



Modelação de Séries Temporais - Métodos Lineares e Não Lineares

Telmo Nuno Martins Machado

Dissertação apresentada à Escola Superior de Tecnologia e de Gestão de Bragança para obtenção do Grau de Mestre em Sistemas de Informação

Orientado por

João Paulo Teixeira e Paula Odete Fernandes

**Bragança
2009**



**Modelação de Séries Temporais – Métodos
Lineares e Não Lineares**

Telmo Nuno Martins Machado

Dissertação para obtenção do Grau de Mestre em
Sistemas de Informação

Dezembro de 2009

Resumo

O principal objectivo do presente trabalho assenta no estudo comparativo entre o modelo baseado na regressão linear e o modelo baseado na tecnologia de redes neuronais artificiais (RNA). A metodologia de regressão linear consiste na aplicação de um modelo univariado sobre a evolução de cada mês individualmente, ao longo do intervalo de tempo definido pela série estudada. Para o modelo de RNA foi estudada a inclusão de variáveis auxiliares.

Neste trabalho foi ainda desenvolvido uma interface gráfica para utilização destes modelos de previsão para utilizadores não especialistas.

A inclusão destas duas metodologias de natureza diferente, tem como finalidade, perceber a respectiva potencialidade aplicada às características peculiares que as séries de turismo evidenciam, tais como, a sazonalidade e a tendência. Recorreu-se para o efeito a séries mensais da procura turística: “Dormidas Mensais Registadas nos Estabelecimentos Hoteleiros - NDH” para Portugal, Região Norte e Região Centro, definidas no período compreendido entre Janeiro de 1990 e Dezembro de 2008. Foi também incluída neste estudo a série turística homóloga, referente ao Reino Unido. Esta série apresenta características ligeiramente diferentes das séries turísticas nacionais.

Os modelos desenvolvidos apresentaram qualidades estatísticas e de ajustamento de precisão elevada, pelo que se procedeu à sua utilização para efeitos previsionais. Neste contexto, efectuou-se uma comparação dos resultados previstos com os reais, para os anos de 2007 e 2008, em relação à série NDH – Portugal, apresentando, para o período em análise, um Erro Percentual Absoluto Médio (EPAM) de 4,8% e 4,2%, respectivamente para o método de regressão linear (RL) e Redes Neuronais Artificiais. Em séries cujo comportamento se verificou mais irregular, como a NDH – Região Centro, as RNA superaram a regressão linear apresentando um EPAM de 8,2%, duas vezes inferior ao conseguido com os métodos lineares.

A inclusão de variáveis auxiliares, nomeadamente indicadores macroeconómicos e do estado do tempo, à entrada da rede neuronal, também foi estudada, apresentando alguns resultados positivos. A título de exemplo, o Índice de Preços no Consumidor (IPC), reduziu o EPAM relativo à série, NDHP-RC, em cerca de 15%.

Palavras-chave: Procura Turística; Séries Temporais; Modelação Linear; Previsão Redes Neuronais Artificiais.

Área temática: Modelação Económica e Financeira.

Abstract

The main purpose of this paper is the comparative study of the linear regression model to the neural network model. The linear regression method is the appliance of an univariate model on every month for the studied time-period.

Using such different nature methodologies helps assessing each one's potential on the specificities of the tourism demand series, such as their seasonality and tendency. So, monthly data series for the tourism demand were used: "Monthly Guest Nights in the Hotels of Portugal (NDH) - North and Center Regions" for the time-period from January 1990 to December 2008 and also the United Kingdom data for the very same time-period. The latter presents a different pattern of data.

Since the developed models presented high statistical accuracy and fitting capabilities they were used for forecasting. We compared forecasted NDH data to the real data for the years 2007 and 2008 so as to measure its prediction capability. The linear regression model got a Mean Absolute Percent Error of 4,8% and the neural network model got a Mean Absolute Percent Error of 4,2%. In series, such as that of the NDH- Center Region, in which data shows more variability, the neural network model scored a Mean Absolute Percent Error of 8,2%, outperforming the linear regression model by two.

Extra auxiliary input variables to the neural network, such as macroeconomic and weather condition data were also used showing some good results. Just to give an example, using the NDH data series for the Center Region, the Consumer Price Index was able to reduce 15% the Mean Absolute Percent Error.

Key words: Tourism Demand; Time Series; Linear Modelling; Forecasting.

Thematic Area: Economic and Financial Modelling.

Agradecimentos

Quero agradecer ao orientador, o Prof. João Paulo Teixeira, pelo apoio concedido a todos os níveis, em todas as fases do projecto.

Não poderia deixar de agradecer à Prof. Paula Odete Fernandes, co-orientadora do projecto, pelo importante contributo dado em torno da problemática da procura turística em Portugal.

Finalmente, agradeço a todos aqueles que contribuíram de forma directa ou indirecta para a realização deste projecto.

Conteúdo

1. Introdução	1
1.1 Objectivos Propostos	2
1.2 Organização do Documento	3
2. Redes Neurais	4
2.1 Fundamentos.....	4
2.2 Arquitectura.....	5
2.3 Funções de Transferência/Activação	6
2.4 Processos de Aprendizagem.....	7
2.5 Algoritmos de Treino	8
2.5.1 Algoritmo de <i>Back-Propagation</i>	9
2.5.2 Algoritmo de Treino <i>Back-Propagation</i> Avançados.....	10
2.6 Critérios de Paragem.....	12
3. Objecto de Estudo	14
3.1 O Turismo Hoje	14
3.2 Importância da Previsão no Turismo	16
3.3 Influências Externas.....	17
3.4 Apresentação das Séries Temporais	18
4. Modelos de Previsão.....	22
4.1 Metodologias	22
4.2 Regressão Linear	23
4.2.1 Modelo Univariado	23
4.2.2 Resultados.....	27
4.3 A Rede Neuronal	29
4.3.1 Variáveis Auxiliares	31
4.3.2 Pré processamento dos Dados	32
4.3.3 Parâmetros pré-definidos	32
4.3.4 Resultados	34
5. Análise dos Resultados	37
5.1 Regressão Linear <i>versus</i> Redes Neurais.....	37
5.1.1 Séries Turísticas Regionais	39
5.1.2 Série Turística do Reino Unido	43

5.2	Influência das Variáveis Auxiliares	44
6.	Interface Gráfica	47
6.1	Ferramenta Utilizada.....	47
6.2	Barra de Menus.....	48
6.3	Ecrã – Sistema de Previsão de Séries Temporais	49
6.3.1	Funcionalidades.....	50
6.4	Ecrã - Visualizar Série de Dados	51
6.5	Ecrã – <i>Neural Network Training</i>	52
6.6	Ecrã – <i>Regression</i>	54
6.7	Ecrã – Definição de Parâmetros	55
6.7.1	Validações	56
6.7.2	Gestão das Matrizes de Parâmetros	57
7.	Conclusões e Perspectivas Futuras	58
7.1	Conclusões	58
7.2	Perspectivas Futuras.....	60
	Bibliografia.....	61

Lista de Tabelas

Tabela 1 Equações das rectas de regressão linear, mensais; valor da estatística <i>t-Student</i> e respectivo valor de prova, para cada coeficiente; coeficiente de correlação.....	24
Tabela 2 Erro Percentual Absoluto Médio de NDHP, no intervalo [Jan-2007:Dez-2008].	28
Tabela 3 Erro percentual mensal e EPAM no intervalo [Jan-2007:Dez-2008] para ambos os modelos, regressão linear (RL) e redes neuronais artificiais (RNA).....	38
Tabela 4 Erro percentual mensal e EPAM no intervalo [Jan-2007:Dez-2008] para ambos os modelos, regressão linear (RL) e redes neuronais artificiais (RNA) relativo à série NDHP-RN.	40
Tabela 5 Erro percentual mensal e EPAM no intervalo [Jan-2007:Dez-2008] para ambos os modelos, regressão linear (RL) e redes neuronais artificiais (RNA) relativo à série NDHP-RC.	41
Tabela 6 Erro percentual mensal e EPAM no intervalo [Jan-2007:Dez-2008] para ambos os modelos, regressão linear (RL) e redes neuronais artificiais (RNA) relativo à série NDH-RU.	43
Tabela 7 EPAM – Variáveis auxiliares.	45
Tabela 8 Dormidas registadas em Portugal (NDHP) no período [Jan-1990:Dez-2008].	63
Tabela 9 Valores previstos para a Série NDHP, no período [Jan-1990:Dez-2008] (RL)	64
Tabela 10 Erro Percentual Absoluto de NDHP, no período [Jan-1990:Dez-2008] (RL)	65
Tabela 11 EPAM em função do n.º de neurónios e algoritmo de treino.....	65
Tabela 12 Tempo de execução em função do n.º de neurónios e algoritmo de treino....	66
Tabela 13 Valores previstos para a Série NDHP, no período [Jan-1991:Dez-2008] (RNA - LM).....	67
Tabela 14 Valores previstos para a Série NDHP, no período [Jan-1991:Dez-2008] (RNA - RP).....	68
Tabela 15 Erro Percentual Absoluto de NDHP, no período [Jan-1990:Dez-2008] (RNA - LM).....	69
Tabela 16 Erro Percentual Absoluto de NDHP, no período [Jan-1991:Dez-2008] (RNA - RP).....	69
Tabela 17 Dormidas registadas em Portugal (NDHP-RN) no período [Jan-1990:Dez-2008]	70
Tabela 18 Valores previstos para a Série NDHP-RN, no período [Jan-1990:Dez-2008] (RL).....	71
Tabela 19 Valores previstos para a Série NDHP-RN, no período [Jan-1991:Dez-2008] (RNA).....	72
Tabela 20 Erro Percentual Absoluto de NDHP-RN, no período [Jan-1990:Dez-2008] (RL).....	73
Tabela 21 Erro Percentual Absoluto de NDHP-RN, no período [Jan-1991:Dez-2008] (RNA).....	73
Tabela 22 Dormidas registadas em Portugal (NDHP-RC) no período [Jan-1990:Dez-2008]	74

Tabela 23 Valores previstos para a Série NDHP-RC, no período [Jan-1990:Dez-2008] (RL).....	75
Tabela 24 Valores previstos para a Série NDHP-RC, no período [Jan-1991:Dez-2008] (RNA).....	76
Tabela 25 Erro Percentual Absoluto de NDHP-RC, no período [Jan-1990:Dez-2008] (RL).....	77
Tabela 26 Erro Percentual Absoluto de NDHP-RC, no período [Jan-1991:Dez-2008] (RNA).....	77
Tabela 27 Dormidas registadas no Reino Unido (NDH-RU) no período [Jan-1990:Dez-2008]	78
Tabela 28 Valores previstos para a Série NDH-RU, no período [Jan-1990:Dez-2008] (RL).....	79
Tabela 29 Valores previstos para a Série NDH-RU, no período [Jan-1991:Dez-2008] (RNA).....	80
Tabela 30 Erro Percentual Absoluto de NDH-RU, no período [Jan-1990:Dez-2008] (RL).....	81
Tabela 31 Erro Percentual Absoluto de NDH-RU, no período [Jan-1991:Dez-2008] (RNA).....	81
Tabela 32 Indicadores da Actividade Turística	81
Tabela 33 Índice de Preços no Consumidor em Portugal (2005=100) no intervalo [Jan-1996:Dez-2008]	82
Tabela 34 Índice de Preços no Consumidor no Reino Unido (2005=100) no intervalo [Jan-1996:Dez-2008]	82
Tabela 35 N.º de desempregados em Portugal (Milhares) no intervalo [Jan-1990:Dez-2008]	83
Tabela 36 N.º de desempregados no Reino Unido (Milhares) no intervalo [Jan-1990:Dez-2008]	83
Tabela 37 N.º de desempregados em Espanha (Milhares) no intervalo [Jan-1990:Dez-2008]	84
Tabela 38 N.º de desempregados em França (Milhares) no intervalo [Jan-1990:Dez-2008]	84
Tabela 39 Valores previstos para a Série NDHP-RC (VA=IPC), no período [Jan-1991:Dez-2008] (RNA)	85
Tabela 40 Percentual Absoluto de NDHP-RC, no período [Jan-1991:Dez-2008] (RNA-IPC).....	86

Lista de Figuras

Figura 1 Neurónio Artificial, R: n.º de elementos do vector de entrada [Demuth & Mark, 2002: 2-28].	5
Figura 2 Camada de Neurónios [Demuth & Mark, 2002:2-29].	6
Figura 3 Função Sigmoidal Logarítmica [Demuth & Mark, 2002:2-26].	6
Figura 4 Função Sigmoidal Tangente Hiperbólica [Demuth & Mark, 2002:2-26].	7
Figura 5 Função linear [Demuth & Mark, 2002:2-26].	7
Figura 6 Processo de ajuste dos parâmetros da RNA em aprendizagem supervisionada.	8
Figura 7 Rede neuronal (MLP) do tipo <i>feed-forward</i> genérica [Demuth & Mark, 2002:5-7].	9
Figura 8 N.º de empregados a tempo inteiro e <i>part-time</i> ligados directamente ao turismo em Portugal [Eurostat; 2009].	15
Figura 9 Proveitos totais dos estabelecimentos hoteleiros em Portugal [INE; 2009].	15
Figura 10 N.º absoluto de estabelecimentos hoteleiros em Portugal [Eurostat; 2009].	16
Figura 11 Dormidas dos turistas nas Unidades de Alojamento, em Portugal, no intervalo [Jan-90:Dez-08].	19
Figura 12 Dormidas dos turistas nas Unidades de Alojamento, na Região Norte de Portugal – NDHP-RN, no intervalo [Jan-90:Dez-08].	20
Figura 13 Dormidas dos turistas nas Unidades de Alojamento, na Região Centro de Portugal - NDHP-RC, no intervalo [Jan-90:Dez-08].	20
Figura 14 Dormidas dos turistas nas Unidades de Alojamento, no Reino Unido – NDH-RU, no intervalo [Jan-90:Dez-08].	21
Figura 15 Representação das rectas de regressão para os meses de Janeiro, Fevereiro e Março, no intervalo [1990:2006].	25
Figura 16 Representação das rectas de regressão para os meses de Abril, Maio e Junho, no intervalo [1990:2006].	25
Figura 17 Representação das rectas de regressão para os meses de Julho, Agosto e Setembro, no intervalo [1990:2006].	26
Figura 18 Representação das rectas de regressão para os meses de Outubro, Novembro e Dezembro, no intervalo [1990:2006].	26
Figura 19 Representação da série NDHP e a respectiva previsão para os anos de 2007 e 2008.	27
Figura 20 Representação gráfica dos valores previstos sobre a totalidade do conjunto NDHP (RL).	28
Figura 21 EPAM – Algoritmos de Treino/n.º de Neurónios.	33
Figura 22 Tempo de Execução (s - segundos) – Algoritmos de Treino/n.º de Neurónios.	34
Figura 23 Representação da série NDHP e a respectiva previsão para os anos de 2007 e 2008 (RNA – LM/RP).	35
Figura 24 Representação gráfica dos valores previstos sobre a totalidade do conjunto NDHP (RNA - RP).	36
Figura 25 Representação da série e a respectiva previsão para os anos de 2007 e 2008, pelos modelos RL e RNA.	38
Figura 26 Representação gráfica dos valores previstos da série NDHP-RN (RL).	40
Figura 27 Representação gráfica dos valores previstos da série NDHP-RN (RNA).	41

Figura 28	Representação gráfica dos valores previstos da série NDHP-RC (RL).	42
Figura 29	Representação gráfica dos valores previstos da série NDHP-RC (RNA).	42
Figura 30	Representação gráfica dos valores previstos da série NDH-RU (RL).	44
Figura 31	Representação gráfica dos valores previstos da série NDH-RU (RNA).	44
Figura 32	EPAM – Variáveis Auxiliares.	46
Figura 33	Barra de Menus.	48
Figura 34	Sistema de Previsão de Séries Temporais.	50
Figura 35	Ecrã – Visualizar Séries de Dados.	52
Figura 36	Ecrã – <i>Neural Network Training</i>	54
Figura 37	Ecrã – <i>Regression</i>	55
Figura 38	Ecrã – Definição de Parâmetros.....	56

Lista de Equações

Equação 1 Saída efectiva.....	8
Equação 2 Ajuste do peso sináptico.....	8
Equação 3 Ajuste do peso <i>bias</i>	9
Equação 4 Variação de peso sináptico.....	10
Equação 5 Erro na camada escondida.....	10
Equação 6 Erro em unidades de saída.....	10
Equação 7 Actualização dos pesos com <i>momentum</i>	11
Equação 8 Forma Matricial para a Série de Dados.....	23
Equação 9 Regressão Linear.....	23
Equação 10 Erro Percentual Absoluto Médio.....	27
Equação 11 Conjunto de treino.....	29
Equação 12 Conjunto de validação.....	30
Equação 13 Conjunto de teste.....	30
Equação 14 Normalização dos dados.....	32

Abreviaturas e Acrónimos

<i>IA</i>	Inteligência Artificial
<i>RNA's</i>	Redes Neurais Artificiais
<i>NN</i>	<i>Neural Networks</i>
<i>RL</i>	Regressão Linear
<i>SLP</i>	<i>Single Layer Perceptron</i>
<i>MLP</i>	<i>Multilayer Perceptron</i>
<i>FFN</i>	<i>Feed-Forward Networks</i>
<i>RP</i>	<i>Resilient Back-Propagation</i>
<i>LM</i>	<i>Levenberg Marquardt</i>
<i>CG</i>	<i>Conjugate Gradient</i>
<i>MSE</i>	<i>Mean Squared Error</i>
<i>EPAM</i>	Erro Percentual Absoluto Médio
<i>MSE</i>	<i>Mean Squared Error</i>
<i>GUI</i>	<i>Graphical User Interface</i>
<i>NDH</i>	Número de Dormidas em Hotéis
<i>NDHP</i>	Número de Dormidas em Hotéis de Portugal
<i>NDHP-RN</i>	Número de Dormidas em Hotéis de Portugal (Região Norte)
<i>NDHP-RC</i>	Número de Dormidas em Hotéis de Portugal (Região Centro)
<i>NDH-RU</i>	Número de Dormidas em Hotéis (Reino Unido)
<i>PENT</i>	Plano Estratégico Nacional do Turismo
<i>PIB</i>	Produto Interno Bruto
<i>INE</i>	Instituto Nacional de Estatística
<i>IPC</i>	Índice de Preços no Consumidor
<i>NDPT</i>	Número de Desempregados em Portugal
<i>NDES</i>	Número de Desempregados em Espanha
<i>NDFR</i>	Número de Desempregados em França
<i>NDRU</i>	Número de Desempregados no Reino Unido
<i>TMM</i>	Temperatura Máxima Média

Capítulo 1

1. Introdução

Este capítulo apresenta o contexto necessário para a compreensão do presente estudo. O âmbito onde se insere, a inspiração que segue e a motivação que o realiza, serão abordados nos parágrafos seguintes.

No contexto da disciplina de Inteligência Artificial, surgem vários paradigmas, na tentativa de implementar as capacidades cognitivas dos seres humanos, em máquinas. Interessa especialmente para este estudo, a metáfora designada por conexionista. Esta metáfora vem inspirar-se nas propriedades que se conhecem do sistema nervoso central, apresentando o processo de aprendizagem como o resultado da interação de um grande número de unidades de processamento fortemente interligadas entre si [Costa e Simões, 2004].

Esta inspiração é sustentada pelas vantagens óbvias, como o processamento paralelo, a tolerância a falhas, a imunidade ao ruído¹, que permitem a resolução de problemas complexos, em diversas áreas de aplicação [Haykin, 1999].

As redes neuronais têm sido aplicadas com sucesso em áreas como o reconhecimento de padrões, o controlo de processos e a modelação de séries temporais. Vários autores têm utilizado com êxito as redes neuronais para a previsão de séries temporais de diferentes naturezas, e.g., previsão de índices económicos, previsão de resultados desportivos ou mesmo a previsão de índices associados à procura turística [Bloom, 2003]. O desenvolvimento deste tipo de aplicações depende essencialmente dos registos históricos do índice que se pretende modelar, sendo portanto necessário dispor de um conjunto de informação tanto maior quanto maior a complexidade do comportamento da série alvo. Uma das limitações da previsão de séries temporais é a falta de habilidade para prever alterações que não se encontrem nos registos históricos [Law, 1999]. Uma forma de resolver o problema pode passar pela inclusão, à entrada da rede, de conhecimento adicional susceptível de influenciar a série alvo de modelação.

O turismo, sendo um factor estruturante da dinâmica da economia de qualquer país, se for implementado de uma forma sustentada pode vir a proporcionar benefícios a longo prazo, reflectindo-se na criação de postos de trabalho e na captação de divisas para os diferentes sectores da actividade económica a que está ligado.

À semelhança de outros países, no decurso das últimas décadas, também em Portugal o turismo se transformou num fenómeno singular apesar da situação económica e política lhe serem, muitas vezes, desfavoráveis.

¹ Entenda-se, imunidade a dados fragmentados ou não correlacionados.

Neste sentido e dado o crescimento substancial deste sector de actividade em Portugal, será de todo proveitoso desenvolver modelos susceptíveis de serem utilizados para efectuar previsões fiáveis da procura turística, dado que esta assume um papel relevante no processo de planeamento e de tomada de decisões, tanto no contexto do sector público como do privado. Actualmente, na área da previsão, 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, baseadas em modelos de regressão linear univariada, às mais complexas, como modelos que combinam algoritmos genéticos com redes neuronais artificiais [Fernandes, 2005].

Assim sendo, o propósito do presente trabalho prende-se com a descrição e comparação entre dois modelos desenvolvidos. Um modelo de regressão linear univariado mensal, tendo por base o método dos mínimos quadrados ordinários e o outro modelo baseado em redes neuronais artificiais (RNA), que tira partido da sua capacidade para modelar problemas não lineares, com frequência associados à previsão de séries temporais.

Por um lado, a regressão linear, apesar de eficiente, reduz-se a modelar problemas de natureza linear, mostrando ser uma metodologia muito limitada, em relação ao elenco de problemas que se pode aplicar. As redes neuronais artificiais acarretam um processo mais pesado em termos de cálculo, no entanto mostram ser eficazes na modelação de problemas não lineares, com relação matemática desconhecida.

A série temporal, que serviu de referência ao estudo foi: “Dormidas Mensais Registadas nos Estabelecimentos Hoteleiros em Portugal”. Estes dados estão disponíveis, no EUROSTAT [Eurostat, 2009]. Os dados observados reportam-se ao período compreendido entre Janeiro de 1990 e Dezembro de 2008, correspondendo assim a um total 228 observações mensais, 19 para cada mês, 12 para cada ano.

1.1 Objectivos Propostos

O principal objectivo da presente dissertação assenta no desenvolvimento dos modelos de previsão da série temporal e estudo comparativo entre o modelo baseado na regressão linear e o modelo baseado na técnica de redes neuronais artificiais (RNA). Adicionalmente, desenvolveu-se um sistema de computação que sustente os modelos desenvolvidos com a possibilidade do utilizador introduzir variações aos modelos ou usar novos dados (séries temporais ou novos dados de entrada – no caso de RNA).

O sistema deve possibilitar o carregamento de uma determinada série², guardada num ficheiro de texto, escolher o método de previsão, inicializar os parâmetros gerais necessários à organização dos dados e à metodologia associada, efectuar uma simulação ou uma previsão sobre o conjunto de dados introduzido e calcular o erro associado, no caso de simulação.

Propõe-se também um estudo comparativo sobre os resultados das duas metodologias de natureza diferente, a regressão linear e as redes neuronais artificiais, para a modelação de um mesmo problema – séries de procura turística, com toda a sua especificidade de que são características, como a sazonalidade e tendência.

² Neste documento designada muitas vezes por *dataset* (conjunto de dados).

Efectua-se também um estudo sobre as redes neuronais artificiais, sobre as variáveis que definem esta metodologia, nomeadamente, a sua arquitectura, a sua dinâmica de aprendizagem e capacidade de generalização.

Pretende-se ainda, estudar a evolução da procura turística, bem como todas as variáveis que a possam influenciar, como por exemplo, os índices económicos.

Finalmente, pretende-se criar uma aplicação gráfica que facilite o interface do utilizador com o sistema desenvolvido, permitindo de forma intuitiva e flexível, tirar partido das funcionalidades implementadas.

1.2 Organização do Documento

Este documento está organizado em sete capítulos.

Depois do capítulo dedicado à introdução, aborda-se com alguma profundidade a metodologia de redes neuronais, os seus fundamentos, a sua arquitectura bem como os mecanismos subjacentes ao processo de aprendizagem. O capítulo três destina-se à apresentação do objecto de estudo, a importância da modelação da procura turística, bem como as suas principais influências. Os modelos de previsão implementados são descritos com detalhe, no capítulo 4. Esta secção descreve as metodologias utilizadas. O capítulo cinco é dedicado à análise de resultados e conclusões. Neste capítulo vão comparar-se os resultados obtidos em ambas as metodologias. O capítulo seis faz a apresentação da aplicação gráfica e suas funcionalidades. Segue-se ainda um capítulo para referir as conclusões gerais e perspectivas futuras para desenvolvimentos desta natureza.

Capítulo 2

2. Redes Neurais

Este capítulo é dedicado à metodologia de redes neurais artificiais, à sua origem, aos seus fundamentos e à sua definição como mecanismo inteligente na resolução de problemas.

2.1 Fundamentos

Não raras as vezes, o ser humano, no seu caminho pela procura de conhecimento, inspira-se na natureza, para criar soluções de engenharia elegantes e robustas na resolução de problemas.

O funcionamento do sistema nervoso central, mais propriamente do cérebro, tem sido alvo de constantes estudos, na tentativa de perceber a sua dinâmica na aquisição de conhecimento. Os primeiros esforços com resultados, foram obtidos pelo espanhol, Santiago Ramón y Cajal. Este histologista ganhou o prémio Nobel da medicina de 1906 porque descobriu pela primeira vez que o cérebro é composto por milhares de milhões de neurónios distintos e que comunicam entre si através de conexões designadas por sinapses [Wikipédia, 2009].

Não é sem receios que hoje podemos falar sobre o funcionamento do nosso cérebro, no entanto, sabe-se que cada neurónio recebe impulsos de outros neurónios através das dendrites, processa esses impulsos, podendo desencadear disparos de sinais de natureza electroquímica através do seu axónio, para outros neurónios. A natureza destes sinais é eléctrica ao longo do axónio até à sinapse, local onde é convertida para um sinal de natureza química [Costa e Simões, 2004]. Esta conversão é assegurada pelos neurotransmissores.

De forma geral, o funcionamento do sistema nervoso compreende três fases. Através de receptores específicos, o cérebro recebe estímulos do exterior, processa o sinal recebido e gera uma resposta para os órgãos actuadores recebendo de seguida um feedback dessa mesma resposta e assim sucessivamente.

Quando uma criança começa a dar os primeiros pontapés numa bola de futebol, realiza um conjunto de experiências que são positivas ou negativas, consoante o objectivo que se pretenda. Em cada tentativa, existe associado um processo mental, que dá origem a um determinado resultado. No caso de ser o desejado, a criança irá reforçar as ligações cerebrais que lhe deram origem, porque observa que atingiu com prazer o objectivo de

acertar na bola. Este processo vai repetir-se em cada experiência, aprimorando desta forma a técnica de rematar uma bola de futebol.

Portanto, a metáfora conexionista, admite que a aquisição de conhecimento, bem como o seu armazenamento, deriva directamente do ajuste dos pesos sinápticos que determinada experiência proporciona [Haykin, 1999].

2.2 Arquitectura

No sentido de reproduzir de forma simplificada o neurónio biológico, o fisiologista Warren McCulloch e o Lógico Walter Pitts, criaram em 1943 um modelo para o neurónio artificial [McCulloch & Pitts, 1943].

Conforme apresentado na Figura 1, o neurónio artificial possui um vector de entradas, p e um vector de pesos sinápticos W , uma entrada b designada por *bias* ou desvio, que representa um determinado limiar e uma função de activação f . Esta representação voltará a ser abordada na secção abaixo, no tema sobre algoritmos de treino e processos de aprendizagem. Desta forma, a saída a , é o resultado da aplicação da função f ao produto Wp somado ao desvio b .

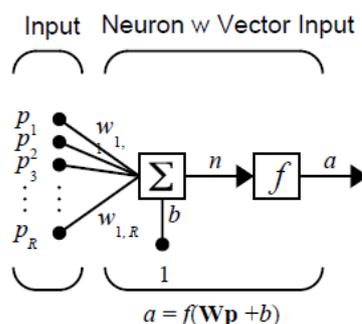


Figura 1 Neurónio Artificial, R : n.º de elementos do vector de entrada [Demuth & Mark, 2002: 2-28].

A arquitectura de uma rede neuronal artificial define a forma como as unidades de processamento, os neurónios, estão organizadas e como interagem entre si. Desta forma a arquitectura está estritamente relacionada com o algoritmo de treino da rede.

Existem duas grandes categorias de arquitecturas de rede: As redes *feed-forward* (*FFN*), caracterizadas pelo facto de possuírem um fluxo cálculo unidireccional, sem ciclos. Esta categoria pode ainda ser classificada como *Single Layer Perceptron* (*SLP*), se possuírem apenas uma camada (Figura 1 e Figura 2), ou *Multi Layer Perceptron* (*MLP*), quando se trata de redes com mais de uma camada (Figura 7). O número de camadas está directamente relacionado com a capacidade para resolver problemas mais complexos. Especialmente, quando o conjunto de dados de entrada possui dimensões elevadas, é particularmente útil incluir pelo menos uma camada oculta. As *FFN*'s podem ainda ser parcialmente ou totalmente conectadas, se cada neurónio de uma camada está ligado a todos os neurónios da camada adjacente seguinte [Haykin, 1999].

Existem também redes neuronais cíclicas ou recorrentes, cuja saída de determinado neurónio pode contribuir como entrada para o mesmo neurónio, ou para outros elementos que não se situem necessariamente na camada adjacente seguinte.

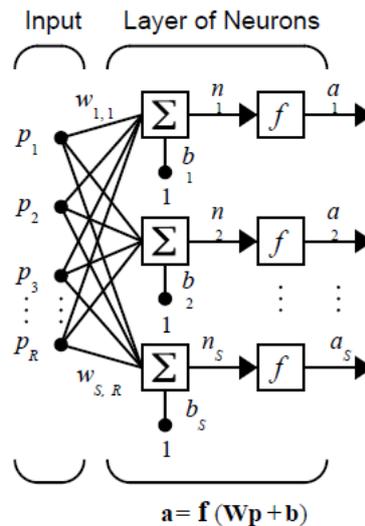


Figura 2 Camada de Neurónios [Demuth & Mark, 2002:2-29].

2.3 Funções de Transferência/Activação

A função de activação representada por $f(v)$, define a amplitude da saída de um neurónio num intervalo finito. A função sigmoideal logarítmica, (Figura 3) com a sua típica forma de um S , é a função de activação mais utilizada em projectos de RNA's. Esta função crescente descreve um balanceamento gracioso entre o comportamento linear e não linear [Haykin, 1999].

Em determinados algoritmos pode ser vantajoso, utilizar uma função de activação anti-simétrica permitindo reduzir o número de iterações de treino necessárias, à convergência pretendida, como é o caso da função tangente hiperbólica mostrada na Figura 4. Esta também é uma função sigmoideal, mas com variação simétrica entre -1 e 1 [Haykin, 1999].

A função de activação linear, representada na Figura 5, é frequentemente utilizada, na camada de saída da rede neuronal. Esta função de activação tem como vantagem a sua simplicidade de cálculo.

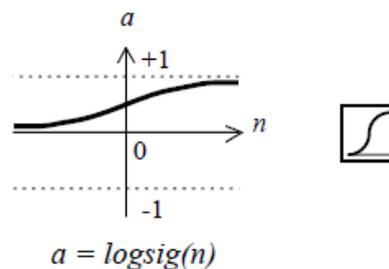


Figura 3 Função Sigmoideal Logarítmica [Demuth & Mark, 2002:2-26].

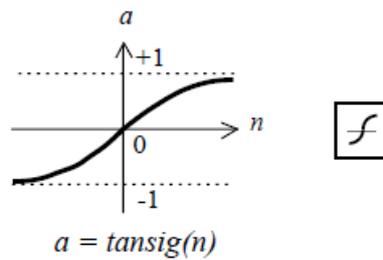


Figura 4 Função Sigmoide Tangente Hiperbólica [Demuth & Mark, 2002:2-26].

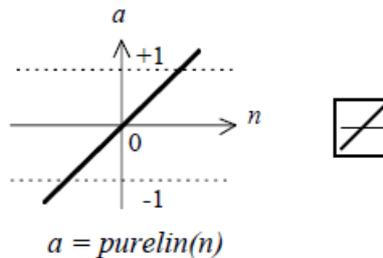


Figura 5 Função linear [Demuth & Mark, 2002:2-26].

2.4 Processos de Aprendizagem

Já foi referido em secções anteriores que a aprendizagem baseia-se na alteração iterativa das estruturas internas de um determinado agente, permitindo que este se adapte ao seu ambiente.

A aprendizagem é definida como,

'...o processo pelo qual um determinado agente altera de forma automática as suas estruturas internas, de modo a realizar melhor tarefas do seu interesse, nas condições impostas pelo ambiente' [Costa e Simões, 2004:195].

Do ponto de vista técnico, a aprendizagem é descrita como o processo de ajustar os pesos sinápticos, W , da rede neuronal segundo uma determinada regra denominada por algoritmo de aprendizagem.

O processo de aprendizagem pode ser supervisionado ou não supervisionado, dependendo da forma como o processo de treino é assistido ou não por uma saída alvo. Em redes supervisionadas, de particular interesse para o nosso estudo, existe um conjunto de dados de treino e valores alvo que é fornecido à rede, numa primeira fase, de modo que o ajuste do peso de cada ligação seja actualizado de forma a produzir a saída desejada. Estes dados são fundamentalmente conjuntos de pares entrada/saída, consistindo cada par, por um sinal de entrada e a correspondente resposta desejada. Estas instâncias são designadas neste documento por conjunto de amostras de treino.

Mais detalhadamente, no processo de aprendizagem supervisionada, a rede neuronal é alimentada com um conjunto de dados de treino (amostras) e calcula a saída, com base em pesos inicialmente aleatórios, comparando-os com a saída desejada. Se a saída obtida, for diferente da desejada, os pesos dos neurónios são actualizados com base numa fórmula especificada no algoritmo de treino. Este processo repete-se

iterativamente até que a saída desejada seja aproximada, ou até que seja verificado algum critério de paragem (abordados ainda neste capítulo) (Figura 6).

Depois desta etapa, a rede neuronal possui uma estrutura capaz de classificar dados inteiramente novos (relacionados com a mesma experiência), através da sua capacidade de generalização.

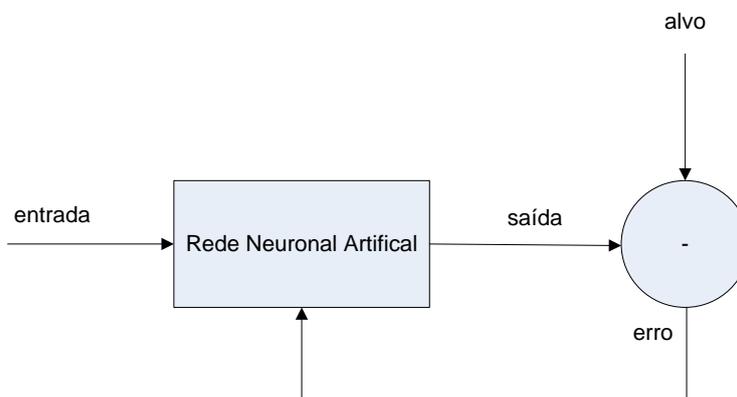


Figura 6 Processo de ajuste dos parâmetros da RNA em aprendizagem supervisionada.

2.5 Algoritmos de Treino

Como já foi referido na secção anterior, está subjacente ao processo de aprendizagem das redes neuronais, um algoritmo de treino, cuja finalidade é ajustar os pesos dos neurónios, de modo que a aproximação à saída desejada em cada iteração, produza um erro tendencialmente menor.

Essencialmente, esta actualização é efectuada através da aplicação iterativa de regras de aprendizagem, traduzidas por fórmulas matemáticas.

Partindo da estrutura apresentada na Figura 1, e efectuando o cálculo da saída efectiva, a_j , o ajuste do peso do neurónio, j , com R entradas, é dado por, W_{ji} Equação 2).

Equação 1 Saída efectiva.

$$a_j = f\left(b_j + \sum_{i=1}^R W_{ji} p_i\right)$$

Numa primeira fase, os pesos W_{ji} , juntamente com a variável b_j , designada por *bias*, são inicializados com valores aleatórios. Depois de se efectuar o cálculo da saída, aplicando a função de activação, f , à entrada efectiva, os pesos e *bias*, são actualizados respectivamente pela Equação 2 e Equação 3. Desta forma a variação do peso, depende da diferença entre o valor da variável que se pretende aproximar, t (*target*) e o valor obtido, y .

Equação 2 Ajuste do peso sináptico.

$$W_{ji} = W_{ji} + \alpha (t - a_i) p_i, \quad \alpha = 0:1$$

Equação 3 Ajuste do peso *bias*.

$$b_j = b_j + \alpha t - a_i, \quad \alpha = 0:1$$

Ao ajuste do peso sináptico, é ainda associada uma taxa de aprendizagem, α . Este parâmetro controla a velocidade de convergência do processo. Quanto menor for este parâmetro, menores serão as variações dos pesos neuronais, entre duas iterações de treino. Por outro lado se a taxa de aprendizagem for demasiado grande, os saltos no espaço de procura são tais, que a rede pode tornar-se instável ou oscilatória.

Este processo iterativo é interrompido por um determinado critério de paragem definido previamente. Em cada iteração é calculado o erro associado à aproximação efectuada, podendo este, no caso de apresentar um valor aceitável, interromper o treino. Os critérios de paragem serão ainda abordados neste capítulo, com mais detalhe.

O algoritmo apresentado nesta secção permite apenas treinar redes neuronais de apenas uma camada, permitindo apenas a resolução de problemas simples, de natureza linear. Quando se pretende uma solução para um problema mais complexo, é necessária uma arquitectura multi-camada (MLP), bem como também um algoritmo adequado a esta topologia.

2.5.1 Algoritmo de *Back-Propagation*

As redes neuronais MLP, têm sido aplicadas com êxito para resolver problemas complexos, através de processos de treino supervisionados com o algoritmo de treino popular designado por, algoritmo de *Back-Propagation* (BP).

Este algoritmo permite, após o cálculo do erro, propagar esse mesmo erro para as camadas predecessoras, de modo a dividi-lo pelos pesos que mais contribuíram para ele, recorrendo ao gradiente da função de erro.

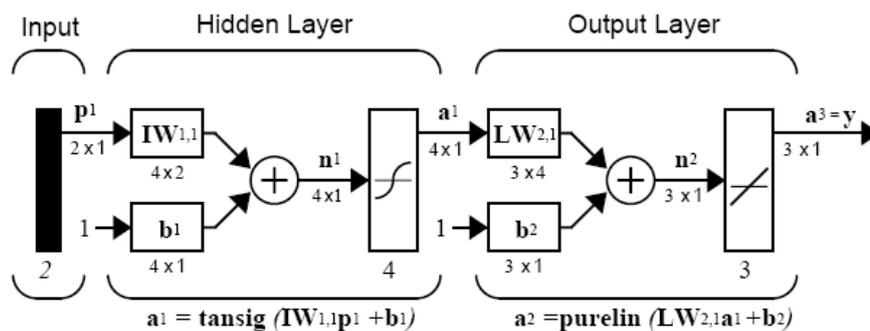


Figura 7 Rede neuronal (MLP) do tipo *feed-forward* genérica [Demuth & Mark, 2002:5-7].

Numa rede com várias camadas (Figura 7), a saída de um determinado neurónio i contribui como entrada para o neurónio j , que por sua vez vai influenciar vários neurónios na camada seguinte. Isto dificulta a tarefa de determinar a contribuição individual de cada neurónio para o erro.

Devido ao facto de não se possuírem valores para a saída esperada nas camadas intermédias ou escondidas, o ajuste dos pesos sinápticos não pode ser efectuada. Durante o processo de propagação (para a frente), os pesos são fixos, sendo apenas no

sentido inverso (retropropagação) que os pesos são ajustados segundo uma regra de correcção de erro [Haykin, 1999].

A propagação progressiva do erro para trás, permite que a actualização dos pesos seja efectuada apenas com base em informação local.

A variação do peso sináptico entre o neurónio i e j é dado pela Equação 4, δ_j representa o erro a retropropagar na camada escondida como um gradiente local (Equação 5) e y_i o sinal de entrada do neurónio j .

Equação 4 Variação de peso sináptico.

$$\Delta W_{ji} = \alpha \times \delta_j \times y_i$$

Equação 5 Erro na camada escondida.

$$\delta_j = \sum \delta_{kj} \times W_{kj} \times f' a_j$$

A Equação 6, representa o erro a retropropagar nas unidades de saída, onde ϵ_k expressa o erro efectivo (diferença entre a saída obtida e a saída esperada).

Equação 6 Erro em unidades de saída.

$$\delta_k = \epsilon_k \times f' a_k$$

Num primeiro passo é calculado o erro na camada de saída δ_k , actualizando-se, W_{kj} depois é calculado o erro na camada intermédia δ_j (dependente de δ_k), actualizando-se W_{ji} e assim sucessivamente, percorrendo todas as camadas da rede neuronal.

A regra de aprendizagem utilizada por este algoritmo de treino é designada por gradiente descendente. Esta regra utiliza, como mecanismo de controlo, a variação da função de erro em relação ao vector de pesos. Desta forma considera-se que o algoritmo de *Back-Propagation*, convergiu convenientemente para a solução desejada, quando a variação do erro por iteração de treino for suficientemente pequena. Este facto levanta uma questão importante, no que se refere ao tempo de convergência. A taxa de aprendizagem constante, leva-o a ziguezaguear no espaço da superfície de procura, tornando o processo de convergência mais lento [Haykin, 1999].

No sentido de resolver esta e outras limitações deste algoritmo de treino, foram criadas algumas variantes no sentido de diminuir o tempo de convergência e aumentar a performance do processo de treino.

2.5.2 Algoritmo de Treino *Back-Propagation* Avançados

O problema de determinar a taxa de aprendizagem óptima em cada iteração é resolvido introduzindo, no cálculo da variação dos pesos, uma constante positiva designada por *momentum*. Partindo da Equação 4, a regra Delta é generalizada através da adição de um termo que tem um efeito estabilizador no processo de convergência (Equação 7). Para uma determinada iteração n , a actualização dos pesos não será função apenas do erro, mas também do valor dos pesos em $n-1$, pelo factor N . Este termo funciona como uma

memória acelerando as descidas do gradiente, evitando em alguns casos que o processo de aprendizagem termine num mínimo local da superfície de erro.

Equação 7 Actualização dos pesos com *momentum*.

$$\Delta w_{ji}(n) = \eta \times \Delta w_{ji}(n-1) + \alpha \times \delta_j(n) \times y_i(n)$$

Com o objectivo de melhorar ainda mais a performance, foram surgindo outros algoritmos de treino, que apesar de se basearem no algoritmo de *Back-Propagation*, introduzem novos conceitos, nomeadamente heurísticas e técnicas de optimização.

Resilient Back-Propagation (RP)

Este algoritmo foi criado com o objectivo de eliminar as pequenas variações dos pesos neuronais, introduzidas pela magnitude das derivadas parciais. Desta forma utiliza apenas o sinal da derivada para actualizar os pesos.

Também conhecido como *Rpro*, este algoritmo, baseia-se no sinal do gradiente descendente da função *performance*, em vez de no próprio gradiente descendente, como é habitual. Este aspecto permite que continue a haver uma evolução no ajuste dos parâmetros da rede em zonas em que o gradiente da função tenha uma amplitude muito baixa, como sucede nas extremidades das funções sigmoidais usadas na camada escondida. Este algoritmo tem sido utilizado por vários autores [Fernandes e Teixeira, 2008].

Este algoritmo é geralmente mais rápido em relação a outros algoritmos baseados no gradiente descendente e bastante modesto nos requisitos de memória e computação [Demuth & Mark, 2002].

Informação completa e detalhada sobre o funcionamento deste algoritmo é dada em [Riedmiller & Braun, 1993].

Conjugate Gradient (CG)

Em algoritmos baseados no gradiente conjugado, a taxa de aprendizagem é ajustada em cada iteração. A procura é efectuada ao longo da direcção do gradiente de modo a minimizar a função de performance.

Os algoritmos de gradiente conjugado são usualmente mais rápidos que o algoritmo RP. Apesar de necessitarem de mais memória, oferecem uma boa escolha para redes que possuam uma estrutura com um número elevado de pesos [Demuth & Mark, 2002].

Existem variantes do algoritmo do gradiente conjugado, cuja variação assenta na forma como é calculada a procura do gradiente. Cada fórmula leva o nome dos seus criadores. Destacam-se os algoritmos de *Powell-Beale*, *Fletcher-Reeves* e *Polak-Ribiere*.

Dentro desta categoria existe ainda o algoritmo *Scaled Conjugate Gradient*. Este algoritmo requer mais iterações para convergir, no entanto o número de computações em cada iteração é significativamente reduzido [Demuth & Mark, 2002].

Método de *Newton*

Este método é uma alternativa de otimização que se baseia na computação da matriz *Hessiana*³, que prevê o cálculo da derivada de segunda ordem do erro. No entanto, este método é complexo e computacionalmente exigente. Existem outros métodos, baseados neste que não efectuam o cálculo da matriz *Hessiana*, fazendo apenas uma aproximação desta em cada iteração⁴.

Este método tem como vantagem a convergência em poucas iterações, sendo eficiente em redes neuronais de dimensões reduzidas.

One Step Secant

Este algoritmo é uma alternativa, computacionalmente mais modesta, do método de *Newton*. Este método não guarda completamente a matriz *Hessiana*, assume que em cada iteração a matriz anterior é a matriz identidade [Demuth & Mark, 2002].

Levenberg-Marquardt

Este algoritmo, semelhante aos métodos de *Newton*, foi criado com o intuito de aproximar a matriz *Hessiana*, com recurso à matriz *Jacobiana*. Esta última matriz contém em cada linha um exemplo particular do conjunto de treino, sendo computacionalmente menos complexa que a matriz *Hessiana*.

Este algoritmo converge em relativamente poucas iterações, embora cada uma delas necessite de mais memória.

Se a rede neuronal utilizada não for demasiado grande, este algoritmo tem mostrado ser uma solução bastante eficiente.

2.6 Critérios de Paragem

Para que seja possível aplicar a metodologia de redes neuronais a um determinado problema, é necessário possuir um conjunto de dados (pares entrada/saída) de tal forma vasto que permita a sua segmentação em três conjuntos: conjunto de treino, de validação e de teste. Reserva-se para o processo de treino os primeiros dois conjuntos, sendo que estes devem ser representativos das situações diversas possíveis de ocorrer no problema a modelar. O conjunto de teste é utilizado para qualificar o poder de generalização do processo de treino.

O processo de treino é interrompido quando se atingem níveis de *performance* desejáveis na função de erro. Habitualmente, a função de erro é a função, erro quadrático médio (MSE)⁵. O treino pode ainda ser interrompido quando se atinge um limite de iterações previamente definido. Existem outros critérios de paragem, nomeadamente, a utilização de um conjunto de validação, que permite impedir que o processo de treino provoque uma perda do poder de generalização na classificação da rede neuronal. Denomina-se habitualmente por validação cruzada.

Amiúde, acontece que quando durante o processo de treino, o ajuste dos pesos dos neurónios é de tal maneira específico para classificar os dados de treino, o processo de

³ A matriz *Hessiana* (em honra ao matemático alemão *Ludwig Otto Hesse*) é constituída pelas derivadas de segunda ordem dos erros em função dos pesos sinápticos.

⁴ Designados *Quasi-Newton*.

⁵ Do inglês *Mean Squared Error*.

aprendizagem perde capacidade de generalização em relação aos dados de teste. Este fenómeno designa-se por *overfitting*. Desta forma a validação cruzada permite que ao fim de cada iteração de treino, seja calculado o erro não só para o conjunto de treino, mas também para o conjunto de validação. Quando, durante este processo, o erro no conjunto de validação deixa de convergir, ou aumentar, o treino é interrompido.

Esta solução apresenta alguns problemas. Um deles é a grande probabilidade de ficar retido em mínimos locais do conjunto de validação. Outro problema reside na determinação do tamanho do conjunto de validação. Se for demasiado grande o treino é prejudicado por falta de informação, se for muito pequeno, o treino será interrompido precocemente.

Outra solução baseia-se na redução da complexidade da rede neuronal, nomeadamente do número de camadas ou neurónios, limitando a capacidade da rede para memorizar demasiados padrões. Porém esta solução é “arriscada”, uma vez que se pode estar, inadvertidamente, a limitar o poder de convergência para a solução desejada.

É importante referir também que a estrutura da RNA (número de nós e camadas) deve ser suficiente para modelar adequadamente a complexidade do problema, mas não maior que isso, porque obriga à utilização de uma maior quantidade de dados para treinar a rede de forma a poder ajustar correctamente todos os pesos e *bias*.

Capítulo 3

3. Objecto de Estudo

A procura turística, traduzida pelo número de dormidas, de residentes e não residentes, em estabelecimentos hoteleiros ou similares, em Portugal, constitui-se como o objecto de estudo alvo da aplicação das metodologias de modelação referidas em secções anteriores. Trata-se de uma série de dados fornecida pelo *Eurostat*, com os valores absolutos mensais desde Janeiro de 1990 a Dezembro de 2008, que reflectem a procura turística em todas as regiões de Portugal – Região Norte, Região Centro, Região de Lisboa, Alentejo, Algarve, Ilhas dos Açores e Madeira.

Nas secções seguintes pretende-se fazer uma abordagem geral da importância que o turismo assume na economia portuguesa, bem como as principais motivações por detrás da sua modelação. Pretende-se ainda analisar as principais influências externas ao turismo em Portugal, no sentido de encontrar informação relevante que auxilie a modelação da sua procura.

3.1 O Turismo Hoje

O sector turístico, numa perspectiva macro económica, tem emergido como um dos mais dinâmicos e líderes da economia de muitos países. Na década de 90, foi considerado como a principal fonte de receitas de economias como a Italiana e a Espanhola [Goeldner *et al.*, 1999] e como um dos principais sectores empregador nas economias do Canadá, Japão e Reino Unido [Hawkins *et al.*, 1992].

Também em Portugal à semelhança de outros países, no decurso das últimas décadas, o turismo transformou-se num fenómeno singular apesar da situação económica e política lhe serem, muitas vezes, desfavoráveis. Dados estatísticos de 2007 indicam que o turismo contribuiu com 10,4% do PIB (para Portugal), e a proporção do Valor Acrescentado gerado pelo turismo no PIB da economia portuguesa foi de 4,6% em 2005, 4,9% em 2006 e 5,1% em 2007, [INE, 2008; Turismo de Portugal, 2009a]. Ainda, segundo as mesmas fontes, a procura turística, medida pelo consumo, ultrapassou os 15 mil milhões de euros, em 2006, mais 1,3 mil milhões que no ano anterior. E no período entre 2000 e 2006, o consumo turístico teve um crescimento médio anual de 3,7%, correspondendo no final do período a mais 3 mil milhões de euros, comparativamente com o ano de 2000. Em 2007, o crescimento situou-se nos 11,4%, ano em que o consumo turístico ultrapassou os 17 mil milhões de euros.

Desta forma, o mercado turístico em Portugal assume-se como um sector verdadeiramente estratégico para a economia, na medida em que gera riqueza e emprego para o país. Portugal possui as características necessárias – condições climatéricas, recursos naturais e culturais, essenciais na criação de uma oferta turística rica e diversificada. Este facto deve motivar o investimento na diversidade e qualificação da oferta, bem como na sua promoção. A Figura 8 mostra a tendência claramente positiva nos últimos 15 anos, da influência do turismo na criação de emprego. Neste período o número de pessoas que possui um emprego a tempo inteiro ou *part-time*, directamente relacionado com o turismo, quase duplicou.

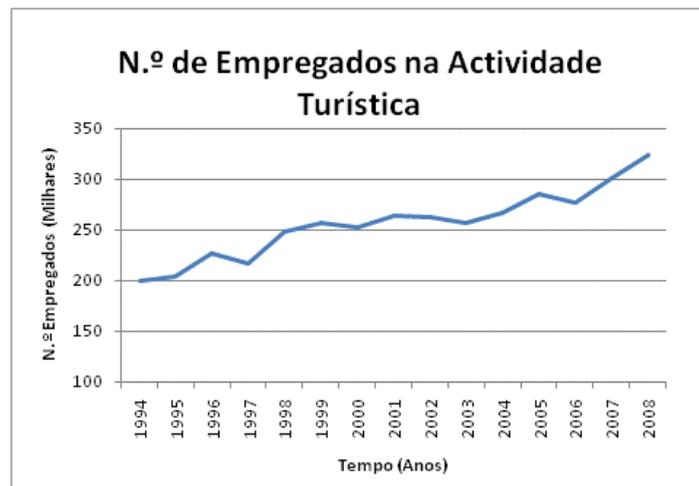


Figura 8 N.º de empregados a tempo inteiro e *part-time* ligados directamente ao turismo em Portugal [Eurostat; 2009].

Além da capacidade para gerar emprego, o sector turístico em Portugal contribui em larga medida para o Produto Interno Bruto (PIB) nacional. Os proveitos totais que os estabelecimentos hoteleiros obtiveram nos últimos 7 anos, espelham bem como os lucros em torno desta actividade não param de crescer (Figura 9). Estes indicadores revelam a importância de se continuar a investir neste sector, incrementando a sua influência claramente positiva e impulsionadora da economia de Portugal.

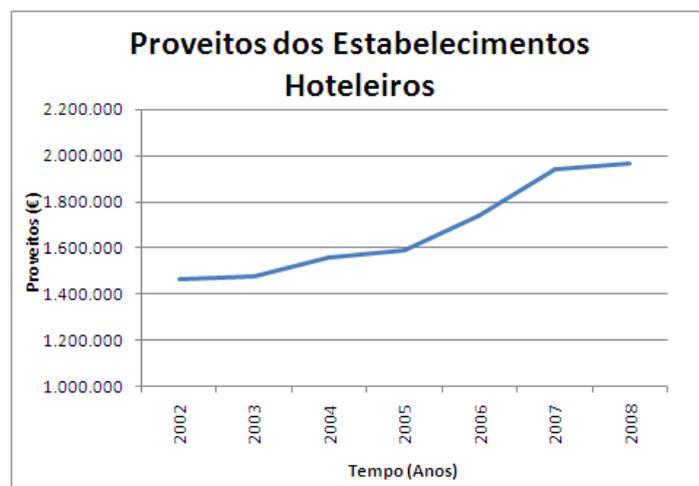


Figura 9 Proveitos totais dos estabelecimentos hoteleiros em Portugal [INE; 2009].

O investimento no sector turístico tem se evidenciado cada vez mais em Portugal. Nos últimos 15 anos o n.º de estabelecimentos hoteleiros não parou de crescer, acentuando-se esse mesmo crescimento a partir do ano 2000. Desde então, até 2007, foram criados cerca de 300 novos estabelecimentos.

Segundo os objectivos estratégicos enunciados no Plano Estratégico Nacional do Turismo [PENT; 2006], Portugal ambiciona crescer ao ano, 5% no número de turistas, atingindo os 20 milhões de turistas em 2015. Neste mesmo ano, a actividade turística em Portugal representará 15% do PIB nacional contribuindo também com uma considerável fatia no que diz respeito à criação de emprego.

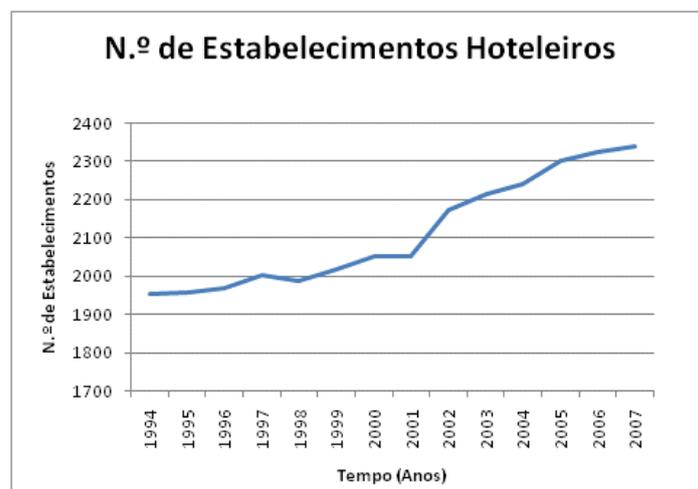


Figura 10 N.º absoluto de estabelecimentos hoteleiros em Portugal [Eurostat; 2009].

O investimento no turismo em Portugal prevê a criação de sistemas de incentivo (mecanismos de apoio ao financiamento), a promoção de eventos internacionais que projectem a imagem de Portugal no estrangeiro, a participação em feiras de turismo, a criação de agências regionais de promoção e formação de recursos humanos especializados, com escolas de hotelaria, estágios curriculares em empresas e organismos do sector e ainda bolsas de emprego.

Destaca-se ainda o investimento em pólos e produtos turísticos, criando uma oferta qualificada e diferenciada relativamente aos demais destinos turísticos. Prevê-se também a valorização turística dos recursos naturais e do património das regiões [Turismo de Portugal, 2009a].

3.2 Importância da Previsão no Turismo

Motivada, pelo crescente investimento no turismo em Portugal, a modelação deste sector, no sentido de estimar a sua evolução, está a assumir um papel cada vez mais importante. Projectar tendências, prever alterações e modelar comportamentos, pode ser fundamental para sustentar determinadas opções no investimento de qualquer sector, incluindo o sector turístico. Ainda, e no sentido de minimizar os potenciais impactos negativos no turismo, os agentes decisores, associados a este sector devem ter não só um conhecimento profundo do mesmo mas também das mudanças que afectam as tendências no turismo. Faulkner e Valério (2000) destacam a importância dos decisores se preocuparem com a redução do risco associado à tomada de decisão.

Assim, torna-se pertinente aos decisores turísticos planearem cuidadosamente as acções a tomar para atrair e manter os turistas numa região, e se necessário, aumentar as infraestruturas. Logo, a previsão da procura turística é de extrema importância na gestão turística devido sobretudo ao carácter perecível do produto/serviço turístico e ao facto de as fases de produção e de consumo do produto/serviço turístico verificarem-se em simultâneo. Por forma a ir ao encontro destas características singulares dos produtos/serviços turísticos os decisores turísticos devem dar uma resposta imediata mediante a observação dos níveis de oferta e de procura para conseguirem reagir em sintonia com o mercado. Uma previsão efectiva da procura turística fornece informação atempada para que os decisores turísticos possam tomar as respectivas decisões. A previsão da procura turística apresenta-se como um pilar estratégico ao permitir, atempadamente, informação relevante para que os decisores turísticos consigam antecipar potenciais crescimentos e riscos num futuro próximo para um país, região ou economia específica [Hu, 2002].

Diversos trabalhos científicos têm sido realizados e publicados na área do turismo, nos últimos tempos, e são unânimes em considerar a relevância que a procura turística assume no processo de planeamento, tomada de decisão e controlo do sector turístico [Witt & Witt, 1995; Fernandes, 2005; Yu & Schwartz, 2006; Fernandes *et al.* 2008].

3.3 Influências Externas

Não seria possível efectuar um estudo do turismo em Portugal sem compreender realmente quais as principais influências. No contexto da previsão de um determinado índice, visto que a mesma se baseia essencialmente nas características e padrões presentes nos seus dados históricos, pode ser importante adicionar alguma heurística extra.

Dadas as características da oferta turística em Portugal, estritamente ligada ao sol e mar (praia, turismo náutico, golfe, *touring* paisagístico), uma das principais influências está relacionada com as condições meteorológicas e climáticas. Pode constatar-se, na próxima secção, que a procura turística, traduzida pelo número de dormidas em estabelecimentos hoteleiros, é pautada por uma visível sazonalidade, com picos de procura no Verão e quedas abruptas na direcção dos meses mais frios. Desta forma o clima assume-se como uma das principais influências externas da procura turística.

De salientar que o *Travel and Tourism Competitiveness Report 2009* coloca Portugal na 17ª posição no ranking do *Travel and Tourism Competitiveness Index*, composto por uma lista de 133 países. No contexto da União Europeia (UE27), Portugal apresenta-se na 10ª posição, colocando-se à frente de países como a Irlanda, Bélgica, Itália ou a Grécia [Turismo de Portugal, 2009b]. Assim, com esta posição favorável, o turismo para Portugal, sem dúvida que é também um factor de equilíbrio das contas públicas, tendo as receitas turísticas um peso significativo na Balança Corrente. O turismo assume, deste modo, uma importância estratégica na economia portuguesa, gerando receitas significativas e revelando um conjunto de vantagens competitivas, em regra, não alcançáveis por outras actividades [Turismo de Portugal, 2009a].

Neste sentido, a procura turística está também directamente relacionada com alguns indicadores do estado da economia. No caso dos turistas de proveniência interna, os residentes, interessa-nos os indicadores económicos nacionais, como o índice de preços no consumidor ou a taxa de desemprego, pois ambos permitem medir a situação

económica de um país. Quando a proveniência é estrangeira, em que os principais mercados emissores situam-se na Europa, concentrando-se em quatro países, a saber: Alemanha, Espanha, França e Reino Unido, já interessa por e.g. o índice de preços relativo ou o estado da economia desses mesmos países.

No presente trabalho optamos por trabalhar os seguintes indicadores, por nos parecerem aqueles que permitem influenciar a procura turística e pelo facto de se ter conseguido obter os dados para o período considerado:

- *Índice de preços no consumidor em Portugal (hotéis, restaurantes e serviços relacionados com férias e alojamento), no período entre Janeiro de 1996 e Dezembro de 2008* [Eurostat, 2009]. Optou-se por este indicador dado que tem por finalidade medir a evolução no tempo dos preços de um conjunto de bens e serviços considerados representativos da estrutura de consumo da população residente em Portugal;

- *Temperatura máxima média em Lisboa, com valores mensais de Janeiro de 1990 a Dezembro de 2008* [ECA, 2009]. Apesar de ser referente apenas a Lisboa geofísica, Também, para o intervalo de tempo estudado, não foram encontrados valores das temperaturas nos vários pontos do país ou valores mensais com a média nacional;

- *Taxa de desemprego para os principais mercados estratégicos (Espanha, França, Portugal e Reino Unido) no período entre Janeiro de 1990 e Dezembro de 2008*. Optou-se por inserir esta variável macro económica no modelo porque nos parece ser um indicador relevante da situação económica de um país, podendo ser um dos indicadores indirectos da propensão para o consumo e procura de bens e serviços não essenciais, como é o caso dos serviços prestados no sector turístico. Na verdade é utilizada a variável Número de Desempregados, mas como as variáveis são normalizadas à entrada da RNA, a taxa de desemprego pode, sem cometer um erro significativo, considerar-se proporcional ao Número de Desempregados.

3.4 Apresentação das Séries Temporais

Na Figura 11, pode observar-se a série temporal “Dormidas Mensais Registadas nos Estabelecimentos Hoteleiros em Portugal”. Os dados recolhidos consideram o período compreendido entre Janeiro de 1990 e Dezembro de 2008, correspondendo a 228 dados mensais ao longo de 19 anos.

Optou-se por estudar esta série, uma vez que tem sido considerada como significativa da actividade turística, devido às suas características e uma vez que contempla os visitantes que usufruíram das capacidades turísticas de Portugal nesse período.

Analisando o seu comportamento verifica-se a existência de oscilações periódicas em torno de um valor médio crescente, sugerindo 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.

A Figura 11 mostra que o turismo em Portugal, em particular a procura de estabelecimentos hoteleiros, segue um comportamento bastante típico, evidenciando para cada ano, um pico no mês de Agosto, seguido de uma queda abrupta até Janeiro e novamente uma subida na direcção dos meses de Verão. Esta característica, conhecida como sazonalidade, resulta da associação que existe entre a procura turística e as condições climáticas que a favorecem. Segundo Fúster (1989), a sazonalidade é uma das características do fenómeno turístico. Butler (2001) refere que a sazonalidade pode apresentar-se sob duas formas: a natural e a institucional. Relativamente à primeira,

pode resultar de alterações nas condições climáticas; já a segunda, pode resultar de aspectos relacionados com as instituições sociais, religiosas, étnicas e culturais, sendo menos previsível que a sazonalidade natural.

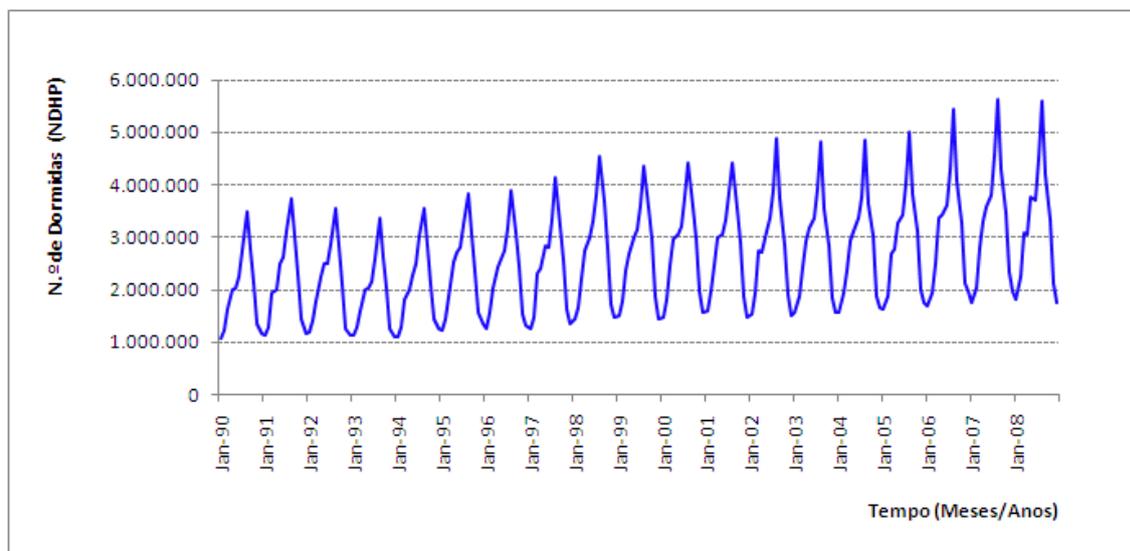


Figura 11 Dormidas dos turistas nas Unidades de Alojamento, em Portugal, no intervalo [Jan-90:Dez-08].

Também se pode constatar, pela análise da figura anterior, que existe uma tendência crescente ao longo do intervalo de tempo abrangido neste estudo, sendo esta manifestamente mais evidente em meses com maior valor absoluto de dormidas. A variação é mais acentuada para os valores da série relativos ao mês de Agosto, se compararmos com a evolução dos valores da série relativos ao mês de Janeiro ao longo dos anos. A tendência registada é resultado do crescimento económico e dos investimentos no sector turístico, que se têm verificado em Portugal nos últimos anos. Verifica-se, no entanto, que esta tendência não é tão linear quanto isso. O início da década de noventa apresenta uma tendência decrescente, que se prolonga até o ano de 1994. O mesmo aconteceu na transição do ano 1998 para 2000. Estas tendências decrescentes coincidem com períodos de recessão ou abrandamento económico. A série apresenta ainda um crescimento muito acentuado entre 2005 e 2007. Este crescimento pode ser resultado do evento desportivo, EURO2004, pois Portugal teve uma projecção turística bastante expressiva além fronteiras, e esses aumentos podem ser fruto desta projecção.

A série, aqui designada por NDHP, representa a soma de um conjunto de sub-séries regionais relativas às seguintes zonas: Norte, Centro, Lisboa, Alentejo, Algarve e Ilhas. As séries das regiões Norte (NDHP-RN) e Centro (NDHP-RC) também serão consideradas como objecto de estudo desta tese. As suas características peculiares, nomeadamente a sua tendência não tão linear relativamente à série mais abrangente NDHP, serão essenciais para avaliar os modelos propostos.

As Figuras 12 e 13 representam respectivamente o número de dormidas em hotéis nas regiões norte e centro. A série NDHP-RC apresenta especialmente um comportamento pouco linear, na medida em que se verifica um crescimento súbito do seu valor no ano de 2003.

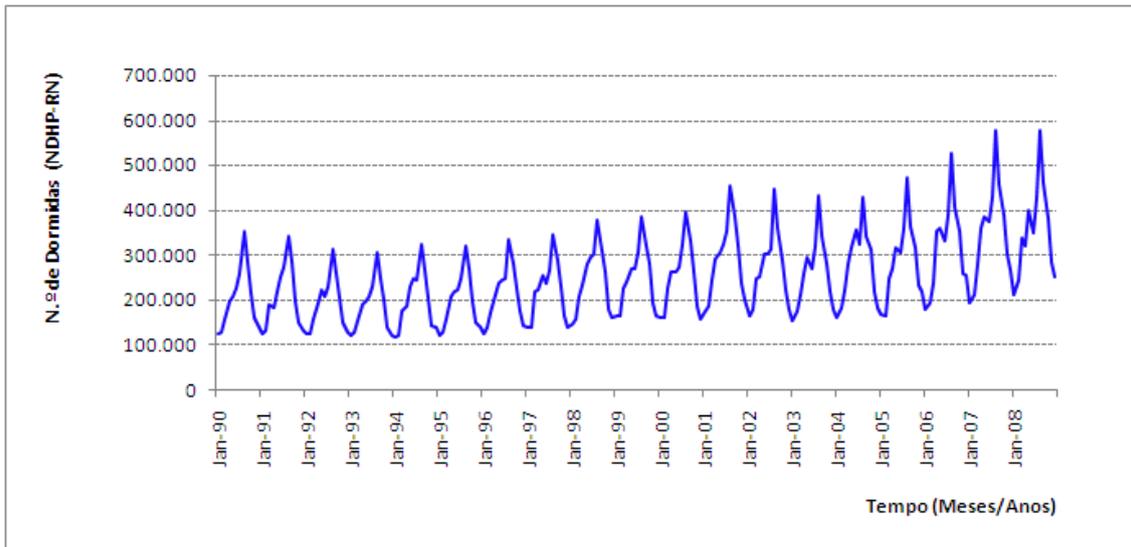


Figura 12 Dormidas dos turistas nas Unidades de Alojamento, na Região Norte de Portugal – NDHP-RN, no intervalo [Jan-90:Dez-08].

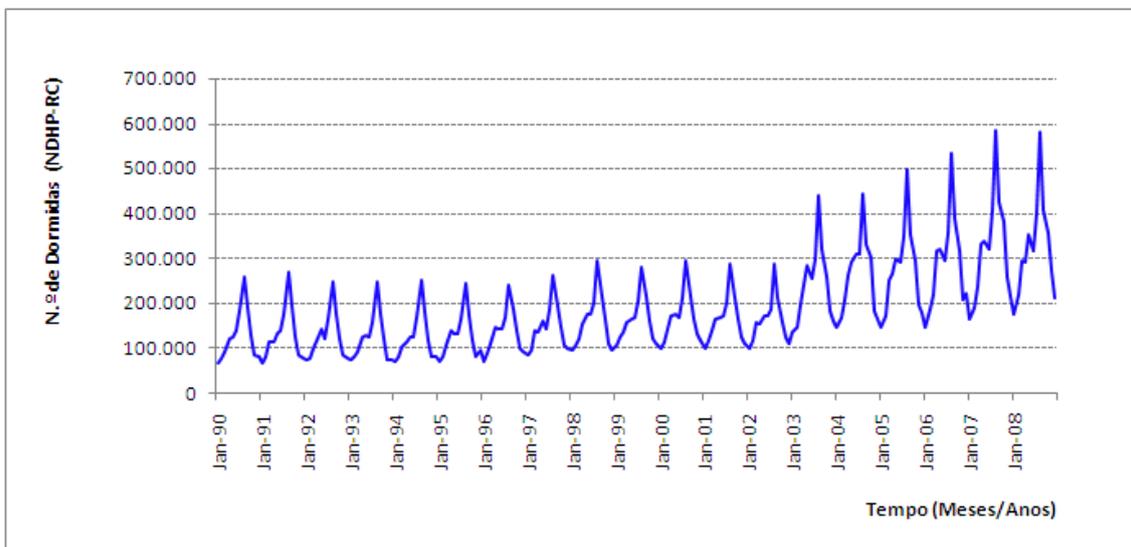


Figura 13 Dormidas dos turistas nas Unidades de Alojamento, na Região Centro de Portugal - NDHP-RC, no intervalo [Jan-90:Dez-08].

Foi também incluída neste estudo a série que representa o número de dormidas em hotéis e estabelecimentos similares, no Reino Unido (NDH-RU). Esta série apresenta um comportamento diferente das séries de procura turística em Portugal. Apesar de se perceber a sua sazonalidade, o seu comportamento ao longo do período compreendido entre Janeiro de 1990 e Dezembro de 2008, varia de forma não linear em amplitude e tendência (Figura 14). Pode observar-se também que o ano 2000 foi um ponto de viragem no comportamento desta série.

O facto de se estudarem séries de natureza diferente, permite validar qualitativamente e avaliar de forma mais concisa os sistemas de modelação linear e não linear, abordados nesta dissertação.

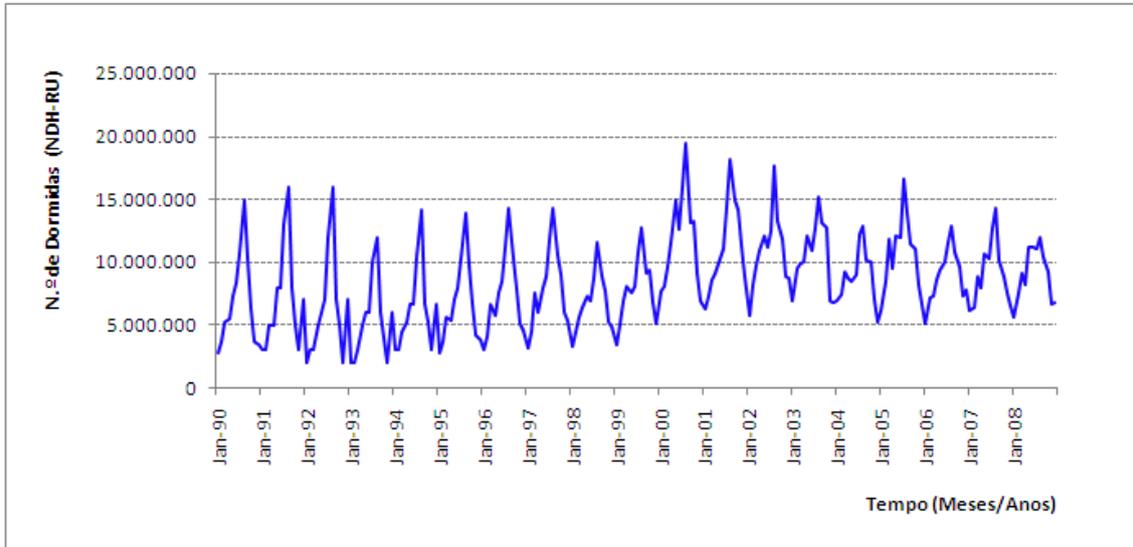


Figura 14 Dormidas dos turistas nas Unidades de Alojamento, no Reino Unido – NDH-RU, no intervalo [Jan-90:Dez-08].

Capítulo 4

4. Modelos de Previsão

Este capítulo faz a abordagem das metodologias desenvolvidas, fazendo uma descrição detalhada dos procedimentos subjacentes à sua aplicação sobre o objecto de estudo já introduzido. Será feita uma descrição dos modelos de regressão linear univariada e redes neuronais artificiais, bem como dos respectivos resultados sobre os conjuntos de dados da série alvo da modelação.

4.1 Metodologias

Foram implementadas duas metodologias de natureza diferente, com objectivo de modelar o comportamento da série apresentada na secção anterior. Uma vez que a série apresenta uma tendência crescente ao longo do intervalo de tempo em estudo, pressupõe-se que a aplicação da metodologia de regressão linear univariada, poderá apresentar bons resultados. Observando a evolução de cada mês individualmente, ao longo do intervalo estudado, verifica-se que a sua tendência é linearmente crescente.

Pretende-se aplicar também a metodologia baseada em redes neuronais artificiais, apresentada no capítulo 2 como uma ferramenta de natureza não linear, caracterizada pela sua, relativamente elevada, capacidade para extrair padrões e modelar funções.

Os tipos de redes neuronais habitualmente mais usadas, por terem melhores resultados, dividem-se entre redes lineares, redes recorrentes, redes radiais e redes *feed-forward* multi-camada. Estas últimas são mesmo as mais utilizadas devido ao sucesso da sua utilização nas mais diversas áreas desde a Engenharia, Economia, Medicina e etc., desde que em 1986 Rumelhart e McClelland apresentaram um algoritmo capaz de as treinar denominado de algoritmo de *Back-Propagation* [Rumelhard & McClelland, 1986].

De uma forma geral, a metodologia baseada em redes neuronais, aplicada neste trabalho, pressupõe várias fases. Numa primeira fase é necessário organizar o conjunto de dados de entrada em subconjunto de treino validação e teste. Considere-se que os conjuntos de dados de entrada constituem não só os registos históricos da série, como também variáveis auxiliares susceptíveis de influenciar a evolução da série. De seguida, procede-se ao pré-processamento dos dados de entrada. O próximo passo é criar uma rede neuronal *feed-forward* seleccionando os seus parâmetros de forma experimental. Finalmente testa-se o poder de generalização da rede com um conjunto de dados desconhecidos, não incluídos no processo de treino.

As próximas secções farão uma descrição detalhada das metodologias aqui introduzidas.

4.2 Regressão Linear

Nesta secção será abordada a modelação por regressão linear, bem como todos os procedimentos, em termos de organização de dados e cálculo, necessários para a sua aplicação. Serão também apresentados graficamente os resultados da modelação da série NDHP (a série no capítulo anterior não foi assim identificada), utilizando esta técnica.

4.2.1 Modelo Univariado

De forma a prever a evolução do valor mensal de dormidas, precedeu-se à determinação das rectas que definem a tendência para cada mês individualmente, criando assim, doze sub-modelos lineares, um para cada mês do ano. Assim sendo, a actual variável independente n , será alvo de uma transformação, passando a ser definida no intervalo $1:i$, com n a representar o número de observações e i o número de anos [Machado, Teixeira e Fernandes, 2009].

Considere-se, S , o vector que contém os valores da série representada na Figura 11, (Tabela 8 dos anexos). A partir de S , define-se a matriz Y_{ij} , (Equação 8).

Equação 8 Forma Matricial para a Série de Dados.

$$Y_{ij} = \begin{bmatrix} S_1 & S_2 & \cdots & S_j \\ S_{j+1} & S_{j+2} & \cdots & S_{2j} \\ \vdots & \vdots & \ddots & \vdots \\ S_{(i-1)j+1} & S_{(i-1)j+2} & \cdots & S_{ij} \end{bmatrix}$$

Onde, j representa o número de meses, e varia entre 1 - Janeiro e 12 - Dezembro.

Para cada coluna da matriz de $Y_{i,j}$, são calculados os coeficientes de regressão linear simples utilizando para a sua estimação o método dos mínimos quadrados ordinários⁶. Pretende-se assim minimizar a soma do quadrado dos erros observados de forma a obter as equações do tipo [Johnston & Dinardo, 2000]:

Equação 9 Regressão Linear.

$$Y_j = \beta_0 + \beta_1 X_j + \varepsilon_j$$

Onde, β_0 , representa a ordenada na origem e β_1 , o declive da recta. Sendo, X_j , a variável independente n para cada mês j , a variar no intervalo $1:i$ e o termo ε , os erros ou resíduos do modelo.

⁶ Coeficientes calculados com base na função *regress* da *toolbox* do *matlab*. Esta função retorna os coeficientes de regressão na forma, $[\beta_1; \beta_0]$, para as entradas Y_i, X_i .

À semelhança do que acontece com outros modelos, também o modelo de regressão tem por base alguns pressupostos, sabendo-se que os principais pressupostos⁷ recaem sobre a componente do erro do modelo, ε_j , e a validação é realizada através dos resíduos [Gujarati, 2003].

O modelo de regressão até aqui definido, será aplicado ao intervalo [Jan-90:Dez-06], de forma a reservar os dois últimos anos para efectuar a previsão e avaliar o modelo. Desta forma as rectas de regressão Y_j , serão definidas para X_j a variar no intervalo [1:17]. Uma vez definidas as rectas de regressão (Tabela 1), procede-se à reconstrução do restante intervalo, calculando a projecção de Y_j , para $X_j = [18:19]$. Ainda de referir que os valores obtidos para os parâmetros dos modelos são significativamente diferentes de zero [Machado, Teixeira e Fernandes, 2009].

Tabela 1 Equações das rectas de regressão linear, mensais; valor da estatística *t-Student* e respectivo valor de prova, para cada coeficiente; coeficiente de correlação.

Y_j	β_0 [<i>t-Student</i> ;valor de prova]	β_1 [<i>t-Student</i> ;valor de prova]	$\beta_0+\beta_1X_j$	r
Y_{Jan}	1 008 736 [31,679; 0,0000]	40 017 [12,877; 0,0000]	1 008 736+40 017 X	0,96975
Y_{Fev}	1 165 255 [24,584; 0,0000]	51 619 [11,159; 0,0000]	1 165 255+51 619 X	0,95642
Y_{Mar}	1 607 897 [19,029; 0,0000]	64 758 [7,854; 0,0000]	1 607 897+64 758 X	0,92493
Y_{Abr}	1 901 811 [20,729; 0,0000]	74 502 [8,321; 0,0000]	1 901 811+74 502 X	0,91674
Y_{Mai}	2 106 436 [25,259; 0,0000]	77 885 [9,5701; 0,0000]	2 106 436+77 885 X	0,94753
Y_{Jun}	2 217 542 [25,311; 0,0000]	84 188 [9,847; 0,0000]	2 217 542+84 188 X	0,94679
Y_{Jul}	2 641 517 [28,347; 0,0000]	93 323 [10,262; 0,0000]	2 641 517+93 323 X	0,95264
Y_{Ago}	3 193 696 [33,259; 0,0000]	118 627 [12,659; 0,0000]	3 193 696+118 627 X	0,96705
Y_{Set}	2 630 092 [30,520; 0,0000]	78 564 [9,342; 0,0000]	2 630 092+78 564 X	0,94403
Y_{Out}	1939 979 [26,554; 0,0000]	76 480 [10,727; 0,0000]	1939 979+76 480 X	0,95486
Y_{Nov}	1 225 761 [22,966; 0,0000]	50 934 [9,779; 0,0000]	1 225 761+50 934 X	0,93976
Y_{Dez}	1 019 961 [24,912; 0,0000]	44 875 [11,231; 0,0000]	1 019 961+44 875 X	0,94826

Nas Figuras 15, 16, 17 e 18, apresentam-se as rectas sobre a dispersão que define a procura turística de cada mês, ao longo de 17 anos. Esta representação permite criar uma perspectiva diferente de abordar o problema, simplificando a sua análise.

⁷ [i] Assume-se que o erro tem média nula e variância constante; [ii] Assume-se que os elementos do erro são independentes entre si; [iii] O erro deve ter uma distribuição Normal para que se possa usar a inferência estatística e generalizar à população os resultados obtidos a partir da amostra [Gujarati, 2003].

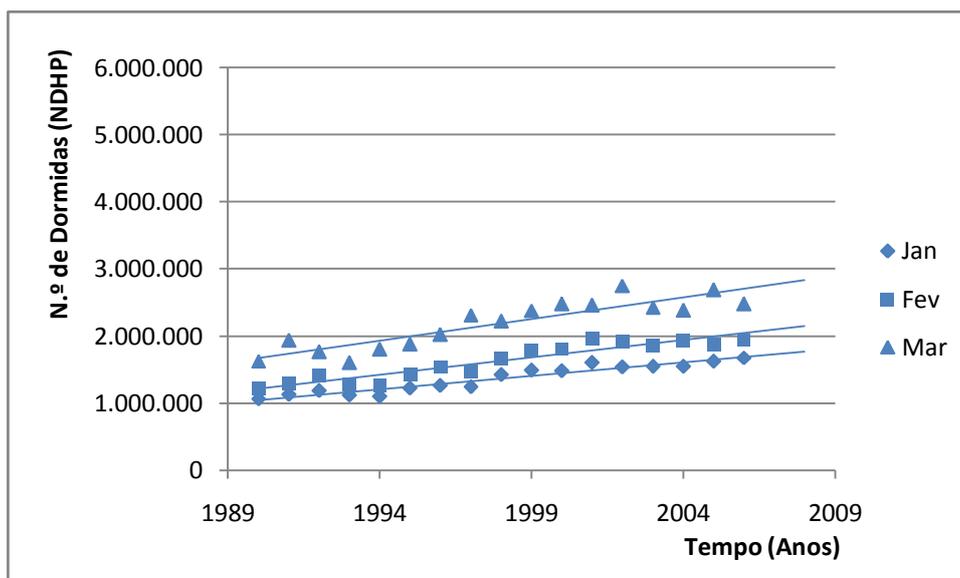


Figura 15 Representação das rectas de regressão para os meses de Janeiro, Fevereiro e Março, no intervalo [1990:2006].

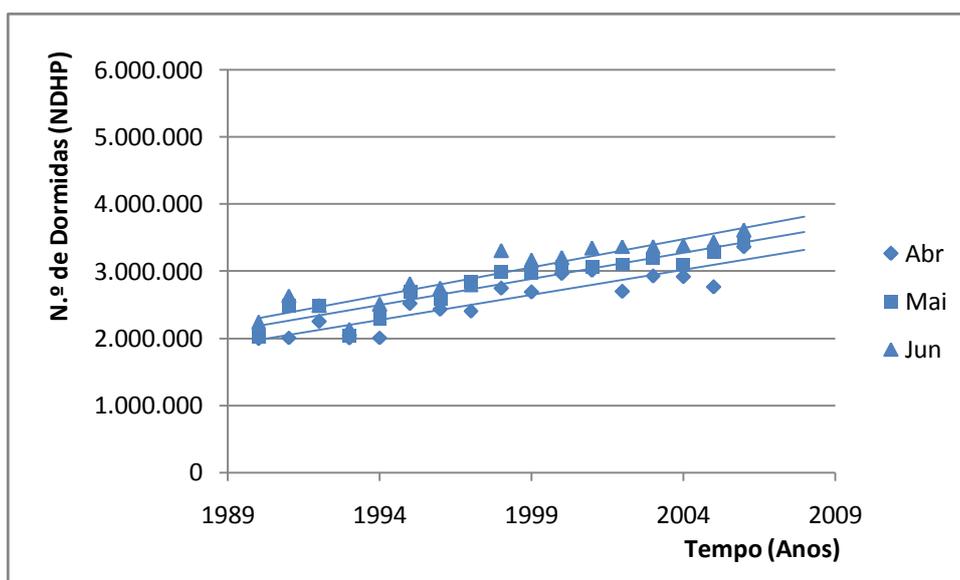


Figura 16 Representação das rectas de regressão para os meses de Abril, Maio e Junho, no intervalo [1990:2006].

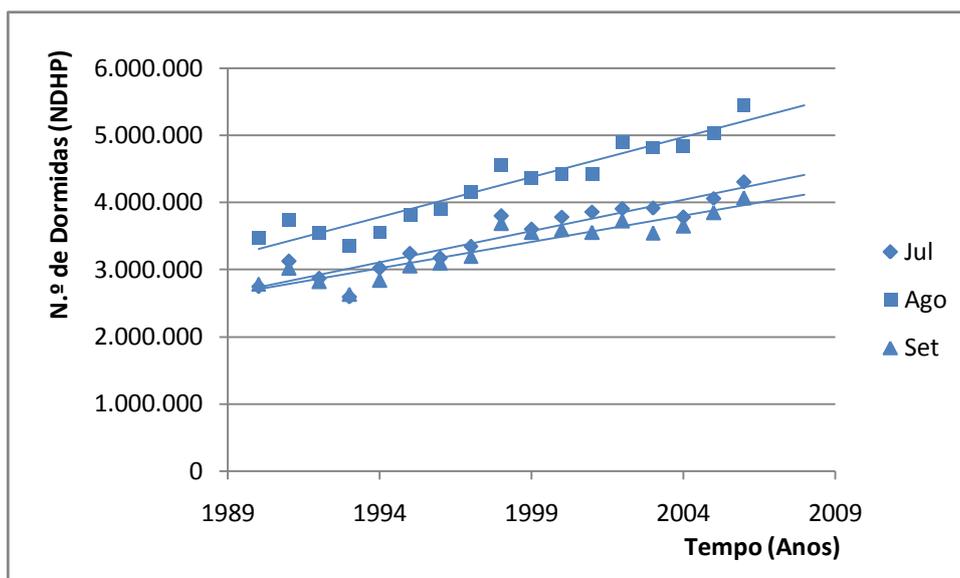


Figura 17 Representação das rectas de regressão para os meses de Julho, Agosto e Setembro, no intervalo [1990:2006].

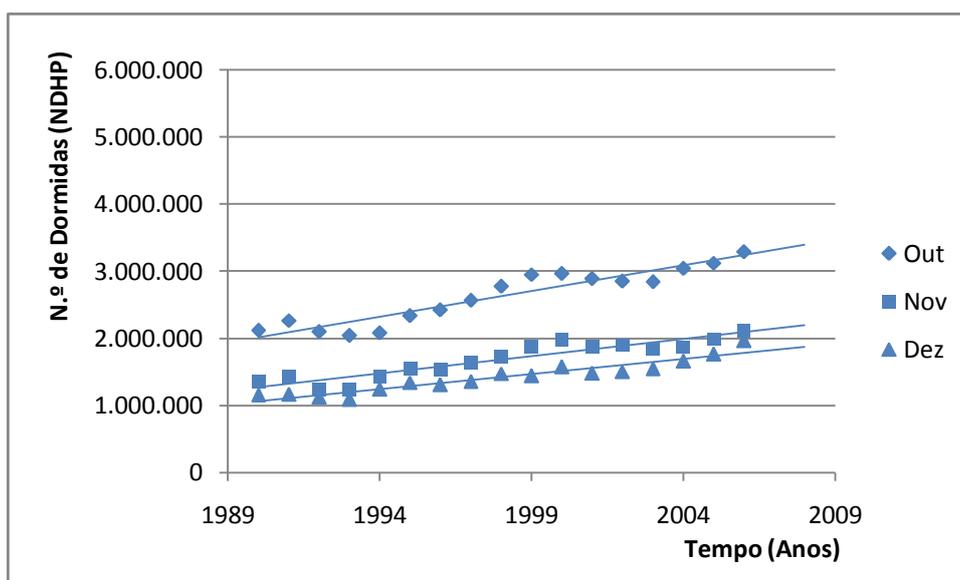


Figura 18 Representação das rectas de regressão para os meses de Outubro, Novembro e Dezembro, no intervalo [1990:2006].

Efectuada a projecção dos períodos referentes aos anos de 2007 e 2008, procedeu-se à inversão da transformação inicial definindo-se novamente a variável independente no intervalo [Jan-07:Dez-08], obtendo-se assim Y'_j com a previsão da série NDHP, conforme ilustrado na Figura 19.

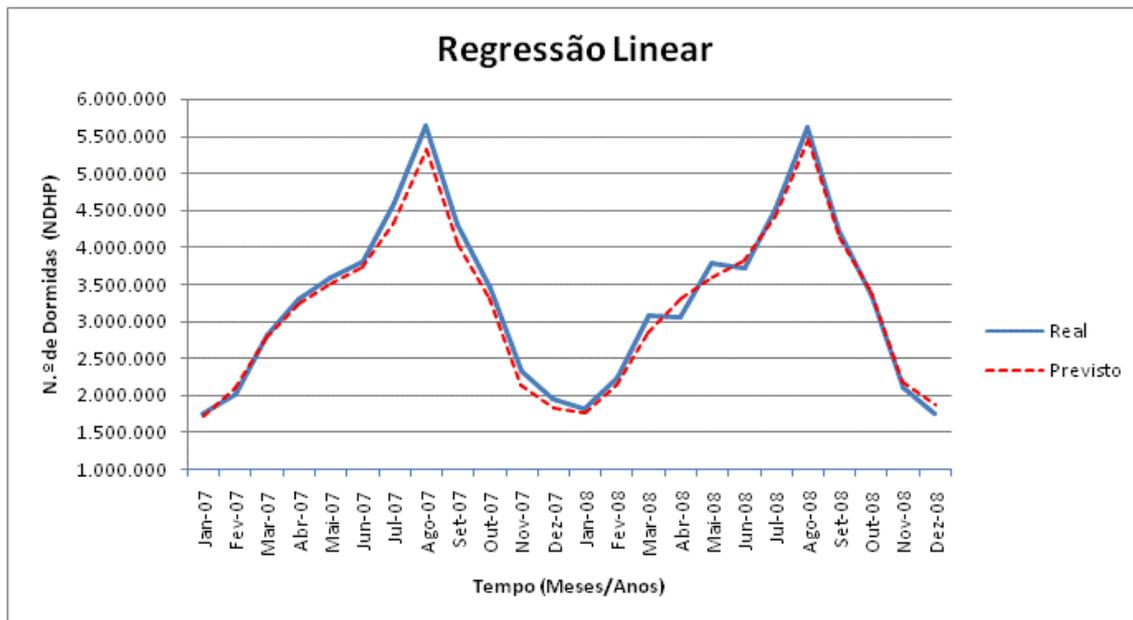


Figura 19 Representação da série NDHP e a respectiva previsão para os anos de 2007 e 2008.

4.2.2 Resultados

De forma a validar a qualidade de ajustamento dos resultados obtidos foi calculado o erro percentual absoluto médio definido como:

Equação 10 Erro Percentual Absoluto Médio

$$EPAM = \frac{1}{N} \sum_{i=1}^n \left| \frac{Y'_i - Y_i}{Y_i} \right|$$

Onde, Y'_i representa o conjunto de valores previstos e Y_i os valores reais. No presente caso, para $N = 24$, o erro obtido foi de 4,16% (Tabela 2).

O valor obtido, utilizando esta técnica de regressão, corresponde, a um resultado satisfatório, se tivermos em conta a simplicidade do método. Tendo por base a classificação do MAPE⁸, ou EPAM, proposto por [Lewis, 1982], pode mesmo dizer-se que o modelo produziu previsões de precisão elevada, uma vez que para o EPAM⁹ obteve-se um valor inferior a 10%.

⁸ Do inglês *Mean Absolute Percentage Error*.

⁹ Um erro absoluto em percentagem pode também ser designado por erro relativo.

Tabela 2 Erro Percentual Absoluto Médio de NDHP, no intervalo [Jan-2007:Dez-2008].

<i>Meses</i>	<i>2007</i>	<i>2008</i>	<i>EPAM</i>
<i>Jan</i>	1,94%	2,94%	
<i>Fev</i>	3,43%	4,16%	
<i>Mar</i>	1,53%	7,81%	
<i>Abr</i>	1,71%	8,27%	
<i>Mai</i>	1,88%	5,20%	
<i>Jun</i>	1,92%	2,56%	
<i>Jul</i>	6,04%	2,55%	4,16%
<i>Ago</i>	5,58%	3,05%	
<i>Set</i>	6,11%	2,40%	
<i>Out</i>	3,98%	1,61%	
<i>Nov</i>	8,07%	3,82%	
<i>Dez</i>	6,62%	6,73%	

O principal problema deste método assenta no facto de ser incapaz de modelar determinados comportamentos de natureza estocástica. Todo o conjunto de amostras que não segue o padrão da maioria, será classificado com um erro acentuado, nomeadamente, as variações identificadas nos meses de Março e Abril de 2008, que registam um erro absoluto de 7,81% e 8,27%, respectivamente (Tabela 10, do Anexo). É também superior à média o erro associado aos meses de Julho, Agosto e Setembro de 2007, de respectivamente 6,04%, 5,58% e 6,11%, resultado da inversão de tendência nos últimos dois anos de NDHP, face à tendência geral dos meses anteriores (Tabela 10, do Anexo).

Projectando Y'_i em todo o domínio de NDHP, consegue ter-se uma ideia mais abrangente dos resultados deste método.

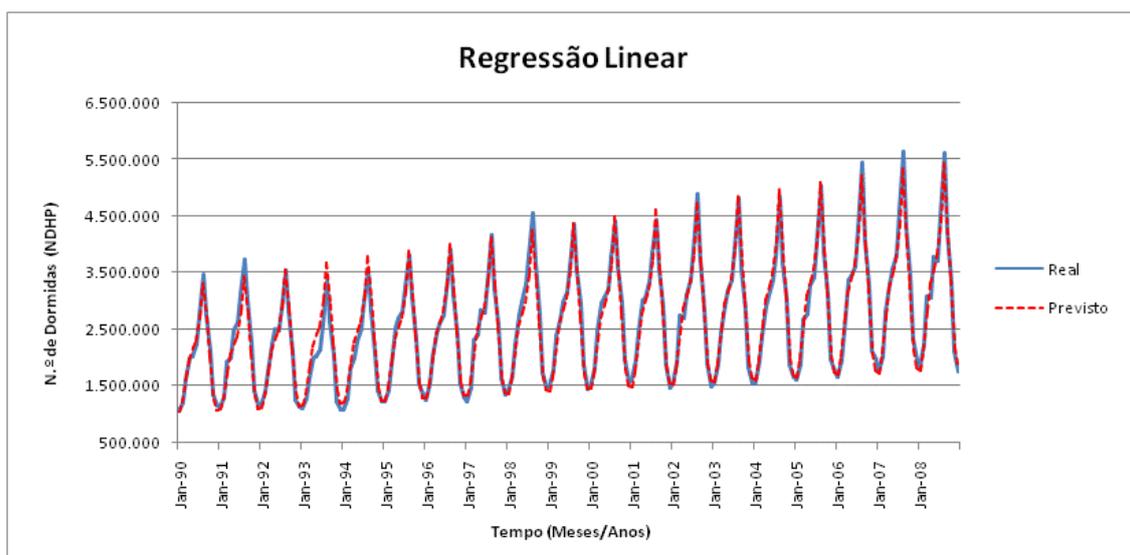


Figura 20 Representação gráfica dos valores previstos sobre a totalidade do conjunto NDHP (RL).

Como se pode observar, na Figura 20, o erro é mais acentuado no ano de 1994, com um erro relativo médio de 12,2% (Tabela 10, do Anexo), sendo o período que registou uma estimativa menos satisfatória.

Também se pode verificar da mesma observação, que existem meses que seguem uma tendência mais próxima do linear, como é o caso do mês de Janeiro, e outros que descrevem um comportamento não tão linear. Março é o mês que ao longo de todo o intervalo, apresenta um maior desvio em relação à tendência calculada (Tabela 10, do Anexo).

4.3 A Rede Neuronal

A abordagem por esta metodologia consiste em considerar as séries apresentadas, constituídas por 228 amostras distribuídas no intervalo que decorre desde Janeiro de 1990 e Dezembro de 2008.

A organização dos dados foi ajustada ao método de previsão mensal, método este designado, *one-step-ahead forecast*. Desta forma, para prever o valor relativo ao mês i da série, inclui-se na entrada da rede o conjunto de valores $i-p:i-1$, em que p define a janela temporal, ou seja, o número de amostras do passado a incluir na previsão de um determinado mês i .

A janela temporal p , foi definida com exactamente 12 amostras, ou seja, são utilizados os valores de um ano para prever o mês imediatamente a seguir, razão pela qual, não existe previsão para o primeiro ano. A dimensão da janela temporal definida, tem produzido bons resultados em trabalhos semelhantes [Fernandes e Teixeira, 2008].

A segmentação do conjunto de dados em subconjuntos de treino, validação e teste, foi efectuada segundo uma percentagem para cada subconjunto. O conjunto de treino deve reunir o maior número possível de amostras, sem prejudicar obviamente os outros conjuntos, de forma a representar convenientemente o universo das observações. Para o efeito seleccionou-se 67% das amostras, portanto, 136 observações (Equação 11). O conjunto de validação, utilizado como critério de paragem durante o processo de treino, de forma a evitar *overfitting*¹⁰, inclui 22% das amostras, portanto, 48 observações, correspondentes a uma selecção aleatória¹¹ no intervalo que decorre entre Janeiro de 1990 e Dezembro de 2006 (Equação 12). Desta forma, o conjunto de validação reúne amostras de diferentes anos, garantindo uma representação transversal da série. O conjunto de dados de teste fica com o restante 11%, constituindo o alvo da previsão (Equação 13). As últimas 24 observações correspondem portanto aos valores da série para os anos de 2007 e 2008. A escolha destas proporções permite também reservar para o processo de teste os últimos dois anos da série.

Equação 11 Conjunto de treino.

$$S_{tr} = 0,67S$$

¹⁰ Perda de capacidade de generalização aquando da classificação do conjunto de teste, devido ao facto de estar demasiado adaptado ao conjunto de treino.

¹¹ Função *rand* da *toolbox* do Matlab.

Equação 12 Conjunto de validação.

$$S_v = 0,22S$$

Equação 13 Conjunto de teste.

$$S_{ts} = 0,11S$$

A arquitectura é definida por uma rede neuronal *feed-forward*, com três camadas, sendo o número de neurónios da camada de entrada definido pela variável p . O número de neurónios da camada escondida foi fixado com base num estudo empírico, que fez variar o número de unidades no intervalo [2:20], com passos de dois. Optou-se por incluir 6 neurónios na camada escondida. A secção abaixo, ‘Parâmetros pré-definidos’ justificará a escolha do número de neurónios na camada escondida.

A camada de saída consiste em apenas 1 neurónio com o resultado da previsão de um determinado mês.

As funções de activação são *logsig*¹² e *purelin*¹³, respectivamente para a camada escondida e para a de saída.

Os algoritmos de treino com os quais se obteve melhor performance foram os algoritmos *Levenberg-Marquardt* e *Resilient Back-Propagation*¹⁴, baseados na retropropagação do erro [Riedmiller & Braun, 1993]. Estes dois algoritmos registaram a melhor performance dentre o conjunto de algoritmos enumerado na secção abaixo.

O processo de treino da rede é um processo iterativo que é interrompido por um dos critérios de paragem: atingir um erro mínimo desejado, atingir um número de iterações de treino ou validação cruzada. No presente estudo foi utilizada a função de performance MSE, que mede o erro quadrático médio e interrompe o processo sempre que é atingido o valor desejado. Utilizou-se um conjunto de validação, já descrito, para paragem antecipada com a finalidade de evitar a perda de generalização por parte da rede, aquando da classificação dos dados de teste. O MSE foi inicializado com um valor de 0,002, o que significa que o critério de paragem accionado será sempre a validação cruzada. O parâmetro que define o número máximo de iterações de treino (também designado por número de épocas), é inicializado a 200. Não é desejável que este critério interrompa o processo de treino, quando se possuem critérios de paragem mais avançados, por isso inicializa-se com um valor relativamente alto. No entanto, por uma questão de tempo de computação, pode eventualmente ser necessário parar o processo.

Ao critério de paragem baseado na validação cruzada, está associado um parâmetro que define o número máximo de iterações sucessivas, nas quais a função de performance no conjunto de validação não foi melhorada. Por omissão esta variável é inicializada a 6 iterações. Isto significa que se ao longo do processo de treino, as actualizações sobre os pesos não resultarem numa melhoria da performance, ao longo de 6 iterações consecutivas, o processo de treino da iteração i é interrompido e o estado do processo de treino volta à iteração $i-6$. Desta forma garante-se que as últimas iterações não contribuirão para a perda de generalização da rede neuronal em relação ao conjunto de teste. Refira-se que o conjunto de teste não é usado no processo de treino.

¹² Função de transferência sigmoideal logarítmica.

¹³ Função de transferência linear.

¹⁴ *Toolbox, Neural Networks* do Matlab.

O processo de treino foi repetido 50 vezes, com diferentes inicializações dos pesos, uma vez que diferentes valores iniciais dos parâmetros da rede levam a diferentes mínimos locais. Das várias repetições, foi seleccionada a rede com menor erro relativo (Erro Percentual Absoluto Médio - EPAM), apresentado na (Equação 10).

4.3.1 Variáveis Auxiliares

Uma das limitações dos modelos de previsão de séries temporais é a sua incapacidade para prever alterações que não se encontram de alguma forma presentes no seu histórico. No contexto específico da metodologia de redes neuronais, será impossível para a rede prever características de uma determinada série, se não tiver a oportunidade de lidar, no processo de treino com padrões de alguma forma semelhantes. Desta forma surge a necessidade de dotar o sistema de previsão de um mecanismo que permita a inclusão de variáveis auxiliares no conjunto de entrada da rede neuronal [Law, 1999].

Estudou-se o efeito de determinadas variáveis susceptíveis de influenciar a série alvo da modelação. Assim, a entrada da RNA, poderá incluir nas suas amostras não só os dados históricos da série, mas também variáveis correlacionadas. A correlação é calculada sempre que o utilizador seleccione uma determinada variável. Cada variável constitui um vector de valores compreendidos no intervalo de tempo definido pela série que se pretende modelar.

Para o efeito foi seleccionado:

- O índice de preços no consumidor em Portugal, compreendido no intervalo que decorre entre Janeiro de 1996 e Dezembro de 2008. Este índice refere-se apenas aos preços de hotéis, restaurantes e serviços relacionados com férias e alojamento [Eurostat, 2009];
- A taxa de desemprego nacional e dos principais mercados estratégicos (Reino Unido, Espanha e França) no período entre Janeiro de 1990 e Dezembro de 2008 [Eurostat, 2009];
- Uma variável representativa do estado do tempo em Portugal – temperatura máxima média em Lisboa, com valores mensais de Janeiro de 1990 a Dezembro de 2008 [ECA, 2009].

O mecanismo para incluir os vectores de variáveis auxiliares no processo de treino da rede neuronal é relativamente simples, adicionando à entrada o seu valor em paralelo com os dados históricos da série.

Desta forma, se for incluída uma variável auxiliar, e.g. temperatura média em Lisboa no intervalo de tempo que decorre entre Janeiro 1990 e Dezembro 2008, valor relativo ao mês i da série, será previsto, contando à entrada com o conjunto de valores $S_i - p, \dots, S_i - 2, S_i - 1, v_i$, em que p define a janela temporal e v_i o valor da variável auxiliar. Generalizando para mais de uma variável, a entrada da rede fica organizada do seguinte modo $S_i - p, \dots, S_i - 2, S_i - 1, v_j^1, \dots, v_j^2, v_j^1$, sendo que v_j^i corresponde ao valor da variável auxiliar j relativo ao mês i . Deste modo, não estão a ser incluídos os valores históricos das variáveis auxiliares. De forma geral, o aumento acentuado da procura turística no mês de Agosto, deve-se às excelentes condições meteorológicas nesse mesmo mês, e não necessariamente à sequência formada pelos últimos p meses. A mesma explicação pode não ser tão linear em relação às variáveis de índole económica, no entanto manteve-se o mesmo critério de organização.

4.3.2 Pré processamento dos Dados

O pré-processamento dos dados de entrada compreende o conjunto de operações realizadas sobre o *dataset*, imediatamente antes de estes alimentarem a rede neuronal.

Basicamente, este tipo de pré-processamento normaliza o conjunto de entrada no intervalo $[-1:1]$ sem alterar a sua variância dos dados originais. A Equação 14 apresenta a fórmula aplicada ao conjunto de entrada x , para um limite desejado compreendido no intervalo $[y_{\min}, y_{\max}]$. Este processamento é apenas aplicado às variáveis de entrada designadas por auxiliares.

Equação 14 Normalização dos dados.

$$y = \left[\frac{y_{\max} - y_{\min}}{x_{\max} - x_{\min}} (x - x_{\min}) \right] + y_{\min}$$

4.3.3 Parâmetros pré-definidos

Uma das questões mais importantes quando se define uma rede neuronal, é determinar o valor dos seus principais parâmetros. Uma rede neuronal *feed-forward*, pode apresentar diferentes arquiteturas, com mais ou menos neurónios, recorrendo a diferentes algoritmos de treino ou critérios de paragem. Dependendo do problema em causa, uma determinada rede pode comprometer a sua performance se não for convenientemente definida.

Recorrendo a uma técnica puramente empírica, realizaram-se um conjunto de experiências, de forma a justificar a fixação de determinados parâmetros da rede neuronal, nomeadamente o número de neurónios na camada escondida, o algoritmo de treino e o número máximo de validações falhadas em cada iteração de treino.

Fixando-se os parâmetros que definem a organização dos conjuntos de dados de entrada, bem como os parâmetros da rede neuronal cuja variação não influencia preponderantemente a performance da rede, fez-se variar o algoritmo de treino e o número de neurónios na camada oculta. A taxa de aprendizagem não foi incluída uma vez que os algoritmos de treino utilizados fazem variar o seu valor ao longo do processo. O valor definido para este parâmetro, que pode variar no intervalo $[0:1]$, conta apenas como uma inicialização.

A Figura 21 representa graficamente o erro associado a cada algoritmo de treino, referente ao melhor registo de entre um conjunto de 50 processos de treino, realizados nas mesmas condições, sem a inclusão de variáveis auxiliares. O processo foi repetido devido às alterações produzidas pela inicialização aleatória dos pesos. A performance apresentada é resultado do erro percentual absoluto médio medido nos conjuntos de treino e validação, não sendo desta forma incluído o conjunto de teste. O número de neurónios varia no intervalo $[2:20]$, com passos de duas unidades. Como se pode observar, o algoritmo de treino *Levenberg-Marquardt*, seguido do *Resilient-Back-Propagation* registam as melhores performances.

Quanto ao número de neurónios, pode constatar-se que a partir de 14 neurónios, não se verifica melhoria na performance (Tabela 11). Um valor inferior a 6 nós, também não parece uma boa opção.

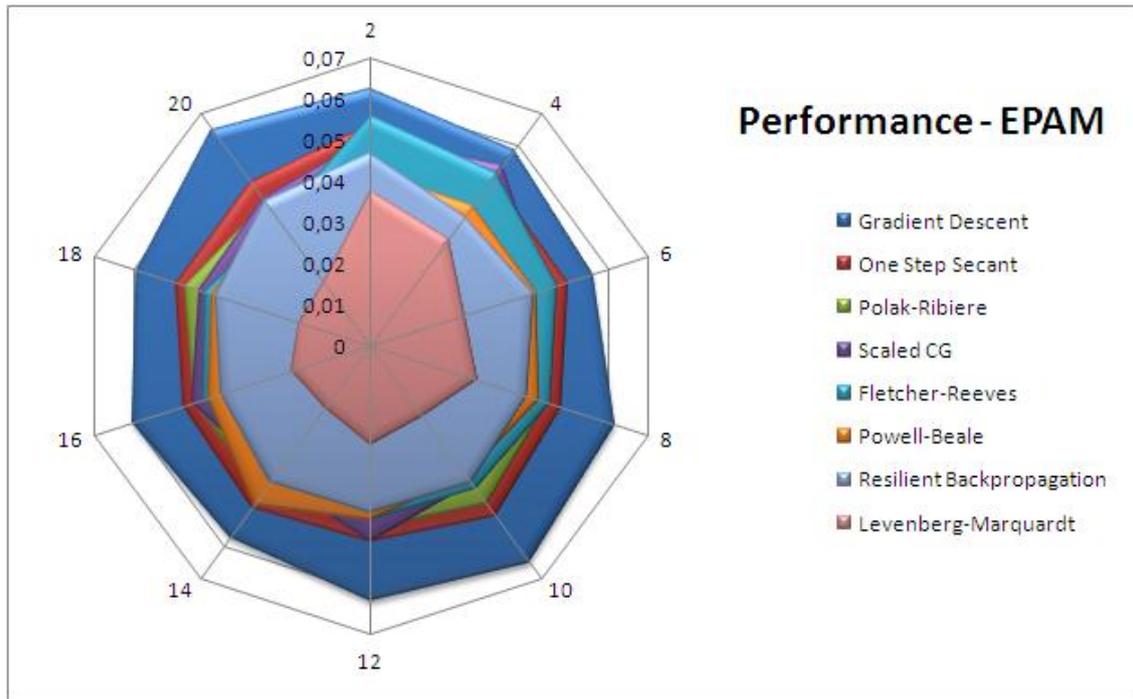


Figura 21 EPAM – Algoritmos de Treino/n.º de Neurónios.

O tempo de computação pode também ser um critério relevante na escolha de determinada arquitectura ou algoritmo de treino. A Figura 22 mostra que os algoritmos que convergem mais rapidamente são também aqueles que apresentam melhor performance. Em relação ao número de neurónios, observa-se que estes não estabelecem nenhuma relação facilmente perceptível com o tempo de execução. De uma maneira geral, um maior número de neurónios deveria ser computacionalmente mais pesado, no entanto, o poder de convergência também tem tendência a aumentar com o incremento do número de unidades, e por consequência, o tempo de processamento tende a diminuir. No caso do *Gradient Descent*, verifica-se o contrário, sendo que o aumento do número de neurónios provoca um significativo atraso no tempo de execução Tabela 12.

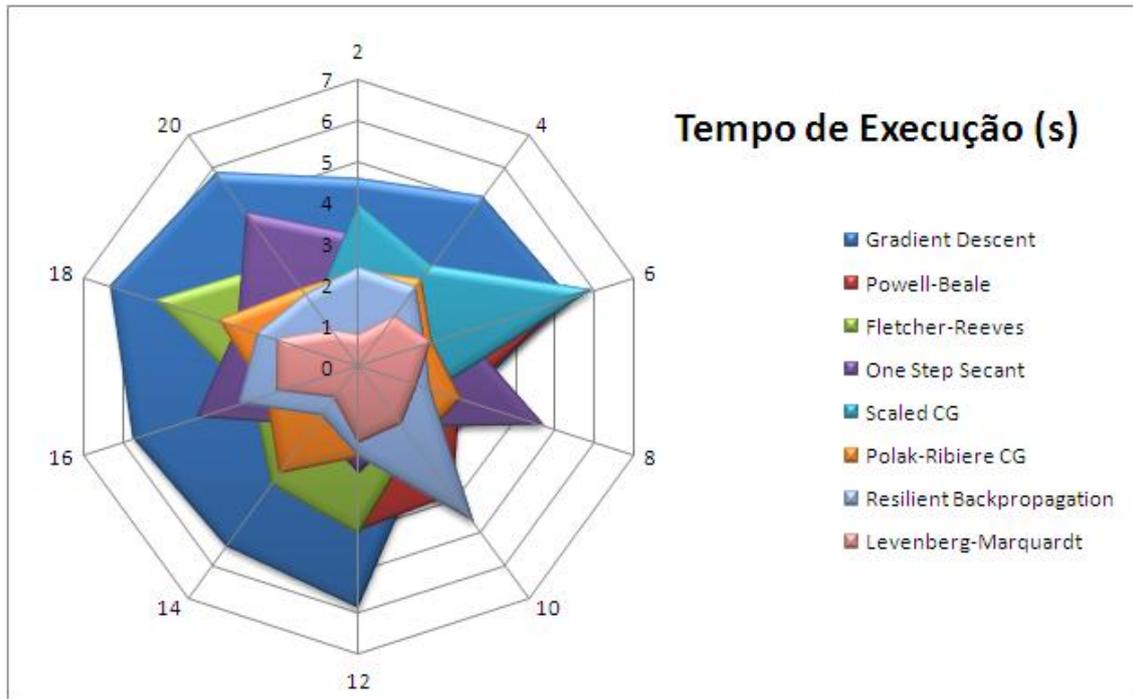


Figura 22 Tempo de Execução (s - segundos) – Algoritmos de Treino/n.º de Neurónios.

O procedimento apresentado permite uma visão geral da performance dos diferentes algoritmos, para este objecto de estudo específico. Convém referir que o carácter deste processo, não é vinculativo em relação à escolha do algoritmo de treino ou à determinação do número de neurónios na camada escondida, sendo apenas válido nas condições impostas pelos parâmetros que não se fizeram variar. Por exemplo, a dimensão do conjunto de validação, bem como o número de iterações falhadas em cada processo de treino, no contexto dos critérios de paragem baseados na validação cruzada, pode ser insuficiente para interromper um algoritmo que converge em relativamente poucas iterações, como é o caso do algoritmo de *Levenberg-Marquardt*. O facto de se aumentar o número de neurónios, pode também conceder à rede uma capacidade de tal forma grande para convergir sobre o conjunto de treino, que compromete a capacidade de generalização do modelo. Se o conjunto de validação não for suficientemente representativo do problema, então, poder-se-á estar a prejudicar o processo de treino, tanto mais quanto maior for o número de neurónios.

4.3.4 Resultados

Nesta secção será testada a capacidade de generalização da rede neuronal sobre o conjunto de dados de teste. Este procedimento permite qualificar todo o processo de treino, não só a selecção dos parâmetros, como os critérios de paragem utilizados.

Como foi referido, foram reservadas 11% das amostras para testar a rede. Este conjunto corresponde aos valores da série relativos aos anos de 2008 e 2009.

Para o efeito foi utilizada a rede neuronal *feed-forward*, com a estrutura acima referida aplicando os dois melhores algoritmos de treino (*Levenberg-Marquardt* e *Resilient-Back-Propagation*). O EPAM de cada algoritmo é de respectivamente 0,052 e 0,048 para LM e RP (Tabela 15 e Tabela 16).

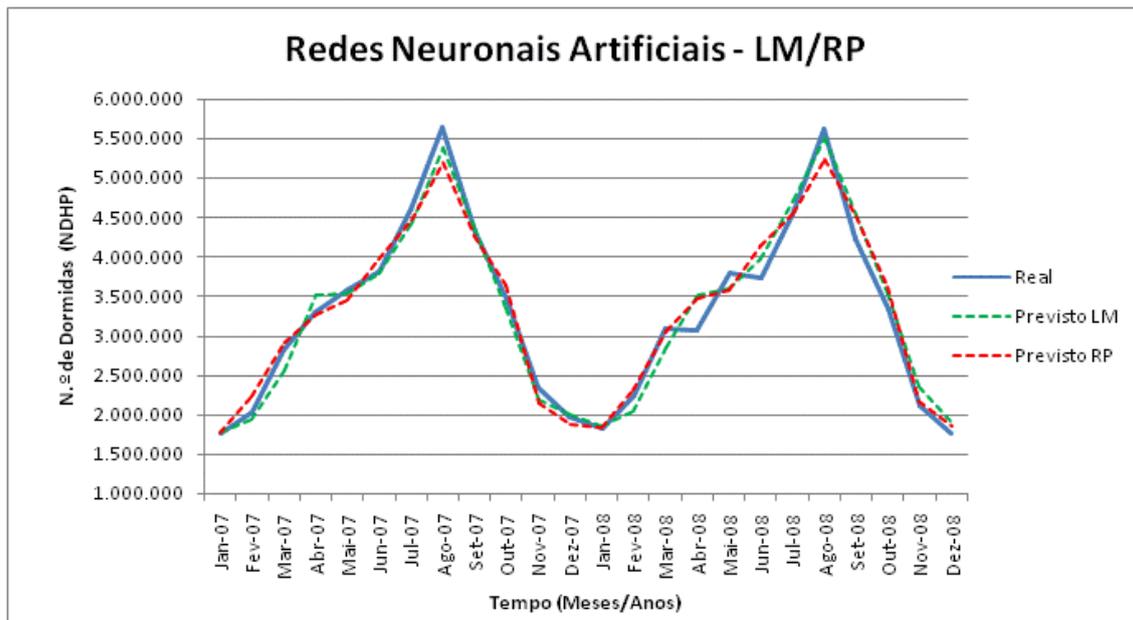


Figura 23 Representação da série NDHP e a respectiva previsão para os anos de 2007 e 2008 (RNA – LM/RP).

Representando a modelação sobre todo o conjunto de NDHP, observa-se que qualquer um dos algoritmos, permite aproximar de perto a evolução da série, com um erro mínimo. O erro subjacente à modelação efectuada pelo algoritmo de treino LM (EPAM = 0,052), sob o conjunto de treino, com excepção do ano de 2000, foi ligeiramente superior ao erro obtido pelo algoritmo de treino RP (EPAM = 0,038). Em relação ao conjunto de dados de teste, o algoritmo LM consegue superar o RP no ano de 2007, com um erro de apenas 0,036, perdendo em relação ano de 2008, apresentando um erro de 0,068, enquanto o RP consegue modelar este ano com um erro de apenas 0,053.

Em relação à capacidade de modelação destes algoritmos é interessante notar que ambos apresentam a mesma dificuldade em modelar os meses de Março e Abril. O algoritmo LM apresenta um máximo na sua função de erro médio em Abril de 0,081, seguido de Março com 0,074. O algoritmo de treino RP, apresenta um comportamento semelhante. Nesta altura do ano a procura por estabelecimentos hoteleiros segue um comportamento aparentemente estocástico, ou ainda não correlacionado com as variáveis usadas, difícil de modelar (Tabela 15 e Tabela 16).

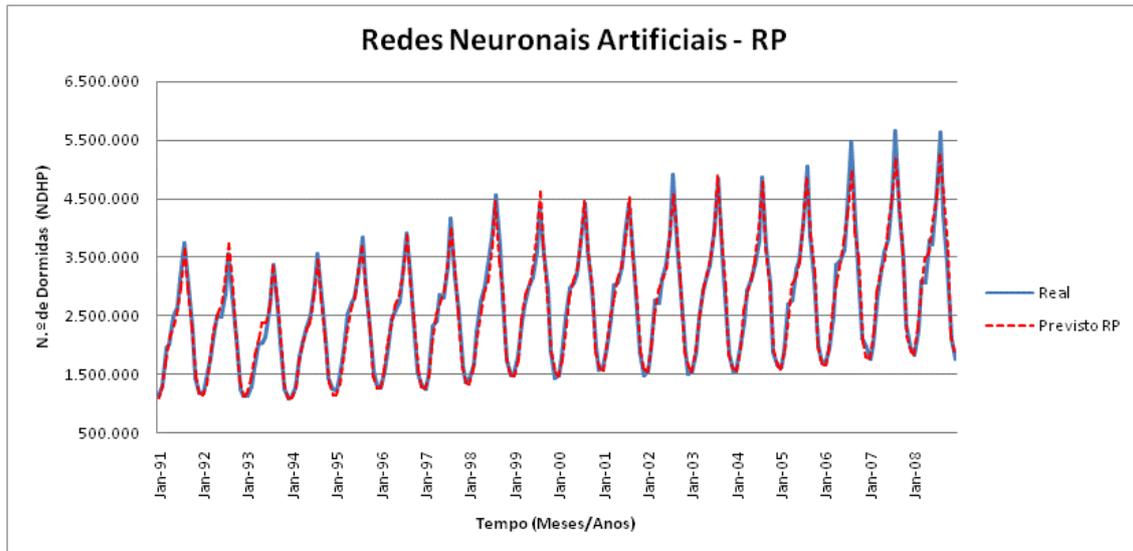


Figura 24 Representação gráfica dos valores previstos sobre a totalidade do conjunto NDHP (RNA - RP).

Referira-se que os resultados obtidos não incluem variáveis auxiliares. Devido à qualidade de ajustamento obtida, com recurso apenas aos valores históricos da série, a inclusão das variáveis seleccionadas, não acrescentou qualquer melhoria na performance da rede. A aplicação destas variáveis, será guardada para uma secção posterior na qual se irá modelar uma série de natureza ligeiramente diferente.

Capítulo 5

5. Análise dos Resultados

As secções seguintes são dedicadas ao estudo comparativo das metodologias abordadas, a regressão linear e as redes neuronais artificiais. Serão abordadas as vantagens de cada uma das metodologias e tiradas algumas conclusões acerca da dependência dos modelos em relação ao objecto de estudo.

5.1 Regressão Linear *versus* Redes Neuronais

Nas condições definidas em secções anteriores, Contrapõem-se agora ambos os modelos, de forma a determinar qual se adapta melhor à natureza da série NDHP.

Comparando as últimas N amostras ($N = 24$), referentes aos anos de 2007 e 2008, observa-se que ambos os modelos apresentam qualidades de ajustamento bastante satisfatório com um EPAM de 4,16% e 4,84% respectivamente para os modelos de regressão linear e redes neuronais artificiais (Tabela 3).

Verifica-se que o erro foi mais acentuado em amostras que não seguem o padrão da maioria, como é o caso das variações identificadas nos meses de Março, Abril e Maio de 2008, que registam um erro absoluto de quase 7%. Este erro deve-se ao comportamento irregular, de difícil modelação, da série para estes meses.

Tabela 3 Erro percentual mensal e EPAM no intervalo [Jan-2007:Dez-2008] para ambos os modelos, regressão linear (RL) e redes neuronais artificiais (RNA).

<i>Meses</i>	<i>2007</i>		<i>2008</i>		<i>EPAM</i>	
	<i>RL</i>	<i>RNA</i>	<i>RL</i>	<i>RNA</i>	<i>RL</i>	<i>RNA</i>
<i>Jan</i>	1,94%	0,06%	2,94%	0,49%		
<i>Fev</i>	3,43%	9,88%	4,16%	3,11%		
<i>Mar</i>	1,53%	3,54%	7,81%	0,52%		
<i>Abr</i>	1,71%	1,33%	8,27%	13,49%		
<i>Mai</i>	1,88%	3,60%	5,20%	5,65%		
<i>Jun</i>	1,92%	4,29%	2,56%	10,86%	4,16%	4,84%
<i>Jul</i>	6,04%	2,90%	2,55%	0,66%		
<i>Ago</i>	5,58%	8,16%	3,05%	6,75%		
<i>Set</i>	6,11%	0,89%	2,40%	6,97%		
<i>Out</i>	3,98%	5,12%	1,61%	8,20%		
<i>Nov</i>	8,07%	7,97%	3,82%	2,51%		
<i>Dez</i>	6,62%	3,83%	6,73%	5,44%		

De forma geral, os resultados não mostram uma grande distinção no desempenho entre os modelos estudados. Como se pode observar na Figura 25, ambos os modelos acompanham o comportamento da série original, sendo neste caso, o modelo baseado em regressão linear a conseguir, mesmo que de forma ligeira, a melhor aproximação.

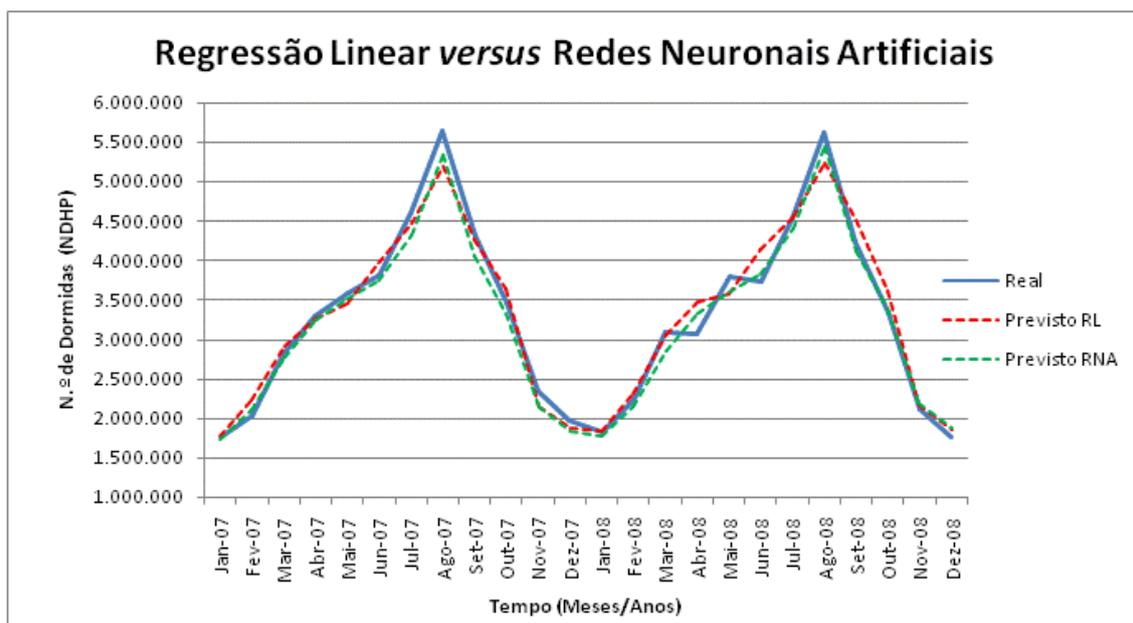


Figura 25 Representação da série e a respectiva previsão para os anos de 2007 e 2008, pelos modelos RL e RNA.

A Tabela 10 mostra também que o erro é mais acentuado no ano de 1994, com o modelo RL, com um EPAM para esse ano de 12,2%, sendo o período que registou a estimativa menos satisfatória.

O modelo baseado em RNA, acompanha de perto o comportamento da série, nomeadamente, para os anos de 1993/1994 e 2000/2001, onde a tendência da série se

inverteu ligeiramente. Nestes períodos a performance baseada em redes neuronais é cerca do dobro da performance obtida através da regressão linear. Isto deve-se fundamentalmente ao facto de se estar perante uma modelação sobre dados de treino, e portanto a aproximação é tanto melhor quanto maior o número de iterações do algoritmo de treino. No limite, se o processo de treino não fosse interrompido, a aproximação seria ainda maior.

Também se pode verificar da mesma observação, que existem meses que seguem uma tendência mais próxima do linear, como é o caso do mês de Janeiro, e outros que descrevem um comportamento não tão linear. Março é o mês que ao longo de todo o intervalo, apresenta um maior desvio em relação à tendência calculada.

Os valores obtidos correspondem a um resultado satisfatório em ambos os modelos. Atendendo à classificação do MAPE, ou EPAM, proposto por [Lewis, 1982], pode mesmo dizer-se que ambos os modelos produziram previsões de precisão elevada, uma vez que para o EPAM se obteve um valor inferior a 10%.

Dada a complexidade da aplicação do modelo de redes neuronais, até agora não se têm mostrado argumentos que justifiquem a sua aplicação. O modelo de regressão linear mostra ser mais eficiente e elegante na modelação da série NDHP.

A secção seguinte mostrará porque as redes neuronais podem sair a ganhar na modelação deste tipo de séries.

5.1.1 Séries Turísticas Regionais

Esta secção é dedicada ao estudo comparativo dos modelos apresentados, aplicados às séries da região norte e centro de Portugal.

Região Norte

Contida na série NDHP, a série que regista a procura turística na região norte de Portugal, segue um comportamento, não tão linear quanto a série NDHP, apresentando nos últimos três anos um crescimento mais acentuado (Tabela 17). Por esta razão, o modelo de regressão univariado não vai além de uma modelação com um erro de 11,88% para os anos de 2007 e 2008. O modelo baseado em redes neuronais artificiais, consegue uma performance melhor, de 6,89% (Tabela 4). Relativamente ao ano de 2007, esta metodologia não linear apresenta uma performance de 4,38%, quase três vezes melhor à performance do modelo de regressão linear, com um EPAM de 12,48%.

Tabela 4 Erro percentual mensal e EPAM no intervalo [Jan-2007:Dez-2008] para ambos os modelos, regressão linear (RL) e redes neurais artificiais (RNA) relativo à série NDHP-RN.

<i>Meses</i>	<i>2007</i>		<i>2008</i>		<i>EPAM</i>	
	<i>RL</i>	<i>RNA</i>	<i>RL</i>	<i>RNA</i>	<i>RL</i>	<i>RNA</i>
<i>Jan</i>	6,43%	3,20%	13,31%	5,27%		
<i>Fev</i>	7,96%	4,07%	18,40%	11,40%		
<i>Mar</i>	8,67%	3,81%	22,13%	12,14%		
<i>Abr</i>	13,67%	1,95%	0,37%	19,89%		
<i>Mai</i>	11,22%	2,55%	11,70%	2,57%		
<i>Jun</i>	10,12%	4,48%	1,00%	10,45%	11,88%	6,89%
<i>Jul</i>	14,26%	7,77%	12,24%	5,29%		
<i>Ago</i>	15,24%	4,79%	13,14%	0,58%		
<i>Set</i>	15,43%	2,33%	14,28%	9,86%		
<i>Out</i>	12,35%	1,04%	8,61%	14,17%		
<i>Nov</i>	17,69%	7,92%	9,70%	11,03%		
<i>Dez</i>	16,76%	8,66%	10,56%	10,03%		

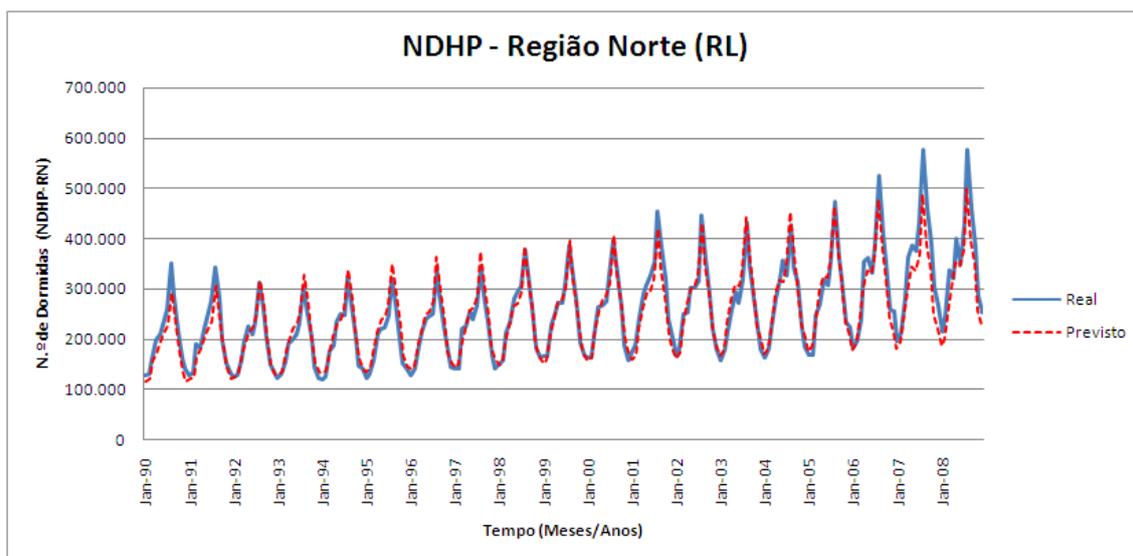


Figura 26 Representação gráfica dos valores previstos da série NDHP-RN (RL).

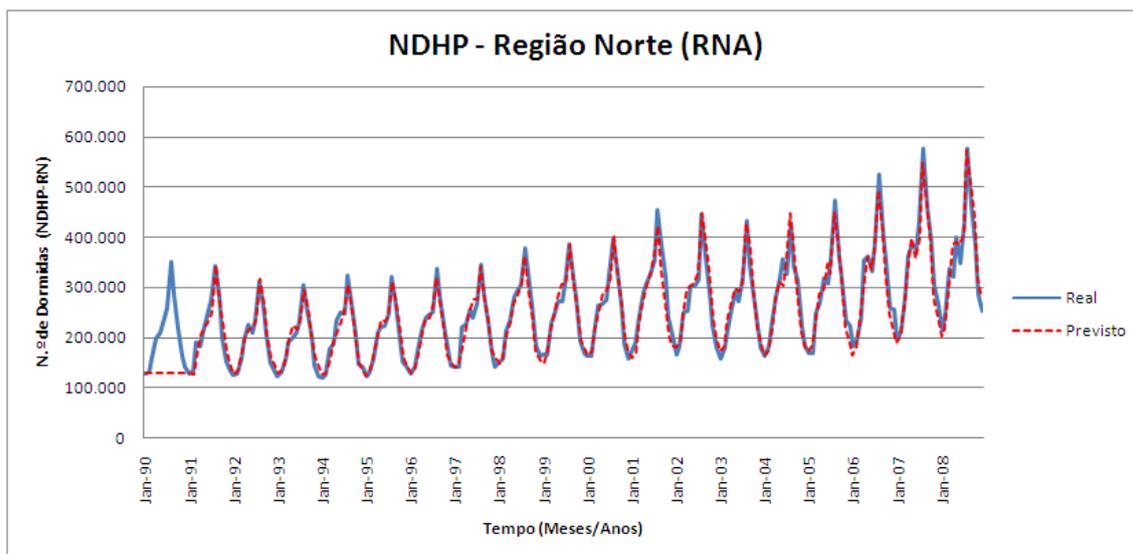


Figura 27 Representação gráfica dos valores previstos da série NDHP-RN (RNA).

Região Centro

A procura turística na zona centro, compreendida no mesmo intervalo de tempo, segue um comportamento pouco linear, especialmente a partir do ano de 2003 (Tabela 22).

Aplicando as metodologias de regressão linear e redes neuronais artificiais à série NDHP-RC (Região Centro) obteve-se um erro respectivamente de 16,47% e 8,21%. Estes resultados são bem elucidativos da capacidade das redes neuronais artificiais para modelar comportamentos não lineares. A performance das redes neuronais é o dobro da obtida com regressão linear (Tabela 5).

Tabela 5 Erro percentual mensal e EPAM no intervalo [Jan-2007:Dez-2008] para ambos os modelos, regressão linear (RL) e redes neuronais artificiais (RNA) relativo à série NDHP-RC.

Meses	2007		2008		EPAM	
	RL	RNA	RL	RNA	RL	RNA
Jan	10,40%	19,73%	12,79%	16,96%		
Fev	9,79%	4,80%	18,20%	5,26%		
Mar	7,63%	17,75%	22,02%	18,73%		
Abr	19,24%	2,17%	4,94%	0,05%		
Mai	14,10%	7,05%	14,55%	0,48%		
Jun	12,40%	6,39%	8,51%	4,95%	16,47%	8,21%
Jul	22,78%	12,30%	19,38%	3,14%		
Ago	22,01%	9,34%	18,81%	18,30%		
Set	21,60%	12,26%	15,18%	2,53%		
Out	25,96%	4,62%	17,67%	9,32%		
Nov	23,87%	11,10%	25,74%	3,88%		
Dez	15,87%	1,78%	11,89%	4,18%		

Devido à mudança brusca a partir do ano de 2003, fazendo subir o valor absoluto da série, o cálculo da regressão linear no ano de 2002 apresenta um EPAM de 29,18%. Este salto a meio da série contribui substancialmente para aumentar o erro absoluto médio nos extremos da série. Desta forma, o erro é acentuado nos anos de 1990/1992, com um erro respectivo de 28,01% e 21,15%. No outro extremo encontram-se os anos de 2007 e 2008 com um erro de 17,14% e 15,81% respectivamente (Tabela 25).

As redes neuronais lidaram bem com esse salto, conseguindo uma boa performance no conjunto de teste, especialmente no ano de 2008. O facto de se ter utilizado um conjunto de validação aleatório com 48 amostras (22%), permitiu incluir no processo de treino amostras de diferentes amplitudes, preparando a rede para a modelação dos dois últimos anos Figura 29.

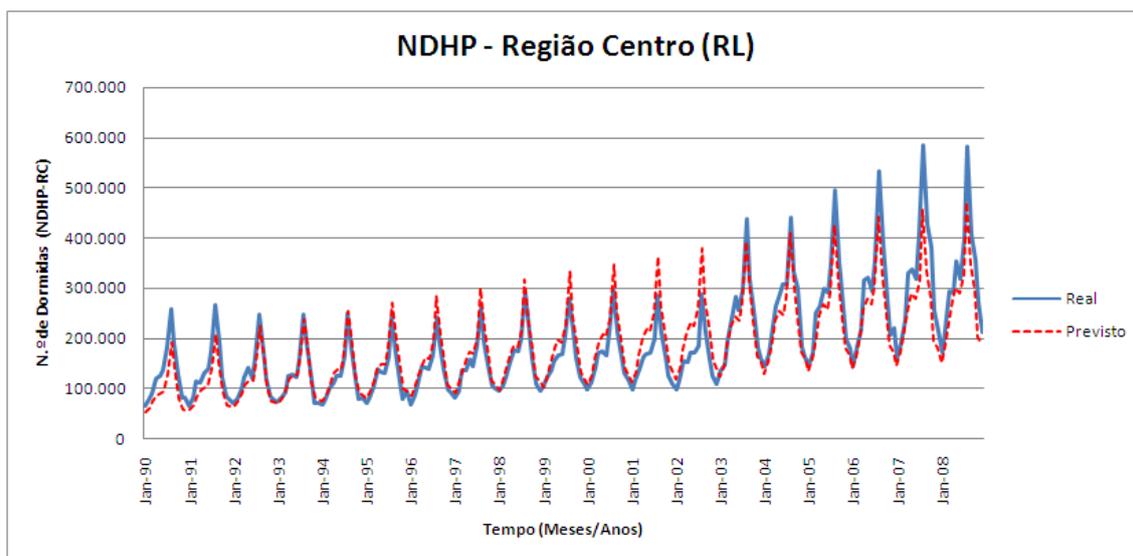


Figura 28 Representação gráfica dos valores previstos da série NDHP-RC (RL).

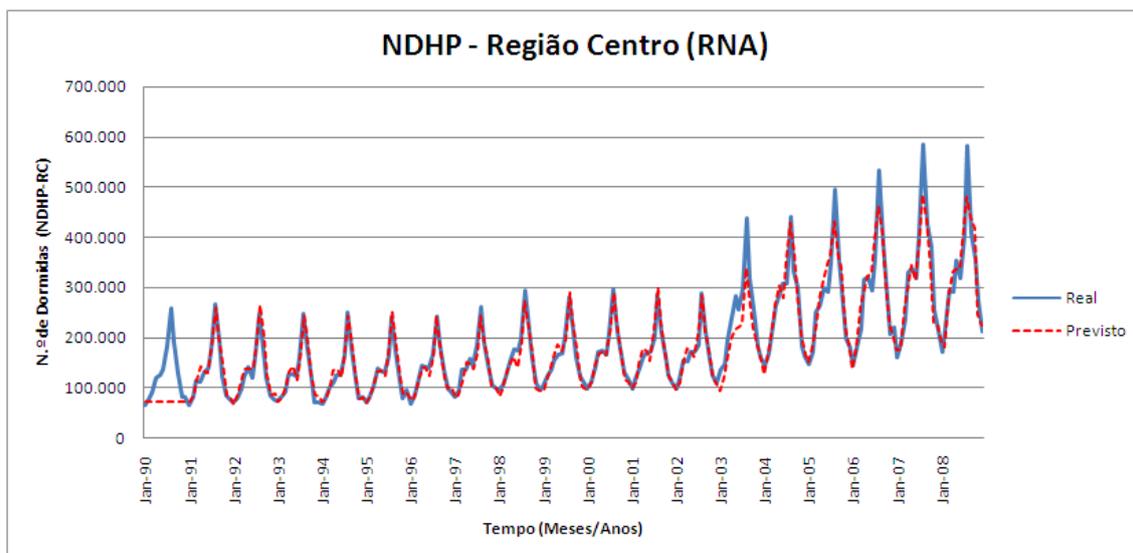


Figura 29 Representação gráfica dos valores previstos da série NDHP-RC (RNA).

5.1.2 Série Turística do Reino Unido

A série turística do Reino Unido (NDH-RU) apresenta um comportamento ligeiramente diferente das séries de procura turística estudadas em secções anteriores. Esta série apresenta uma quebra na sua aparente linearidade, após o ano 2000. Por esta mesma razão, o erro absoluto médio do modelo de regressão linear nos anos de 2007 e 2008 é relativamente elevado, atingindo o valor de 24,77%.

A modelação do mesmo período, aplicando redes neuronais, consegue um erro de 8,96%, cerca de um terço do erro em relação aos métodos lineares (Tabela 6).

Observando os resultados da modelação sobre todo o intervalo de tempo, deste Janeiro de 1990 a Dezembro, verifica-se ainda que os anos de 1993 e 2000 registam um erro médio de 37,79% e 21,52% respectivamente (Tabela 30 e Tabela 31).

Tabela 6 Erro percentual mensal e EPAM no intervalo [Jan-2007:Dez-2008] para ambos os modelos, regressão linear (RL) e redes neuronais artificiais (RNA) relativo à série NDH-RU.

<i>Meses</i>	<i>2007</i>		<i>2008</i>		<i>EPAM</i>	
	<i>RL</i>	<i>RNA</i>	<i>RL</i>	<i>RNA</i>	<i>RL</i>	<i>RNA</i>
<i>Jan</i>	15,07%	4,06%	31,66%	1,51%		
<i>Fev</i>	45,16%	3,28%	26,97%	18,98%		
<i>Mar</i>	22,03%	18,40%	24,26%	6,91%		
<i>Abr</i>	34,88%	15,49%	35,64%	18,59%		
<i>Mai</i>	11,61%	6,39%	9,30%	4,37%		
<i>Jun</i>	11,61%	3,61%	4,92%	0,50%	24,77%	8,96%
<i>Jul</i>	7,72%	7,09%	26,22%	22,13%		
<i>Ago</i>	2,95%	8,52%	23,92%	10,00%		
<i>Set</i>	28,64%	18,08%	28,83%	4,33%		
<i>Out</i>	49,34%	9,31%	49,39%	2,76%		
<i>Nov</i>	28,71%	10,29%	52,15%	9,43%		
<i>Dez</i>	12,62%	5,26%	10,87%	5,71%		

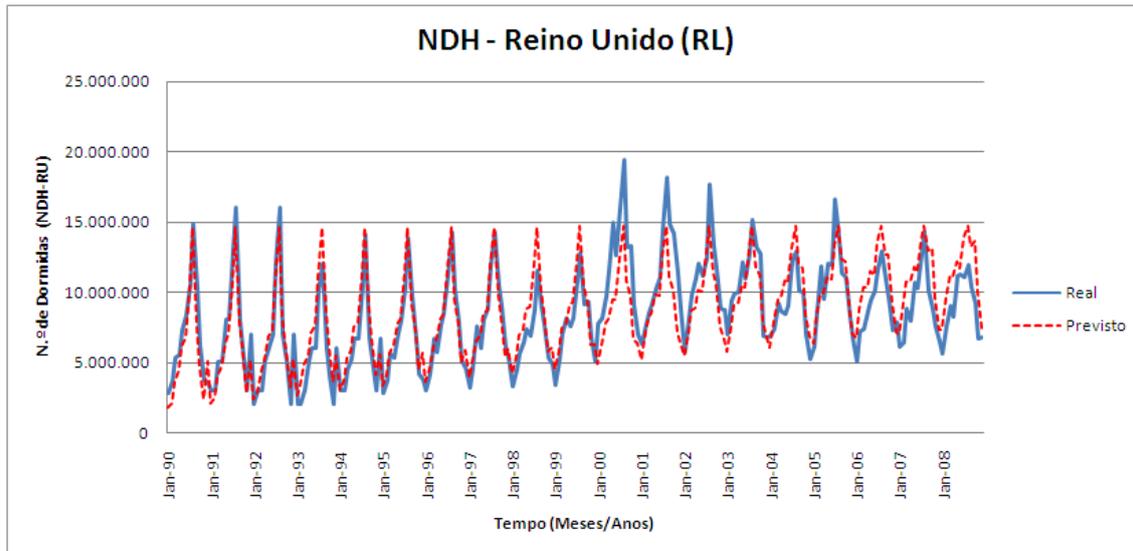


Figura 30 Representação gráfica dos valores previstos da série NDH-RU (RL).

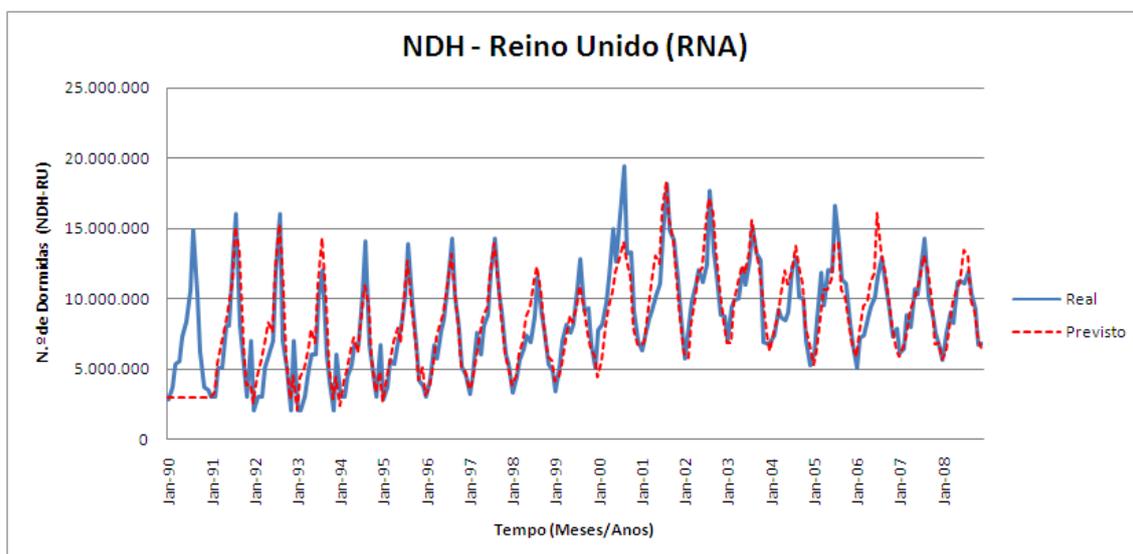


Figura 31 Representação gráfica dos valores previstos da série NDH-RU (RNA).

Os resultados apresentados nas secções até aqui foram obtidos sem recorrer à inclusão de qualquer variável auxiliar. Na próxima secção será testada a influência das variáveis auxiliares.

5.2 Influência das Variáveis Auxiliares

No sentido de perceber a influência das variáveis auxiliares referidas, foram efectuados treinos da rede neuronal nas mesmas condições com a excepção de incluir à entrada da rede neuronal também essas mesmas variáveis.

Desta forma, para cada série, foram efectuados testes à influência de cada uma das variáveis auxiliares da seguinte forma:

- De forma individual para o Índice de Preços no Consumidor em Portugal (IPC), número de desempregados em Portugal (NDPT), Espanha (NDES), França (NDFR) e Reino Unido (NDRU) e Temperatura Máxima Média em Lisboa (TMM);
- Em conjuntos, agrupando à entrada todos os *datasets* relativos ao número de desempregados (ND) e finalmente agrupando todas as variáveis auxiliares (TODAS).

Na Tabela 7 pode observar-se o Erro Percentual Absoluto Médio em cada uma das séries para os anos de 2007 e 2008 (conjunto de teste), com recurso à inclusão das diferentes combinações, relativas às variáveis auxiliares referidas.

Tabela 7 EPAM – Variáveis auxiliares.

	<i>Nenhuma</i>	<i>IPC</i>	<i>NDPT</i>	<i>NDES</i>	<i>NDFR</i>	<i>NDRU</i>	<i>TMM</i>	<i>ND</i>	<i>TODAS</i>
<i>NDHP-RN</i>	6,9%	7,7%	6,6%	9,1%	7,1%	8,2%	7,3%	11,2%	10,0%
<i>NDHP-RC</i>	8,2%	7,1%	7,8%	8,5%	6,8%	8,3%	7,1%	9,3%	12,0%
<i>NDHP</i>	4,8%	5,3%	4,3%	5,0%	5,3%	6,1%	4,6%	10,0%	10,0%

A Figura 32 faz a representação gráfica dos resultados obtidos, permitindo uma visão geral das influências dos conjuntos de dados utilizados.

Uma primeira análise permite concluir que não é expressiva a diferença entre os resultados obtidos com a inclusão das variáveis auxiliares e os resultados obtidos apenas com recurso ao histórico da série. Verifica-se no entanto, que a inclusão das variáveis auxiliares em simultâneo parece degradar a performance da modelação. No caso da série NDHP o erro dispara de 4,8% para 10,0%, quando se inclui à entrada o ND de todos os países ou mesmo todas as variáveis auxiliares.

De forma individual os indicadores que parecem introduzir informação relevante no processo de modelação são o NDPT e o IPC. A inclusão do indicador de desemprego em Portugal introduz uma melhoria na performance em todas as séries apresentadas, conseguindo mesmo um mínimo de 4,3% na série NDHP. O Índice de Preços no Consumidor em Portugal apresenta um bom desempenho quando incluído à entrada, junto com a série NDHP-RC, melhorando a performance da rede em 15%, em relação à melhor performance obtida até agora.

O indicador do estado do tempo, TMM, parece também influenciar positivamente a performance da rede na modelação das séries NDHP e NDHP-RC.

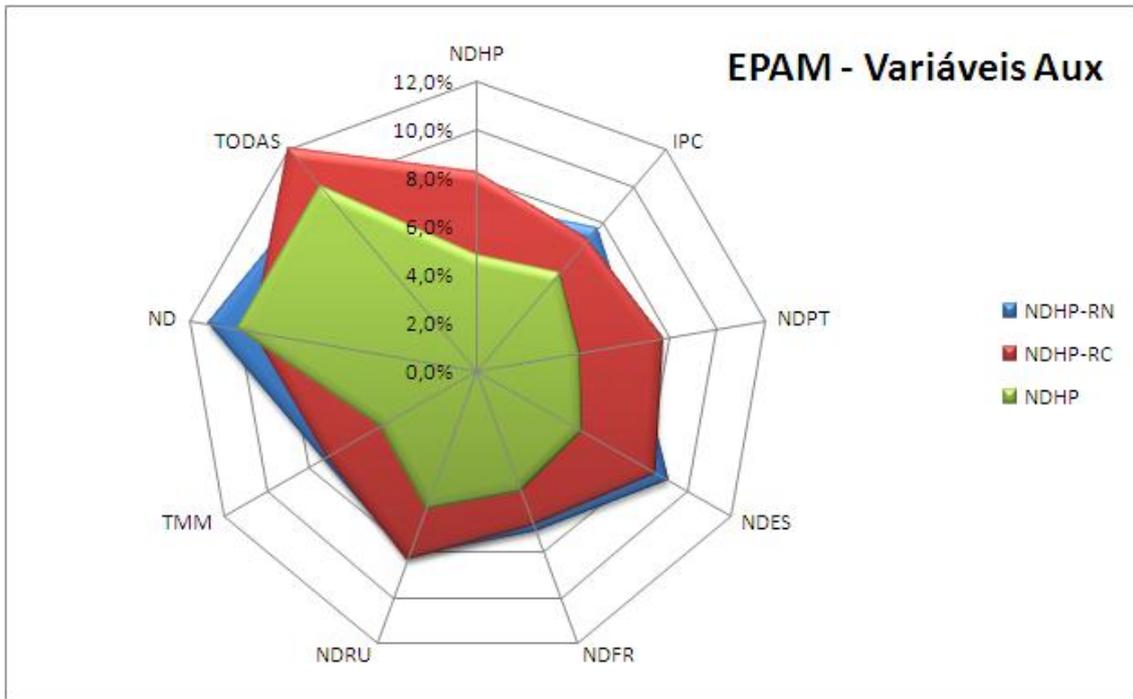


Figura 32 EPAM – Variáveis Auxiliares.

Capítulo 6

6. Interface Gráfica

A criação de uma interface gráfica, reveste-se de especial importância na forma como permite a um utilizador realizar facilmente uma determinada simulação, mesmo não possuindo conhecimentos específicos e detalhados de modelação de séries temporais ou de redes neuronais artificiais. Ao mesmo tempo, permite também a um utilizador mais interessado no funcionamento das RNA's, interagir directamente com os parâmetros que definem a rede neuronal, fazendo-os variar dentro de um determinado intervalo e observando os respectivos resultados, podendo inclusivamente guardar a estrutura com todas as variáveis que definem uma RNA, podendo mesmo reutiliza-la posteriormente.

A aplicação criada é também flexível ao ponto de permitir o estudo de um vasto conjunto de séries, bem como de avaliar o efeito de determinadas variáveis auxiliares, sobre essas mesmas séries.

Este capítulo, faz uma descrição sobre os ecrãs e suas funcionalidades que constituem não só o interface gráfico desenvolvido, como também alguns dos ecrãs que é possível invocar a partir destes. Estes ecrãs pertencem à *toolbox* do Matlab¹⁵ e são uma importante ferramenta para acompanhar o processo de treino da rede neuronal.

6.1 Ferramenta Utilizada

O interface gráfico, tal como todo o motor do sistema de previsão de séries temporais, foi concebido na linguagem de programação designada, Matlab. Esta ferramenta constitui um programa de computador orientado ao cálculo científico baseado em operações matemáticas sobre matrizes. O Matlab é uma ferramenta bastante flexível, na medida em que pode ser utilizado para resolver qualquer problema de natureza técnica.

Dentre as vantagens, desta ferramenta, destacam-se a sua facilidade de uso, as demonstrações incluídas, a independência da plataforma, as biblioteca de funções pré-definidas bem como a possibilidade de criar facilmente interfaces gráficos [Chapman, 2002].

A escolha desta ferramenta assenta no facto de possuir uma biblioteca de funções exclusivamente dedicada às redes neuronais artificiais. Esta *toolbox*, possui um conjunto de funções capazes de lidar com toda a especificidade associada à utilização de redes neuronais artificiais, como a definição da sua arquitectura, controlo dos processos de aprendizagem e generalização.

¹⁵ Que abrevia *MATrix LABoratory* – Laboratório de Matrizes.

O Matlab possui ainda uma ferramenta designada *guide*, que permite o desenvolvimento de GUI's¹⁶.

6.2 Barra de Menus

Na barra de menus (Figura 33), encontram-se os controlos que permitem manipular as variáveis de entrada no sistema. É permitida a selecção e o carregamento da série em estudo, bem como também de variáveis que se julguem importantes para a previsão, nomeadamente indicadores económicos, aqui genericamente designados por indicadores auxiliares. É também possível carregar uma rede neuronal, previamente criada e guardada, de modo que o utilizador possa realizar determinados testes, sem ter a necessidade de executar o moroso processo de treino. Neste contexto, entenda-se por rede neuronal, uma estrutura com a definição de todos os parâmetros de uma rede treinada, para referir apenas alguns, a arquitectura, o algoritmo de treino e as matrizes com os pesos sinápticos. Uma rede treinada numa sessão anterior cuja performance foi relevante, pode ser guardada e utilizada posteriormente sobre um conjunto de dados que se pretenda classificar, basta que este possua uma estrutura semelhante.

Sempre que o utilizador introduza uma variável susceptível de influenciar o processo de previsão da série, é calculada de imediato a correlação que existe entre esta variável e a série, permitindo ao utilizador ter uma ideia da vantagem que terá em carregar determinada variável.

A partir do menu 'Ficheiro' é também possível guardar os resultados de uma sessão. Entenda-se por resultados, a matriz de resultados da previsão bem como a rede neuronal que lhe deu origem.

A barra de menus permite também o acesso aos ecrãs de ajuda ao utilizador da aplicação.

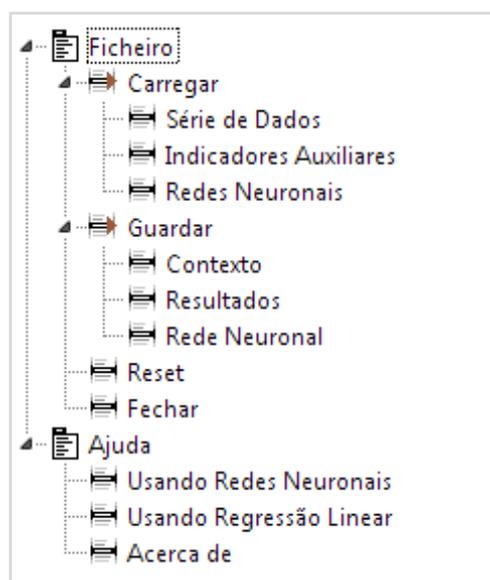


Figura 33 Barra de Menus.

¹⁶ Do inglês *Graphical User Interface*.

6.3 Ecrã – Sistema de Previsão de Séries Temporais

O interface principal apresentado na Figura 34, serve de base a todas as operações a que esta aplicação se destina e o seu *layout* está organizado em três blocos. O bloco de definição das variáveis, bloco de processamento dos dados e bloco de apresentação de resultados.

A definição de variáveis compreende, a selecção/carregamento do *dataset* e possíveis indicadores auxiliares, a selecção do método que se pretende utilizar, tipo de previsão e data de projecção, no caso de se tratar de uma previsão do tipo ‘Real’, ou simplesmente de uma simulação sobre um dos métodos.

O método de previsão varia entre o recurso à tecnologia de redes neuronais e a regressão linear. A previsão pode ser do tipo ‘Simulação’, se for relativa a dados que apesar de pertencerem ao *dataset*, não são incluídos nos processos de treino/cálculo, consoante se trate, respectivamente, do método de redes neuronais ou regressão linear. A previsão é designada ‘Real’ se for relativa a dados posteriores, não presentes no *dataset*, com limite máximo definido em ‘Projecção’.

O bloco de definição inclui também, o carregamento em memória dos parâmetros infra referidos (Figura 38).

O bloco de processamento de dados difere consoante o método de previsão. No caso do método de redes neuronais, o processamento é efectuado em duas fases, o treino da rede e a simulação/projecção dos dados a que a rede não teve acesso durante o processo de treino.

O processo de treino pode ser repetido tantas vezes quantas as definidas em ‘Processos de Treino’ do ecrã ‘Parâmetros’, sendo apresentado o seu resultado em *run-time*, no ecrã *Neural Network Training* (Figura 36). Em cada processo de treino é calculado o erro em relação à saída desejada, sendo este registado na lista de contexto.

O processamento do método de regressão linear, é relativamente mais simples, não possui nenhum processo de treino, sendo necessário apenas calcular a regressão.

O bloco de apresentação de resultados é composto por um gráfico que apresenta o resultado da simulação/projecção da série, apresentado os valores reais e os previstos para os intervalos de tempo definidos. No caso de se tratar de uma projecção, não existem valores reais, uma vez que se está a prever para o futuro, não sendo portanto neste caso, apresentado o valor do erro (EPAM), na lista de contexto.

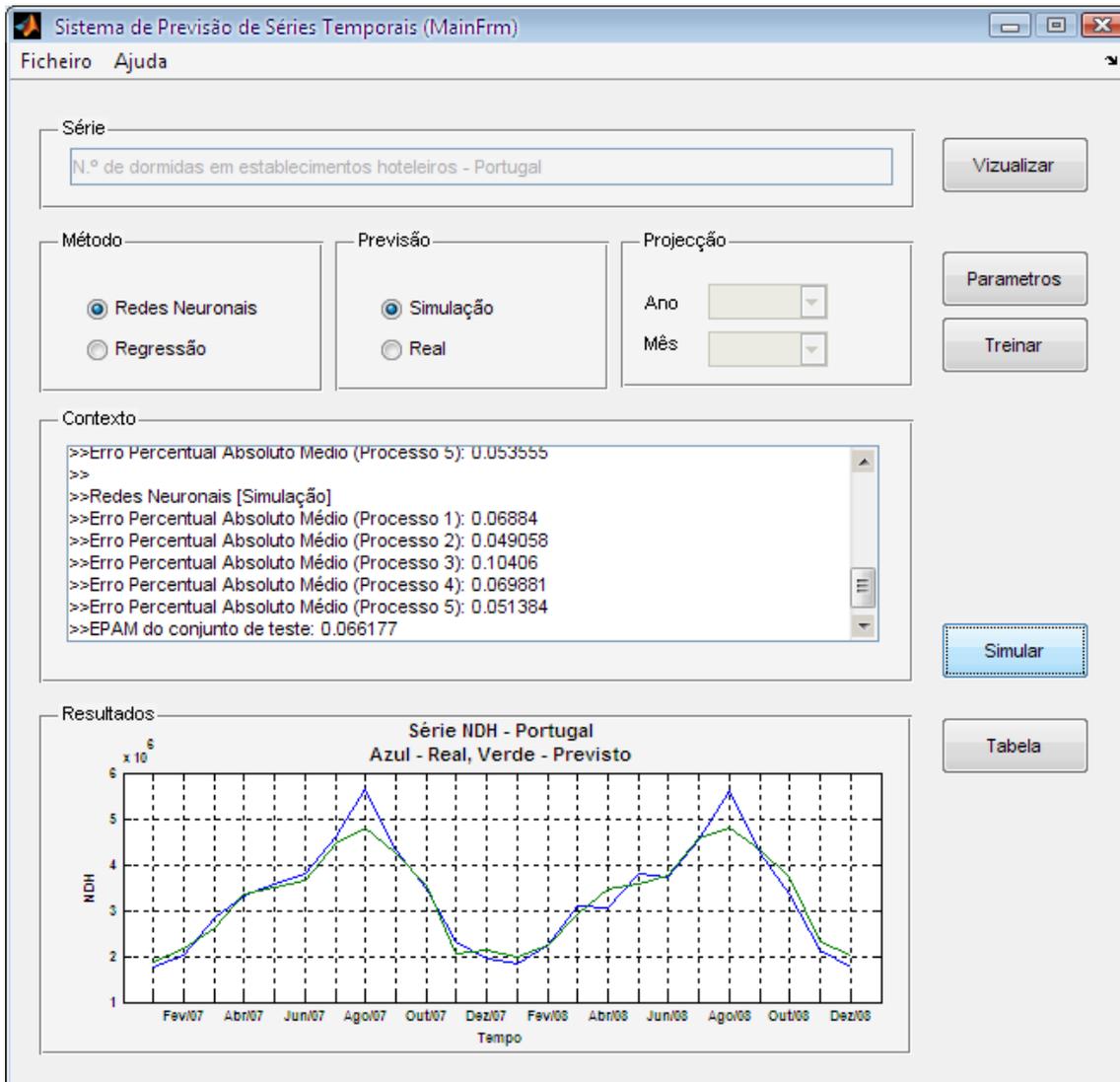


Figura 34 Sistema de Previsão de Séries Temporais.

6.3.1 Funcionalidades

Quando a aplicação é iniciada, todos os controlos estão desabilitados até que o utilizador carregue a série de dados. É possível também carregar séries (uma ou mais) de variáveis auxiliares que o utilizador considere relevantes. Após a selecção de uma variável auxiliar, é calculada a correlação linear¹⁷ entre esta e a série de dados a modelar, sendo possível confirmar a sua inclusão ou simplesmente descartá-la.

Antes de iniciar o processo de modelação, o utilizador pode ainda carregar uma rede neuronal já treinada, podendo passar directamente ao teste da rede sobre o conjunto de dados definido.

Convém referir que, em termos de *workflow* de processamento, podemos dividir esta aplicação em quatro tipos de processamento básicos.

O primeiro, aquele que está definido por omissão, utiliza o método de redes neuronais para simular o conjunto de teste inserido no ecrã 'Parâmetros', não sendo necessário

¹⁷ O resultado da correlação varia entre -1 (oposição exacta) e 1 (total concordância).

definir uma projecção, uma vez que o limite máximo da previsão é sempre definido pelos limites do *dataset*.

Seleccionando o mesmo método é possível optar por uma previsão do tipo ‘Real’, permitindo projectar para o futuro, a previsão da série. Neste caso o utilizador terá de seleccionar o ano e o mês de projecção, limitado a dois anos, a partir do último mês da série, para evitar a degradação de performance.

Existe uma variante para ambos os tipos de execução, sempre que se carregue inicialmente uma rede neuronal previamente testada e guardada. Neste caso é apenas necessário definir os parâmetros relativos ao conjunto de dados. Uma vez que a rede já existe, todos os parâmetros gerais relativos à definição da rede neuronal são desabilitados. O processo de treino também é suprimido, pois, subentende-se que se guardou uma rede neuronal convenientemente treinada.

Optando pelo método de regressão linear, não será necessário definir os parâmetros (Figura 38), uma vez que se trata do cálculo da regressão em ordem ao mês ao longo tempo. No entanto, é possível simular a previsão para os últimos dois anos ou projectá-la até dois anos a frente, à semelhança da projecção com redes neuronais.

No caso de se utilizar o método de regressão linear, não é possível seleccionar o mês de projecção, uma vez que o cálculo é feito ao ano, unidade mínima da variável independente.

6.4 Ecrã - Visualizar Série de Dados

Neste ecrã é possível visualizar a série de dados, de um determinado país em forma matricial e gráfica, para cada mês ao longo do intervalo de tempo, neste caso de Janeiro de 1990 a Dezembro de 2008. O controlo que permite abrir este ecrã apenas está habilitado quando um *dataset* é carregado em memória.

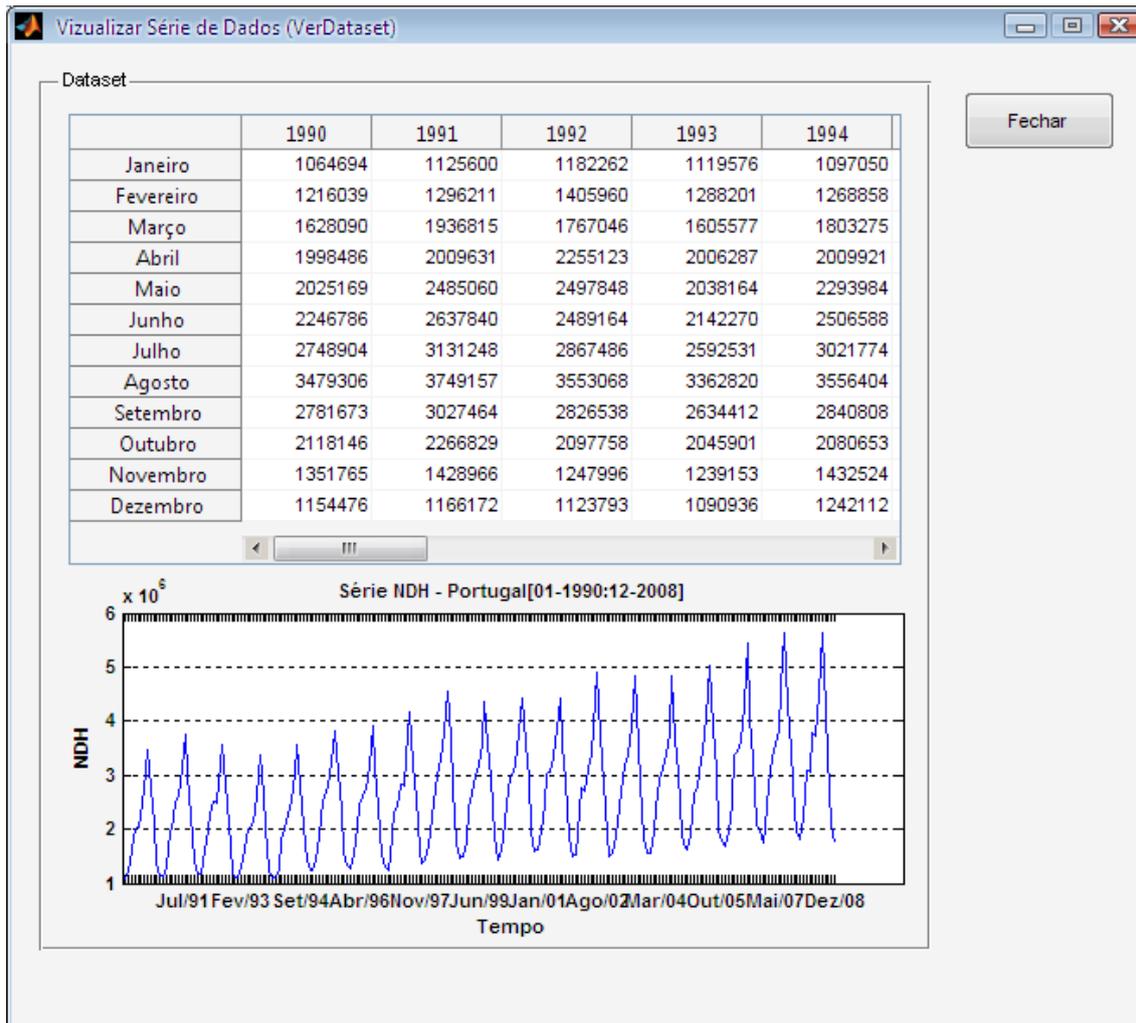


Figura 35 Ecrã – Visualizar Séries de Dados.

6.5 Ecrã – Neural Network Training

Este ecrã permite visualizar todas as iterações do processo de treino para uma determinada configuração, bem como a representação gráfica das curvas de performance, estado do treino e regressão. Normalmente tem este aspecto, podendo variar em função do algoritmo de treino utilizado, dado que o progresso de cada algoritmo opera sobre um conjunto diferente de variáveis. Um exemplo disso é o parâmetro de adaptação μ , característico do algoritmo de treino Levenberg-Marquardt.

Neste ecrã pode observar-se a arquitectura da rede neuronal, as funções de treino e performance, o progresso das principais variáveis que definem o processo de treino e os respectivos gráficos.

No painel 'Progresso', a primeira barra representa do número de épocas ou iterações do processo de treino. Cada iteração de treino representa um determinado ajuste dos pesos sinápticos, no sentido de minimizar o erro na classificação do conjunto de treino. Se nenhum outro critério de paragem for verificado, o processo de treino vai decorrer ao longo das 200 iterações que estão definidas por omissão.

De seguida, pode observar-se o tempo que cada processo de treino consome. Este valor é tanto menor quanto maior a velocidade de convergência do algoritmo, representando o tempo decorrido até que o primeiro critério de paragem se verifique.

A barra de performance apresenta a evolução do erro no conjunto de treino, em relação à saída desejada, no sentido do valor do parâmetro definido em 'Critério de Erro' do ecrã 'Parâmetros' (0,002).

Já foi referida a importância de existir um conjunto de validação para interromper o processo de treino mantendo o poder de generalização. Por omissão o parâmetro da rede neuronal que define o número máximo de iterações consecutivas sem incremento na função de performance sobre o conjunto de validação. Deste modo, sempre que o erro no conjunto de validação não diminua, ao longo de um número de iterações sucessivas, manifestando uma perda do poder de generalização sobre este conjunto, o processo de ajuste dos pesos sinápticos é interrompido.

Este ecrã permite ainda representar graficamente a performance, o estado do treino, bem como as rectas de regressão para os conjuntos de treino, teste, validação e global (todos os conjuntos).

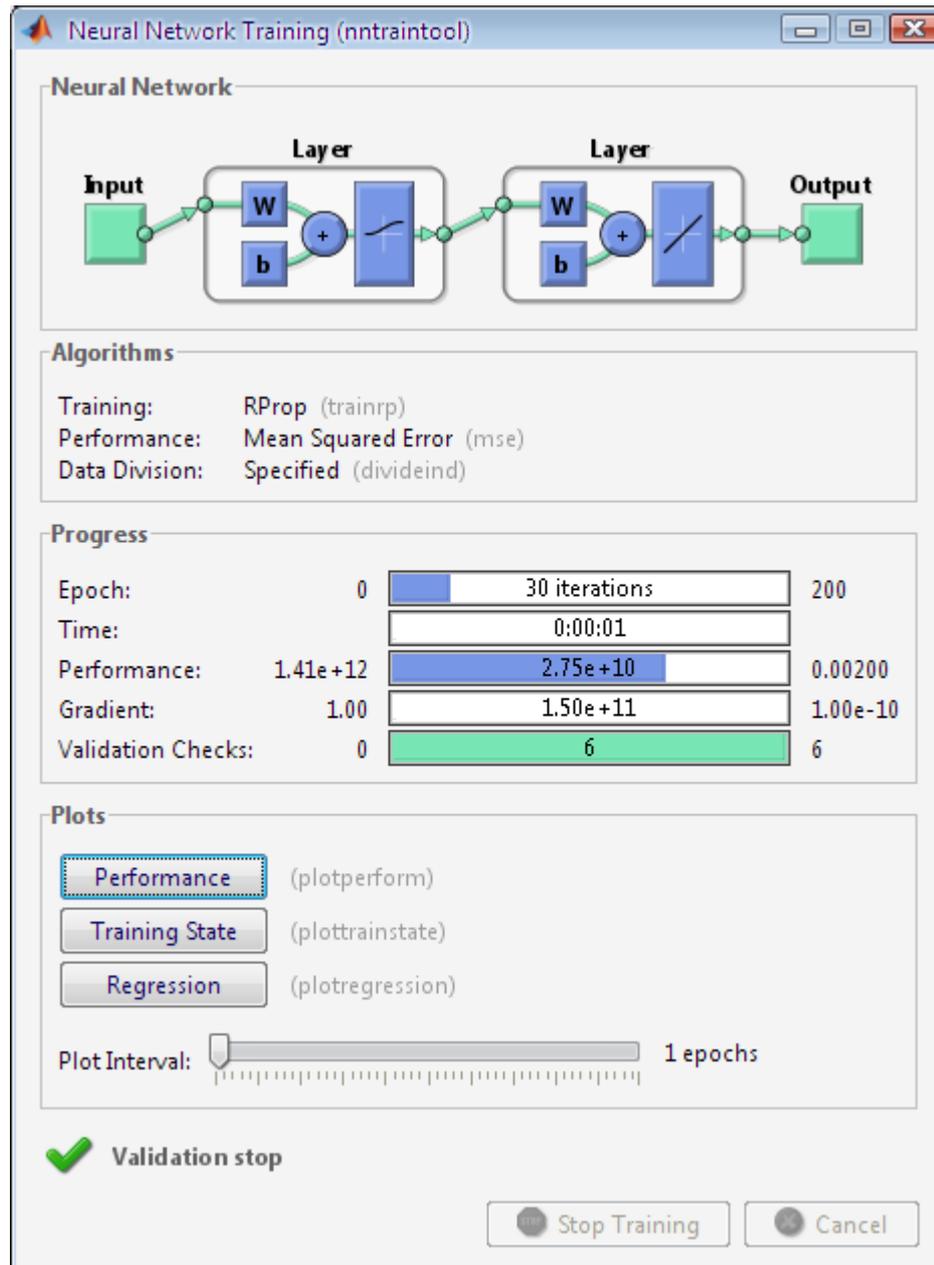


Figura 36 Ecrã – *Neural Network Training*.

6.6 Ecrã – *Regression*

Este ecrã faz a representação gráfica da regressão linear dos valores reais (T -*target*) face aos previstos (Y -*output*), relativos aos conjuntos de treino, validação, teste e ao conjunto global dos dados. É também calculado o coeficiente de correlação (R) que existe entre os valores previstos Y reais T , para os diferentes conjuntos, permitindo assim ter uma qualificação da modelação. Obviamente, que quanto mais alinhadas estão as amostras com a recta ($Y = T$), melhor performance se obteve.

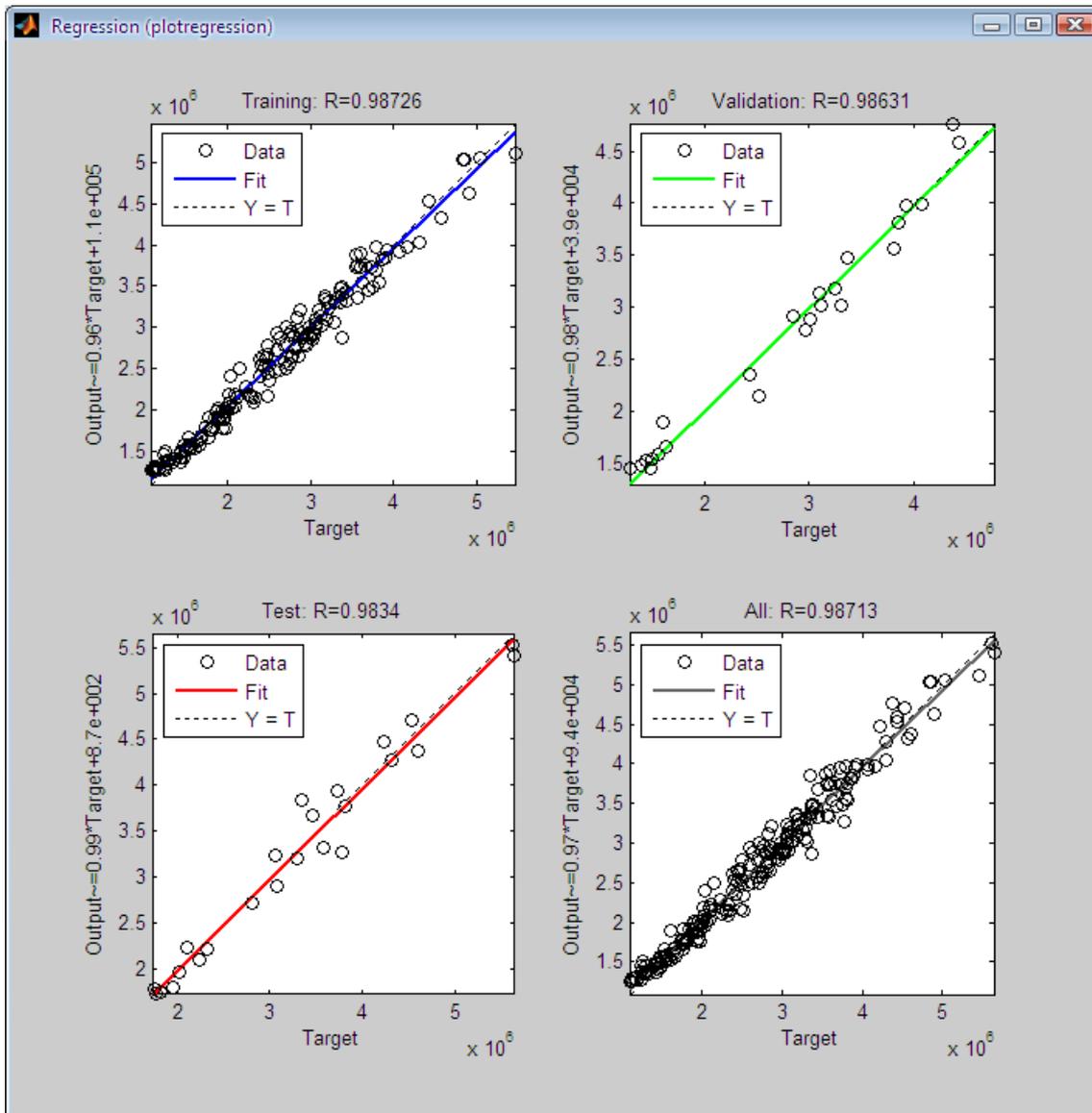


Figura 37 Ecrã – Regression.

6.7 Ecrã – Definição de Parâmetros

Neste ecrã são definidas as variáveis que serão utilizadas na organização do conjunto de dados, rede neuronal e processo de treino.

O campo ‘Janela Temporal’ define o número de amostras N , que no instante n , são utilizadas para prever a amostra relativa a, $n+1$. Esta variável pode também ser vista como a memória sobre a qual a rede neuronal opera ao longo do seu processo de aprendizagem.

O grupo definido em ‘Conjunto de Validação’, permite seleccionar a forma como será organizado o conjunto utilizado para validação no processo de treino. A opção sequencial compreende o número de amostras definidas em percentagem, pelo campo ‘Conjunto de Validação [%]’, extraídas do *dataset* de forma sequencial, imediatamente a seguir ao respectivo conjunto de treino e antes do conjunto de teste. O conjunto de

validação aleatório corresponde à mesma percentagem, extraída de forma aleatória, do conjunto formado pela soma das amostras dos conjuntos de treino e validação.

Nos parâmetros que definem a rede neuronal, o campo ‘N.º de neurónios’, indica o número de nós na camada escondida. O grupo designado funções de activação permite definir a função de transferência respectivamente para a camada escondida e camada de saída. O campo ‘Algoritmo’ define o algoritmo de treino. No campo, ‘Taxa de Aprendizagem’, é definida uma constante, que pode variar no intervalo [0,1], representando o valor da variação dos ajustes dos pesos sinápticos (este parâmetro só é usado nos algoritmos em que a taxa de aprendizagem não varia). No campo, ‘Critério de Erro [MSE]’, é definido o valor desejado para a função de performance *Mean Square Error*. O processo de treino é interrompido quando o MSE atinge o valor indicado.

O campo, ‘Processos de Treino’, permite definir o n.º de sessões que se pretende realizar, mantendo todos os parâmetros, excepto os valores referentes à inicialização inicial dos pesos sinápticos.

Foram também definidos controlos para o carregamento de valores pré-definidos, bem como para guardar determinadas configurações.

Figura 38 Ecrã – Definição de Parâmetros.

6.7.1 Validações

Os campos do ecrã são validados aquando do preenchimento, impedindo o utilizador de inserir caracteres num campo que deve receber um valor numérico.

Os valores inseridos também devem seguir determinadas regras. Não podem ser inseridos valores negativos, os valores expressos em percentagem devem ser respeitados e existem limites máximos para os restantes campos.

A soma das percentagens relativas aos conjuntos de treino validação e teste devem ser igual a 1, abrangendo a totalidade das amostras do *dataset*.

A confirmação dos campos será efectuada com sucesso, apenas quando todos os campos estiverem convenientemente preenchidos.

Todas as excepções são assinaladas com uma mensagem de erro.

6.7.2 Gestão das Matrizes de Parâmetros

São guardadas permanentemente duas matrizes de dados, designadas por, matriz de dados pré-definidos e a matriz de estado. Estas matrizes incluem respectivamente o grupo de variáveis que define o conjunto de dados ou *dataset*, e o grupo de variáveis que define a rede neuronal utilizada.

Entenda-se por dados pré-definidos o conjunto formado pelos parâmetros definidos previamente e guardados antes de inicio de sessão na aplicação. Os dados pré-definidos são carregados por grupo, permitindo ao utilizador, e.g., carregar apenas as variáveis do painel 'Rede Neuronal'. A matriz de estado armazena as variáveis que o utilizador define numa determinada sessão, podendo estas ser carregadas numa sessão posterior.

Numa determinada sessão, ficarão carregados em memória os valores das variáveis que o utilizador confirmar com 'Ok', no entanto, este conjunto não persistirá em outras sessões, a menos que seja guardado.

Este mecanismo foi criado, não só para auxiliar a introdução dos parâmetros pelo utilizador, mas também para permitir guardar configurações com as quais se obtiveram resultados satisfatórios.

Capítulo 7

7. Conclusões e Perspectivas Futuras

7.1 Conclusões

O estudo realizado em torno da influência das variáveis auxiliares sobre as séries turísticas tem a sua limitação na dificuldade em definir e mesurar a relação causa efeito do problema da procura turística. Além disso um estudo desta natureza, em termos computacionais, torna-se muito dispendioso.

Perante as condições e os pressupostos em que se baseia o estudo do efeito das variáveis auxiliares, pode concluir-se que alguns indicadores macro económicos, como o índice de preços no consumidor ou o número de desempregados, são susceptíveis de causar alguma influência nas séries de procura turística aqui analisadas.

Em relação ao modelo de regressão linear, conclui-se que é uma boa opção para modelar a série NDHP. Computacionalmente pouco exigente, demonstra uma qualidade de ajustamento bastante satisfatória.

Em séries cujo comportamento é não linear, ou parcialmente linear, como é o caso da série NDHP-RC, as redes neuronais superam o método linear aqui aplicado, demonstrando ser uma opção vantajosa. A sua capacidade para extrair determinados padrões e modelar comportamentos aparentemente estocásticos, motivam a sua aplicação.

A escolha dos parâmetros que definem a rede neuronal, como a sua arquitectura e algoritmo de treino, também são determinantes na obtenção de resultados satisfatórios. A definição do processo de treino e a forma como é gerido, permite também tirar o máximo partido dos dados de que se dispõem para a modelação de séries.

O estudo sobre os parâmetros que definem não só a estrutura mas também o funcionamento de uma rede neuronal artificial foi conclusivo especialmente em relação ao algoritmo de treino e à estrutura da rede - determinação do número de nós na camada escondida da rede. O algoritmo de treino *Resilient Back-Propagation* bem como o algoritmo de *Levenberg-Marquardt*, mostraram uma performance acima da média em relação aos demais algoritmos de treino disponibilizados. Também se conclui que o número de nós na camada escondida é bastante importante para a rede modelar comportamentos complexos. No entanto, se a rede neuronal for capacitada de uma estrutura tal, que permita modelar com demasiado detalhe o conjunto de treino, e não possuir critérios de paragem convenientes, a performance da rede sobre o conjunto de

teste será seriamente comprometida. Em relação à arquitectura da rede, a inclusão de 6 neurónios na camada escondida foi suficiente, não se evidenciando a necessidade de aumentar a complexidade da rede.

De uma forma geral, pode concluir-se que ambas as metodologias aqui desenvolvidas são opções válidas na modelação de séries temporais, desde que sejam aplicadas a um objecto de estudo adequado. Para séries cujo comportamento aparente ser linear, o modelo univariado apresentado será suficiente para conseguir resultados satisfatórios. Em séries mais complexas, para a modelação das quais se pretenda incluir determinadas influências externas, as redes neuronais são uma opção bastante flexível, robusta e eficiente.

7.2 Perspectivas Futuras

Esta dissertação está inserida numa área de investigação constante, sendo portanto natural a procura progressiva por soluções não só eficazes, como cada vez mais eficientes. A tarefa da modelação/previsão, é sempre complexa uma vez que lhe está subjacente o que designamos vulgarmente por incerteza.

No campo da modelação linear, perante séries de procura turística, poder-se-á pensar em aplicar a regressão linear multivariada. As séries de natureza turística apresentam, não raras as vezes, um comportamento não linear composto por segmentos lineares. Desta forma, este método é bastante simples e pode produzir resultados satisfatórios.

Relativamente às redes neuronais artificiais, poder-se-á apostar mais no pré-processamento dos dados, convertendo a entrada para o domínio logarítmico e/ou usando as diferenças. A utilização de algoritmos genéticos para otimizar os parâmetros que definem a rede neuronal. Esta metodologia inspira-se na evolução das espécies para eleger a aptidão natural dos concorrentes, que neste caso seriam os vectores de valores para os vários parâmetros. Na definição de uma rede neuronal existe um grande número de parâmetros que se pode fazer variar, em busca da melhor rede. Os algoritmos genéticos poderão dar um importante contributo em busca da melhor solução para este problema.

Já no domínio do problema a modelar, poder-se-á, em trabalhos futuros, efectuar um estudo mais abrangente acerca das influências externas de uma determinada série. No caso da procura turística, poder-se-á estudar na influência de mega-eventos, e.g. campeonato europeu de futebol, conferências internacionais e eventos culturais. A organização e o pré-processamento dos dados à entrada da rede neuronal, neste caso, serão preponderantes.

Finalmente, uma perspectiva bastante útil seria o desenvolvimento de um mecanismo que permita actualizar os *datasets*, em tempo real, a partir de um interface de dados e metadados estatísticos.

Bibliografia

- Bloom, J. Z. (2003, July 27). Tourist market segmentation with linear and non-linear techniques. *Tourism Management*.
- Butler, R. W. (2001). *Seasonality in Tourism: Issues and Implications*. Oxford: Pergamon.
- Chapman, S. J. (2002). *Programação em Matlab para Engenheiros*. Austrália: Thomson.
- Cortez, P., & Neves, J. (2000). *Redes Neurais Artificiais*. Braga - Universidade do Minho.
- Costa, E., & Simões, A. (2004). *Inteligência Artificial*. FCA - Editora Informática, Lda.
- Demuth, H., & Mark, B. (2002). *Neural Network Toolbox User's Guide*. The MathWorks, Inc.
- ECA. (2009). Retrieved 11 2009, from EUROPEAN CLIMATE ASSESSMENT & DATASET (ECA&D): <http://eca.knmi.nl>.
- Eurostat. (2009). Retrieved Novembro 2009, from Statistical Office of the European Communities: <http://epp.eurostat.ec.europa.eu>
- Faulkner, B. and Valerio, P. (2000). "An integrative approach to tourism demand forecasting". In C. Ryan and S. Page (Eds), "Tourism Management: Toward the new millennium". Oxford, UK: Pergamon, Elsevier Science, Ltd.
- Fernandes, P. O. (2005). *Modelling, Prediction and Behaviour Analysis of Tourism Demand in the North of Portugal*. Valladolid University - Spain.
- Fernandes, P. O., & Teixeira, J. P. (2008). Modelação da Procura Turística: Um estudo comparativo entre Redes Neurais Artificiais e a Metodologia de Box-Jenkins. *XII Congresso Internacional de Investigación en Ciências Administrativas (ACACIA)*. Universidad de Baja Califórnia.
- Fernandes, Paula O.; Teixeira, João P.; Ferreira, João M.. and Azevedo, Susana G.; (2008a); "Modelling Tourism Demand: A Comparative Study between Artificial Neural Networks and the Box-Jenkins Methodology"; *Romanian Journal of Economic Forecasting*, No 5(3), pp. 30/50.
- Fuster, L. F. (1989). *Introducción a la Teoría y Técnica del Turismo*. Madrid: Alianza.
- Gujarati, D. N. (2003). *Basic Econometrics* (4 ed.). McGraw-Hill.
- Goeldner, C., Ritchie, B. and McIntosh, R. (1999). "Tourism: Principles, practices, philosophies". 8th. New York, NY: John Wiley & Sons.
- Hawkins, D. and Ritchie, B. (1992). "World travel and tourism review: indicators, trends and forecasts". Wallingford, Oxon, UK: C.A.B. International.
- Haykin, S. (1999). *Neural Networks: A Comprehensive foundation*. Prentice Hall, Inc.
- Hu, C. (2002). "Advanced tourism demand forecasting: artificial Neural Network and Box-Jenkins Modeling", Ph.D Thesis, Purdue University.
- INE. (n.d.). Retrieved Novembro 2009, from Instituto Nacional de Estatística: <http://www.ine.pt>
- Johnston, J., & Dinardo, J. (2000). *Métodos Económicos* (4 ed.). McGraw-Hill.
- Law, R. (1999, February 4). Back-propagation learning in improving the accuracy of

- neural network-based tourism demand forecasting. *Tourism Management* .
- Lewis, C. D. (1982). *Industrial and Business Forecasting Method*. London: Butterworth Scientific.
- Machado, T. N., Teixeira, J. P., & Fernandes, P. O. (2009). Modelação da Procura Turística em Portugal: Regressão Linear versus Redes Neurais Artificiais. *INVTUR 2010 - Investigação em Turismo: Estado da Arte e Perspectivas de Futuro*. (aceite para publicação)
- Machado, T. N., Teixeira, J. P., & Fernandes, P. O. (2009). Modelação Linear da Procura Turística em Portugal. *XXXIII Congresso Internacional de Economia Aplicada, ASEPELT'2009*.
- Matlab ST. (2009). *Statistics Toolbox User's Guide* . TheMathWorks, Inc.
- McCulloch, W. S., & Pitts, W. (1943). *A Logical Calculus of the Ideas Immanent in Nervous Activity*. Bulletin of Mathematical Biophysics.
- Riedmiller, M., & Braun, H. (1993). A direct adaptive method for faster backpropagation learning: The RPROP algorithm. *Proceedings of the IEEE International Conference on Neural Networks*.
- Rumelhard, D. E., & McClelland, J. L. (1986). *Parallel Distributed Processing – Explorations in the Microstructure of Cognition* (Vol. 1). The Massachusetts Institute of Technology Press.
- PENT - Plano Estratégico Nacional do Turismo*. 2006 Lisboa.
- Turismo de Portugal*. (2009a). Retrieved Abril 13, 2009, from Conta Satélite do Turismo - Turismo: Actividade Económica 2000 2007: <http://www.turismodeportugal.pt/>.
- Turismo de Portugal*. (2009b). Índice de Competitividade Viagens e Turismo 2009: <http://www.turismodeportugal.pt/>
- Wikipédia*. (n.d.). Retrieved 11 26, 2009, from http://pt.wikipedia.org/wiki/Santiago_Ram%C3%B3n_y_Cajal
- Witt, Stephen F. and Witt, Christine A.; (1995); “Forecasting tourism demand: a review of empirical research”; *International Journal of Forecasting*; No 11, pp.447/475.
- Yu, Gongmei and Schwartz, Zvi; (2006); “Forecasting Short Time-Series Tourism Demand with Artificial Intelligence Models”; *Journal of Travel Research*; No 45, pp. 194/203.

Apêndice A

Tabela 8 Dormidas registadas em Portugal (NDHP) no período [Jan-1990:Dez-2008]

	<i>Jan</i>	<i>Fev</i>	<i>Mar</i>	<i>Abr</i>	<i>Mai</i>	<i>Jun</i>
1990	1064694,00	1216039,00	1628090,00	1998486,00	2025169,00	2246786,00
1991	1125600,00	1296211,00	1936815,00	2009631,00	2485060,00	2637840,00
1992	1182262,00	1405960,00	1767046,00	2255123,00	2497848,00	2489164,00
1993	1119576,00	1288201,00	1605577,00	2006287,00	2038164,00	2142270,00
1994	1097050,00	1268858,00	1803275,00	2009921,00	2293984,00	2506588,00
1995	1227438,00	1428667,00	1878477,00	2522411,00	2705234,00	2812203,00
1996	1265604,00	1547819,00	2027335,00	2436365,00	2596425,00	2747914,00
1997	1239368,00	1478048,00	2314002,00	2402960,00	2844861,00	2799762,00
1998	1427813,00	1669745,00	2226679,00	2744398,00	2989225,00	3309335,00
1999	1491872,00	1781463,00	2380190,00	2686488,00	2984547,00	3167089,00
2000	1484802,00	1805077,00	2481530,00	2966987,00	3064380,00	3201352,00
2001	1608556,00	1964165,00	2457972,00	3011554,00	3064136,00	3349125,00
2002	1537139,00	1926842,00	2752279,00	2703439,00	3108798,00	3371407,00
2003	1550720,00	1869726,00	2429128,00	2925012,00	3195173,00	3365166,00
2004	1549992,00	1934294,00	2382340,00	2918348,00	3105531,00	3383239,00
2005	1622322,00	1878765,00	2686969,00	2764258,00	3287038,00	3432908,00
2006	1676355,00	1947106,00	2484546,00	3367859,00	3440228,00	3616770,00
2007	1763320,00	2024858,00	2816553,00	3299232,00	3575744,00	3806145,00
2008	1822603,00	2239183,00	3078757,00	3063921,00	3782832,00	3721954,00
	<i>Jul</i>	<i>Ago</i>	<i>Set</i>	<i>Out</i>	<i>Nov</i>	<i>Dez</i>
1990	2748904,00	3479306,00	2781673,00	2118146,00	1351765,00	1154476,00
1991	3131248,00	3749157,00	3027464,00	2266829,00	1428966,00	1166172,00
1992	2867486,00	3553068,00	2826538,00	2097758,00	1247996,00	1123793,00
1993	2592531,00	3362820,00	2634412,00	2045901,00	1239153,00	1090936,00
1994	3021774,00	3556404,00	2840808,00	2080653,00	1432524,00	1242112,00
1995	3246540,00	3821775,00	3057635,00	2339465,00	1554192,00	1342805,00
1996	3173069,00	3898713,00	3102936,00	2426589,00	1533598,00	1306920,00
1997	3347065,00	4163378,00	3199612,00	2568769,00	1636918,00	1355540,00
1998	3800426,00	4562245,00	3686793,00	2778732,00	1733056,00	1476052,00
1999	3608203,00	4369943,00	3554816,00	2951672,00	1886679,00	1442033,00
2000	3789338,00	4424181,00	3599561,00	2970317,00	1977758,00	1575413,00
2001	3866128,00	4426781,00	3559937,00	2887563,00	1886345,00	1480329,00
2002	3909345,00	4900264,00	3731515,00	2853204,00	1909950,00	1504786,00
2003	3920883,00	4833007,00	3551038,00	2843491,00	1843415,00	1548712,00
2004	3788629,00	4850672,00	3654583,00	3041458,00	1871106,00	1660389,00
2005	4066733,00	5034593,00	3850305,00	3122789,00	1985737,00	1765439,00
2006	4305916,00	5456461,00	4072262,00	3287801,00	2111710,00	1969363,00
2007	4599228,00	5644162,00	4307242,00	3454104,00	2330737,00	1957329,00
2008	4530198,00	5619010,00	4224292,00	3339442,00	2112816,00	1754558,00

Tabela 9 Valores previstos para a Série NDHP, no período [Jan-1990:Dez-2008] (RL)

	<i>Jan</i>	<i>Fev</i>	<i>Mar</i>	<i>Abr</i>	<i>Mai</i>	<i>Jun</i>
1990	1048753,67	1216873,61	1672655,41	1976312,94	2184321,25	2301730,02
1991	1088770,95	1268492,25	1737413,56	2050814,46	2262206,11	2385917,58
1992	1128788,24	1320110,88	1802171,71	2125315,99	2340090,96	2470105,13
1993	1168805,52	1371729,52	1866929,85	2199817,51	2417975,81	2554292,69
1994	1208822,80	1423348,16	1931688,00	2274319,03	2495860,66	2638480,25
1995	1248840,09	1474966,79	1996446,15	2348820,55	2573745,51	2722667,80
1996	1288857,37	1526585,43	2061204,29	2423322,07	2651630,36	2806855,36
1997	1328874,66	1578204,07	2125962,44	2497823,60	2729515,21	2891042,91
1998	1368891,94	1629822,71	2190720,59	2572325,12	2807400,06	2975230,47
1999	1408909,23	1681441,34	2255478,74	2646826,64	2885284,91	3059418,03
2000	1448926,51	1733059,98	2320236,88	2721328,16	2963169,76	3143605,58
2001	1488943,79	1784678,62	2384995,03	2795829,68	3041054,61	3227793,14
2002	1528961,08	1836297,25	2449753,18	2870331,21	3118939,46	3311980,70
2003	1568978,36	1887915,89	2514511,32	2944832,73	3196824,31	3396168,25
2004	1608995,65	1939534,53	2579269,47	3019334,25	3274709,16	3480355,81
2005	1649012,93	1991153,17	2644027,62	3093835,77	3352594,01	3564543,37
2006	1689030,22	2042771,80	2708785,76	3168337,29	3430478,86	3648730,92
2007	1729047,50	2094390,44	2773543,91	3242838,82	3508363,71	3732918,48
2008	1769064,78	2146009,08	2838302,06	3317340,34	3586248,56	3817106,03
	<i>Jul</i>	<i>Ago</i>	<i>Set</i>	<i>Out</i>	<i>Nov</i>	<i>Dez</i>
1990	2734839,63	3312323,41	2708656,41	2016459,69	1276694,76	1064836,12
1991	2828162,75	3430950,40	2787220,60	2092940,00	1327629,01	1109711,53
1992	2921485,87	3549577,38	2865784,78	2169420,31	1378563,25	1154586,94
1993	3014808,99	3668204,37	2944348,96	2245900,62	1429497,49	1199462,35
1994	3108132,11	3786831,35	3022913,15	2322380,93	1480431,74	1244337,76
1995	3201455,23	3905458,34	3101477,33	2398861,24	1531365,98	1289213,18
1996	3294778,35	4024085,32	3180041,51	2475341,55	1582300,22	1334088,59
1997	3388101,47	4142712,31	3258605,70	2551821,87	1633234,46	1378964,00
1998	3481424,59	4261339,29	3337169,88	2628302,18	1684168,71	1423839,41
1999	3574747,71	4379966,28	3415734,07	2704782,49	1735102,95	1468714,82
2000	3668070,83	4498593,26	3494298,25	2781262,80	1786037,19	1513590,24
2001	3761393,95	4617220,25	3572862,43	2857743,11	1836971,43	1558465,65
2002	3854717,07	4735847,24	3651426,62	2934223,42	1887905,68	1603341,06
2003	3948040,19	4854474,22	3729990,80	3010703,73	1938839,92	1648216,47
2004	4041363,31	4973101,21	3808554,99	3087184,04	1989774,16	1693091,88
2005	4134686,43	5091728,19	3887119,17	3163664,36	2040708,40	1737967,29
2006	4228009,55	5210355,18	3965683,35	3240144,67	2091642,65	1782842,71
2007	4321332,67	5328982,16	4044247,54	3316624,98	2142576,89	1827718,12
2008	4414655,79	5447609,15	4122811,72	3393105,29	2193511,13	1872593,53

Tabela 10 Erro Percentual Absoluto de NDHP, no período [Jan-1990:Dez-2008] (RL)

	<i>Jan</i>	<i>Fev</i>	<i>Mar</i>	<i>Abr</i>	<i>Mai</i>	<i>Jun</i>	<i>Jul</i>	<i>Ago</i>	<i>Set</i>	<i>Out</i>	<i>Nov</i>	<i>Dez</i>
1990	1,5%	0,1%	2,7%	1,1%	7,9%	2,5%	0,5%	4,8%	2,6%	4,8%	5,6%	7,8%
1991	3,3%	2,1%	10,3%	2,1%	9,0%	9,6%	9,7%	8,5%	7,9%	7,7%	7,1%	4,8%
1992	4,5%	6,1%	2,0%	5,8%	6,3%	0,8%	1,9%	0,1%	1,4%	3,4%	10,5%	2,7%
1993	4,4%	6,5%	16,3%	9,7%	18,6%	19,2%	16,3%	9,1%	11,8%	9,8%	15,4%	10,0%
1994	10,2%	12,2%	7,1%	13,2%	8,8%	5,3%	2,9%	6,5%	6,4%	11,6%	3,3%	0,2%
1995	1,7%	3,2%	6,3%	6,9%	4,9%	3,2%	1,4%	2,2%	1,4%	2,5%	1,5%	4,0%
1996	1,8%	1,4%	1,7%	0,5%	2,1%	2,1%	3,8%	3,2%	2,5%	2,0%	3,2%	2,1%
1997	7,2%	6,8%	8,1%	4,0%	4,1%	3,3%	1,2%	0,5%	1,8%	0,7%	0,2%	1,7%
1998	4,1%	2,4%	1,6%	6,3%	6,1%	10,1%	8,4%	6,6%	9,5%	5,4%	2,8%	3,5%
1999	5,6%	5,6%	5,2%	1,5%	3,3%	3,4%	0,9%	0,2%	3,9%	8,4%	8,0%	1,9%
2000	2,4%	4,0%	6,5%	8,3%	3,3%	1,8%	3,2%	1,7%	2,9%	6,4%	9,7%	3,9%
2001	7,4%	9,1%	3,0%	7,2%	0,8%	3,6%	2,7%	4,3%	0,4%	1,0%	2,6%	5,3%
2002	0,5%	4,7%	11,0%	6,2%	0,3%	1,8%	1,4%	3,4%	2,2%	2,8%	1,2%	6,6%
2003	1,2%	1,0%	3,5%	0,7%	0,1%	0,9%	0,7%	0,4%	5,0%	5,9%	5,2%	6,4%
2004	3,8%	0,3%	8,3%	3,5%	5,5%	2,9%	6,7%	2,5%	4,2%	1,5%	6,3%	2,0%
2005	1,7%	6,0%	1,6%	11,9%	2,0%	3,8%	1,7%	1,1%	1,0%	1,3%	2,8%	1,6%
2006	0,8%	4,9%	9,0%	5,9%	0,3%	0,9%	1,8%	4,5%	2,6%	1,5%	1,0%	9,5%
2007	1,9%	3,4%	1,5%	1,7%	1,9%	1,9%	6,0%	5,6%	6,1%	4,0%	8,1%	6,6%
2008	2,9%	4,2%	7,8%	8,3%	5,2%	2,6%	2,6%	3,1%	2,4%	1,6%	3,8%	6,7%

Tabela 11 EPAM em função do n.º de neurónios e algoritmo de treino

<i>N.º Neurónios/Alg. Treino</i>	<i>RProp</i>	<i>L. Marquardt</i>	<i>G. Descent</i>	<i>Powell Beale</i>
2	0,046665	0,037451	0,062374	0,04218
4	0,038637	0,031824	0,058704	0,041795
6	0,041255	0,023662	0,05622	0,041901
8	0,039647	0,026907	0,061754	0,042607
10	0,041264	0,021225	0,064829	0,039078
12	0,040679	0,024146	0,06163	0,04211
14	0,040997	0,018643	0,057145	0,047418
16	0,037655	0,019873	0,05996	0,040836
18	0,038968	0,017937	0,058882	0,039914
20	0,043128	0,018971	0,064713	0,03985
<i>N.º Neurónios/Alg. Treino</i>	<i>Fletcher-Reeves</i>	<i>Polak-Ribiere</i>	<i>One Step Secant</i>	<i>Scaled C.G.</i>
2	0,055575	0,049345	0,053771	0,050354
4	0,051412	0,048155	0,050344	0,053851
6	0,047066	0,045653	0,049893	0,043767
8	0,045402	0,044643	0,048193	0,044449
10	0,042954	0,047455	0,051027	0,041496
12	0,042395	0,044526	0,047387	0,04761
14	0,039397	0,043922	0,048416	0,0399
16	0,042284	0,044456	0,04718	0,045111
18	0,04178	0,046788	0,0491	0,043257
20	0,037338	0,042822	0,049012	0,043874

Tabela 12 Tempo de execução em função do n.º de neurónios e algoritmo de treino

<i>N.º Neurónios/Alg. Treino</i>	<i>RProp</i>	<i>L. Marquardt</i>	<i>G. Descent</i>	<i>Powell Beale</i>
2	0,02381	0,00734	0,045544	0,00828
4	0,023473	0,014768	0,05093	0,028994
6	0,016878	0,018166	0,05364	0,053823
8	0,018104	0,014889	0,02719	0,025793
10	0,047215	0,017665	0,026885	0,038792
12	0,022417	0,019071	0,058983	0,04023
14	0,015319	0,010005	0,054306	0,019887
16	0,030401	0,020542	0,057168	0,031979
18	0,025508	0,020404	0,062926	0,033803
20	0,021422	0,01046	0,058096	0,036719
<i>N.º Neurónios/Alg. Treino</i>	<i>Fletcher-Reeves</i>	<i>Polak-Ribiere</i>	<i>One Step Secant</i>	<i>Scaled C.G.</i>
2	0,021569	0,022865	0,031581	0,039543
4	0,029607	0,025811	0,022392	0,030231
6	0,024005	0,018692	0,026317	0,059668
8	0,034539	0,026141	0,047009	0,019669
10	0,022865	0,023824	0,018745	0,02825
12	0,041064	0,022342	0,026418	0,017358
14	0,035694	0,03246	0,018197	0,013371
16	0,026485	0,023432	0,041012	0,022143
18	0,051265	0,035429	0,030749	0,017529
20	0,031076	0,023039	0,046107	0,01782

Tabela 13 Valores previstos para a Série NDHP, no período [Jan-1991:Dez-2008] (RNA - LM)

	<i>Jan</i>	<i>Fev</i>	<i>Mar</i>	<i>Abr</i>	<i>Mai</i>	<i>Jun</i>
1991	1105488,50	1242090,49	1652241,17	2121009,14	2114458,45	2381961,33
1992	1183942,17	1342056,73	1913111,29	2144673,38	2542146,77	2758928,67
1993	1236593,91	1402638,53	1780620,34	2315417,27	2505294,67	2543944,69
1994	1155588,57	1289800,98	1633663,71	2109241,45	2106751,80	2271918,40
1995	1165405,21	1314858,69	1810350,20	2141970,64	2428849,06	2667292,47
1996	1286730,52	1442298,01	1939154,80	2614631,92	2760853,52	2903452,35
1997	1321481,82	1552551,05	2032630,68	2586847,54	2674066,37	2864476,18
1998	1327265,73	1528554,28	2293728,65	2578552,15	2892484,99	2984259,73
1999	1474088,83	1712004,63	2261533,78	2870649,48	3071923,90	3402034,51
2000	1527548,94	1795552,85	2397675,49	2842208,25	3078274,50	3305033,93
2001	1529963,98	1831616,08	2542758,43	3058127,85	3140615,27	3353424,94
2002	1620744,89	1956720,74	2524779,41	3140657,80	3118902,49	3444496,58
2003	1623800,87	1948574,19	2675772,53	2895857,38	3175584,10	3500328,76
2004	1623974,65	1855764,45	2470054,39	3069889,55	3224647,31	3498590,05
2005	1629902,10	1934686,28	2407239,83	3092263,11	3155910,22	3505537,51
2006	1719034,58	1914252,83	2644068,87	2954294,34	3394655,45	3631000,84
2007	1776874,95	1942635,93	2548971,54	3513638,83	3525001,24	3783009,19
2008	1851736,83	2037206,89	2848536,12	3517085,56	3596602,49	3963894,18
	<i>Jul</i>	<i>Ago</i>	<i>Set</i>	<i>Out</i>	<i>Nov</i>	<i>Dez</i>
1991	2880721,43	3478743,50	2933821,97	2203266,39	1416593,19	1131423,94
1992	3166564,10	3644485,02	3143113,32	2317597,49	1411923,04	1158688,74
1993	2876106,24	3459166,07	2950507,56	2130977,45	1270043,29	1136885,93
1994	2711543,06	3359629,69	2797094,03	2098022,15	1289904,11	1107425,14
1995	3102463,90	3539251,00	3013785,84	2178530,90	1456464,65	1282675,73
1996	3300637,73	3772799,93	3207284,81	2412897,05	1586277,48	1344382,04
1997	3262786,63	3867523,28	3282578,74	2487247,31	1597084,34	1300838,24
1998	3489291,27	4137046,60	3446335,27	2666614,87	1686320,66	1380776,57
1999	3866887,42	4479132,70	3808155,25	2835267,21	1819884,67	1502895,88
2000	3685704,16	4327715,66	3752038,64	2969887,13	1930345,79	1519031,85
2001	3860776,04	4386021,62	3780127,01	2989738,20	1996686,57	1627374,66
2002	3956526,29	4408634,28	3778595,92	2962080,51	1919996,24	1475938,64
2003	4026577,51	4773739,13	3934376,63	2906735,90	1891088,88	1547129,49
2004	4024181,78	4689815,51	3784899,21	2924599,60	1875348,33	1579255,41
2005	3928296,68	4777325,44	3900432,00	3090199,77	1952674,69	1648680,35
2006	4149966,15	4959356,42	4172329,13	3187349,58	1998485,57	1825331,17
2007	4423206,55	5386486,46	4357672,63	3348878,42	2180881,96	1987900,98
2008	4715884,45	5514432,56	4528392,25	3532712,51	2339966,71	1898081,70

Tabela 14 Valores previstos para a Série NDHP, no período [Jan-1991:Dez-2008] (RNA - RP)

	<i>Jan</i>	<i>Fev</i>	<i>Mar</i>	<i>Abr</i>	<i>Mai</i>	<i>Jun</i>
1991	1086762,31	1242972,83	1720991,74	2142685,98	2330328,82	2521296,69
1992	1145900,23	1326979,81	1901285,64	2251745,13	2535764,28	2720987,52
1993	1163014,28	1426548,46	1841795,51	2113473,82	2372649,29	2370821,31
1994	1089947,81	1277434,72	1709231,15	2055994,03	2286938,81	2385190,37
1995	1131618,41	1335203,90	1858000,02	2219028,19	2602707,19	2842524,17
1996	1275399,69	1549652,45	2070468,22	2426222,59	2688223,69	2834906,12
1997	1264454,97	1500113,94	1993096,88	2516053,28	2728516,60	2926087,81
1998	1327039,46	1574512,93	2139123,62	2586476,42	2881970,33	3107798,63
1999	1462722,13	1743604,24	2417208,05	2874128,74	3033173,78	3273486,52
2000	1457156,68	1738451,12	2331806,42	2807394,44	3107297,83	3298255,62
2001	1555587,65	1925583,25	2482543,68	2808065,25	3123987,55	3336042,39
2002	1541299,77	1994309,75	2505885,51	2926039,64	3121422,81	3309068,22
2003	1548238,53	1875157,03	2499245,35	2972683,14	3177191,91	3484794,00
2004	1557205,30	1955967,93	2582802,62	2911071,50	3144746,79	3418553,72
2005	1579295,52	1914104,94	2491575,99	3031574,39	3134315,47	3422478,01
2006	1660271,88	1921335,18	2603946,53	3110050,41	3334606,61	3776495,25
2007	1764378,76	2224977,53	2916268,07	3255318,73	3446905,27	3969289,15
2008	1831602,20	2308912,93	3062723,21	3477125,79	3569211,29	4126017,12
	<i>Jul</i>	<i>Ago</i>	<i>Set</i>	<i>Out</i>	<i>Nov</i>	<i>Dez</i>
1991	2983116,81	3645677,22	2969212,17	2298466,30	1458951,13	1175784,58
1992	3080735,19	3739528,33	2950884,72	2209978,80	1373614,25	1124074,40
1993	2676739,03	3360918,57	2844955,96	2068129,15	1346384,83	1095214,79
1994	2838700,34	3452252,26	2824175,72	2138205,03	1350364,14	1150684,59
1995	3219308,34	3691031,60	2928201,62	2268896,80	1474181,51	1266549,34
1996	3238132,80	3856488,15	3102645,56	2389346,20	1573122,39	1286462,90
1997	3311656,34	3988637,57	3220470,52	2532009,85	1627815,74	1356273,35
1998	3676457,74	4449729,78	3418611,04	2874900,95	1783442,72	1479065,76
1999	3791112,68	4609282,14	3729080,99	2822077,77	1870915,81	1543964,61
2000	3701308,18	4446296,40	3543182,64	2957174,46	1877249,23	1618395,09
2001	3803271,52	4519573,59	3575027,74	2887984,19	1876773,32	1616718,07
2002	3848431,50	4570957,73	3771158,93	3075950,93	1884193,68	1634924,32
2003	3985515,06	4887183,50	3829328,38	2998672,42	1837923,70	1601132,44
2004	3988011,54	4784242,78	3650655,37	3017537,47	1920895,01	1621954,49
2005	3968831,19	4838242,63	3863012,45	3228852,95	1981269,58	1684385,74
2006	4269983,07	4958035,75	3943897,86	3483294,82	2044857,81	1797642,29
2007	4465803,39	5183870,70	4268942,73	3630900,53	2145065,13	1882347,68
2008	4560238,61	5239457,53	4518712,09	3613208,37	2165760,03	1850040,39

Tabela 15 Erro Percentual Absoluto de NDHP, no período [Jan-1990:Dez-2008] (RNA - LM)

	<i>Jan</i>	<i>Fev</i>	<i>Mar</i>	<i>Abr</i>	<i>Mai</i>	<i>Jun</i>	<i>Jul</i>	<i>Ago</i>	<i>Set</i>	<i>Out</i>	<i>Nov</i>	<i>Dez</i>
1991	1,8%	4,2%	14,7%	5,5%	14,9%	9,7%	8,0%	7,2%	3,1%	2,8%	0,9%	3,0%
1992	0,1%	4,5%	8,3%	4,9%	1,8%	10,8%	10,4%	2,6%	11,2%	10,5%	13,1%	3,1%
1993	10,5%	8,9%	10,9%	15,4%	22,9%	18,7%	10,9%	2,9%	12,0%	4,2%	2,5%	4,2%
1994	5,3%	1,7%	9,4%	4,9%	8,2%	9,4%	10,3%	5,5%	1,5%	0,8%	10,0%	10,8%
1995	5,1%	8,0%	3,6%	15,1%	10,2%	5,2%	4,4%	7,4%	1,4%	6,9%	6,3%	4,5%
1996	1,7%	6,8%	4,3%	7,3%	6,3%	5,7%	4,0%	3,2%	3,4%	0,6%	3,4%	2,9%
1997	6,6%	5,0%	12,2%	7,7%	6,0%	2,3%	2,5%	7,1%	2,6%	3,2%	2,4%	4,0%
1998	7,0%	8,5%	3,0%	6,0%	3,2%	9,8%	8,2%	9,3%	6,5%	4,0%	2,7%	6,5%
1999	1,2%	3,9%	5,0%	6,9%	2,9%	7,4%	7,2%	2,5%	7,1%	3,9%	3,5%	4,2%
2000	2,9%	0,5%	3,4%	4,2%	0,5%	3,2%	2,7%	2,2%	4,2%	0,0%	2,4%	3,6%
2001	4,9%	6,7%	3,4%	1,5%	2,5%	0,1%	0,1%	0,9%	6,2%	3,5%	5,8%	9,9%
2002	5,4%	1,6%	8,3%	16,2%	0,3%	2,2%	1,2%	10,0%	1,3%	3,8%	0,5%	1,9%
2003	4,7%	4,2%	10,2%	1,0%	0,6%	4,0%	2,7%	1,2%	10,8%	2,2%	2,6%	0,1%
2004	4,8%	4,1%	3,7%	5,2%	3,8%	3,4%	6,2%	3,3%	3,6%	3,8%	0,2%	4,9%
2005	0,5%	3,0%	10,4%	11,9%	4,0%	2,1%	3,4%	5,1%	1,3%	1,0%	1,7%	6,6%
2006	2,5%	1,7%	6,4%	12,3%	1,3%	0,4%	3,6%	9,1%	2,5%	3,1%	5,4%	7,3%
2007	0,8%	4,1%	9,5%	6,5%	1,4%	0,6%	3,8%	4,6%	1,2%	3,0%	6,4%	1,6%
2008	1,6%	9,0%	7,5%	14,8%	4,9%	6,5%	4,1%	1,9%	7,2%	5,8%	10,8%	8,2%

Tabela 16 Erro Percentual Absoluto de NDHP, no período [Jan-1991:Dez-2008] (RNA - RP)

	<i>Jan</i>	<i>Fev</i>	<i>Mar</i>	<i>Abr</i>	<i>Mai</i>	<i>Jun</i>	<i>Jul</i>	<i>Ago</i>	<i>Set</i>	<i>Out</i>	<i>Nov</i>	<i>Dez</i>
1991	3,45%	4,11%	11,14%	6,62%	6,23%	4,42%	4,73%	2,76%	1,92%	1,40%	2,10%	0,82%
1992	3,08%	5,62%	7,60%	0,15%	1,52%	9,31%	7,44%	5,25%	4,40%	5,35%	10,07%	0,03%
1993	3,88%	10,74%	14,71%	5,34%	16,41%	10,67%	3,25%	0,06%	7,99%	1,09%	8,65%	0,39%
1994	0,65%	0,68%	5,22%	2,29%	0,31%	4,84%	6,06%	2,93%	0,59%	2,77%	5,74%	7,36%
1995	7,81%	6,54%	1,09%	12,03%	3,79%	1,08%	0,84%	3,42%	4,23%	3,02%	5,15%	5,68%
1996	0,77%	0,12%	2,13%	0,42%	3,54%	3,17%	2,05%	1,08%	0,01%	1,53%	2,58%	1,57%
1997	2,02%	1,49%	13,87%	4,71%	4,09%	4,51%	1,06%	4,20%	0,65%	1,43%	0,56%	0,05%
1998	7,06%	5,70%	3,93%	5,75%	3,59%	6,09%	3,26%	2,47%	7,27%	3,46%	2,91%	0,20%
1999	1,95%	2,13%	1,56%	6,98%	1,63%	3,36%	5,07%	5,48%	4,90%	4,39%	0,84%	7,07%
2000	1,86%	3,69%	6,03%	5,38%	1,40%	3,03%	2,32%	0,50%	1,57%	0,44%	5,08%	2,73%
2001	3,29%	1,96%	1,00%	6,76%	1,95%	0,39%	1,63%	2,10%	0,42%	0,01%	0,51%	9,21%
2002	0,27%	3,50%	8,95%	8,23%	0,41%	1,85%	1,56%	6,72%	1,06%	7,81%	1,35%	8,65%
2003	0,16%	0,29%	2,89%	1,63%	0,56%	3,55%	1,65%	1,12%	7,84%	5,46%	0,30%	3,38%
2004	0,47%	1,12%	8,41%	0,25%	1,26%	1,04%	5,26%	1,37%	0,11%	0,79%	2,66%	2,31%
2005	2,65%	1,88%	7,27%	9,67%	4,65%	0,30%	2,41%	3,90%	0,33%	3,40%	0,22%	4,59%
2006	0,96%	1,32%	4,81%	7,65%	3,07%	4,42%	0,83%	9,13%	3,15%	5,95%	3,17%	8,72%
2007	0,06%	9,88%	3,54%	1,33%	3,60%	4,29%	2,90%	8,16%	0,89%	5,12%	7,97%	3,83%
2008	0,49%	3,11%	0,52%	13,49%	5,65%	10,86%	0,66%	6,75%	6,97%	8,20%	2,51%	5,44%

Tabela 17 Dormidas registadas em Portugal (NDHP-RN) no período [Jan-1990:Dez-2008]

	<i>Jan</i>	<i>Fev</i>	<i>Mar</i>	<i>Abr</i>	<i>Mai</i>	<i>Jun</i>
1990	126826,00	131653,00	188999,00	182290,00	219187,00	251295,00
1991	124194,00	127474,00	157536,00	196087,00	223918,00	207907,00
1992	121469,00	129284,00	154734,00	189142,00	198402,00	207216,00
1993	118606,00	122988,00	175261,00	185525,00	232075,00	248237,00
1994	122480,00	130393,00	156645,00	209263,00	218666,00	222720,00
1995	126910,00	139403,00	172393,00	213973,00	239142,00	245264,00
1996	140430,00	141183,00	219465,00	224382,00	253833,00	238334,00
1997	148218,00	157415,00	209929,00	232767,00	280326,00	296612,00
1998	163696,00	165988,00	228149,00	242744,00	269854,00	270126,00
1999	162389,00	162637,00	226010,00	262865,00	264497,00	273881,00
2000	176690,00	186586,00	245261,00	291395,00	306743,00	325568,00
2001	165653,00	181005,00	249214,00	253274,00	302028,00	301465,00
2002	155527,00	177818,00	214106,00	258519,00	293531,00	271454,00
2003	162900,00	181900,00	224600,00	279800,00	317300,00	355300,00
2004	168100,00	166800,00	247000,00	268500,00	316900,00	307700,00
2005	180700,00	195100,00	237200,00	352600,00	361200,00	331500,00
2006	193600,00	211300,00	281400,00	360300,00	386900,00	374600,00
2007	213398,00	243765,00	337644,00	320364,00	398896,00	347620,00
2008	126826,00	131653,00	188999,00	182290,00	219187,00	251295,00
	<i>Jul</i>	<i>Ago</i>	<i>Set</i>	<i>Out</i>	<i>Nov</i>	<i>Dez</i>
1990	273927,00	341490,00	283378,00	197241,00	152554,00	132802,00
1991	231801,00	312026,00	259023,00	205400,00	149289,00	130963,00
1992	231453,00	304576,00	249583,00	202792,00	141976,00	120748,00
1993	246274,00	322366,00	266094,00	206256,00	144803,00	139706,00
1994	247589,00	320750,00	269433,00	196466,00	152340,00	140643,00
1995	248398,00	336086,00	280769,00	225734,00	175438,00	143163,00
1996	266993,00	345672,00	288409,00	232052,00	166835,00	141349,00
1997	303866,00	377645,00	309700,00	263522,00	180796,00	161273,00
1998	306031,00	385868,00	321248,00	280597,00	193062,00	166990,00
1999	324962,00	397405,00	331155,00	263217,00	186445,00	157210,00
2000	351955,00	452581,00	383793,00	319417,00	238925,00	202351,00
2001	314560,00	444991,00	361181,00	287383,00	221910,00	179766,00
2002	318706,00	433211,00	343534,00	281472,00	219463,00	178439,00
2003	324400,00	426900,00	342100,00	311500,00	221200,00	182800,00
2004	358500,00	472400,00	362200,00	315900,00	233400,00	221300,00
2005	388400,00	524500,00	406500,00	353300,00	258800,00	254700,00
2006	430400,00	576400,00	459000,00	387500,00	303000,00	264700,00
2007	430043,00	575652,00	462421,00	381753,00	283750,00	253297,00
2008	273927,00	341490,00	283378,00	197241,00	152554,00	132802,00

Tabela 18 Valores previstos para a Série NDHP-RN, no período [Jan-1990:Dez-2008] (RL)

	<i>Jan</i>	<i>Fev</i>	<i>Mar</i>	<i>Abr</i>	<i>Mai</i>	<i>Jun</i>
1990	115797,24	119073,10	156564,47	172504,59	195277,25	209796,49
1991	119642,13	123508,29	162473,35	180653,89	203996,11	217261,18
1992	123487,03	127943,49	168382,22	188803,19	212714,97	224725,87
1993	127331,93	132378,68	174291,10	196952,49	221433,83	232190,56
1994	131176,82	136813,87	180199,97	205101,79	230152,69	239655,25
1995	135021,72	141249,07	186108,85	213251,10	238871,54	247119,93
1996	138866,62	145684,26	192017,72	221400,40	247590,40	254584,62
1997	142711,51	150119,45	197926,60	229549,70	256309,26	262049,31
1998	146556,41	154554,65	203835,47	237699,00	265028,12	269514,00
1999	150401,31	158989,84	209744,35	245848,30	273746,98	276978,69
2000	154246,21	163425,03	215653,22	253997,60	282465,83	284443,38
2001	158091,10	167860,23	221562,10	262146,90	291184,69	291908,07
2002	161936,00	172295,42	227470,97	270296,21	299903,55	299372,75
2003	165780,90	176730,62	233379,85	278445,51	308622,41	306837,44
2004	169625,79	181165,81	239288,72	286594,81	317341,26	314302,13
2005	173470,69	185601,00	245197,60	294744,11	326060,12	321766,82
2006	177315,59	190036,20	251106,47	302893,41	334778,98	329231,51
2007	181160,49	194471,39	257015,35	311042,71	343497,84	336696,20
2008	185005,38	198906,58	262924,22	319192,01	352216,70	344160,89
	<i>Jul</i>	<i>Ago</i>	<i>Set</i>	<i>Out</i>	<i>Nov</i>	<i>Dez</i>
1990	227018,24	293495,78	248604,45	182392,31	133779,27	114577,65
1991	235372,19	304970,33	256815,42	191641,33	140581,12	120798,72
1992	243726,15	316444,88	265026,38	200890,34	147382,96	127019,79
1993	252080,10	327919,43	273237,35	210139,35	154184,80	133240,87
1994	260434,06	339393,98	281448,31	219388,36	160986,64	139461,94
1995	268788,01	350868,53	289659,28	228637,38	167788,48	145683,01
1996	277141,97	362343,08	297870,25	237886,39	174590,32	151904,09
1997	285495,93	373817,63	306081,21	247135,40	181392,16	158125,16
1998	293849,88	385292,18	314292,18	256384,41	188194,00	164346,24
1999	302203,84	396766,73	322503,14	265633,42	194995,84	170567,31
2000	310557,79	408241,27	330714,11	274882,44	201797,68	176788,38
2001	318911,75	419715,82	338925,07	284131,45	208599,52	183009,46
2002	327265,71	431190,37	347136,04	293380,46	215401,36	189230,53
2003	335619,66	442664,92	355347,00	302629,47	222203,20	195451,60
2004	343973,62	454139,47	363557,97	311878,49	229005,04	201672,68
2005	352327,57	465614,02	371768,94	321127,50	235806,88	207893,75
2006	360681,53	477088,57	379979,90	330376,51	242608,73	214114,82
2007	369035,49	488563,12	388190,87	339625,52	249410,57	220335,90
2008	377389,44	500037,67	396401,83	348874,53	256212,41	226556,97

Tabela 19 Valores previstos para a Série NDHP-RN, no período [Jan-1991:Dez-2008] (RNA)

	<i>Jan</i>	<i>Fev</i>	<i>Mar</i>	<i>Abr</i>	<i>Mai</i>	<i>Jun</i>
1991	127242,68	125376,45	160016,36	203520,84	220939,93	227982,63
1992	125050,53	133322,38	171072,85	199146,16	216280,18	218235,16
1993	129190,06	133031,22	162112,48	195770,17	224364,75	214811,46
1994	126909,84	128299,06	156492,13	189397,76	211237,18	232372,28
1995	122906,12	134195,38	171395,47	196547,96	233233,03	231801,42
1996	128974,47	134039,77	166499,45	210899,63	236694,94	238075,43
1997	138149,61	143483,58	180498,55	224199,31	251888,31	275691,26
1998	146474,07	159645,7	204642,25	233082,67	265972,09	283753,87
1999	147733,43	164102,53	214999,35	252133,71	278601,9	305732,14
2000	163333,71	173218,11	218105,95	250995,95	289369,6	292642,53
2001	158530,95	168958,38	225074,04	273540,32	304381,42	322350,14
2002	181380,55	188376,31	243262,06	296088,08	302401,27	309163,59
2003	172856,56	179972,08	231691,81	268041,93	297064,27	291863,89
2004	163169,57	171252,59	217556,73	271561,43	309652,57	303849,13
2005	170602,28	182748,78	243777,26	285955,75	288462,75	345227,49
2006	164690,36	181339,92	244985,45	291366,11	359085,19	335355,32
2007	187412,33	202709,14	270677,49	353285,57	396751,38	357820,36
2008	202149,01	215980,68	296642,72	384086,35	388655,96	383957,83
	<i>Jul</i>	<i>Ago</i>	<i>Set</i>	<i>Out</i>	<i>Nov</i>	<i>Dez</i>
1991	247594,79	342376,65	282105,42	229181,9	167859,8	140061,2
1992	234209,28	315881,86	264997,08	203716,88	159300,17	141843,86
1993	228494,58	294167,83	255060,59	206387,17	160465,96	139906,63
1994	250280,69	305729,8	248288,61	203870,97	146660,19	130200,35
1995	243644,73	311061,99	263773,82	206089,56	159992,02	139121,14
1996	266935,77	317790,02	260997,48	206431,68	163325,77	146349,54
1997	274829,2	341314,21	275400,06	229881,72	173741,21	159453,54
1998	308121,23	363632,6	293066,65	235330,54	170616,44	154803,49
1999	305828,32	384601,87	318769,37	263410,32	194099,81	181116,83
2000	308326,34	400147,62	340337,73	277058,49	202146,26	175796,61
2001	360529,94	421928	338125,71	271538,46	202517,73	184065,49
2002	326503,93	445770,69	403196,32	336423,91	236913,76	201091,63
2003	299318,91	428565,56	378446,32	301171,87	219497,97	185482,9
2004	352241,14	446419,42	363200,44	289546,72	208750,45	176915,82
2005	325979,39	450690,2	383818,47	323199,93	217072,35	191567,27
2006	372713,94	492516,56	414895,45	330774,43	244658,26	213573,51
2007	396945,63	548817,65	469711,24	383475,63	278990,81	241774,97
2008	407272,53	572301,38	508029,69	435851,25	315061,12	278706,71

Tabela 20 Erro Percentual Absoluto de NDHP-RN, no período [Jan-1990:Dez-2008] (RL)

	<i>Jan</i>	<i>Fev</i>	<i>Mar</i>	<i>Abr</i>	<i>Mai</i>	<i>Jun</i>	<i>Jul</i>	<i>Ago</i>	<i>Set</i>	<i>Out</i>	<i>Nov</i>	<i>Dez</i>
1990	8,6%	8,3%	1,3%	12,8%	6,1%	7,6%	11,9%	16,5%	12,7%	15,7%	17,5%	18,0%
1991	5,7%	6,2%	14,0%	0,9%	6,9%	13,5%	14,1%	10,7%	9,4%	2,8%	7,8%	9,0%
1992	0,6%	0,4%	6,9%	3,7%	5,0%	8,1%	5,1%	1,4%	2,3%	2,2%	1,3%	3,0%
1993	4,8%	2,4%	12,6%	4,1%	11,6%	12,1%	8,9%	7,7%	9,5%	3,6%	8,6%	10,3%
1994	10,6%	11,2%	2,8%	10,6%	0,8%	3,5%	5,7%	5,3%	5,8%	6,4%	11,2%	0,2%
1995	10,2%	8,3%	18,8%	1,9%	9,2%	11,0%	8,6%	9,4%	7,5%	16,4%	10,1%	3,6%
1996	9,4%	4,5%	11,4%	3,5%	3,5%	3,8%	11,6%	7,8%	6,1%	5,4%	0,5%	6,1%
1997	1,6%	6,3%	9,8%	2,3%	1,0%	10,0%	6,9%	8,1%	6,1%	6,5%	8,7%	11,9%
1998	1,1%	1,8%	2,9%	2,1%	5,5%	9,1%	3,3%	2,0%	1,5%	2,7%	4,1%	1,9%
1999	8,1%	4,2%	8,1%	1,3%	1,4%	2,5%	1,3%	2,8%	0,4%	5,3%	1,0%	2,1%
2000	5,0%	0,5%	4,6%	3,4%	6,8%	3,9%	4,4%	2,7%	0,1%	4,4%	8,2%	12,5%
2001	10,5%	10,0%	9,7%	10,0%	5,1%	10,3%	9,4%	7,3%	11,7%	11,0%	12,7%	9,6%
2002	2,2%	4,8%	8,7%	6,7%	0,7%	0,7%	4,0%	3,1%	3,9%	2,1%	2,9%	5,3%
2003	6,6%	0,6%	9,0%	7,7%	5,1%	13,0%	5,3%	2,2%	3,4%	7,5%	1,2%	9,5%
2004	4,1%	0,4%	6,5%	2,4%	0,0%	11,5%	6,0%	6,4%	6,3%	0,1%	3,5%	10,3%
2005	3,2%	11,3%	0,7%	9,8%	2,9%	4,6%	1,7%	1,4%	2,6%	1,7%	1,0%	6,1%
2006	1,9%	2,6%	5,9%	14,1%	7,3%	0,7%	7,1%	9,0%	6,5%	6,5%	6,3%	15,9%
2007	6,4%	8,0%	8,7%	13,7%	11,2%	10,1%	14,3%	15,2%	15,4%	12,4%	17,7%	16,8%
2008	13,3%	18,4%	22,1%	0,4%	11,7%	1,0%	12,2%	13,1%	14,3%	8,6%	9,7%	10,6%

Tabela 21 Erro Percentual Absoluto de NDHP-RN, no período [Jan-1991:Dez-2008] (RNA)

	<i>Jan</i>	<i>Fev</i>	<i>Mar</i>	<i>Abr</i>	<i>Mai</i>	<i>Jun</i>	<i>Jul</i>	<i>Ago</i>	<i>Set</i>	<i>Out</i>	<i>Nov</i>	<i>Dez</i>
1991	0,3%	4,8%	15,3%	11,6%	0,8%	9,3%	9,6%	0,3%	0,4%	16,2%	10,0%	5,5%
1992	0,7%	4,6%	8,6%	1,6%	3,4%	5,0%	1,0%	1,2%	2,3%	0,8%	6,7%	8,3%
1993	6,4%	2,9%	4,8%	3,5%	13,1%	3,7%	1,3%	3,4%	2,2%	1,8%	13,0%	15,9%
1994	7,0%	4,3%	10,7%	2,1%	9,0%	6,4%	1,6%	5,2%	6,7%	1,2%	1,3%	6,8%
1995	0,3%	2,9%	9,4%	6,1%	6,7%	4,1%	1,6%	3,0%	2,1%	4,9%	5,0%	1,1%
1996	1,6%	3,8%	3,4%	1,4%	1,0%	2,9%	7,5%	5,4%	7,0%	8,6%	6,9%	2,2%
1997	1,6%	1,6%	17,8%	0,1%	0,8%	15,7%	2,9%	1,3%	4,5%	0,9%	4,1%	12,8%
1998	1,2%	1,4%	2,5%	0,1%	5,1%	4,3%	1,4%	3,7%	5,4%	10,7%	5,6%	4,0%
1999	9,8%	1,1%	5,8%	3,9%	3,2%	13,2%	0,1%	0,3%	0,8%	6,1%	0,5%	8,5%
2000	0,6%	6,5%	3,5%	4,5%	9,4%	6,9%	5,1%	0,7%	2,8%	5,3%	8,4%	11,8%
2001	10,3%	9,4%	8,2%	6,1%	0,8%	1,0%	2,4%	6,8%	11,9%	15,0%	15,2%	9,0%
2002	9,5%	4,1%	2,4%	16,9%	0,1%	2,6%	3,8%	0,2%	11,6%	17,1%	6,8%	11,9%
2003	11,1%	1,2%	8,2%	3,7%	1,2%	7,5%	6,1%	1,1%	10,2%	7,0%	0,0%	3,9%
2004	0,2%	5,9%	3,1%	2,9%	2,4%	14,5%	8,6%	4,6%	6,2%	7,0%	5,6%	3,2%
2005	1,5%	9,6%	1,3%	6,5%	9,0%	12,2%	9,1%	4,6%	6,0%	2,3%	7,0%	13,4%
2006	8,9%	7,1%	3,3%	17,4%	0,6%	1,2%	4,0%	6,1%	2,1%	6,4%	5,5%	16,1%
2007	3,2%	4,1%	3,8%	1,9%	2,5%	4,5%	7,8%	4,8%	2,3%	1,0%	7,9%	8,7%
2008	5,3%	11,4%	12,1%	19,9%	2,6%	10,5%	5,3%	0,6%	9,9%	14,2%	11,0%	10,0%

Tabela 22 Dormidas registadas em Portugal (NDHP-RC) no período [Jan-1990:Dez-2008]

	<i>Jan</i>	<i>Fev</i>	<i>Mar</i>	<i>Abr</i>	<i>Mai</i>	<i>Jun</i>
1990	66389,00	78898,00	91836,00	121039,00	125580,00	138110,00
1991	67712,00	81963,00	114931,00	112756,00	130316,00	140715,00
1992	72006,00	78873,00	98200,00	124425,00	141334,00	121020,00
1993	73457,00	82466,00	93210,00	125441,00	127772,00	122687,00
1994	69142,00	80463,00	101582,00	113765,00	125687,00	125656,00
1995	70798,00	81326,00	104727,00	139292,00	133419,00	130530,00
1996	69186,00	89418,00	110697,00	145682,00	142172,00	141044,00
1997	82964,00	95439,00	137757,00	136194,00	159817,00	144019,00
1998	95078,00	106779,00	122126,00	151959,00	176390,00	173863,00
1999	105697,00	123941,00	136214,00	155533,00	165865,00	169182,00
2000	97748,00	112210,00	141973,00	173166,00	173781,00	167906,00
2001	97835,00	117057,00	138851,00	164615,00	168582,00	171690,00
2002	99913,00	118807,00	156803,00	154440,00	172775,00	172701,00
2003	136669,00	146512,00	196309,00	240487,00	282940,00	256314,00
2004	144481,00	169494,00	206316,00	263603,00	290185,00	308510,00
2005	146800,00	172000,00	251100,00	264200,00	299900,00	293000,00
2006	146300,00	184500,00	219800,00	317200,00	320500,00	294000,00
2007	162664,00	190582,00	238453,00	330410,00	337547,00	319452,00
2008	173380,00	218199,00	293540,00	291930,00	353025,00	317896,00
	<i>Jul</i>	<i>Ago</i>	<i>Set</i>	<i>Out</i>	<i>Nov</i>	<i>Dez</i>
1990	183161,00	259879,00	190030,00	127891,00	83646,00	82305,00
1991	175843,00	267754,00	193701,00	123425,00	85675,00	76662,00
1992	163168,00	247192,00	175842,00	121295,00	84867,00	78134,00
1993	158791,00	247527,00	176980,00	118980,00	72739,00	72227,00
1994	166728,00	250555,00	177707,00	116944,00	80985,00	81664,00
1995	164749,00	242433,00	171988,00	116247,00	80925,00	97189,00
1996	166283,00	241940,00	187513,00	137972,00	100324,00	93096,00
1997	185696,00	262815,00	193321,00	147357,00	107827,00	100364,00
1998	200270,00	294081,00	216871,00	162655,00	109382,00	96465,00
1999	203694,00	280780,00	214071,00	161856,00	122468,00	108546,00
2000	211569,00	296264,00	213978,00	162932,00	131786,00	116821,00
2001	200343,00	287122,00	211241,00	163283,00	125344,00	110652,00
2002	185184,00	288336,00	211734,00	158020,00	125915,00	108691,00
2003	297678,00	439293,00	319576,00	257783,00	183431,00	161020,00
2004	308175,00	442413,00	331474,00	300534,00	182155,00	163759,00
2005	348200,00	496700,00	353900,00	294200,00	198200,00	179200,00
2006	358000,00	534200,00	388300,00	316100,00	208200,00	221000,00
2007	407512,00	585003,00	426485,00	381577,00	256977,00	214573,00
2008	403899,00	581096,00	407699,00	357546,00	274371,00	213193,00

Tabela 23 Valores previstos para a Série NDHP-RC, no período [Jan-1990:Dez-2008] (RL)

	<i>Jan</i>	<i>Fev</i>	<i>Mar</i>	<i>Abr</i>	<i>Mai</i>	<i>Jun</i>
1990	52914,88	60524,55	73371,47	85465,12	90834,20	92472,63
1991	58375,34	67077,70	82012,04	96133,43	102546,21	103494,04
1992	63835,79	73630,85	90652,60	106801,74	114258,22	114515,46
1993	69296,25	80184,00	99293,17	117470,04	125970,23	125536,87
1994	74756,71	86737,16	107933,74	128138,35	137682,25	136558,28
1995	80217,16	93290,31	116574,30	138806,66	149394,26	147579,70
1996	85677,62	99843,46	125214,87	149474,97	161106,27	158601,11
1997	91138,07	106396,61	133855,43	160143,28	172818,28	169622,53
1998	96598,53	112949,76	142496,00	170811,59	184530,29	180643,94
1999	102058,99	119502,92	151136,57	181479,90	196242,31	191665,36
2000	107519,44	126056,07	159777,13	192148,21	207954,32	202686,77
2001	112979,90	132609,22	168417,70	202816,51	219666,33	213708,18
2002	118440,35	139162,37	177058,26	213484,82	231378,34	224729,60
2003	123900,81	145715,52	185698,83	224153,13	243090,36	235751,01
2004	129361,26	152268,68	194339,40	234821,44	254802,37	246772,43
2005	134821,72	158821,83	202979,96	245489,75	266514,38	257793,84
2006	140282,18	165374,98	211620,53	256158,06	278226,39	268815,25
2007	145742,63	171928,13	220261,10	266826,37	289938,40	279836,67
2008	151203,09	178481,28	228901,66	277494,68	301650,42	290858,08
	<i>Jul</i>	<i>Ago</i>	<i>Set</i>	<i>Out</i>	<i>Nov</i>	<i>Dez</i>
1990	128896,71	192131,41	139270,76	80824,41	57642,55	55965,00
1991	139825,29	207668,54	150745,94	92688,08	65759,80	63291,40
1992	150753,88	223205,68	162221,12	104551,75	73877,04	70617,79
1993	161682,47	238742,81	173696,29	116415,42	81994,29	77944,19
1994	172611,06	254279,94	185171,47	128279,09	90111,54	85270,59
1995	183539,65	269817,07	196646,65	140142,76	98228,79	92596,99
1996	194468,24	285354,21	208121,82	152006,43	106346,03	99923,38
1997	205396,82	300891,34	219597,00	163870,10	114463,28	107249,78
1998	216325,41	316428,47	231072,18	175733,76	122580,53	114576,18
1999	227254,00	331965,60	242547,35	187597,43	130697,78	121902,57
2000	238182,59	347502,74	254022,53	199461,10	138815,02	129228,97
2001	249111,18	363039,87	265497,71	211324,77	146932,27	136555,37
2002	260039,76	378577,00	276972,88	223188,44	155049,52	143881,76
2003	270968,35	394114,13	288448,06	235052,11	163166,77	151208,16
2004	281896,94	409651,26	299923,24	246915,78	171284,01	158534,56
2005	292825,53	425188,40	311398,41	258779,45	179401,26	165860,96
2006	303754,12	440725,53	322873,59	270643,12	187518,51	173187,35
2007	314682,71	456262,66	334348,76	282506,79	195635,76	180513,75
2008	325611,29	471799,79	345823,94	294370,46	203753,00	187840,15

Tabela 24 Valores previstos para a Série NDHP-RC, no período [Jan-1991:Dez-2008] (RNA)

	<i>Jan</i>	<i>Fev</i>	<i>Mar</i>	<i>Abr</i>	<i>Mai</i>	<i>Jun</i>
1991	72490,857	86021,835	117342,21	143313,38	133349,98	129414,32
1992	67917,701	89324,807	121431,28	137181,93	138623,33	136957,59
1993	74363,522	82655,277	105032,07	132986,16	143558,02	117716,54
1994	72519,979	80049,34	104643,67	136109,03	137069,93	121178,4
1995	69271,543	83019,809	112332,14	133995,61	136828,98	124030,86
1996	77278,391	80450,331	114126,06	140412,53	141823,84	123636,39
1997	80450,482	86891,354	113992,81	151362,26	147721,02	138523,67
1998	83978,076	98401,681	130992,77	151676,86	163002,93	141049,66
1999	95616,603	113784,22	140454,43	170737,78	185686,82	172657,41
2000	98805,907	113776,5	142665,15	169273,96	173415,4	165084,2
2001	99960,39	111408,31	150181,29	177489,33	174521,38	155069,54
2002	98022,247	111979,52	145352,14	177761,99	172829,85	163789,59
2003	95354,092	121134,71	153892,05	196470,49	217005,45	221922,56
2004	126802,55	166733,91	222134,65	264496,95	302074,96	279988,52
2005	150757,48	181787,34	228607,58	276618,82	312737,12	345182,54
2006	139340,04	182532,45	262476,14	285687,9	323078,22	328633,84
2007	175159,33	175637,53	258819,4	310315,61	343081,64	312799,91
2008	190244,73	180559,82	283113,85	330563,44	335940,98	335278,62
	<i>Jul</i>	<i>Ago</i>	<i>Set</i>	<i>Out</i>	<i>Nov</i>	<i>Dez</i>
1991	185927,99	259227,46	199592,68	148146,66	83268,544	79231,902
1992	176209,89	265100,08	201550,06	128555,12	87558,908	88514,287
1993	157260,62	243928,05	184025,63	127957,02	87253,168	83707,37
1994	157143,08	243278,16	183920,36	133281,54	79618,765	80176,619
1995	165360,45	251256,97	181115,78	120413,13	88111,904	90539,647
1996	165833,89	243927,86	177042,88	122533,04	101219,82	103579,16
1997	169363,61	243818,4	189423,66	148516,88	106182,8	101997,19
1998	188650,15	273570,77	221563,47	165912,97	101673,88	95351,622
1999	197123,89	289521,81	230947,94	180014,49	109553,79	99062,304
2000	209043,19	290701,88	231212,9	171166,33	117243,5	114186,87
2001	213978,23	300311,83	229190,09	170823,51	120929,08	114752,82
2002	204647,03	285482,25	216258,94	174520,52	118049,64	111273,87
2003	226664,34	338154,53	263687,62	218847,07	186165,21	154738,43
2004	350910,43	427364,2	364437,7	290677,88	206833,92	166968,7
2005	362667,56	431020,88	360791,11	338041,69	218681,99	180366,25
2006	388790,07	459608,85	405880,3	318569,17	227663,73	201002,88
2007	402023,99	484283,48	435919,77	330716,24	231300,92	224928,66
2008	420302,49	477949,08	437273,18	417128,78	247007,17	223535,06

Tabela 25 Erro Percentual Absoluto de NDHP-RC, no período [Jan-1990:Dez-2008] (RL)

	<i>Jan</i>	<i>Fev</i>	<i>Mar</i>	<i>Abr</i>	<i>Mai</i>	<i>Jun</i>	<i>Jul</i>	<i>Ago</i>	<i>Set</i>	<i>Out</i>	<i>Nov</i>	<i>Dez</i>
1990	20,3%	23,3%	20,1%	29,4%	27,7%	33,0%	29,6%	26,1%	26,7%	36,8%	31,1%	32,0%
1991	13,8%	18,2%	28,6%	14,7%	21,3%	26,5%	20,5%	22,4%	22,2%	24,9%	23,2%	17,4%
1992	11,3%	6,6%	7,7%	14,2%	19,2%	5,4%	7,6%	9,7%	7,7%	13,8%	12,9%	9,6%
1993	5,7%	2,8%	6,5%	6,4%	1,4%	2,3%	1,8%	3,5%	1,9%	2,2%	12,7%	7,9%
1994	8,1%	7,8%	6,3%	12,6%	9,5%	8,7%	3,5%	1,5%	4,2%	9,7%	11,3%	4,4%
1995	13,3%	14,7%	11,3%	0,3%	12,0%	13,1%	11,4%	11,3%	14,3%	20,6%	21,4%	4,7%
1996	23,8%	11,7%	13,1%	2,6%	13,3%	12,4%	17,0%	17,9%	11,0%	10,2%	6,0%	7,3%
1997	9,9%	11,5%	2,8%	17,6%	8,1%	17,8%	10,6%	14,5%	13,6%	11,2%	6,2%	6,9%
1998	1,6%	5,8%	16,7%	12,4%	4,6%	3,9%	8,0%	7,6%	6,5%	8,0%	12,1%	18,8%
1999	3,4%	3,6%	11,0%	16,7%	18,3%	13,3%	11,6%	18,2%	13,3%	15,9%	6,7%	12,3%
2000	10,0%	12,3%	12,5%	11,0%	19,7%	20,7%	12,6%	17,3%	18,7%	22,4%	5,3%	10,6%
2001	15,5%	13,3%	21,3%	23,2%	30,3%	24,5%	24,3%	26,4%	25,7%	29,4%	17,2%	23,4%
2002	18,5%	17,1%	12,9%	38,2%	33,9%	30,1%	40,4%	31,3%	30,8%	41,2%	23,1%	32,4%
2003	9,3%	0,5%	5,4%	6,8%	14,1%	8,0%	9,0%	10,3%	9,7%	8,8%	11,0%	6,1%
2004	10,5%	10,2%	5,8%	10,9%	12,2%	20,0%	8,5%	7,4%	9,5%	17,8%	6,0%	3,2%
2005	8,2%	7,7%	19,2%	7,1%	11,1%	12,0%	15,9%	14,4%	12,0%	12,0%	9,5%	7,4%
2006	4,1%	10,4%	3,7%	19,2%	13,2%	8,6%	15,2%	17,5%	16,8%	14,4%	9,9%	21,6%
2007	10,4%	9,8%	7,6%	19,2%	14,1%	12,4%	22,8%	22,0%	21,6%	26,0%	23,9%	15,9%
2008	12,8%	18,2%	22,0%	4,9%	14,6%	8,5%	19,4%	18,8%	15,2%	17,7%	25,7%	11,9%

Tabela 26 Erro Percentual Absoluto de NDHP-RC, no período [Jan-1991:Dez-2008] (RNA)

	<i>Jan</i>	<i>Fev</i>	<i>Mar</i>	<i>Abr</i>	<i>Mai</i>	<i>Jun</i>	<i>Jul</i>	<i>Ago</i>	<i>Set</i>	<i>Out</i>	<i>Nov</i>	<i>Dez</i>
1991	7,1%	5,0%	2,1%	27,1%	2,3%	8,0%	5,7%	3,2%	3,0%	20,0%	2,8%	3,4%
1992	5,7%	13,3%	23,7%	10,3%	1,9%	13,2%	8,0%	7,2%	14,6%	6,0%	3,2%	13,3%
1993	1,2%	0,2%	12,7%	6,0%	12,4%	4,1%	1,0%	1,5%	4,0%	7,5%	20,0%	15,9%
1994	4,9%	0,5%	3,0%	19,6%	9,1%	3,6%	5,7%	2,9%	3,5%	14,0%	1,7%	1,8%
1995	2,2%	2,1%	7,3%	3,8%	2,6%	5,0%	0,4%	3,6%	5,3%	3,6%	8,9%	6,8%
1996	11,7%	10,0%	3,1%	3,6%	0,2%	12,3%	0,3%	0,8%	5,6%	11,2%	0,9%	11,3%
1997	3,0%	9,0%	17,3%	11,1%	7,6%	3,8%	8,8%	7,2%	2,0%	0,8%	1,5%	1,6%
1998	11,7%	7,8%	7,3%	0,2%	7,6%	18,9%	5,8%	7,0%	2,2%	2,0%	7,0%	1,2%
1999	9,5%	8,2%	3,1%	9,8%	12,0%	2,1%	3,2%	3,1%	7,9%	11,2%	10,5%	8,7%
2000	1,1%	1,4%	0,5%	2,2%	0,2%	1,7%	1,2%	1,9%	8,1%	5,1%	11,0%	2,3%
2001	2,2%	4,8%	8,2%	7,8%	3,5%	9,7%	6,8%	4,6%	8,5%	4,6%	3,5%	3,7%
2002	1,9%	5,7%	7,3%	15,1%	0,0%	5,2%	10,5%	1,0%	2,1%	10,4%	6,2%	2,4%
2003	30,2%	17,3%	21,6%	18,3%	23,3%	13,4%	23,9%	23,0%	17,5%	15,1%	1,5%	3,9%
2004	12,2%	1,6%	7,7%	0,3%	4,1%	9,2%	13,9%	3,4%	9,9%	3,3%	13,5%	2,0%
2005	2,7%	5,7%	9,0%	4,7%	4,3%	17,8%	4,2%	13,2%	1,9%	14,9%	10,3%	0,7%
2006	4,8%	1,1%	19,4%	9,9%	0,8%	11,8%	8,6%	14,0%	4,5%	0,8%	9,3%	9,0%
2007	7,7%	7,8%	8,5%	6,1%	1,6%	2,1%	1,3%	17,2%	2,2%	13,3%	10,0%	4,8%
2008	9,7%	17,2%	3,6%	13,2%	4,8%	5,5%	4,1%	17,8%	7,3%	16,7%	10,0%	4,9%

Tabela 27 Dormidas registradas no Reino Unido (NDH-RU) no período [Jan-1990:Dez-2008]

	<i>Jan</i>	<i>Fev</i>	<i>Mar</i>	<i>Abr</i>	<i>Mai</i>	<i>Jun</i>
1990	2769000,00	3634000,00	5277000,00	5516000,00	7246000,00	8328000,00
1991	3000000,00	3000000,00	5000000,00	5000000,00	8000000,00	8000000,00
1992	2000000,00	3000000,00	3000000,00	5000000,00	6000000,00	7000000,00
1993	2000000,00	2000000,00	3000000,00	5000000,00	6000000,00	6000000,00
1994	2970000,00	2970000,00	4455000,00	5198000,00	6683000,00	6683000,00
1995	2733000,00	3672000,00	5635000,00	5338000,00	7046000,00	7985000,00
1996	2997000,00	4121000,00	6648000,00	5737000,00	7610000,00	8500000,00
1997	3200000,00	4500000,00	7600000,00	6000000,00	8000000,00	8800000,00
1998	3310000,00	4480000,00	5570000,00	6390000,00	7330000,00	6900000,00
1999	3380000,00	5080000,00	6940000,00	8100000,00	7600000,00	8130000,00
2000	7710000,00	8120000,00	9690000,00	11520000,00	14980000,00	12650000,00
2001	6260000,00	7300000,00	8570000,00	9120000,00	10220000,00	11060000,00
2002	5740000,00	8160000,00	9800000,00	10980000,00	12050000,00	11140000,00
2003	6930000,00	9450000,00	9860000,00	10020000,00	12130000,00	10970000,00
2004	6870000,00	7380000,00	9250000,00	8670000,00	8480000,00	9010000,00
2005	6084417,00	8525801,00	11870431,00	9531000,00	12020000,00	11920000,00
2006	5060000,00	7200000,00	7340000,00	8640000,00	9400000,00	10050000,00
2007	6130000,00	6350000,00	8870000,00	7970000,00	10640000,00	10300000,00
2008	5590000,00	7590000,00	9050000,00	8200000,00	11170000,00	11220000,00
	<i>Jul</i>	<i>Ago</i>	<i>Set</i>	<i>Out</i>	<i>Nov</i>	<i>Dez</i>
1990	10599000,00	14905000,00	10319000,00	6228000,00	3698000,00	3482000,00
1991	13000000,00	16000000,00	8000000,00	5000000,00	3000000,00	7000000,00
1992	12000000,00	16000000,00	7000000,00	5000000,00	2000000,00	7000000,00
1993	10000000,00	12000000,00	6000000,00	4000000,00	2000000,00	6000000,00
1994	10396000,00	14109000,00	6683000,00	5198000,00	2970000,00	6683000,00
1995	10334000,00	13880000,00	9779000,00	6659000,00	4140000,00	3799000,00
1996	11192000,00	14308000,00	10280000,00	7866000,00	5080000,00	4565000,00
1997	11800000,00	14300000,00	10500000,00	9000000,00	6000000,00	5300000,00
1998	8740000,00	11600000,00	8900000,00	7700000,00	5270000,00	4870000,00
1999	10750000,00	12780000,00	9130000,00	9310000,00	6730000,00	5070000,00
2000	15980000,00	19450000,00	13180000,00	13260000,00	9070000,00	6930000,00
2001	14470000,00	18170000,00	14870000,00	14140000,00	11370000,00	8860000,00
2002	12460000,00	17680000,00	13260000,00	11760000,00	8840000,00	8700000,00
2003	12680000,00	15160000,00	13180000,00	12680000,00	6900000,00	6820000,00
2004	12200000,00	12870000,00	10080000,00	10040000,00	7020000,00	5170000,00
2005	16600000,00	13870000,00	11390000,00	11070000,00	8220000,00	6660000,00
2006	11600000,00	12920000,00	10850000,00	9620000,00	7290000,00	7810000,00
2007	12760000,00	14330000,00	10130000,00	8850000,00	7530000,00	6510000,00
2008	11040000,00	11910000,00	10360000,00	9180000,00	6650000,00	6730000,00

Tabela 28 Valores previstos para a Série NDH-RU, no período [Jan-1990:Dez-2008] (RL)

	<i>Jan</i>	<i>Fev</i>	<i>Mar</i>	<i>Abr</i>	<i>Mai</i>	<i>Jun</i>
1990	1842830,04	2094466,59	3657064,24	4417862,75	6199078,43	6795470,59
1991	2149339,64	2513487,68	4078647,61	4790335,78	6532980,39	7071963,24
1992	2455849,25	2932508,78	4500230,99	5162808,82	6866882,35	7348455,88
1993	2762358,86	3351529,88	4921814,36	5535281,86	7200784,31	7624948,53
1994	3068868,46	3770550,97	5343397,74	5907754,90	7534686,27	7901441,18
1995	3375378,07	4189572,07	5764981,11	6280227,94	7868588,24	8177933,82
1996	3681887,67	4608593,16	6186564,49	6652700,98	8202490,20	8454426,47
1997	3988397,28	5027614,26	6608147,86	7025174,02	8536392,16	8730919,12
1998	4294906,88	5446635,35	7029731,24	7397647,06	8870294,12	9007411,76
1999	4601416,49	5865656,45	7451314,61	7770120,10	9204196,08	9283904,41
2000	4907926,09	6284677,54	7872897,99	8142593,14	9538098,04	9560397,06
2001	5214435,70	6703698,64	8294481,36	8515066,18	9872000,00	9836889,71
2002	5520945,30	7122719,74	8716064,74	8887539,22	10205901,96	10113382,35
2003	5827454,91	7541740,83	9137648,11	9260012,25	10539803,92	10389875,00
2004	6133964,51	7960761,93	9559231,49	9632485,29	10873705,88	10666367,65
2005	6440474,12	8379783,02	9980814,86	10004958,33	11207607,84	10942860,29
2006	6746983,73	8798804,12	10402398,24	10377431,37	11541509,80	11219352,94
2007	7053493,33	9217825,21	10823981,61	10749904,41	11875411,76	11495845,59
2008	7360002,94	9636846,31	11245564,99	11122377,45	12209313,73	11772338,24
	<i>Jul</i>	<i>Ago</i>	<i>Set</i>	<i>Out</i>	<i>Nov</i>	<i>Dez</i>
1990	10537352,94	14663882,35	7683019,61	4755647,06	2451333,33	5118470,59
1991	10726073,53	14669147,06	7997649,51	5253330,88	2877254,90	5248654,41
1992	10914794,12	14674411,76	8312279,41	5751014,71	3303176,47	5378838,24
1993	11103514,71	14679676,47	8626909,31	6248698,53	3729098,04	5509022,06
1994	11292235,29	14684941,18	8941539,22	6746382,35	4155019,61	5639205,88
1995	11480955,88	14690205,88	9256169,12	7244066,18	4580941,18	5769389,71
1996	11669676,47	14695470,59	9570799,02	7741750,00	5006862,75	5899573,53
1997	11858397,06	14700735,29	9885428,92	8239433,82	5432784,31	6029757,35
1998	12047117,65	14706000,00	10200058,82	8737117,65	5858705,88	6159941,18
1999	12235838,24	14711264,71	10514688,73	9234801,47	6284627,45	6290125,00
2000	12424558,82	14716529,41	10829318,63	9732485,29	6710549,02	6420308,82
2001	12613279,41	14721794,12	11143948,53	10230169,12	7136470,59	6550492,65
2002	12802000,00	14727058,82	11458578,43	10727852,94	7562392,16	6680676,47
2003	12990720,59	14732323,53	11773208,33	11225536,76	7988313,73	6810860,29
2004	13179441,18	14737588,24	12087838,24	11723220,59	8414235,29	6941044,12
2005	13368161,76	14742852,94	12402468,14	12220904,41	8840156,86	7071227,94
2006	13556882,35	14748117,65	12717098,04	12718588,24	9266078,43	7201411,76
2007	13745602,94	14753382,35	13031727,94	13216272,06	9692000,00	7331595,59
2008	13934323,53	14758647,06	13346357,84	13713955,88	10117921,57	7461779,41

Tabela 29 Valores previstos para a Série NDH-RU, no período [Jan-1991:Dez-2008] (RNA)

	<i>Jan</i>	<i>Fev</i>	<i>Mar</i>	<i>Abr</i>	<i>Mai</i>	<i>Jun</i>
1991	2963688,81	3348314,58	5386355,79	6792670,57	8171761,49	9392341,94
1992	2524427,86	4472734,39	5680438,68	6666792,62	8275510,13	7586513,85
1993	2043604,65	4412959,31	4937514,19	6368623,16	7757264,66	6817679,20
1994	2355869,94	4233721,51	4872771,61	6444208,06	7187714,71	6139249,75
1995	2651764,52	4411774,04	5340184,61	6875952,32	7878745,10	6825454,00
1996	3140816,77	3557117,55	5479412,59	7205629,01	8336564,33	8906283,97
1997	3484999,26	4011246,47	5854175,61	7713146,20	8752837,71	9352766,95
1998	3835595,75	4449329,20	6093854,37	7028870,60	8728747,21	9368930,53
1999	4054466,18	4430229,50	5296301,19	7165390,80	8796908,52	8292215,78
2000	4345505,21	5584489,10	8134792,93	9365865,45	10556935,79	11938473,92
2001	6582154,72	6889460,30	9359034,05	11070697,56	13047786,62	12512764,40
2002	5947035,97	5767344,91	8174704,03	10067038,69	11799065,73	12222379,68
2003	6786038,28	6822602,32	9744067,37	11136165,15	12362831,73	11838174,34
2004	6192989,62	7558784,13	8651205,82	10226089,35	11952044,57	11043112,47
2005	5305001,16	6200594,82	8789183,41	10601637,09	10609356,28	11441764,39
2006	5842506,10	7527072,83	9480223,68	9749754,24	11167920,30	11821537,03
2007	5881343,16	6558196,06	7238024,04	9204591,90	9960635,72	10672131,29
2008	5674162,45	6149038,57	8424953,32	9724141,52	10681849,42	11164254,15
	<i>Jul</i>	<i>Ago</i>	<i>Set</i>	<i>Out</i>	<i>Nov</i>	<i>Dez</i>
1991	11211344,85	15138751,29	13356724,99	6710043,62	3819812,46	3930260,79
1992	12031016,67	15341914,69	12675713,37	5687275,88	2808092,43	4501089,65
1993	11071470,47	14364858,79	9727795,02	5226346,92	2864326,49	4363949,71
1994	9135814,22	10906269,06	9512426,92	5537776,46	3315160,77	4775948,37
1995	10643685,46	12696769,03	9502643,67	6895457,68	4183643,46	5039318,92
1996	11003993,17	13160668,49	10894189,18	7949627,30	5191758,83	4809390,33
1997	11841830,30	13908230,30	11139116,27	8780723,44	5941912,67	5184794,38
1998	11268187,44	12265086,63	10410849,99	8108561,68	5868666,98	5535976,09
1999	9450605,49	10909838,94	9487123,12	7394088,21	6449142,89	5891565,34
2000	12877958,29	14010603,59	12795940,92	11468404,80	7939092,43	6701639,77
2001	16394433,75	18342854,06	15590759,84	14242475,41	10767222,42	8523996,48
2002	15703648,25	17202315,70	16145836,76	13339697,02	10394556,70	8123728,77
2003	13291444,42	15585253,36	13521392,20	11783771,43	10019275,92	7456891,89
2004	11928929,90	13738316,90	12354847,43	11054254,85	8040785,26	6769481,76
2005	13982709,42	13976820,62	10602699,63	10005606,69	7927289,27	6646212,41
2006	16055390,70	13231313,95	11789605,67	10287459,06	8036308,45	6310614,61
2007	11855093,69	13109164,76	11961629,06	9674086,60	6755069,99	6852208,44
2008	13483157,37	13100625,03	9911887,47	8926479,27	7276810,45	6345410,11

Tabela 30 Erro Percentual Absoluto de NDH-RU, no período [Jan-1990:Dez-2008] (RL)

	<i>Jan</i>	<i>Fev</i>	<i>Mar</i>	<i>Abr</i>	<i>Mai</i>	<i>Jun</i>	<i>Jul</i>	<i>Ago</i>	<i>Set</i>	<i>Out</i>	<i>Nov</i>	<i>Dez</i>
1990	33,4%	42,4%	30,7%	19,9%	14,4%	18,4%	0,6%	1,6%	25,5%	23,6%	33,7%	47,0%
1991	28,4%	16,2%	18,4%	4,2%	18,3%	11,6%	17,5%	8,3%	0,0%	5,1%	4,1%	25,0%
1992	22,8%	2,2%	50,0%	3,3%	14,4%	5,0%	9,0%	8,3%	18,7%	15,0%	65,2%	23,2%
1993	38,1%	67,6%	64,1%	10,7%	20,0%	27,1%	11,0%	22,3%	43,8%	56,2%	86,5%	8,2%
1994	3,3%	27,0%	19,9%	13,7%	12,7%	18,2%	8,6%	4,1%	33,8%	29,8%	39,9%	15,6%
1995	23,5%	14,1%	2,3%	17,7%	11,7%	2,4%	11,1%	5,8%	5,3%	8,8%	10,7%	51,9%
1996	22,9%	11,8%	6,9%	16,0%	7,8%	0,5%	4,3%	2,7%	6,9%	1,6%	1,4%	29,2%
1997	24,6%	11,7%	13,1%	17,1%	6,7%	0,8%	0,5%	2,8%	5,9%	8,5%	9,5%	13,8%
1998	29,8%	21,6%	26,2%	15,8%	21,0%	30,5%	37,8%	26,8%	14,6%	13,5%	11,2%	26,5%
1999	36,1%	15,5%	7,4%	4,1%	21,1%	14,2%	13,8%	15,1%	15,2%	0,8%	6,6%	24,1%
2000	36,3%	22,6%	18,8%	29,3%	36,3%	24,4%	22,2%	24,3%	17,8%	26,6%	26,0%	7,4%
2001	16,7%	8,2%	3,2%	6,6%	3,4%	11,1%	12,8%	19,0%	25,1%	27,7%	37,2%	26,1%
2002	3,8%	12,7%	11,1%	19,1%	15,3%	9,2%	2,7%	16,7%	13,6%	8,8%	14,5%	23,2%
2003	15,9%	20,2%	7,3%	7,6%	13,1%	5,3%	2,5%	2,8%	10,7%	11,5%	15,8%	0,1%
2004	10,7%	7,9%	3,3%	11,1%	28,2%	18,4%	8,0%	14,5%	19,9%	16,8%	19,9%	34,3%
2005	5,9%	1,7%	15,9%	5,0%	6,8%	8,2%	19,5%	6,3%	8,9%	10,4%	7,5%	6,2%
2006	33,3%	22,2%	41,7%	20,1%	22,8%	11,6%	16,9%	14,1%	17,2%	32,2%	27,1%	7,8%
2007	15,1%	45,2%	22,0%	34,9%	11,6%	11,6%	7,7%	3,0%	28,6%	49,3%	28,7%	12,6%
2008	31,7%	27,0%	24,3%	35,6%	9,3%	4,9%	26,2%	23,9%	28,8%	49,4%	52,1%	10,9%

Tabela 31 Erro Percentual Absoluto de NDH-RU, no período [Jan-1991:Dez-2008] (RNA)

	<i>Jan</i>	<i>Fev</i>	<i>Mar</i>	<i>Abr</i>	<i>Mai</i>	<i>Jun</i>	<i>Jul</i>	<i>Ago</i>	<i>Set</i>	<i>Out</i>	<i>Nov</i>	<i>Dez</i>
1991	1,2%	11,6%	7,7%	35,9%	2,1%	17,4%	13,8%	5,4%	67,0%	34,2%	27,3%	43,9%
1992	26,2%	49,1%	89,3%	33,3%	37,9%	8,4%	0,3%	4,1%	81,1%	13,7%	40,4%	35,7%
1993	2,2%	120,6%	64,6%	27,4%	29,3%	13,6%	10,7%	19,7%	62,1%	30,7%	43,2%	27,3%
1994	20,7%	42,5%	9,4%	24,0%	7,6%	8,1%	12,1%	22,7%	42,3%	6,5%	11,6%	28,5%
1995	3,0%	20,1%	5,2%	28,8%	11,8%	14,5%	3,0%	8,5%	2,8%	3,6%	1,1%	32,6%
1996	4,8%	13,7%	17,6%	25,6%	9,5%	4,8%	1,7%	8,0%	6,0%	1,1%	2,2%	5,4%
1997	8,9%	10,9%	23,0%	28,6%	9,4%	6,3%	0,4%	2,7%	6,1%	2,4%	1,0%	2,2%
1998	15,9%	0,7%	9,4%	10,0%	19,1%	35,8%	28,9%	5,7%	17,0%	5,3%	11,4%	13,7%
1999	20,0%	12,8%	23,7%	11,5%	15,7%	2,0%	12,1%	14,6%	3,9%	20,6%	4,2%	16,2%
2000	43,6%	31,2%	16,0%	18,7%	29,5%	5,6%	19,4%	28,0%	2,9%	13,5%	12,5%	3,3%
2001	5,1%	5,6%	9,2%	21,4%	27,7%	13,1%	13,3%	1,0%	4,8%	0,7%	5,3%	3,8%
2002	3,6%	29,3%	16,6%	8,3%	2,1%	9,7%	26,0%	2,7%	21,8%	13,4%	17,6%	6,6%
2003	2,1%	27,8%	1,2%	11,1%	1,9%	7,9%	4,8%	2,8%	2,6%	7,1%	45,2%	9,3%
2004	9,9%	2,4%	6,5%	17,9%	40,9%	22,6%	2,2%	6,7%	22,6%	10,1%	14,5%	30,9%
2005	12,8%	27,3%	26,0%	11,2%	11,7%	4,0%	15,8%	0,8%	6,9%	9,6%	3,6%	0,2%
2006	15,5%	4,5%	29,2%	12,8%	18,8%	17,6%	38,4%	2,4%	8,7%	6,9%	10,2%	19,2%
2007	4,1%	3,3%	18,4%	15,5%	6,4%	3,6%	7,1%	8,5%	18,1%	9,3%	10,3%	5,3%
2008	1,5%	19,0%	6,9%	18,6%	4,4%	0,5%	22,1%	10,0%	4,3%	2,8%	9,4%	5,7%

Tabela 32 Indicadores da Actividade Turística

	<i>1994</i>	<i>1995</i>	<i>1996</i>	<i>1997</i>	<i>1998</i>	<i>1999</i>	<i>2000</i>	<i>2001</i>
N.º de Empregados	200	204	226	216	248	256	253	264
N.º de Est. Hoteleiros	1952	1956	1969	2002	1987	2016	2049	2051
Proveitos (€)	:	:	:	:	:	:	:	:
	<i>2002</i>	<i>2003</i>	<i>2004</i>	<i>2005</i>	<i>2006</i>	<i>2007</i>	<i>2008</i>	
N.º de Empregados	262	257	266	285	276	301	324	
N.º de Est. Hoteleiros	2171	2214	2239	2300	2324	2339	:	
Proveitos (€)	1466066	1479130	1560947	1591273	1741455	1943590	1964602	

Tabela 33 Índice de Preços no Consumidor em Portugal (2005=100) no intervalo [Jan-1996:Dez-2008]

	<i>Jan</i>	<i>Fev</i>	<i>Mar</i>	<i>Abr</i>	<i>Mai</i>	<i>Jun</i>	<i>Jul</i>	<i>Ago</i>	<i>Set</i>	<i>Out</i>	<i>Nov</i>	<i>Dez</i>
1996	70,49	71,14	71,43	71,94	71,79	72,52	72,52	72,73	72,73	72,81	73,75	72,81
1997	72,66	73,02	73,38	73,60	73,89	74,11	73,96	74,32	74,54	74,61	74,47	74,47
1998	74,76	74,90	75,05	75,41	76,13	76,35	76,63	76,78	77,07	77,00	76,78	76,78
1999	77,00	77,28	77,50	77,72	77,93	78,01	78,30	78,58	78,80	79,09	79,16	79,02
2000	79,16	79,45	79,52	79,88	80,75	81,33	82,20	82,41	82,70	82,41	81,76	81,69
2001	82,27	82,92	82,92	83,64	84,36	84,72	85,52	85,59	85,88	86,46	86,10	85,66
2002	86,02	86,46	87,18	87,97	88,99	89,49	90,29	91,01	91,01	91,87	91,15	91,30
2003	91,51	91,95	91,95	92,74	93,82	94,26	95,27	95,70	95,99	95,20	94,33	94,40
2004	95,20	95,70	95,85	96,79	98,16	104,73	100,54	100,11	99,96	99,53	98,30	98,09
2005	97,72	98,09	98,30	98,81	100,11	100,40	101,48	102,35	101,63	101,12	99,96	100,04
2006	100,07	100,14	100,35	101,07	102,07	102,06	102,79	103,35	103,35	103,46	102,74	102,42
2007	102,78	103,08	103,25	103,79	104,24	104,46	104,83	105,28	105,22	105,36	105,31	105,35
2008	106,31	106,55	107,00	107,22	107,82	108,23	108,57	108,97	109,00	109,74	109,24	109,43

Tabela 34 Índice de Preços no Consumidor no Reino Unido (2005=100) no intervalo [Jan-1996:Dez-2008]

	<i>Jan</i>	<i>Fev</i>	<i>Mar</i>	<i>Abr</i>	<i>Mai</i>	<i>Jun</i>	<i>Jul</i>	<i>Ago</i>	<i>Set</i>	<i>Out</i>	<i>Nov</i>	<i>Dez</i>
1996	86,8	87,2	87,5	88,0	88,3	88,4	87,8	88,3	88,7	88,7	88,7	89,0
1997	88,6	88,8	89,0	89,4	89,6	89,8	89,5	90,0	90,3	90,3	90,4	90,5
1998	89,9	90,3	90,5	91,0	91,5	91,3	90,8	91,2	91,6	91,6	91,7	91,9
1999	91,4	91,5	92,0	92,4	92,7	92,6	92,0	92,3	92,7	92,6	92,7	93,0
2000	92,1	92,4	92,6	92,9	93,2	93,3	92,8	92,8	93,6	93,5	93,7	93,7
2001	92,9	93,1	93,4	94,0	94,7	94,9	94,2	94,5	94,8	94,7	94,5	94,7
2002	94,4	94,5	94,9	95,3	95,5	95,5	95,2	95,5	95,7	95,9	95,9	96,3
2003	95,7	96,0	96,3	96,7	96,7	96,5	96,5	96,8	97,1	97,2	97,2	97,5
2004	97,0	97,2	97,4	97,8	98,1	98,1	97,8	98,1	98,2	98,4	98,6	99,1
2005	98,6	98,8	99,3	99,7	100,0	100,0	100,1	100,4	100,6	100,7	100,7	101,0
2006	100,5	100,9	101,1	101,7	102,2	102,5	102,5	102,9	103,0	103,2	103,4	104,0
2007	103,2	103,7	104,2	104,5	104,8	105,0	104,4	104,7	104,8	105,3	105,6	106,2
2008	105,5	106,3	106,7	107,6	108,3	109,0	109,0	109,7	110,3	110,0	109,9	109,5

Tabela 35 N.º de desempregados em Portugal (Milhares) no intervalo [Jan-1990:Dez-2008]

	<i>Jan</i>	<i>Fev</i>	<i>Mar</i>	<i>Abr</i>	<i>Mai</i>	<i>Jun</i>	<i>Jul</i>	<i>Ago</i>	<i>Set</i>	<i>Out</i>	<i>Nov</i>	<i>Dez</i>
1990	246	242	226	208	204	212	217	222	226	227	227	220
1991	213	205	194	184	184	191	199	205	207	211	208	196
1992	191	192	189	185	181	182	186	193	198	204	212	217
1993	227	236	243	247	245	247	253	260	266	279	289	297
1994	312	318	318	316	315	315	319	321	328	337	343	341
1995	352	353	348	340	329	322	322	323	329	342	350	350
1996	356	358	355	352	343	336	334	336	335	339	342	341
1997	350	353	339	322	308	303	318	323	324	321	317	309
1998	311	300	276	246	227	219	229	236	245	248	245	241
1999	245	242	239	235	230	224	221	215	213	213	215	217
2000	231	230	219	200	190	190	205	208	206	196	194	195
2001	215	223	221	207	200	201	207	210	217	219	222	220
2002	232	235	240	241	242	246	257	270	295	318	335	336
2003	345	345	345	341	333	327	329	330	346	356	359	352
2004	349	345	347	346	346	351	365	375	389	387	390	392
2005	410	416	412	401	396	400	414	428	447	451	451	440
2006	439	430	420	412	405	400	405	414	433	449	462	465
2007	473	471	466	451	436	434	440	445	448	445	441	432
2008	434	429	418	409	407	414	426	437	438	434	435	444

Tabela 36 N.º de desempregados no Reino Unido (Milhares) no intervalo [Jan-1990:Dez-2008]

	<i>Jan</i>	<i>Fev</i>	<i>Mar</i>	<i>Abr</i>	<i>Mai</i>	<i>Jun</i>	<i>Jul</i>	<i>Ago</i>	<i>Set</i>	<i>Out</i>	<i>Nov</i>	<i>Dez</i>
1990	2011	2011	1994	1970	1906	1864	1940	1960	1957	1930	1973	2081
1991	2187	2256	2357	2419	2432	2430	2549	2597	2566	2500	2500	2542
1992	2632	2624	2596	2623	2618	2708	2770	2780	2765	2786	2782	2834
1993	2917	2938	2909	2821	2828	2884	2907	2899	2844	2807	2767	2763
1994	2755	2752	2697	2622	2638	2665	2695	2657	2569	2489	2410	2380
1995	2411	2449	2427	2358	2359	2413	2450	2469	2402	2356	2241	2256
1996	2264	2274	2279	2227	2248	2269	2289	2298	2241	2189	2106	2056
1997	2031	2022	1981	1931	1959	2034	2026	1966	1873	1829	1753	1715
1998	1731	1766	1744	1689	1713	1762	1822	1809	1758	1722	1664	1697
1999	1717	1739	1723	1667	1663	1714	1766	1768	1691	1674	1609	1617
2000	1611	1643	1611	1552	1521	1541	1584	1608	1582	1523	1449	1419
2001	1438	1439	1390	1320	1360	1463	1540	1551	1512	1489	1460	1446
2002	1442	1483	1492	1469	1450	1509	1584	1618	1575	1529	1462	1419
2003	1461	1507	1468	1397	1382	1500	1557	1563	1505	1460	1405	1370
2004	1373	1400	1386	1347	1355	1416	1453	1457	1438	1416	1377	1369
2005	1400	1399	1368	1350	1362	1407	1456	1480	1543	1550	1520	1490
2006	1532	1582	1594	1579	1611	1676	1742	1755	1734	1663	1625	1608
2007	1665	1692	1659	1576	1580	1623	1696	1709	1671	1598	1507	1497
2008	1553	1588	1603	1528	1600	1702	1848	1901	1919	1924	1914	1951

Tabela 37 N.º de desempregados em Espanha (Milhares) no intervalo [Jan-1990:Dez-2008]

	<i>Jan</i>	<i>Fev</i>	<i>Mar</i>	<i>Abr</i>	<i>Mai</i>	<i>Jun</i>	<i>Jul</i>	<i>Ago</i>	<i>Set</i>	<i>Out</i>	<i>Nov</i>	<i>Dez</i>
1990	2108	2102	2073	2059	2025	2007	1983	1996	2005	2028	2014	2013
1991	2013	2018	2006	1990	1968	1993	2023	2052	2101	2139	2124	2126
1992	2154	2179	2209	2223	2208	2240	2262	2293	2382	2470	2529	2584
1993	2671	2735	2808	2791	2785	2877	2921	2925	2994	3040	3055	3064
1994	3136	3149	3144	3132	3099	3104	3055	3045	3077	3068	3067	3027
1995	3019	3022	2988	2924	2902	2923	2917	2933	2990	2971	2976	2909
1996	2972	2999	2996	2939	2903	2921	2899	2888	2949	2937	2918	2829
1997	2859	2869	2836	2804	2774	2794	2746	2738	2796	2778	2752	2675
1998	2661	2636	2612	2567	2536	2549	2509	2525	2541	2523	2482	2394
1999	2364	2299	2236	2155	2100	2111	2099	2131	2142	2110	2107	2056
2000	2122	2102	2057	1992	1943	1932	1931	1942	1952	1956	1945	1882
2001	1901	1940	1915	1846	1852	1819	1816	1863	1864	1883	1912	1911
2002	2036	2138	2079	2052	2015	2027	2081	2126	2144	2145	2156	2141
2003	2225	2234	2223	2155	2094	2122	2133	2165	2182	2185	2185	2180
2004	2230	2219	2191	2199	2140	2132	2096	2128	2123	2101	2108	2058
2005	2112	2101	2084	2045	1935	1855	1779	1761	1755	1808	1855	1861
2006	1928	1939	1941	1895	1833	1783	1756	1774	1765	1796	1824	1812
2007	1868	1860	1840	1797	1741	1741	1742	1808	1825	1871	1935	1977
2008	2112	2190	2220	2323	2381	2440	2472	2593	2731	2976	3212	3436

Tabela 38 N.º de desempregados em França (Milhares) no intervalo [Jan-1990:Dez-2008]

	<i>Jan</i>	<i>Fev</i>	<i>Mar</i>	<i>Abr</i>	<i>Mai</i>	<i>Jun</i>	<i>Jul</i>	<i>Ago</i>	<i>Set</i>	<i>Out</i>	<i>Nov</i>	<i>Dez</i>
1990	2185	2157	2119	2062	1992	1957	1975	2033	2096	2140	2146	2149
1991	2156	2150	2137	2107	2080	2077	2132	2225	2305	2363	2387	2406
1992	2429	2426	2389	2347	2313	2296	2339	2421	2503	2574	2615	2656
1993	2690	2695	2684	2658	2625	2617	2679	2779	2881	2970	3006	3027
1994	3047	3033	2991	2925	2855	2802	2809	2867	2929	2963	2951	2941
1995	2924	2872	2810	2758	2692	2602	2587	2691	2788	2851	2916	2955
1996	3004	3034	2974	2877	2804	2762	2787	2896	2998	3050	3080	3082
1997	3083	3060	2979	2880	2807	2773	2791	2880	2967	3016	3035	3010
1998	2990	2964	2877	2770	2690	2648	2674	2785	2873	2905	2936	2928
1999	2919	2904	2827	2715	2623	2574	2578	2663	2704	2687	2689	2651
2000	2622	2586	2482	2362	2271	2214	2243	2354	2389	2378	2375	2347
2001	2334	2305	2228	2139	2075	2037	2089	2211	2275	2310	2354	2357
2002	2347	2325	2264	2203	2161	2147	2221	2361	2439	2475	2522	2546
2003	2567	2560	2493	2416	2353	2292	2321	2391	2467	2565	2625	2666
2004	2709	2680	2596	2494	2441	2386	2437	2590	2597	2658	2705	2658
2005	2663	2604	2536	2484	2474	2429	2495	2663	2655	2706	2742	2729
2006	2815	2790	2699	2592	2522	2463	2524	2676	2584	2566	2542	2519
2007	2627	2605	2502	2385	2314	2231	2288	2411	2347	2337	2302	2233
2008	2277	2220	2176	2143	2111	2072	2110	2255	2250	2324	2414	2457

Tabela 39 Valores previstos para a Série NDHP-RC (VA=IPC), no período [Jan-1991:Dez-2008] (RNA)

	<i>Jan</i>	<i>Fev</i>	<i>Mar</i>	<i>Abr</i>	<i>Mai</i>	<i>Jun</i>
1991	68815,93	69645,50	89818,18	121886,38	132916,58	136382,15
1992	66595,60	78737,07	96937,31	113148,24	142378,12	138203,87
1993	71008,40	75419,14	90384,29	122432,38	139882,61	122102,40
1994	72692,75	73670,90	90182,10	120136,57	132306,88	122761,15
1995	69074,52	74415,90	90621,22	115546,00	136470,61	125390,32
1996	72414,83	76255,61	107337,48	134551,00	133026,02	132589,36
1997	72345,96	84578,86	113335,48	146029,26	137961,35	141378,67
1998	78690,44	99729,51	126702,38	140415,57	155196,62	151592,69
1999	88867,05	106905,56	132750,12	155979,05	173621,34	169931,19
2000	101766,91	115622,16	138180,17	162129,94	173263,20	171725,59
2001	101226,22	112762,09	151210,93	177852,63	176943,42	173145,45
2002	103631,36	119302,27	152292,38	180379,17	176996,27	178522,52
2003	108088,03	141747,82	172741,44	185513,75	214744,01	214894,78
2004	151444,51	174252,28	220460,20	263464,84	289421,54	286216,90
2005	161841,89	189452,44	233479,99	276816,81	302863,40	320552,37
2006	162916,03	199300,67	256818,73	275413,71	306689,11	314246,60
2007	178447,20	202040,62	246744,13	302648,34	311022,60	315549,02
2008	180315,87	210805,79	264454,44	305805,23	305952,74	319481,97
	<i>Jul</i>	<i>Ago</i>	<i>Set</i>	<i>Out</i>	<i>Nov</i>	<i>Dez</i>
1991	194499,42	237396,04	196541,02	119980,05	89770,83	82005,72
1992	184567,96	245044,22	191464,11	114929,67	86689,46	76519,66
1993	173184,56	224264,30	178174,77	115358,26	85098,40	75436,15
1994	173516,71	225530,33	181008,32	110630,82	79449,71	75027,24
1995	177502,02	227445,88	176238,82	109792,58	81979,58	78516,72
1996	178768,11	234047,00	182656,75	123769,04	94724,04	89505,04
1997	180179,84	235945,65	204399,67	139352,55	105828,68	93346,61
1998	201925,67	260845,01	217301,20	158383,62	116233,85	100330,55
1999	218278,68	294865,73	233812,12	168452,74	117744,56	103264,78
2000	217487,87	285874,69	241884,83	173499,16	127499,04	114899,27
2001	229855,41	301623,11	242266,40	178114,67	136994,57	121056,19
2002	226546,61	294560,36	244815,62	182703,14	137184,45	125687,94
2003	229669,06	319196,43	263576,26	196289,01	171251,03	159738,35
2004	330244,46	467209,38	352562,29	284077,94	202335,60	190088,40
2005	336160,84	470921,03	362439,11	311470,32	204402,39	197353,67
2006	358845,75	506938,28	377409,02	318195,56	211044,13	205417,60
2007	362880,15	519832,91	399471,32	348123,54	220506,65	237274,91
2008	382333,62	547085,55	424555,77	376362,68	265350,76	239387,08

Tabela 40 Percentual Absoluto de NDHP-RC, no período [Jan-1991:Dez-2008] (RNA-IPC)

	<i>Jan</i>	<i>Fev</i>	<i>Mar</i>	<i>Abr</i>	<i>Mai</i>	<i>Jun</i>	<i>Jul</i>	<i>Ago</i>	<i>Set</i>	<i>Out</i>	<i>Nov</i>	<i>Dez</i>
1991	3,7%	11,7%	2,2%	0,7%	5,8%	1,3%	6,2%	8,7%	3,4%	6,2%	7,3%	0,4%
1992	1,6%	3,9%	15,7%	0,3%	9,3%	1,8%	5,0%	8,5%	1,2%	6,9%	1,2%	0,2%
1993	1,4%	4,4%	8,0%	1,6%	1,0%	0,9%	6,1%	9,3%	1,3%	4,9%	0,3%	3,5%
1994	1,0%	10,7%	3,2%	4,2%	3,5%	0,1%	9,3%	8,9%	2,3%	7,0%	9,2%	3,9%
1995	0,1%	7,5%	10,8%	1,6%	8,6%	0,2%	6,5%	9,2%	0,8%	6,1%	1,2%	3,9%
1996	2,3%	6,2%	2,5%	3,4%	0,3%	1,6%	8,5%	3,5%	6,2%	6,5%	17,1%	7,9%
1997	4,6%	5,4%	2,4%	0,2%	3,0%	0,2%	8,4%	2,5%	9,0%	1,0%	5,5%	0,3%
1998	5,2%	4,5%	8,0%	3,1%	2,9%	5,3%	8,7%	0,7%	12,4%	7,5%	7,8%	0,0%
1999	6,5%	0,1%	8,7%	2,6%	1,6%	2,3%	9,0%	0,3%	7,8%	3,6%	7,6%	7,0%
2000	3,7%	6,7%	1,4%	4,2%	4,5%	1,5%	6,8%	1,8%	13,0%	7,2%	4,1%	5,9%
2001	3,6%	0,5%	6,5%	2,7%	1,8%	3,1%	8,6%	1,8%	13,2%	9,3%	4,0%	3,6%
2002	5,9%	1,9%	9,7%	9,6%	5,0%	4,0%	13,1%	2,6%	15,9%	11,9%	9,4%	13,6%
2003	8,2%	19,3%	10,2%	20,1%	24,3%	24,4%	24,0%	10,7%	24,5%	24,2%	36,0%	47,0%
2004	10,8%	18,9%	12,3%	9,6%	2,3%	11,7%	10,9%	6,4%	10,3%	10,2%	10,3%	18,1%
2005	12,0%	11,8%	13,2%	5,0%	4,4%	3,9%	9,1%	6,4%	9,3%	3,6%	12,2%	20,5%
2006	11,0%	15,9%	2,3%	4,2%	2,3%	7,3%	3,1%	2,1%	6,6%	8,2%	6,5%	14,6%
2007	22,0%	9,5%	12,3%	4,6%	3,0%	7,3%	1,4%	2,7%	2,9%	10,1%	5,9%	7,4%
2008	10,9%	10,6%	10,9%	7,4%	9,4%	0,0%	6,2%	6,5%	0,5%	1,4%	3,3%	11,6%