

# XII CONGRESO INTERNACIONAL DE INVESTIGACIÓN EN CIENCIAS ADMINISTRATIVAS (ACACIA)

## **Modelação da Procura Turística: Um estudo comparativo entre Redes Neurais Artificiais e a Metodologia de Box-Jenkins**

### **MESA DE TRABAJO: FINANZAS Y ECONOMÍA**

Paula Odete Fernandes (responsable); [pof@ipb.pt](mailto:pof@ipb.pt); Departamento de Economia e Gestão  
João Paulo Teixeira; [joaopt@ipb.pt](mailto:joaopt@ipb.pt); Departamento de Electrotecnia  
(Telf.: +351.273.303103; Fax: +351.273.313051)  
Escola Superior de Tecnologia e de Gestão (ESTiG)  
Instituto Politécnico de Bragança (IPB)  
Campus de Sta. Apolónia, Apartado 134  
5301-857 Bragança, Portugal

João José Ferreira; [jjmf@ubi.pt](mailto:jjmf@ubi.pt); Departamento de Gestão e Economia  
Susana Garrido Azevedo; [sazevedo@ubi.pt](mailto:sazevedo@ubi.pt); Departamento de Gestão e Economia  
(Telf.: +351.275.319600; Fax: +351.275.319601)  
Pólo IV - Edifício Ernesto Cruz  
6200-209 Covilhã, Portugal

CETYS, Universidad de Baja California.  
13-16 Mayo de 2008

## **Modelação da Procura Turística: Um estudo comparativo entre Redes Neuronais Artificiais e a Metodologia de Box-Jenkins**

### **ABSTRACT**

The present research aims to explore and to evidence the utility of the methodology of Artificial Neural Networks (ANN) in the analysis of tourism demand as an alternative to the Box-Jenkins methodology. The first methodology has arising interest in the economic and business area since several researches have verified that methodology presents a valid alternative to classical methods of forecasting allowing giving answer to situations in which the traditional ones will be of difficult to apply (Thawornwong & Enke, 2004). According to Hill *et al.* (1996) and Hansen *et al.* (1999) ANN show capacity to improve the time-series forecasts through of additional information analysis decreasing their dimension and reducing their complexity. For that, each one of the referred methodologies focused in the treatment, analysis and modeling of the tourism time-series: *Monthly Guest Nights in Hotels* registered between January 1987 to December 2006, since it is one of the variables that better explain the effective tourism demand. The Study was performed for two regions of Portugal: North region and Centre region. Considering the results, and according to the Criteria of MAPE for model evaluation proposed by Lewis (1982), the ANN model presented acceptable statistical qualities and adjustments satisfied. Being so, it is adequate not only for the modelling but also to the prediction of times series, when compared to the model performed by Box-Jenkins methodology.

We intended also to evaluate the performance and competitiveness of the tourism destinations - North region and Center region of Portugal - by main origin markets and to analyse how it is distributed their portfolio of origin markets for the period of 1997 to 2006. The Market Share Analysis tool proposed by Faulkner (1997) was applied and it was observed an high dependency of the domestic market for both regions.

**Keywords:** Artificial Neural Networks; ARIMA Models; Time Series Forecasts; Tourism Destinations; Competitiveness, Market Share.

**JEL-Codes:** C01; C02; C22; C45.

## **Modelação da Procura Turística: Um estudo comparativo entre Redes Neuronais Artificiais e a Metodologia de Box-Jenkins**

### **RESUMO**

O presente estudo pretende explorar e evidenciar a utilidade da metodologia das Redes Neuronais Artificiais como uma alternativa à metodologia de Box-Jenkins, na análise da procura turística. A primeira metodologia tem vindo a suscitar interesse na área das ciências económicas e empresariais, pois pelos trabalhos de investigação realizados tem-se verificado que a mesma apresenta uma alternativa válida a métodos clássicos de previsão, conseguindo dar resposta a situações que pelos métodos clássicos seriam de difícil tratamento (Thawornwong & Enke, 2004). Hill *et al.* (1996) e Hansen *et al.* (1999), referem que as ANN mostram capacidade para melhorar a previsão de séries temporais através da análise de informação adicional, diminuindo a sua dimensão e reduzindo a sua complexidade. Para tal, cada uma das metodologias referidas centrou-se no tratamento, análise e modelação da série temporal de turismo: “Dormidas Mensais nos Estabelecimentos Hoteleiros”, registadas no período de Janeiro de 1987 a Dezembro de 2006, uma vez que é uma das variáveis que melhor traduz a procura efectiva. O estudo foi realizado para as regiões Norte e Centro de Portugal. Os resultados obtidos, e tendo por base a classificação do MAPE proposto por Lewis (1982), revelaram que o modelo obtido, utilizando a metodologia das Redes Neuronais Artificiais, apresentou qualidades estatísticas e de ajustamento satisfatórias evidenciando ser adequado para a modelação e previsão da série de referência, quando comparado com o modelo produzido pela metodologia de Box-Jenkins.

Pretendeu-se ainda, com este estudo, avaliar o desempenho e a competitividade dos destinos turísticos - Região Norte e Região Centro, de Portugal - por principais mercados emissores e analisar como se encontra distribuída a sua carteira de mercados emissores, para o período de 1997 a 2005. Utilizou-se para o efeito o instrumento de análise proposto por Faulkner (1997), tendo-se observado uma grande dependência do mercado interno, para ambas as regiões.

**Palavras-Chave:** Redes Neuronais Artificiais; Modelos ARIMA; Previsão de Séries Temporais; Destinos Turísticos; Competitividade, Quotas de Mercado.

**JEL:** C01; C02; C22; C45.

## 1. Introdução

Inúmeros estudos empíricos têm sido realizados e publicados na área do turismo, nos últimos tempos, e são unânimes ao considerar que no processo de planeamento, tomada de decisões e controlo do sector do turismo a previsão da procura turística assume um papel importante (Witt & Witt, 1995; Wong, 2002; Fernandes, 2005; Yu & Schwartz, 2006).

Na área da previsão, actualmente, encontra-se disponível uma grande multiplicidade de métodos que têm vindo a emergir para fazer face às mais variadas situações, apresentando características e metodologias diferentes, que vão desde as abordagens mais simples às mais complexas. Os modelos de previsão de *Box-Jenkins* pertencem à família de modelos algébricos designados por modelos ARIMA e permitem efectuar previsões com base numa dada série temporal estacionária. A metodologia considera que uma série temporal real constitui uma provável realização, de um determinado processo estocástico. O objectivo da análise é identificar o modelo que melhor represente o processo estocástico desconhecido subjacente e que seja uma boa representação da realização, isto é, da série temporal real. Uma outra metodologia que tem tido inúmeras aplicações nas mais diversas áreas do conhecimento e na área da previsão tem sido aplicada como uma alternativa aos modelos clássicos, são os modelos baseados nas redes neuronais artificiais. Estes modelos, não lineares surgiram como tentativa de reprodução do funcionamento do cérebro humano, sendo o complexo sistema de neurónios biológicos a sua principal fonte de inspiração.

Com a presente investigação pretende-se explorar e evidenciar a utilidade da metodologia das Redes Neuronais Artificiais como uma alternativa à metodologia de Box-Jenkins, na análise da procura turística e avaliar o desempenho e a competitividade dos destinos turísticos por principais mercados emissores. A primeira metodologia tem vindo a suscitar interesse na área das ciências económicas e empresariais, pois pelos trabalhos de investigação realizados tem-se verificado que a mesma apresenta uma alternativa válida a métodos clássicos de previsão, conseguindo dar resposta a situações que pelos métodos clássicos seriam de difícil tratamento (Thawornwong & Enke, 2004). Hill *et al.* (1996) e Hansen *et al.* (1999) referem que as ANN mostram capacidade para melhorar a previsão de séries temporais através da análise de informação adicional, diminuindo a sua dimensão e reduzindo a sua complexidade. Para tal, cada uma das metodologias referidas centrou-se no tratamento, análise e modelação da série temporal de turismo: “Dormidas Mensais nos Estabelecimentos Hoteleiros”. A série Dormidas Mensais de Turistas é considerada significativa da actividade turística, devido às suas características, e contempla o número de visitantes que usufruíram das capacidades turísticas. O estudo foi realizado para duas regiões de Portugal: Norte e Centro. Assim, a análise desenvolvida nesta investigação terá

como séries de estudos as Dormidas Mensais Registadas na Região Norte [RN] e as Dormidas Mensais Registadas na Região Centro [RC]. Os dados observados consideram o período compreendido entre Janeiro de 1987 e Dezembro de 2006, correspondendo assim a 240 observações mensais ao longo de 20 anos.

A presente investigação está estruturada da seguinte forma: depois da introdução, apresentar-se-á na segunda secção as metodologias aplicadas, ou seja, as redes neuronais artificiais e a metodologia de Box-Jenkins. Seguidamente procede-se à apresentação e análise da série temporal “Dormidas mensais de turistas” para as regiões em estudo construindo-se modelos e elabora-se a previsão da procura turística para os anos 2005 e 2006. Na secção três avaliar-se-á a competitividade entre os destinos turísticos analisados. Por último, serão apresentadas as conclusões e desenvolvimentos futuros.

## **2. Redes Neuronais Artificiais versus Metodologia de Box-Jenkins**

### **2.1. Metodologias Utilizadas**

A metodologia proposta por Box e Jenkins, em 1970, permite efectuar uma análise ao comportamento das séries temporais, com base num duplo estudo e conjunto: por um lado uma componente autorregressiva que se estabelece em função da história estatística passada das variáveis consideradas e, por outro lado, um tratamento dos factores aleatórios ou estocásticos, especificado por processos de médias móveis. Estes modelos, devido ao seu esquema de delineamento e resolução operativa, permitem a incorporação de análises de carácter sazonal, o isolamento da componente tendência e, inclusive aprofundar as inter-relações entre estas componentes que se integram na evolução das séries em estudo (Parra & Domingo, 1987; Chu, 1998). Os modelos introduzidos por *Box* e *Jenkins* descrevem exclusivamente séries estacionárias, ou seja, com média e variância constantes no tempo e autocovariância dependente apenas do grau de desfasamento entre as variáveis, devendo começar-se por verificar ou provocar a estacionaridade da série (Pulido, 1989). Trata-se dos chamados modelos ARIMA (*Autoregressive Integrated Moving Average*), bastante adequados para previsões de curto prazo e no caso de séries que contenham variações sazonais (Witt & Witt, 1995).

Assim, para se aplicar a metodologia de *Box-Jenkins* é necessário em primeiro lugar identificar a série e remover a não estacionaridade, pelo que se deve aplicar uma ou mais transformações sobre os valores da série de modo a obter uma outra série (transformada dos valores originais) estacionária. Tais transformações, apesar de conservarem a estrutura geral da série, exercem efeitos consideráveis sobre o conjunto de dados que podem tornar o seu estudo mais cómodo, alterando a sua escala (eventualmente diminuindo a sua

amplitude), atenuando assimetrias, eliminando possíveis *outliers*, diminuindo resíduos e atingindo finalmente os objectivos em causa: estabilizando variâncias e linearizando tendências (Otero, 1993; Fernandes & Cepeda, 2000). Após a identificação da série deve estimar-se os seus parâmetros e posteriormente efectuar a avaliação do ajustamento e, caso necessário, deve tentar-se encontrar um novo modelo que melhor descreva o fenómeno em causa. Finalmente proceder à fase de previsão.

Neste sentido, o modelo ARIMA ( $p,d,q$ ), em que:  $p$  corresponde à ordem do processo Autorregressivo (AR);  $d$  é o número de diferenças ou integrações; e  $q$  corresponde à ordem do processo Médias Móveis (MA), vem representado pela seguinte expressão (Murteira *et al.*, 1993; Zou & Yang, 2004):

$$(1 - \phi_1 B - \dots - \phi_p B^p)(1 - B)^d Y_t = (1 - \theta_1 B - \dots - \theta_q B^q) e_t \quad [1]$$

ou ainda de uma forma mais resumida:

$$\phi_p(B) \nabla^d Y_t = \theta_q(B) e_t \quad [2]$$

As aplicações usuais dos modelos ARIMA realizam-se com dados trimestrais, mensais ou mesmo semanais, diários ou horários, ou seja, num contexto de previsão de curto prazo. Para tais efeitos, os modelos ARIMA tratam de captar o comportamento sazonal, de uma maneira idêntica ao tratamento da componente regular (ou não sazonal) da série. Nas aplicações não é frequente trabalhar só com um modelo ARIMA ( $p,d,q$ ), mas sim com o produto dos modelos: ARIMA( $p,d,q$ )( $P,D,Q$ )<sub>s</sub> em que a primeira parte corresponde à parte regular e a segunda à parte sazonal, correspondente à seguinte expressão (Murteira *et al.*, 1993; Zou & Yang, 2004):

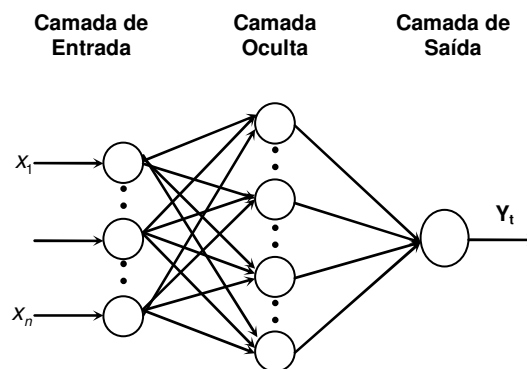
$$\phi_p(B) \Phi_P(B^s)(1 - B)^d (1 - B^s)^D Y_t = \theta_q(B) \Theta_Q(B^s) e_t \quad [3]$$

As previsões com o modelo ARIMA, em função dos dados históricos, vêm dadas pela função de previsão:

$$Y_t^*(m) = E\{Y_{t+m} / Y_t, Y_{t-1}, Y_{t-2}, \dots\} \quad [4]$$

Uma outra metodologia que tem merecido alguma atenção, por parte da comunidade científica nos últimos anos, apresentando alguns avanços na área de conhecimento das ciências empresariais, assenta na utilização das redes neuronais artificiais (RNA). As RNA são modelos que se destacam dentro de uma ampla área do conhecimento da inteligência artificial. Estes têm como base modelos matemáticos de arquitectura semelhante à do cérebro humano. Uma rede neuronal encontra-se composta por um conjunto de neurónios artificiais, nós ou um conjunto de unidades de processamento, ligadas entre si - as conexões

entre as unidades de processamento denominam-se por *sinapses* -, que processam e transmitem informação através de funções de activação, sendo as mais utilizadas as funções lineares e as sigmoidais - logística e tangente hiperbólica - (Rodrigues, 2000; Fernandes, 2005). Ainda de referir que os neurónios de uma rede estão estruturadas em distintas camadas (mais conhecidas por camada de entrada, intermédia ou oculta e de saída), sendo as mais comuns, para a previsão de séries temporais, as multi-camadas ou também designadas por MLP<sup>1</sup> (Bishop, 1995), por forma que um neurónio de uma camada encontra-se ligada com os neurónios da camada seguinte às que pode enviar informação, Figura 1, (Fernandes, 2005). As redes, dependendo da forma como estão ligadas entre as diferentes camadas podem classificar-se em redes recorrente<sup>2</sup> (*feedback*) e redes não recorrentes<sup>3</sup> (*feedforward*).



**Figura 1.** Estrutura de uma Rede Neuronal Artificial *Feedforward*.

Da especificação da rede neuronal também faz parte uma função de erro e um algoritmo para determinar o valor dos parâmetros que minimizam a função de erro. Desta forma, existem dois conceitos centrais: a parte física da rede, ou seja, a sua arquitectura, e o procedimento algorítmico que determina o seu funcionamento, ou seja, a forma como a rede se modifica consoante os dados fornecidos pelo ambiente (Haykin, 1999).

É importante ainda referir que para as RNA aprenderem com a experiência têm que ser submetidas a um processo designado de treino, para o que existem diferentes algoritmos de treino, sendo um dos mais utilizados na previsão de séries temporais o algoritmo *Backpropagation*<sup>4</sup> ou suas variantes, distribuindo-se os mesmos por duas classes (i) supervisionados e (ii) não supervisionados (Haykin, 1999). Para o primeiro caso, durante o processo de treino existe um ‘professor’ que fornece um conjunto de casos de treino, e um

<sup>1</sup> Do inglês *Multi-Layer Perceptron*.

<sup>2</sup> As conexões permitem que a informação regresse a lugares por onde já passou e admite-se também ligações intra-camada (laterais) (Fernandes, 2005).

<sup>3</sup> A informação flui de forma unidireccional de uma camada para outra, desde a camada de entrada para a camada oculta e posteriormente para a camada de saída (Fernandes, 2005).

<sup>4</sup> Este algoritmo procura o mínimo da função de erro no espaço de procura dos pesos das ligações entre os neurónios, baseando-se em métodos de gradiente descendente. A combinação de pesos que minimiza a função de erro é considerada a solução para o problema de aprendizagem. A descrição do algoritmo pode ser analisada em Rumelhart and McClelland (1986) e em Haykin (1999).

caso de treino consiste num vector de entrada  $X$  e no correspondente vector de saída  $Y$ , passando a aprendizagem pela minimização do erro da saída, o que é feito através do ajustamento dos pesos das conexões de acordo com uma determinada regra. No segundo caso, existe um conjunto de entradas, pelo que o algoritmo de treino tenta agrupar os dados a partir de padrões por estes apresentados, seguindo assim uma regra de auto-organização (Haykin, 1999; Fernandes, 2005).

Resumidamente um valor produzido por uma rede *feedforward*, com uma camada oculta, pode ser expressa da seguinte forma (Fernandes & Teixeira, 2007):

$$Y_t = b_{2,1} + \sum_{j=1}^n \alpha_j f \left( \sum_{i=1}^m \beta_{ij} y_{t-i} + b_{1,j} \right) \quad [5]$$

Onde,

$m$ , número de nós na camada de entrada;

$n$ , número de nós na camada oculta;

$f$ , função de activação sigmoideal;

$\{\alpha_j, j = 0, 1, \dots, n\}$ , vector de pesos que liga os nós da camada oculta aos da camada de saída;

$\{\beta_{ij}, i = 0, 1, \dots, m; j = 1, 2, \dots, n\}$ , pesos que relacionam os nós da camada de entrada aos da camada oculta;

$b_{2,1}$  y  $b_{1,j}$ , indicam os pesos dos termos independentes (*bias*) associados a cada nó da camada de saída e da camada oculta, respectivamente.

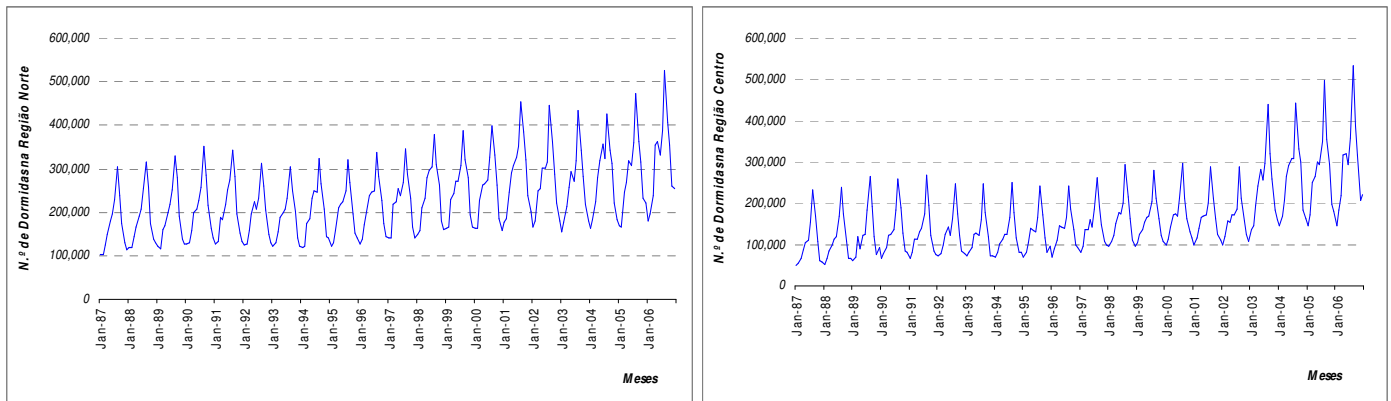
A equação indica também a utilização duma função de activação linear na camada de saída.

## 2.2. Apresentação e Análise do Comportamento das Séries Temporais

A série Dormidas Mensais de Turistas é considerada como significativa da actividade turística, devido às suas características, e contempla visitantes que usufruíram das capacidades turísticas, neste caso das regiões Norte e Centro de Portugal.

Assim, a análise que se vai levar a cabo terá como séries de estudos as Dormidas Mensais Registadas na Região Norte [RN] e as Dormidas Mensais Registadas na Região Centro [RC]. Os dados observados consideram o período compreendido entre Janeiro de 1987 e Dezembro de 2006, correspondendo assim a 240 observações mensais ao longo de 20 anos (ver Anexo A, Tabelas A.1 e A.2). Os valores das séries foram facultados pelo Instituto Nacional de Estatísticas de Portugal (INE).

Na Figura 2, apresentam-se as duas séries, pelo que da análise do seu comportamento verifica-se, facilmente, a existência de oscilações irregulares sugerindo uma não estabilização da média e a presença de sazonalidade (valores máximos nos meses de Verão e mínimos nos meses de Inverno), isto é, os valores das dormidas vêm em função da altura do ano.

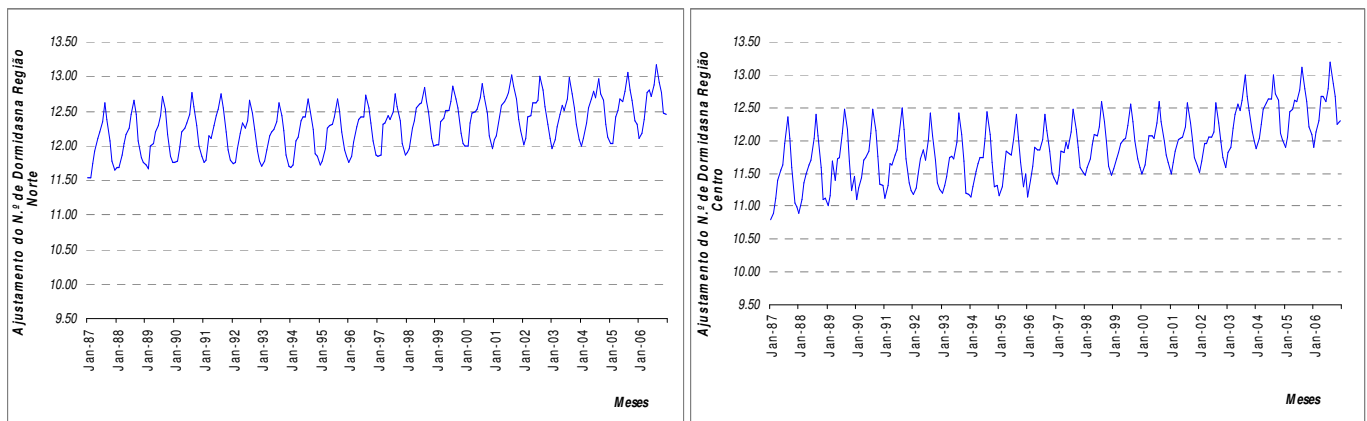


**Figura 2.** Cronograma das séries para o período de 1987:01 a 2006:12.

## 2.3. Construção dos Modelos

### 2.3.1. Modelo ARIMA

Para se aplicar a metodologia de Box-Jenkins existe a necessidade, numa primeira fase, de converter as séries temporais em séries estacionárias. Assim, no sentido de estabilizar a variância das séries, estas transformaram-se aplicando, a cada uma, o logaritmo natural: LRN e LRC, respectivamente para a região Norte e para a região Centro.



**Figura 3.** Cronograma das séries transformadas para o período de 1987:01 a 2006:12.

Através da análise da Figura 3, pode observar-se que as séries continuam não estacionárias mas conseguiu-se alguma estabilização em termos de variância, notando-se ainda uma tendência crescente e a existência de movimentos periódicos. Assim, na continuação do estudo das séries toda a análise terá como base as séries transformadas e o período

compreendido entre Janeiro de 1987 e Dezembro de 2004. Os anos de 2005 e 2006, vão ser considerados apenas para analisar o desempenho do modelo construído, ou seja, servirão como conjunto de teste.

Uma vez que com a transformação efectuada, aplicação do logaritmo natural, não se conseguiu converter as séries em estacionárias, deve aplicar-se mais uma transformação - aplicar diferenças<sup>5</sup>.

A estacionarização das séries em estudo, foi conseguida após a aplicação de uma diferenciação simples  $[\nabla Y_t = Y_t - Y_{t-1} = (1-B)Y_t]$  e uma diferença sazonal  $[\nabla_s Y_t = Y_t - Y_{t-s} = (1-B^s)Y_t]$ . O mesmo é dizer que se aplicaram transformações sucessivas e diferenças entre as observações separadas pelo período sazonal (de 12 em 12 meses), transformando-se as séries anteriores em novas séries. Assim, os resultados das novas séries, que servirão de base para a aplicação da metodologia de Box-Jenkins vêm dados pelas expressões, região Norte [6] e região Centro [7]:

$$(1-B^{12})(1-B)LRN_t \quad [6]$$

$$(1-B^{12})(1-B)LRC_t \quad [7]$$

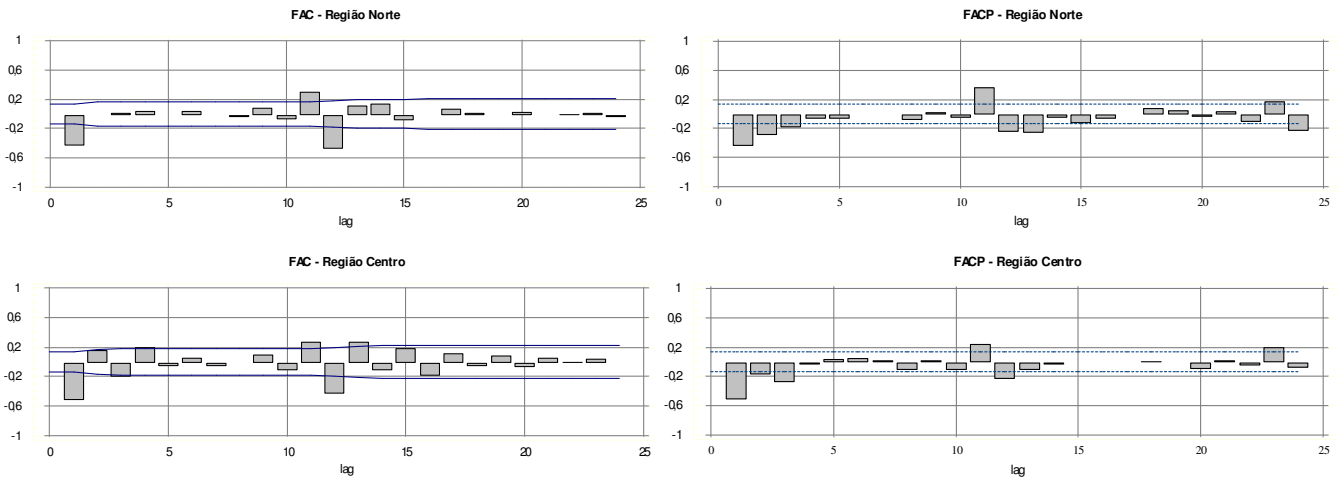
A fase seguinte assenta na identificação dos modelos. Este processo tem por base a análise dos correlogramas das Funções de Autocorrelação (FAC) e Autocorrelação Parcial (FACP). A identificação das componentes sazonal e não sazonal é feita separadamente recorrendo aos figurinos teóricos (Otero, 1993; Fernandes, 2005).

Observando FAC e FACP para as duas séries, Figura 4, estimadas após diferenciações simples e sazonal, considerando um intervalo de confiança de 95%, parecem sugerir, para ambas as séries:

- (i) um processo, para a componente não sazonal, ARMA (0,1) visto que o primeiro coeficiente da estimativa da FAC, para ambas as séries, é significativo decaindo os restantes para zero e os valores iniciais da FACP são significativos e a decair de uma forma exponencial;
- (ii) relativamente à componente sazonal, as FAC e FACP estimadas sugerem igualmente um processo ARMA (0,1) atendendo aos valores da FAC estimada para o *lag* 12 e 24 (sendo o primeiro significativo, enquanto o segundo é sem qualquer expressão) e aos valores da FACP para os mesmos *lags*, ambos significativos.

---

<sup>5</sup> É aconselhável minimizar as diferenciações dos dados (para evitar a sobre-diferenciação ou *overdifferencing*), pois diferenciar origina aumento da variância do erro da previsão (Murteira *et al.*, 1993; González, 1999).



**Figura 4.** FAC e FACP estimadas das séries após diferenciações simples e sazonal, para as duas regiões.

A análise efectuada anteriormente sugere, para ambas as séries, os mesmos modelos  $M1 = ARIMA(0,1,1) \times (0,1,1)_{12}$  e  $M2 = ARIMA(1,1,1) \times (1,1,1)_{12}$ .

Uma vez identificados os modelos ARIMA, que melhor se ajustam às séries, existe a necessidade de determinar os valores dos parâmetros das funções lineares que os definem. O método utilizado para a estimação dos parâmetros  $\phi$  e  $\theta$  é o dos mínimos quadrados, obtendo-se os seguintes resultados:

**Tabela 1.** Estimativas obtidas para os coeficientes dos modelos.

MODELOS ARIMA	MODELOS REGIÕES	PARÂMETROS	LAGS	COEFICIENTE	DESVIO PADRÃO	RÁCIO t	P-VALUE	DESVIO PADRÃO DOS RESÍDUOS
<b>M1</b>	<b>Região Norte (MRN<sub>1</sub>)</b>	Médias Móveis	1	0,654218	0,0534728	12,2346	0,000000	0,0574563
		Médias Móveis	12	0,757521	0,0446032	16,9835	0,000000	
	<b>Região Centro (MRC<sub>1</sub>)</b>	Médias Móveis	1	0,602289	0,0548320	10,9842	0,000000	
		Médias Móveis	12	0,662380	0,0520395	12,7284	0,000000	
<b>M2</b>	<b>Região Norte (MRN<sub>2</sub>)</b>	Autorregressivo	1	0,132364	0,104493	1,26673	0,206742	0,0573292
		Médias Móveis	1	0,733003	0,070979	10,327	0,000000	
		Autorregressivo	12	-0,125477	0,095449	-1,31459	0,190167	
	Médias Móveis	12	0,703627	0,066186	10,6309	0,000000		
	<b>Região Centro (MRC<sub>2</sub>)</b>	Autorregressivo	1	0,008005	0,117814	0,067954	0,945891	
		Médias Móveis	1	0,600721	0,094128	6,38196	0,000000	
Autorregressivo		12	-0,012083	0,110839	-0,109013	0,894630		
		Médias Móveis	12	0,658766	0,080228	8,21113	0,000000	

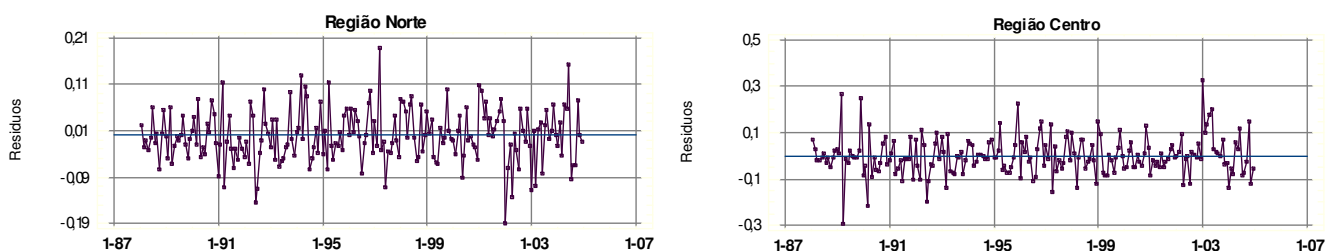
A análise da significância estatística dos coeficientes estimados para o modelo 1 (M1), para as duas séries, revela que os mesmos são, significativamente, diferentes de zero, ao nível de significância de 5%, ou seja, os rácios t para os parâmetros estimados, levam a concluir

que ambos os coeficientes são estatisticamente significativos, o mesmo é dizer que os valores para o rácio  $t$ , em valor absoluto, são superiores a 1,96 para cada parâmetro estimado, pelo que se pode dizer que os coeficientes são estatisticamente significativos e devem permanecer do modelo (Tabela 1). O mesmo não se verifica para o modelo 2 (M2), visto que se comprova que os coeficientes associados à componente AR(1) e AR(12) no permite rejeitar a hipótese de nulidade do parâmetro teórico, ou seja, os valores do  $t$  estatístico inferiores a 1,96 permitem concluir que os coeficientes não são estatisticamente significativos, pelo que e tendo em conta o princípio da parcimónia tais parâmetros devem ser excluídos dos modelos.

Quanto à invertibilidade, das duas componentes - sazonal e não sazonal - para ambos os modelos verificam as condições de invertibilidade, dado que as estimativas dos parâmetros das componentes das médias móveis são, em módulo, inferiores à unidade. Por natureza os processos autorregressivos são invertíveis.

Dado que o modelo M2 apresentou características frágeis, não avança na análise e apenas se continuará com a análise para o modelo M1 (para ambas as regiões), sendo este o modelo seleccionado para a metodologia de Box-Jenkins.

Assim, uma vez avaliada a qualidade estatística do modelo, interessa avaliar a qualidade do ajustamento, este assenta na análise dos respectivos resíduos. Com efeito, se este explica correctamente a série em causa, os resíduos estimados deverão ter um comportamento análogo ao de um ruído branco (*white noise*).



**Figura 5.** Gráfico dos resíduos para o modelo M1, para as duas regiões.

Pela análise à Figura 5, e para a região Norte pode observar-se alguns atípicos para os anos de 1992, 1997, 2001, 2002 e 2004, verificando-se ainda algumas flutuações nos meses de Março e Abril. Esta última ocorrência pode dever-se ao facto do feriado móvel da Páscoa. Relativamente aos resíduos correspondentes ao ano de 1992 (Julho e Agosto), podem vir justificados pela Guerra do Golfo e para o ano de 1997, mesmos meses, justificados pela instabilidade do mercado Russo e o conflito dos Balcãs. Para o ano de 2001, o comportamento dos resíduos pode fundamentar-se no facto da cidade do Porto ter

sido a Capital Europeia da Cultura e também pelo casco histórico da cidade de Guimarães e o Alto Douro Vinhateiro, terem sido classificados pela UNESCO como Património Cultural Mundial; tendo estes dois factores, sem dúvida, despertado a curiosidade dos turistas, quer nacionais quer estrangeiros, para se deslocarem à região Norte. Após o conhecimento da decisão da UEFA em que Portugal seria o país organizador do Campeonato Europeu de Futebol - EURO2004 - e depois da promoção, agressiva, que se fez do país junto dos países europeus, a partir de 2002, pode justificar o comportamento dos resíduos para o ano de 2002 e 2003. No ano de 2004 e para os meses de Maio e Junho, coincidente com o evento desportivo, o comportamento dos resíduos justifica-se pela realização do evento desportivo EURO2004, referindo-se que 5 dos 10 estádios de futebol encontram-se na região Norte.

Tendo ainda por base a Figura 5 e efectuando agora a análise para a região Centro, para os anos de 1989, 1990 e 1997, o comportamento dos resíduos podem justificar-se pelo feriado móvel da Páscoa, dado verificar-se nos meses de Março e Abril. Para Junho de 1992, pode apontar-se a Guerra do Golfo, levando a que os turistas seleccionem a região Centro para passarem as suas férias e em Janeiro de 2003, pode fundamentar-se no facto de nos últimos anos as autarquias da Região Centro terem vindo a apostar mais na divulgação e realização de eventos culturais a par com algum investimento na criação de melhores condições à prática de desportos de Inverno, nomeadamente a prática do *sky* e do *snow board*, que vão atraindo pessoas à região, essencialmente nos meses de Inverno.

Assim e uma vez comprovada a adequação dos resíduos do modelo M1, para as duas regiões, levou-se a cabo a análise global dos resíduos utilizando a estatística de *Box-Pierce*. Assim, para o modelo da Região Norte e para o *lag* 24, o valor de  $Q=16,6893$  e  $p\text{-value}=0,780268$ ; para o modelo da Região Centro e para o *lag* 24, o valor de  $Q=25,5231$  e  $p\text{-value}=0,272722$ . Pode então concluir-se que se aceita a ideia de que os resíduos dos modelos estimados seguem um ruído branco dado que os valores dos  $p\text{-value}$  associados ao contraste de *Box-Pierce* são diferentes de zero.

Resumindo, e tendo em atenção os diferentes critérios analisados para a avaliação dos modelos, pode dizer-se que os modelos, para cada uma das regiões, vêm expressos pelas seguintes equações:

$$MRN_1 = \nabla \nabla_{12} LRN = (1 - 0,654218B)(1 - 0,757521B^{12})e_t \quad [8]$$

$$[t_1 = 12,2346] \quad [t_{12} = 16,9835]$$

$$MRC_1 = \nabla \nabla_{12} LRC = (1 - 0,602289B)(1 - 0,662380B^{12})e_t \quad [9]$$

$$[t_1 = 10,9842] \quad [t_{12} = 12,7284]$$

Cabe sublinhar que a prova conclusiva de que o modelo seleccionado para captar o comportamento de uma série é a previsão, determinando-se assim a efectividade do estudo. Esta será realizada no ponto 2.4.

### **2.3.2. Modelo das Redes Neurais Artificiais**

O modelo de RNA seleccionado para o estudo de cada uma das séries DRN e DRC foi do tipo multi-camada, onde se utilizaram três camadas: de entrada, oculta e de saída e com uma estrutura tipo *feedforward*. Na camada oculta foi utilizada a função de activação sigmoidal logística [Logsig]. Na camada de saída utilizou-se a função de activação linear, por ser a que melhores resultados proporciona para este tipo de arquitecturas. No treino da rede utilizou-se o algoritmo *Resilient Backpropagation*, uma variante do algoritmo de treino *backpropagation*. A selecção deste algoritmo deve-se ao facto de ter produzido resultados satisfatórios em trabalhos realizados pelos autores Fernandes (2005) e Fernandes e Teixeira (2007a). As redes, utilizadas neste estudo, têm a seguinte arquitectura: 12 nós na camada de entrada, correspondentes aos últimos 12 valores da série, 4 nós na camada oculta e 1 na camada de saída, correspondente à previsão do valor do mês seguinte, ou seja (1-12;4;1). A estimação/previsão produzida foi efectuada mensalmente, isto é, *one-step-ahead forecast*. O processo de treino para actualização dos pesos fez-se utilizando o método *batch training*.

As séries temporais com os dados originais foram divididas em três conjuntos distintos, a saber: treino (as primeiras 216 observações para a série DRN e 216 observações para a DRC, considerando que as observações contempladas para a validação não foram consideradas no treino); validação (12 observações, correspondentes ao ano de 2004 para a série DRN; para a série DRC as observações: Janeiro de 1999, Fevereiro de 2004, Março de 2002, Abril de 1996, Maio de 2003, Junho de 2000, Julho de 1998, Agosto de 2004, Setembro de 1997, Outubro de 2001, Novembro de 1994 e Dezembro de 2003; optou-se por extrair estas observações para a série DRC por se acreditar ser uma 'boa' representação do conjunto total, dado o comportamento da mesma e por conhecimento do fenómeno em análise por parte dos autores) e de teste (24 observações, correspondentes aos anos de 2005 e 2006).

Cabe sublinhar que foi realizado um pré-processamento dos dados de entrada e dos dados de saída correspondente apenas a uma normalização entre -1 e 1, para ambas as séries. Após este processamento, cada uma das séries foi treinada com a introdução de mais variáveis nos modelos, o valor mais elevado das séries mais a média dos dados observados, numa primeira etapa. Dado não se terem obtido resultados satisfatórios, numa segunda etapa além de se ter utilizado as variáveis referidas anteriormente, incluiu-se ainda

no modelo a deriva - diferença - dos picos. Continuou-se a não se conseguir obter resultados satisfatórios, para o conjunto de validação, para ambas as séries, pelo que se teve de optar por um outro tipo de pré-processamento, a passagem para o domínio logarítmico, tendo-se conseguido melhorias nos resultados finais produzidos para as séries, embora para a série DRC as melhorias não foram significativas. Uma vez que já se tinha resolvido - minimizado - o problema para a série DRN houve a necessidade de experimentar mais um pré-processamento para a DRC, no sentido de 'limpar' a série. Assim, optou-se por aplicar, à série no domínio logarítmico, uma diferença simples e outra sazonal, ou seja, aplicaram-se transformações sucessivas e diferenças entre as observações separadas pelo período sazonal (de 12 em 12 meses), tendo-se obtido resultados mais satisfatórios, transformando-se a série DRC numa nova série. Deste modo as novas séries que serviram de base a todo o estudo foram: DRN no domínio logarítmico e DRC no domínio logarítmico com a aplicação de diferenças, uma simples e outra sazonal.

Para cada uma das situações descritas anteriormente realizaram-se 250 sessões de treino, seleccionando os resultados da melhor sessão de treino e escolhendo as RNA com melhores resultados no conjunto de validação, para cada uma das séries. De referir ainda que foi utilizado o conjunto de validação, para cada uma das séries, para interromper as iterações de aprendizagem quando a *performance* nesse conjunto não melhorava ao fim de 5 iterações sucessivas. A realização de várias sessões de treinos justifica-se porque em cada sessão de treino os valores iniciais dos pesos são diferentes, chegando-se a soluções também diferentes, podendo ter performances significativamente diferentes. O critério que esteve na base para escolher o melhor modelo, para cada uma das séries em análise, foi a raiz do erro quadrático médio (REQM<sup>6</sup>) na comparação dos resultados obtidos pela rede com os valores observados.

As diferentes escolhas experimentadas e descritas nos parágrafos anteriores tiveram por base os trabalhos de investigação realizados por Faraway e Chatfield (1998), Thawornwong e Enke (2004), Fernandes (2005) e Fernandes e Teixeira (2007).

## 2.4. Previsão da Procura Turística: Análise dos Resultados

Nesta secção vão ser analisados os resultados para o conjunto de teste (anos 2005 e 2006), comparando os valores observados com os valores previstos, para as duas séries e utilizando as duas metodologias. Posteriormente, foram também analisadas as previsões

---

<sup>6</sup>

$$REQM = \sqrt{\frac{\sum_{t=1}^n (A_t - P_t)^2}{n}}; \text{ com } : A_t, \text{ valores reais no período } t; P_t, \text{ valores previstos no período } t; n, \text{ total de observações utilizadas.}$$

produzidas para os anos de 2005 e 2006 e comparadas com as dormidas mensais ocorridas ao longo dos mesmos. De referir que a previsão para os meses dos anos de 2005 e de 2006 foi realizada não utilizando na entrada nenhum valor observado relativo ao ano em causa. Em sua vez são usados nas entradas correspondentes aos meses desse ano os valores previstos anteriormente para esse ano. As equações [4] e [5] foram as utilizadas para o cálculo das previsões para cada uma das metodologias utilizadas, Box-Jenkins e Redes Neurais Artificiais, respectivamente, tendo por base ainda o processo inverso das transformações efectuadas.

Com esta análise pretendeu-se verificar se os modelos encontrados continuam a acompanhar as oscilações das séries e a produzir previsões aceitáveis para a Procura Turística, para as regiões em estudo.

Assim, no sentido de observar se o modelo escolhido produz erros de previsão aceitáveis vão ser calculados os seguintes critérios dos erros de previsão: erro percentual absoluto (EPA) e o erro percentual absoluto médio (EPAM), dados pelas equações:

$$EPA = \left| \frac{Y_t - P_t}{Y_t} \right|; Y_t, \text{valor observado e } P_t, \text{valor previsto.} \quad [10]$$

$$EPAM = \frac{1}{n} \sum_{t=1}^n \left| \frac{Y_t - P_t}{Y_t} \right|; Y_t, \text{valor observado e } P_t, \text{valor previsto.} \quad [11]$$

Como critério de análise da qualidade dos valores previstos com cada um dos modelos teve-se por base a classificação do MAPE<sup>7</sup>, ou EPAM, proposto por Lewis (1982) e que se apresenta na tabela seguinte.

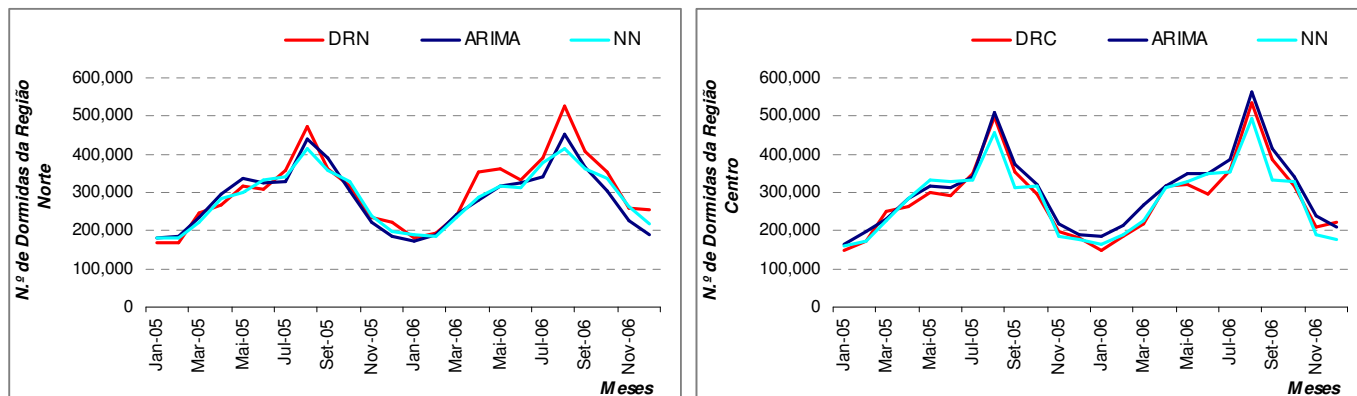
**Tabela 2.** Critério MAPE para Avaliação de um Modelo, Lewis (1982).

MAPE (%)	Classificação das Previsões
<10	Precisão Elevada
10-20	Boa Precisão
20-50	Razoável Precisão
>50	Pouco Fiáveis

No sentido de avaliar a capacidade preditiva do modelo, realizaram-se previsões para os anos de 2005 e 2006, que podem ser observadas na Figura 6 e Tabela A.3, em Anexo. Efectuando uma análise a esta figura, pode observar-se que os valores estimados, pelos modelos, acompanham o comportamento da série original, ou seja, os modelos obtidos conseguem acompanhar as oscilações da série do número de Dormidas Mensais quer para

<sup>7</sup> Do inglês *Mean Absolute Percentage Error*.

a Região Norte quer para a Região Centro de Portugal. No entanto, para ambas as regiões, verificou-se algum afastamento, significativo, nalguns meses entre os valores previstos e os realmente observados o que permite dizer que o modelo não conseguiu incorporar alguns factos decorridos nos anos em análise.



**Figura 6.** Representação gráfica das Dorminas Mensais, para as Regiões, para os anos 2005 e 2006.

Na Tabela 3 apresentam-se os valores do erro percentual absoluto (EPA) e o erro percentual absoluto médio (EPAM). Da análise dos valores de erro e tendo por base também os critérios estabelecidos por Lewis (1982) apresentados na Tabela 2, poder-se-á dizer que os modelos conseguiram produzir previsões de precisão elevada para o ano de 2005, uma vez que o EPAM apresenta valores inferiores a 10%, para cada um dos modelos. No entanto, para o ano de 2006 enquanto o modelo das Redes Neurais Artificiais continuou apresentar valores muito satisfatórios para ambas as regiões, inferiores a 10%, o mesmo não acontece quando se analisam os valores do modelo ARIMA. Embora apresentando valores satisfatórios, que se enquadram no intervalo que permite classificar que as previsões são de 'Boa Precisão', quando comparados com os do modelo das Redes Neurais Artificiais, os mesmos ficaram ligeiramente agravados. Calculando-se o EPAM, para o conjunto de teste (integrando os anos de 2005 e 2006), para cada uma das regiões, observou-se para a Região Norte que o modelo ARIMA apresentou um valor de 9,39% e o modelo de Rede Neuronal Artificial de 7,79%. Valores semelhantes foram também produzidos para a Região Centro, 9,48% e 7,80%, para o modelo ARIMA e de Redes Neurais Artificiais, respectivamente. Este facto revela-se interessante, dado que por exemplo os modelos de redes neuronais artificiais construídos para cada uma das regiões tiveram pré-processamentos diferentes, embora tendo-se utilizado a mesma rede. Será interessante continuar aplicar esta metodologia em estudos futuros, no sentido de se observar se os modelos construídos continuam a apresentar o mesmo comportamento.

De referir ainda que o facto da existência de alguns valores registados para o EPA, para os anos 2005 e 2006 e para ambas as regiões, serem superiores a 10% e a 20%,

consequência dos modelos terem revelado alguma dificuldade em fazer boas previsões, quando surgiram acontecimentos que fizeram alterar significativamente os valores observados, embora continuando a classificar-se como previsões fiáveis. Estes factos podem ser consequência por exemplo da elevada promoção que se tem vindo a fazer às regiões em análise junto de mercados internacionais, também as autarquias locais têm vindo a apostar mais na divulgação e realização de eventos culturais, a realização de feiras comerciais temáticas, entre outros eventos. Para a região Norte tem-se investido na divulgação de alguns destinos turísticos como Parque Natural do Douro Internacional e o Alto Douro Vinhateiro, e na região Centro a apostada tem-se centrado na divulgação a par com algum investimento na criação de melhores condições à prática de desportos de Inverno, nomeadamente a prática do *sky* e do *snow board*, que vão atraindo pessoas à região, essencialmente nos meses de Inverno. Todos estes factores, já que não foram incorporados nos modelos, leva a que os mesmos tenham alguma dificuldade em produzir as previsões que conduzam a um EPA muito reduzido, pelo que existe a necessidade de criar mecanismos que permitam minimizar os erros, como por exemplo trabalhar com variáveis de intervenção.

**Tabela 3.** EPA e EPAM, para os anos 2005 e 2006, para as Regiões.

Meses	Região Norte				Região Centro			
	2005		2006		2005		2006	
	ARIMA (EPA)	RNA (EPA)	ARIMA (EPA)	RNA (EPA)	ARIMA (EPA)	RNA (EPA)	ARIMA (EPA)	RNA (EPA)
<b>Janeiro</b>	7.4%	8.5%	3.9%	4.8%	11.9%	8.2%	25.8%	12.6%
<b>Fevereiro</b>	11.2%	9.0%	3.8%	5.8%	15.9%	1.1%	14.9%	1.9%
<b>Março</b>	4.5%	10.7%	2.3%	1.1%	8.8%	9.9%	21.3%	2.5%
<b>Abril</b>	10.0%	6.0%	20.2%	18.9%	7.4%	8.1%	0.2%	1.9%
<b>Mai</b>	5.8%	5.0%	12.6%	12.4%	6.1%	10.5%	9.6%	2.6%
<b>Junho</b>	5.8%	8.5%	2.3%	5.8%	6.6%	12.7%	19.0%	19.4%
<b>Julho</b>	7.8%	5.0%	11.9%	3.2%	1.5%	4.9%	8.0%	1.8%
<b>Agosto</b>	6.9%	11.8%	14.1%	21.0%	2.3%	8.0%	5.6%	8.0%
<b>Setembro</b>	7.5%	1.4%	9.7%	10.6%	6.2%	11.3%	7.3%	14.8%
<b>Outubro</b>	4.1%	4.0%	13.4%	4.6%	8.6%	7.4%	7.3%	4.0%
<b>Novembro</b>	5.4%	1.8%	13.0%	2.0%	8.9%	6.5%	13.5%	9.0%
<b>Dezembro</b>	15.8%	11.0%	26.1%	14.2%	5.8%	0.8%	4.8%	19.5%
<b>EPAM</b>	<b>7.7%</b>	<b>6.9%</b>	<b>11.1%</b>	<b>8.7%</b>	<b>7.5%</b>	<b>7.4%</b>	<b>11.4%</b>	<b>8.2%</b>

Pela análise efectuada anteriormente, observou-se que existe uma diferença reduzida entre os valores obtidos para o EPAM, com os dois modelos construídos com as diferentes metodologias e para ambas a regiões. No entanto, pode inferir-se que os modelos da Redes Neurais Artificiais apresentaram qualidades estatísticas e de ajustamento satisfatórias, evidenciando serem adequados para a modelação e previsão da série de referência, quando comparados com os modelos produzidos pela metodologia de Box-Jenkins, ou seja, a metodologia das Redes Neurais Artificiais podem ser consideradas uma alternativa à metodologia clássica de Box-Jenkins, na análise da Procura Turística.

### 3. Competitividade entre os Destinos Turísticos Norte e Centro de Portugal

#### 3.1. Preferência dos Destinos Turísticos

A capacidade de atracção de alguns destinos turísticos e a sua acessibilidade são preferidos a outros pelas correntes turísticas, por isso importa avaliar cada destino turístico em termos das preferências que exercem sobre os movimentos turísticos e a sua evolução temporal, (Cunha, 2003).

Mariotti formulou uma teoria sobre os centros de atracção turística que permite analisar em que medida um determinado destino turístico é mais preferido relativamente aos restantes e de que modo estes se comportam perante a evolução da procura turística. Este grau ou índice de preferência (*IP*) determina-se relacionando o número de turistas/hóspedes, chegados a uma determinada localidade, ao longo de um período (*T*) e o número total de turistas/hóspedes, registados em todos os estabelecimentos do país, durante o mesmo período (*Tt*) e pode ser traduzido mediante a seguinte expressão:

$$IP = \frac{T}{Tt} \quad [12]$$

Paralelamente à análise da preferência turística importa observar a importância relativa do turismo em cada região e a sua capacidade para suportar acréscimos adicionais da procura turística, pelo que este indicador é obtido pela relação entre o número de turistas/hóspedes que visita um destino e a respectiva população total e designa-se por índice de saturação turística (*IST*), (Cunha, 1997, 2003), dado pela expressão:

$$IST = \frac{N.^{\circ}Turistas}{PopulaçãoTotal} \quad [13]$$

De referir que um destino com um índice de saturação turística inferior a um outro tem uma importância turística menor mas possui maior capacidade de crescimento potencial (Cunha, 2003).

Assim, seguindo a linha de raciocínio apresentada anteriormente e aplicando as fórmulas [12] e [13], na Tabela 4 pode observar-se os resultados obtidos para os dois indicadores que permitem efectuar uma avaliação da procura turística e de alguns dos efeitos que a mesma provoca. Saliente-se que as conclusões, que serão apresentadas ao longo deste ponto, apenas são indicativas da tendência que se tem vindo a assistir nos últimos anos, atendendo ao período em análise - 1997 e 2006.

Analisando os valores, para o índice de preferência, apresentados na tabela seguinte, e comparando os dois destinos turísticos - norte e centro - verifica-se a preferência dos turistas pela região Norte, em 1997, mantendo-se essa mesma preferência passados 9

anos. Poder-se-á dizer que a região Norte tem vindo a exercer uma atracção mais eficaz sobre os turistas enquanto que a região Centro perdeu alguma capacidade de atracção sobre os mesmos, situação que tende a modificar-se uma vez que se registou um aumento no índice de preferência. De salientar que a diferença entre as duas regiões tende a atenuar-se levando a acreditar que os turistas procuram a tranquilidade do campo, as riquezas termais, a cultura, a redescoberta da natureza, e a procura por uma oferta turística direccionada para a saúde e bem-estar, que são mais-valias encontrada nestes destinos.

**Tabela 4.** Índice de Preferência e Índice de Saturação Turística, para as regiões em análise.

	DESTINOS TURÍSTICOS			
	PORTUGAL	CONTINENTE	REGIÃO NORTE	REGIÃO CENTRO
<b>Ano de 1997</b>				
Hóspedes (n.º)	8.751.547	7.887.056	1.456.891	967.967
População* (n.º)	10.107.916	9.623.101	3.585.454	1.752.829
<i>Índice de Preferência</i>	----	0.90	0.17	0.11
<i>Índice de Saturação Turística</i>	0.87	0.82	0.41	0.55
<b>Ano de 2006</b>				
Hóspedes (n.º)	12.376.900	10.975.400	2.144.000	1.874.400
População* (n.º)	10.599.095	101.102.71	3.744.341	2.385.891
<i>Índice de Preferência</i>	----	0.89	0.17	0.15
<i>Índice de Saturação Turística</i>	1.17	1.09	0.57	0.79

**Nota:** \* Estimativas da População Residente aferidas para os resultados dos Censos 2001.

Na Tabela 4, pode-se ainda observar que relativamente ao índice de saturação turística em 1997, nenhum destino apresentou valores superiores a um, contudo em 2006 quer para Portugal quer para o Continente o índice registou valores superiores a um implicando um certo impacto negativo do ponto de vista social e ambiental. Quanto às regiões Norte e Centro, ambas apresentam valores inferiores a um, embora a região Centro registasse de 1997 para 2006 um agravamento do índice que poderá provocar uma menor satisfação do turista nesta região. Para fazer face a esta situação as entidades governamentais locais ou regionais devem implementar estratégias que levem a que se aumente o número de turistas sem provocar impactos negativos para a região, tais como: (i) criar animação cultural através de programas e actividades desenvolvidas nos principais centros históricos da região, recuperando tradições antigas e introduzindo actividades mais recentes e modernas que atraiam os jovens, podendo passar também pela introdução de feiras temáticas e artistas de rua; (ii) renovar as vantagens competitivas da oferta turística, enriquecendo-a com produtos mais qualificados e elevando os níveis de qualidade de serviço, nomeadamente apostando numa vertente turística direccionada para a saúde e bem-estar; (iii) estimular e valorizar a oferta de produtos turísticos que correspondem a potencialidades da Região, designadamente a segmentos da procura ligados ao turismo desportivo (golfe, desportos náuticos, náutica de recreio, sky), ao turismo de natureza (incluindo o segmento

ligado ao turismo científico) e ao turismo de congressos, incentivos e de negócios; (iv) dinamizar e enriquecer o calendário de animação turística e cultural, nomeadamente através do desenvolvimento de actividades de animação com conteúdos inovadores e duração mais alargada; (v) participar em feiras anuais, nos mercados mais promissores; *workshops*; sessões de informação e seminários dirigidos a profissionais do sector e criadores de opinião.

### 3.2. Análise da Quota de Mercado

O cálculo da quota de mercado muitas das vezes é o processo mais comum para analisar a competitividade entre diferentes destinos turísticos, utilizando diferentes factores que a explique, tais como: índice de preferência, o nível de preços relativos, o custo da viagem física e financeira, pressão competitiva exercida pela publicidade dos destinos concorrentes, entre outros (Fernandes & González, 2007; Fernandes & Teixeira, 2007b). No entanto, e porque nem sempre é fácil obter dados estatísticos para as diferentes variáveis, na medida em que posteriormente se possa efectuar uma análise comparativa, para os diferentes destinos turísticos, opta-se por utilizar uma variável (dormidas) em que a sua contabilização está uniformizada. Estes valores absolutos, deverão transformar-se em quotas de mercado, para que se possa proceder a uma comparabilidade.

Faulker (1997) desenvolveu um instrumento de análise designado por Análise da Quota de Mercado [AQM], para avaliar o desempenho e a competitividade dos destinos turísticos. Assim, tendo como base a avaliação dos mercados emissores, a AQM é analisada sob duas dimensões:

- (i) um índice de desvio da quota de mercado, designado por [B] (eq.[14]);
- (ii) um índice de variação da quota de mercado, designado por [C] (eq.[15]).

Relativamente ao índice [B], este é estático e na óptica de análise do destino, reflecte a comparação da quota de mercado do destino *i* do mercado emissor *j* com a quota de mercado média do destino *i* no conjunto de origens em análise num determinado momento. Quanto ao segundo índice, [C], é um indicador dinâmico que na óptica de análise do destino relaciona, num determinado intervalo de referência, a taxa de crescimento dos fluxos turísticos a partir do mercado emissor *j* para o destino *i* com a taxa de crescimento dos fluxos turísticos a partir do mercado emissor *j* para o conjunto dos destinos concorrentes em estudo, num determinado período de análise (Faulkner, 1997; Águas *et al.*, 2003; Fernandes & González, 2007).

Para o cálculo dos índices Faulkner (1997) propôs as seguintes expressões:

$$B_{ijk} = \frac{\left( X_{ijk} / \sum_{i=1}^n X_{ijk} \right)^{-1}}{\left( \sum_{j=1}^n X_{ij(k)n} / \sum_{i=j}^n \sum_{j=i}^n X_{ij(k)n} \right)} \quad [14]$$

onde,

$B_{ik}$ , índice de desvio da quota de mercado para o destino  $i$  no ano  $k$ ;

$X_{ijk}$ , número de visitantes para o destino  $i$  da origem/mercado emissor  $j$ , no ano  $k$ ;

$n$ , número de mercados emissores/origens e destinos.

$$C_{ijk} = \left[ \left( X_{ijk} / X_{ij1} \right) - 1 \right] - \left[ \left( X_{jk} / X_{j1} \right) - 1 \right] \quad [15]$$

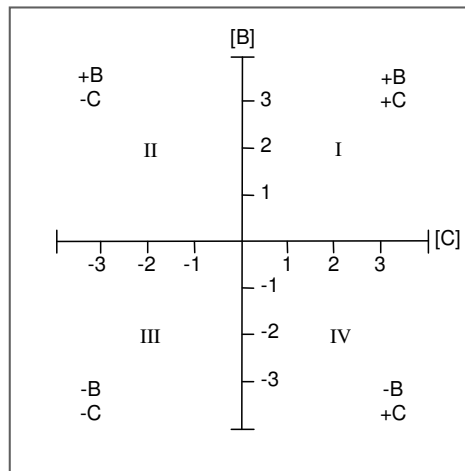
com,

$C$ , índice de variação da quota de mercado;

$X_{ijk}$ , número de visitantes para o destino  $i$  da origem/mercado emissor  $j$ , no ano  $k$ ;

$X_{jk}$ , total número de visitantes do mercado emissor/origem  $j$  no ano  $k$ ;

$1 \dots k$ , do ano 1 ao ano  $k$ .



**Figura 7.** Tipologias de Quotas de Mercado, (Adaptado de Faulkner, 1997:29).

Deste modo, de acordo com os valores obtidos para cada um dos índices, estes podem ser representados graficamente num sistema de eixos ortogonal, em que [C] representado no eixo horizontal e [B] no eixo vertical, permitindo a determinação de quatro quadrantes (Figura 7).

Tendo por base o clássico modelo do ciclo-de-vida do produto pode ainda ser estabelecida uma correspondência, que permite identificar os seguintes mercados apresentando as seguintes características (Faulkner, 1997; Águas *et al.*, 2003; Fernandes & González, 2007):

- i. Mercados em crescimento (Quadrante I: + B; + C; Sucesso), origens em que o destino já detém uma quota de mercado acima da média e em que se encontra a ganhar quota de mercado;
- ii. Mercados em maturidade (Quadrante II: + B; - C; Incerteza), origens em que o destino detém uma quota de mercado acima da média mas, em que se começa a perder quota de mercado;
- iii. Mercados em declínio (Quadrante III: - B; - C; Insucesso), origens em que o destino goza de uma quota de mercado abaixo da média e em que se encontra a perder quota de mercado;
- iv. Mercados emergentes (Quadrante IV: - B; + C; Entrada), origens em que o destino desfruta de uma quota de mercado abaixo da média e em que se encontra a ganhar quota de mercado.

Antes de proceder com a respectiva análise comparativa entre os dois destinos turísticos em estudo - região Norte e região Centro -, cabe referir que para o cálculo da AQM entrou-se em linha de conta com os seguintes elementos:

[a] horizonte temporal: 1997 e 2006;

[b] variável em análise: dormidas efectuadas nos estabelecimentos hoteleiros;

[c] mercado receptor/destino turístico: Região Norte e Região Centro;

[d] mercados emissores/origens: optou-se, com base nos dados disponíveis pelo INE, por considerar os países da União Europeia que têm um peso significativo na procura turística de Portugal, sendo eles: Alemanha (A), Espanha (E), França (F), Itália (I), Países Baixos (PB), Portugal (P) e Reino Unido (RU).

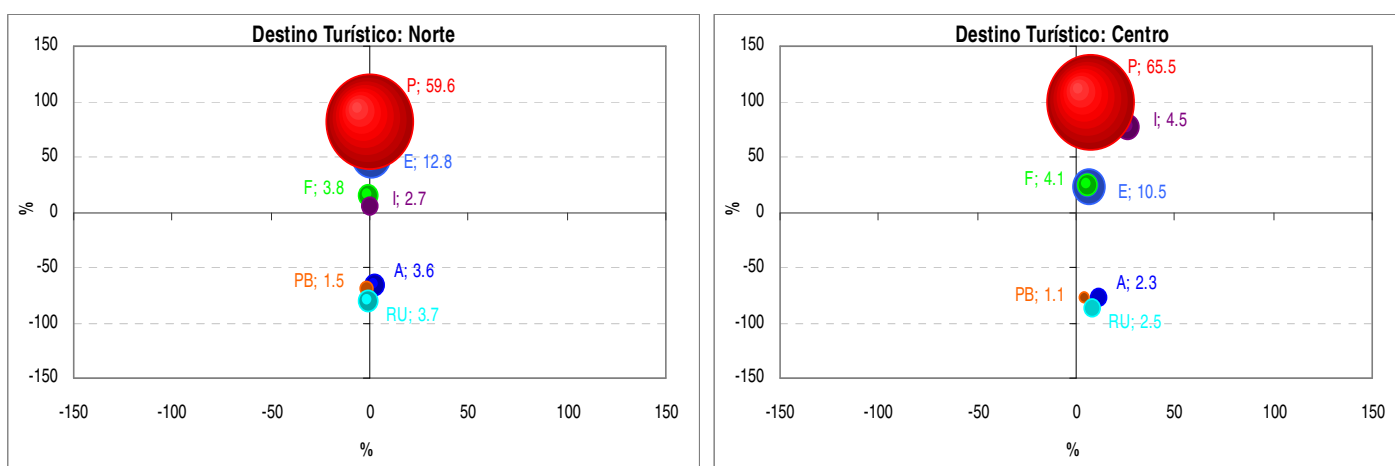
Os resultados obtidos para o Desvio de Quota de Mercado e para a Variação da Quota de mercado, apresentam-se nas Tabela 7, o qual constitui as coordenadas para os pontos origens/destinos representado na Figura 8, permitem observar a evolução do destino por principais origens.

**Tabela 7.** Tabela Resumo da AQM, para as Regiões Norte e Centro.

Dormidas_1997		Dormidas_2006		Mercado Emissor	Índice B (%)		Índice C (%)	
Norte (n.º)	Centro (n.º)	Norte (n.º)	Centro (n.º)		Norte	Centro	Norte	Centro
140.117	38.164	136.900	81.300	<b>Alemanha</b>	-65.4	-77.5	2.59	11.61
232.856	109.034	491.600	367.500	<b>Espanha</b>	50.4	23.2	0.46	6.26
114.027	62.268	145.500	144.100	<b>França</b>	14.6	24.3	-1.12	5.90
73.386	15.626	102.500	157.500	<b>Itália</b>	5.1	76.9	0.23	25.72
57.875	23.227	56.400	38.000	<b>Países Baixos</b>	-69.3	-77.3	-2.01	3.90
1.626.512	857.846	2.292.400	2.297.600	<b>Portugal</b>	81.4	99.2	-0.35	7.33
126.673	36.558	143.600	87.800	<b>Reino Unido</b>	-80.7	-87.0	-0.52	8.30

Assim, da análise à Figura 8 pode observar-se, para os dois destinos turísticos - Norte e Centro -, uma elevada dependência da origem Portugal e um desvio de quota de mercado positivo nos mercados emissores geograficamente mais próximos (Portugal, Espanha, França e Itália). Para o destino turístico - região Norte - onde apenas os mercados emissores Espanha e Itália é que se encontram no quadrante de sucesso e as origens França e Portugal encontram-se no quadrante de incerteza. Ainda, este destino turístico face aos mercados emissores Países Baixos e Reino Unido considera-se em declínio. Relativamente à origem Alemanha, embora o destino turístico possua uma quota de mercado abaixo da média mas encontra-se a ganhar quota de mercado, considerando-se assim um mercado emergente. Dado que as principais origens, deste destino, se encontram no segundo quadrante, esta região face a estes mercados emissores considera-se de incerteza, ou seja, mercados emissores em que o destino detém uma quota de mercado acima da média mas, em que se começa a perder quota de mercado.

Efectuando a mesma análise para o destino turístico - região Centro -, este apresenta resultados mais satisfatórios, quando comparados com o destino Norte, pois os mercados emissores (Portugal, Espanha, Itália e França) com maior peso posicionam-se no quadrante de sucesso, isto é, mercados emissores em que o destino já detém uma quota de mercado acima da média e em que se encontra a ganhar quota de mercado, considerando-se por isso um mercado em crescimento. As demais origens, Alemanha, Reino Unido e Países Baixos, encontram-se no quadrante de entrada, ou seja, o destino turístico é um mercado emergente para estes mercados emissores.



**Figura 8.** Análise das Origens: Destinos Turísticos Norte e Centro.

Comparando os dois destinos turísticos, poder-se-á referir que o destino Centro é uma região promissora, uma vez que existe um predomínio das situações de ganho de quota de mercado. Esta situação poderá ser reflexo do desenvolvimento, por parte das autarquias locais, de um conjunto de acções no domínio das infra-estruturas turísticas, acessibilidades

e equipamentos de animação turística, desportiva e de lazer; de valorização do património histórico e da oferta cultural.

#### **4. Conclusão e Desenvolvimentos Futuros**

Tal como tem acontecido noutros países, em que o turismo tem sido uma actividade geradora de riqueza e cada dia que passa tem assumido um papel significativo na economia do país, Portugal também não fica alheio a este facto.

Neste contexto, também as organizações, públicas ou privadas, particularmente ligadas ao sector turístico, implementadas nas regiões em estudo (Regiões Norte e Centro de Portugal) devem preocupar-se em construir mecanismos que lhes permitam obter uma visão antecipada da evolução da procura turística, no sentido de criar condições favoráveis aos visitantes que se deslocam a esses destinos turísticos.

Deste modo, pretendia-se com a presente investigação explorar e evidenciar a utilidade da metodologia das RNA como uma alternativa à metodologia de Box-Jenkins; e construir modelos, com estas duas metodologias, que permitissem analisar e prever o comportamento da Procura Turística para as Regiões em estudo. Apresentaram-se e analisaram-se os dados avançados sobre futuros fluxos turísticos, nacionais e internacionais, de dormidas de turistas, para os anos de 2005 e 2006 e compararam-se com os valores efectivamente observados. Relativamente ao modelo construído com a metodologia de Box-Jenkins, para as duas regiões em análise, constatou-se que o modelo  $ARIMA(0,1,1) \times (0,1,1)_{12}$ , foi o que melhor se ajustou ao comportamento das séries de referência, para ambas as regiões, e permitiu produzir previsões para a variável procura turística. Os modelos construídos, embora com pré-processamentos distintos, com a metodologia das RNA, tiveram por base uma estrutura do tipo *feedforward*, foram treinadas com o algoritmo de *Resilient Backpropagation* tendo-se utilizado a função de activação sigmoidal logística, com 4 neurónios na camada oculta. Cada valor da série depende directamente dos doze valores precedentes. As previsões foram efectuadas mensalmente. Os modelos obtidos com a metodologia das RNA apresentaram valores bastante satisfatórios, seguindo de perto o comportamento das séries que estiveram na base deste estudo.

Assim, e dada análise efectuada, concluiu-se que os modelos obtidos, para as duas metodologias e para ambas as regiões, são válidos para o conjunto de dados que lhe serviu de suporte e apresentaram qualidades estatísticas e de ajustamento satisfatórias evidenciando serem adequados para a modelação e previsão das séries de referência. Embora, os modelos construídos com a metodologia das RNA terem evidenciado ser

superiores aos construídos com a metodologia de Box-Jenkins, permitiu inferir que podem ser uma alternativa à metodologia de Box-Jenkins. Dado que os modelos revelaram alguma dificuldade em fazerem boas previsões em presença de alguns acontecimentos sugere-se a sua inclusão futura no modelo utilizando-se para tal, por exemplo variáveis de intervenção. Este é um desafio que os autores se propõem fazer em futuras investigações, no sentido de se conseguirem obter previsões mais próximas das que venham a ser registadas e assim conseguir uma maior precisão dos modelos.

Foi ainda propósito do presente estudo avaliar o desempenho dos destinos turísticos - Região Norte e Região Centro - por principais mercados emissores e analisar como se encontra distribuída as suas carteiras de mercados emissores, para o período de 1997 a 2006. Pela Análise da Quota de Mercado observou-se, para os dois destinos turísticos uma elevada dependência da origem Portugal e um desvio de quota de mercado positivo nos mercados emissores geograficamente mais próximos (Portugal, Espanha, França e Itália). Verificou-se que a Região Centro apresentou resultados mais satisfatórios, quando comparados com o destino Norte, tendo-se concluído que é uma região promissora, uma vez que existe um predomínio das situações de ganho de quota de mercado. Esta situação poderá ser reflexo do desenvolvimento, por parte das autarquias locais, de um conjunto de acções no domínio das infra-estruturas turísticas, acessibilidades e equipamentos de animação turística, desportiva e de lazer; de valorização do património histórico e da oferta cultural.

## Referências Bibliográficas

- Águas**, Paulo; **Grade**, Ana e **Sousa**, Peter. (2003). “Competitividade Turística Regional: Avaliação nos Principais Mercados, 1991-2001”. *Revista Portuguesa de Estudos Regionais*. N.º1, pp.29/52.
- Bishop**, C. M.. (1995). “*Neural Networks for pattern recognition*”. Oxford University Press. Oxford. London.
- Chu**, Fong-Lin. (1998). “Forecasting Tourist Arrivals: nonlinear sine wave or ARIMA?”. *Journal of Travel Research*. Vol. 36; pp.79/84.
- Cunha**, Licínio. (1997). “*Economia e Política do Turismo*”. Editora McGraw-Hill de Portugal, Lda.
- Cunha**, Licínio. (2003). “*Introdução ao Turismo*”. Editorial Verbo, Lisboa/São Paulo.
- Faraway**, Julian and **Chatfield**, Chris. (1998). “Time series forecasting with neural networks: a comparative study using the airline data”. *Applied Statistics*. N.º47, pp.231/250.
- Faulkner**, B. (1997). “A Model for the Evaluation of National Tourism Destination Marketing Programs”. *Journal of Travel Research*. N.º35 (3), pp. 23/32.
- Fernandes**, Paula O. e **Cepeda**, Francisco J.T.. (2000). “Aplicação da Metodologia de Box-Jenkins à Série Temporal de Turismo: Dormidas Mensais na Região Norte de Portugal”. *Actas do VII Congresso da Associação Portuguesa de Desenvolvimento Regional*. Volume I. Universidade dos Açores; Ponta Delgada, Açores, Portugal; pp. 261/272. ISBN:972-97825-8-X.
- Fernandes**, Paula O. and **Teixeira**, João Paulo. (2007a). “A new approach to modelling and forecasting monthly overnights in the Northern Region of Portugal”. *Proceedings of the 4th International Finance Conference*. Université de Cergy; Hammamet, Medina, Tunísia.
- Fernandes**, Paula O. e **Teixeira**, João Paulo. (2007b). “Comportamento da Procura Turística: Região Norte de Portugal”. *IX Seminário Luso-Espanhol de Economia Empresarial; Inovação e Educação: Ensino do Empreendedorismo*. ISBN:978-972-8790-76-9.
- Fernandes**, Paula O. y **González**, Luis B.. (2007). “Análisis de la Cuota del Mercado Turístico en Portugal”. *Actas XVII Jornadas Hispano-Lusas de Gestión Científica, Conocimiento, Innovación y Emprendedores: Camino al Futuro*. Universidad de la Rioja; Logroño, Espanha; Editora Ayala Calvo, J.C. y grupo de investigación FEDRA. pp. 2444/2455. ISBN:978-84-690-3573-3.
- Fernandes**, Paula Odete. (2005). “*Modelling, Prediction and Behaviour Analysis of Tourism Demand in the North of Portugal*”. Ph.D. Thesis in Applied Economy and Regional Analysis. Valladolid University - Spain.
- González**, Carlos Sánchez. (1999). “*Métodos Econometricos*”. 1ª Edición; Editorial Ariel. Barcelona.
- Hansen**, J. V., **Mcdonald**, J. B. and **Nelson**, R. D. (1999). “Time series prediction with genetic-algorithm designed neural networks: an empirical comparison with modern statistical models”. *ComputIntell*. N.º15, pp. 171/184.
- Haykin**, Simon. (1999). “*Neural Networks. A comprehensive foundation*”. New Jersey, Prentice Hall.
- Hill**, T.; **O’connor**, M. and **Remus**, W. (1996). “Neural network models for time series forecasts”. *Management Science*. Vol. 42 (7), pp. 1082/1092.
- INE**. Anuários Estatísticos do Turismo de 1987 e 2006. Lisboa.

- Lewis**, C.D. (1982). *“Industrial and Business Forecasting Method”*. Butterworth Scientific. London.
- Murteira**, Bento J.F.; **Müller**, Daniel A. e **Turkman**, K. Feridun. (1993). *“Análise de sucessões cronológicas”*. McGraw-Hill; Lisboa.
- Otero**, José M<sup>a</sup>. (1993). *“Econometría - series temporales y predicción”*. Editorial AC; Madrid.
- Parra**, S. B. y **Domingo**, J. U.. (1987). “Análisis de series temporales de turismo de la Comunidad Valenciana”. *Estadística Española*. N.º114, pp.111/132.
- Pulido**, Antonio. (1989). *“Predicción Económica y Empresarial”*. Ediciones Pirámide; Madrid.
- Rodrigues**, Pedro João S.. (2000). *“Redes neuronais aplicadas à segmentação e classificação de leucócitos em imagens”*. Dissertação de Mestrado em Engenharia Electrónica e Telecomunicações. Universidade de Aveiro.
- Rumelhart**, D. E. and **McClelland**, J. L.. (1986). *“Parallel Distributed Processing: Explorations in the Microstructure of Cognition”*. Volume 1: Foundations. The Massachusetts Institute of Technology Press, Cambridge.
- Thawornwong**, S. and **Enke**, D. (2004). “The adaptive selection of financial and economic variables for use with artificial neural networks”. *Neurocomputing*. N.º6, pp. 205/232.
- Witt**, Stephen F. and **Witt**, Christine A.. (1995). “Forecasting tourism demand: a review of empirical research”. *International Journal of Forecasting*. N.º 11, pp.447/475.
- Wong**, K. F.. (2002). “Introduction: Tourism Forecasting State of the Art”. *Journal of Travel and Tourism Marketing*; N.º 13 (1/2), pp.1/3.
- Yu**, Gongmei and **Schwartz**, Zvi. (2006). “Forecasting Short Time-Series Tourism Demand with Artificial Intelligence Models”. *Journal of Travel Research*. N.º 45, pp. 194/203.
- Zou**, Hui and **Yang**, Yuhong. (2004). “Combining time series models for forecasting”. *International Journal of Forecasting*. Vol. 20, pp. 69/84.

## ANEXO A

**Tabela A.1. Valores da Série Original, para o período entre 1987:01 e 1996:12, Região Norte (cont.).**

MESES \ ANOS	ANOS									
	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996
Janeiro	102.447	118.011	122.217	126.671	126.826	124.194	121.469	118.606	122.480	126.910
Fevereiro	102.123	117.547	116.837	129.802	131.653	127.474	129.284	122.988	130.393	139.403
Março	125.401	142.687	160.658	158.701	188.999	157.536	154.734	175.261	156.645	172.393
Abril	150.042	167.118	169.326	197.757	182.290	196.087	189.142	185.525	209.263	213.973
Mai	180.430	189.823	199.158	207.876	219.187	223.918	198.402	232.075	218.666	239.142
Junho	197.113	207.729	218.595	227.159	251.295	207.907	207.216	248.237	222.720	245.264
Julho	229.293	254.523	252.634	257.633	273.927	231.801	231.453	246.274	247.589	248.398
Agosto	304.847	315.113	329.014	351.500	341.490	312.026	304.576	322.366	320.750	336.086
Setembro	238.542	258.287	278.074	284.867	283.378	259.023	249.583	266.094	269.433	280.769
Outubro	173.503	174.359	189.664	216.286	197.241	205.400	202.792	206.256	196.466	225.734
Novembro	130.187	137.933	138.683	162.062	152.554	149.289	141.976	144.803	152.340	175.438
Dezembro	114.229	128.774	127.730	139.683	132.802	130.963	120.748	139.706	140.643	143.163
<b>TOTAL</b>	<b>2.048.157</b>	<b>2.211.904</b>	<b>2.302.590</b>	<b>2.459.997</b>	<b>2.481.642</b>	<b>2.325.618</b>	<b>2.251.375</b>	<b>2.408.191</b>	<b>2.387.388</b>	<b>2.546.673</b>

Fonte: Elaboração própria com base nos dados facultados pelo INE.

**Tabela A.1. Valores da Série Original, para o período entre 1997:01 e 2006:12, Região Norte.**

MESES \ ANOS	ANOS									
	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006
Janeiro	140.430	148.218	163.696	162.389	176.690	165.653	155.527	162.900	168.100	180.700
Fevereiro	141.183	157.415	165.988	162.637	186.586	181.005	177.818	181.900	166.800	195.100
Março	219.465	209.929	228.149	226.010	245.261	249.214	214.106	224.600	247.000	237.200
Abril	224.382	232.767	242.744	262.865	291.395	253.274	258.519	279.800	268.500	352.600
Mai	253.833	280.326	269.854	264.497	306.743	302.028	293.531	317.300	316.900	361.200
Junho	238.334	296.612	270.126	273.881	325.568	301.465	271.454	355.300	307.700	331.500
Julho	266.993	303.866	306.031	324.962	351.955	314.560	318.706	324.400	358.500	388.400
Agosto	345.672	377.645	385.868	397.405	452.581	444.991	433.211	426.900	472.400	524.500
Setembro	288.409	309.700	321.248	331.155	383.793	361.181	343.534	342.100	362.200	406.500
Outubro	232.052	263.522	280.597	263.217	319.417	287.383	281.472	311.500	315.900	353.300
Novembro	166.835	180.796	193.062	186.445	238.925	221.910	219.463	221.200	233.400	258.800
Dezembro	141.349	161.273	166.990	157.210	202.351	179.766	178.439	182.800	221.300	254.700
<b>TOTAL</b>	<b>2.658.937</b>	<b>2.922.069</b>	<b>2.994.353</b>	<b>3.012.673</b>	<b>3.481.265</b>	<b>3.262.430</b>	<b>3.145.780</b>	<b>3.330.700</b>	<b>3.438.700</b>	<b>3.844.500</b>

Fonte: Elaboração própria com base nos dados facultados pelo INE.

**Tabela A.2. Valores da Série Original, para o período entre 1987:01 e 1996:12, Região Centro (cont.).**

ANOS MESES	ANOS									
	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996
Janeiro	48.413	53.251	60.593	66.389	67.712	72.006	73.457	69.142	70.798	69.186
Fevereiro	53.932	66.257	70.923	78.898	81.963	78.873	82.466	80.463	81.326	89.418
Março	67.949	84.982	118.949	91.836	114.931	98.200	93.210	101.582	104.727	110.697
Abril	88.730	97.751	88.999	121.039	112.756	124.425	125.441	113.765	139.292	145.682
Mai	103.595	112.881	122.323	125.580	130.316	141.334	127.772	125.687	133.419	142.172
Junho	111.331	120.029	126.325	138.110	140.715	121.020	122.687	125.656	130.530	141.044
Julho	154.594	167.631	182.117	183.161	175.843	163.168	158.791	166.728	164.749	166.283
Agosto	233.117	240.183	263.974	259.879	267.754	247.192	247.527	250.555	242.433	241.940
Setembro	168.602	176.127	190.951	190.030	193.701	175.842	176.980	177.707	171.988	187.513
Outubro	106.730	107.174	118.864	127.891	123.425	121.295	118.980	116.944	116.247	137.972
Novembro	62.249	67.058	75.367	83.646	85.675	84.867	72.739	80.985	80.925	100.324
Dezembro	58.618	67.540	94.352	82.305	76.662	78.134	72.227	81.664	97.189	93.096
<b>TOTAL</b>	<b>1.257.860</b>	<b>1.360.864</b>	<b>1.513.737</b>	<b>1.548.764</b>	<b>1.571.453</b>	<b>1.506.356</b>	<b>1.472.277</b>	<b>1.490.878</b>	<b>1.533.623</b>	<b>1.625.327</b>

Fonte: Elaboração própria com base nos dados facultados pelo INE.

**Tabela A.2. Valores da Série Original, para o período entre 1997:01 e 2006:12, Região Centro.**

ANOS MESES	ANOS									
	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006
Janeiro	82.964	95.078	105.697	97.748	97.835	99.913	136.669	144.481	146.800	146.300
Fevereiro	95.439	106.779	123.941	112.210	117.057	118.807	146.512	169.494	172.000	184.500
Março	137.757	122.126	136.214	141.973	138.851	156.803	196.309	206.316	251.100	219.800
Abril	136.194	151.959	155.533	173.166	164.615	154.440	240.487	263.603	264.200	317.200
Mai	159.817	176.390	165.865	173.781	168.582	172.775	282.940	290.185	299.900	320.500
Junho	144.019	173.863	169.182	167.906	171.690	172.701	256.314	308.510	293.000	294.000
Julho	185.696	200.270	203.694	211.569	200.343	185.184	297.678	308.175	348.200	358.000
Agosto	262.815	294.081	280.780	296.264	287.122	288.336	439.293	442.413	496.700	534.200
Setembro	193.321	216.871	214.071	213.978	211.241	211.734	319.576	331.474	353.900	388.300
Outubro	147.357	162.655	161.856	162.932	163.283	158.020	257.783	300.534	294.200	316.100
Novembro	107.827	109.382	122.468	131.786	125.344	125.915	183.431	182.155	198.200	208.200
Dezembro	100.364	96.465	108.546	116.821	110.652	108.691	161.020	163.759	179.200	221.000
<b>TOTAL</b>	<b>1.753.570</b>	<b>1.905.919</b>	<b>1.947.847</b>	<b>2.000.134</b>	<b>1.956.615</b>	<b>1.953.319</b>	<b>2.918.012</b>	<b>3.111.099</b>	<b>3.297.400</b>	<b>3.508.100</b>

Fonte: Elaboração própria com base nos dados facultados pelo INE.

**Tabela A.3. Valores Previstos para os Modelos, para o período entre 2005:01 e 2006:12.**

	Região Norte				Região Centro			
	Ano 2005		Ano 2006		Ano 2005		Ano 2006	
	ARIMA	RNA	ARIMA	RNA	ARIMA	RNA	ARIMA	RNA
	Janeiro	180.579	173.626	182.389	189.349	164.330	158.907	184.061
Fevereiro	185.481	187.654	181.870	183.731	199.311	173.894	212.067	187.964
Março	235.924	242.683	220.635	234.591	228.927	226.225	266.561	225.229
Abril	295.228	281.367	284.692	285.916	283.693	285.479	317.873	311.094
Mai	335.197	315.652	301.171	316.248	318.323	331.353	351.116	328.840
Junho	325.419	323.895	333.732	312.298	312.205	330.240	349.908	351.012
Julho	330.532	342.132	340.731	376.036	342.810	331.239	386.499	351.520
Agosto	440.017	450.663	416.740	414.580	508.285	456.970	564.198	491.349
Setembro	389.361	367.067	357.019	363.306	375.764	313.744	416.813	330.920
Outubro	302.841	305.864	328.557	337.129	319.465	315.830	339.159	328.694
Novembro	220.912	225.089	237.594	264.057	215.853	185.341	236.406	189.486
Dezembro	186.379	188.159	196.989	218.612	189.648	177.854	210.464	177.856
<b>TOTAL</b>	<b>3.427.870</b>	<b>3.403.851</b>	<b>3.382.119</b>	<b>3.495.853</b>	<b>3.458.614</b>	<b>3.287.076</b>	<b>3.835.125</b>	<b>3.438.730</b>