014	0,006
06-12-2006	0,092	<b>-0,111</b>	0,022	<b>0,856</b>	0,072	0,081	-0,034
29-12-2006	-0,022	<b>-0,228</b>	0,153	<b>0,970</b>	0,060	0,068	0,008

Data	p-values (Hipótese nula: elasticidade = 0)						
	PSI(-1)	CAC(-1)	AEX(-1)	BEL(-1)	EUR(-1)	XOI(-1)	NAS(-1)
05-01-1999	0,160	0,446	<b>0,010</b>	<b>0,000</b>	0,384	0,204	0,194
19-10-1999	0,616	0,714	<b>0,068</b>	<b>0,000</b>	<b>0,018</b>	0,194	<b>0,056</b>
09-08-2000	0,570	0,178	0,974	<b>0,000</b>	0,440	0,314	<b>0,020</b>
04-06-2001	<b>0,086</b>	0,434	<b>0,074</b>	<b>0,000</b>	0,248	0,406	0,886
25-03-2002	0,188	0,110	0,258	<b>0,000</b>	0,472	<b>0,036</b>	<b>0,070</b>
16-01-2003	0,286	0,158	0,368	<b>0,000</b>	<b>0,022</b>	<b>0,078</b>	<b>0,000</b>
28-10-2003	0,140	<b>0,012</b>	0,306	<b>0,000</b>	0,418	<b>0,064</b>	<b>0,022</b>
10-08-2004	0,440	<b>0,032</b>	0,228	<b>0,000</b>	0,400	0,180	<b>0,014</b>
19-05-2005	0,206	<b>0,026</b>	0,632	<b>0,000</b>	0,336	<b>0,046</b>	<b>0,078</b>
24-02-2006	0,108	<b>0,026</b>	0,146	<b>0,000</b>	0,748	0,426	0,232
06-12-2006	0,354	<b>0,022</b>	0,102	<b>0,000</b>	0,326	0,320	0,794
29-12-2006	0,600	<b>0,032</b>	0,114	<b>0,000</b>	0,388	0,360	0,216

As células pintadas correspondem a *p-values* inferiores a 10%, mostrando os pontos em que as elasticidades são estatisticamente significativas. A escolha dos pontos (datas) das tabelas foi efectuada de forma a manter uma distância constante (200 dias de negociação) entre os pontos, para cobrir uniformemente os dados da amostra. Além disso, incluiu-se o primeiro e o último ponto da amostra. As conclusões que se retiram de seguida são generalizáveis a toda a amostra.

Em primeiro lugar, o mercado Nasdaq parece ter tido uma maior importância no comportamento dos mercados Euronext no passado do que tem hoje em dia. Esta situação pode eventualmente estar relacionada com o refrear do entusiasmo que se criou em relação às expectativas de evolução dos preços das acções das empresas do sector da “Nova Economia” (principais constituintes do Nasdaq). Os mercados Euronext, fortemente influenciados por este entusiasmo, atingiram máximos históricos no ano 2000, tendo iniciado subsequentemente trajectórias de correcção que seguiram os movimentos do Nasdaq. Nos últimos anos, contudo, pode ter-se assistido ao encerramento deste ciclo de entusiasmo. Como consequência, a informação que os investidores retiram actualmente do Nasdaq poderá não ser tão relevante para os seus investimentos, numa base diária, como o foi no passado.

Em segundo lugar, ainda que com intensidades distintas, todas as variáveis parecem ser relevantes para explicar o comportamento dos mercados Euronext, pelo menos em alguns momentos da amostra. Verifica-se que quando o impacto da taxa de juro e do preço do petróleo é significativo, o sinal desse impacto é por vezes oposto ao inicialmente esperado, contrariando assim a relação teórica entre estas variáveis, no prazo de um dia.

Em terceiro lugar, os mercados Euronext parecem estar relacionados entre si. Os resultados sugerem que o mercado francês é o que mais influencia os restantes, principalmente num período mais recente. Os mercados menos influentes são em geral o português e o belga. Este sistema de relações é consentâneo com a dimensão relativa dos vários mercados: os de maior dimensão são os que mais impacto provocam nos restantes e, simultaneamente, os que menos se deixam influenciar. As estatísticas da Euronext<sup>17</sup>

---

<sup>17</sup> Volume transaccionado em mercado secundário, associado ao livro central de ofertas. [www.euronext.com](http://www.euronext.com).

relativamente ao ano de 2006 mostram que em termos de volume transaccionado, o mercado francês é de longe o mais relevante, com um volume de cerca de 1578 mil milhões de euros, o que representa uma quota de 65%. Em segundo plano aparece o mercado holandês, com uma quota de 29% e, num patamar inferior, surgem os mercados belga e português, com quotas de 4% e 2%, respectivamente.

Em relação às consequências para as estratégias de investimento, os resultados sugerem que o mercado francês tem um impacto negativo nos restantes mercados Euronext, *ceteris paribus*. Esta constatação não garante, contudo, que se o mercado francês valorizar (desvalorizar) num determinado dia, os investidores possam esperar sempre uma desvalorização (valorização) nos outros mercados Euronext no dia de negociação seguinte. As elasticidades são parciais e não totais, e por isso os movimentos finais dos mercados dependem também do comportamento das restantes variáveis.

Quando as variáveis são convertidas em rendibilidades as previsões obtidas pelo modelo podem ser melhoradas, tal como refere Gately (1996). O quadro seguinte mostra os resultados obtidos na previsão com as variáveis estacionárias.

Quadro 7. Resultados da 2ª experiência<sup>15</sup>

<i>in sample</i>	<b>PSI20</b>	<b>CAC40</b>	<b>AEX</b>	<b>BEL20</b>
<i>R<sup>2</sup></i>	6,6%	10,1%	9,8%	7,3%
<i>RMSE (raiz do erro quadrático médio)</i>	0,92%	1,34%	1,42%	1,13%
<i>MAE (erro médio absoluto)</i>	0,67%	0,98%	0,99%	0,80%

<i>out sample</i>	<b>PSI20</b>	<b>CAC40</b>	<b>AEX</b>	<b>BEL20</b>
<i>RMSE (raiz do erro quadrático médio)</i>	0,66%	0,84%	0,84%	0,87%
<i>MAE (erro médio absoluto)</i>	0,66%	0,86%	0,84%	0,88%
<i>Acertos na direcção do mercado</i>	73,2%	62,0%	62,0%	60,6%
<i>Estatística de Pesaran-Timmermann</i>	3,59 *	1,94 **	1,88 **	2,11 **

<i>Rendibilidade seguindo os conselhos da rede</i>	16,66%
<i>Índice de Sharpe</i>	3,73

		<b>Mercado real</b>			
		<b>1°</b>	<b>2°</b>	<b>3°</b>	<b>4°</b>
<b>Previsão da rede</b>	<b>1°</b>	36%	25%	27%	12%
	<b>2°</b>	21%	25%	25%	30%
	<b>3°</b>	33%	18%	32%	18%
	<b>4°</b>	11%	33%	16%	40%

*Matriz das classificações*  
(PCC = 33%)

Pode verificar-se que os resultados melhoraram com a 2ª experiência: o teste de Pesaran-Timmermann permitiu obter resultados estatisticamente significativos para os 4 mercados, a rendibilidade e o índice de *Sharpe* melhoraram, e a percentagem de casos correctamente classificados também aumentou, de 31% para 33%.

Esta rede permite obter uma rendibilidade 2,26 vezes superior à do mercado Euronext com melhor *performance* no período de previsão.

A figura seguinte mostra as forças de conexão sinápticas e os parâmetros de enviesamento associados à rede, que resultaram da aplicação do algoritmo de treino de Levenberg-Marquardt.

Quadro 8. Forças de conexão sinápticas e enviesamentos obtidos na 2ª experiência

$$W_{mn}^{(1)} = \begin{bmatrix} 10,7782 & 29,5347 & 41,7535 & 2,9362 & -31,5948 & 49,5683 & -27,4058 & -18,3768 & 19,5107 & 24,1883 & -30,0251 & 42,9992 & -44,6159 & -20,7774 & -35,0681 \\ 35,7599 & 22,9187 & 2,6636 & -5,1394 & 15,9197 & 20,9411 & 27,1836 & -7,1858 & 36,7925 & 1,7643 & 0,7041 & 39,9902 & -16,6784 & -38,8482 & -15,2093 \\ 3,2241 & -3,1323 & -28,0509 & -23,2311 & -9,662 & 15,3755 & -18,2799 & -0,6163 & 10,1334 & 27,723 & 12,1504 & -19,2694 & -23,8208 & -8,3034 & -0,3823 \\ 11,5788 & -13,5554 & 37,1825 & 27,576 & -29,2215 & -20,8089 & 4,9734 & 7,0536 & 26,0581 & -40,0162 & -5,3925 & -6,282 & 18,5437 & 27,0063 & 17,1655 \\ 2,0386 & -12,0408 & -0,7783 & -15,3596 & -3,5931 & 13,1055 & -28,2057 & 20,4493 & -19,2698 & 4,4367 & -12,3977 & -1,5783 & -8,8701 & 7,1561 & 15,2723 \\ -13,7328 & -14,2941 & 12,9454 & -7,7439 & -34,5032 & -12,7927 & 0,945 & -41,3799 & -18,1457 & -24,7938 & 18,5359 & 4,4258 & 6,9504 & -16,7914 & -39,8922 \\ -23,6623 & 21,7338 & 0,0075 & 14,6099 & -4,7738 & -7,9426 & -4,1239 & -2,6431 & 1,1198 & 0,6996 & -22,8019 & -6,3744 & 5,4939 & -14,5308 & -6,275 \end{bmatrix}$$

$$W_{kn}^{(2)} = \begin{bmatrix} 0,0351 & 0,1848 & 0,1599 & 0,1239 \\ -0,0041 & 0,0193 & 0,0189 & 0,0039 \\ 0,249 & 1,2998 & 1,869 & 1,9787 \\ 0,1199 & 0,0084 & -0,036 & 0,0342 \\ 0,0556 & 0,0942 & 0,1276 & 0,0776 \\ 0,0523 & 0,047 & 0,0517 & 0,0298 \\ 0,06 & 0,0834 & 0,0971 & 0,0748 \\ -0,0194 & -0,0257 & -0,0272 & -0,0107 \\ -0,0295 & -0,0506 & -0,0484 & -0,0346 \\ -0,0399 & -0,0726 & -0,0798 & -0,061 \\ -0,0498 & -0,0781 & -0,084 & -0,0579 \\ -0,0691 & -0,1175 & -0,1247 & -0,0992 \\ -0,0798 & -0,1413 & -0,1461 & -0,1331 \\ 0,8116 & 2,5045 & 2,8068 & 2,8684 \\ -0,9058 & -2,3108 & -2,71 & -2,4374 \end{bmatrix}$$

$$b_k^{(1)} = \begin{bmatrix} -6,2154 \\ -2,1603 \\ -6,648 \\ -4,5227 \\ 3,6718 \\ -1,5803 \\ 0,6918 \\ 1,3912 \\ 0,3014 \\ 0,8621 \\ -1,0591 \\ 1,2258 \\ -1,3259 \\ -7,527 \\ -7,284 \end{bmatrix}$$

$$b_n^{(2)} = \begin{bmatrix} 0,04 \\ 0,0826 \\ 0,0533 \\ 0,0582 \end{bmatrix}$$

## 5. CONCLUSÕES

Neste estudo efectuou-se uma análise empírica com o objectivo de observar o comportamento de uma rede na detecção dos mercados Euronext com proporcionam a maior rendibilidade diária. A rede, com uma tipologia 7-15-4, com função de propagação aditiva, função de activação logística nos neurónios da camada oculta, e função de activação linear nos neurónios da camada de *output*, foi treinada com o algoritmo de Levenberg-Marquardt.

As variáveis dependentes utilizadas foram as cotações de fecho dos índices PSI20 (Lisboa), CAC40 (Paris), AEX (Amesterdão) e BEL20 (Bruxelas), e as de *input* foram as cotações de fecho destes índices desfasadas 1 dia, o valor da taxa de juro Euribor a 3 meses, o valor de fecho do índice petrolífero XOI convertido em euros, e o valor de fecho do índice Nasdaq100, também desfasados 1 dia.

Numa 1ª experiência foram utilizados os dados em nível (cotações de fecho), no sentido de tentar tirar partido de uma das maiores vantagens dos modelos neuronais: a não necessidade de se efectuar hipóteses relativamente à distribuição dos dados. Esta experiência permitiu concluir que um investidor que seguisse os conselhos da rede para seleccionar os mercados Euronext mais rentáveis, conseguiria obter nos primeiros meses de 2007 uma rendibilidade de 11,63%, superior a qualquer rendibilidade proporcionada pelas 5 estratégias de investimento alternativas consideradas: *buy and hold* dos índices PSI20, CAC40, AEX, BEL20 e NXT100. Não foram considerados custos de transacção, e assumiu-se que o investidor hipotético conseguiria tomar posições longas ou curtas nos índices ao preço de fecho dos mesmos.

Para contrariar a crítica da *black box* que se faz amiúde aos modelos neuronais, foi efectuada uma experiência de *bootstrap* para avaliar a significância estatística das variáveis de *input*, e tentar perceber como a rede relaciona as diversas variáveis. Conclui-se que as variáveis independentes são relevantes para prever o comportamento das dependentes, pelo menos em determinados pontos da amostra. Mais especificamente, verificou-se que, em primeiro lugar, o principal factor explicativo para um determinado índice de mercado é o próprio índice desfasado um dia.

Em segundo lugar, existem determinados momentos em que algumas das outras variáveis independentes contribuem significativamente para explicar o índice, sobrepondo-se inclusive à capacidade explicativa do próprio índice desfasado.

Em terceiro lugar, e ao contrário do inicialmente esperado, o preço do petróleo e a taxa de juro nem sempre influenciam de forma negativa o valor dos índices, numa base diária, contrariando assim a relação teórica entre estas variáveis, no prazo de um dia.

Em quarto lugar, o mercado Nasdaq parece ter tido uma maior importância no comportamento dos mercados Euronext no passado, até cerca de 2004, do que tem hoje em dia. Esta situação pode eventualmente estar relacionada com o refrear do entusiasmo que se

criou no passado em relação às expectativas de evolução dos preços das acções das empresas do sector da “Nova Economia”.

Em quinto lugar, os resultados sugerem que o mercado francês é o que mais influencia os restantes, principalmente no período mais recente, e os mercados menos influentes são o português e o belga. Este sistema de relações é consentâneo com a dimensão relativa dos vários mercados. Considerando que os testes de qualidade do modelo produziram resultados razoáveis, e que o procedimento de *bootstrap* utilizado mostrou que as influências mútuas dos vários mercados Euronext são significativas do ponto de vista estatístico, conclui-se que estes estão significativamente interligados. Com efeito, a utilização de 4 neurónios na camada de *output* mostra que os mercados reagem da mesma forma a grande parte da informação de *input*, seguindo caminhos distintos apenas na fase terminal da rede.

Relativamente às consequências para as estratégias de investimento, os resultados sugerem que a informação que os investidores podem retirar actualmente do Nasdaq não é eventualmente tão relevante para os seus investimentos, numa base diária, como o foi no passado. Além disso, o mercado francês parece ter um impacto negativo nos restantes mercados Euronext, *ceteris paribus*. Esta constatação não garante, contudo, que se o mercado francês valorizar (desvalorizar) num determinado dia, os investidores possam esperar sempre uma desvalorização (valorização) nos outros mercados Euronext no dia de negociação seguinte. As elasticidades são parciais e não totais, e por isso os movimentos finais dos mercados dependem da conjugação de todas as variáveis.

Para tentar melhorar os resultados da 1ª experiência, realizou-se uma 2ª experiência em que se converteu as cotações de fecho em rendibilidades. Os resultados melhoraram efectivamente, e um investidor que seguisse os conselhos da rede neuronal conseguiria desta vez uma rendibilidade de 16,66%, mais do dobro da rendibilidade proporcionada pela melhor das 5 estratégias alternativas (o *buy and hold* do PSI20 seria a melhor estratégia alternativa nos primeiros meses de 2007, proporcionando uma rendibilidade de 7,37%). Verificou-se ainda que as outras estatísticas utilizadas para avaliar a qualidade do modelo neuronal proporcionaram resultados aceitáveis. Por exemplo, o teste de Pesaran-Timmermann permitiu concluir que o sucesso obtido pela rede na previsão da direcção dos

movimentos dos mercados Euronext não foi devido ao acaso; de igual forma a matriz das classificações permitiu concluir que o sucesso do modelo na classificação diária de cada mercado em 1º melhor, 2º, 3º ou 4º melhor mercado, também não foi fruto do “acaso máximo” ou do “acaso proporcional”.

## REFERÊNCIAS

- Beck, N., King, G., & Zeng, L. (2000). Improving Quantitative Studies of International Conflict: A Conjecture. *American Political Science Review*. Vol. 94, No. 1 March 2000.
- Demuth, H., Beale, M., & Hagan, M. (2006). *Neural Network Toolbox: For Use with MATLAB*. The MathWorks, Inc.
- Efron, B., & Tibshirani, R. (1993). *An Introduction to the Bootstrap*. Chapman and Hall, New York. Z.
- Eisinga, R., Franses, P., & Dijk, D. (1997). Timing of Vote Decision in First and Second Order Dutch Elections 1978-1995: Evidence from Artificial Neural Networks. *Oxford Journal, Political Analysis*.
- Feldman, J. A., & Ballard, D. H. (1982). Connectionist models and their properties. *Cognitive Science*, 6, 205-254.
- Fu, L. (1994). *Neural Networks in Computer Intelligence*. New York: McGraw-Hill.
- Gately, E. (1996). *Neural Networks for Financial Forecasting*. John Wiley & Sons. New York.
- Gorr, L. (1994). Research prospective on neural network forecasting. *International Journal of Forecasting*, 10, 1-4.
- Gurney, K. (1997). *An Introduction to neural networks*. CRC Press.
- Green, H. & Pearson, M. (1994). Neural nets for foreign exchange trading. In G. J. Deboeck (Ed.), *Trading on the Edge* (pp. 123-129). New York: John Wiley & Sons.
- Hagan, M.T., & M. Menhaj (1994). Training feed-forward networks with the Marquardt algorithm. *IEEE Transactions on Neural Networks*, Vol. 5, No. 6, 1999, pp. 989-993.
- Hagan, M.T., H.B. Demuth, & M.H Beale (1996). *Neural Network Design*. Boston. MA: PWS Publishing.
- Hornik, K., Stinchcombe, M., & White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural Networks*, 2(5), 359-366.
- Hornik, K. (1991). Approximation Capabilities of Multilayer Feedforward Networks. *Neural Networks*, 4, 251-257.
- Kröse, B., & Smagt, P. (1996). *An Introduction to Neural Networks*. Eighth edition. The University of Amsterdam.
- Law, R. & Pine, R. (2004). Tourism Demand Forecasting for the Tourism Industry: A Neural Network Approach. In G. Peter Zang, *Neural Networks in Businesses Forecasting*. Chapter VI. IRM Press.
- Lawrance, R. (1997). *Using Neural Networks to Forecast Stock Market Prices*. University of Manitoba.
- Li, L., Pang, W., & Yu, W. (2004). Forecasting Short-Term Exchange Rates: A Recurrent Neural Network Approach. In G. Peter Zang, *Neural Networks in Businesses Forecasting*. Chapter X. IRM Press.
- Massoumi, E., A. Khotanzad, & A. Abaye (1994). Artificial Neural Networks for Some Macroeconomic Series: A First Report, *Econometric Reviews*, 13(1).
- McCluskey, P. (1993). *Feedforward and Recurrent Neural Networks and Genetic Programs for Stock Market and Time Series Forecasting*. Brown University.



- McNelis, P. D. (2005). *Neural Networks in Finance: Gaining Predictive Edge in the Market*. Elsevier Academic Press.
- Nargundkar, S., & Priestley, J. (2004). Assessment of Evaluation Methods for Prediction and Classifications of Consumer Risk in the Credit Industry. In G. Peter Zang, *Neural Networks in Businesses Forecasting*. Chapter XIV. IRM Press.
- Neves, J. C., & Vieira, A. (2004). Estimating Banruptcy Using Neural Networks Trained with Hidden Layer Learning Vector Quantization. Working Paper, Departamento de Gestão, ISEG, UTL.
- Pan, H., Tilakaratne, C., & Yearwood, J. (2005). Predicting Australian Stock Market Index Using Neural Networks Exploiting Dynamical Swings and Intermarket Influences. *Journal of Research and Practice in Information Technology*, Vol. 37, No. 1, February 2005.
- Pesaran, M., & A. Timmermann (1992). A Simple Nonparametric Test of Predictive Performance. *Journal of Business and Economics Statistics*, 10: 461-465.
- Reed, R., & Marks, R. (1999). *Neural Smithing: supervised learning in feedforward artificial neural networks*. The MIT Press.
- Reis, E. (2001). *Estatística Multivariada Aplicada*. 2ª Edição revista e corrigida. Edições Sílabo.
- Roman, J., & Jameel, A. (1996). Backpropagation and Recurrent Neural Networks in Financial Analysis of Multiple Stock Market Returns. Proceedings of the 29th Annual Hawaii International Conference on System Sciences.
- Saad, E., Prokhorov, D., & Wunsch, D. (1998). Comparative Study of Stock Trend Prediction Using Time Delay, Recurrent and Probabilistic Neural Networks. *IEEE Transactions on Neural Networks*, Vol 9, No. 6, November 1998.
- Shachmurove, Y. (2002). *Applying Artificial Neural Networks to Business, Economics and Finance*. CARESS Working Papers. UCLA Department of Economics.
- Shachmurove, Y. & Witkowska, D. (2000). Utilizing Artificial Neural Network Model to Predict Stock Markets. CARESS Working Papers. UCLA Department of Economics.
- Stanford University's Computer Science Education. (<http://www-cse.stanford.edu/classes/sophomore-college/projects-00/neural-networks/Biology/index.html>).
- Tahai, A., Walczak, S., & Rigsby, J.T. (1998). Improving artificial neural network performance through input variables selection. In P. Siegel, K. Omer, A. deKorvin & A. Zebda (Eds.), *Applications of Fuzzy Sets and The Theory of Evidence to Accounting II* (pp. 277-292). Stamford, CT: JAI Press.
- Thawornwong, S., & Enke, D. (2004). Forecasting Stock Returns with Artificial Neural Networks. In G. Peter Zang, *Neural Networks in Businesses Forecasting*, Chapter III, IRM Press.
- Walczak, S. & Cerpa, N. (1999). Heuristic Principles for the design of artificial neural networks. *Information And Software Technology*, 41(2), 109-119.
- Walczak, S. (2001). An empirical analysis of data requirements for financial forecasting with neural networks. *Information And Software Technology*, 41(2), 109-119.
- Walczak, Steven (2004), Forecasting Emerging Market Indexes with Neural Networks. In G. Peter Zang, *Neural Networks in Businesses Forecasting*, Chapter IV (pp. 84), IRM Press.
- White, H. (1990). Connectionist nonparametric regression: Multilayer feedforward networks can learn arbitrary mappings. *Neural Networks*, 3(5), 535-549.
- Widrow, B., Rumelhart, D. E., & Lehr, M. A. (1994). Neural Networks: Applications in industry, business and science. *Communications of the ACM*, 37(3), 93-105.
- Zhang, Y., Akkaladevi, S., Vachtsevanos, G., & Lin, T. (2002). Granular neural web agents for stock prediction. *Soft Computing* 6 (2002) 406 – 41. Springer-Verlag.