

# ESTUDOS I



FACULDADE de ECONOMIA da UNIVERSIDADE do ALGARVE

# ESTUDOS I

---

**Cidadania, Instituição e Património**

**Economia e Desenvolvimento Regional**

**Finanças e Contabilidade**

**Gestão e Apoio à Decisão**

**Modelos Aplicados à Economia e à Gestão**

**A Faculdade de Economia da Universidade do Algarve**



Faculdade de Economia da Universidade do Algarve

2004

## COMISSÃO EDITORIAL

António Covas  
Carlos Cândido  
Duarte Trigueiros  
Efigénio da Luz Rebelo  
João Albino da Silva  
João Guerreiro  
Paulo M.M. Rodrigues  
Rui Nunes

---

## FICHA TÉCNICA

### **Faculdade de Economia da Universidade do Algarve**

Campus de Gambelas, 8005-139 Faro  
Tel. 289817571 Fax. 289815937  
E-mail: ccfeua@ualg.pt  
Website: www.ualg.pt/feua

### ***Título***

Estudos I - Faculdade de Economia da Universidade do Algarve

### ***Autor***

Vários

### ***Editor***

Faculdade de Economia da Universidade do Algarve  
Morada: Campus de Gambelas  
Localidade: FARO  
Código Postal: 8005-139

### ***Compilação e Design Gráfico***

Susy A. Rodrigues

### ***Revisão de Formatação e Paginação***

Lídia Rodrigues

### ***Fotolitos e Impressão***

Serviços Gráficos da Universidade do Algarve

### ***ISBN***

972-99397-0-5 - Data: 26.10.2004

### ***Depósito Legal***

218279/04

### ***Tiragem***

500 exemplares

### ***Data***

Novembro 2004

**RESERVADOS TODOS OS DIREITOS  
REPRODUÇÃO PROIBIDA**

# Combinação de previsões com modelos autoregressivos lineares e não lineares – aplicações ao turismo <sup>1</sup>

**Paulo M.M. Rodrigues**

*Faculdade de Economia, Universidade do Algarve*

**Pedro Gouveia <sup>2</sup>**

*Escola Superior de Gestão, Hotelaria e Turismo, Universidade do Algarve*

## Resumo

As metodologias de combinação de previsões exploram relações de complementaridade entre diversos tipos de modelos econométricos. O seu desenvolvimento e crescente utilização resultam do facto destas apresentarem melhor performance, em termos de previsão, do que os modelos individuais que as compõem (vide, *inter alia*, Bates e Granger, 1969 e Stock e Watson, 2003). Este artigo é pioneiro na utilização de metodologias de combinação de previsões a séries do turismo. Neste artigo são utilizados modelos lineares (quer modelos só com componente sazonal determinística, quer modelos sazonais autoregressivos) e modelos mais complexos e de evolução recente, captadores da não linearidade determinística (como é o caso dos modelos Periódicos Autoregressivos [PAR]) ou da não-linearidade estocástica (como é o caso dos modelos *Self Exciting Threshold Autoregressive* [SETAR]) das séries. Procura-se ainda, pelo recurso a modelos com componente sazonal, contribuir para a ruptura da prática dominante de utilização de séries sazonalmente ajustadas (ou seja, de séries expurgadas da informação sazonal). Por outro lado, através de uma metodologia de selecção da ordem da parte autoregressiva sugerida em Rodrigues e Gouveia (2004) procura-se conjugar o princípio da parcimónia com a salvaguarda de ausência de autocorrelação, por forma a que cada um dos modelos utilizados apresente boa performance em termos de previsão.

**Palavras-chave:** Combinação de previsões, modelos periódicos, modelos *threshold*, não-linearidade, sazonalidade.

## Abstract

Forecast combination methodologies explore complementary relations between different types of econometric models. The development and growing use of forecast combinations result from the fact that these present better forecast performance than the individual models on which the combinations are based (see, *inter alia*, Bates and Granger, 1969 and Stock and Watson, 2003). This paper looks to the application of forecast combinations to tourism data. Linear models (models with only deterministic seasonal component and seasonal autoregressive models) and more complex models, such as models that capture deterministic and/or stochastic nonlinearity of the series, are employed. We hope this study will motivate a break in the typical use of seasonally

---

<sup>1</sup> Os autores agradecem à Fundação para a Ciência e Tecnologia o apoio financeiro concedido no âmbito do programa POCTI/ECO/49266/2002 (FEDER).

<sup>2</sup> Doutorando em Métodos Quantitativos, especialidade de Econometria, na Faculdade de Economia da Universidade do Algarve.

adjusted data. On the other hand, using the lag selection approach suggested by Rodrigues and Gouveia (2004), the application of a principle of parsimony with adequately specified models is ideal, in order to ensure that the models used display a good forecast performance.

**Keywords:** Combining forecasts, periodic models, threshold models, non-linearity, seasonality.

## 1. Introdução

Os modelos de séries temporais envolvem a análise estatística da evolução passada da variável de interesse, partindo do princípio que as observações passadas incorporam informação relativa a um vasto conjunto de variáveis que influenciam o fenómeno em estudo (e, em muitos casos, difícil de identificar). Cada tipo de modelo é captor apenas de uma parte da realidade.

A interdependência entre séries temporais torna a utilização de modelos univariados (dominante ao nível da modelação e previsão em séries temporais) por vezes redutora. Uma forma de contornar este problema pode consistir em explorar as relações de complementaridade existentes entre diferentes modelos e assim obter uma melhor aproximação ao processo gerador de dados. No campo da previsão, desde o estudo de Bates e Granger (1969), têm sido desenvolvidas diversas metodologias de combinação de previsões. De um modo geral, tem sido possível encontrar evidência de que as previsões obtidas, mesmo com base em metodologias simples de combinação de previsões, apresentam vantagens em relação às geradas por modelos isolados (*vide*, Stock e Watson, 2003).

Neste artigo procuram-se utilizar diferentes modelos e diversas metodologias de combinação de previsões no campo da previsão de séries do turismo. Nesse sentido, são utilizados modelos lineares simples apenas com componente determinística (*dummies* sazonais e/ou tendência), modelos Sazonais Autoregressivos [SAR] e modelos não-lineares mais complexos e de evolução mais recente na literatura econométrica, como é o caso dos modelos Periódicos Autorregressivos (PAR) e dos modelos *Self Exciting Threshold Autoregressive* (SETAR).

Procura-se, também, contribuir para a ruptura com a prática dominante de utilização de séries sazonalmente ajustadas. Com efeito, boa parte das aplicações na área do turismo baseiam-se em dados sazonalmente ajustados de frequência anual (*vide*, por exemplo, Kulendran e Witt, 2001, p. 294). Durante muito tempo a sazonalidade foi considerada uma componente perturbadora das séries temporais que dificultava a análise das componentes ciclo e tendência. Neste contexto, era privilegiado o recurso a séries sazonalmente ajustadas, expurgadas da componente sazonal. Actualmente é reconhecido, que a sazonalidade é uma componente importante das séries temporais, que deve ser tida em conta ao nível da modelação e não ignorada ou removida como foi corrente nas últimas décadas (*vide* Ghysels, 1994 e Ghysels, Osborn e Rodrigues, 2000).

Este artigo encontra-se estruturado da seguinte forma. Na Secção 2 é identificada a série de dados que sustenta o estudo de previsão e analisadas as suas propriedades. Na Secção 3 procede-se a uma breve descrição de cada um dos tipos de modelos

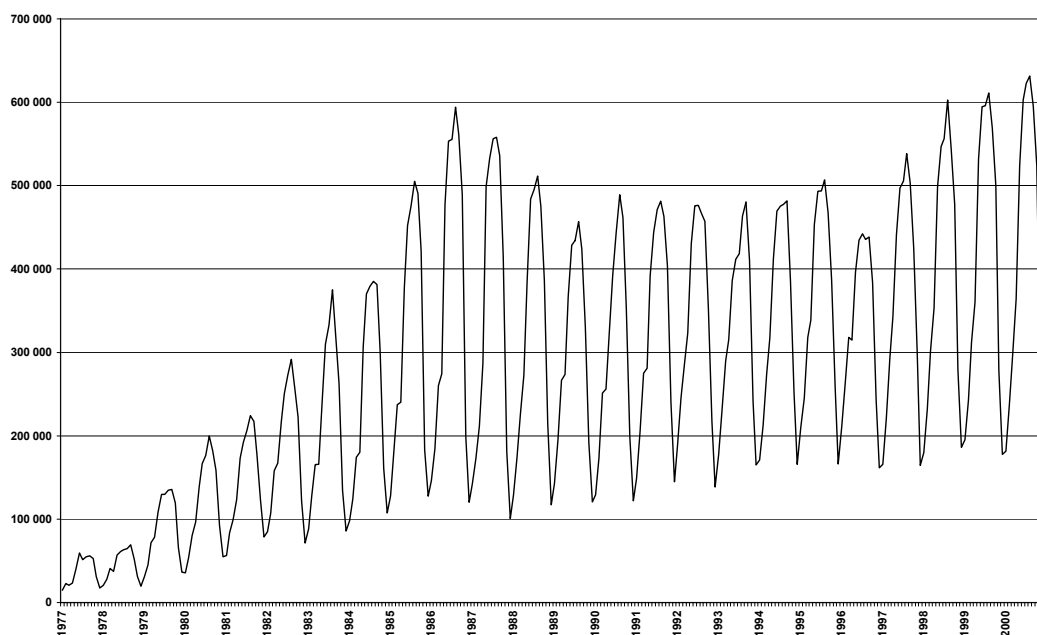
utilizados. No que diz respeito ao estudo de previsão, na Secção 4 são discutidos os métodos de previsão utilizados nos modelos lineares e nos modelos com não-linearidade determinística e estocástica; são apresentadas as estatísticas de avaliação e é feita a comparação da performance de previsão dos diferentes modelos e propostas as metodologias de combinação de previsões. Na secção 5 é desenvolvido um estudo de previsão com base nos modelos apresentados na Secção 3 para a série considerada. Finalmente, na secção 6, são apresentadas as principais conclusões que o desenvolvimento do trabalho permitiu alcançar.

## 2 - Série de dados e suas propriedades estatísticas

### 2.1 Descrição da série

Neste artigo, são utilizados dados mensais (não ajustados sazonalmente) de Janeiro de 1977 a Dezembro de 2000, relativos às dormidas de hóspedes do Reino Unido no Algarve (Figura 2.1), facultados pelo Instituto Nacional de Estatística (INE), Direcção Regional do Algarve. Em todos os modelos considerados foram utilizadas as observações relativas ao período de 1977-2000.

Figura 2.1: Dormidas de hóspedes oriundos do Reino Unido



### 2.2 Propriedades da série

A sazonalidade pode ser determinística ou estocástica estacionária ou não estacionária (vide, *inter alia*, Franses, 1996 e Ghysels e Osborn, 2001). Quando os fenómenos apresentam um comportamento sazonal estável em torno da tendência determinística, as *dummies* sazonais explicam a quase totalidade da variabilidade das

séries. Neste caso, estamos perante o que habitualmente se designa por sazonalidade determinística.

Quando um processo apresenta pelo menos uma raiz unitária sazonal na sua representação autorregressiva, estamos perante um processo sazonal integrado (vide Hylleberg *et al.*, 1990). Este tipo de processos estão associados a fenómenos de persistência geradores de alterações lentas mas permanentes nos padrões sazonais.

No caso particular das séries mensais, um passeio aleatório sazonal como o que se segue,  $y_t = y_{t-12} + \varepsilon_t$  apresenta raízes unitárias nas frequências,  $0, \pi, \frac{3\pi}{2}, \frac{\pi}{2}, \frac{\pi}{3}, \frac{5\pi}{3}, \frac{2\pi}{3}, \frac{4\pi}{3}, \frac{\pi}{6}, \frac{5\pi}{6}$  e  $\frac{7\pi}{6}$ . Neste caso, o polinómio  $(1 - L^{12})$ , onde L representa o operador de desfasamento temporal ( $L^k y_t = y_{t-k}$ ) pode ser factorizado (cf. Hylleber, *et al.* (1990), Beaulieu e Miron (1993) e Smith e Taylor (1999)) em função das suas raízes unitárias sazonais  $\exp(\pm i2\pi k/S)$ ,  $k=0, \dots, 6$ , de modo a obter-se a regressão teste,

$$\Delta_{12} y_t = \sum_{i=1}^{12} \pi_i z_{i,t-1} + \sum_{j=1}^p \beta_j \Delta_{12} y_{t-j} + u_t$$

onde

$$z_{1,t} = \sum_{j=0}^{11} y_{t-j}, \quad z_{2,t} = \sum_{j=0}^{11} \cos[(j+1)\pi] y_{t-j},$$

$$z_{i,t} = \sum_{j=0}^{11} \cos[(j+1)\varpi_k] y_{t-j}, \quad z_{i+1,t} = -\sum_{j=0}^{11} \sin[(j+1)\varpi_k] y_{t-j}$$

$i = 3, 5, 7, 9, 11$ ,  $\varpi_k = 2\pi k/12$ ,  $k=1, \dots, 5$ , e  $\Delta_{12} y_t = (1 - L^{12})y_t = y_t - y_{t-12}$ . Note que  $z_{j,t}$ ,  $j = 1, \dots, 12$ , correspondem a combinações lineares de desfasamentos de  $y_t$  de modo que em cada combinação seja isolada uma determinada raiz unitária. Por exemplo,  $z_{1,t}$  visa avaliar se existe ou não uma raiz unitária na frequência zero.

Os coeficientes dos  $\pi_j$ ,  $j = 1, \dots, 12$ , são os parâmetros relevantes para efeitos de teste às raízes unitárias. Consequentemente, para testar a hipótese nula de passeio aleatório sazonal contra a alternativa de estacionaridade em pelo menos uma frequência, HEGY e Smith and Taylor (1999), *inter alia*, sugeriram:  $t_1$  (unilateral) para a exclusão de  $z_{1,t-1}$ ;  $t_2$  (unilateral) para a exclusão de  $z_{2,t-1}$ ;  $t_k$  (unilateral) para a exclusão de  $z_{k,t-1}$  e  $t_{k+1}$  (bilateral) para a exclusão de  $z_{k+1,t-1}$ , para  $k = 3, 5, 7, 9, 11$ ; sugeriram também os testes conjuntos  $F_{[k]}$  para a exclusão simultânea de  $z_{k,t-1}$  e  $z_{k+1,t-1}$ ,  $k = 3, 5, 7, 9, 11$ . Ghysels, Lee e Noh (1994), Taylor (1998), Smith and Taylor (1998, 1999) consideram também os testes,  $F_{[2 \dots 12]}$ , para a exclusão de  $z_{2,t-1}$  e  $\{z_{k,t-1}, z_{k+1,t-1}\}_{k=3}^{12}$ , e  $F_{[1 \dots 12]}$ , para a exclusão de  $z_{1,t-1}$ ,  $z_{2,t-1}$  e  $\{z_{k,t-1}, z_{k+1,t-1}\}_{k=3}^{12}$ .

Para uma abordagem mais detalhada sobre estes testes e suas propriedades veja Rodrigues e Taylor (2004).

A Tabela 2.1 apresenta os resultados da aplicação do teste HEGY mensal à série relativa às dormidas de hóspedes do Reino Unido no Algarve.

**Tabela 2.1: Resultados do Teste HEGY à Presença de Raízes Unitárias**

Frequência	Teste HEGY
0	-1.836*
$\pi$	-1.590*
$\frac{\pi}{2}$ e $\frac{3\pi}{2}$	6.271*
$\frac{5\pi}{3}$ e $\frac{7\pi}{6}$	7.969
$\frac{\pi}{6}$ e $\frac{11\pi}{6}$	4.735*
$\frac{2\pi}{3}$ e $\frac{4\pi}{3}$	3.556*
$\frac{\pi}{3}$ e $\frac{5\pi}{3}$	4.081*

**Nota:** \* indica a não rejeição da presença de raiz unitária a um nível de significância de 5%.

Da análise dos resultados desta tabela verifica-se que unicamente o par de raízes complexas, relativas às frequências  $5\pi/3$  e  $7\pi/6$  é rejeitado. Portanto, verifica-se forte evidência de um padrão sazonal não estacionário.

### 3. Descrição dos modelos

#### 3.1 Modelos Sazonais Determinísticos

Se considerarmos que o padrão sazonal de uma série é bem aproximado apenas por componentes sazonais determinísticas é possível recorrer a um modelo como o que se segue,

$$y_t = \sum_{i=1}^S \alpha_i D_{it} + \sum_{j=1}^S \beta_j D_{jt} t + x_t \quad (3.1)$$

$$\phi(L)\Delta_1 x_t = \theta(L)\varepsilon_t$$

onde  $D_{it}$  corresponde à *dummy* sazonal, que assume o valor um na estação  $i$  e zero nos outros casos;  $\Delta_1 x_t$  é gerado por um processo autoregressivo de médias móveis estacionário e invertível, onde  $\phi(L)$  e  $\theta(L)$  representam polinómios em  $L$  de ordem  $p$  e  $q$ , respectivamente; e  $\varepsilon_t \sim iid(0, \sigma^2)$ .

### 3.2 Modelos Sazonais Autoregressivos

Quando as relações causais entre variáveis são difíceis de fundamentar, uma alternativa consiste em procurar explicar o fenómeno em estudo com base unicamente na sua evolução passada. Com efeito, a evolução histórica incorpora informação acerca de inúmeras variáveis que influenciam o fenómeno, pelo que os modelos autorregressivos são frequentemente utilizados na literatura econométrica para efeitos de previsão.

O modelo Sazonal Autorregressivo (SAR) com componentes determinísticas pode ser representado, para a frequência mensal pela seguinte expressão:

$$y_t = \sum_{i=1}^s \alpha_i D_{it} + \sum_{j=1}^s \beta_j D_{jt} + x_t \quad (3.2)$$

$$\phi(L)\Delta_{12}x_t = \theta(L)\varepsilon_t$$

Para uma abordagem mais exaustiva aos modelos autorregressivos e sazonal autorregressivo veja, *inter alia*, Hamilton (1994), Franses (1996) e Ghysels e Osborn (2001).

### 3.3 Modelos Periódicos Autorregressivos

Os modelos Periódicos Autorregressivos (PAR) constituem um conjunto de modelos alternativos utilizados na modelação de séries sazonalmente não ajustadas. Estes modelos permitem a alteração dos parâmetros do modelo de acordo com as estações do ano. Nos modelos PAR, as observações de cada estação do ano são modeladas por um regime autorregressivo específico.

Em termos gerais, um modelo  $PAR(12,p)$  pode ser representado pela seguinte expressão,

$$\Delta_k y_t = \sum_{j=1}^{12} (u_j + \beta_j t) D_{jt} + \sum_{j=1}^{12} \sum_{i=1}^{p_j} \alpha_{ij} D_{jt} \Delta_k y_{t-i} + \varepsilon_t, \quad (3.3)$$

onde  $p = \max\{p_1, \dots, p_{12}\}$ ,  $p_j$  representa a ordem da parte autorregressiva do modelo da estação  $j$ ,  $\varepsilon_t \sim iid(0, \sigma^2)$  e  $D_{jt}$  são *dummies* sazonais que assumem o valor um na estação  $j$  e zero nas outras estações. Os modelos PAR são modelos não lineares, onde a não-linearidade é introduzida de forma determinística.

Numa fase prévia à aplicação dos modelos PAR importa verificar se nos dados em análise há evidência de periodicidade, de forma a evitar o recurso a modelos sobreparametrizados. Note que num simples  $PAR(12,1)$  não restrito sem componente determinística são estimados 24 parâmetros e num  $PAR(12,p)$ ,  $12 + \sum_{j=1}^p 12$ , parâmetros.

Uma forma de verificar a presença de periodicidade consiste em testar a hipótese nula

$$H_0 : \alpha_{ij} = \alpha_i$$

contra a hiptese alternativa,

$$H_1 : \alpha_{ij} \neq \alpha_i$$

com  $j = 1, 2, \dots, 12$  e  $i = 1, 2, \dots, p$ . O teste  hiptese nula dado por,

$$F_{PAR} = \frac{(RSS_{H_0} - RSS_{H_1}) / (S - 1)p}{RSS_{H_1} / (T - (S + Sp))} \sim F_{[(S-1)p, T-(S+Sp)]} \quad (3.4)$$

onde  $RSS_{H_0}$  e  $RSS_{H_1}$  representam a soma dos quadrados dos resduos sob a hiptese nula e alternativa, respectivamente.  possvel encontrar mais desenvolvimentos, em termos dos testes  periodicidade e  heteroscedasticidade sazonal em Boswijk and Franses (1996) e Rodrigues e Gouveia (2004). Neste ultimo artigo foi possvel encontrar evidncia estatstica de periodicidade na srie utilizada no presente trabalho, validando assim a aplicao dos modelos PAR a esta srie.

A frequncia mensal da sries utilizada neste artigo sugere o recurso a um modelo  $PAR(12, p)$ , onde  $p = \max \{p_1, \dots, p_{12}\}$ ,

$$\Delta_k y_t = \sum_{i=1}^{12} \delta_i D_{it} + \sum_{i=1}^{12} \sum_{j=1}^{p_i} \beta_{ij} \Delta_k y_{t-j} + \varepsilon_t. \quad (3.5)$$

Porm os 12 regimes de um modelo deste tipo tendem a gerar problemas de sobreparametrizao, como j foi referido. Neste contexto,  aconselhvel o recurso  metodologia de reduo do modelo proposta por Rodrigues e Gouveia (2004), segundo a qual podem ser definidos trs regimes correspondentes  estaes determinadas em funo da percentagem de dormidas de cada ms. A classificao dos dados obtida por Rodrigues e Gouveia (2004) utilizada na construo de um modelo  $PAR(3, p^*)$  para a srie em anlise consta da tabela 3.2.

**Tabela 3.2: Classificao da srie: Reino Unido**

poca baixa	poca intermdia	poca alta
Nov. a Mar.	Abr., Mai, Out.	Jun. a Set.

### 3.4 Modelos SETAR Sazonais

Os modelos SETAR constituem uma classe importante de modelos no lineares que se baseiam na ideia de que a evoluo de um processo segue um conjunto de regimes autorregressivos, onde o valor da srie e o regime vigente em cada perodo, dependem do valor assumido por uma varivel limite. Estes modelos revelam-se particularmente adequados para a modelao das caractersticas do ciclo econmico, como a assimetria entre as fases de expanso e contraco (c.f. Hamilton, 1989) o desemprego (c.f. Caner and Hansen, 2001), as sries financeiras (c.f. Franses e van Dijk, 2000) a Teoria da Paridade dos Poderes de Compra (c.f. Enders e Granger, 1998 e Gouveia e Rodrigues, 2002), as taxas de juro (c.f. Enders e Granger, 1998), e a

estabilidade monetária no Mecanismo de Taxas de Câmbio do Sistema Monetário Europeu (Gouveia, Rebelo e Rodrigues, 2003).

Consideremos, para efeitos de exposição, um modelo SETAR com dois regimes e componente sazonal dado por,

$$\Delta_k y_t = \begin{cases} \sum_{j=1}^S \alpha_{1j} D_{jt} + \sum_{j=1}^S \lambda_{1j} D_{jt} t + \phi_1 \Delta_k y_{t-1} + e_{1t} & \text{if } q_{t-d} \leq \gamma \\ \sum_{j=1}^S \alpha_{2j} D_{jt} + \sum_{j=1}^S \lambda_{2j} D_{jt} t + \phi_2 \Delta_k y_{t-1} + e_{2t} & \text{if } q_{t-d} > \gamma \end{cases} \quad (3.6)$$

onde  $q_{t-d}$  representa a variável limite e  $d$  o parâmetro de desfasamento. Por seu turno, o parâmetro de limite,  $\gamma$  é determinado a partir dos valores da variável limite,  $q_{t-d}$ , que geralmente assume a forma de uma combinação de desfasamentos da variável dependente, vide, *inter alia*, Hansen (1997, p. 10).

Embora os modelos SETAR sazonais possuam uma estrutura semelhante à dos modelos PAR, a natureza da sua não linearidade é estocástica (dado que o parâmetro limite é endógeno ao modelo).

Considerando  $\beta_i = (\alpha_{i,j}, \lambda_{i,j}, \phi_i)$ ,  $i = 1, 2$ ;  $j = 1, \dots, S$  e  $\beta = (\beta'_1, \beta'_2)$  o estimador obtido pelo método dos mínimos quadrados ordinários (MMQO) vem dado por

$$\hat{\beta} = \left( \sum_{t=1}^n x_t(\gamma) x_t(\gamma)' \right)^{-1} \left( \sum_{t=1}^n x_t(\gamma) \Delta_k y_t \right). \quad (3.8)$$

No processo de estimação, os parâmetros  $d$  e  $\gamma$  são sequencialmente estimados da seguinte forma,

$$\hat{\gamma} = \arg \min_{\gamma \in \Gamma, d \in S} \hat{\sigma}^2(\gamma, d) \quad (3.9)$$

onde  $\hat{\sigma}^2(\gamma, d)$  corresponde ao estimador da variância para cada par  $(\gamma, d)$  obtido por recurso aos MMQO,  $\Gamma$  representa o conjunto de todos os possíveis valores do parâmetro limite (ou seja, o conjunto de possíveis valores da variável limite,  $q_{t-d}$ ) e  $s = \{1, 2, \dots, p\}$  é o conjunto de todos os valores possíveis do parâmetro de desfasamento,  $d$ . Segundo Chan (1993), o estimador  $\hat{\gamma}$ , obtido de acordo com (3.9) é fortemente consistente. A importância da endogeneização do parâmetro de desfasamento,  $d$ , para efeitos da estimação do modelo é demonstrada através de diversos estudos de Monte Carlo; e.g. Gouveia (2000) e Rodrigues e Gouveia (2001).

A pertinência da aplicação dos modelos SETAR depende da existência de diferenças significativas entre os coeficientes dos dois regimes. Os testes à linearidade em modelos SETAR para variáveis estacionárias foram objecto de desenvolvimentos em vários estudos, veja, por exemplo, Caner e Hansen (2001), Chan (1990), Chan e Tong (1990) e Hansen (1997). Estes autores propõem, sob a hipótese nula de linearidade o seguinte teste,

$$F(\gamma) = n \left( \frac{\tilde{\sigma}^2 - \hat{\sigma}^2(\gamma)}{\hat{\sigma}^2(\gamma)} \right) \quad (3.10)$$

onde  $n$  representa o nmero de observaes e  $\tilde{\sigma}^2$  e  $\hat{\sigma}^2(\gamma)$ , representam respectivamente, os estimadores das varincias da varivel residual resultante dos modelos linear autorregressivo e SETAR.

Note no entanto, que a presena de um parmetro perturbador,  $\gamma$ , sob a hiptese alternativa que no est presente sob a hiptese nula, conduz a distribuices no-*standard* e no-centrais (invalidando a utilizao da teoria assinttica tradicional) aconselhando assim  obteno de valores crticos com base no mtodo do *bootstrap*. Este fenmeno  conhecido na literatura econmica por problema de Davis (1977); veja, tambm, Andrews e Ploberger (1994) e Hansen (1996).

Para uma anlise mais exaustiva dos modelos SETAR veja entre outros Tong (1990), Hansen (1997, 2000), Caner e Hansen (2001) e Rodrigues e Gouveia (2001).

## 4. Previso

### 4.1. Previso com modelos lineares e modelos no-lineares determinsticos

De acordo com Box e Jenkins (1970), o valor da previso de uma determinada srie temporal que minimiza a Mdia do Quadrado dos Erros de Previso (MQEP),

$$MQEP = E[e_{t+h|t}^2] = E[(y_{t+h} - \hat{y}_{t+h|t})^2] \quad (4.1)$$

 dado pelo valor esperado condicional de  $\hat{y}_{t+h|t} = E[y_{t+h} | \Omega_t] = F(y_{t+h}; \theta | \Omega_t)$ , onde  $\Omega_t$  representa a informao histrica da srie temporal no momento  $t$  e representa a funo associada ao modelo linear em causa. No caso linear a obteno de previses para  $h$  perodos pode ser realizada por substituio recursiva.

### 4.2. Previso com modelos SETAR

A obteno de previses no caso dos modelos no-lineares  mais complexa e computacionalmente mais morosa do que no caso dos modelos lineares. Quando  utilizado o mtodo dos mnimos quadrados ordinrios na obteno de previses pontuais, os valores futuros da srie so calculados com base no valor esperado condicional (aos valores histricos da srie), i.e.,

$$\hat{y}_{t+h|t} = E[y_{t+h} | \Omega_t]. \quad (4.2)$$

O valor da previso para  $h=1$  pode ser facilmente obtido pela expresso,

$$\hat{y}_{t+1|t} = E[y_{t+1} | \Omega_t] = F(y_t; \theta). \quad (4.3)$$

No entanto, sempre que  $h > 1$ , o calculo da previso torna-se mais complexo. Com efeito, por exemplo, para  $h = 2$ , temos,

$$\hat{y}_{t+2|t} = E[y_{t+2} | \Omega_t] = E[F(y_{t+1}; \theta | \Omega_t)]. \quad (4.4)$$

Ora, no caso dos modelos SETAR,

$$E[F(y_{t+1}; \theta | \Omega_t)] \neq F[E(y_{t+1}; \theta | \Omega_t)] = F(\hat{y}_{t+1}; \theta). \quad (4.5)$$

Esta impossibilidade de permutar os operadores  $E$  e  $F$  no caso dos modelos SETAR (e dos modelos não-lineares em geral) tem como consequência a impossibilidade de estabelecer relações recursivas entre diferentes horizontes de previsão em contraponto com o que acontece com os modelos lineares.

Quando esta impossibilidade de permutação é negligenciada, estamos perante um método de previsão designado na literatura por “naive” (vide, *inter alia*, Franses e van Dijk, 2000) dado pela expressão,

$$\hat{y}_{t+h|t} = F(\hat{y}_{t+2|t}; \theta) \quad (4.6)$$

e que é gerador de previsões enviesadas.

A obtenção de previsões com modelos não-lineares está geralmente associado a processos de integração numérica baseadas na relação de Chapman-Kolmorov. Neste contexto, a previsão para  $h$  períodos pode ser obtida de uma forma recursiva a partir da expressão,

$$\hat{y}_{t+h|t} = E[y_{t+h} | \Omega_t] = \int_{-\infty}^{\infty} E[y_{t+h} | y_{t+h-1}] g(y_{t+h-1} | \Omega_t) dy_{t+h-1} \quad (4.6)$$

onde  $g(y_{t+h-1} | \Omega_t)$  é a distribuição condicional de  $y_{t+h-1}$  dado  $\Omega_t$ .

Uma vez que, de uma forma geral, a expressão (4.6) não é conhecida torna-se necessário recorrer a técnicas de computação intensiva. Uma possível abordagem consiste no recurso ao método de Monte Carlo segundo o qual a previsão para  $h$  períodos pode ser obtida com base na expressão,

$$\hat{y}_{t+h|t}^{mc} = \left( \frac{1}{k} \sum_{i_h=1}^k \dots F \left( \frac{1}{k} \sum_{i_3=1}^k F \left( \frac{1}{k} \sum_{i_2=1}^k F \left( \hat{y}_{t+1|t} + \varepsilon_{i_2}; \theta \right) + \varepsilon_{i_3}; \theta \right) \dots \right) + \varepsilon_{i_h}; \theta \right) \quad (4.7)$$

onde  $k$  representa o número de replicações utilizado e  $\varepsilon_{i_j}$  (com  $i=1, \dots, K$ ,  $j=1, \dots, h$ ) os valores aleatórios obtidos a partir da presumível distribuição de  $\varepsilon_{i_j}$ .

No caso da distribuição de  $\varepsilon_{i_j}$  ser desconhecida é possível recorrer ao método do *bootstrap* baseado em,

$$\hat{y}_{t+h|t}^{boot} = \left( \frac{1}{k} \sum_{i_h=1}^k \dots F \left( \frac{1}{k} \sum_{i_3=1}^k F \left( \frac{1}{k} \sum_{i_2=1}^k F \left( \hat{y}_{t+1|t} + \hat{\varepsilon}_{i_2}; \theta \right) + \hat{\varepsilon}_{i_3}; \theta \right) \dots \right) + \hat{\varepsilon}_{i_h}; \theta \right). \quad (4.8)$$

Neste caso,  $\hat{\varepsilon}_{ij}$ ,  $j = 1, \dots, h$ , são valores aleatórios, com reposição, obtidos a partir dos resíduos do modelo estimado. Lin e Granger (1994) e Clements e Smith (1997) comparam os três métodos de obtenção de previsões para horizontes temporais superiores a um e obtêm evidência empírica da vantagem associada ao método do *bootstrap*.

### 4.3 Comparação da performance em termos de previsão

Para efeitos de avaliação da performance de cada um dos modelos, em termos de previsão, são retidas  $m+h$  observações, onde  $h$  representa o horizonte temporal de previsão, de modo a avaliarem-se e compararem-se  $m$  vezes, de forma sequencial a  $h$  períodos para a frente (vide, *inter alia*, Franses e van Dijk, 2000). Ou seja, a capacidade de previsão de um modelo para um horizonte de  $h$  períodos pode deste modo ser testada para os últimos  $[m/12]$  anos. Neste estudo considerou-se  $m = 36$  (3 anos), 48 (4 anos) e 60 (5 anos).

Tendo por base esta metodologia, são utilizados três critérios para efeitos de comparação de previsões:

- i) A Raiz Quadrada da Média do Quadrado dos Erros de Previsão (RQMQEP),

$$RQMQEP = \sqrt{\frac{1}{h} \sum_{j=1}^h (y_{t+j} - \hat{y}_{t+j|t+j-1})^2} . \quad (4.9)$$

Para efeitos de avaliação da qualidade das previsões e no enquadramento da metodologia associada à janela de previsão móvel, torna-se pertinente recorrer à média da Raiz Quadrada da Média do Quadrado dos Erros de Previsão (MRQMQEP), dada pela expressão,

$$MRQMQEP = \frac{1}{m} \sum_{i=1}^m \sqrt{\frac{1}{h} \sum_{j=1}^h (y_{t-m+i+j} - \hat{y}_{t-m+i+j|t-m+i+j-1})^2} . \quad (4.10)$$

- ii) O Erro de Previsão Absoluto (EPA),

$$EPA = |y_{t-m+i+h} - \hat{y}_{t-m+i+h|t-m+i}| . \quad (4.11)$$

No contexto da utilização da janela de previsão móvel recorre-se à média dos Erros de Previsão Absolutos (MEPA) dada pela expressão,

$$MEPA = \frac{1}{m} \sum_{i=1}^m |y_{t-m+i+h} - \hat{y}_{t-m+i+h|t-m+i}| . \quad (4.12)$$

- iii) O Erro de Previsão Absoluto Percentual (EPAPerc),

$$EPA_{PERC} = \left| \frac{y_{t+h} - \hat{y}_{t+h|t}}{y_{t+h}} \right|. \quad (4.13)$$

Tal como no caso dos Erros de Previsão Absolutos, ao nível da janela de previsão, a medida a utilizar será a Média dos Erros de Previsão Absolutos Percentuais ( $MAPE_{PERC}$ ), i.e.,

$$MEPA_{PERC} = \frac{1}{m} \sum_{j=1}^m \left| \frac{y_{t-m+i+h} - \hat{y}_{t-m+i+h+j|t-m+i}}{y_{t-m+i+h}} \right|. \quad (4.14)$$

#### 4.4. Metodologia de combinação de previsões

As combinações lineares constituem uma das metodologias de crescente utilização na literatura com vista à obtenção de previsões (c.f., Bates e Granger, 1969; Stock e Watson, 2003), por recurso a,

$$w_{i,h,t} = \left(1/RQM_{QEP}_{i,h,t}\right)^\lambda / \sum_{j=1}^K \left(1/RQM_{QEP}_{i,h,t}\right)^\lambda \quad (4.15)$$

$$w_{i,h,t} = \left(1/MEPA_{i,h,t}\right)^\lambda / \sum_{j=1}^K \left(1/MEPA_{i,h,t}\right)^\lambda \quad (4.16)$$

onde  $k$  representa o número de modelos utilizados e  $w_{i,h,t}$  o peso do modelo  $i$  na previsão para o período  $h$ .

Neste caso, o peso de cada modelo é função de  $\lambda$ . Quando  $\lambda=0$ , todos os modelos têm igual peso na combinação. Um aumento de  $\lambda$  significa um aumento da ênfase nos modelos com melhor performance relativa. Na medida em que os modelos que apresentam melhor performance em termos de previsão no curto prazo nem sempre correspondem aos que possuem uma melhor performance no longo prazo, pode tornar-se pertinente o ajustamento da combinação de previsões a este tipo de fenómeno como se segue,

$$w_{i,j,t} = \begin{cases} \left(1/RQM_{QEP}_{i,h/2,t}\right)^\lambda / \sum_{j=1}^K \left(1/RQM_{QEP}_{i,h/2,t}\right)^\lambda & \text{se } j \leq h/2 \\ \left(1/RQM_{QEP}_{i,h,t}\right)^\lambda / \sum_{j=1}^K \left(1/RQM_{QEP}_{i,h,t}\right)^\lambda & \text{se } j > h/2 \end{cases} \quad (4.17)$$

$$w_{i,j,t} = \begin{cases} \left(1/MEPA_{i,h/2,t}\right)^\lambda / \sum_{j=1}^K \left(1/MEPA_{i,h/2,t}\right)^\lambda & \text{se } j \leq h/2 \\ \left(1/MEPA_{i,h,t}\right)^\lambda / \sum_{j=1}^K \left(1/MEPA_{i,h,t}\right)^\lambda & \text{se } j > h/2 \end{cases}. \quad (4.18)$$

Com efeito, as expressões (4.17) e (4.18) correspondem a uma adaptação das expressões (4.15) e (4.16) de forma a diferenciar o peso de cada modelo até  $h/2$  e de  $h/2$  a  $h$ .

Para além deste tipo de metodologias ainda se consideraram neste estudo combinações baseadas na média e na mediana da totalidade ou de parte do conjunto dos modelos individuais.

## **5. Estudo de Previsão**

### **5.1. Descrição do estudo**

A séries de dados que suporta o estudo de previsão (vide Secção 2) respeita ao período 1977-1998 sendo considerado um horizonte temporal de previsão de  $h = 24$  (de Janeiro de 1999 a Dezembro de 2000).

No domínio da modelação e previsão, a selecção da ordem da parte autorregressiva desempenha um papel crucial. Neste artigo, seguimos a metodologia proposta em Rodrigues e Gouveia (2004), onde, a partir de uma ordem máxima da parte autorregressiva suficientemente elevada (*e.g.*  $p = 36$ ) são eliminadas, sucessivamente, as variáveis menos significativas (procurando, no entanto, garantir em todas as fases a ausência de autocorrelação). Com base nesta metodologia é possível (em contraponto ao que tem sido tradicional, sobretudo no campo da aplicação de modelos PAR e SETAR) que cada regime apresente uma ordem distinta da parte autorregressiva. Esta abordagem permite a construção de modelos mais parcimoniosos com vantagens ao nível da previsão (vide Rodrigues e Gouveia, 2004).

Por outro lado, os métodos para obtenção de previsões correspondem aos descritos na secção 4.2 e são diferenciados para o caso dos modelos AR e PAR (lineares e com não-linearidade determinística) e SETAR (com não linearidade estocástica).

A selecção da ordem da parte autorregressiva e a obtenção de previsões teve por base métodos de computação intensiva. Para cada um dos modelos discriminados nas Tabelas 5.1 e 5.2, todos os cálculos foram realizados em GAUSS 3.2 e tiveram por base rotinas desenvolvidas pelos autores.

Tabela 5.1: Lista dos Modelos Utilizados

Modelo	Descrição	Filtro
1	ARP com 12 <i>dummies</i> sazonais e 1 tendência	Níveis
2	ARP com 12 <i>dummies</i> sazonais e 1 tendência	$\Delta_1\Delta_{12}$
3	ARP com 12 <i>dummies</i> sazonais e 1 tendência	$\Delta_{12}$
4	ARP com 12 <i>dummies</i> sazonais e 1 tendência	$\Delta_1$
5	ARP com 12 <i>dummies</i> sazonais e 1 tendência	$(1-L)(1+L)$
6	ARP com 3 <i>dummies</i> sazonais e 1 tendência	Níveis
7	ARP com 3 <i>dummies</i> sazonais e 1 tendência	$\Delta_{12}$
8	ARP com 3 <i>dummies</i> sazonais e 1 tendência	$\Delta_1$
9	ARP com 3 <i>dummies</i> sazonais e 1 tendência	$(1-L)(1+L)$
10	ARP com 3 <i>dummies</i> sazonais e 3 tendência	Níveis
11	ARP com 3 <i>dummies</i> sazonais e 3 tendência	$\Delta_{12}$
12	ARP com 3 <i>dummies</i> sazonais e 3 tendência	$\Delta_1$
13	ARP com 3 <i>dummies</i> sazonais e 3 tendência	$(1-L)(1+L)$
14	Modelo Determinístico com 12 <i>dummies</i> e 12 tendências	Níveis
15	Modelo Determinístico com 12 <i>dummies</i> e 12 tendências	$\Delta_1\Delta_{12}$
16	Modelo Determinístico com 12 <i>dummies</i> e 12 tendências	$\Delta_{12}$
17	Modelo Determinístico com 12 <i>dummies</i> e 12 tendências	$\Delta_1$
18	Modelo Determinístico com 12 <i>dummies</i> e 12 tendências	$(1-L)(1+L)$
19	PAR3 com 3 <i>dummies</i> e 3 tendências	Níveis
20	PAR3 com 3 <i>dummies</i> e 3 tendências	$\Delta_{12}$
21	PAR3 com 3 <i>dummies</i> e 3 tendências	$\Delta_1$
22	PAR3 com 3 <i>dummies</i> e 3 tendências	$(1-L)(1+L)$
23	PAR2 com 2 <i>dummies</i> e 1 tendência	Níveis
24	PAR2 com 2 <i>dummies</i> e 1 tendência	$\Delta_{12}$
25	PAR2 com 2 <i>dummies</i> e 1 tendência	$\Delta_1$
26	PAR2 com 2 <i>dummies</i> e 1 tendência	$(1-L)(1+L)$
27	PAR2 com 2 <i>dummies</i> e 2 tendências	Níveis
28	PAR2 com 2 <i>dummies</i> e 2 tendências	$\Delta_{12}$
29	PAR2 com 2 <i>dummies</i> e 2 tendências	$\Delta_1$
30	SEATAR com 3 <i>dummies</i> e 1 tendência- método "naive"	$\Delta_{12}$
31	SEATAR com 3 <i>dummies</i> e 1 tendência- método "Monte Carlo"	$\Delta_{12}$
32	SEATAR com 3 <i>dummies</i> e 1 tendência- método "Bootstrap"	D12

**Tabela 5.2: Lista das Combinações Consideradas**

Modelo	Descrição
33 Média	
34 Combi1	Expressão (14.15), $\lambda = 0$
35 Combi2	Expressão (14.15), $\lambda = 1$
36 Combi3	Expressão (14.15), $\lambda = 1.5$
37 Combi4	Expressão (14.15), $\lambda = 2$
38 Combi5	Expressão (14.15), $\lambda = 5$
39 Combi6	Expressão (14.16), $\lambda = 0$
40 Combi7	Expressão (14.16), $\lambda = 1$
41 Combi8	Expressão (14.16), $\lambda = 1.5$
42 Combi9	Expressão (14.16), $\lambda = 2$
43 Combi10	Expressão (14.16), $\lambda = 5$
44 Combi11	Expressão (14.17), $\lambda = 1$
45 Combi12	Expressão (14.18), $\lambda = 1$
46 Combi13	Média dos melhores 10 modelos (com base no MRQMSEP para $h=24$ )
47 Combi14	Média dos melhores 15 modelos (com base no MRQMSEP para $h=24$ )
48 Combi15	Mediana dos melhores 10 modelos (com base no MRQMSEP para $h=24$ )
49 Combi16	Mediana dos melhores 15 modelos (com base no MRQMSEP para $h=24$ )
50 Combi17	Média dos melhores 10 modelos (com base no MEPA para $h=24$ )
51 Combi18	Média dos melhores 15 modelos (com base no MEPA para $h=24$ )
52 Combi19	Mediana dos melhores 10 modelos (com base no MEPA para $h=24$ )
53 Combi20	Mediana dos melhores 15 modelos (com base no MEPA para $h=24$ )
54 Combi21	Mediana de Todas as combinações

## 5.2. Resultados do estudo de previsão<sup>1</sup>

De uma forma geral e para o período pós amostral todos os modelos utilizados apresentam boa performance em termos de previsão. Este resultado, que pode estar associado à metodologia de selecção da ordem da parte autorregressiva e é directamente observável na Tabela 5.5.

Deste estudo, resulta também evidência da vantagem do recurso a metodologias de combinação de previsões. Na tabela 5.3 regista-se que em relação a todos os horizontes temporais considerados (de  $h = 1$  a  $h = 24$ ) os diferentes tipos de metodologias de combinação de previsões encontram-se geralmente entre as previsões com menor RQMSEP. Neste âmbito, é particularmente notória a vantagem da metodologia de combinação sugerida em (4.18), que permite a diferenciação do peso de cada um dos modelos de acordo com o curto ( $h = 12$ ) e o longo prazo ( $h = 24$ ).

<sup>1</sup> Neste artigo apresentamos unicamente os resultados obtidos com base no RQMSEP. Relativamente aos resultados obtidos para as estatísticas consideradas em (4.12) e (4.14), estes podem ser solicitados aos autores.

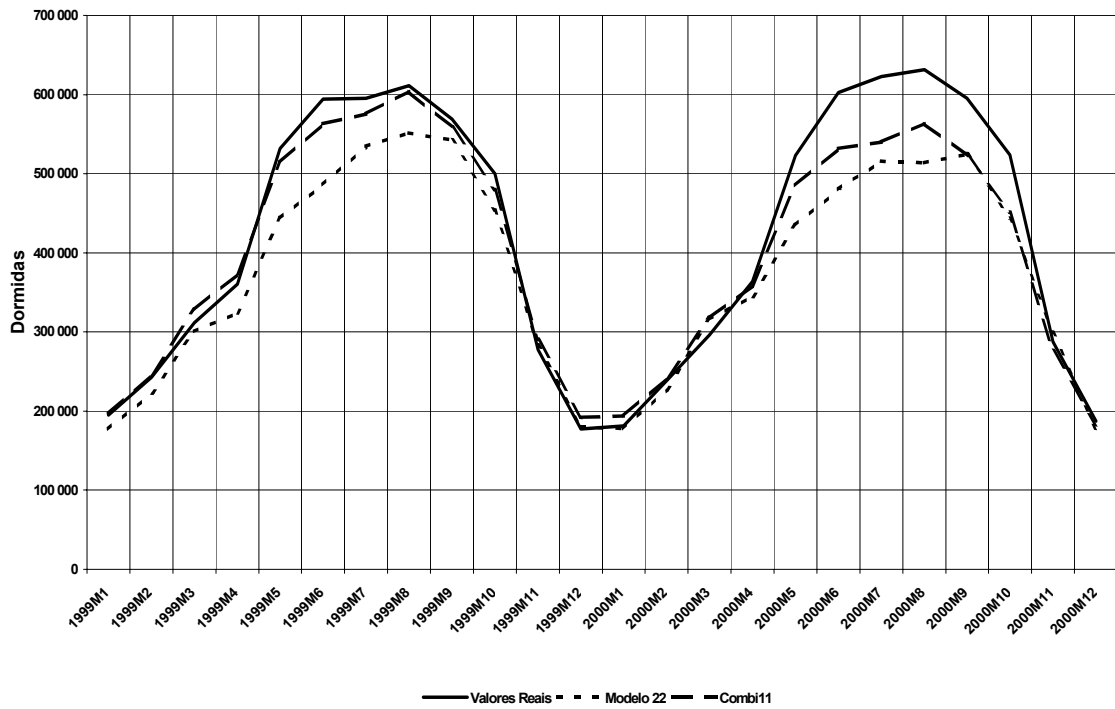
A Figura 5.1 permite comparar os resultados da previsão do melhor modelo (Modelo 44 que corresponde à combinação 11- vide tabela 5.2) com o modelo que apresentou pior performance em termos de previsão (modelo 22) em relação ao período pós amostral. Por seu turno, a figura 5.2 apresenta essa comparação em termos de erro percentual. Neste caso, desvios de sinal positivo e negativo representam respectivamente sobre e sub previsão. Em ambas as figuras é ainda notória uma deterioração da qualidade das previsões com o aumento do horizonte temporal de previsão.

**Tabela 5.3: Ranking dos modelos utilizados de acordo com o RQMQEP (período pós-amostral)**

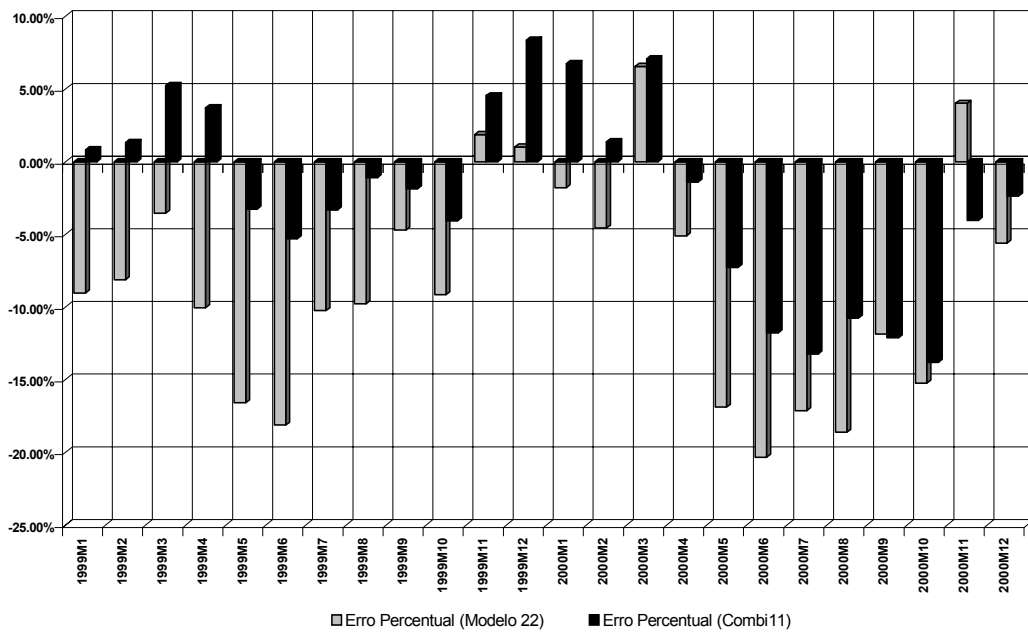
Posição /Horizonte de Previsão	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1	50	4	46	45	24	10	44	44	44	44	44	44	44	44	44	44	30	30	30	30	30	30	30	30
2	46	47	50	35	20	44	10	30	30	2	30	30	30	30	30	30	44	44	18	18	18	18	18	18
3	28	51	16	54	44	24	30	10	10	30	2	10	10	10	24	24	24	10	44	2	2	2	17	17
4	44	44	54	39	7	30	24	2	2	10	10	18	24	24	10	6	10	6	10	17	17	17	2	14
5	4	50	45	36	50	6	6	24	23	18	18	2	18	23	6	10	6	24	2	14	14	14	14	2
6	24	46	35	50	46	28	2	6	27	24	24	24	23	6	42	54	41	38	6	44	23	10	10	10
7	47	38	38	46	47	46	1	23	6	14	14	14	6	34	43	37	40	33	23	23	10	23	23	23
8	51	34	34	38	51	50	50	27	18	6	6	6	33	38	37	42	39	34	1	10	44	44	44	44
9	38	33	33	33	10	51	46	18	24	23	23	23	38	33	54	43	54	39	17	6	6	33	34	33
10	34	45	39	34	28	47	51	1	1	17	27	38	34	27	39	39	38	40	38	38	38	34	38	38
11	33	27	36	40	6	1	47	38	33	15	38	34	27	39	33	34	34	41	33	34	34	38	33	34
12	27	54	4	4	52	54	27	34	34	27	34	33	39	54	38	38	33	35	34	33	33	6	6	6
13	53	16	48	41	48	2	23	33	38	38	33	17	54	35	34	33	42	37	14	39	1	15	35	35
14	7	35	52	48	53	52	32	54	54	34	17	27	35	40	40	40	37	42	39	35	35	35	15	39
15	49	39	13	52	49	48	38	47	39	33	1	39	40	43	36	35	43	43	35	1	39	1	39	1
16	54	53	40	7	8	38	34	51	14	1	15	54	43	37	35	41	35	36	40	36	36	39	1	36
17	48	49	5	3	54	33	33	50	35	54	54	1	37	42	41	36	36	1	36	40	40	36	36	15
18	52	24	53	16	12	34	54	46	40	39	39	35	42	36	23	23	45	54	41	43	42	40	40	40
19	45	20	49	49	13	49	31	39	36	35	35	40	36	41	45	45	47	18	54	42	43	31	45	45
20	10	10	51	53	9	53	28	35	47	40	40	36	41	45	50	50	51	45	43	37	37	32	43	43
21	39	52	47	13	3	39	39	40	51	36	36	43	45	47	46	46	50	47	42	41	45	45	42	42
22	35	48	41	5	38	40	35	36	46	43	43	42	51	51	52	48	46	51	37	45	27	43	37	37
23	21	36	7	51	34	8	8	14	50	42	42	37	47	50	48	52	48	46	45	54	41	42	41	41
24	23	40	3	47	33	41	36	43	41	37	37	41	50	46	51	51	52	50	47	27	54	37	27	27
25	6	13	9	20	39	35	40	42	43	41	41	45	46	1	47	47	28	23	51	47	15	41	54	54
26	13	3	20	9	40	36	48	37	42	45	45	47	1	52	8	8	23	48	46	51	47	27	47	47
27	3	5	6	43	35	12	52	41	37	47	47	51	52	48	49	49	8	52	50	46	51	54	51	51
28	40	41	8	42	41	9	12	45	45	51	51	15	48	8	53	53	1	8	27	50	32	47	50	50
29	36	23	43	37	45	15	45	52	15	50	50	50	2	49	3	27	49	2	48	15	31	51	46	46
30	16	7	42	8	36	43	41	48	52	46	46	8	53	27	12	53	53	52	48	50	50	31	48	48
31	5	9	37	6	1	42	18	15	48	52	48	52	3	28	12	9	12	49	8	52	46	46	32	52
32	8	28	12	23	30	37	43	8	17	48	52	48	14	3	9	1	9	12	31	8	48	48	48	49
33	9	6	24	12	15	45	42	28	8	8	8	8	49	12	1	28	13	9	32	53	52	52	52	53
34	41	1	10	19	43	13	37	49	53	49	49	49	53	26	28	3	27	20	15	49	8	49	49	8
35	12	8	44	27	42	20	49	53	49	53	53	53	26	18	26	26	3	15	24	31	49	53	53	31
36	20	43	19	44	37	7	53	17	12	26	26	26	28	9	13	13	20	28	53	32	53	8	8	9
37	15	42	23	1	4	3	9	12	9	3	3	3	12	29	29	29	15	27	49	29	12	29	29	12
38	43	37	21	10	23	27	15	9	7	28	28	28	17	13	7	25	18	17	12	9	9	9	9	26
39	42	12	27	15	27	4	14	7	3	12	12	12	9	2	25	20	26	13	9	12	29	12	12	29
40	37	21	15	24	11	23	7	26	26	9	9	9	29	15	18	15	29	14	29	24	26	26	26	32
41	25	15	28	28	16	32	13	3	25	29	29	29	15	25	20	7	2	3	13	26	25	25	25	25
42	11	19	1	26	26	31	17	13	29	7	13	13	13	14	15	18	7	29	25	25	13	13	13	13
43	1	11	22	30	2	18	3	25	28	13	7	7	7	7	16	16	25	25	26	13	3	24	24	24
44	2	22	30	25	5	14	26	29	13	32	25	25	25	17	2	2	16	31	20	3	24	3	3	3
45	19	25	26	21	25	25	25	32	32	31	20	20	20	20	17	4	17	32	3	20	20	20	20	20
46	22	2	25	11	29	26	4	31	31	25	16	16	16	16	4	5	14	26	28	7	7	16	16	16
47	26	32	11	29	19	11	29	4	20	20	31	4	4	4	5	17	31	7	7	28	16	7	28	28
48	18	26	29	22	32	17	20	20	4	16	32	5	5	5	14	14	32	16	16	16	28	28	7	7
49	31	30	32	32	31	16	5	16	16	4	4	31	31	31	31	31	4	5	5	5	5	5	5	5
50	30	31	31	31	18	5	16	5	5	5	5	32	32	32	32	32	5	4	4	4	4	4	22	22
51	32	29	2	2	14	29	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11
52	29	18	18	18	17	19	19	19	19	22	22	22	22	22	22	22	22	22	22	11	11	11	19	19
53	14	14	14	14	22	22	22	22	22	19	19	19	19	19	19	19	19	19	19	19	19	19	11	11
54	17	17	17	17	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21

**Nota:** Para a discrição dos modelos veja as Tabelas 5.1 e 5.2.

**Figura 5.1: Previsão para o período: 1999-2000  
(Dormidas de Hóspedes do Reino Unido no Algarve)**



**Figura 5.2: Erro Percentual de Previsão do período 1999-2000  
(Dormidas de Hóspedes do Reino Unido no Algarve)**



## 6. Conclusão

Baseado na ideia introduzida por Bates e Granger (1969) de que a exploração das relações de complementaridade, em termos de poder explicativos, de diversos tipos de modelos, neste artigo foi possível obter evidência da superioridade de algumas metodologias de combinação de previsões para o caso da série relativa às dormidas de hóspedes do Reino Unido no Algarve.

A metodologia de combinação de previsões baseada na diferenciação do peso de cada modelo em função da performance de curto e longo prazo, sugerida na secção 4 foi a que obteve melhor performance em termos de previsão

Com este estudo de previsão foi ainda possível aplicar com vantagem a metodologia de selecção da ordem sugerida em Rodrigues e Gouveia (2004) que permite a obtenção de modelos parcimoniosos, com salvaguarda de ausência de autocorrelação, o que acaba por se traduzir num vasto conjunto de modelos com boa performance em termos de previsão.

Por fim, e ainda que neste artigo a aplicação esteja limitada a uma série temporal, os resultados obtidos deixam boas perspectivas no que toca à obtenção de previsões de qualidade para um conjunto de variáveis importantes na área do turismo.

## Referncias

- Andrews, W.K. Ploberger (1994), Optimal Tests when a Nuisance Parameter is Present Only Under the Alternative, *Econometrica*, 62, 1383-1414.
- Bates, J.M. e Granger, C.W.J. (1969), The Combination of Forecasts, *Operations Research Quarterly* 20, 451-468.
- Beaulieu, J.J. e J. A. Miron (1993), Seasonal Unit Roots in Agregate U.S. Data, *Journal of Econometrics*, 55, 305-328.
- Boswijk, H. P. e P.H. Franses (1996), Unit Root in Periodic Autorressions, *Journal of Time Series Analysis*, 17, 221-245.
- Box, G.E.P. e G.M. Jenkins (1970), *Time Series Analysis, Forecasting and Control*, San Francisco: Holden-Day.
- Caner, M. e B. Hansen (2001), Threshold Autoregression with a Unit Root, *Econometrica*, 69, 1555-1596.
- Chan, K.S. (1990), Testing for Threshold Autoregression, *The Annals of Statistics*, 18, 1886-1894.
- Chan, K.S. e H. Tong (1990), On Likelihood Ratio Tests for Threshold Autoregression, *Journal of Royal Statistical Society, B*, 52, 469-479.
- Clemen, R.T. (1989), Combining Forecasts: a Review and Annotated Bibliography, *International Journal of Forecasting*, 5, 559-583.
- Clements, M.P. and D.F. Hendry, (1997), 'An Empirical Study of Seasonal Unit Roots in Forecasting', *International Journal of Forecasting*, 13, 341-56.
- Clements, M. P. e Smith, J. (1999), A Monte Carlo study of the forecasting performance of empirical SETAR models, *Journal of Applied Econometrics*, 14, 123-141.
- Davies, R.B. (1977), Hypothesis Testing when a Nuisance Parameter is Present Only Under the Alternative, *Biometrika*, 64, 247-254
- Enders, W. e C.W.J. Granger (1998), Unit-root Tests and Asymmetric Adjustment with an Example Using the Term Structure of Interest Rates, *Journal of Business and Economic Statistics* 16, 304-311.
- Franses, P.H. (1996), *Periodicity and Stochastic Trends in Economic Time Series*, Oxford: Oxford University Press.
- Franses, P.H. e D. van Dijk (2000), *Nonlinear Time Series Models in Empirical Finance*, Cambridge: Cambridge University Press.
- Gouveia, P.M.D.C. (2000) *A No-Linearidade em Variveis Econmicas: Os Modelos SETAR*, Dissertao de Mestrado; Faculdade de Economia da Universidade do Algarve.
- Gouveia, P.M.D.C. e P.M.M. Rodrigues (2002), Cointegrao em Modelos SETAR – Uma Aplicao  Teoria da Paridade dos Poderes de Compra in Carvalho, L., Brilhante, F. e Rosado, F. (eds.) *Novos Rumos em Estatística – Actas do IX Congresso Anual da Sociedade Portuguesa de Estatística*, 273-284.
- Gouveia, P.M.D.C. E. L. Rebelo e P.M.M. Rodrigues (2003), Assimetria no Mecanismo de Taxas de Cmbio do SME, in Brito, P, Figueiredo, A., Sousa, F., Teles, P., Rosado, F. (eds.) *Literacia e Estatística, - Actas do X Congresso Anual da Sociedade Portuguesa de Estatística, Edioes SPE*, 325-336.
- Gouveia, P.M.D.C.B. e P.M.M. Rodrigues, P.M.M. (2004), Threshold Cointegration and the P.P.P. Hypothesis, *Journal of Applied Statistics*, 31, 115-127.

- Ghysels, E. (1994), 'On Economics and Econometrics of Seasonality' in C.A. Sims (ed.), *Advances in Econometrics - Sixth World Congress of the Econometric Society*, Cambridge: Cambridge University Press, Vol. 1, Chap. 7: 257-316.
- Ghysels, E., D.R. Osborn e P.M.M. Rodrigues (2000) Seasonal Nonstationarity and Near-nonstationarity. In *A Companion to Theoretical Econometrics*, Ed. Badi Baltagi, Blackwells, 655-677.
- Hamilton, J. D. (1994), *Time Series Analysis*, Princeton, NJ: Princeton University Press.
- Hansen, B.E. (1996), Inference when a Nuisance Parameter is Not Identified Under the Null Hypothesis, *Econometrica*, 64, 413-430.
- Hansen, B.E. (1997), Inference in TAR Models, *Studies in Non-linear Dynamics and Econometrics*, 2, 1-14.
- Hylleberg, S., R.F. Engle, Engle, C.W.J. Granger, and B. S. Yoo (1990), Seasonal Intergration and Cointegration, *Journal of Econometrics* 69, 5-25.
- Kennedy, E. (1999), 'Seasonality in Irish Tourism, 1973-1995', *Tourism Economics*, 5: 25-47.
- Kulendran, N. e Witt, S.F. (2001). Cointegration Versus Least Squares Regression. *Annals of Tourism Research*, 28.
- Rodrigues, P.M.M. e P.M.D.C. Gouveia (2001), A Note on Residual Based Tests for Threshold Cointegration, Documento de Trabalho, Faculdade de Economia, Universidade do Algarve.
- Rodrigues, P.M.M. e P.M.D.C. Gouveia (2002), Avaliação de Testes de Raízes Unitárias em Modelos SETAR, *Actas do IV Encontro de Economistas de Língua Portuguesa*.
- Rodrigues, P.M.M. e P.M.D.C. Gouveia (2004), An Application of PAR Models for Tourism Forecasting, *Tourism Economics*, no prelo.
- Rodrigues, P.M.M. e A.M. R. Taylor (2004) Asymptotic Distributions for Regression-based Seasonal Unit Root Tests in a Near-Integrated Model. *Econometric Theory*, no prelo.
- Stock, J.H. e M.W. Watson (1999), A Comparison of Linear and Nonlinear Univariate Models for Forecasting Macroeconomic Time Series, Ch. 1 in R. Engle and H. White (eds.), *Cointegration, Causality and Forecasting: A Festschrift for Clive W.J. Granger*. Oxford: Oxford University Press, 1-44.