

**UNIVERSIDADE ABERTA**



UNIVERSIDADE  
**AbERTA**  
[www.uab.pt](http://www.uab.pt)

**Modelação de Séries Temporais  
de Dados Oceanográficos**

**Bruno Vasconcelos Oliveira Monteiro**

**Mestrado em Bioestatística e Biometria**

2017



**UNIVERSIDADE ABERTA**



UNIVERSIDADE  
**AbERTA**  
[www.uab.pt](http://www.uab.pt)

**Modelação de Séries Temporais  
de Dados Oceanográficos**

**Bruno Vasconcelos Oliveira Monteiro**

**Mestrado em Bioestatística e Biometria**

**Dissertação de mestrado orientada pela  
Prof. Doutora Clara Cordeiro  
e coorientado pela  
Prof. Doutora Maria do Rosário Ramos**


2017



### Resumo

O aquecimento global, e em particular o dos oceanos, é um assunto de extrema importância, atendendo à atualidade das alterações climáticas e da consequente necessidade de adaptação das populações face a este fenómeno. As séries temporais de variáveis físicas, como por exemplo a temperatura à superfície do mar (SST), permitem estudar e interpretar estes fenómenos ao procurar modelos matemáticos para os descrever e prever, quando possível.

Nesta dissertação são utilizadas técnicas estatísticas de análise de séries temporais no estudo da SST. Foram selecionadas latitudes representativas das regiões Norte, Centro e Sul de Portugal, e fixadas a duas localizações geográficas, uma próxima da costa e outra distante da costa. É dada ênfase ao fenómeno do afloramento costeiro e foi proposto um índice de afloramento costeiro relativo (IAR), para as localizações geográficas com a mesma latitude.

Para cada região, foram explorados modelos de séries temporais para a série IAR, nos quais se incluem os modelos de Decomposição Clássica e de Holt-Winters, e ainda os modelos ARIMA sobre a série residual, resultante do ajustamento de uma tendência linear e uma componente sazonal à série temporal inicial. A seleção do melhor modelo foi auxiliada por medidas de ajustamento disponíveis para este tipo de análise. Toda a componente computacional deste trabalho foi realizada com recurso ao programa estatístico .

Foi identificado um padrão sazonal de doze meses para as três regiões e foi ainda possível observar o efeito do afloramento costeiro, mais evidente nos meses mais quentes do ano, em que, devido a este fenómeno, o aumento da SST é mais evidente nas localizações distantes da costa relativamente às localizações costeiras com a mesma latitude.

Por fim são apontadas algumas direções para a continuidade deste estudo.


Palavras-chave: Afloramento Costeiro; Modelos de Alisamento Exponencial; Modelos ARIMA; Modelos de Decomposição Clássica; Temperatura à Superfície do Mar.



### Abstract

Global warming, in particular that of the oceans, is a matter of extreme importance, taking into account the current climate change and the consequent adaptation of the population due to this phenomenon. Time series of physical variables, such as the sea surface temperature (SST), allow to study and interpret these phenomena by looking for mathematical models to describe them and to predict them, whenever possible.

In this dissertation, statistical techniques of time series analysis are used in the study of the SST. Latitudes were selected to represent the North, Centre and South regions of Portugal, and two geographic points were established, one near the coast and another farther from the coast. Emphasis is given to the phenomenon of upwelling and a relative upwelling index (IAR) has been proposed, based on the geographic points at the same latitude.

For each region, time series models are explored in the IAR series, namely the Classical Decomposition and Holt-Winters models, as well as the ARIMA models of the residual series, after the linear trend and seasonal component is fitted to the original time series. The selection of the best model was aided by the accuracy measures available for this type of analysis. All the computational work was performed with software .

A seasonal pattern of twelve months was identified in three regions and it was possible to observe the upwelling effect, most evident in the warmest months of the year, where, because of this effect, the increase in SST is more evident in distant coastal locations relative to coastal locations at the same latitude.

Finally, some directions are pointed out for the continuity of this study.

Keywords: *ARIMA* Models; Classic Decomposition Models; Exponential Smoothing Models; Sea Surface Temperature; Upwelling.



**Ao José Manuel pelo seu apoio incondicional.**

**À Paula e à Joseane pelo carinho e suporte.**

**À Serena pela inspiração nesta jornada.**



### Agradecimentos

Os meus primeiros agradecimentos vão para as minhas mães e para o José Manuel que me acompanharam nesta jornada com amor, dedicação e cuidado. E ainda pela vossa paciência durante estes dois anos em que me dediquei quase exclusivamente a este novo caminho.

Agradeço ao meu pai pelo exemplo e conhecimento no âmbito da pesquisa científica.

Ao meu irmão e cunhada pela ajuda constante e apoio incondicional.

Aos meus sobrinhos por trazerem a ideia de um futuro melhor e a vontade de renovar sempre.

À minha Serena que partiu das nossas vidas, mas que sempre servirá de inspiração no caminho do bem, do amor e da dedicação.

Aos meus queridos amigos e familiares. À Shirley, Martinha, Andreia, Mami, Isabel, Mário entre outros, que carinhosamente ofereceram a sua compreensão pela minha falta de tempo.

Às minhas orientadoras, Prof. Doutora Clara Cordeiro e Prof. Doutora Maria do Rosário Ramos, que me dedicaram muitas horas e paciência durante este ano. E ainda pelas sugestões e disponibilidade constantes que contribuíram para a elaboração coerente deste trabalho.

E finalmente, à Universidade Aberta por oferecer os meios necessários à minha progressão académica e pessoal.




## Índice

<b>Capítulo I - Introdução .....</b>	<b>1</b>
1. Introdução .....	3
<b>Capítulo II - Séries Temporais .....</b>	<b>5</b>
2.1. Introdução.....	7
2.1.1. O conceito de série temporal .....	7
2.1.2. Componentes de uma série temporal .....	8
2.1.3. Os objetivos do estudo de uma série temporal .....	13
2.2. Conceitos fundamentais em séries temporais .....	13
2.2.1. Processo estocástico .....	13
2.2.2. Processo estacionário .....	14
2.2.3. Funções de autocovariância, autocorrelação e autocorrelação parcial.....	15
2.2.4. Exemplos de processos estocásticos .....	18
2.3. Metodologia utilizada na análise de uma série temporal.....	23
2.3.1. Decomposição clássica .....	24
2.3.2. Modelos de alisamento exponencial .....	25
2.3.3. Modelos <i>ARIMA</i> .....	28
2.4. Critérios e medidas para a escolha do modelo .....	29
2.4.1. Critérios AIC, AICc e BIC.....	30
2.4.2. Medidas de erro .....	31
2.5. Testes de hipóteses.....	31
2.5.1. Testes de independência.....	32
2.5.2. Teste de raízes unitárias.....	33
<b>Capítulo III - Estudo de Séries Temporais de Dados Oceanográficos .....</b>	<b>37</b>
3.1. Introdução.....	39

## Modelação de Séries Temporais de Dados Oceanográficos

---

3.2	A temperatura à superfície do mar .....	41
3.2.1	Introdução .....	41
3.2.2	As regiões em estudo .....	41
3.2.3	Análise da SST por região e localização .....	42
3.3	O fenómeno do Afloramento.....	46
3.3.1	Introdução .....	46
3.3.2	Índice de Afloramento Relativo .....	47
3.3.3	A análise do IAR para cada região .....	48
3.3.4	Modelação da IAR de cada região .....	51
3.3.5	Interpretação dos resultados com base no melhor modelo .....	61
3.4	Discussão.....	63
<b>Capítulo IV - Conclusão e Trabalhos Futuros.....</b>		<b>65</b>
4.	Conclusão e trabalhos futuros .....	67
<b>Anexos .....</b>		<b>69</b>
Anexo I – Dados Iniciais.....		71
Anexo II – Comandos do  .....		77
<b>Bibliografia .....</b>		<b>83</b>

### Índice de figuras

Figura 2.1: Atividade elétrica cardíaca de uma pessoa saudável.....	7
Figura 2.2: Número anual de passageiros do Aeroporto de Heathrow, Reino Unido. ....	7
Figura 2.3: Concentração de dióxido de carbono (CO <sub>2</sub> ).....	8
Figura 2.4: Variação da extensão de gelo do mar Ártico para o mês de junho. ....	8
Figura 2.5: Temperatura superficial do mar registada pela Boia de Faro Costeira. ....	10
Figura 2.6: Consumo trimestral de gás no Reino Unido.....	11
Figura 2.7: Número lince capturados em armadilhas anualmente no Canadá. ....	12
Figura 2.8: Exemplos de correlograma para um processo puramente aleatório (à esquerda), e uma série temporal com sazonalidade (à direita).....	17
Figura 3.1: Mapa de Portugal com as seis localizações em estudo. ....	41
Figura 3.2: Representação gráfica da SST para as regiões (a) Norte, (b) Centro e (c) Sul. ....	43
Figura 3.3: Representações gráficas da ACF e PACF. ....	44
Figura 3.4: Caixa de bigodes da SST para as regiões (a) Norte, (b) Centro e (c) Sul. ....	45
Figura 3.5: Afloramento costeiro no hemisfério Norte.....	46
Figura 3.6: Representação gráfica do IAR.....	48
Figura 3.7: Representação da ACF e PACF da IAR.....	49
Figura 3.8: Caixa de bigodes do IAR. ....	50
Figura 3.9: Decomposição clássica para a séries temporais IAR. ....	52
Figura 3.10: ACF e PACF da componente irregular. ....	53
Figura 3.11: Representação gráfica dos modelos de Holt-Winters do tipo aditivo (a preto) e do tipo multiplicativo (a vermelho) e das séries temporais do IAR das regiões Norte, Centro e Sul. ....	54
Figura 3.12: ACF e PACF dos resíduos. ....	56
Figura 3.13: Representação gráfica do IAR e dos respetivos modelos da tendência linear+índices sazonais (a preto).....	57

## Modelação de Séries Temporais de Dados Oceanográficos

---

Figura 3.14: ACF e PACF das séries temporais $IAR^{-T-S}$ .....	58
Figura 3.15: Gráficos dos resíduos, ACF, Papel de Probabilidade e valores-p do teste de Ljung-Box. ....	60
Figura 3.16: Gráficos dos índices sazonais.....	62

**Índice de quadros**

Quadro 3.1 – Valores da ACF para  $k$  igual a 6, 12, 18 e 24.....44

Quadro 3.2 – Índices sazonais dos modelos de decomposição aditivo e multiplicativo. ....53

Quadro 3.3 – Resultados do alisamento exponencial de Holt – Winters.....55

Quadro 3.4 – Declive e termo independente do modelo de regressão linear.....56

Quadro 3.5 – Índices Sazonais do modelo de regressão linear.....57

Quadro 3.6 – Critérios de seleção AIC, AICc e BIC dos modelos ARIMA. ....59

Quadro 3.7 – Parâmetros estimadas para os modelos  $AR(1)$ .....59

Quadro 3.8 – MAE dos modelos. ....61



### Lista geral de abreviaturas, siglas e acrónimos

#### *Área das ciências do mar*

GOES - Satélites Geoestacionários de Órbita Terrestre (do inglês *Geostationary Orbiting Earth Satellites*)

IH - Instituto Hidrográfico

MODIS - Espectrómetro de Imagem de Moderada Resolução (do inglês *Moderate Resolution Imaging Spectroradiometer*)

NASA - Administração Nacional da Aeronáutica e do Espaço (do inglês *National Aeronautic and Space Administration*)

NOAA - Administração Nacional dos Oceanos e da Atmosfera (do inglês *National Oceanic and Atmospheric Administration*)

NSIDC - Centro Nacional de Observações da Neve e do Gelo dos Estados Unidos (do inglês *National Snow and Ice Data Center*)

SST - Temperatura à Superfície do Mar (do inglês *Sea Surface Temperature*)

#### *Área de séries temporais*

ACF - Função de Autocorrelação (do inglês *Autocorrelation Function*)

ADF - Dickey-Fuller Aumentado (do inglês *Augmented Dickey-Fuller*)

AIC - Critério de Informação de Akaike (do inglês *Akaike's Information Criterion*)

AICc - Critério de Informação de Akaike Corrigido (do inglês *Corrected Akaike's Information Criterion*)

AR - Processo Autorregressivo (do inglês *Autoregressive*)

ARIMA - Processo Autorregressivo e de Médias Móveis integrado (do inglês *Autoregressive Integrated Moving Average*)

BIC - Critério de Informação Bayesiano (do inglês *Bayesian Information Criterion*)

DF - Dickey-Fuller

IA - Índice de Afloramento

IAR - Índice de Afloramento Relativo

MA - Processo de Média Móvel (do inglês *Moving Average*)

MAE - Erro Absoluto Médio (do inglês *Mean Absolute Error*)

MAPE - Erro Percentual Absoluto Médio (do inglês *Mean Absolute Percentage Error*)

MdAE - Erro Absoluto Mediano (*Median Absolute Error*)

MSE - Erro Quadrático Médio (do inglês *Mean Square Error*)

PACF - Função de Autocorrelação Parcial (do inglês *Partial Autocorrelation Function*)

RMSE - Desvio Padrão Empírico (Amostral) Generalizado (do inglês *Root Mean Square Error*)

SARIMA - Processo Autorregressivo e de Médias Móveis integrado Sazonal (do inglês *Seasonal Autoregressive Integrated Moving Average*)

SES - Alisamento Exponencial Simples (do inglês *Simple Exponential Smoothing*)

SMA - Médias Móveis Simples (do inglês *Simple Moving Average*)

$C_t$  - Componente Cíclica no instante  $t$

$I_t$  - Componente Irregular no instante  $t$

$S_t$  - Componente Sazonal no instante  $t$

$T_t$  - Componente Tendência no instante  $t$

$h$  - Horizonte de previsão

$k$  - Desfasamento entre dois instantes de uma série temporal

$\gamma_k$  - Função de autocovariância para desfasamento  $k$

$\rho_k$  - Função de autocorrelação para desfasamento  $k$

$\phi_{kk}$  - Função de autocorrelação parcial para desfasamento  $k$

$\theta$  - Parâmetro de médias móveis

$\phi$  - Parâmetro autorregressivo

$B$  - Operador de atraso

$\Theta(B)$  - Polinómio autorregressivo

$\Phi(B)$  - Polinómio de médias móveis

$b_t$  - Taxa de crescimento no instante  $t$

$s_t$  - Índice de sazonalidade do instante  $t$

$l_t$  - Nível médio no instante  $t$

$\alpha$  - Constante de alisamento associada ao nível médio

$\beta$  - Constante de alisamento associada à tendência

$\gamma$  - Constante de alisamento associada à sazonalidade

$d$  - Ordem que representa o número de diferenças necessárias para obter uma série estacionária

$D$  - Ordem que representa o número de diferenças sazonais necessárias para obter uma série estacionária relativamente à sazonalidade

$p$  - Ordem associada ao processo autorregressivo

$P$  - Ordem associada à parte sazonal do processo autorregressivo

$q$  - Ordem associada ao processo de médias móveis

$Q$  - Ordem associada à parte sazonal do processo de médias móveis



**Capítulo I - Introdução**



### 1. Introdução

A análise de séries temporais é de extrema importância para a compreensão da evolução ao longo do tempo de características e mais genericamente de fenómenos que se enquadram numa grande diversidade de áreas do conhecimento. Uma destas áreas é a oceanografia que beneficia das técnicas de análise de séries temporais que foram sendo desenvolvidas por investigadores, na medida em que possibilitam, através de uma diversidade de modelos, descrever o comportamento de fenómenos neste domínio.

As camadas superficiais dos oceanos têm um papel importante na manutenção da temperatura global, contudo, em várias regiões planetárias, a temperatura à superfície do mar (SST) tem sofrido um aumento significativo nas últimas décadas (Castro, 2003). O estudo deste aumento e das suas possíveis causas é muito relevante, assim como a identificação de modelos que permitam prever a sua evolução futura. Na atualidade é disponibilizada ao público uma grande quantidade de informação sobre as características físicas dos oceanos (temperatura, salinidade, nível do mar, entre outras) obtida *in situ* ou por satélites.

Nesta dissertação utilizaram-se técnicas de análise exploratória e modelos de séries temporais para descrever o comportamento da SST no horizonte temporal compreendido entre janeiro de 1982 e dezembro de 2013, nas regiões Norte, Centro e Sul de Portugal, registadas em duas localizações, uma mais perto, a menos de 20 km da costa, e outra mais afastada, entre 135 e 150 km da costa. Foi igualmente objeto de interesse o fenómeno denominado por afloramento costeiro (é frequente o uso do termo em inglês, *Upwelling*), que é recorrente nas localizações mais próximas da costa portuguesa e responsável por temperaturas superficiais do mar mais baixas que o esperado nos meses mais quentes do ano. Neste contexto foi proposta uma medida comparativa, o Índice de Afloramento Relativo (IAR) que possibilita uma interpretação em termos relativos, da variação da temperatura devida a este fenómeno.

Os dados da SST das seis localizações geográficas são dados de satélite, fornecidos pelo *E.U. Copernicus Marine Service Information*.

Os **objetivos principais** deste trabalho foram:


- Descrever a evolução da SST ao longo do horizonte temporal em estudo, nas suas componentes mais elementares;

## Capítulo I - Introdução

---

- Estudar o fenómeno do afloramento costeiro nas três regiões através da análise de uma medida proposta neste trabalho, o Índice de Afloramento Relativo (*IAR*) de cada região;
- Identificar de entre vários modelos o que melhor se ajusta aos dados do *IAR* para cada uma das regiões, recorrendo a medidas de avaliação da qualidade do ajustamento.
- Confrontar os resultados obtidos com este índice com outros estudos já publicados.

No **capítulo 2** serão introduzidos os principais conceitos e propriedades associados à análise de séries temporais, bem como os modelos que são utilizados posteriormente na componente prática, da análise de dados. Apresentam-se os modelos de decomposição clássica, modelos de alisamento exponencial e a metodologia de Box-Jenkins, os modelos ARIMA. São apresentadas ainda as metodologias utilizadas para a descrição inicial dos dados e para o estudo e avaliação da qualidade dos modelos ajustados. Descrevem-se as funções de autocorrelação e autocorrelação parcial, testes para a autocorrelação, estacionaridade e teste de raízes unitárias. Em relação às medidas de ajustamento que apoiam a seleção do modelo, estas incluem os critérios de Informação de Akaike (AIC), de Akaike corrigido (AICc), bayesiano (BIC) e medidas de erro, como o erro absoluto médio (MAE).

No **capítulo 3** é realizada a análise de dados propriamente dita. É feito um estudo exploratório para ilustrar e descrever as seis séries temporais da SST relativas à combinação região *vs* distância à costa. Na secção seguinte é aprofundado o estudo do *IAR*. São ajustados diversos modelos para descrever o comportamento deste índice, sendo escolhido aquele que apresenta os melhores valores das medidas de ajustamento. Esta procura de modelo é realizada para o *IAR* calculado nas três regiões. A análise e tratamento dos dados foi realizado com recurso ao programa estatístico .

O **capítulo 4** apresenta uma síntese dos resultados e as considerações mais relevantes sobre o trabalho realizado. São ainda indicados alguns pontos em aberto e sugestões para trabalhos futuros.

**Capítulo II - Séries Temporais**

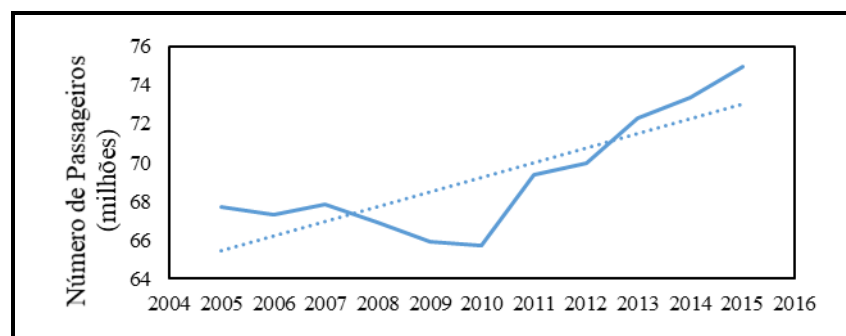
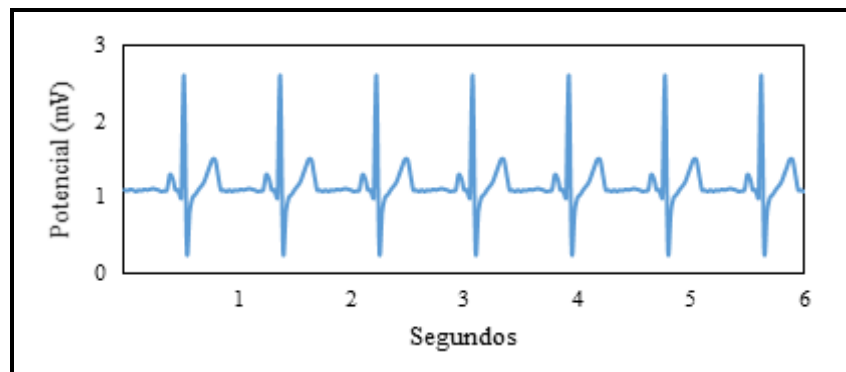


### 2.1. Introdução

#### 2.1.1. O conceito de série temporal

Define-se como série temporal a sequência cronológica de observações de uma variável de interesse (Montgomery *et al.*, 2008).

A série temporal pode ser **contínua** ou **discreta**. É contínua quando o valor da variável é registado continuamente no tempo. O registo dos batimentos cardíacos (Figura 2.1) é um exemplo de série temporal contínua. Diz-se que é uma série temporal discreta quando o valor da variável é registado em intervalos de tempo regulares. O número anual de passageiros no Aeroporto de Heathrow (Figura 2.2) é um exemplo de série temporal discreta.



As séries temporais contínuas podem ser transformadas em discretas por um processo de discretização, passando a ser utilizados na análise os valores observados em intervalos tempo regulares (minutos, dias, meses, anos, etc.).

<sup>1</sup> Fonte: Autor.

<sup>2</sup> Fonte: Autor a partir de dados obtidos em <http://www.heathrow.com/company/investor-centre/results-and-performance/traffic-statistics>.

### 2.1.2. Componentes de uma série temporal

Quando se pretende descrever o comportamento de uma série temporal o processo mais comum, é decompor a série temporal em componentes de natureza distinta: a componente **tendência**, a componente **sazonal**, a componente **cíclica** e a componente **irregular** (ou **residual**).

#### Componente tendência

A *componente tendência* ( $T_t$ ,  $t = 1, 2, \dots, n$ ) pode ser interpretada, como o aumento (Figura 2.3), ou a diminuição (Figura 2.4), de médio ou longo prazo, dos valores médios de uma série temporal (Chatfield, 2003).

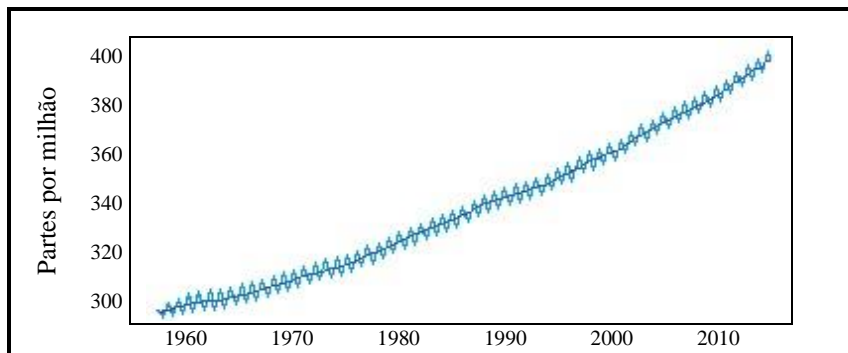


Figura 2.3: Concentração de dióxido de carbono ( $CO_2$ ).<sup>3</sup>

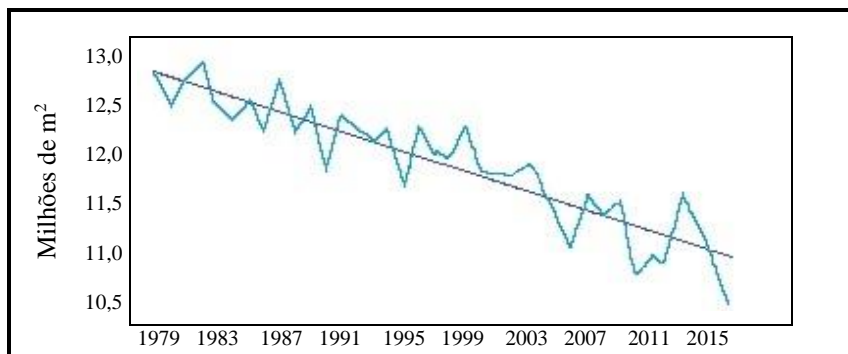


Figura 2.4: Variação da extensão de gelo do mar Ártico para o mês de junho.<sup>4</sup>

Na Figura 2.3 pode ser observado o aumento progressivo da concentração de  $CO_2$  na atmosfera ao longo dos anos, ou seja, estamos na presença de uma série temporal com uma tendência crescente. Na Figura 2.4 observa-se uma diminuição da extensão do gelo no mar

<sup>3</sup> Fonte: NOAA - Earth System Research Laboratory Global Monitoring Division - Mauna Loa Observatory. <http://www.esrl.noaa.gov/gmd/obop/mlo/> .

<sup>4</sup> Fonte: National Snow and Ice Data Centre (NSIDC). <https://nsidc.org/arcticseaicenews/category/analysis/>.

Ártico no mês de junho entre os anos de 1979 e 2016, indicando a existência de uma tendência decrescente neste período.

A componente tendência pode seguir um comportamento linear, exponencial ou de outro tipo. A existência de uma tendência pode ser consequência direta de diferentes fatores, como nos exemplos anteriores, o aumento da concentração média do CO<sub>2</sub> ao longo dos anos pode ser causado pelo aumento do consumo de combustíveis fósseis, assim como a diminuição da extensão média do gelo no mar Ártico ao longo dos anos pode ser uma consequência direta do aquecimento global.

### *Estimação*

Esta componente pode ser estimada pelo **método das médias móveis** ou suavização usado para obter uma série temporal sem esta componente e que será utilizado no método de modelação *ARIMA* (explicado posteriormente). Este consiste em estimar  $T_t$  através de uma sucessão de médias móveis, em que cada observação é substituída pela média dela própria com as que lhe estão mais próximas. Para uma série temporal  $X_1, X_2, \dots, X_n$  e com  $K$  que pode ser qualquer número inteiro obtemos a sucessão de médias móveis ( $\bar{X}_t$ ) associadas, em que  $d + 1 \leq t \leq n - d$ :

$$\begin{cases} \bar{X}_t = \frac{1}{K} \sum_{j=-d}^d X_{t+j} & \text{Se } K \text{ for ímpar e } K = 2d + 1, \\ \bar{X}_t = \frac{1}{K} \left[ \left( \frac{X_{t-d}}{2} \right) + \sum_{j=1-d}^{d-1} X_{t+j} + \left( \frac{X_{t+d}}{2} \right) \right] & \text{Se } K \text{ for par e } K = 2d. \end{cases} \quad (1)$$

A nova série temporal  $\bar{X}_t$  possui uma variância menor que a série temporal original. Para que este procedimento seja esclarecedor é importante que a tendência  $T_t$  seja uma função em que os valores próximos possam ser considerados iguais, ou seja, que varie lentamente (como seria de esperar para as variáveis consideradas neste estudo) (Alpuim, 2003). Esta técnica será aplicada para estimar a tendência dos modelos de decomposição clássica que serão descritos posteriormente.

Na presença de uma tendência linear (com base na análise gráfica das observações) considera-se um modelo de regressão linear. A componente é estimada pelo **método dos mínimos quadrados** que em certas condições fornece estimadores lineares centrados.

A tendência linear estimada será:

$$\hat{T}_t = b_0 + b_1 t, \quad (2)$$

em que o coeficiente  $b_0$  e  $b_1$  são estimados pelo método dos mínimos quadrados (Alpuim, 2003). Este método será utilizado para estimar e extrair a tendência linear das séries temporais antes de se proceder à modelação dos modelos do tipo *ARIMA*, explicados posteriormente neste trabalho.

### Componente sazonal

A *componente sazonal* ( $S_t, t = 1, 2, \dots, n$ ) pode ser definida pelas mudanças dos valores de uma série temporal que seguem um padrão, assumindo valores semelhantes em intervalos regulares de tempo, geralmente iguais, ou inferiores, a um ano (Figura 2.5).

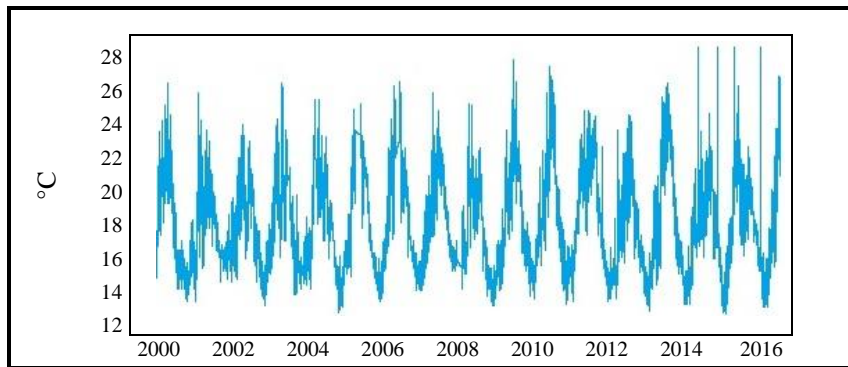


Figura 2.5: Temperatura superficial do mar registrada pela Bóia de Faro Costeira.<sup>5</sup>

Na Figura 2.5 apresenta-se a série temporal da temperatura do mar à superfície, sendo os dados registados a cada 10 minutos pela Boia de Faro Costeira, entre abril de 2000 e julho de 2016. Pode-se observar um comportamento periódico, em que os valores da série temporal se comportam de forma análoga (de doze em doze meses), com temperaturas mais baixas no inverno e mais altas no verão. A repetição deste padrão é um indicativo de que a série temporal apresenta um padrão sazonal.

Dois tipos distintos da componente sazonal podem ser identificados: o **tipo aditivo**, em que a amplitude sazonal não varia de acordo com o valor médio dos dados; o **tipo multiplicativo**, em que a amplitude sazonal varia proporcionalmente ao aumento, ou diminuição, do valor médio dos dados (Figura 2.6) (Ehlers, 2009).

---

<sup>5</sup> Fonte: Instituto Hidrográfico (IH). <http://www.hidrografico.pt/boias-ondografo.php>.

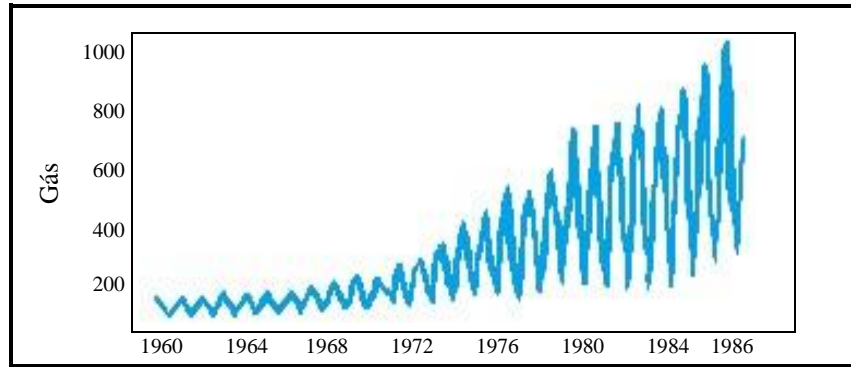


Figura 2.6: Consumo trimestral de gás no Reino Unido.<sup>6</sup>

A Figura 2.6 apresenta o consumo de gás no Reino Unido entre o primeiro trimestre de 1960 e o quarto trimestre de 1986. A análise do gráfico da série temporal, sugere que a componente sazonal seja do tipo multiplicativo, pois a amplitude sazonal aumenta proporcionalmente ao aumento do valor médio do consumo de gás (verificando-se também a presença de uma tendência crescente).

### *Estimação*

A componente sazonal  $S_t$  é uma função periódica com período  $s$ , ou seja,  $S_{t+s} = S_t$ . Os **Índices Sazonais**  $s_i$  ( $i = 1, 2, \dots, s$ ) são os  $s$  diferentes valores estimados da função  $S_t$  e o seu somatório é nulo. Estes são estimados pelo cálculo da média dos elementos pertencentes a cada fase:

$$\hat{s}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} X_{(j-1)s+i}, \quad (3)$$

com  $i = 1, 2, \dots, s$  e  $n_i$  é o número de observações para cada período (Alpuim, 2003).

### **Componente cíclica**

A **componente cíclica** ( $C_t$ ,  $t = 1, 2, \dots, n$ ) pode ser definida pelas mudanças dos valores médios das observações de uma série temporal que seguem um padrão e assumem valores semelhantes em intervalos não regulares de tempo (geralmente com períodos superiores a dois anos) (Hyndman e Athanasopoulos, 2013). A Figura 2.7 é um exemplo de uma série temporal com este tipo de padrão.

<sup>6</sup> Fonte: Ehlers - Análise de Séries Temporais 2009, capítulo 1, página 5, figura (d).

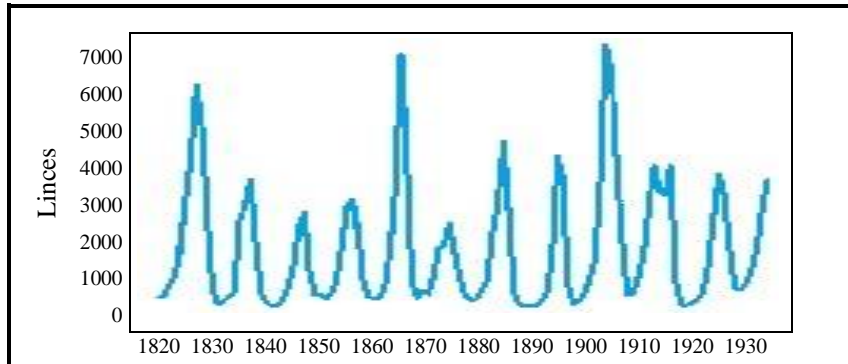


Figura 2.7: Número lince capturados em armadilhas anualmente no Canadá.<sup>7</sup>

Na Figura 2.7 pode ser observada oscilação cíclica do número de lince capturados no Canadá entre 1821 e 1934. Segundo Stenseth *et al.* (1998), o número de lince capturados corresponde entre 5% a 9% da população total, e a oscilação populacional está associada à relação predador-presa entre esta espécie e a população de lebres do qual se alimenta, com uma oscilação populacional alternada por um aumento da população, com período de aproximadamente 6 anos, seguido por um declínio populacional nos 4 anos consecutivos. A interpretação do gráfico sugere a presença de uma oscilação periódica de 10 anos para a população de lince no Canadá. Estes valores para a periodicidade cíclica foram relatados em vários estudos destas séries temporais, sendo o primeiro trabalho desta análise realizado por Moran (1953).

A componente cíclica pode ser confundida com a componente sazonal, embora ambas tenham diferentes características. A componente sazonal tem um período fixo, já a componente cíclica é composta por flutuações que não estão associadas a um período fixo e geralmente com ciclos mais longos com uma duração superior aos dois anos (Hyndman e Athanasopoulos, 2013).

### Componente irregular

A **componente irregular** ( $I_t$ ,  $t = 1, 2, \dots, n$ ), também designada por componente residual, pode ser definida pelas mudanças aleatórias não explicadas pelas outras componentes. Parte importante da análise de séries temporais centra-se na estimação desta componente com o auxílio de modelos de séries temporais. Esta componente pode ser estimada removendo as outras componentes (estimadas por métodos estatísticos) da série original (Chatfield, 2003).

<sup>7</sup> Fonte: Ehlers - Análise de Séries Temporais 2009, capítulo 1, página 5, figura (b).

### 2.1.3. Os objetivos do estudo de uma série temporal

O principal objetivo da análise de séries temporais é o desenvolvimento de modelos matemáticos que expliquem de uma maneira plausível os dados de um determinado fenómeno (Shumway e Stoffer, 2014).

Segundo Chatfield (2003) os objetivos gerais da análise de séries temporais podem ser classificados como:

- Descritivos – Utilização de gráficos e técnicas de modelação que permitam descrever o processo em estudo e as suas diferentes componentes, assim como encontrar *outliers* que possam influenciar a análise.
- Explicativos – Encontrar processos de causa e efeito entre duas ou mais séries temporais, em que a variação de uma série temporal explique a variação das outras. Permitindo conhecer o mecanismo que gera as variações de uma série temporal.
- Preditivos – Prever futuros valores para uma variável a partir da análise e modelação de uma série temporal.
- De controlo – Aumentar o controlo sobre os valores de uma determinada variável, ou seja, compreender os mecanismos e parâmetros de uma variável de forma a poder manipulá-los para que os seus valores futuros se aproximem dos valores desejados.

A análise dos dados permite compreender o comportamento do fenómeno em estudo, assim como prever valores futuros com base neste conhecimento. A maioria dos valores observados estão sujeitos a perturbações de natureza aleatória e o seu estudo implica geralmente a utilização de técnicas estatísticas.

## 2.2. Conceitos fundamentais em séries temporais

### 2.2.1. Processo estocástico

Define-se como *processo estocástico* uma família de variáveis aleatórias que variam no tempo (ou em outras dimensões como o espaço) definida por  $\{X_t, t \in \mathbf{T}\}$ , em que  $X_t$  é uma variável aleatória (ou um conjunto ou família de variáveis aleatórias) com contradomínio denominado por Espaço de Estados e representado por  $S$ .  $\mathbf{T}$  é o conjunto de índices ordenados, denominado por Espaço de Parâmetros (Alpuim, 2003).

Podemos formalmente definir uma série temporal como o conjunto de observações de um processo estocástico  $\{X_t: t \in \mathbf{T}\}$ , sendo  $\mathbf{T} = \mathbb{R}$ , para as séries temporais contínuas, ou  $\mathbf{T} = \{t_1, t_2, \dots, t_n\}$  com valores consecutivos de  $t$  equidistantes, para séries temporais discretas (Ehlers, 2009).

### 2.2.2. Processo estacionário

Um *processo estacionário* é um processo estocástico, em que os valores da variável aleatória distribuem-se no tempo em torno de um valor médio constante e é presumida uma certa independência entre os valores observados em diferentes instantes  $t$ . Neste tipo de processos a média e a variância são constantes, sendo assim não existe a necessidade de se conhecer o valor na origem da série (Alpuim, 2003).

Para um processo estacionário  $\{X_t, t \in \mathbf{T}\}$  com variância finita  $\forall t \in \mathbf{T}$  (Alpuim, 2003):

Valor médio  $\mu_t = E(X_t) ,$

Variância  $\sigma_t^2 = Var(X_t) ,$

Covariância  $\gamma_{t_1, t_2} = cov(X_{t_1}, X_{t_2}) = E[X_{t_1} X_{t_2}] - \mu_{t_1} \mu_{t_2} ,$

Correlação  $\rho_{t_1, t_2} = \frac{\gamma_{t_1, t_2}}{\sigma_{t_1} \sigma_{t_2}} = \frac{cov(X_{t_1}, X_{t_2})}{\sqrt{Var(X_{t_1}) Var(X_{t_2})}} .$

Segundo Shumway e Stoffer (2014), consideram-se dois tipos de estacionaridade, uma forte, em que os processos são denominados por estritamente estacionários, e outra fraca, em que os processos são denominados por estacionários para a covariância ou de 2ª ordem.

**Processos estritamente estacionários** são definidos como qualquer processo estocástico  $\{X_t, t \in \mathbf{T}\}$  em que o comportamento probabilístico de qualquer sequência de variáveis aleatórias  $\{X_{t_1}, X_{t_2}, \dots, X_{t_n}\}$  é idêntico ao de qualquer outra sequência  $\{X_{t_1+h}, X_{t_2+h}, \dots, X_{t_n+h}\}$ , com o mesmo número de elementos  $n$  e a distância temporal  $h = 0, \pm 1, \pm 2, \dots$ , entre as duas sequências. Portanto, um processo estacionário é estritamente estacionário se a distribuição de probabilidades conjuntas de  $X_{t_1}, \dots, X_{t_k}$  for igual à de  $X_{t_1+h}, \dots, X_{t_n+h}$  (Shumway e Stoffer, 2014).

Num processo estritamente estacionário, a média e a variância são constantes para  $\forall t \in \mathbf{T}$  e a covariância entre  $X_{t_1}$  e  $X_{t_2}$  é igual à covariância entre  $X_{t_1+h}$  e  $X_{t_2+h}$ , para todo o  $h$  (Alpuim, 2003).

A condição de estacionaridade restrita implica o conhecimento de todas as distribuições marginais, o que torna difícil ser verificada na prática, e além disso não é aplicável à maioria das séries temporais (Alpuim, 2003).

O conceito de *processos estacionários para a covariância (ou de 2ª ordem)* é mais conveniente para a análise de séries temporais. Um processo estocástico é considerado estacionário de 2ª ordem,  $\forall t \in \mathbf{T}$ , se satisfaz as seguintes condições:  $E(X_t^2) < +\infty$ , a média  $\mu_t = E(X_t)$  é constante e a autocovariância  $\gamma_{t_1, t_2}$  entre os instantes  $t_1$  e  $t_2$  é função de  $|t_2 - t_1|$ . Ambos, o valor médio e a autocovariância, são independentes do instante  $t$  (Shumway e Stoffer, 2014).

Para um processo estacionário de 2ª ordem temos:

- $Var(X_t) = \gamma_{(t,t)} = \gamma(0) = \sigma_t^2 = \sigma^2$ ,
- $\rho_{t_1, t_2} = \rho_{t_1+h, t_2+h} = \rho_{t_2-t_1}$ .

Muitas técnicas de análise de séries temporais baseiam-se nessa propriedade. É importante salientar que um processo estritamente estacionário com variância finita é também um processo estacionário de 2ª ordem.

### 2.2.3. Funções de autocovariância, autocorrelação e autocorrelação parcial

A *função de autocovariância* e a *função de autocorrelação* (ACF) são respetivamente as funções de covariância e de correlação entre realizações de um processo estocástico observadas em diferentes horizontes de tempo e desfasadas em  $k$  unidades de tempo (Ehlers, 2009).

A função de autocovariância em séries estacionárias pode ser reduzida a uma só função:

$$cov(X_t, X_{t+k}) = \gamma_{t, t+k} = \gamma_{|t-t+k|} = \gamma_k.$$

O estimador da função de autocovariância para um processo estacionário de 2ª ordem ( $\gamma_{t, t+k}$ ) é dado pela seguinte equação:

$$\hat{\gamma}_k = \frac{1}{n} \sum_{t=1}^{n-k} (X_t - \bar{X})(X_{t-k} - \bar{X}).$$

É aconselhado estimar o  $\hat{\gamma}_k$  para apenas os primeiros  $\frac{n}{4}$  valores de  $k$ , pois quanto maior é o valor de  $k$  maior será o desvio entre o valor estimado ( $\hat{\gamma}_k$ ) e o valor da função de autocovariância ( $\gamma_k$ ) (Alpuim, 2003).

A função de autocorrelação (ACF) também pode ser reduzida a uma única função para séries temporais com estacionaridade de pelo menos 2ª ordem:

$$\rho_k = \text{corr}(X_t, X_{t+k}) = \rho_{t,t+k} = \frac{\gamma_k}{\gamma_0}.$$

O estimador da função de autocorrelação para um processo estacionário de 2ª ordem ( $\rho_{t,t+k}$ ) é dado por:

$$\hat{\rho}_k = \frac{\hat{\gamma}_k}{\gamma_0} = \frac{\sum_{t=1}^{n-k} (X_t - \bar{X})(X_{t-k} - \bar{X})}{\sum_{t=1}^n (X_t - \bar{X})^2}. \quad (4)$$

Segundo Alpuim (2003) as funções de autocovariância e autocorrelação têm as seguintes propriedades:

$$\gamma_0 = \text{Var}(X_t).$$

$|\gamma_k| \leq \gamma_0$ , pois a desigualdade de Cauchy-Schwarz  $|E(XY)| \leq \sqrt{E(X^2)E(Y^2)}$  tem como consequência  $|\rho_k| \leq 1$

$$\gamma_k = \gamma_{-k}.$$

As funções  $\gamma_k$  e  $\rho_k$  são semidefinidas positivas.

A **função de autocorrelação parcial (PACF)** que pode ser expressa como a parte da autocorrelação para o desfasamento  $k$ , que não é explicada pela correlação de valores de  $k$  de ordem igual ou inferior a  $k - 1$ , ou seja, é a contribuição da correlação do desfasamento sem a participação dos coeficientes dos demais desfasamentos (Ehlers, 2009). A PACF,  $\phi_{kk}$ , é dada pela seguinte equação:

$$\rho_k = \sum_{i=1}^k \phi_{ki} \rho_{k-i} = \phi_{k1} \rho_{k-1} + \phi_{k2} \rho_{k-2} + \dots + \phi_{k(k-1)} \rho_1 + \phi_{kk}, k > 1.$$

Temos então que calcular o conjunto de parâmetros  $\phi_{ki}$  que satisfaz o seguinte sistema de equações:

$$\rho_j = \sum_{i=1}^k \phi_{ki} \rho_{j-i} = \phi_{k1} \rho_{j-1} + \phi_{k2} \rho_{j-2} + \dots + \phi_{kk} \rho_{j-k}, j = 1, 2, \dots, k. \quad (5)$$

A análise do correlograma, é uma ferramenta muito útil na identificação de várias características de uma série temporal pois permite verificar se esta possui um comportamento puramente aleatório ou se é possível identificar uma periodicidade sazonal. A interpretação da ACF para os primeiros desfasamentos, em conjunto com a PACF, permite tirar as seguintes conclusões: se a série temporal possui um comportamento puramente aleatório; se é não estacionária; se possui um padrão sazonal, e caso exista qual a sua periodicidade; existência de *outliers* (Ehlers, 2009). A autocorrelação  $\rho_k$  (Figura 2.8) assume valores entre 1 e -1 e as linhas tracejadas em azul representam os limites do intervalo de confiança ( $\pm 1,96/\sqrt{n}$ ) para o qual se podem considerar estatisticamente nulas as autocorrelações contidas no mesmo (Ehlers, 2009). Assim, utiliza-se um gráfico da ACF (Figura 2.8) com as correlações para diferentes valores de  $k$ , do qual se pode tirar conclusões sobre as características de uma série temporal.

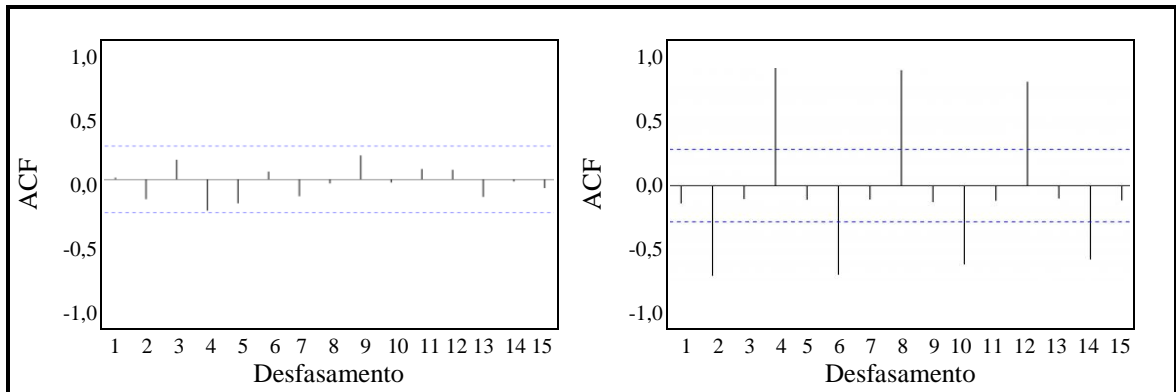


Figura 2.8: Exemplos de correlograma para um processo puramente aleatório (à esquerda), e uma série temporal com sazonalidade (à direita).<sup>8</sup>

Na Figura 2.8 temos dois exemplos de correlogramas: um processo puramente aleatório (correlograma à esquerda), para o qual a autocorrelação para todos os desfasamentos  $k = 1, 2, \dots, 15$  é próxima de 0; uma série temporal da produção trimestral de cerveja na Austrália (correlograma à direita), que sugere a presença de uma sazonalidade, sendo  $s = 4$ . A sazonalidade é identificada pela correlação muito forte (próxima de 1) para o desfasamento

<sup>8</sup> Fonte: Hyndman e Athanasopoulos (2013).

$k = 4$ , o que indica que valores distantes anualmente são muito próximos, e pela oscilação periódica da correlação para diferentes defasamentos  $k$ . Os picos de correlações positivas são observados para defasamentos  $k = 4 \times i$  e  $i \in \mathbb{N}$  e os de correlações negativas para defasamentos  $k = 2 \times i$ . Verifica-se que defasamentos múltiplos de 4 diminuem progressivamente para valores maiores de  $k$ .

#### 2.2.4. Exemplos de processos estocásticos

##### Sequência aleatória

*Sequência aleatória* é a denominação dada a séries temporais discretas compostas por uma sequência de variáveis aleatórias  $\varepsilon_t$ , independentes. Trata-se de um processo puramente aleatório com média,  $\mu = E(\varepsilon_t)$ , e variância,  $\sigma_\varepsilon^2 = Var(\varepsilon_t)$ , constantes e covariância igual a 0 entre duas variáveis distantes entre si em  $k$  unidades, ou seja,  $\gamma_k = Cov(\varepsilon_t, \varepsilon_{t+k}) = 0$  para qualquer  $k = \pm 1, 2, \dots$  (Ehlers, 2009).

##### Passeio aleatório

Tal como referido por Alpuim (2003) *passeio aleatório* é a denominação dada a séries temporais em que o processo estocástico é o das somas parciais, dadas por:

$$X_t = X_{t-1} + \varepsilon_t = \varepsilon_0 + \varepsilon_1 + \dots + \varepsilon_t = \varepsilon_0 + \sum_{j=1}^t \varepsilon_j = \sum_{j=0}^t \varepsilon_j.$$

Em que  $\varepsilon_t$  é a componente irregular associada ao momento  $t \in \mathbb{N}_0$ , com média  $\mu_\varepsilon = 0$ , variância ( $\sigma_\varepsilon^2$ ) constante e  $X_0 = \varepsilon_0$  é o primeiro valor da série.

A sua média será:

$$E(X_t) = \frac{\sum_{i=0}^t X_i}{t+1} = \frac{\sum_{i=0}^t (\sum_{j=0}^i \varepsilon_j)}{t+1} = \sum_{i=0}^t (E(\varepsilon_i)) = t\mu_\varepsilon = 0.$$

A variância será:

$$Var(X_t) = (t+1)\sigma_\varepsilon^2.$$

E a covariância será:

$$Cov(X_{t_1}, X_{t_2}) = \min\{t_1 + 1, t_2 + 1\} \sigma_\varepsilon^2.$$

Como a variância não é constante este tipo de processo estocástico não é estacionário. A sua variação é consequência da soma sucessiva de variações aleatórias que em alguns casos podem parecer com uma tendência, mas que na verdade não são.

### Processos de médias móveis

Segundo Ehlers (2009) uma série temporal  $X_t$  é considerada um *processo de média móvel (MA)*, com ordem  $q$ , se:

$$X_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} + \mu_t.$$

$\mu_t$  é a média da série temporal  $X_t$ ,  $\varepsilon_t$  é um processo discreto puramente aleatório (com média  $\mu_\varepsilon = 0$  e variância  $\sigma_\varepsilon^2$  constante) e os parâmetros  $\theta_i$  do modelo são números reais ( $\mathbb{R}$ ) com  $i = 1, \dots, q$ . A média de  $X_t$  é dada por:

$$E(X) = E(\varepsilon_t) + \sum_{j=1}^q \theta_j E(\varepsilon_{t-j}) = \mu_t.$$

Se  $\mu_t = 0$  teremos  $E(X_t) = 0$  e a variância será:

$$Var(X_t) = (1 + \theta_1^2 + \dots + \theta_q^2) \sigma_\varepsilon^2.$$

A função autocovariância (com  $\theta_0 = 1$ ):

$$\gamma_k = Cov(X_t, X_{t+k}) = \begin{cases} 0 & k > q \\ \sigma_\varepsilon^2 \sum_{j=0}^{q-k} \theta_j \theta_{j+k} & k = 0, \dots, q \\ \gamma_{-k} & k < 0. \end{cases}$$

E a função autocorrelação:

$$\rho_k = \begin{cases} 1 & k = 0 \\ \frac{\sum_{j=0}^{q-k} \theta_j \theta_{j+k}}{\sum_{j=0}^q \theta_j^2} & k = 1, \dots, q \\ 0 & k > q \\ \rho_{-k} & k < 0. \end{cases}$$

Pode-se dizer que esta série temporal é estacionária de 2ª ordem, com média e variância constantes e a autocovariância independente do instante  $t$ , para qualquer valor possível de  $\theta_1, \dots, \theta_q$ . Se a distribuição do processo puramente aleatório,  $\varepsilon_t$ , for normal (Gaussiana) podemos dizer que o processo é estritamente estacionário (Ehlers, 2009).

Para o caso particular de  $q = 1$  com  $\mu_t = E(X_t) = 0$ , temos o seguinte MA (1):

$$X_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \mu_t = \varepsilon_t + \theta_1 \varepsilon_{t-1}.$$

Com variância:

$$\text{Var}(X_t) = (1 + \theta_1^2) \sigma_\varepsilon^2.$$

E função autocovariância:

$$\gamma_k = \text{Cov}(X_t, X_{t+k}) = \begin{cases} 0 & |k| > 1 \\ \sigma_\varepsilon^2(1 + \theta_1^2) & k = 0 \\ \sigma_\varepsilon^2 \theta_1 & k = \pm 1. \end{cases}$$

E a função de autocorrelação será:

$$\rho_k = \begin{cases} 0 & |k| > 1 \\ 1 & k = 0 \\ \frac{\theta_1}{1 + \theta_1^2} & k = \pm 1. \end{cases}$$

Este processo é estacionário para qualquer valor de  $\theta_1$ . Como pode existir mais de um valor de  $\theta_1$  que conduz à mesma autocorrelação utiliza-se a noção de invertibilidade para restringir a um único modelo para cada função de autocorrelação (Ehlers, 2009).

Portanto o modelo deve ser invertível, ou seja, sabendo os valores de  $\varepsilon_t$  podem-se obter os respectivos valores de  $X_t$  e inversamente, sabendo os valores de  $X_t$  podem-se calcular os valores de  $\varepsilon_t$ .

Temos então que:

$$X_t = \varepsilon_t + \theta_1 \varepsilon_{t-1}.$$

Pode ser escrita como:

$$\varepsilon_t = X_t - \theta_1 X_{t-1} + \theta_1^2 X_{t-2} - \theta_1^3 X_{t-3} + \dots$$

No caso de  $\theta_1$  ser maior que 1, ou menor que -1, a invertibilidade não se verifica, valores de  $X_t$  mais distantes teriam uma influência maior sobre o valor de  $\varepsilon_t$ , o que não é verdadeiro. Assim temos que  $|\theta_1| < 1$  é uma condição necessária à invertibilidade e que a variação de  $X_t$  é mais influenciada pelos valores de  $X_{t-h}$  mais próximos do que de valores mais distantes.

Para outros valores de  $q$  usamos um operador de atraso  $B$  definido por:

$$B^j X_t = X_{t-j}, \text{ para todo } j.$$

Definimos o polinómio  $\Theta(B)$ , de ordem  $q$  em  $B$ , por:

$$\Theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q .$$

Então temos que:

$$X_t = (1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q) \varepsilon_t = \Theta(B) \varepsilon_t .$$

Existem  $2^q$  possíveis modelos para  $\Theta(B) = 0$  que terão a mesma função de autocorrelação, mas somente um terá as raízes da equação fora do círculo unitário. Ou seja, o processo MA( $q$ ) é invertível, quando todas as  $q$  soluções  $(\theta_1, \dots, \theta_q)$  de  $\Theta(B) = 0$  sejam  $|\theta_i| > 1$  e  $i = 1, \dots, q$  (Box *et al.*, 2008; Ehlers, 2009).

### Processos autorregressivos

Um processo estocástico  $\{X_t, t \in \mathbf{T}\}$  é considerado um **processo autorregressivo (AR)**, com ordem  $p$ , se:

$$X_t = \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + \varepsilon_t .$$

Os processos autorregressivos são modelos úteis para explicar séries temporais em que o valor da variável, em um determinado instante  $t$ , é dependente dos valores das variáveis anteriores até à variável  $t - p$  e do processo puramente aleatório ( $\varepsilon_t$ ) associado.  $\varepsilon_t$  é um processo discreto puramente aleatório com média  $\mu_\varepsilon = 0$  e variância  $\sigma_\varepsilon^2$  constantes. Os parâmetros  $\phi_i$  do modelo são números reais ( $\mathbb{R}$ ) com  $i = 1, \dots, p$  (Ehlers, 2009).

O caso mais simples de um modelo do tipo AR é o de ordem  $p = 1$  em que o modelo assume a forma:

$$X_t = \phi_1 X_{t-1} + \varepsilon_t , \tag{6}$$

que é equivalente a:

$$X_t = \phi_1^{r+1} X_{t-r-1} + \sum_{j=0}^r \phi_1^j \varepsilon_{t-j} .$$

Quando  $X_t$  for um processo estacionário com média  $\mu_t$  e variância  $\sigma_t^2$  finita temos:

$$E \left[ X_t - \sum_{j=0}^r \phi_1^j \varepsilon_{t-j} \right]^2 = \phi_1^{2r+2} E(X_{t-r-1}^2) = \phi_1^{2r+2} \sigma_t^2 .$$

Com  $|\alpha| < 1$  e  $r \rightarrow \infty$  teremos  $\phi_1^{2r+2} \rightarrow 0$  assim podemos reescrever:

$$X_t = \varepsilon_t + \phi_1 \varepsilon_{t-1} + \phi_1^2 \varepsilon_{t-2} + \dots ,$$

semelhante a um processo de médias móveis de ordem infinita. Pode-se utilizar o operador de atraso  $B$  a partir da equação:

$$BX_t = X_{t-1} \Rightarrow X_t = \phi_1 BX_t + \varepsilon_t \Leftrightarrow X_t = \frac{\varepsilon_t}{(1 - \phi_1 B)} .$$

Pode-se reescrever a equação como:

$$X_t = \frac{\varepsilon_t}{(1 - \phi_1 B)} = \varepsilon_t + \phi_1 \varepsilon_{t-1} + \phi_1^2 \varepsilon_{t-2} + \dots$$

Assim temos a média  $E(X_t) = 0$  e a variância:

$$Var(X_t) = \frac{\sigma_\varepsilon^2}{1 - \phi_1^2} .$$

A função de autocovariância:

$$\gamma_k = Cov(X_t, X_{t+k}) = E(X_t X_{t+k}) \text{ para qualquer } k = 0, 1, 2, \dots$$

E a função de autocorrelação:

$$\rho_k = \phi_1^k .$$

Como a média e a variância são constantes e a autocovariância é independente do instante  $t$  o modelo autorregressivo com ordem  $p = 1$  e  $|\phi_1| < 1$  é um processo estacionário (Ehlers, 2009).

Para ordem  $p$ , usando o conceito de que o processo autorregressivo pode ser expresso como um processo  $MA$  infinito, teremos:

$$X_t = \psi_0 \varepsilon_t + \psi_1 \varepsilon_{t-1} + \psi_2 \varepsilon_{t-2} + \dots = \psi(B) \varepsilon_t .$$

Assim se  $\sum_j \psi_j^2 < \infty$ , o modelo é estacionário.

Usando o operador de atraso  $B$ :

$$X_t = \phi_1 B + \dots + \phi_p B^p + \varepsilon_t \Leftrightarrow (1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) X_t = \varepsilon_t .$$

Que pode ser reescrito como:

$$\phi(B) X_t = \varepsilon_t \Leftrightarrow X_t = \frac{\varepsilon_t}{\phi(B)} = \psi(B) \varepsilon_t .$$

Como no caso do modelo de médias móveis, temos  $2^p$  possíveis soluções para  $\phi(B) = 0$ , e somente um dos modelos tem todas as raízes desta equação fora do círculo unitário, tornando o modelo invertível (Box *et al.*, 2008; Ehlers, 2009).

Para  $p = 1$  temos  $\phi(B) = 1 - \alpha_1 B = 0 \Leftrightarrow B = 1/\alpha_1$ , portanto para  $|B| > 1$  temos que  $|\alpha_1| < 1$ , como demonstrado anteriormente.

O processo é estacionário para o modelo em que  $\Phi(B) = 0$  tenha as raízes da equação fora do círculo unitário. Ou seja, o processo  $AR(p)$  é invertível, quando todas as  $p$  soluções  $(\alpha_1, \dots, \alpha_p)$  de  $\Phi(B) = 0$  sejam  $|\alpha_i| > 1$  e  $i = 1, \dots, p$  (Box *et al.*, 2008; Ehlers, 2009).

Para um processo estacionário  $AR(p)$  com  $E(X_t) = 0$  e variância  $\sigma_t^2$  teremos as seguintes funções de autocovariância e de autocorrelação:

$$\gamma_k = E(X_t X_{t+k}) = \sum_{i=1}^p \phi_i \gamma_{k-i}, k > 0 ,$$

$$\rho_k = \sum_{i=1}^p \alpha_i \rho_{k-i}, k > 0 .$$

### 2.3. Metodologia utilizada na análise de uma série temporal

É usual iniciar a análise de uma série temporal através da sua representação gráfica. Esta análise gráfica é útil na medida em que se tem uma ideia global do comportamento da série temporal, ou seja, da existência de movimentos periódicos e de outras características, que nos ajudarão a escolher os modelos apropriados.

Na sequência da secção 2.1.2., em que foram apresentadas as componentes básicas de uma série temporal, prossegue-se agora com os conceitos e metodologias que permitem modelar essas componentes, avaliar a qualidade do seu ajuste e significado estatístico.

### 2.3.1. Decomposição clássica

Os métodos de decomposição clássica são utilizados no estudo de séries temporais cujo objetivo é decompor a série temporal nas suas componentes de maneira a estudá-las separadamente e obter uma melhor compreensão da mesma. Estes métodos têm origem na década de 1920 e constituem a base dos métodos usados atualmente para a análise de séries temporais (Hyndman e Athanasopoulos, 2013).

Definimos por  $\{X_t: t \in \mathbb{R}\}$  a série temporal e por  $T_t, S_t, C_t$  e  $I_t$  as respectivas componentes: *tendência, sazonal, cíclica e irregular*.

A decomposição pode ser aditiva:

$$X_t = T_t + S_t + C_t + I_t, \quad 1 \leq t \leq n,$$

ou multiplicativa:

$$X_t = T_t \times S_t \times C_t \times I_t, \quad 1 \leq t \leq n.$$

Em muitos casos a componente cíclica não existe, ou poderá ser incorporada pela componente tendência, assim simplifica-se a decomposição clássica em:

$$X_t = T_t + S_t + I_t, \quad 1 \leq t \leq n, \quad (7)$$

ou

$$X_t = T_t \times S_t \times I_t, \quad 1 \leq t \leq n. \quad (8)$$

O **modelo aditivo** (7) é usado quando a componente sazonal é do tipo aditivo, ou seja, a amplitude sazonal mantém-se constante para qualquer valor do nível médio da série temporal. O **modelo multiplicativo** (8) é usado quando a componente sazonal é do tipo multiplicativo, ou seja, a amplitude sazonal depende do nível médio da série temporal (Hyndman e Athanasopoulos, 2013).

Uma alternativa para o modelo multiplicativo é a transformação dos dados da série temporal usando uma transformação que resulte numa série temporal cuja componente sazonal seja estável ao longo do tempo, podendo assim usar-se o modelo aditivo. A transformação logarítmica é uma das possíveis transformações utilizadas e a sua equivalência pode ser demonstrada pelo seguinte:

$$X_t = S_t \times T_t \times I_t \Rightarrow \log X_t = \log S_t + \log T_t + \log I_t, \quad 1 \leq t \leq n.$$

### 2.3.2. Modelos de alisamento exponencial

Os métodos de *alisamento exponencial* têm a sua origem na década de 1940, com os estudos do pioneiro Robert Goodell Brown. Estes métodos foram propostos pelo mesmo nos finais da década de 1950. Na mesma década, e em estudos independentes, Charles Carter Holt desenvolveu outros métodos de alisamento exponencial (1957), sendo os seus modelos testados na década seguinte por Peter Winters, o que conferiu a esses modelos a designação de modelos de alisamento exponencial de Holt-Winters (Hyndman e Athanasopoulos, 2013).

Os métodos de alisamento exponencial são ferramentas que permitem criar modelos que explicam o padrão comportamental de uma determinada série temporal. Estas técnicas assumem que os valores discrepantes (*outliers*) são determinados pela aleatoriedade, assim com o alisamento suaviza-se o efeito dos mesmos permitindo a identificação do padrão básico da série temporal (Morettin e Toloí, 2006).

Nesta metodologia são atribuídos pesos diferentes às observações, sendo que previsões baseadas nestes modelos são as médias ponderadas das observações, onde o peso atribuído a cada observação decai exponencialmente das observações mais recentes às mais antigas, ou seja, quanto mais recente for uma observação maior será o peso associado à mesma. Estes métodos são amplamente utilizados uma vez que geram modelos simples, que se adequam a uma grande variedade de séries temporais e que possuem uma capacidade de previsão confiável (Hyndman e Athanasopoulos, 2013).

#### Alisamento exponencial simples

O *alisamento exponencial simples (SES)* é um método que usa valores exponencialmente alisados  $l_t$ , e  $t = 1, 2, \dots, n$ , estes são denominados por nível da série e são a média ponderada dos dados, em que o peso das observações mais recentes é maior e decai exponencialmente para valores mais antigos (Morettin e Toloí, 2006).

O nível ( $l_t$ ) é obtido pela seguinte equação:

$$l_t = \alpha X_t + (1 - \alpha)l_{t-1} .$$

A constante de alisamento  $\alpha \in [0,1]$  está associada ao nível ( $l_t$ ) da série temporal.

A previsão para  $h$  passos à frente é igual ao  $l_t$  da última observação:

$$\hat{X}_{t+h} = l_t, \quad \forall h > 0 .$$

A previsão é atualizada a cada nova observação, corrigindo estimativas de  $\hat{X}_t$  baseadas em observações anteriores.

Este método é de simples aplicação e flexível, é possível adequar o valor de  $\alpha$  para obter um melhor ajustamento. Uma das suas desvantagens é não ser aplicável para séries que tenham tendência ou sazonalidade.

### Alisamento exponencial de Holt

O *alisamento de Holt* é recomendado para séries temporais que apresentam tendência linear (positiva ou negativa). Esta técnica apresenta semelhanças ao alisamento exponencial simples, mas com o acréscimo da constante de alisamento  $\beta$  (taxa de crescimento) que está associada à componente tendência (Morettin e Toloí, 2006).

Três equações caracterizam este modelo (Hyndman e Athanasopoulos, 2013; Morettin e Toloí, 2006).

$$\begin{aligned} \text{Nível:} & \quad l_t = \alpha X_t + (1 - \alpha)(l_{t-1} + b_{t-1}), \\ \text{Tendência:} & \quad b_t = \beta(l_t - l_{t-1}) + (1 - \beta)b_{t-1}, \\ \text{Previsão:} & \quad \hat{X}_{t+h} = l_t + hb_t, \quad \forall h > 0. \end{aligned}$$

Em que as constantes de alisamento,  $\alpha$  e  $\beta$ , pertencem ao intervalo  $[0,1]$ .

Este modelo apresenta as mesmas vantagens dos modelos anteriores acrescentando a possibilidade de modelar séries temporais com tendência. Por outro lado, a desvantagem principal é a dificuldade em determinar os valores mais adequados para as constantes de alisamento (Morettin e Toloí, 2006).

### Alisamento exponencial de Holt-Winters

Os modelos de *alisamento exponencial de Holt-Winters* são adequados para séries temporais que apresentem tendência e sazonalidade, podendo ser do tipo aditivo ou multiplicativo. Esta técnica apresenta semelhanças ao alisamento de Holt com o acréscimo da constante de alisamento  $\gamma$  associada à sazonalidade (Morettin e Toloí, 2006).

Estes modelos são caracterizados por três constantes de alisamento  $\alpha$ ,  $\beta$  e  $\gamma$  (estimadas pelo método dos mínimos quadrados), associadas ao nível médio ( $l_t$ ), à taxa de crescimento ( $b_t$ ) e à sazonalidade ( $s_t$ ).

As quatro equações que caracterizam o modelo aditivo são as seguintes (Hyndman e Athanasopoulos, 2013; Morettin e Toloi, 2006):

$$\begin{aligned}
 \text{Nível:} \quad & l_t = \alpha(X_t - s_{t-s}) + (1 - \alpha)(l_{t-1} + b_{t-1}), \\
 \text{Tendência:} \quad & b_t = \beta(l_t - l_{t-1}) + (1 - \beta)b_{t-1}, \\
 \text{Sazonalidade:} \quad & s_t = \gamma(X_t - l_{t-1} - b_{t-1}) + (1 - \gamma)s_{t-s}, \\
 \text{Previsão:} \quad & \hat{X}_{t+h} = l_t + hb_t + s_{t-s+h}, \quad 0 < h \leq s \\
 & \hat{X}_{t+h} = l_t + hb_t + s_{t-2s+h}, \quad s < h \leq 2s \\
 & \dots \\
 & \hat{X}_{t+h} = l_t + hb_t + s_{t-is+h}, \quad s(i-1) < h \leq is \text{ e } i \in \mathbb{N}.
 \end{aligned}$$

Sendo  $s$  o período da sazonalidade. Por exemplo: para dados mensais temos  $s = 12$  e para dados trimestrais temos  $s = 4$ .

Para o modelo multiplicativo temos as equações (Hyndman e Athanasopoulos, 2013; Morettin e Toloi, 2006):

$$\begin{aligned}
 \text{Nível:} \quad & l_t = \alpha \frac{X_t}{s_{t-s}} + (1 - \alpha)(l_{t-1} + b_{t-1}), \\
 \text{Tendência:} \quad & b_t = \beta(l_t - l_{t-1}) + (1 - \beta)b_{t-1}, \\
 \text{Sazonalidade:} \quad & s_t = \gamma \frac{X_t}{(l_{t-1} + b_{t-1})} + (1 - \gamma)s_{t-s}, \\
 \text{Previsão:} \quad & \hat{X}_{t+h} = (l_t + hb_t)s_{t-s+h}, \quad 0 < h \leq s \\
 & \hat{X}_{t+h} = (l_t + hb_t)s_{t-2s+h}, \quad s < h \leq 2s \\
 & \dots \\
 & \hat{X}_{t+h} = (l_t + hb_t)s_{t-is+h}, \quad s(i-1) < h \leq is \text{ e } i \in \mathbb{N}.
 \end{aligned}$$

As constantes de alisamento  $\alpha$ ,  $\beta$  e  $\gamma$  pertencem ao intervalo  $[0,1]$ . À semelhança do alisamento exponencial de Holt, a cada nova observação  $\hat{X}_t$  é atualizada (Hyndman e Athanasopoulos, 2013; Morettin e Toloi, 2006).

Este modelo apresenta as mesmas vantagens dos modelos anteriores acrescentando a possibilidade de modelar séries temporais que apresentem sazonalidade. A desvantagem principal é a dificuldade em identificar propriedades estatísticas como a média e variância de previsão, o que torna difícil a determinação de intervalos de confiança (Morettin e Toloi, 2006).

### 2.3.3. Modelos *ARIMA*

Os modelos *autorregressivos e de médias móveis integrados (ARIMA)* são uma metodologia muito utilizada na análise de séries temporais. Em 1970 os professores George E. P. Box e Gwilyn M. Jenkins publicaram o livro *Time Series: Analysis, Forecasting and Control* (existem várias edições posteriores), com uma metodologia que reuniu diferentes métodos já existentes de maneira a produzir modelos capazes de descrever o comportamento de uma grande variedade de séries temporais de forma parcimoniosa (o número de parâmetros a estimar é pequeno) e cujos resultados têm grande poder preditivo quando comparados a outros métodos (Morettin e Toloi, 2006).

Este método tem como base a ideia de que cada valor da série temporal pode ser explicado pelos seus valores anteriores, ou seja, de que existe uma certa dependência entre valores próximos no tempo. Estes modelos derivam da combinação de três processos: o *processo autorregressivo (AR)*, o *processo de integração (I)* e o *processo de médias móveis (MA)*.

Os modelos  $AR(p)$ ,  $MA(q)$  e  $ARMA(p, q)$  impõem a condição de estacionaridade, o que implica que séries temporais que não sejam estacionárias tenham que ser alteradas, removendo a fonte de não estacionaridade (como a tendência ou a sazonalidade). Nos modelos *ARIMA* temos uma série temporal  $X_t$  não estacionária que é transformada em outra série temporal  $W_t$  estacionária, por uma ou mais diferenças, e que pode ser descrita por um modelo misto. Neste caso temos a diferenciação que reduz a não estacionaridade integrada no modelo (Ehlers, 2009).

$$W_t^d = \nabla^d X_t = (1 - B)^d X_t.$$

Sendo o *processo autorregressivo de médias móveis integrado (ARIMA(p, d, q))* dado pela equação:

$$\Phi(B)(1 - B)^d X_t = \Theta(B)I_t.$$

A diferenciação permite transformar a série original numa série estacionária, e a condição de invertibilidade também é imposta a este tipo de modelos (Ehlers, 2009).

### Modelos *SARIMA*

A equação anterior é adequada para modelos não sazonais e pode ser eficaz para séries que possuam uma tendência que será retirada pela diferenciação. Para séries temporais que possuem uma componente sazonal, utilizam-se modelos *ARIMA* sazonais, designados por modelos *autorregressivos e de médias móveis integrado sazonal (SARIMA)*. Nestes modelos para tornar a série estacionária é necessário utilizar uma ou mais diferenciações múltiplas do padrão sazonal  $s$ . A diferença sazonal é representada por  $\nabla_s^D$ , em que  $D$  é o número de diferenciações sazonais. Os dois tipos de diferenciação podem ser usados simultaneamente  $\nabla^d \nabla_s^D$  (Ehlers, 2009).

Os Modelos *ARIMA* sazonais podem ser obtidos pela inclusão no modelo *ARIMA*( $p, d, q$ ) da componente sazonal obtendo um modelo *SARIMA* de ordem  $(p, d, q)(P, D, Q)_s$  definido por:

$$\Phi(B)\Phi(B^s)\nabla^d\nabla_s^D X_t = \Theta(B)\Theta(B^s)e_t.$$

Os valores das ordens  $d$  e  $D$  removem respetivamente a tendência e a sazonalidade e geralmente não serão maiores que 1 para tornar a série temporal estacionária. As restantes ordens  $p, P, q$  e  $Q$  são determinadas com base nas funções de autocorrelação e autocorrelação parcial.

#### 2.4. Critérios e medidas para a escolha do modelo

As várias metodologias apresentadas podem ser usadas na modelação de uma série temporal. Desta forma podemos ajustar para as mesmas observações mais do que um modelo utilizando diferentes metodologias (modelos de decomposição clássica, Holt-Winters, *ARIMA* entre outros) assim como para a mesma metodologia com diferente estrutura ou parâmetros (por exemplo: decomposição clássica aditiva ou multiplicativa, modelos *ARIMA* com diferentes ordens  $p, d$  e  $q$ ). Como vários modelos podem ser ajustados, torna-se necessário definir critérios e medidas que permitam comparar a qualidade do ajustamento, levando à escolha do modelo mais adequado para descrever a série temporal em estudo.

### 2.4.1. Critérios AIC, AICc e BIC

Quando temos vários modelos *ARIMA* ou *SARIMA* de diferentes ordens pode-se determinar o mais adequado com base num critério. Segundo Ehlers (2009), o **critério de informação de Akaike (AIC)** é o critério de seleção mais utilizado em séries temporais. Foi proposto por Akaike (1973, 1974) que sugere a escolha de modelos com ordens  $p$  e  $q$  que minimizem este critério. A sua vantagem é a possibilidade de comparar modelos muito diferentes (desde que sejam ajustados aos mesmos dados). Uma das desvantagens do critério AIC é a tendência (principalmente para um número pequeno de observações) de selecionar modelos que contenham muitos parâmetros. Para contornar essa desvantagem foram criados outros critérios que são modificações do AIC que penalizam a inclusão de parâmetros adicionais. Dois dos mais utilizados são o **critério de informação de Akaike corrigido (AICc)** (proposto por Hurvick e Tsai em 1989) e o **critério informação bayesiano (BIC)** (desenvolvido por Schwarz e publicado em 1978). O primeiro é uma versão do AIC com uma correção para um número finito de observações. O segundo penaliza mais um número maior de parâmetros do que o AIC. Sendo assim ambos os critérios selecionam ou os mesmos modelos que o critério AIC, ou modelos com menor número de parâmetros (Ehlers, 2009; Hyndman e Athanasopoulos, 2013; Morettin & Toloi, 2006).

As fórmulas associadas a estes critérios são (Hyndman e Athanasopoulos, 2013):

$$AIC = n \log\left(\frac{SSE}{n}\right) + 2(k + 2), \quad (9)$$

$$AICc = AIC + \frac{2(k + 2)(k + 3)}{n - K - 3}, \quad (10)$$

$$BIC = N \log\left(\frac{SSE}{N}\right) + (k + 2) \log(n). \quad (11)$$

Sendo  $SSE = \sum_{i=1}^N e_i^2$  é a soma dos erros quadrados,  $n$  é o número de observações e  $k$  é número de parâmetros do modelo.

Neste trabalho, os três critérios serão utilizados para escolha do melhor modelo *ARIMA*. No caso de os critérios selecionarem modelos diferentes, o modelo escolhido será aquele com menor valor para pelo menos dois dos critérios de seleção.

### 2.4.2. Medidas de erro

As medidas de erro são usadas para verificar a precisão de previsão de modelos, assim como para comparar diferentes modelos. Segundo Hyndman e Koehler (2006) as medidas de erro mais usadas para comparar métodos diferentes para dados de uma mesma série temporal são o *Erro Quadrático Médio (MSE)*, *Desvio Padrão Empírico (Amostral) Generalizado (RMSE)*, o *Erro Absoluto Médio (MAE)* e o *Erro Absoluto Mediano (MdAE)*. Os dois primeiros eram muito populares devido à sua relevância teórica para a modelação estatística, mas são mais sensíveis a *outliers* que os erros absolutos médio e mediano. Ainda segundo estes autores as medidas de erro baseadas na percentagem são muito eficazes para comparar modelos entre diferentes séries temporais. Uma das medidas mais utilizadas é o *Erro Percentual Absoluto Médio (MAPE)*.

A utilização do MAPE não é adequada para dados em que o valor zero não signifique o valor mínimo, não é, portanto, uma medida de comparação adequada para a temperatura ou séries com valores negativos (Hyndman e Athanasopoulos, 2013).

Neste trabalho a medida MAE foi escolhida como principal medida de comparação do ajustamento de modelos gerados por diferentes métodos. O MAE é uma das medidas mais populares para a comparação de diferentes modelos para a mesma série temporal pela facilidade de compreensão e cálculo (Hyndman e Athanasopoulos, 2013; Mentzer e Bahn, 1995).

Esta medida expressa a média dos erros absolutos calculada por:

$$MAE = \frac{1}{n} \sum |X_t - \hat{X}_t|. \quad (12)$$

### 2.5. Testes de hipóteses

Depois de identificar/ajustar um modelo que explique a série temporal em estudo torna-se necessário verificar se os seus resíduos são independentes e se são um processo estacionário, sendo assim um processo puramente aleatório.

### 2.5.1. Testes de independência

Diferentes testes podem ser usados para verificar a independência dos resíduos. Quando observamos os valores das autocorrelações num correlograma pretendemos identificar se para cada desfasamento  $k$  a autocorrelação dos resíduos é estatisticamente nula. Segundo Hyndman (2013) este método equivale à realização de vários testes de hipóteses em que cada um possui individualmente uma pequena possibilidade de gerar um resultado falso positivo. Para um número grande de desfasamentos será provável que pelo menos um dos testes produza um falso positivo e que se conclua que os resíduos não sejam independentes. Para contornar este problema criaram-se testes denominados por testes de *portmanteau* que testam um grupo de autocorrelações para diferentes desfasamentos.

Dois dos testes de *portmanteau* mais conhecidos são os testes **de Box-Pierce e o de Ljung-Box**. Ambos são usados para verificar se a autocorrelação dos resíduos é nula no horizonte temporal (Ehlers, 2009).

Ambos os testes testam a hipótese nula de independência da variável, ou seja, que a autocorrelação entre os valores da variável em diferentes instantes é nula, contra a hipótese alternativa de que a autocorrelação é diferente de zero:

$$\begin{aligned} H_0: \rho_1 = \rho_2 = \dots = \rho_m = 0 \\ H_1: \rho_k \neq 0, \quad \text{para pelo menos um } k \in \{1, \dots, m\}. \end{aligned}$$

Estes testam se um grupo de  $m$  autocorrelações é significativamente diferente de zero através da estatística de teste  $Q$ .

A estatística  $Q$  do teste de Box-Pierce é dada por (Hyndman e Athanasopoulos, 2013):

$$Q = n \sum_{i=1}^m \rho_i^2.$$

O teste de Box-Pierce não é indicado para amostras de pequenas dimensões, pois neste caso estas afastam-se da distribuição qui-quadrado, sendo preferível utilizar nestes casos o teste de Ljung-Box (Ehlers, 2009).

A estatística  $Q^*$ , do teste de Ljung-Box, é uma versão mais precisa da estatística  $Q$ , e é dada por (Hyndman e Athanasopoulos, 2013):

$$Q^* = n(n+2) \sum_{i=1}^m (n-i)^{-1} \rho_i^2. \quad (13)$$

Hyndman e Athanasopoulos (2013) sugerem a utilização de  $m = 2s$  para séries temporais com sazonalidade, sendo o valor máximo de  $m$  não maior do que  $n/5$  e de  $m = 10$  para séries sem sazonalidade.

O teste de Ljung-Box é uma versão mais atualizada e com maior precisão do teste de Box-Pierce que pode ser utilizada para amostras de menores dimensões, por estes motivos foi o teste de independência escolhido neste trabalho. Se  $Q^* > \chi_{\alpha}^2$ , para o nível de significância de 0,05, rejeita-se a hipótese nula e conclui-se que os resíduos do modelo não são independentes.

### 2.5.2. Teste de raízes unitárias

Diferentes testes podem ser utilizados para a verificar a presença (ou ausência) de uma raiz do polinómio autorregressivo dentro do círculo unitário e, portanto, de não estacionariedade. Estes testes permitem identificar se há ou não necessidade de uma diferenciação para tornar a série temporal estacionária (Hyndman e Athanasopoulos, 2013).

Entre eles os mais utilizados são os *testes de Dickey-Fuller* e *Phillips-Perron*. O teste de Phillips-Perron é uma generalização do teste de Dickey-Fuller para resíduos que não apresentem variância constante e que estejam autocorrelacionados, ambos testam a hipótese de que exista pelo menos uma raiz dentro do círculo unitário e que, portanto, a estacionariedade não seja verificada (Hamilton, 1994). O *teste de Dickey-Fuller aumentado* foi escolhido neste trabalho por ser um dos mais populares (Hyndman e Athanasopoulos, 2013) e pelo facto da componente irregular em que este teste será aplicado apresentar variância constante e da autocorrelação ser nula para a maior parte dos desfasamentos.

A hipótese nula testada é a de que o polinómio autorregressivo possui pelo menos uma raiz unitária, ou seja, de que a série não é estacionária, contra a hipótese alternativa de que todas as raízes do processo estarão fora do círculo unitário levando à conclusão de que a série é estacionária (Morettin e Toloí, 2006):

$$H_0: \sum_{i=1}^p \phi_i = 1 \Leftrightarrow \delta = 0$$

$$H_1: \sum_{i=1}^p \phi_i < 1 \Leftrightarrow \delta < 0.$$

Este teste é baseado na regressão da variável em relação aos seus valores desfasados ( $\Delta X_t = X_t - X_{t-1}$ ) e no teste dos parâmetros associados. Este verifica se pelo menos uma das raízes se encontra fora do círculo unitário quando a série temporal pode ser explicada por um dos três modelos (Harris e Sollis, 2003):

$$X_t = \phi_1 X_{t-1} + \varepsilon_t,$$

$$X_t = a_0 + \phi_1 X_{t-1} + \varepsilon_t,$$

$$X_t = a_0 + a_1 t + \phi_1 X_{t-1} + \varepsilon_t.$$

Onde  $\phi_1$  representa as raízes da série,  $a_0$  é constante e  $a_1$  é o termo determinístico. A hipótese nula é neste caso  $H_0: \phi_1 = 1$  em oposição a  $H_1: \phi_1 < 1$ , o que é equivalente a dizer que  $H_0: \phi_1 - 1 = 0$  em oposição a  $H_1: \phi_1 - 1 < 0$ . Substituindo  $\delta = \phi_1 - 1$  obtemos as seguintes equações:

$$\Delta X_t = X_t - X_{t-1} = \delta X_{t-1} + \varepsilon_t,$$

$$\Delta X_t = a_0 + \delta X_{t-1} + \varepsilon_t,$$

$$\Delta X_t = a_0 + a_1 t + \delta X_{t-1} + \varepsilon_t.$$

E as hipóteses passam a ser  $H_0: \delta = 0$  em oposição a  $H_1: \delta < 0$  (Harris e Sollis, 2003).

Este teste parte do pressuposto de que os diferentes valores dos resíduos do modelo sejam independentes. Caso os resíduos apresentem autocorrelação e o número de observações seja suficientemente grande utiliza-se o teste aumentado de Dickey-Fuller. Assim, para processos AR( $p$ ) e  $\delta = \phi_1 + \phi_2 + \dots + \phi_p - 1$ , temos as seguintes equações aumentadas:

$$\Delta X_t = \delta X_{t-1} + \sum_{i=1}^{p-1} \phi_i \Delta X_{t-i} + \varepsilon_t, \quad (14)$$

$$\Delta X_t = a_0 + \delta X_{t-1} + \sum_{i=1}^{p-1} \phi_i \Delta X_{t-i} + \varepsilon_t,$$

$$\Delta X_t = a_0 + at + \delta X_{t-1} + \sum_{i=1}^{p-1} \alpha_i \Delta X_{t-i} + \varepsilon_t.$$

Se o polinómio  $\Phi(B)$  possuir uma raiz unitária implica que  $\Phi(1) = 1 - \phi_1 - \phi_2 - \dots - \phi_p = 0$ , ou seja, que  $\sum_{i=1}^p \phi_i = 1$  o que equivale a  $\delta = 0$ .

O modelo pode ser ainda aumentado para incluir processos de médias móveis na componente aleatória. Os valores críticos para este teste foram calculados e tabelados por Dickey e Fuller utilizando o método de simulação Monte Carlo (Harris e Sollis, 2003).



**Capítulo III - Estudo de Séries Temporais de Dados Oceanográficos**



### 3.1. Introdução



Os oceanos têm um papel primordial no sistema climático do planeta. Os grandes volumes de água do mar permitem armazenar, transportar e equilibrar grandes quantidades de calor solar (Bollmann *et al.*, 2010; McLellan, 1965). Alterações a longo prazo na temperatura superficial dos oceanos têm consequências nos ecossistemas marinhos, afetando espécies de plantas, animais e seres unicelulares. Padrões migratórios, reprodutivos, aumento dos níveis metabólicos e da cadeia alimentar de várias espécies podem ser influenciados por variações da temperatura. Espécies vulneráveis (como os corais por exemplo) entram em risco e outras espécies oportunistas podem ter um crescimento populacional exagerado (Bollmann *et al.*, 2010).

A área superficial dos oceanos retém grandes quantidades de energia proveniente do sol, tendo um papel fundamental na manutenção da temperatura global. O aumento da temperatura global de cerca 0,5°C no último século, assim como a previsão do aumento entre 1°C a 5,8°C para o século atual (previsão feita por um painel de cientistas escolhidos pelas Nações Unidas) implica na maior evaporação da água nos oceanos, aumento de precipitação e intempéries climáticas (Castro, 2003). O aumento da temperatura global reflete-se na componente térmica dos oceanos com o consequente aumento da temperatura das águas. No verão absorve energia na forma de temperatura e no inverno liberta esta mesma energia, reduzindo assim grandes variações da temperatura planetária (Bollmann *et al.*, 2010; McLellan, 1965).

A **unidade de medida** para a temperatura do mar em oceanografia são os °C (graus Celsius), medida proposta em 1742 pelo astrónomo Anders Celsius. Esta medida é baseada nos pontos de evaporação considerados de 0°C e o ponto de ebulição de 100°C para a pressão de 1 atm (pressão atmosférica padrão). A precisão padrão na determinação da temperatura é de 0,01°C. A temperatura pode também ser expressa na escala Kelvin ( $0^{\circ} \text{K} = -273,15^{\circ}\text{C}$ , considerado o zero absoluto) ou na escala Fahrenheit ( $x^{\circ}\text{C} = \frac{y^{\circ}\text{F}-32}{1,8}$ ) (Mamayev, 1975). A temperatura do mar varia entre os -2°C e os 30°C, a existência de temperaturas inferiores a 0°C é possível pelo ponto de congelação da água salgada ser inferior ao da água doce (Castro, 2003). A maior parte desta variação ocorre na superfície do mar e tem o limite inferior delimitado pelo ponto de congelação da água. A maior parte da água oceânica está a baixas

temperaturas, com quase metade abaixo dos 2°C e apenas uma pequena fração com temperatura superior a 6°C. A **temperatura à superfície do mar** (SST) sofre grandes variações tanto no tempo quanto no espaço, variando ao decorrer do ano e com a latitude, sendo mais quente próxima do equador e mais fria nas regiões do Ártico e da Antártida (Castro, 2003; McLellan, 1965).

A SST pode ser obtida por métodos de obtenção *in situ* ou **satélite**. Nos métodos *in situ*: procede-se à recolha da água do mar (a partir de navios) em recipientes apropriados, e faz-se a medição direta utilizando termómetros; em estações hidrográficas remotas utilizam-se termómetros que são colocados entre a profundidade de 0,5 a 3 metros; a partir de navios, ou aviões, mede-se a temperatura da superfície do mar estimada por termómetros de leitura da radiação ao nível dos infravermelhos (Castro, 2003; McLellan, 1965). Os satélites são usados desde o final da década de 60 para a determinação remota da SST a nível global. Equipamentos nos satélites detetam variações da radiação na região dos infravermelhos (ou outras frequências que possam estar relacionadas à temperatura) emitida pela superfície do mar, obtendo assim uma estimativa da temperatura. Desde 2000, a NASA (*National Aeronautic and Space Administration*) fornece informação sobre a SST a nível global obtida pelos SST satélites observacionais providos com o instrumento MODIS (*Moderate Resolution Imaging Spectroradiometer*). O NOAA (*National Oceanic and Atmospheric Administration*) disponibiliza dados relativos à SST a nível global diários (com intervalos de algumas horas) recolhidos por satélites geo-estacionários GOES (*Geostationary Orbiting Earth Satellites*).

Neste capítulo serão utilizadas as metodologias de análise temporal (introduzidas no Capítulo 2) para dados oceanográficos que serão descritos a seguir. Também serão apresentados os resultados e interpretações desta análise e os modelos mais ajustados para descrever algumas das séries temporais deste estudo. Foi utilizada a versão 3.3.2 do programa estatístico . Utilizaram-se os pacotes suplementares *astsa* (Stoffer, 2016), *tseries* (Trapletti *et al.*, 2017) e o *forecast* (Hyndman *et al.*, 2017). Os comandos utilizados encontram-se no Anexo II – Comandos do .

### 3.2 A temperatura à superfície do mar

#### 3.2.1 Introdução

Compreender os fenômenos da variação da temperatura à superfície do mar ao longo do tempo contribui significativamente para a compreensão do sistema climático global do nosso planeta (Bindoff *et al.* 2007). Alterações da SST podem ter efeitos profundos no clima global devido à interação entre os oceanos e a atmosfera. O aumento da temperatura superficial dos oceanos tem como consequência direta o aumento da evaporação. A análise da humidade a nível global demonstra o aumento desta sobre a superfície oceânica e terrestre (Hartmann *et al.*, 2013). O aumento de vapor de água sobre os oceanos interfere diretamente com os sistemas atmosféricos associados, implicando o aumento de precipitação e alteração de padrões de tempestades.

#### 3.2.2 As regiões em estudo

Neste trabalho são estudadas as séries temporais correspondentes a seis localizações (Figura 3.1).

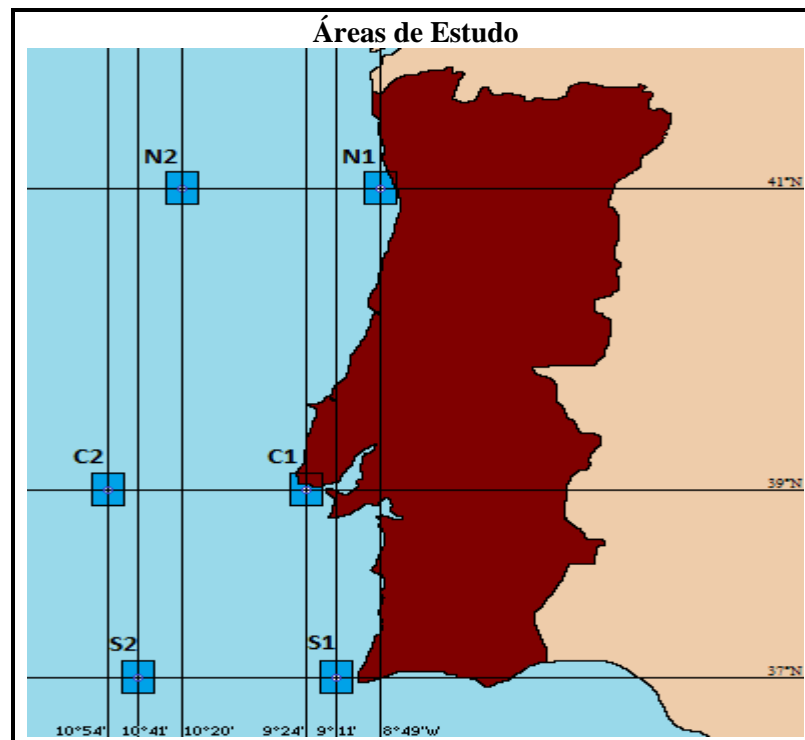


Figura 3.1: Mapa de Portugal com as seis localizações em estudo.<sup>9</sup>

<sup>9</sup> Fonte: Autor.

- N1 - Coordenadas de latitude 41°8'1" N e longitude 8°49'4" W com uma resolução de 20,90 km (eixo Este Oeste) por 27,75 km (eixo Este Oeste eixo Norte Sul).
- N2 - Coordenadas de latitude 41°9'50" N e longitude 10°19'51" W com uma resolução de 20,89 km (eixo Este Oeste) por 27,75 km (eixo Este Oeste eixo Norte Sul).
- C1 - Coordenadas de latitude 38°38'56" N e longitude 9°23'43" W com uma resolução de 21,67 km (eixo Este Oeste) por 27,75 km (eixo Norte Sul).
- C2 - Coordenadas de latitude 38°40'11" N e longitude 10°54'11" W com uma resolução de 21,67 km (eixo Este Oeste) por 27,75 km (eixo Norte Sul).
- S1 - Coordenadas de latitude 37°4'9" N e longitude 9°10'37" W com uma resolução de 22,14 km (eixo Este Oeste) por 27,75 km (eixo Norte Sul).
- S2 - Coordenadas de latitude 37°5'9" N e longitude 10°40'56" W com uma resolução de 22,14 km (eixo Este Oeste) por 27,75 km (eixo Norte Sul).

As três localizações costeiras possuem centros da grelha a menos de 20 km da costa portuguesa e as três oceânicas possuem centros da grelha entre 135 e 150 km da costa, e latitudes semelhantes às anteriores. Por conveniência definiram-se as localizações próximas da costa por **N1, C1 e S1** e as localizações distantes da costa por **N2, C2 e S2**.

Os dados deste estudo utilizam informação proveniente do E.U. Copernicus Marine Service Information. Foram retirados do produto GLOBAL\_REANALYSIS\_PHYS\_001\_011, produzida pelo Centro Euro-Mediterraneo sui Cambiamenti Climatici (CMCC), fornecido no catálogo do website <http://marine.copernicus.eu/>. As observações correspondem às médias mensais da SST entre os anos de **janeiro 1982 e dezembro 2013** à profundidade de 0,4942 metros.

#### 3.2.3 Análise da SST por região e localização

Na análise de séries temporais a representação gráfica dos dados ao longo do tempo é de extrema importância. Esta análise inicial permite muitas vezes identificar padrões sazonais, a presença de uma tendência crescente ou decrescente, alterações estruturais, *outliers* entre outros comportamentos importantes para a descrição da série e para determinar quais os passos seguintes na análise e modelação das séries temporais em estudo (Ehlers, 2009).

A análise gráfica das séries temporais para as seis regiões em estudo (Figura 3.2) evidencia a existência de um padrão sazonal nas seis localizações.

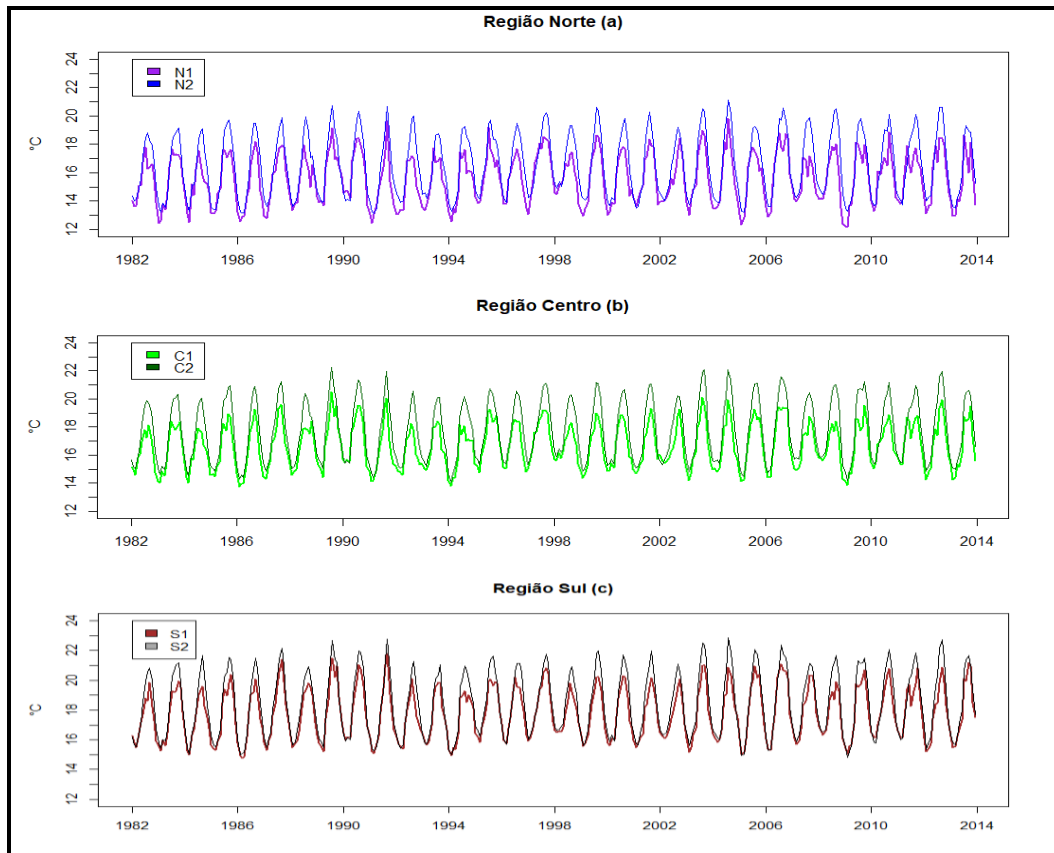


Figura 3.2: Representação gráfica da SST para as regiões (a) Norte, (b) Centro e (c) Sul.<sup>10</sup>

As séries temporais seguem um padrão bem definido e não se verifica à priori alterações comportamentais significativas no horizonte temporal. Verificam-se temperaturas mais altas para as séries temporais da região Sul, seguidas pelas da região Centro e com temperaturas mais baixas para as da região Norte. É possível observar uma diferença de comportamento nos períodos de temperaturas mais altas entre cada série temporal próxima e distante da costa, para cada região.

Para completar e confirmar a análise anterior, recorreu-se à representação e observação dos gráficos da ACF e da PACF (Figura 3.3) das séries temporais, estimadas com o auxílio das equações (4) e (5). Observa-se que as séries temporais não possuem um comportamento puramente aleatório, pois os valores da autocorrelação não são estatisticamente nulos (estão fora do intervalo de confiança de  $\pm 0,1021$ , para o nível de significância de 0,05) para todos os desfasamentos  $k$ .

<sup>10</sup> Fonte: Autor.

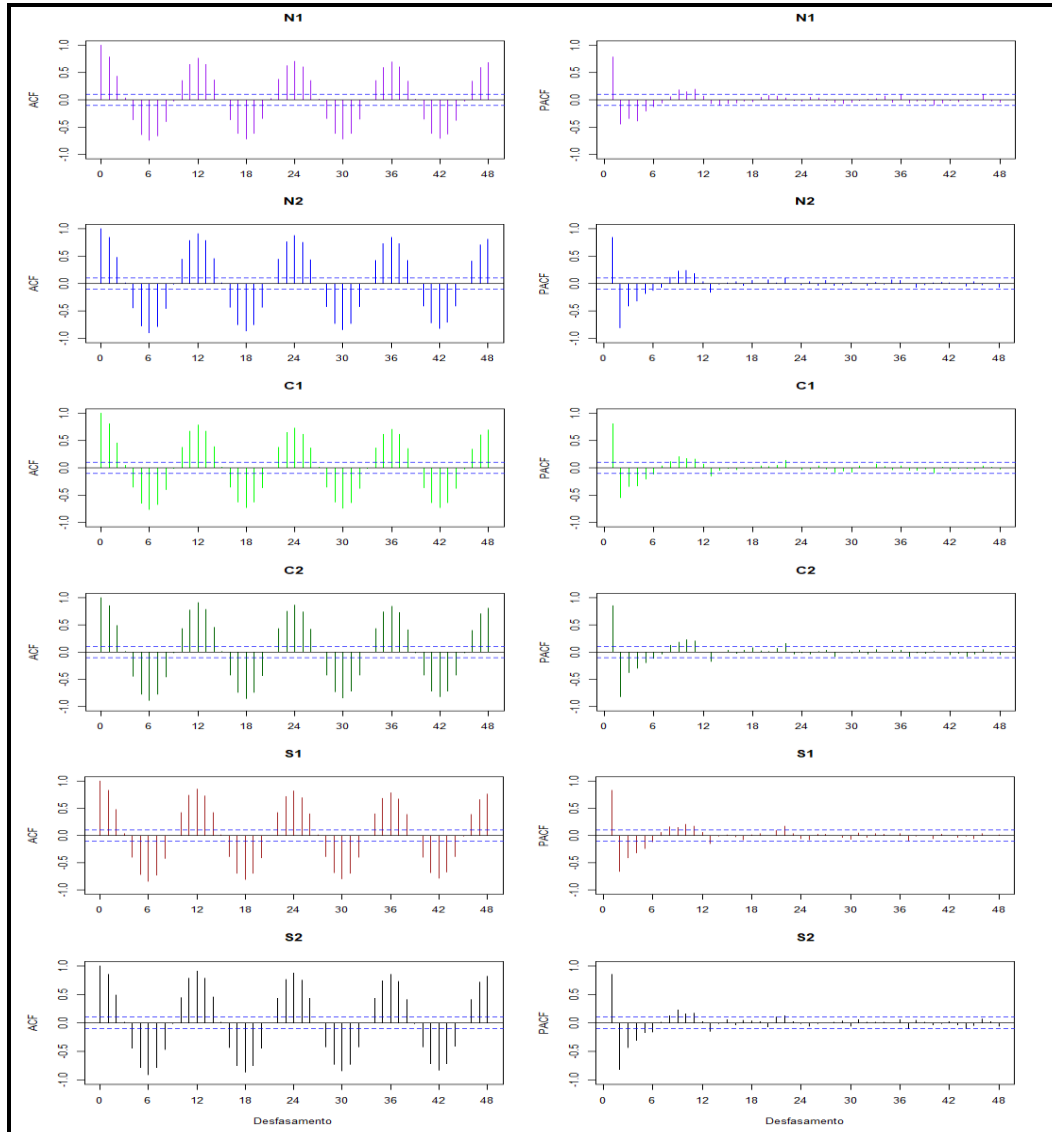


Figura 3.3: Representações gráficas da ACF e PACF.<sup>11</sup>

A sazonalidade é identificada pela forte autocorrelação para desfasamentos com  $k = 12$  (ou  $k = 12 \times i$  e  $i \in \mathbb{N}$ ) observada para as seis séries temporais (Quadro 3.1).

Quadro 3.1 – Valores da ACF para  $k$  igual a 6, 12, 18 e 24.

Região	$K = 6$	$K = 12$	$K = 18$	$K = 24$
Norte 1	-0,738	0,757	-0,714	0,700
Norte 2	-0,892	0,913	-0,862	0,871
Centro 1	-0,759	0,785	-0,722	0,730
Centro 2	-0,888	0,905	-0,848	0,866
Sul 1	-0,835	0,852	-0,805	0,821
Sul 2	-0,903	0,911	-0,866	0,875

<sup>11</sup> Fonte: Autor.

A autocorrelação segue um padrão sazonal oscilatório com valores máximos para  $k = 12 \times i$  e mínimos para  $k = 6 \times i$ . Ou seja, os valores da SST distantes anualmente possuem uma forte autocorrelação positiva e os distantes em meio ano uma forte autocorrelação negativa. Por este motivo conclui-se que as séries temporais em estudo não são estacionárias. As oscilações observadas nos gráficos da ACF tem uma frequência periódica associada a doze desfasamentos ( $s = 12$ ) o que é um indicativo da frequência das flutuações sazonais das séries temporais em estudo, ou seja, da existência de um padrão sazonal acentuado.

A observação dos diagramas de extremos e quartis (Figura 3.4) demonstra um padrão sazonal semelhante para as três regiões.

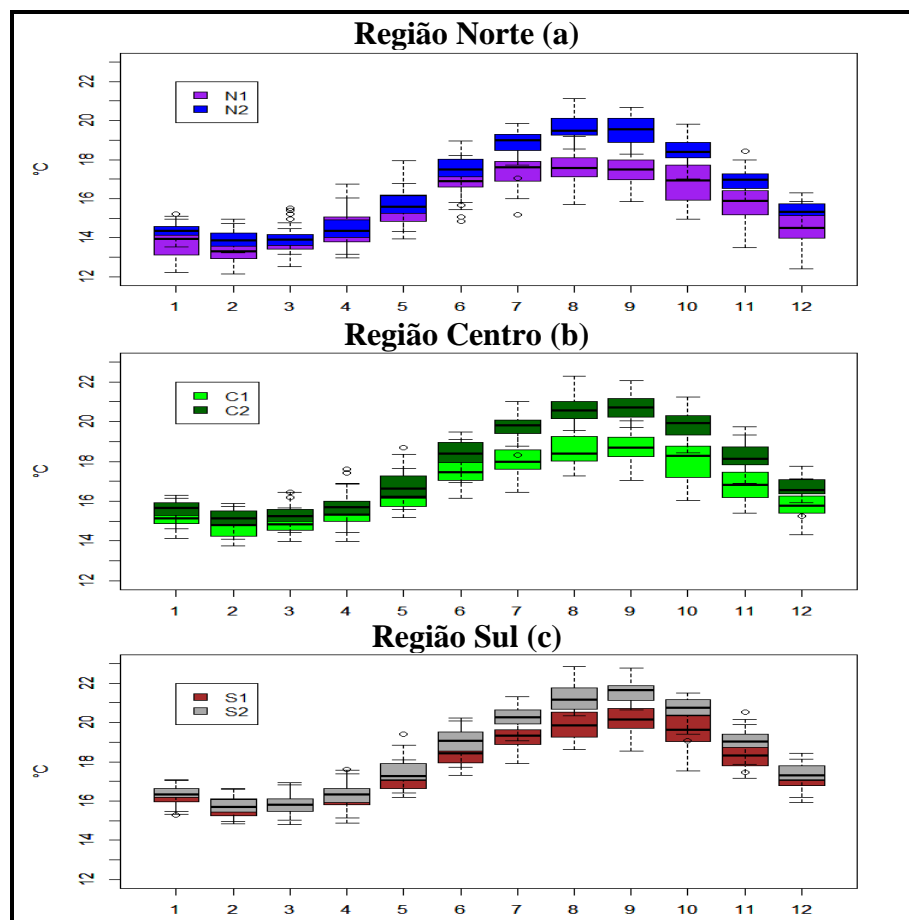


Figura 3.4: Caixa de bigodes da SST para as regiões (a) Norte, (b) Centro e (c) Sul.<sup>12</sup>

Os valores da SST aumentam progressivamente a partir do mês de fevereiro até aos seus valores máximos nos meses de agosto e setembro, quando decrescem novamente até aos seus valores mínimos no mês de fevereiro. Nos meses mais quentes a SST das localizações

<sup>12</sup> Fonte: Autor.

costeiras não sofre o mesmo padrão de aumento da SST das respectivas localizações distantes da costa. Esta diferença foi observada nos gráficos das séries em estudo e pode ter como causa o efeito do afloramento costeiro evidenciado na região costeira portuguesa descrito em outros estudos relativos à costa portuguesa e região Ibérica Atlântica (Goela *et al*, 2016; Rubal *et al*, 2013).

Em suma, as séries temporais em estudo possuem um comportamento bem definido, com um forte padrão sazonal anual cuja frequência periódica é  $s = 12$ . Para as mesmas localizações (ou próxima, ou distante da costa) as temperaturas são mais altas na região Sul, seguidas pelas da região Centro e com temperaturas mais baixas na região Norte. Foi possível identificar um comportamento ao longo do tempo semelhante para as mesmas localizações nas três regiões, embora exista uma diferença no padrão sazonal entre as duas localizações de cada região.

### 3.3 O fenómeno do Afloramento

#### 3.3.1 Introdução

O fenómeno do *afloramento costeiro* (Figura 3.5) é caracterizado pela subida de águas mais profundas e frias para níveis mais superficiais, causado pela ação de ventos persistentes e paralelos à costa.

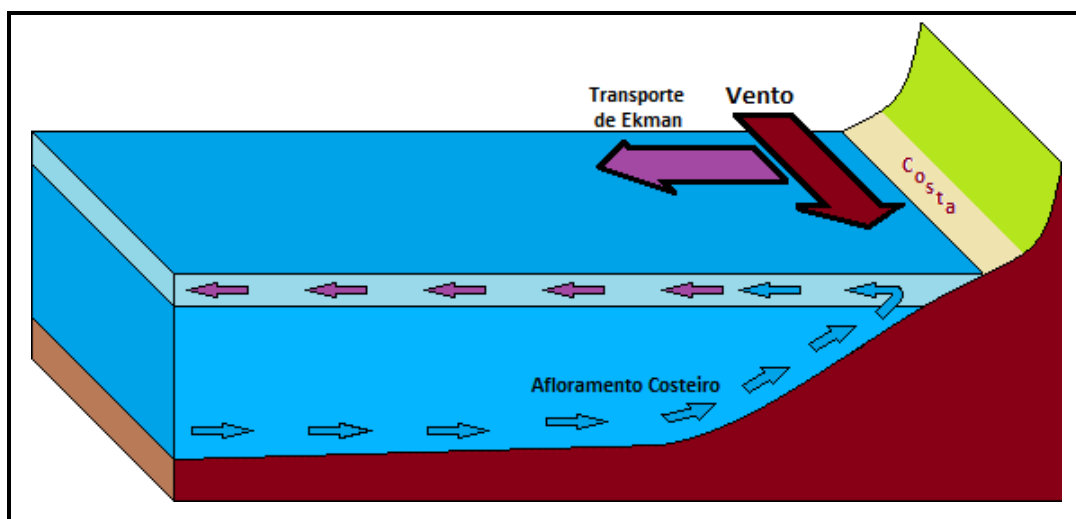


Figura 3.5: Afloramento costeiro no hemisfério Norte.<sup>13</sup>

<sup>13</sup> Fonte: Autor.

A costa portuguesa está localizada na parte norte do cinturão de alta pressão subtropical do hemisfério Norte e do conhecido sistema de afloramento da Corrente das Canárias que é influenciado por dois sistemas atmosféricos (o anticiclone dos Açores e a depressão da Islândia). Este fenómeno, embora relativamente pouco conhecido, é explicado pela *teoria de Ekman* (1905). Segundo esta teoria um fenómeno de divergência (com afloramento de águas subsuperficiais frias) ocorre na costa leste de um oceano se o vento que exerce atrito na superfície da água estiver orientado paralelamente à costa na direção do equador. Se o vento se desloca na direção oposta (na direção dos polos) o fenómeno de convergência é observado (com o afundamento das águas superficiais) (Fiúza *et al.*, 1982). A teoria de Ekman enuncia que quando a superfície da água sofre a ação horizontal do atrito causado pelo vento gera-se uma corrente perpendicular à direção do vento, para a direita no Hemisfério Norte e para a esquerda no Hemisfério Sul, a direção perpendicular da corrente é consequência do movimento de rotação da Terra (Cushman-Roisin e Beckers, 2011).

#### 3.3.2 Índice de Afloramento Relativo

O fenómeno do afloramento costeiro é de difícil quantificação, para isso foram criados índices de afloramento costeiro que permitem descrever o padrão sazonal da temperatura das águas próximas da costa explicado por este fenómeno. Entre diferentes métodos de contabilização do fenómeno do afloramento costeiro podem-se calcular *Índice de Afloramento (IA)*. Estes índices podem ser calculados a partir de dados do transporte de Ekman ou por valores de dados da SST. No primeiro caso os índices de afloramento costeiro são iguais ao valor negativo do transporte de Ekman, calculado a partir dos componentes da velocidade dos ventos a 10 m acima da superfície do mar, do coeficiente de arrasto, da densidade da água e do ar (Gonzalez-Nuevo *et al.*, 2014). No segundo caso os índices são obtidos pela diferença entre a SST da região costeira (que sofre o efeito do afloramento costeiro) e a SST da região oceânica (em que este fenómeno não acontece) para a mesma latitude (Álvarez *et al.*, 2011).

Este trabalho propõe uma nova expressão para o cálculo do índice costeiro denominada por *índice de afloramento relativo (IAR)* que tem a seguinte expressão:

$$IAR = \frac{SST_1 - SST_2}{SST_2}.$$

Em que  $SST_1$  e  $SST_2$  são respetivamente as séries temporais das localizações próximas e distantes da costa. Considera-se que esta nova abordagem facilita a comparação dos resultados entre zonas geográficas com diferentes latitudes e é de fácil interpretação.

#### 3.3.3 A análise do IAR para cada região

Para cada uma das regiões foram definidos os seguintes índices relativos:  $IAR_{Norte}$ ,  $IAR_{Centro}$  e  $IAR_{Sul}$ . A análise gráfica destas séries temporais (Figura 3.6) permitiu identificar a presença de um padrão sazonal, de acordo com a figura abaixo.

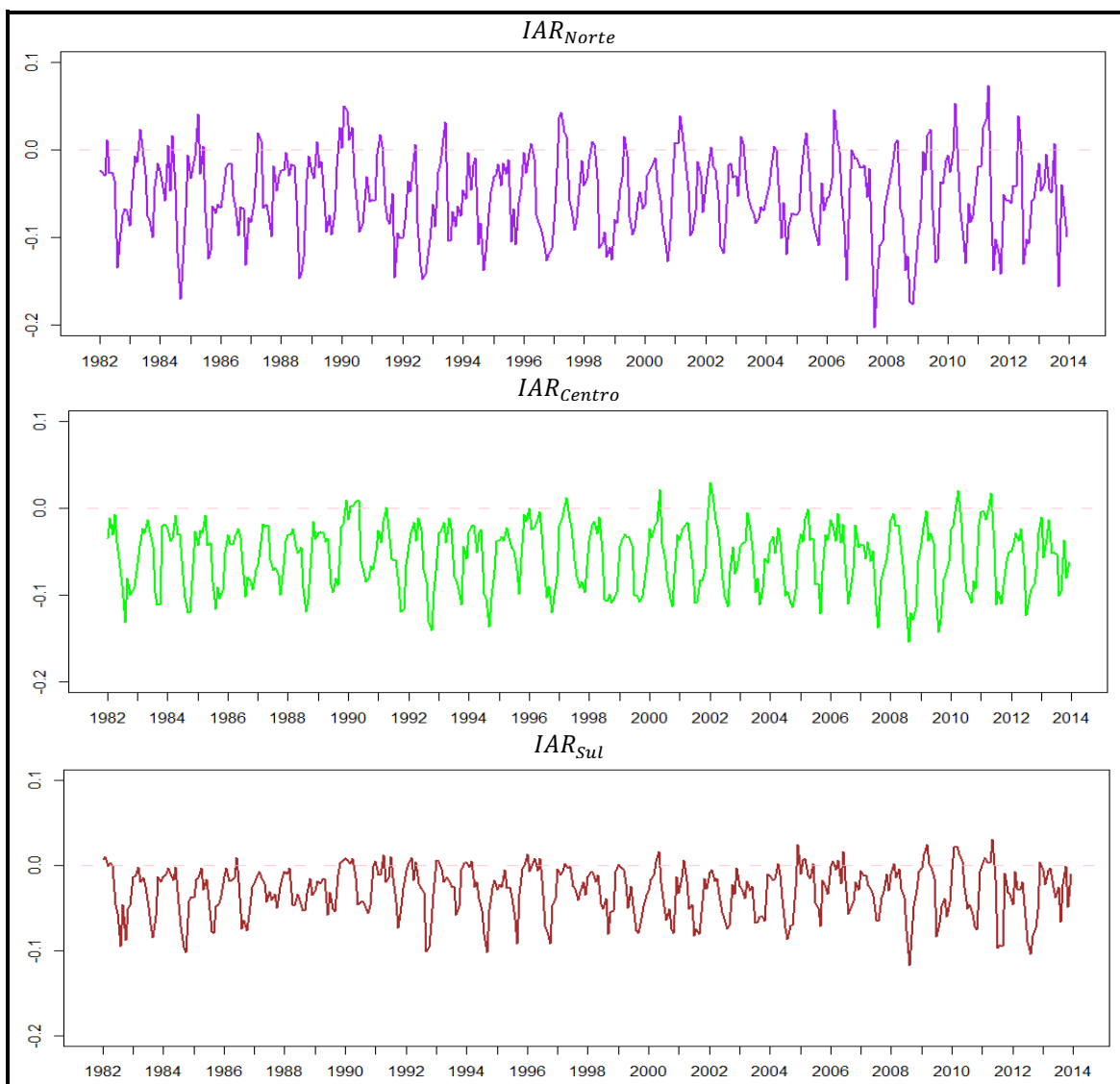


Figura 3.6: Representação gráfica do IAR.<sup>14</sup>

<sup>14</sup> Fonte: Autor.

A linha horizontal representa o valor em que ambas as séries temporais (costeira e distante da costa) possuem valores iguais da *SST*, e que, portanto, o efeito do afloramento costeiro é nulo. Valores negativos distantes de zero indicam que o efeito do afloramento costeiro é evidente. Valores positivos indicam o fenômeno oposto, o do afundamento costeiro causado por ventos que se dirigem para Norte.

Para as três séries temporais observa-se uma oscilação periódica entre valores mais próximos de zero nos meses mais frios e valores negativos nos meses mais quentes. Estes valores evidenciam que o fenômeno do afloramento costeiro é mais evidente nos meses de verão. Não é possível verificar a presença (ou ausência) de uma tendência a longo prazo.

Para confirmar a presença da sazonalidade, procedeu-se à análise dos gráficos da ACF e da PACF das séries temporais das três regiões (Figura 3.7).

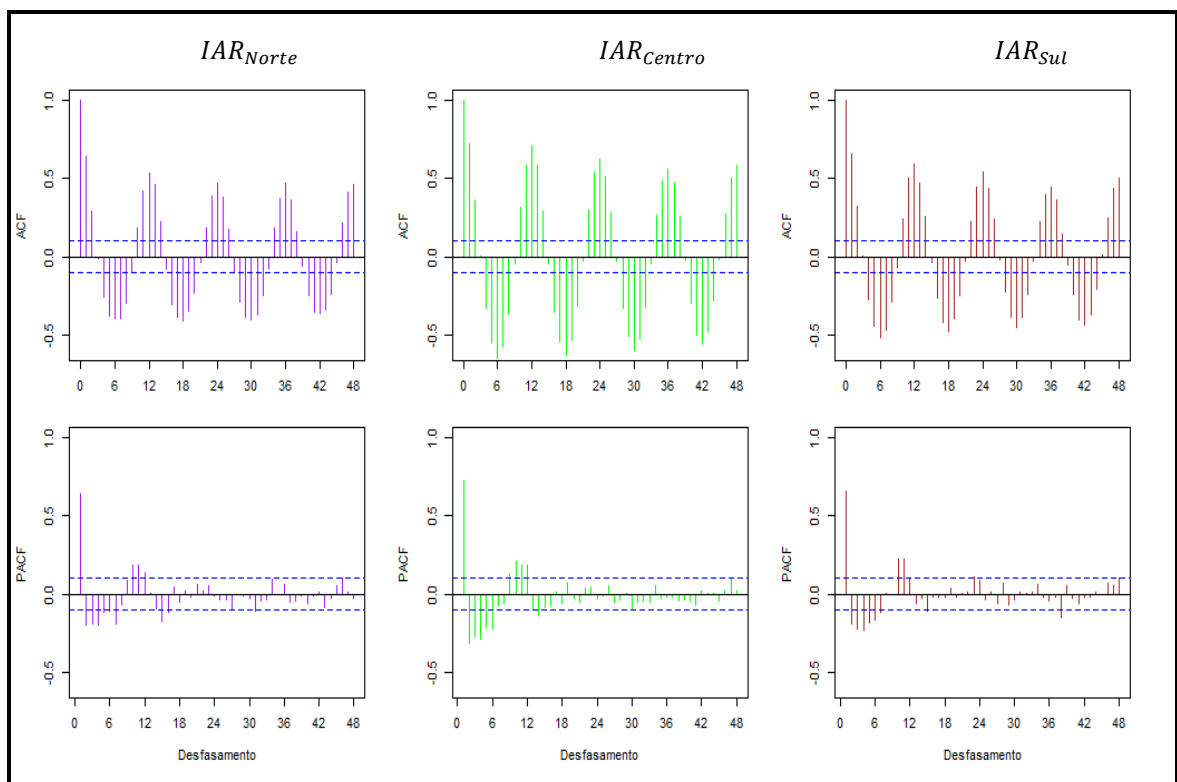


Figura 3.7: Representação da ACF e PACF da IAR.<sup>15</sup>

Os gráficos da ACF e PACF indicam que a existência de autocorrelação para a maioria dos desfasamentos  $k$ .

<sup>15</sup> Fonte: Autor.

Nos diagramas de extremos e quartis (Figura 3.8) podemos observar a evolução anual do IAR para cada uma das três regiões.

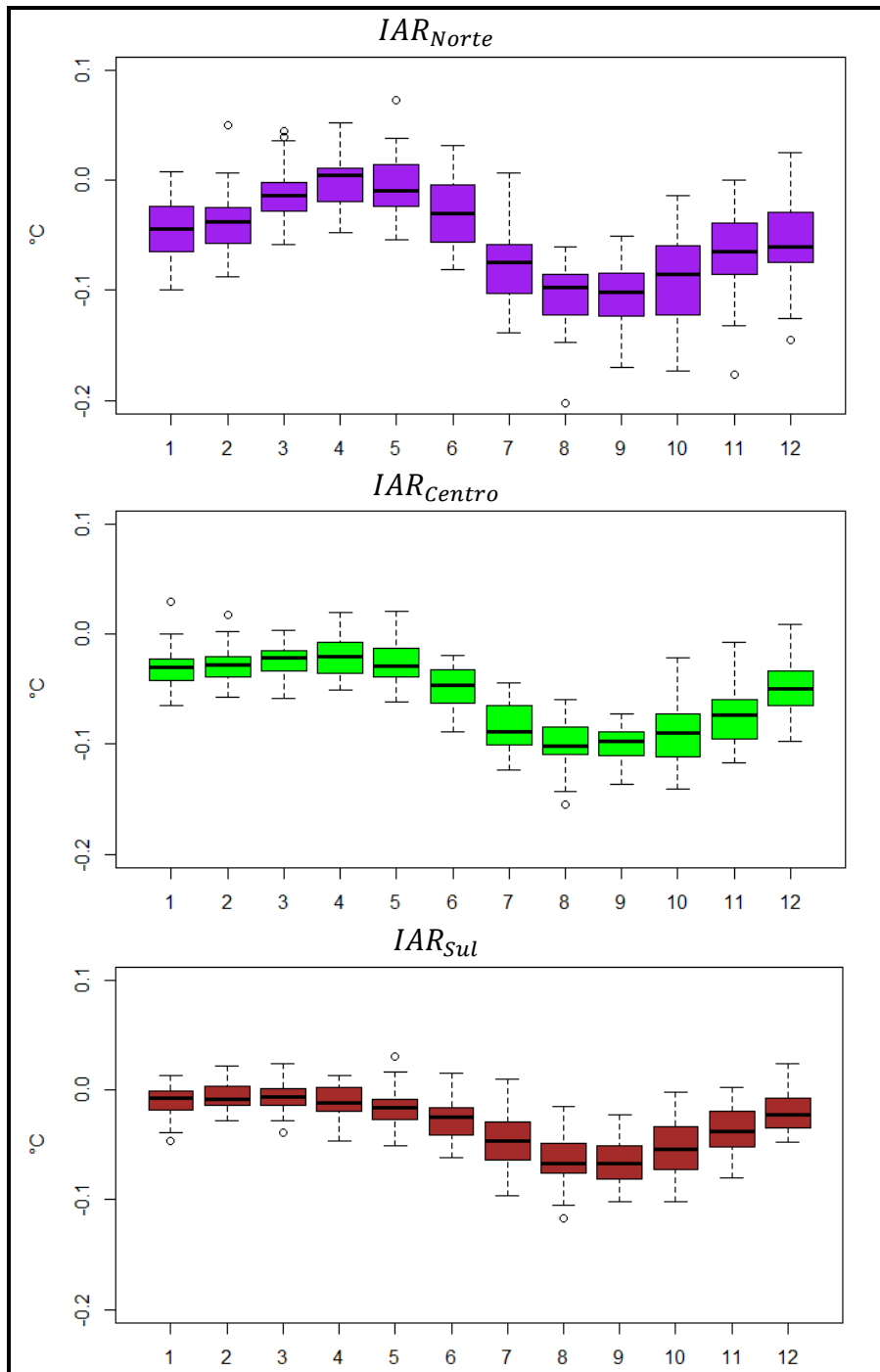


Figura 3.8: Caixa de bigodes do IAR.<sup>16</sup>

O padrão sazonal do IAR na região Norte é mais evidente, seguido pela região Centro e Sul. Na região Norte pode ser observado uma oscilação mais acentuada do IAR em que

<sup>16</sup> Fonte: Autor.

o fenómeno do afloramento costeiro é muito baixo no mês de abril a partir do qual aumenta visivelmente até ao valor máximo da mediana no mês de setembro. O padrão sazonal é menos acentuado nas regiões Centro e Sul, embora seja visível que o afloramento costeiro é mais intenso no verão, com as medianas menores nos meses de agosto e setembro (o que indica maior intensidade do afloramento costeiro) e menos intenso para os meses de março e de abril. Podem ser observados *outliers* nas três regiões (sete para a região Norte, três para a região Centro e quatro para a região Sul).

#### 3.3.4 Modelação do IAR de cada região

A próxima fase consiste em encontrar o modelo que melhor descreve o comportamento do IAR em cada região. Recorrendo às metodologias descritas no capítulo anterior, ajustaram-se modelos de decomposição clássica, alisamento exponencial de Holt-Winters e *ARIMA*. O melhor modelo será aquele que possuir o menor valor da medida de erro MAE.

No caso do modelo de Holt-Winters do tipo multiplicativos dados têm que ser diferentes de zero. Assim, para contornar esta situação é adicionada uma constante à série temporal de valor igual a um menos o mínimo da série, ou seja, a transformação a executar será:

$$IAR^* = IAR - \min(IAR) + 1. \quad (15)$$

#### *Modelos de decomposição clássica*

Ajustaram-se dois modelos de decomposição clássica (Figura 3.9), um do tipo aditivo (7) e do outro tipo multiplicativo (8). O ajuste do modelo do tipo multiplicativo foi feito por não ser bem claro se a amplitude sazonal varia proporcionalmente com o valor médio dos dados. A tendência foi estimada e retirada pelo método das médias móveis (1). Da série obtida retirando-se a tendência procedeu-se ao cálculo dos índices sazonais (3). A componente irregular ( $I_t$ ) deste modelo é obtida pela extração das componentes tendência e sazonalidade estimadas.

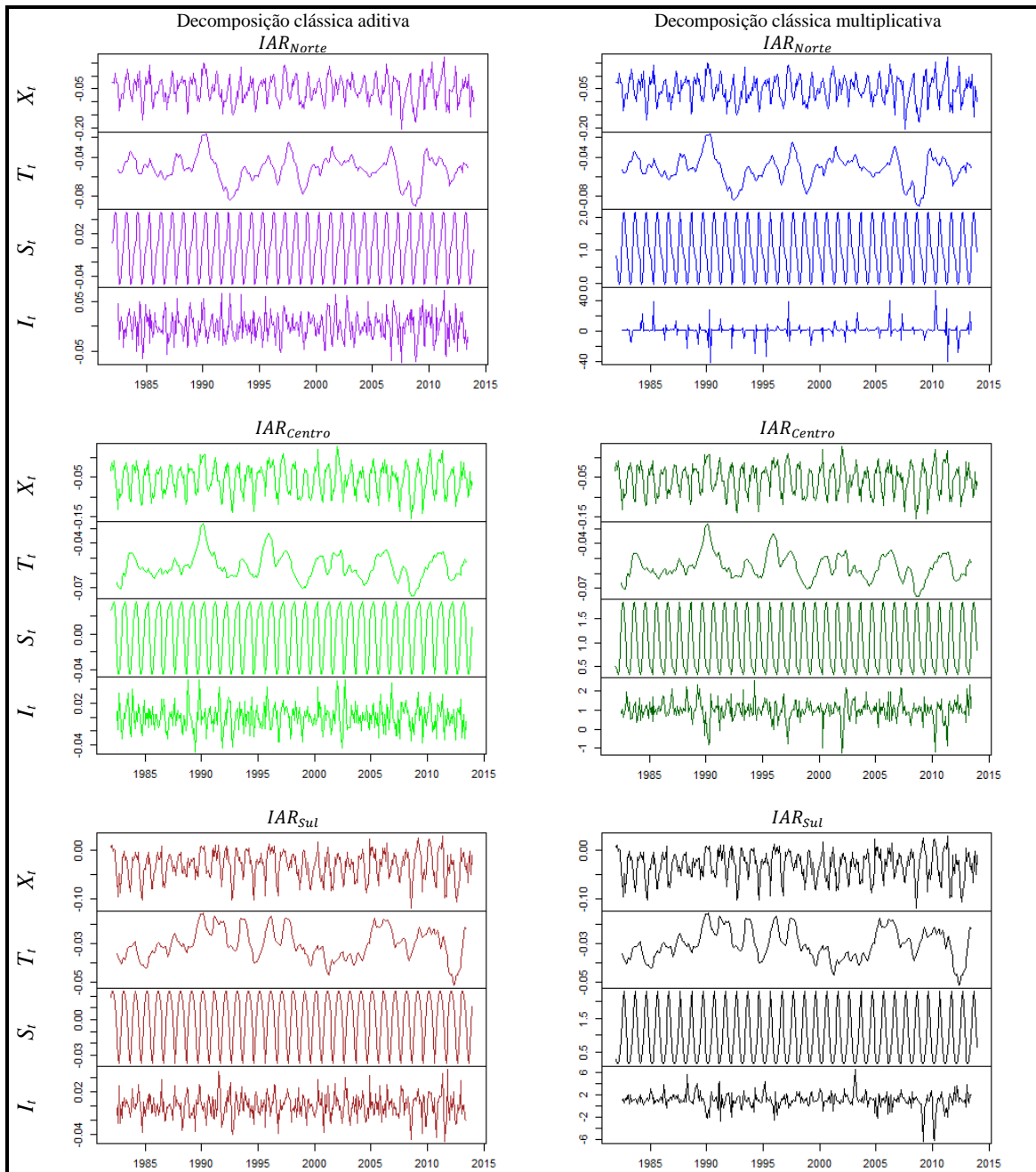


Figura 3.9: Decomposição clássica para a séries temporais IAR.<sup>17</sup>

O valor de um índice sazonal ( $s_i$ ) da série temporal  $IAR$  é um indicativo da intensidade do afloramento costeiro do respetivo mês, em termos relativos.

Os índices sazonais obtidos na decomposição clássica aditiva e multiplicativa são apresentados no Quadro 3.2.

<sup>17</sup> Fonte: Autor.

Quadro 3.2 – Índices sazonais dos modelos de decomposição aditivo e multiplicativo.

	Aditivo			Multiplicativo		
	$IAR_{Norte}$	$IAR_{Centro}$	$IAR_{Sul}$	$IAR_{Norte}$	$IAR_{Centro}$	$IAR_{Sul}$
$S_1$	0,0069	0,0265	0,0194	0,8204	0,5043	0,2986
$S_2$	0,0119	0,0282	0,0241	0,6820	0,4853	0,1746
$S_3$	0,0420	0,0325	<b>0,0242</b>	0,1140	0,4127	<b>0,1672</b>
$S_4$	<b>0,0503</b>	<b>0,0361</b>	0,0207	<b>-0,0245</b>	<b>0,3280</b>	0,2642
$S_5$	0,0490	0,0303	0,0131	0,0364	0,4450	0,5764
$S_6$	0,0234	0,0083	0,0046	0,6033	0,8737	0,8314
$S_7$	-0,0267	-0,0299	-0,0174	1,5943	1,5541	1,5721
$S_8$	<b>-0,0521</b>	<b>-0,0447</b>	-0,0331	<b>2,1081</b>	1,8230	2,0967
$S_9$	-0,0507	<b>-0,0447</b>	<b>-0,0370</b>	2,1422	<b>1,8507</b>	<b>2,3110</b>
$S_{10}$	-0,0372	-0,0364	-0,0236	1,7334	1,6643	1,8453
$S_{11}$	-0,0144	-0,0138	-0,0060	1,2349	1,2247	1,2101
$S_{12}$	-0,0025	-0,0077	-0,0110	0,9554	0,8342	0,6523

A inspeção dos gráficos das ACF e PACF da componente irregular do modelo de decomposição clássica do tipo aditivo e do tipo multiplicativo (Figura 3.10) mostra a existência de autocorrelações significativas para alguns dos desfasamentos  $k$ .

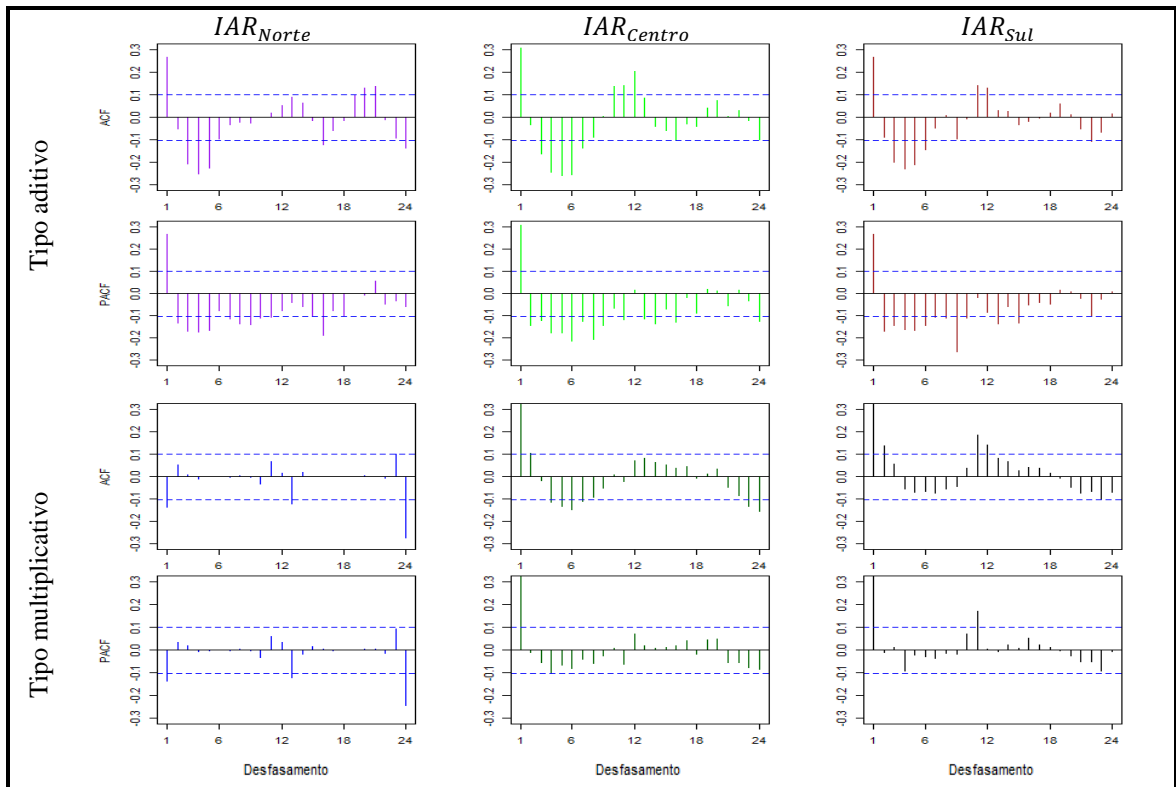


Figura 3.10: ACF e PACF da componente irregular.<sup>18</sup>

<sup>18</sup> Fonte: Autor.

Podemos completar a informação através do teste Ljung-Box (13). Os resultados deste teste (valores- $p \leq 0,0013$  para  $m = 24$ ) permitem para o nível de significância de 0,05 rejeitar a hipótese nula de que a autocorrelação conjunta da componente irregular seja zero para estes desfasamentos. Assim, conclui-se que os resíduos destes modelos não são independentes.

#### *Modelos de alisamento exponencial de Holt-Winters*

Foram ajustados os modelos de *Holt-Winters* (Cap. II - 2.3.2), um do tipo aditivo às séries temporais  $IAR_{Região}$  e outro do tipo multiplicativo para as séries temporais  $IAR^*_{Região}$ , esta última sujeita à transformação definida em (15). Na Figura 3.11 pode-se observar o ajustamento dos modelos.

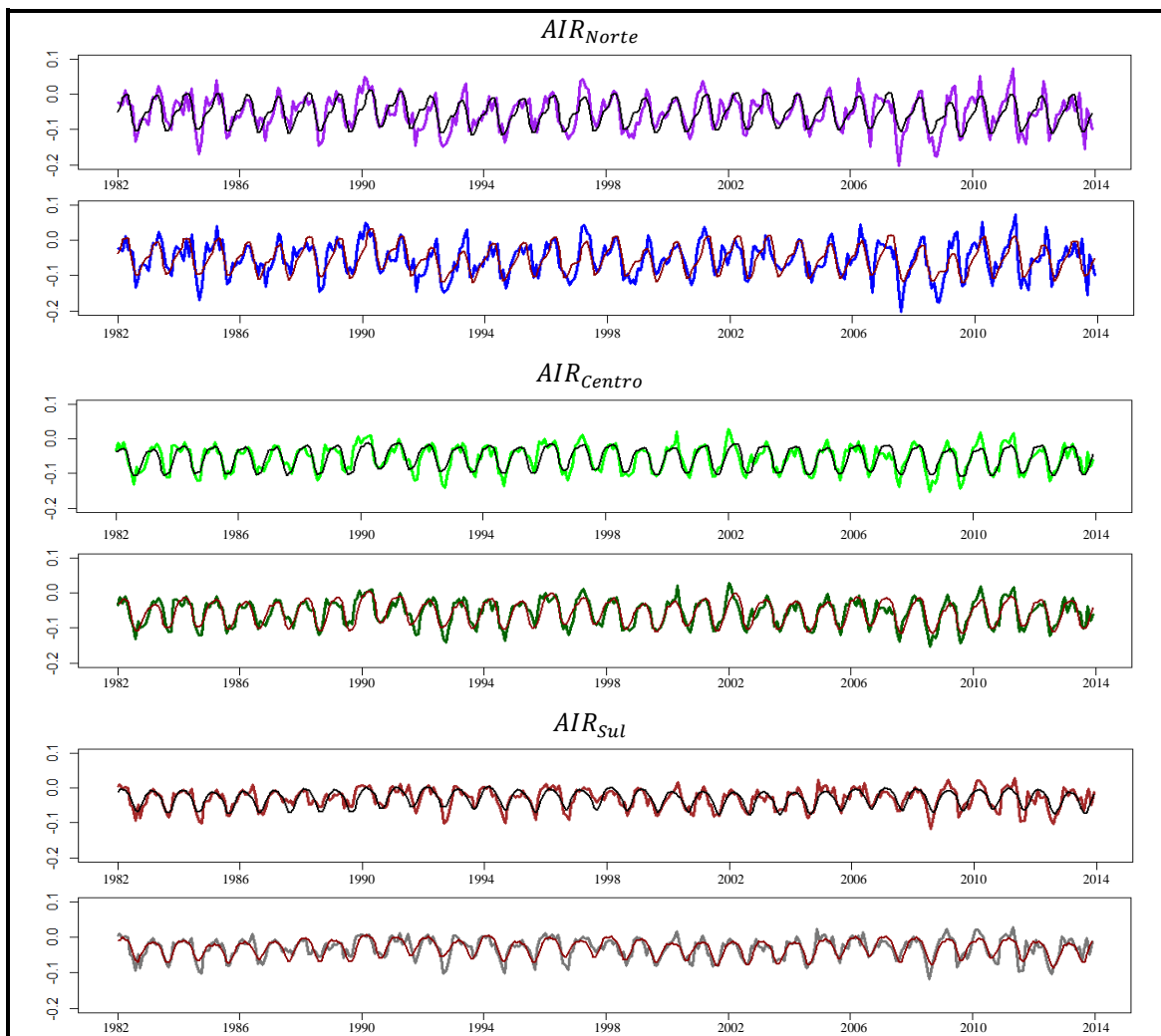


Figura 3.11: Representação gráfica dos modelos de Holt-Winters do tipo aditivo (a preto) e do tipo multiplicativo (a vermelho) e das séries temporais do IAR das regiões Norte, Centro e Sul.<sup>19</sup>

<sup>19</sup> Fonte: Autor.

A análise dos gráficos sugere um bom ajustamento em ambos os modelos.

Os resultados do alisamento exponencial de Holt-Winters do tipo aditivo dos  $IAR$  e do tipo multiplicativo dos  $IAR^*$  das três regiões encontram-se no Quadro 3.3.

Quadro 3.3 – Resultados do alisamento exponencial de Holt – Winters.

	Aditivo			Multiplicativo		
	$IAR^*_{Norte}$	$IAR^*_{Centro}$	$IAR^*_{Sul}$	$IAR^*_{Norte}$	$IAR^*_{Centro}$	$IAR^*_{Sul}$
$\alpha$	0,0459	0,0552	0,0595	0,1264	0,1457	0,1462
$\beta$	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001
$\gamma$	0,0417	0,0466	0,0418	0,0001	0,0001	0,0002
$l_0$	-0,0527	-0,0636	-0,0308	1,1610	1,0990	1,0874
$b_0$	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
$s_{1_0}$	0,0257	0,0079	0,0042	1,0180	1,0078	1,0036
$s_{2_0}$	0,0521	0,0298	0,0127	1,0448	1,0286	1,0140
$s_{3_0}$	0,0494	0,0353	0,0228	1,0434	1,0346	1,0182
$s_{4_0}$	0,0419	0,0321	0,0245	1,0380	1,0322	1,0224
$s_{5_0}$	0,0120	0,0276	0,0266	1,0109	1,0242	1,0182
$s_{6_0}$	0,0028	0,0274	0,0198	1,0042	1,0225	1,0188
$s_{7_0}$	-0,0004	0,0086	0,0110	0,9969	1,0069	1,0105
$s_{8_0}$	-0,0136	-0,0157	-0,0084	0,9870	0,9859	0,9942
$s_{9_0}$	-0,0381	-0,0367	-0,0273	0,9689	0,9676	0,9793
$s_{10_0}$	-0,0519	-0,0425	-0,0369	0,9552	0,9584	0,9662
$s_{11_0}$	-0,0527	-0,0448	-0,0334	0,9532	0,9583	0,9688
$s_{12_0}$	-0,0272	-0,0289	-0,0156	0,9795	0,9729	0,9858

A constante  $\beta$  é aproximadamente igual a zero nos dois tipos de modelo e para as três regiões o que é um indicativo de que  $b_t$  não muda no horizonte temporal. A constante  $\gamma$  também é aproximadamente nula para o modelo de Holt-Winters do tipo multiplicativo, o que é um indicativo de que  $s_t$  é constante no horizonte temporal para este tipo de modelo.

Os gráficos das ACF e PACF dos resíduos (Figura 3.12) são importantes na verificação autocorrelação nula dos resíduos e na interpretação do seu comportamento. Para os três modelos foi identificada uma forte autocorrelação para o primeiro desfaseamento.

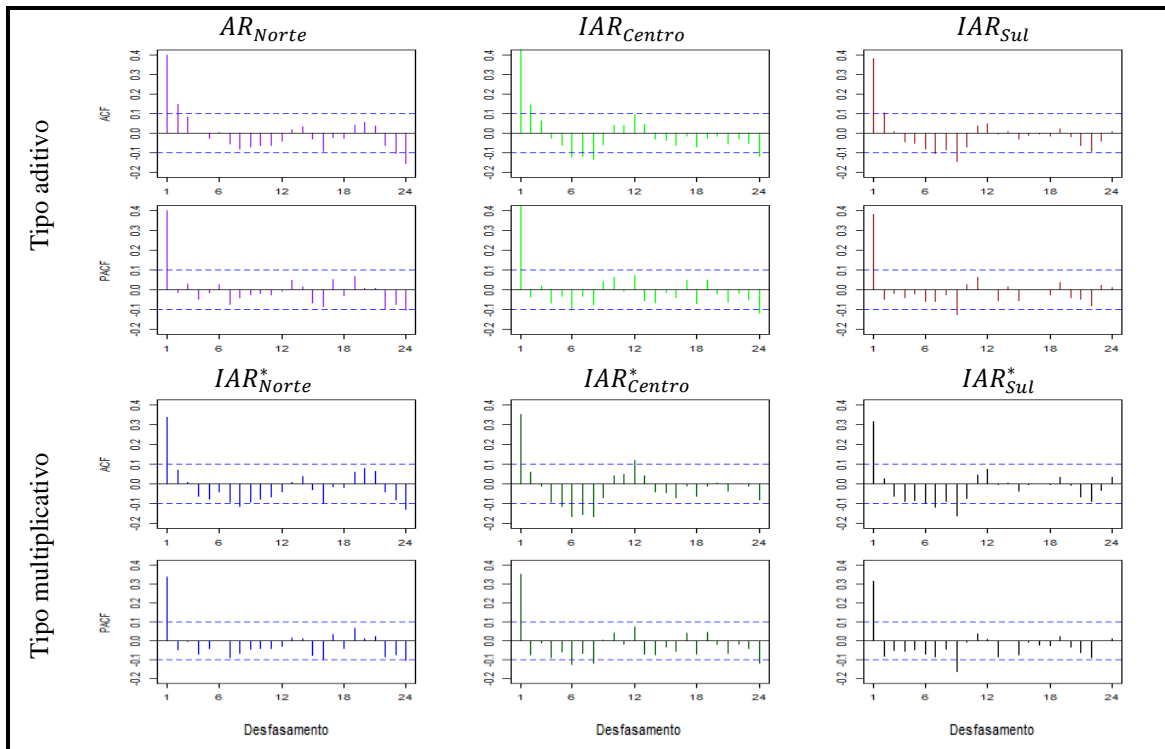


Figura 3.12: ACF e PACF dos resíduos.<sup>20</sup>

Os resultados do teste de Ljung-Box (valores- $p \leq 3,1530 \times 10^{-14}$  com  $m = 24$  para todos os modelos) indicam que os resíduos não são independentes.

Para contornar o facto de existir uma autocorrelação forte poder-se-ia ajustar os resíduos por um modelo  $AR(1)$ , pois a ACF parece diminuir exponencialmente (com oscilação) e a PACF anula-se para defasagens maiores que um (Alpuim, 2003).

### Modelos ARIMA

Antes de se ajustar os modelos  $ARIMA$  para a  $IAR$  de cada região, foram estimadas as componentes tendência (modelo linear) e  $S_t$  (pelo cálculo dos índices sazonais  $s_i$ ) como indicado Cap. II - 2.1.2.

A estimação da tendência é obtida com o auxílio de uma regressão linear dada pela equação (2). Os resultados podem ser observados no Quadro 3.4.

Quadro 3.4 – Declive e termo independente do modelo de regressão linear.

	$IAR_{Norte}$	$IAR_{Centro}$	$IAR_{Sul}$
Declive	$-2,1731 \times 10^{-5}$	$-1,3795 \times 10^{-5}$	$-4,2303 \times 10^{-6}$
Termo Independente	-0,0481	-0,0529	-0,0293

<sup>20</sup> Fonte: Autor.

Foi retirada a tendência linear obtendo-se uma série temporal sem a tendência dada por:

$$IAR_{região_t}^{-T} = IAR_{região_t} - b_0 - b_1t.$$

A partir da nova série  $IAR_{região_t}^{-T}$  calcularam-se os índices sazonais  $s_i$  (Quadro 3.5).

Quadro 3.5 – Índices Sazonais.

	$s_1$	$s_2$	$s_3$	$s_4$	$s_5$	$s_6$	$s_7$	$s_8$	$s_9$	$s_{10}$	$s_{11}$	$s_{12}$
Norte	0,007	0,012	0,041	0,050	0,048	0,023	-0,024	-0,052	-0,052	-0,036	-0,014	-0,004
Centro	0,026	0,029	0,032	0,036	0,030	0,008	-0,029	-0,045	-0,045	-0,035	-0,014	-0,007
Sul	0,019	0,024	0,024	0,021	0,013	0,004	-0,017	-0,034	-0,036	-0,022	-0,007	-0,011

Retiram-se os índices sazonais  $s_i$  obtendo-se as séries temporais:

$$IAR_{região_t}^{-T-S} = IAR_{região_t}^{-T} - s_i.$$

Temos então o modelo ajustado ao  $IAR$  de cada região dado por:

$$\widehat{IAR}_{região_t} = s_i + b_0 + b_1t + \widehat{IAR}_{região_t}^{-T-S} \text{ para } \forall t \in \mathbf{T}.$$

Na Figura 3.13 podem ser visualizados os modelos ajustados às séries temporais  $AIR$  das diferentes regiões.

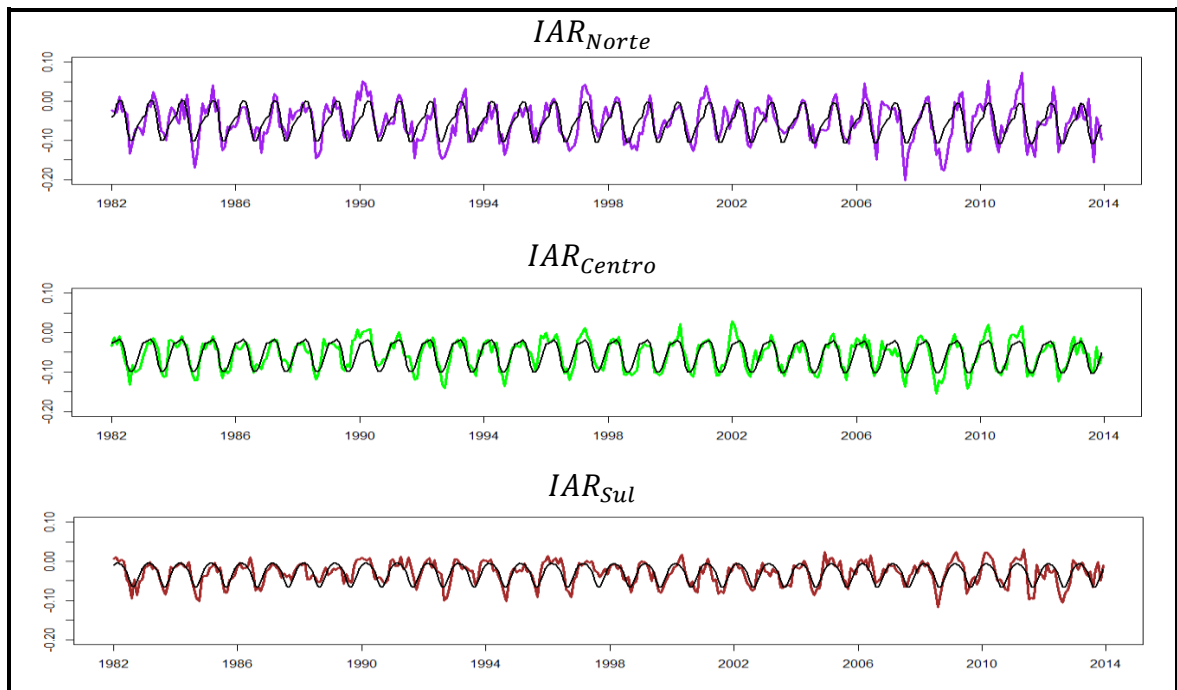


Figura 3.13: Representação gráfica do IAR e dos respectivos modelos da tendência linear+índices sazonais (a preto).<sup>21</sup>

<sup>21</sup> Fonte: Autor.

Foram calculados os valores da ACF e PACF da  $IAR^{-T-S}$ . Tanto a análise gráfica (Figura 3.14) quanto os resultados do teste (os valores-p são de  $4,77 \times 10^{-15}$  na região Norte, menor que  $2,20 \times 10^{-16}$  na região Centro e de  $8,43 \times 10^{-14}$  na região Sul para  $m = 10$ ) indicam que os resíduos não são independentes.

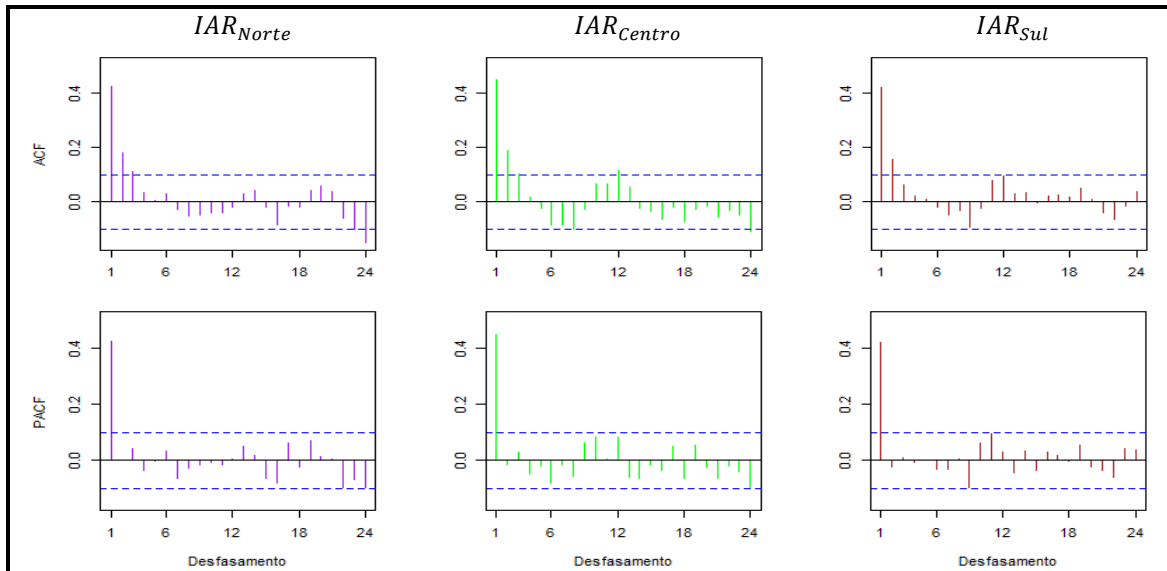


Figura 3.14: ACF e PACF das séries temporais  $IAR^{-T-S}$ .<sup>22</sup>

Os valores da ACF e da PACF sugerem que as séries temporais do  $IAR$  sem a tendência linear e a sazonalidade são processos do tipo autorregressivo (as ACF decaem exponencialmente) e com ordem  $p = 1$  (as PACF anulam-se para valores de  $k > 1$ ). Estas observações indicam assim que um modelo do tipo  $AR(1)$  pode ser o mais adequado para descrever os dados.

O teste de Dickey-Fuller aumentado (14) foi realizado para verificar a existência (ou não) de pelo menos uma raiz dentro do círculo unitário e assim determinar a ausência (ou presença) de estacionaridade. Os valores-p para as componentes  $IAR_t^{-T-S}$  são iguais a 0,01, o que permite ao nível de significância de 0,05 rejeitar a hipótese nula de que  $\phi_1 = 1$ . Aceita-se então que  $H_1: \phi_1 < 1$  é verdadeiro, ou seja, todas as raízes estão fora do círculo unitário e que, portanto, ambos os processos são estacionários. Com base nestes resultados não há necessidade de fazer diferenciação ( $d$ ).

<sup>22</sup> Fonte: Autor.

### Capítulo III – Estudo de Séries Temporais de Dados Oceanográficos

Ajustaram-se vários modelos com diferentes ordens  $p, d, q$ , para cada uma das três séries temporais. O método utilizado foi o dos mínimos quadrados para estimar os parâmetros do modelo.

No Quadro 3.6 pode observar-se os diferentes modelos utilizados e os valores dos critérios de seleção AIC ( 9), AICc ( 10) e BIC ( 11). O modelo selecionado encontra-se a verde.

Quadro 3.6 – Critérios de seleção AIC, AICc e BIC dos modelos ARIMA.

ARIMA ( $p, d, q$ )	$IAR_{Norte}^{-T-S}$			$IAR_{Norte}^{-T-S}$			$IAR_{Norte}^{-T-S}$		
	AIC	AICc	BIC	AIC	AICc	BIC	AIC	AICc	BIC
(1,0,0)	-1638,12	-1638,05	-1626,26	-1976,66	-1976,60	-1964,81	-2028,15	-2028,09	-2016,30
(0,1,0)	-1628,69	-1628,63	-1616,84	-1966,76	-1966,70	-1954,91	-2020,16	-2020,09	-2008,31
(0,0,1)	-1507,75	-1507,74	-1503,80	-1851,51	-1851,50	-1847,57	-1895,11	-1895,10	-1891,16
(1,1,0)	-1538,64	-1538,61	-1530,74	-1876,93	-1876,90	-1869,04	-1923,01	-1922,98	-1915,11
(1,0,1)	-1636,12	-1636,01	-1620,31	-1974,78	-1974,68	-1958,98	-2026,36	-2026,25	-2010,55
(0,1,1)	-1574,35	-1574,32	-1566,46	-1903,70	-1903,67	-1895,80	-1958,62	-1958,59	-1950,72
(2,0,0)	-1636,12	-1636,01	-1620,31	-1974,77	-1974,66	-1958,96	-2026,35	-2026,24	-2010,54
(0,2,0)	-1142,49	-1142,48	-1138,54	-1492,11	-1492,10	-1488,17	-1532,30	-1532,29	-1528,35
(0,0,2)	-1632,84	-1632,73	-1617,04	-1971,41	-1971,30	-1955,60	-2024,89	-2024,79	-2009,09
(1,1,1)	-1627,99	-1627,93	-1616,15	-1965,74	-1965,67	-1953,89	-2016,99	-2016,93	-2005,14

O modelo que apresenta melhor ajuste segundo os três critérios de seleção para os  $IAR^{-T-S}$  das três regiões é o modelo ARIMA (1,0,0), que corresponde a um modelo autorregressivo de primeira ordem.

No Quadro 3.7 podem ser observados os valores estimados dos parâmetros e o respetivo erro padrão para cada um dos modelos AR(1), definidos pela equação ( 6).

Quadro 3.7 – Parâmetros estimadas para os modelos AR(1).

	Região Norte	Região Centro	Região Sul
$\phi_1$	0,4226	0,4488	0,4183
$\sigma$	0,0285	0,0184	0,0172

Os modelos propostos são os modelos AR(1) ,  $\forall t \in \mathbf{T}$ :

$$\widehat{IAR}_{Norte_t}^{-T-S} = 0,4226 \widehat{IAR}_{Norte_{t-1}}^{-T-S} + e_t,$$

$$\widehat{IAR}_{Centro_t}^{-T-S} = 0,4488 \widehat{IAR}_{Centro_{t-1}}^{-T-S} + e_t,$$

$$\widehat{IAR}_{Sul_t}^{-T-S} = 0,4183 \widehat{IAR}_{Sul_{t-1}}^{-T-S} + e_t.$$

Na Figura 3.15 pode analisar-se os resíduos e verificar se os modelos são adequados.

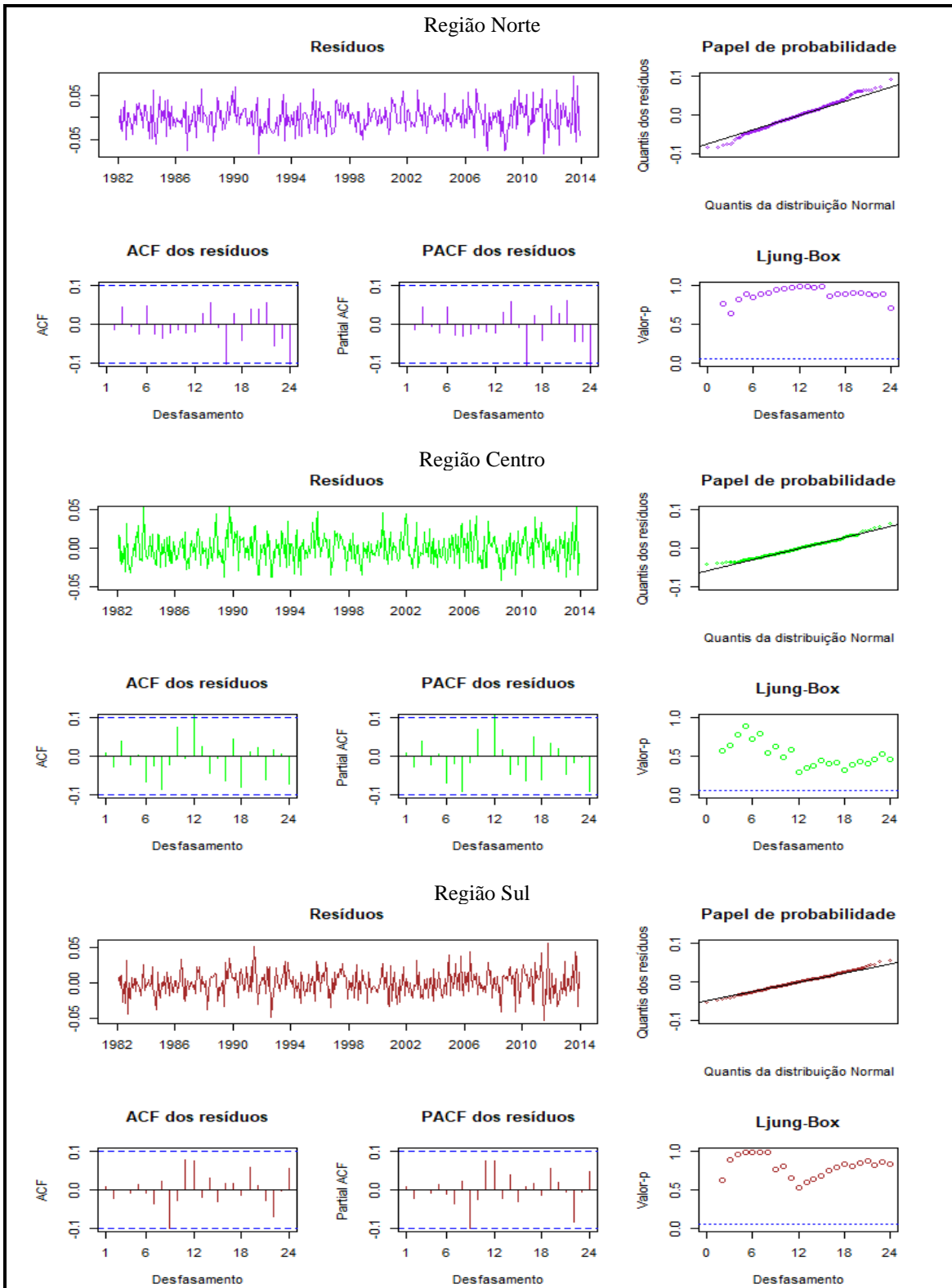


Figura 3.15: Gráficos dos resíduos, ACF, Papel de Probabilidade e valores-p do teste de Ljung-Box.<sup>23</sup>

<sup>23</sup> Fonte: Autor.

Os gráficos dos resíduos dos modelos  $AR(1)$  não sugerem a existência de um padrão e em geral a variância mantém-se constante no horizonte temporal das três regiões. Através da observação do gráfico de papel de probabilidade pode-se concluir que os resíduos das três séries temporais têm uma distribuição aproximadamente normal. Não se verificam autocorrelações significativas para nenhum dos 24 primeiros desfasamentos a partir dos gráficos das ACF e PACF, este resultado foi confirmado pelo teste de Ljung-Box (para  $m = 10$  o valor-p é igual a 0,9666 na região Norte, 0,4878 na região Centro e 0,8124 para a região) para o nível de significância de 0,05.

Em suma, os modelos para as séries temporais do  $IAR$  são dados por:

$$\widehat{IAR}_{Norte_t} = s_i - 0,0481 - 2,1731 \times 10^{-5} t + 0,4226 \widehat{IAR}_{Norte_{t-1}} + e_t,$$

$$\widehat{IAR}_{Centro_t} = s_i - 0,0529 - 1,3795 \times 10^{-5} t + 0,4488 \widehat{IAR}_{Centro_{t-1}} + e_t,$$

$$\widehat{IAR}_{Sul_t} = s_i - 0,0293 - 4,2303 \times 10^{-6} t + 0,4183 \widehat{IAR}_{Sul_t} + e_t.$$

Os valores dos  $s_i$  estão no Quadro 3.5.

Observando-se os coeficientes  $b_1$  da tendência linear (Quadro 3.4) verifica-se que os seus valores são aproximadamente nulos, o que leva a concluir que segundo este modelo os valores médios das séries do  $IAR$  não variam no horizonte temporal.

### 3.3.5 Interpretação dos resultados com base no melhor modelo

Comparando as medidas de erro MAE (12) (Quadro 3.8) podemos determinar o modelo que melhor se ajusta aos dados.

Quadro 3.8 – MAE dos modelos.

Modelo	$IAR_{Norte}$	$IAR_{Centro}$	$IAR_{Sul}$
Dec. Clássica aditivo	0,02159	0,01413	0,01276
Dec. Clássica multiplicativo	0,02348	0,01444	0,01312
Holt-Winters aditivo	0,02560	0,01614	0,01487
Holt-Winters multiplicativo	0,02478	0,01606	0,01452
$AR(1) + S_t + b_0 + b_1 t$	0,02222	0,01480	0,01352

Com base no quadro anterior, conclui-se que o modelo de decomposição clássica do tipo aditivo é modelo com menor valor de MAE, para as três regiões.

Realizando a representação dos índices sazonais (Figura 3.16) para as três regiões pode-se tirar conclusões quanto ao padrão do afloramento costeiro no decorrer do ano.

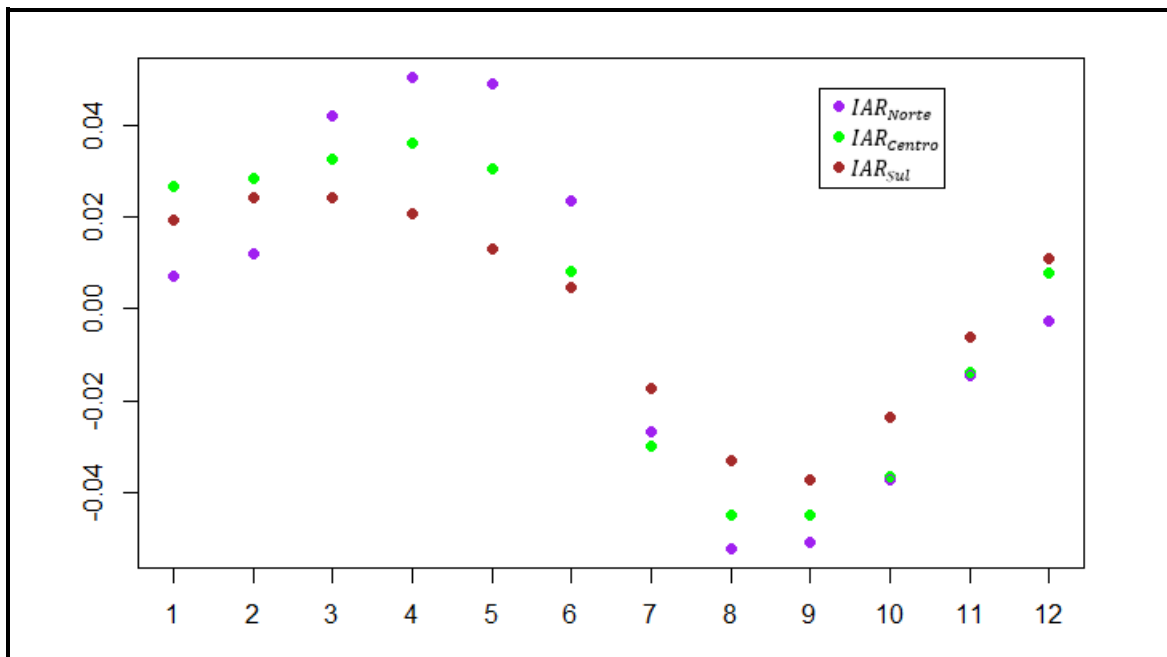


Figura 3.16: Gráficos dos índices sazonais.<sup>24</sup>

Os valores de  $s_i$  positivos indicam um  $IAR$ , para o respectivo mês, mais próximo de zero e consequentemente uma menor intensidade do afloramento costeiro. Em contrapartida, valores negativos mais distantes de zero implicam valores do  $IAR$  mais baixos e um afloramento costeiro mais intenso. Verifica-se um padrão semelhante nas três regiões, sendo o afloramento costeiro mais intenso no verão, começa a diminuir a partir do mês de setembro e atinge o seu valor mínimo no mês de março para a região Sul e abril para as regiões Norte e Centro. Os meses em que o fenômeno é mais intenso são os seguintes: na região Norte o mês de agosto; na região Centro os meses de agosto e setembro; na região Sul o mês de setembro.

Pode-se ainda concluir que o padrão sazonal do  $IAR$  da região Norte é mais pronunciado, seguido da região Centro e menos intenso na região Sul em que os valores dos índices sazonais se mantêm mais próximos do valor médio.

A interpretação deste fenômeno é importante para áreas do conhecimento ligadas à oceanografia e biologia.

<sup>24</sup> Fonte: Autor.

### 3.4 Discussão

Neste capítulo foram estudadas as séries temporais relativas à temperatura à superfície do mar (SST) para as regiões Norte, Centro e Sul de Portugal. As séries temporais em estudo são constituídas por 384 observações mensais, de janeiro de 1982 a dezembro de 2013. Em cada região, são consideradas duas localizações geográficas (próxima e afastada da costa) com a finalidade de estudar e caracterizar a SST para cada uma das regiões e também, comparar a SST entre as regiões. Além de se estudar a evolução da SST, estudou-se o fenómeno do afloramento costeiro.

Em suma, as séries temporais de SST possuem um padrão sazonal bem definido, com uma frequência periódica de 12 meses. Para cada região, as localizações geográficas próximas da costa têm uma evolução da SST menos pronunciada nos meses mais quentes (SST mais baixas) do que as distantes da costa, para a mesma latitude. Estes resultados evidenciaram a presença do fenómeno do afloramento costeiro e justificam o seu estudo neste trabalho. Relativamente ao fenómeno do afloramento costeiro, foi proposta uma nova expressão para o seu cálculo que designamos por Índice de Afloramento Relativo (*IAR*). Foram ajustados vários modelos para o *IAR* de cada região. A medida de erro MAE foi utilizada para comparar e determinar qual o modelo mais adequado para representar os dados em estudo, para cada região. Entre os modelos estudados, o modelo de decomposição clássica do tipo aditivo foi o selecionado. A partir dos resultados obtidos podemos concluir que estes possuem um padrão sazonal (menos evidente que o das SST) com periodicidade de 12 meses, sendo mais pronunciado na região Norte e menos evidente na região Sul. Em particular, o afloramento costeiro é mais frequente a partir da primavera aumentando de intensidade até atingir o seu valor máximo nos meses de agosto e setembro nas três regiões.



**Capítulo IV - Conclusão e Trabalhos Futuros**



### 4. Conclusão e trabalhos futuros

A presente dissertação teve como finalidade analisar a evolução da SST e do fenómeno do afloramento costeiro em três regiões de Portugal, para um período de 32 anos utilizando modelos de séries temporais. Os modelos selecionados para os *IAR* são diferentes na sua abordagem, ou seja, com e sem método analítico para a estimação da tendência e atribuem pesos diferentes às várias observações da série. A análise da série temporal da SST envolveu duas localizações geográficas para cada uma das regiões, num total de seis localizações geográficas. Foi possível identificar que as localizações mais próximas da costa têm valores mais baixos da SST em comparação aos das respetivas localizações mais distantes, em particular nos meses mais quentes do ano. Este resultado vai de encontro a estudos já publicados, que justificam esta ocorrência (Álvarez *et al.*, 2011; Fiúza *et al.*, 1982; Goela *et al.*, 2016). Na costa portuguesa os ventos favoráveis ao afloramento costeiro são mais intensos nos meses mais quentes do ano. Os ventos vindos do Norte deslocam-se na direção do equador paralelamente à costa criando uma corrente superficial na direção do oceano (à direita da direção do vento) gerando assim um fenómeno do afloramento costeiro responsável por valores da SST mais baixas em relação às localizações distantes da costa com mesma latitude (Fiúza *et al.*, 1982). Este fenómeno traz à superfície do mar águas mais frias e ricas em nutrientes que são de extrema importância ao desenvolvimento de diversas espécies marinhas e do seu ciclo reprodutivo. As águas frias ricas em nutrientes que chegam à superfície são essenciais ao desenvolvimento do fitoplâncton, e por isso locais onde o afloramento costeiro ocorre são muito produtivos, permitindo uma população muito desenvolvida de plâncton e de espécies que dele se alimentam (como por exemplo a sardinha) (Fiúza *et al.*, 1982). Quantificar este fenómeno, o seu padrão e a sua evolução no tempo, em conjunto com o conhecimento das espécies que dele dependem, é de extrema importância para a atividade económica da pesca e para o conhecimento e preservação dos ecossistemas da nossa costa.

A análise temporal da SST e do fenómeno do afloramento costeiro são dois tópicos de estudo muito importantes no enquadramento das alterações climáticas e na necessidade de adaptação das populações a estas, quer em termos ambientais, sociais ou económicos. Este trabalho pode ser um ponto de partida que visa servir de inspiração para estudos futuros. Considerar os dados atualizados (mais recentes) poderia ajudar a encontrar um modelo

## Conclusão e Trabalhos Futuros

---

também ele atualizado e aprofundar a análise. Uma possibilidade é fazer uma análise de tendência para saber se existe algum movimento persistente na série, complementando com testes de tendência. Uma outra possibilidade para estudos futuros seria a construção de modelos preditivos que permitissem prever valores futuros do *IAR* ou da *SST* relativos às localizações geográficas. Seria ainda interessante fazer uma análise da correlação espacial entre regiões e, caso fosse significativa, fazer uma análise mais completa da distribuição geográfica do *IAR*.

**Anexos**



## Anexos

### Anexo I – Dados Iniciais

Observações da SST em °C para a localização geográfica N1:												
Ano	Jan	Fev	Mar	Abr	Mai	Jun	Jul	Ago	Set	Out	Nov	Dez
1982	14,026	13,656	13,711	15,089	15,116	16,750	17,820	16,296	16,364	16,593	15,551	13,904
1983	12,480	12,637	13,743	13,402	14,636	16,409	17,700	17,252	17,288	17,221	16,964	15,702
1984	14,172	13,134	12,526	15,233	14,424	16,246	17,545	16,651	15,843	15,352	15,202	14,644
1985	13,159	13,151	13,317	14,552	14,852	17,220	17,628	17,049	17,416	17,653	15,992	14,624
1986	13,221	12,582	12,925	12,959	14,374	14,839	16,573	17,558	18,199	17,508	14,726	14,085
1987	12,971	12,816	13,677	14,657	15,910	16,098	17,147	17,716	17,889	17,771	16,116	14,883
1988	14,056	13,335	13,874	13,877	15,224	16,902	17,985	17,020	16,206	14,965	16,574	15,464
1989	14,333	13,906	14,051	13,698	16,397	17,520	18,001	19,166	16,956	17,113	16,484	15,777
1990	14,608	14,722	14,747	14,174	16,873	17,419	18,416	18,412	17,938	17,094	15,956	13,696
1991	13,091	12,454	13,080	13,786	14,806	15,436	16,411	17,596	19,637	15,379	14,589	13,763
1992	13,069	13,105	13,391	13,339	15,089	16,836	16,880	17,203	17,003	15,057	14,582	14,113
1993	13,575	13,333	13,800	14,384	15,825	17,724	16,721	16,817	17,054	15,513	14,973	14,024
1994	13,231	12,556	13,526	13,163	14,609	17,481	17,016	17,635	15,914	16,142	16,080	15,492
1995	14,303	13,841	13,951	14,728	16,032	16,829	19,200	17,659	17,109	16,388	16,915	15,566
1996	14,932	13,822	13,871	15,527	15,916	16,772	17,201	17,667	16,824	15,573	14,702	13,512
1997	13,065	14,021	15,504	16,730	17,122	18,082	17,712	18,523	18,373	18,250	17,212	15,662
1998	14,599	14,457	15,223	15,158	15,674	16,681	16,452	17,263	17,479	16,040	15,446	13,822
1999	13,263	12,981	13,475	13,889	15,880	16,982	17,614	18,616	18,530	17,379	15,781	14,177
2000	13,009	13,256	13,916	13,793	15,959	16,901	17,633	17,772	17,696	15,817	14,368	14,838
2001	14,264	13,610	14,145	14,580	15,082	16,894	17,145	18,378	17,908	17,861	15,456	13,757
2002	13,958	13,946	14,074	14,237	14,826	15,590	15,169	16,343	17,547	18,442	17,020	15,217
2003	13,928	13,006	14,223	14,825	15,118	17,259	17,983	19,007	18,631	17,166	15,609	14,088
2004	13,515	13,474	13,628	13,960	15,512	18,199	17,611	19,847	17,745	16,507	15,179	14,022
2005	13,068	12,361	12,865	14,240	16,450	17,417	17,784	17,490	17,084	16,122	16,362	14,507
2006	13,570	12,903	13,164	15,206	16,676	18,081	18,797	17,802	17,495	18,768	18,448	15,866
2007	14,954	14,168	13,945	14,349	14,656	17,041	16,919	15,696	17,154	16,435	15,333	14,519
2008	14,163	14,183	14,163	15,015	16,320	16,910	17,858	17,560	18,028	15,820	13,507	12,413
2009	12,201	12,155	13,685	13,751	15,523	18,158	18,043	17,268	16,693	17,943	16,485	15,207
2010	14,044	13,316	13,742	16,090	15,899	16,568	17,061	16,482	18,877	17,383	15,719	14,255
2011	14,064	13,795	14,105	15,843	17,928	16,610	15,992	17,315	17,726	16,801	16,307	15,217
2012	14,180	13,154	13,704	13,745	16,412	17,935	16,710	18,490	18,406	17,975	16,267	14,573
2013	14,033	12,978	12,990	14,159	13,952	15,052	18,628	17,702	15,993	18,122	16,209	13,752

## Anexos

Observações da SST em °C para a localização geográfica N2:

Ano	Jan	Fev	Mar	Abr	Mai	Jun	Jul	Ago	Set	Out	Nov	Dez
1982	14,368	14,012	14,130	14,916	15,535	17,183	18,455	18,830	18,288	17,924	16,669	14,949
1983	13,665	13,256	13,843	13,599	14,299	16,355	18,210	18,668	18,805	19,145	17,705	15,944
1984	14,512	13,716	13,297	15,157	15,129	15,978	18,371	18,866	19,101	17,653	16,283	14,731
1985	13,605	13,412	13,409	13,985	15,268	17,153	18,986	19,486	19,695	18,867	17,250	15,592
1986	14,159	13,237	13,190	13,170	14,599	15,675	17,738	19,469	19,479	18,781	16,966	15,277
1987	14,142	13,658	14,053	14,383	15,744	17,239	18,291	19,140	19,863	18,108	16,916	15,357
1988	14,394	13,654	13,916	14,297	15,472	17,223	19,072	19,952	18,842	17,051	17,161	15,572
1989	14,703	14,371	13,919	13,978	16,611	18,601	19,857	20,717	18,788	18,430	16,805	15,383
1990	14,578	14,016	14,115	14,014	16,455	17,993	19,774	20,318	19,600	18,146	16,468	14,563
1991	13,879	13,223	13,147	13,548	14,752	16,489	17,841	19,237	20,683	18,020	16,130	15,310
1992	14,520	14,250	13,890	14,010	15,476	16,733	18,428	19,847	19,969	17,550	16,605	15,732
1993	14,471	14,619	14,169	14,623	15,622	17,174	18,662	18,778	18,342	17,015	16,007	15,177
1994	13,875	13,302	13,575	13,801	14,863	17,636	19,079	19,237	18,461	18,282	17,337	16,282
1995	14,762	14,255	14,188	15,352	16,281	17,312	19,411	19,734	18,345	18,371	17,992	16,309
1996	14,987	14,206	13,930	15,410	16,084	18,127	18,754	19,474	18,784	17,835	16,690	15,216
1997	14,249	14,324	14,962	16,046	16,789	17,818	18,818	19,877	20,234	19,799	17,911	15,865
1998	15,220	14,951	15,398	15,014	15,576	17,244	18,528	19,310	19,289	18,280	17,383	15,801
1999	14,408	14,176	14,094	14,345	15,640	17,113	19,079	20,621	20,356	18,503	16,585	15,207
2000	13,881	13,665	14,232	14,058	16,104	17,554	18,696	19,440	19,844	18,131	16,061	15,261
2001	14,148	13,516	13,614	14,430	15,290	17,709	19,015	20,254	19,256	18,112	15,905	14,817
2002	14,538	14,252	14,032	14,497	15,162	16,519	17,052	18,543	19,180	18,778	17,278	15,713
2003	14,349	13,741	14,014	14,725	15,716	18,268	19,279	20,497	20,330	18,637	16,702	15,146
2004	14,342	14,060	13,912	13,907	15,512	18,948	19,592	21,114	20,154	18,075	16,374	15,129
2005	14,115	13,304	13,242	14,172	16,135	17,851	19,129	19,242	18,964	18,108	17,001	15,591
2006	14,370	13,628	13,614	14,533	16,496	18,118	19,837	19,695	20,563	19,633	18,449	16,027
2007	15,114	14,465	14,234	14,607	15,491	17,414	19,483	19,698	19,867	18,446	17,117	15,510
2008	14,938	14,664	14,460	14,902	16,136	18,143	19,338	20,365	20,511	19,135	16,415	14,517
2009	13,526	13,273	13,712	14,076	15,269	17,737	19,366	19,824	19,102	18,626	17,157	15,402
2010	14,139	13,667	13,768	15,280	16,302	17,851	19,027	18,949	20,103	18,960	17,014	14,877
2011	14,343	14,059	13,751	15,317	16,703	18,079	18,555	19,277	20,095	19,588	17,193	16,142
2012	15,065	14,012	14,292	14,348	15,806	18,023	19,234	20,600	20,607	19,087	17,241	15,164
2013	14,248	13,625	13,515	14,225	14,602	15,824	18,504	19,334	18,947	18,877	17,324	15,267

## Anexos

Observações da SST em °C para a localização geográfica C1:												
Ano	Jan	Fev	Mar	Abr	Mai	Jun	Jul	Ago	Set	Out	Nov	Dez
1982	15,135	14,989	14,578	15,837	16,047	16,972	17,799	17,262	18,130	17,182	16,092	14,734
1983	14,121	14,016	14,867	14,557	15,479	16,959	18,408	18,080	17,796	18,095	18,354	17,136
1984	15,453	14,409	14,040	15,749	15,669	16,800	17,911	17,754	17,624	16,716	16,350	15,710
1985	14,603	14,636	14,476	15,202	15,499	17,465	18,219	17,760	18,928	18,744	16,879	15,779
1986	14,714	13,736	13,971	13,978	15,223	16,358	17,647	18,362	19,254	18,452	16,039	15,260
1987	14,454	14,283	15,018	15,465	16,851	17,224	17,970	19,321	19,603	18,212	16,864	16,046
1988	15,396	14,603	14,863	14,966	15,927	17,135	17,859	17,938	17,783	17,497	18,434	16,765
1989	15,772	15,311	14,924	14,444	16,647	17,765	18,991	20,474	18,769	19,493	17,839	16,922
1990	15,731	15,472	15,672	15,536	17,652	17,954	18,984	19,525	19,491	18,400	16,673	15,193
1991	14,879	14,146	14,219	14,890	15,774	16,653	17,743	19,401	20,032	17,012	16,005	15,399
1992	15,206	14,892	14,816	14,572	16,138	17,085	17,445	18,259	17,871	16,052	15,749	15,317
1993	15,434	15,058	14,952	15,256	16,337	17,978	17,963	18,422	18,213	16,386	16,172	15,403
1994	14,220	13,792	14,409	14,348	15,566	18,200	17,480	18,068	17,033	17,127	16,995	17,065
1995	15,355	15,201	14,759	15,953	16,710	17,407	19,063	19,246	18,359	18,496	18,818	17,025
1996	16,142	15,138	15,050	16,118	16,837	18,111	18,551	18,436	18,373	16,964	15,999	14,813
1997	14,983	15,473	16,466	17,614	17,613	18,523	18,676	19,218	19,174	18,986	17,482	16,222
1998	15,689	15,585	16,192	15,760	16,232	17,030	17,476	18,067	18,276	17,237	16,774	15,386
1999	14,844	14,343	14,570	15,187	16,536	17,331	17,809	19,000	18,836	17,963	16,602	15,597
2000	14,890	14,859	15,424	15,121	17,012	17,979	18,843	18,791	18,483	17,138	15,858	15,765
2001	14,988	14,709	14,852	15,263	15,576	17,240	17,490	18,718	19,314	18,563	16,672	15,704
2002	16,073	15,602	15,440	15,547	15,793	16,245	16,458	17,360	18,733	19,243	17,352	15,791
2003	15,076	14,213	14,655	15,699	16,147	17,953	18,335	20,085	19,616	18,662	16,814	15,428
2004	14,995	15,060	14,822	15,020	15,960	18,076	18,072	19,938	18,977	17,697	16,228	15,674
2005	14,783	14,124	14,277	15,488	16,723	18,212	18,566	19,237	18,535	18,670	17,963	15,966
2006	15,085	14,434	14,446	15,947	16,905	19,103	19,456	19,202	19,362	19,343	19,322	16,688
2007	15,606	15,039	14,925	15,104	15,522	17,419	17,572	17,377	18,749	18,319	17,255	16,300
2008	15,811	15,744	15,613	15,938	16,558	17,494	18,247	17,712	18,444	17,324	15,416	14,329
2009	14,151	13,881	14,740	14,663	16,043	18,541	18,575	17,741	17,973	19,521	18,272	16,392
2010	15,479	15,058	15,509	16,793	16,991	17,228	17,939	18,382	18,840	18,184	16,310	16,108
2011	15,870	15,399	15,308	16,911	18,684	18,446	17,190	18,204	18,635	18,800	17,130	16,231
2012	15,162	14,243	14,825	15,026	16,666	17,858	17,402	19,333	19,925	18,812	17,423	15,894
2013	15,366	14,258	14,432	15,394	15,171	16,132	18,778	18,393	18,604	19,464	17,343	15,578

## Anexos

Observações da SST em °C para a localização geográfica C2:

Ano	Jan	Fev	Mar	Abr	Mai	Jun	Jul	Ago	Set	Out	Nov	Dez
1982	15,677	15,164	15,046	15,960	16,706	18,228	19,537	19,890	19,724	19,090	17,812	16,202
1983	15,072	14,665	15,223	14,989	15,689	17,436	19,263	20,019	20,036	20,354	18,737	17,461
1984	15,830	14,971	14,545	15,872	16,163	17,321	19,395	19,795	20,051	19,003	17,466	16,148
1985	15,258	15,028	14,896	15,327	16,194	18,198	19,898	20,092	20,817	20,928	18,649	16,621
1986	15,182	14,323	14,586	14,440	15,592	16,931	18,868	20,461	20,912	20,140	17,703	16,437
1987	15,462	14,865	15,301	15,807	17,185	18,317	19,360	20,741	21,232	20,243	18,065	16,673
1988	15,885	15,059	15,220	15,661	16,800	17,935	19,672	20,367	19,882	18,803	18,721	17,389
1989	16,229	15,744	15,358	15,031	17,251	19,481	21,025	22,270	20,606	19,913	18,174	16,773
1990	15,936	15,441	15,629	15,421	17,488	19,106	20,500	21,337	21,201	19,716	17,939	15,975
1991	15,273	14,735	14,434	14,871	16,155	17,680	18,885	20,619	21,976	19,317	18,123	16,471
1992	15,868	15,310	15,070	15,152	16,324	17,539	18,778	19,943	20,548	18,687	17,416	16,507
1993	15,889	15,470	15,205	15,888	16,530	18,393	19,504	20,118	20,087	18,437	16,910	16,363
1994	14,622	14,076	14,692	15,113	16,021	18,666	19,394	20,176	19,723	19,098	18,310	17,744
1995	15,962	15,713	15,336	16,322	17,290	18,244	20,049	20,732	20,377	19,625	18,947	17,289
1996	16,137	15,522	15,397	16,371	16,906	18,669	19,842	20,559	20,187	19,284	17,790	16,086
1997	15,302	15,648	16,451	17,406	17,819	18,957	19,865	20,871	21,125	20,746	19,351	17,386
1998	16,305	15,886	16,447	16,237	16,396	17,919	19,527	20,228	20,288	19,349	18,686	17,044
1999	15,524	14,894	15,012	15,720	17,118	18,107	19,789	21,150	21,133	19,990	18,126	16,523
2000	15,251	15,382	15,740	15,365	16,653	18,746	20,015	20,537	20,641	19,324	17,152	16,260
2001	15,564	15,070	15,165	15,516	16,167	18,528	19,627	20,992	21,049	20,271	17,932	15,929
2002	15,608	15,332	15,651	15,924	16,505	17,715	18,325	19,573	20,204	20,150	18,770	16,953
2003	15,781	14,824	15,240	15,788	16,569	18,851	20,297	21,846	22,071	20,551	17,850	16,480
2004	15,610	15,566	15,632	15,362	16,584	19,341	20,100	22,053	21,317	19,977	17,921	16,464
2005	15,244	14,696	14,461	15,511	17,370	18,868	20,346	21,074	21,091	20,129	18,523	16,576
2006	15,281	14,738	15,007	16,041	17,602	19,470	20,788	21,591	21,395	20,793	19,707	17,480
2007	16,291	15,710	15,842	15,711	16,545	18,314	19,767	20,164	20,445	19,764	18,358	17,026
2008	16,041	15,840	15,946	16,250	17,273	18,947	20,510	20,951	20,981	19,882	17,408	15,242
2009	14,760	14,208	14,781	15,238	16,519	19,315	20,640	20,702	20,609	21,257	19,748	17,455
2010	15,867	15,373	15,454	16,465	17,281	18,419	19,847	20,403	21,161	19,839	17,989	16,428
2011	15,939	15,465	15,518	16,853	18,361	19,343	19,352	20,131	20,946	20,625	18,249	17,113
2012	15,950	14,858	15,266	15,612	17,070	19,110	19,861	21,571	21,987	20,636	18,742	16,361
2013	15,533	15,116	14,969	15,598	16,011	17,000	19,841	20,468	20,584	20,200	18,873	16,608

## Anexos

Observações da SST em °C para a localização geográfica S1:												
Ano	Jan	Fev	Mar	Abr	Mai	Jun	Jul	Ago	Set	Out	Nov	Dez
1982	16,286	15,826	15,491	16,462	17,135	17,677	18,806	18,614	19,849	18,237	17,287	15,930
1983	15,573	15,245	15,964	15,623	16,462	17,776	19,256	19,190	19,260	19,930	19,553	17,881
1984	16,263	15,259	14,970	16,302	16,690	17,691	19,010	19,332	19,595	18,260	17,526	16,769
1985	15,653	15,368	15,282	15,885	16,269	18,492	19,397	18,958	19,807	20,381	18,853	17,008
1986	15,790	14,933	14,787	14,869	16,185	17,915	19,045	19,190	20,094	19,181	17,370	16,864
1987	15,815	15,299	15,971	16,255	17,640	18,407	19,621	20,555	21,388	19,906	18,339	17,137
1988	16,275	15,496	15,758	15,905	16,535	18,118	19,079	19,321	19,793	19,560	19,024	17,174
1989	16,296	15,779	15,490	15,224	17,429	18,658	20,618	21,500	20,218	20,965	18,874	17,257
1990	16,436	16,019	16,178	16,148	17,800	18,793	20,002	21,052	20,787	19,621	17,955	16,900
1991	16,161	15,225	15,062	15,745	16,240	17,846	19,464	21,038	21,789	19,048	17,852	16,669
1992	16,192	15,748	15,561	15,384	17,164	18,003	18,872	20,098	19,104	17,544	17,170	16,734
1993	16,449	15,823	15,648	15,878	16,929	18,890	19,557	19,862	19,819	18,145	17,372	16,650
1994	15,328	14,950	15,493	15,348	16,344	18,825	18,906	19,273	18,535	19,003	18,520	18,116
1995	16,447	16,193	15,815	16,865	17,567	18,290	19,878	20,053	19,647	19,910	19,816	17,836
1996	17,051	15,950	15,727	16,763	17,479	19,251	20,196	19,608	19,499	18,640	17,951	16,335
1997	15,929	16,107	16,808	17,603	18,101	19,417	19,560	20,515	20,842	20,489	19,033	17,195
1998	16,610	16,492	16,575	16,550	16,864	17,750	18,724	19,794	19,225	18,528	18,037	17,173
1999	16,253	15,569	15,745	16,258	17,333	18,394	19,344	20,187	20,204	19,341	17,969	16,698
2000	15,773	15,598	16,108	16,007	17,549	18,783	19,628	20,289	20,224	18,899	17,253	16,830
2001	15,966	15,463	15,771	16,108	16,378	18,180	18,347	19,539	20,180	19,605	18,459	17,098
2002	16,515	16,217	16,103	16,232	16,800	17,404	17,892	18,855	19,588	20,087	18,384	17,034
2003	16,116	15,183	15,486	16,301	16,778	18,975	19,471	20,985	21,015	20,074	17,963	17,042
2004	16,386	16,155	15,834	16,049	16,620	18,953	19,070	20,857	20,329	19,145	18,176	17,482
2005	16,086	15,025	15,160	16,040	17,653	19,522	20,000	20,950	20,164	20,377	18,877	16,989
2006	16,067	15,376	15,284	16,626	17,586	20,213	20,638	21,086	20,672	20,615	20,140	17,690
2007	16,807	16,085	15,688	15,913	16,707	18,308	18,628	19,177	20,342	20,295	19,549	17,773
2008	16,890	16,615	16,308	16,589	17,353	18,654	19,254	18,713	19,895	19,274	17,356	16,058
2009	15,429	15,030	15,581	15,637	17,056	19,746	19,537	19,700	20,153	20,674	18,661	17,395
2010	16,510	16,184	16,151	17,031	17,778	18,487	19,329	20,403	20,777	19,633	17,756	17,176
2011	16,617	16,158	16,160	17,375	19,409	19,720	18,280	19,257	19,763	20,821	18,593	17,429
2012	16,101	15,215	15,429	15,731	17,406	18,409	18,477	19,818	20,863	19,837	18,288	16,990
2013	16,240	15,462	15,554	16,229	16,678	17,319	20,101	19,930	21,176	20,946	18,474	17,522

## Anexos

Observações da SST em °C para a localização geográfica S2:

Ano	Jan	Fev	Mar	Abr	Mai	Jun	Jul	Ago	Set	Out	Nov	Dez
1982	16,178	15,666	15,513	16,416	17,129	18,532	19,926	20,572	20,810	19,986	18,210	16,715
1983	15,789	15,451	16,003	15,936	16,707	18,222	20,247	20,616	21,055	21,190	19,818	18,235
1984	16,480	15,412	15,016	16,507	16,982	17,735	19,903	20,832	21,669	20,334	18,295	17,419
1985	16,274	15,621	15,505	15,936	16,743	18,792	20,197	20,558	21,511	21,402	19,741	17,647
1986	16,010	14,980	15,061	15,131	16,418	17,748	19,427	20,727	21,470	20,773	18,447	17,302
1987	16,162	15,482	16,084	16,496	18,026	19,228	20,250	21,434	22,138	20,963	18,824	17,411
1988	16,385	15,653	15,815	16,668	17,344	18,748	19,902	20,406	20,895	20,231	19,324	17,800
1989	16,796	16,080	15,829	15,477	17,703	19,820	21,302	22,666	21,369	21,213	18,823	17,161
1990	16,296	15,928	16,149	16,012	18,059	19,702	20,899	21,967	21,891	20,782	18,842	16,910
1991	16,086	15,400	15,239	15,562	16,568	18,103	19,267	21,346	22,779	20,558	18,748	17,132
1992	16,289	15,720	15,422	15,660	17,100	18,400	19,389	20,755	21,249	19,407	18,424	17,314
1993	16,354	15,749	15,721	16,202	17,170	19,232	20,075	20,353	21,057	19,063	17,467	16,598
1994	15,289	14,977	15,417	15,758	16,664	19,595	19,962	20,963	20,643	20,041	19,031	18,438
1995	16,927	16,562	16,225	16,959	18,048	18,763	20,664	21,418	21,639	20,595	19,891	17,841
1996	16,829	16,067	15,695	16,632	17,585	19,097	20,552	21,124	21,142	20,528	18,845	16,995
1997	15,988	16,289	16,929	17,563	18,158	19,433	20,045	21,088	21,732	21,148	19,410	17,808
1998	16,839	16,636	16,710	16,886	17,049	18,584	19,725	20,594	20,904	19,618	19,040	17,306
1999	16,241	15,598	15,831	16,621	18,250	18,863	20,300	21,860	21,950	20,624	18,868	17,298
2000	16,263	15,853	16,299	15,938	17,261	19,180	20,624	21,689	21,317	20,418	18,746	17,049
2001	16,513	15,680	15,669	16,404	17,262	19,055	20,002	21,108	21,954	20,757	18,738	17,591
2002	16,657	16,300	16,413	16,477	17,393	18,423	19,077	20,378	21,085	20,554	19,022	17,082
2003	16,523	15,620	16,107	16,612	17,283	19,455	20,869	22,509	22,368	21,365	19,222	17,228
2004	16,598	16,450	16,076	16,009	16,816	19,718	20,584	22,842	21,863	20,598	18,656	17,065
2005	16,256	14,934	15,043	16,225	17,932	19,477	20,917	22,060	21,721	20,737	19,382	17,574
2006	16,081	15,303	15,358	16,533	17,979	19,897	21,000	22,360	21,795	21,520	20,534	18,208
2007	16,910	16,284	15,867	16,297	17,198	18,997	19,917	20,509	21,147	20,933	19,763	18,324
2008	17,067	16,578	16,540	16,659	17,849	19,379	20,853	21,196	21,659	20,241	17,877	16,175
2009	15,480	14,831	15,206	15,607	17,088	20,064	21,324	21,127	21,176	21,504	19,858	18,178
2010	16,601	15,842	15,795	16,804	17,686	18,968	20,361	21,381	22,038	21,157	19,204	17,294
2011	16,617	16,005	16,090	17,329	18,837	20,058	20,234	21,243	21,845	21,013	19,209	18,041
2012	16,883	15,328	15,868	16,191	17,751	19,616	20,313	22,134	22,735	21,421	19,290	16,918
2013	16,321	15,814	15,655	16,280	17,119	18,000	20,637	21,340	21,668	20,974	19,420	17,697

**Anexo II – Comandos do **

```
#####
#Pacotes do R utilizados:
#####
library("astsa", lib.loc="~/R/win-library/3.2")
library("forecast", lib.loc="~/R/win-library/3.3")
library("tseries", lib.loc="~/R/win-library/3.3")
#####
#Dados das séries temporais
#####
Dados30 <- read.csv (file="C:/.../Dados30.csv", header = TRUE, sep=";")
#####
#Transformação dos dados nas seis séries temporais em estudo
#####
N1 = ts (Dados30$Norte1, start = c(1982), frequency=12)
#Gráfico
plot(N1, xlab = "", ylab = "°C", main = "Região Norte (a)", col = "purple", ylim = c(12,24),
yaxt = "n", xaxt = "n", lwd = 2)
#####
#Cálculo e gráficos da ACF e PACF até o desfasamento 48
#####
acfN1 = acf (N1, plot = F, lag.max = 48); pacfN1 = pacf (N1, plot = F, lag.max = 48)
acfN1$lag = acfN1$lag*12; pacfN1$lag = pacfN1$lag*12
plot (acfN1, main = "N1", xlab = "Desfasamento")
plot (pacfN1, main = "N1", xlab = "Desfasamento")
#####
#Ciclo anual
#####
N1m = matrix(N1, ncol=12, byrow = TRUE)
N1mo = apply (N1m, 2, mean); summary (N1mo)
plot (N1mo, type = "l", lwd = 2, ylab = "°C")
#####
```

```
#Diagrama de extremos e quartis por mês
boxplot (N1~cycle(N1), main = "Região Norte (a)", ylab = "°C", col = "purple")
#####
#Séries temporais relativas
#####
N3 = (N1-N2) / N2
#####
#Modelagem
#####
#Decomposição Clássica dos tipos Aditivo e Multiplicativo
#####
dec.N3A=decompose(N3); dec.N3M=decompose(N3,type="multiplicative")
#####
#Cálculo da série temporal estimada
fit.dec.N3A<-dec.N3A$seasonal+dec.N3A$trend
fit.dec.N3M<-dec.N3M$seasonal*dec.N3M$trend
#####
#Medidas de erro
accuracy(fit.dec.N3A,C3); accuracy(fit.dec.N3M,C3)
#####
#Índices sazonais
dec.N3A$figure; dec.N3M$figure; plot(dec.N3A)
#####
#Teste de Ljung-Box
Box.test(dec.N3A$random,lag=24,type='Ljung-Box')
#####
#Transformação para que as séries temporais não tenham valores negativos
```

```
N4=N3+(1-min(N3))

#####

#Alisamento exponencial de Holt - Winters

#####

hw.N3A=hw(N3, seasonal = "additive"); accuracy(hw.N3A)

hw.N3M=hw(N4,seasonal="multiplicative"); accuracy((hw.N3M$fitted+min(N3)-1),N3)

#####

#Constantes de alisamento, nível médio, tendência, índices sazonais e índices AIC e BIC

summary(hw.N3M)

#####

#Gráfico das séries temporais originais e respectivas séries alisadas

plot(N3, lwd = 3, col = "purple"); lines(hw.N3A$fitted, lwd=2)

#####

#Teste de Ljung-Box

Box.test (hw.N3A$random, lag=24, fitdf=length(hw.N3M$model$par), type='Ljung-Box')

#####

#Modelo ARIMA

#####

#Estimação da tendência linear

t=c(1:length(N3)); Tt.N3<-lm(N3~t)$coef

#Retirada da tendência linear

N3sT<-N3-(Tt.N3[1]+Tt.N3[2]*t)

#####
```

```
#Índices sazonais

N3m<- matrix(N3sT, ncol=12,byrow=TRUE)

N3.si<-apply(N3m,2,mean) ## com média ~ 0

#Retirada dos índices sazonais

for (i in 1:12) N3m[,i]<-N3m[,i]-N3.si[i]

N3.TS<-ts(as.vector(t(N3m)),start=c(1982,1), frequency=12)

for (i in 1:12) C3m[,i]<-C3m[,i]-C3.si[i]

#####

#Teste aumentado de Dickey-Fuller:

adf.test (N3.TS, k=0)

#####

#Gráficos dos modelos ajustados

plot(N3, col = "purple", lwd=3)

lines(N3 - N3.TS, lwd=2)

#####

#Geração de modelos ARIMA com diferentes ordens e respectivos AIC e BIC:

N100=Arima(N3.TS,order=c(1,0,0), seasonal=list(order=(c(0,0,0)), period=12))

N100$aic; N100$bic

#####

#Modelos escolhidos são AR(1)

N3.ar<-ar(N3.TS) ; N3.ar; sqrt(N3.ar$var.pred)

#Resíduos

r.N3.ar=N3.ar$resid
```

```
#####  
#Modelo ajustado  
fit.N3=fitted(N3.ar)+N3.si+Tt.N3[1]+Tt.N3[2]*t #Ou  
fit.N3<-N3-N3.ar$resid  
#####  
#Medidas de comparação  
accuracy(fit.N3,N3)  
#####  
#Ljung-Box  
k = 24 ;lb= matrix(0,23,2)  
for (i in 2:k) {lbN3 = Box.test(r.N3.ar,fitdf=1,lag=i,type='Ljung-Box')  
  lb[i-1,1:2] = c(i,round(lbN3$p.value,5)) }; print(lb)  
Box.test(r.N3.ar,lag=10,fitdf=1,type="Ljung-Box")  
#####  
#Gráficos dos modelos escolhidos  
#####  
attach(mtcars); layout(matrix(c(1,1,1,1,2,2,3,3,4,4,5,5), 2, 6, byrow = TRUE))  
#Gráfico dos resíduos  
plot(r.N3.ar,main="Resíduos")  
#####  
#Gráfico do papel de probabilidade  
qqnorm(r.N3.ar ); qqline(r.N3.ar)  
#####
```



### Bibliografia

ALPUIM, T. – *Séries Temporais*. Lisboa, Portugal: Associação dos Estudantes da Faculdade de Ciências de Lisboa, 2003.

ÁLVAREZ, Ines; GÓMEZ-GESTEIRA, Moncho; deCASTRO, Maite; LORENZO, Maria Nieves; CRESPO, Alejandro J. C.; DIAS, João Miguel. *Comparative analysis of upwelling influence between the western and northern coast of the Iberian Peninsula*. Continental Shelf Research. Vol. 31, Issue 5, 2011. Pages 388-399 ISSN: 0278-4343.

BINDOFF, N. L. *et al.* - *Observations: Oceanic Climate Change and Sea Level*. In: *Climate Change 2007: The Physical Science Basis*. [Em linha]. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change [Solomon, S. *et al.*]. Cambridge, Reino Unido and New York: Cambridge University Press, NY, USA, 2007.

Disponível em [https://www.ipcc.ch/publications\\_and\\_data/ar4/wg1/en/ch5.html](https://www.ipcc.ch/publications_and_data/ar4/wg1/en/ch5.html).

BOLLMANN, Moritz - *World Ocean Review 2010: living with the oceans*. Hamburg, Alemanha: Maribus, 2010. 232 p., ISBN: 978-3-86648-012-4.

Disponível em [http://worldoceanreview.com/wp-content/downloads/wor1/WOR1\\_english.pdf](http://worldoceanreview.com/wp-content/downloads/wor1/WOR1_english.pdf).

BOX, George E. P.; JENKINS, Gwilym M.; REISEL, Gregory C. - *Time Series Analysis, Forecasting and Control* Hoboken, New Jersey, U.S.A.: John Wiley & Sons, INC., 4th edition, 2008. 784 p., ISBN 0-13-060774-6.

CASTRO, Peter - *Marine Biology*. 4th edition. U.S.A.: McGraw-Hill Companies, 2003. 470 p., ISBN: 978-0-072-85290-5.

CHATFIELD, Christopher – *The Analysis of Time Series: An Introduction*. 6th edition. U.K.: Chapman and Hall, 2003. 352 p., ISBN: 978-1-58488-317-3.

COPERNICUS - *Global-Reanalysis-Phys-001-011, Product User Manual*. Copernicus Marine Environment Monitoring Service, Version 2.0., Issue 3.1, 2016. 42 p..

Ficheiro em PDF disponível em:

<http://cmems-resources.cls.fr/documents/PUM/CMEMS-GLO-PUM-001-025-011-017.pdf>.

## Bibliografia

---

COPERNICUS - *GLOBAL OCEAN PHYSICS REANALYSIS CGLORYS (1982-2013)*. [Em linha] CMEMS - Copernicus Marine Environment Monitoring Service. [Download da base de dados no dia 1 de junho 2016].

Disponível em [http://marine.copernicus.eu/services-portfolio/access-to-products/?option=com\\_csw&view=details&product\\_id=GLOBAL\\_REANALYSIS\\_PHYS\\_001\\_011](http://marine.copernicus.eu/services-portfolio/access-to-products/?option=com_csw&view=details&product_id=GLOBAL_REANALYSIS_PHYS_001_011).

CUSHMAN-ROISIN, Benoit; BECKERS, Jean-Marie - *Introduction to Geophysical Fluid Dynamics, Physical and Numerical Aspects*. Academic Press Inc, 2<sup>nd</sup> edition, 2011. 875p. ISBN 10:0120887592.

EHLERS, Ricardo S. - *Análise de Séries Temporais* [Em linha]. Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo – USP, 2009. [Consult. 15 Abril 2016]. Capítulos 1, 2, 3, 4, 5, 6 e comandos do R.

Disponível em <http://www.icmc.usp.br/ehlers/stemp/>.

FIÚZA, Armando F. de G.; MACEDO, M. E.; GUERREIRO, M. R. - *Climatological space and time variation of the portuguese coastal upwelling*. Oceanologica acta Vol. 5, Issue 1, 1982, pages 31-40.

GOELA, Priscila Costa; CORDEIRO, Clara; DANCHENKO, Sergei; ICELY, John; CRISTINA, Sónia; NEWTON, Alice - *Time series analysis of data for sea surface temperature and upwelling components from the southwest coast of Portugal*. Journal of Marine Systems. Vol. 163, November 2016, pages 12–22. ISSN:0924-7963.

GONZALEZ-NUEVO, Gonzalo; GAGO, Jesus; cabanas, Jose M. - *Upwelling index: a powerful tool for marine research in the NW Iberian upwelling system*. Journal of Operational Oceanography. Vol. 7, Issue 1, December 2014, pages 47- 57. ISSN: 1755-876X

HAMILTON, James D. – *Times Series Analysis*. Princeton University Press, Princeton, New Jersey, 1994. 799 p., ISBN 0-691-04289-6.

HARTMANN, D.L. *et al.* - *Observations: Atmosphere and Surface*. In: *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* [Stocker, T.F. *et al.*]. Cambridge, Reino Unido and New York: Cambridge University Press, NY, USA, 2013.

Disponível em <http://www.ipcc.ch/report/ar5/wg1/>.

## Bibliografia

---


HARRIS, Richard; SOLLIS, Robert - *Applied Time Series Modelling and Forecasting*. Chichester, U.K.: John Wiley & Sons Ltd, 2003. 302 p., ISBN: 0-470-84443-4.

HEATHROW AIRPORT - *Monthly traffic statistics up to June 2016, excluding Gatwick, Stansted, Edinburgh, Naples, Aberdeen, Glasgow and Southampton*. [Em linha]. Heathrow Airport REINO UNIDO, 2016. [Consult. 22 Julho 2016].

Disponível em <http://www.heathrow.com/company/investor-centre/results-and-performance/traffic-statistics>.

HYNDMAN, Rob J.; KOEHLER, Anne B. - *Another look at measures of forecast accuracy*. International Journal of Forecasting, Vol. 22, Issue 4, February 2006, pages 679-688.

HYNDMAN, Rob J.; ATHANASOPOULOS, George - *Forecasting: principles and practice*. OTexts, October 2013. 292 p., ISBN: 978-0987507105.

HYNDMAN, Rob J.; O'Hara-Wild, M.; Bergmeir, C.; Razbash, S.; Wang, E. - *Forecasting Functions for Time Series and Linear Models*.  package forecast Version 8.0. February 2017.

GONZALEZ-NUEVO, Gonzalo; GAGO, Jesus; CABANAS, Jose M. *Upwelling index: a powerful tool for marine research in the NW Iberian upwelling system*. Journal of Operational Oceanography, Vol. 7, Issue 1, 2014, pages 47 – 57, ISSN: 1755-876X.

Disponível em <http://dx.doi.org/10.1080/1755876X.2014.11020152>.

MAMAYEV, O.I. - *Temperature-Salinity Analysis of World Ocean Waters*. Elsevier Science Ltd, May 1975. 384 p., ISBN: 978-0444412515.

McLELLAN, Hugh J. - *Elements of Physical Oceanography*. Londres, Reino Unido: Pergamon Press, 1965. 160 p., ISBN: 0-08-011320-6.

MENTZER, John T.; KAHN, Kenneth B. *Forecasting Technique Familiarity, Satisfaction, Usage, and Application*. Journal of Forecasting, Vol. 14, 1995, pages 465-476.

MONTGOMERY, Douglas C.; JENNINGS, Cheryl L.; KULAHCI, Murat - *Introduction to Time Series Analysis and Forecasting*. John Wiley & Sons, Inc. 1st edition, 2008. 472 p., ISBN: 978-0-471-65397-4.

## Bibliografia

---


MORAN, Patrick A. M. - *The statistical analysis of the Canadian Lynx cycle*. Australian Journal of Zoology, Vol. 1, Issue 3, 1953, pages 291-298, ISSN: 0004-959X.


MORETTIN, Pedro A.; TOLOI Clelia M.C. - *Análise de Séries Temporais*. 2<sup>a</sup> ed. São Paulo SP, Brasil: Editora Edgard Blucher, 2006. 564 p., ISBN: 85-212-0389-6.

RUBAL, Marcos; VEIGA, Puri; CACABELOS, Eva; MOREIRA, Juan; SOUSA-PINTO, Isabel - *Increasing sea surface temperature and range shifts of intertidal gastropods along the Iberian Peninsula*. Journal of Sea Research, Volume 77, March 2013, Pages 1–10.

SHUMWAY, Robert H.; STOFFER, David S. - *Time Series Analysis and its Applications, with R Examples*. EZ, third edition, Free texts in Statistics, 2014. 201 p., ISBN: 978-1-4419-7865-3.

STENSETH, Nils C.; FALCK, Wilhelm; CHAN, Kung-Sik.; BJØRNSTAD, Ottar N.; O'DONOGHUE, Mark; TONG, Howell; BOONSTRA, Rudy; BOUTIN, Stan; KREBS, Charles J.; YOCCOZA, Nigel G. - *From patterns to processes: Phase and density dependencies in the Canadian lynx cycle*. Proceedings of the National Academy of Sciences of the United States of America, Volume 95, December 1998, pages 15430-15435.

STOFFER, D. - *astsa: Applied Statistical Time Series Analysis*. Proceedings of the National Academy of Sciences of the United States of America, *Series Analysis*.  package *astsa* version 1.7, 4th Edition, repository CRAN. December 2016.

TRAPLETTI, Adrian; HORNIK, Kurt; LEBARON, Blake. - *Time Series Analysis and Computational Finance*.  package *tseries* version 0.10-41, repository CRAN. May 2017.