



ESCOLA NAVAL

talant de biefaire



Miguel Gonçalves Picoto

Caracterização das variações do nível médio do mar em Portugal

Dissertação para obtenção do Grau de Mestre em
Ciências Militares Navais, na especialidade de Marinha



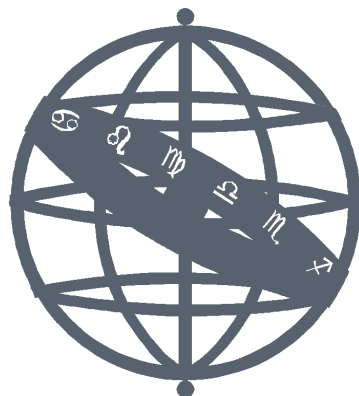
Alfeite

2022



ESCOLA NAVAL

talant de bi-faire



Miguel Gonçalves Picoto

*Caracterização das variações do nível médio do mar em
Portugal*

Dissertação para obtenção do Grau de Mestre em
Ciências Militares Navais, na especialidade de Marinha

Orientação de: Dora Cristina Nunes Paulino Carinhas

Co-orientação de: Carlos Miguel Reis Silva de Oliveira e Lemos

O Aluno Mestrando,

O Orientador,

Miguel Picoto

Dora Carinhas

Alfeite

2022

Agradecimentos

Deixo um especial agradecimento aos meus orientadores, Professora Dora Carinhas e Comandante Oliveira e Lemos. À professora Dora pela ajuda constante, paciência e disponibilidade sem as quais esta dissertação não seria possível. Ao Comandante Oliveira e Lemos, pela sua sabedoria, disponibilidade e pela constante vontade em alcançar a perfeição. Agradeço a ambos por terem abraçado este projeto.

Quero agradecer à Escola de Hidrografia e Oceanografia e à Divisão de Oceanografia, do Instituto Hidrográfico, pelo acolhimento e pela ajuda que deram na elaboração deste trabalho, em especial, pela disponibilidade dos registos maregráficos imprescindíveis à sua realização.

Expresso o meu sincero agradecimento a todos aqueles que constituem o meu curso de entrada na Escola Naval, CMG Monteiro de Barros, por todo o suporte nos bons e maus momentos. Foi um enorme prazer fazer esta caminhada ao vosso lado.

À minha namorada, pela paciência, enorme compreensão, apoio indispensável e energia positiva com que me brindou constantemente.

Por fim, à minha família, em especial à minha mãe, que foi a chave do meu sucesso! Obrigado por serem quem são e por nunca terem duvidado das minhas capacidades e me terem apoiado incondicionalmente.

Resumo

As alterações climáticas e as suas consequências são um dos temas dominantes na atualidade. Uma das consequências mais evidentes do aquecimento global é a tendência de subida do nível médio do mar (NMM), associada à fusão dos glaciares nas regiões polares e à expansão térmica da água dos oceanos. Esta tendência tem vindo a ser estudada desde há várias décadas a nível global, utilizando séries maregráficas longas e, mais recentemente, outros métodos de observação, por exemplo altimetria por satélite.

O Instituto Hidrográfico (IH) desenvolve trabalhos dedicados ao estudo da subida do nível médio do mar em colaboração com entidades nacionais e internacionais. Neste trabalho estendem-se os estudos já efetuados, no IH, sobre a subida do nível do mar utilizando modelos de regressão e modelos de séries temporais. Também se apresentam as tendências de subida do nível médio do mar, as flutuações devidas a fatores sazonais, e ainda se efetuam previsões da tendência de subida, através da análise e processamento de três séries de registos maregráficos. Neste trabalho foram utilizados registos maregráficos dos portos de Viana do Castelo (compreendidos entre os anos 1984 e 2015), Sines (compreendidos entre os anos 1997 e 2020) e Funchal (compreendidos entre os anos 2005 e 2020) devido à sua localização e importância.

Os resultados obtidos nesta dissertação, evidenciam acentuadas variações anuais e semianuais, e a existência de uma subida generalizada do NMM na costa portuguesa. A série de registos de Viana do Castelo apresenta uma tendência de subida de 1,97 mm/ano, enquanto que, as séries de Sines e Funchal, apresentam valores de 4,27 mm/ano e 4,03 mm/ano, respetivamente. Se as condições atuais forem mantidas, podemos antever que, até ao final do século, o NMM esteja, aproximadamente, 43 cm superior ao atual, no pior dos três cenários analisados.

Palavras-chave: modelos de séries temporais, nível médio do mar, registos maregráficos, sazonalidade, tendência do nível do mar

Abstract

Climate change and the associated consequences are one of the most dominant themes of our time. One of the most obvious consequences of global warming is the mean sea level rise, associated with melting glaciers in the polar regions and the thermal expansion of ocean water. This trend has been studied for several decades at the global level using long tide gauge series and, more recently, with different observation methods like satellite altimetry.

The Portuguese Hydrographic Institute has been developping studies on the sea level rise in collaboration with national and international entities. This work seeks to extend the studies already carried out by the Portuguese Hydrographic Institute on the mean sea level rise, using regression and time series models. Following this, efforts are made to present the rising trend of the mean sea level, the periodic fluctuations due to seasonal factors, and also to make forecasts of the rising trend, through the analysis and processing of three series of tide gauge records. This work used tide gauge records from the ports of Viana do Castelo (data between 1984 and 2015), Sines (data between 1997 and 2020) and Funchal (data between 2005 and 2020) due to their location and importance.

The results obtained in this dissertation, show a pronounced annual and semi-annual variations, as well as the existence of a generalized rise in the mean sea level along the Portuguese coast. The series of records from Viana do Castelo present a rising trend of 1,97 mm/year. As for, the series of Sines and Funchal, it, shows significantly higher values, respectively, 4,27 mm/year and 4,03 mm/year. If the predictions made do not change considerably we can expect that by the end of the century, the mean sea level will be, approximately, 43 cm above the present, in the worst of the three scenarios analysed.

Keywords: time series models, mean sea level, tide gauge records, seasonality, sea level trend

Índice

Introdução	1
1 Revisão da Literatura	5
1.1 O Nível Médio do Mar	5
1.2 Variações	7
1.2.1 Variações Orbitais	9
1.2.2 Fenómenos Eustáticos	13
1.2.3 Fenómenos Isostáticos	14
1.3 Causas da Atual Subida do Nível Médio do Mar	15
1.4 Variações Sazonais	17
1.5 Métodos de Observação do Nível Médio do Mar	19
1.5.1 Marégrafos e Registos Maregráficos	19
1.5.2 Altimetria por Satélite	21
1.5.3 Outros Métodos de Reconstrução do Nível Médio do Mar . .	24
1.6 A Evolução do Nível Médio do Mar	26
1.6.1 Período Holoceno	26
1.6.2 Os Últimos Dois Mil Anos	28
1.6.3 Um Novo Paradigma no Século XX	32
1.6.4 O Século XX em Portugal	34
2 Fundamentação Teórica	37
2.1 Definições	37
2.2 Teste de Hipótese	37
2.3 Série Temporal	40
2.3.1 Componentes de Uma Série Temporal	40
2.3.2 Tendência	41
Ajuste de Uma Função Polinomial	41
2.3.3 Sazonalidade	42
Ajuste de Uma Função de Tendência Cíclica	43
2.4 Estacionariedade	43

2.5	Funções de Autocovariância, Autocorrelação e Autocorrelação Parcial	46
2.6	Ruído Branco	47
2.7	Estimação de Parâmetros Característicos dos Processos Estacionários	47
2.8	Processos Estacionários e Lineares	48
2.8.1	Processo Autoregressivo de Ordem p	48
2.8.2	Processos de Médias Móveis de Ordem q	49
2.8.3	Processos Mistos Autoregressivos e de Médias Móveis	50
2.9	Processos Não Estacionários e Lineares	51
2.9.1	Estabilização de Média e Sazonalidade	53
	Operador Diferença	54
	Operador Diferença Sazonal	54
2.9.2	Estabilização da Variância	55
2.9.3	Modelos Integrados Mistos Não Sazonais e Sazonais (ARIMA e SARIMA)	56
2.10	Modelação ARIMA de Séries Temporais	57
2.10.1	Identificação	57
2.10.2	Estimação	59
2.10.3	Métodos de Diagnóstico dos Modelos	59
	Avaliação da Qualidade Estatística dos Modelos	60
	Avaliação da Qualidade de Ajustamento	61
2.11	CrITÉrios de Seleção de Modelos	62
2.11.1	Análise de Variância (ANOVA) e Teste de <i>Wilcoxon</i>	63
2.11.2	CrITÉrios AIC e BIC	65
	CrITÉrio AIC	65
	CrITÉrio BIC	65
2.11.3	Comparação de ResÍduos	66
	Coeficiente de Determinação	66
	<i>Root Mean Square Error (RMSE)</i>	66
3	Análise das Séries de Dados Maregráficos	67
3.1	Análise dos Registos Maregráficos de Sines	68
3.1.1	Análise Descritiva dos Dados	68
3.1.2	Avaliação da Tendência e Sazonalidades	70
3.1.3	Avaliação da aplicação de modelos ARIMA	78
3.1.4	Seleção e Previsão	84
3.2	Análise dos Registos Maregráficos de Viana do Castelo	86
3.2.1	Análise Descritiva dos Dados	86
3.2.2	Avaliação da Tendência e Sazonalidade	87

3.2.3	Avaliação da aplicação de modelos ARIMA	93
3.2.4	Seleção e Previsão	96
3.3	Análise dos Registos Maregráficos do Funchal	98
3.3.1	Análise Descritiva dos Dados	98
3.3.2	Avaliação da Tendência e Sazonalidade	99
3.3.3	Avaliação da aplicação de modelos ARIMA	104
3.3.4	Seleção e Previsão	107
3.4	Semelhança Estatística entre a Sazonalidade nos Diferentes Registos Maregráficos Estudados	109
Conclusão		111
3.5	Recomendações Futuras	115
Bibliografia		117
Apêndices		123
A Avaliação dos Modelos Identificados - Registos Maregráficos de Sines		123
A.1	$SARIMA(0, 1, 1)(1, 1, 0)_{12}$	123
A.2	$SARIMA(0, 1, 1)(0, 1, 1)_{12}$	125
A.3	$SARIMA(0, 1, 1)(1, 1, 1)_{12}$	127
A.4	$SARIMA(11, 1, 0)$	129
B Avaliação dos Modelos Identificados - Registos Maregráficos de Viana do Castelo		133
B.1	$SARIMA(0, 1, 1)(0, 1, 1)_{12}$	133
B.2	$SARIMA(1, 0, 0)(2, 0, 0)_{12}$	135
C Avaliação dos Modelos Identificados - Registos Maregráficos do Funchal		139
C.1	$SARIMA(1, 0, 0)(2, 1, 0)_{12}$	139
C.2	$SARIMA(0, 0, 2)(0, 1, 1)_{12}$	141
D Código Utilizado na Análise e Processamento das Séries Tempo- rais		145
E Código Utilizado na Avaliação da Semelhança Estatística da Sa- zonalidade Entre as Três Estações Maregráficas		153

Anexos	155
I Participação em: 7.^a Jornadas de Engenharia Hidrográfica/ 2^aJornadas Luso-Espanholas de Engenharia Hidrográfica - Resumo Alargado	155
II Participação em: Jornadas do Mar'22 - Artigo Submetido	161

Lista de Figuras

1.1	Planos de referência. Linha Azul- NMA; Linha Vermelha - ZH (adaptado de IH, 2020)	6
1.2	Exemplificação dos três ciclos de Milankovitch: (a) excentricidade da órbita terrestre; (b) obliquidade da elíptica; (c) precessão equinocial (adaptado de Webb, 2021)	10
1.3	Afélio e periélio (Moura, 2011)	11
1.4	Distribuição de registos maregráficos da base de dados da <i>PSMSL</i> . Marégrafos com séries de dados maregráficos superiores a 60 anos (gráfico superior), marégrafos com séries de dados maregráficos inferiores a 30 anos (gráfico inferior) (adaptado de Cazenave e Nerem, 2004)	21
1.5	Distribuição dos dados altimétricos recolhidos ao longo de um ciclo de 10 dias pelo satélite Topex/Poseidon (extraído de Cazenave e Nerem, 2004)	22
1.6	Distribuição geográfica das tendências do NMM entre 1992-2008, obtidas através de altimetria pelos satélites T/P e Jason 1 (extraído de Cazenave e Nerem, 2004)	23
1.7	(B) reconstrução do NMM relativo a partir de duas localizações diferentes, Sand Point e Tump Point, no estado norte-americano da Carolina do Norte desde 100 a.C. (C) reconstrução do NMM após aplicadas as correções isostáticas nas duas localizações. (extraído de Kemp et al., 2011)	29
1.8	Comparação entre reconstruções do NMM durante o Holoceno tardio após aplicadas correções isostáticas na Carolina do Norte e evidências de NMM noutras regiões do planeta (extraído de Kemp et al., 2011).	31
1.9	Evolução do NMM global entre 1860 a 2009. A linha vermelha apresenta uma reconstrução realizada por Church e White, 2006. Linha azul e linha preta representam reconstruções realizadas, respetivamente, através da análise de registos maregráficos e medições altimétricas realizadas por (adaptado de Church e White, 2011)	32

1.10	Variação do NMM global entre 1990 e 2009. Linha Vermelha representa a reconstrução feita através de dados altimétricos; Linha Azul representa a reconstrução feita com séries de dados marégrafos (extraído de Church e White, 2011)	33
1.11	NMM no marégrafo de Cascais (1882-2008). Reconstrução realizada a partir de séries temporais da média mensal do NMM com sobreposição da média móvel de 10 anos (extraído de Antunes e Taborda, 2009) .	35
2.1	Intervalo de rejeição de H_0 para um teste bilateral quando se assume a normalidade da variável aleatória (adaptado de: shorturl.at/fjnS0 , consultado em 19 de maio de 2022)	38
2.2	Determinação de α e β (adaptado de: shorturl.at/gHOS1 , consultado em 20 de maio de 2022)	39
2.3	Ajuste de uma função polinomial de grau um (adaptado de: shorturl.at/lvwZ0 , consultado em 26 de maio de 2022)	42
2.4	Gráfico da esquerda: Série estacionária. Gráfico da direita: Série não estacionária (adaptado de: https://directions.pt/inteligencia/machine-learning-na-previsao-de-series-temporais , consultado em 15 de abril de 2022)	45
2.5	correlograma das funções FAC e FACP de um processo AR(2), (adaptado de Santos, 2011)	48
2.6	correlograma das funções FAC e FACP de um processo AM(2), (adaptado de Santos, 2011)	49
2.7	correlograma das funções FAC e FACP de um processo ARMA(1,1), (adaptado de Santos, 2011)	51
3.1	Localização da estação maregráfica do porto de Sines ($\varphi = 37^{\circ}56,89'N$ e $\lambda = 008^{\circ}53,27'W$) (fonte: IH, 2020)	69
3.2	Histograma das frequências dos NMm em Sines no período 1997 - 2020 (Picoto et al., 2022a e Picoto et al., 2022b)	69
3.3	Linha preta: evolução dos NMm's em Sines no período em análise; linha vermelha: modelo ajustado (Picoto et al., 2022a e Picoto et al., 2022b)	71
3.4	Coeficientes do modelo polinomial de grau um estimado	71
3.5	Coeficientes do modelo polinomial de grau dois estimado	72
3.6	Comportamento dos níveis médios ao longo do ano em Sines no período 1997 a 2020 (Picoto et al., 2022a e Picoto et al., 2022b	73

3.7	Linha preta: série de dados em análise; linha vermelha: modelos ajustado (modelo 3)	74
3.8	Coeficientes do modelo do modelo 3	74
3.9	Linha preta: série de dados em análise; linha vermelha: modelos ajustado (modelo 4)	75
3.10	Coeficientes do modelo do modelo 4	75
3.11	correlogramas da FAC e FACP residual do modelo selecionado . . .	77
3.12	Coeficientes do modelo 4 com a adição do coeficiente β_2	78
3.13	Coeficientes dos modelos com sazonalidade trimestral e quadrimestral: (A) modelo com 3 sinusoides; (B) modelo com 4 sinusoides. . .	78
3.14	Registos maregráficos de Sines, primeiros 230 meses.	79
3.15	Correlogramas da série de Sines: ACF (gráfico superior) e ACFP (gráfico inferior)	79
3.16	(A) Diferenciação simples; (B) Diferenciação sazonal	80
3.17	Correlogramas da série estacionária - FAC (gráfico superior) e FACP (gráfico inferior)	80
3.18	Correlogramas dos resíduos do modelo $SARIMA(1, 1, 1)(0, 1, 1)_{12}$: FAC (gráfico superior) e FACP (gráfico inferior)	82
3.19	Histograma da frequências dos resíduos do modelo	82
3.20	Ajustamento das previsões aos meses não considerados na análise para construção dos modelos.	84
3.21	Previsão dos NMm's para os próximos 10 anos	85
3.22	Localização da estação maregráfica do porto de Viana do Castelo ($\varphi = 41^\circ 41', 1'N$ e $\lambda = 008^\circ 50', 28'W$) (fonte: IH, 2020)	86
3.23	Linha preta: evolução dos NMm's em Viana do Castelo no período em análise; linha vermelha: modelo ajustado	87
3.24	Coeficientes do modelo polinomial de grau um estimado para Viana do Castelo	88
3.25	Coeficientes do modelo polinomial de grau dois estimado para a série de registos de Viana do Castelo	88
3.26	Comportamento dos níveis médios ao longo do ano, em Viana do Castelo no período 1984 a 2015	89
3.27	Coeficientes do modelo de tendência linear e cíclica	90
3.28	Linha preta: série de dados em análise (Viana do Castelo); linha vermelha: modelo ajustado (modelo 3)	91
3.29	Correlogramas dos resíduos do modelo:FAC (gráfico superior) e FACP (gráfico inferior)	91

3.30	Coeficientes do modelo 3 com a adição do coeficiente β_2	92
3.31	Coeficientes dos modelos com sazonalidade trimestral e quadrimestral: (A) modelo com 3 sinusoides; (B) modelo com 4 sinusoides. . .	93
3.32	Correlogramas dos resíduos do modelo $SARIMA(1, 0, 0)(0, 1, 1)_{12}$: FAC (gráfico superior) e FACP (gráfico inferior)	94
3.33	Histograma das frequências dos resíduos do modelo escolhido	95
3.34	Ajustamento das previsões aos meses não considerados na análise para construção dos modelos para Viana do Castelo	96
3.35	Previsão dos NMm's para os próximos 10 anos para Viana do Castelo	97
3.36	Ajustamento de um modelo polinomial de grau um aos últimos 6,5 anos de registos em Viana do Castelo	98
3.37	Localização da estação maregráfica do porto do Funchal ($\varphi = 32^\circ 38, 64' N$ e $\lambda = 016^\circ 54, 78' W$) (fonte: IH, 2020)	98
3.38	Linha preta: evolução dos NMm's no Funchal no período em análise; linha vermelha: modelo ajustado	99
3.39	Coeficientes do modelo de polinomial de grau um estimado para o Funchal	100
3.40	Coeficientes do modelo polinomial de grau dois estimado para a série de registos do Funchal	100
3.41	Comportamento dos níveis médios ao longo do ano, no Funchal no período de 2005 a 2020	101
3.42	Coeficientes do modelo de tendência linear e tendência cíclica	101
3.43	Linha preta: série de dados em análise (Funchal); linha vermelha: modelo ajustado (modelo 3)	102
3.44	Correlogramas dos resíduos do modelo: FAC (gráfico superior) e FACP (gráfico inferior)	103
3.45	Coeficientes do modelo 3 com a adição do coeficiente β_2	104
3.46	Coeficientes dos modelos com sazonalidade trimestral e quadrimestral: (A) modelo com 3 sinusoides; (B) modelo com 4 sinusoides. . .	104
3.47	Correlogramas dos resíduos do modelo $SARIMA(1, 0, 0)(0, 1, 1)_{12}$: FAC (gráfico superior) e FACP (gráfico inferior)	105
3.48	Histograma dos resíduos do modelo escolhido	106
3.49	Ajustamento das previsões aos meses não considerados na análise para construção dos modelos para o Funchal	107
3.50	Previsão dos NMm's para os próximos 10 anos para o Funchal . . .	108
3.51	Ajustamento de um modelo polinomial de grau um aos últimos 3 anos de registos da estação maregráfica do Funchal	109

3.52	Níveis médios mensais característicos dos níveis médios das três séries de registos maregráficos	109
3.53	Níveis médios mensais característicos (pontos azuis) com ajustamento dos níveis médios diários (vermelho), (fonte: Carinhas et al., 2021) .	111
A.1	Correlogramas dos resíduos do modelo $SARIMA = (0, 1, 1)(1, 1, 0)_{12}$: FAC (gráfico superior) e FACP (gráfico inferior)	124
A.2	Histograma das frequências dos resíduos do modelo	125
A.3	Correlogramas dos resíduos do modelo $SARIMA(0, 1, 1)(0, 1, 1)_{12}$: FAC (gráfico superior) e FACP (gráfico inferior)	126
A.4	Histograma de frequências dos resíduos do modelo	127
A.5	Correlogramas dos resíduos do modelo $SARIMA = (0, 1, 1)(1, 1, 1)_{12}$: FAC (gráfico superior) e FACP (gráfico inferior)	128
A.6	Histograma das frequências dos resíduos do modelo	129
A.7	Correlogramas dos resíduos do modelo $SARIMA = (11, 1, 1)$: FAC (gráfico superior) e FACP (gráfico inferior)	131
A.8	Histograma das frequências dos resíduos do modelo	132
B.1	Correlogramas dos resíduos do modelo $SARIMA = (0, 1, 1)(0, 1, 0)_{12}$: FAC (gráfico superior) e FACP (gráfico inferior)	134
B.2	Histograma dos resíduos do modelo	135
B.3	Correlogramas dos resíduos do modelo $SARIMA = (1, 0, 0)(2, 0, 0)_{12}$: FAC (gráfico superior) e FACP (gráfico inferior)	136
B.4	Histograma dos resíduos do modelo	137
C.1	Correlogramas dos resíduos do modelo $SARIMA = (1, 0, 0)(2, 1, 0)_{12}$: FAC (gráfico superior) e FACP (gráfico inferior)	140
C.2	Histograma dos resíduos do modelo	141
C.3	Correlogramas dos resíduos do modelo $SARIMA = (0, 0, 2)(1, 1, 0)_{12}$: FAC (gráfico superior) e FACP (gráfico inferior)	142
C.4	Histograma dos resíduos do modelo	143

Lista de Tabelas

1.1	Fenómenos que originaram variações no NMM ao longo dos últimos 96 mil anos (adaptado de Da Silva et al., 2014)	9
2.1	Erros associados à aceitação, ou rejeição, de H_0	39
2.2	Estimadores de parâmetros característicos	47
2.3	Processos estacionários lineares (adaptado de Murteira et al., 2000)	51
2.4	Parâmetros de λ (Murteira et al., 2000)	55
2.5	Transformações a aplicar a série não estacionária	56
2.6	Condições de Estacionariedade e invertibilidade (adaptado de Murteira et al. (2000))	61
3.1	Estatísticas descritivas dos NMm's (mm) de Sines	70
3.2	R^2 e $RMSE$ dos quatro modelos em análise	76
3.3	Coefficientes do modelos $SARIMA(1, 1, 1)(0, 1, 1)_{12}$	81
3.4	Intervalos de confiança dos coeficientes do modelo	81
3.5	Critérios AIC e BIC; $RMSE$	83
3.6	$RMSE$ e nº de coeficientes dos modelos 4 e $SARIMA(1, 1, 1)(0, 1, 1)_{12}$	85
3.7	Estatísticas descritivas dos NMm's (mm) de Viana do Castelo	87
3.8	R^2 e $RMSE$ dos modelo em análise	90
3.9	Coefficientes do modelos $SARIMA(1, 0, 0)(0, 1, 1)_{12}$	93
3.10	Intervalos de confiança do modelo identificado	94
3.11	Critérios AIC e BIC; $RMSE$	95
3.12	$RMSE$ e nº de coeficientes dos modelos 3 e $SARIMA(1, 0, 0)(0, 1, 1)_{12}$	96
3.13	Estatísticas descritivas dos NMm's (mm) do Funchal	99
3.14	R^2 e $RMSE$ dos modelo em análise	102
3.15	Coefficientes do modelos $SARIMA(1, 0, 0)(0, 1, 1)_{12}$	105
3.16	Intervalos de confiança do modelo identificado	105
3.17	Critérios AIC e BIC; $RMSE$	106
3.18	$RMSE$ e nº de coeficientes dos modelos 3 e $SARIMA(1, 0, 0)(0, 1, 1)_{12}$	107
3.19	Cota adicionada aos registos maregráficos (PSMSL, s.d.)	110
A.1	modelos identificados	123

A.2	Coeficientes do modelos $SARIMA(0, 1, 1)(1, 1, 0)_{12}$	123
A.3	Intervalos de confiança dos coeficientes	124
A.4	Coeficientes do modelos $SARIMA(0, 1, 1)(0, 1, 1)_{12}$	125
A.5	Intervalos de confiança dos coeficientes	126
A.6	Coeficientes do modelos $SARIMA(0, 1, 1)(1, 1, 1)_{12}$	127
A.7	Intervalos de confiança dos coeficientes	128
A.8	Coeficientes do modelos $SARIMA(11, 1, 0)$	130
A.9	Intervalos de confiança dos coeficientes	131
B.1	modelos identificados	133
B.2	Coeficientes do modelos $SARIMA(0, 1, 1)(0, 1, 1)_{12}$	133
B.3	Intervalos de confiança dos coeficientes	134
B.4	Coeficientes do modelos $SARIMA(1, 0, 0)(2, 0, 0)_{12}$	135
B.5	Intervalos de confiança dos coeficientes	136
C.1	modelos identificados	139
C.2	Coeficientes do modelos $SARIMA(1, 0, 0)(2, 1, 0)_{12}$	139
C.3	Intervalos de confiança dos coeficientes	140
C.4	Coeficientes do modelos $SARIMA(0, 0, 2) \times (0, 1, 1)_{12}$	141
C.5	Intervalos de confiança dos coeficientes	142

Lista de Abreviaturas

AIC	A kaike I nformation C riterion
BIC	B ayesian I nformation C riterion
BMmin	N ível da M aré A stronómica M ais B aixa
FAC	F unção A uto C orrelação
FACP	F unção A uto C orrelação P arcial
NMA	N ível M édio A dotado
NMM	N ível M édio do M ar
NMm	N ível M édio m ensal
LI	L imite I nferior
LS	L imite S uperior
PSMSL	P ermanent S ervice for M ean S ea L evel
T/P	T opex/ P oseidon
ZH	Z ero H idrográfico
RLR	R evised L ocal R eference
RMSE	R oot M ean S quare E rror

Lista de Símbolos

X_t	série temporal	mm
H_0	hipótese nula	
H_1	hipótese alternativa	
α	nível de significância	
ω	frequência angular	rad
T_t	tendência da série temporal	mm/ano
S_t	sazonalidade	mm
ϵ_t	componente puramente aleatória	mm
R^2	coeficiente de determinação	
$E()$	valor esperado	
μ	média da população	
$\tilde{\mu}$	mediana da população	
γ_K	função de autocovariância	
ρ_K	função de autocorrelação	
ϕ_{kk}	função de autocorrelação parcial	
σ^2	variância da população	
$Var()$	variância da população	
$Cov()$	covariância da população	
$\phi_p(B)$	polinómio autoregressivo de ordem p	
$\theta_q(B)$	polinómio de médias móveis de ordem q	
Y_t	filtro linear de X_t	
$\nabla^d X_t$	operador de diferença posterior de ordem d de X_t	
$\nabla_s^D X_t$	operador de diferença sazonal posterior de ordem D de X_t	
$T(X_t)$	transformação de <i>Box-Cox</i>	
n	dimensão da amostra	
m	grau do polinómio	
ζ	coeficiente do polinómio	
j	número de sinusoides do modelo	

Introdução

Desde sempre o Homem procurou instalar-se nas zonas ribeirinhas, por forma a beneficiar das ricas fontes de alimento constituídas pela vida marinha, da produtividade agrícola das planícies adjacentes a rios, estuários e oceanos e da utilização do meio aquático como via de comunicação. Na atualidade, a coexistência do Homem no litoral é ainda mais notória pois, para além dos fatores aludidos, é também no litoral que os grandes centros populacionais estão localizados. No caso de Portugal, a percentagem da população residente no litoral tem apresentado aumentos consistentes. Atualmente, estima-se que, aproximadamente, 83% da população portuguesa viva no litoral (Mendes et al., 2020).

No entanto, as alterações climáticas que se têm vindo a manifestar recentemente, têm como uma das consequências mais evidentes a subida do nível médio do mar (NMM), associada à fusão dos glaciares nas regiões polares, em particular sobre a Antártida e a Gronelândia.

A tendência de subida do NMM tem sido estudada desde há várias décadas a nível global, com os estudos a identificarem tendências de subida do NMM a oscilar entre 1,5-3,2 mm/ano durante o século XX e inícios do século XIX (Church e White, 2011; Boski, 2013). No entanto, conforme Picoto et al. (2022a), esta subida é difícil de quantificar essencialmente porque: 1) Desconhece-se o comportamento geral desta subida, ou seja, desconhece-se o modelo (ou função) aplicável às recentes variações do NMM; 2) Os valores da tendência de subida apresentam diferenças notáveis de local para local; 3) O número de estações maregráficas, com séries de observações suficientemente longas para que permitam inferir tendências significativas é bastante diminuto, e a sua distribuição pelo globo é assimétrica. Não obstante estas limitações, os recentes avanços tecnológicos, nomeadamente, na área da medição altimétrica do NMM têm vindo a mitigar a maioria destes inconvenientes (Picoto et al., 2022a).

Em Portugal existem várias séries de registos maregráficos provenientes de marégrafos localizados ao longo do seu litoral. Tanto os registos maregráficos de Cascais como de Lagos evidenciam uma tendência de aumento do NMM (Picoto

et al., 2022a); no entanto, as análises das séries de dados maregráficos sugerem que o NMM não subiu com uma tendência consistente ao longo do tempo (Picoto et al., 2022a). Desde finais do século XIX até cerca de 1915-1920, o NMM teve, em Cascais, uma tendência negativa, $-5\pm 0,4$ mm/ano (Picoto et al., 2022a). Só posteriormente a este período é que se nota uma subida do NMM, sendo que a atual tendência de subida em Cascais é de $1,44\pm 0,85$ mm/ano, obtida a partir da análise dos registos maregráficos nos últimos 30 anos (Mendes et al., 2020).

Nas últimas décadas tem-se verificado uma subida generalizada do NMM a nível global. Se as atuais estimativas não sofrerem alterações significativas é expectável que, o NMM, no final do século XXI, esteja, aproximadamente, 34 cm acima do atual (Jevrejeva et al., 2008). Esta realidade permite antever que, nas próximas décadas, se intensifique a erosão nas áreas costeiras, recuo da linha de costa, aumento da frequência de inundações nas zonas ribeirinhas, destruição de infraestruturas, etc. (Taborda e Dias, 1988). Por exemplo, a nível nacional estima-se que em 2100, aproximadamente, 1,29% do território continental de Portugal esteja suscetível de sofrer inundações com elevada frequência (Rocha et al., 2018).

Portugal tem uma orla costeira densamente povoada, onde se desenvolvem a maioria das atividades socioeconómicas do país e é natural que, cada vez mais, haja uma elevada preocupação, tanto por parte dos governantes como da população, com questões relacionadas com a subida do NMM e com o impacte que poderá causar. Não obstante esta preocupação, verifica-se uma certa inércia por parte das entidades competentes em tomarem as medidas necessárias para que num futuro a médio e longo prazo se minimize ou mesmo evite os impactes negativos desta problemática. Como tal, com este trabalho, pretende-se dar um contributo para a análise da evolução do NMM, através da estimação da tendência de subida e obtenção de previsões de elevação do NMM em Portugal, como forma de sensibilizar a comunidade sobre este assunto.

Os espaços marítimos sob jurisdição nacional constituem um património que é importante salvaguardar e proteger; como só protegemos o que conhecemos, o Instituto Hidrográfico (IH) tem um papel preponderante no conhecimento dos oceanos, revelando-se fundamental para a prossecução dos interesses nacionais (IH, 2021). A missão do IH decorre da sua lei orgânica: “O IH tem por missão fundamental assegurar atividades relacionadas com as ciências e técnicas do mar, tendo em vista a sua aplicação na área militar, e contribuir para o desenvolvimento do País nas áreas científica e de defesa do ambiente marinho”. Este objetivo é concretizado através da observação do meio marinho, com carácter pontual ou permanente.

Na área da oceanografia, tanto a nível nacional como internacional, destaca-se por manter um sistema de monitorização permanente que integra uma rede de boias (condições de agitação marítima), uma rede de marégrafos costeiros (variações do nível do mar), uma rede de estações meteorológicas costeiras e uma rede de radares HF costeiros (correntes superficiais ao largo) (IH, 2021). Na área das marés, o IH recolhe, analisa e fornece informação de marés e níveis de água para as mais diversas finalidades, nomeadamente: segurança à navegação, estabelecimento de níveis de referência em Hidrografia, execução de projetos de engenharia costeira, investigação científica, pesca e atividades recreativas (IH, 2021).

Neste trabalho, pretende-se dar continuidade aos estudos já efetuados no IH sobre a subida do NMM utilizando um conjunto de técnicas estatísticas, habitualmente, utilizadas para modelar séries de dados, nomeadamente, modelos de regressão e modelos de séries temporais. Foram selecionados três registos maregráficos, dois provenientes de portos continentais: os registos maregráficos do porto de Viana do Castelo compreendidos entre os anos 1984 e 2015 e os do porto de Sines entre os anos 1997 e 2020; e registos maregráficos de um porto insular, os do porto do Funchal entre os anos de 2005 e 2020. Os registos das estações maregráficas, selecionadas, foram disponibilizados pelo IH - Divisão de Oceanografia.

São propostos os seguintes objetivos específicos para os três registos maregráficos selecionados:

1. Avaliar a atual tendência de subida do NMM;
2. Obter uma estimativa da tendência de subida do NMM;
3. Caracterizar as variações sazonais do NMM;
4. Realizar previsões do comportamento do NMM.

Os modelos de regressão serão utilizados para identificar eventuais variações temporais da tendência e da amplitude das variações sazonais. A caracterização dessas variações sazonais será ainda complementada com a identificação dos níveis médios mensais característicos por forma a identificar as sucessivas evoluções que se manifestam ao longo do ano.

A obtenção de previsões do comportamento do NMM será feita através do modelo que apresente o melhor ajustamento a cada uma das séries. Para isso, além dos modelos de regressão, serão ainda analisados os modelos de média móvel integrada autorregressiva sazonal (SARIMA).

Esta dissertação encontra-se estruturada em quatro capítulos. O primeiro capítulo dedicado à revisão bibliográfica do tema e de estudos anteriores realizados nesta área.

O segundo capítulo apresenta a metodologia utilizada no estudo das séries de dados maregráficos. No decorrer do capítulo são descritos os fundamentos teóricos dos métodos e processos necessários para a análise de uma série temporal.

No terceiro capítulo são descritas as séries de dados maregráficos que foram utilizados neste trabalho, bem como os procedimentos pelo qual passaram até poderem ser analisadas. Além disso, são apresentados os resultados obtidos através do uso da metodologia descrita no segundo capítulo.

O quarto capítulo apresenta as conclusões que se podem retirar das observações mencionadas no capítulo terceiro, sugestões para a aplicação no IH e ainda recomendações para trabalhos futuros.

Capítulo 1

Revisão da Literatura

1.1 O Nível Médio do Mar

Na nossa vida quotidiana utilizamos inconscientemente níveis de referência (“*data*”¹). Por exemplo, quando dizemos que um determinado objeto tem x metros de altura, o solo torna-se o nível de referência a partir do qual foi possível fazer a medição da altura. No entanto, existem numerosas situações em que não podemos considerar o solo como nível de referência. Se nos tivermos a referir à altura de uma montanha ou de um edifício que está localizado num plano inclinado, naturalmente que precisamos de um referencial diferente do solo. Neste caso, necessitamos de um ponto bem definido para que possa servir de referencial da medição (IOC, 2002).

Para descrever rigorosamente a cota e o nível da maré num determinado local, e relacioná-los com as cotas e níveis de maré noutros locais, torna-se necessário estabelecer níveis de referência. No caso das marés, o Zero Hidrográfico (ZH) é o plano de referência convencionado, situado abaixo do Nível da Maré Astronómica Mais Baixa (BMmin) ao qual estão referidas as sondas e linhas batimétricas apresentadas nas cartas náuticas, bem como as previsões de marés publicadas nas Tabelas de Marés do IH (IH, 2020). A sua implementação permite que a prática de navegação seja feita em segurança. Com exceção de Cascais e do porto de Lisboa, o ZH em Portugal Continental está estabelecido em 2,00 metros abaixo do nível médio adotado (NMA). Nos portos de Lisboa e Cascais está 2,08 metros abaixo. Nos Arquipélagos dos Açores e da Madeira, o ZH situa-se 1,00 e 1,40 metros abaixo do NMA, respetivamente (IH, 2020).

Conforme ilustrado na Figura 1.1, no caso de cotas ou altitudes, o NMA é o nível de referência. Corresponde ao nível médio em Cascais determinado com base em 56 anos de dados maregráficos centrados em 1910 (Da Silva et al., 2014).

¹modelo matemático teórico que representa a superfície da terra.

Para cada ponto da costa existem diferenças entre o ZH e o *datum* altimétrico, isto é, o NMA. Nas Tabelas de Maré estão indicadas as relações entre o ZH e o *datum* altimétrico para vários locais da costa, sendo estas imutáveis ao longo do tempo (IOC, 2002).

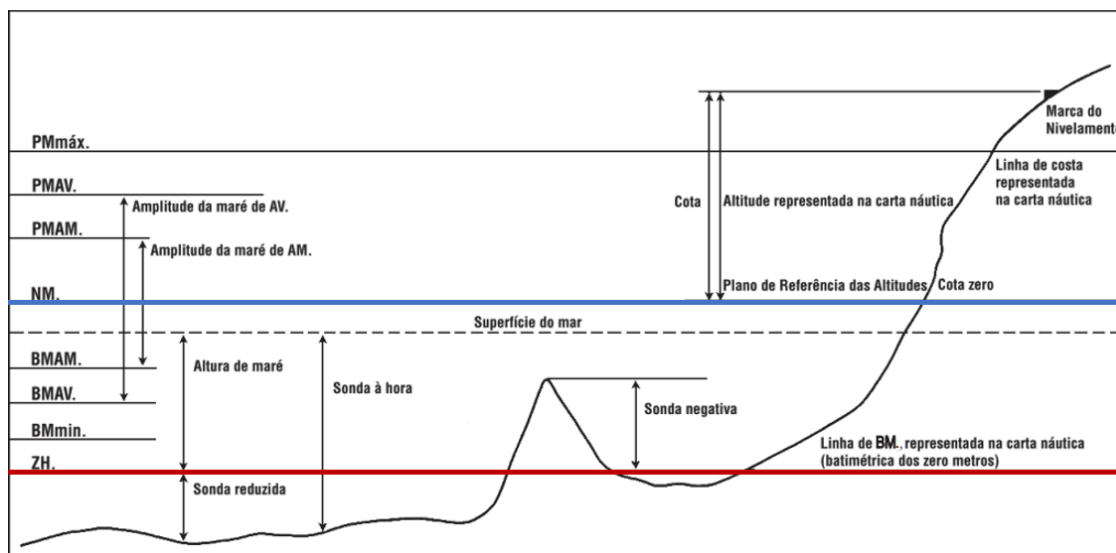


FIGURA 1.1: Planos de referência. Linha Azul- NMA; Linha Vermelha - ZH (adaptado de IH, 2020)

O NMA é materializado através de marcas de nivelamento, cuja a sua cota indica o afastamento vertical relativamente ao NMA. Estas marcas existem por todo o território nacional. Para que isto tenha sido possível, foi realizado o transporte da cota vertical em Cascais para todo o país, através de nivelamento de precisão realizado pelo Instituto Geográfico Português (Martins e Reis, 2007).

Ao contrário do NMA, o nível médio do mar (NMM) varia lentamente ao longo do tempo, por causas que serão descritas mais adiante. Como o NMM subiu desde o período em que foi estabelecido o NMA, é seguro admitir que o nível da maré observada é em geral superior ao nível da maré prevista (Martins e Reis, 2007). Deste modo, o problema está em como definir o "nível médio", sendo que este é variável.

A obtenção do NMM é realizada a partir de longas séries de observações horárias (ou por vezes tri-horárias). A partir destas, calculam-se os níveis médios diários e, com base nestes, os níveis médios mensais. Posteriormente, o nível médio anual é obtido através da média ponderada dos valores mensais. Consuante a janela temporal, o nível médio pode ser, designadamente, diário, mensal, anual ou secular,

não se devendo considerar os fenómenos que provoquem variações de período inferior à escala escolhida (Da Silva et al., 2014). Por exemplo, a utilização do nível médio mensal permite eliminar a variância da altura da superfície das águas, acima do ZH, provocada pela maré e agitação marítima. Se for considerado o nível médio anual, eliminam-se também as flutuações causadas pelas variações sazonais da pressão atmosférica ao longo do ano (Da Silva et al., 2014).

Os registos do marégrafo de Cascais destacam-se por serem suficientemente longos, de tal forma que permitem estudar tendências seculares. O início dos seus registos data de 1882, no entanto, nem sempre foram realizados da mesma forma. Ao longo do período de medição foram realizadas mudanças de critérios de medição e análise (Da Silva et al., 2014). Por exemplo, os primeiros registos apenas se referiam às alturas da água na preia-mar e baixa-mar, e só mais tarde passaram a ser efetuados registos horários da altura (ou nível) da maré. De acordo com Da Silva et al. (2014), pensa-se que os registos das alturas horárias da maré iniciaram-se por volta de 1938, não havendo registo da data exata. Uma vez que houve, de alguma forma, alteração no processo de obtenção do nível médio anual, não se pode deixar de questionar a sua consistência em termos de qualidade. Em 1938 foi adotado o nível médio atual, com base nos registos disponíveis até à data. Mais tarde, nos inícios da década de cinquenta, procedeu-se à calibração da marca altimétrica de referência, mantendo-se, no entanto, o NMA (Martins e Reis, 2007). O NMA mantém-se até aos dias de hoje, uma vez que a sua atualização obrigaria à alteração dos valores numéricos de todos os nivelamentos, sondas e linhas batimétricas representadas nas cartas náuticas, efetuados a partir do atual NMA.

1.2 Variações

Como já apresentado por Picoto et al. (2022a), existe uma diversidade de fatores que influenciam o NMM e, em geral, fazem-se sentir no espaço e no tempo, com maior ou menor intensidade e interligação. A compreensão de muitos destes processos é, em muitos casos, complexa, e apenas em alguns, como as marés, devidamente compreendidos. Em contrapartida, as variações interanuais são muito menos previsíveis. Em estudos de análise das variações do NMM, cujo objetivo é provar a relação entre estas e o atual aquecimento global, associado a atividades antrópicas, são muitas vezes omitidos ou subvalorizados os fenómenos subjacentes que apesar de conhecidos, não existe a quantificação necessária do seu impacto (Dias e Taborda, 1988 e Picoto et al., 2022a).

As variações orbitais têm, seguramente, influência na evolução do NMM. As variações na órbita terrestre têm sido apontadas como um dos principais fenômenos de variação do clima e consequentemente do NMM ao longo do período Quaternário, que se iniciou há, aproximadamente, 2,6 milhões de anos (Hansel e McKay III, 2010).

Conforme Rovere et al. (2016), após o último período glaciário e, essencialmente, durante o período Holoceno, que se iniciou há 11,7 mil anos e perdura até ao presente, existem essencialmente dois grandes fenômenos, a nível global, que desencadeiam variações de longas escalas de tempo no NMM devido às interações Criosfera-Oceano, os fenômenos eustáticos e isostáticos. Os fenômenos eustáticos provocam variações no volume e na massa dos oceanos, enquanto os fenômenos isostáticos estão relacionados com a elevação ou subsidência da crosta terrestre, provocando uma variação relativa do NMM (Rovere et al., 2016; Picoto et al., 2022a).

As flutuações do NMM causadas por fenômenos glácio-eustáticos estão fortemente associadas ao aquecimento e arrefecimento global que consequentemente levam à fusão ou formação das calotas polares e glaciares (Rovere et al., 2016; Picoto et al., 2022a). Estes processos desenvolvem-se em escalas temporais de vários milénios. Durante os períodos glaciários, o nível do mar desce porque a humidade que é evaporada dos oceanos não é repostada, ficando aprisionada sob a forma de gelo nas calotas polares. De acordo com Rovere et al. (2016), em períodos interglaciários, com temperaturas mais amenas, dá-se o recuo dos glaciares e, consequentemente, o aumento do NMM devido à reposição da água que estava retida nas calotas (Picoto et al., 2022a).

Os movimentos isostáticos, que derivam dos fenômenos eustáticos, também desempenham um papel fundamental nas variações do NMM. Estas variações podem não ser evidentes em locais que anteriormente se encontravam cobertos por calotas polares. Ao longo destes locais, ocorrem fenômenos de recuperação isostática da carga glaciária (Vieira, 1981). Por outras palavras, existe uma elevação destas zonas. Esta elevação ocorre porque nos períodos em que esses locais ficaram cobertos de gelo, passaram a suportar o seu peso, fazendo com que houvesse uma ligeira subsidência (Rovere et al., 2016). Posteriormente, após vários milénios, esses territórios deixaram de estar cobertos de gelo originando a sua elevação, através de um processo que se designa por equilíbrio isostático. Os fenômenos isostáticos originam movimentos verticais de elevação ou subsidência de áreas da crosta terrestre onde há recuperação isostática, influenciando assim a variação do NMM (Vieira, 1981; Rovere et al., 2016).

As ações antrópicas têm sido apontadas como um novo mecanismo que potencia a subida do NMM, por coincidirem com seu recente aumento. No entanto, estes fatores, bem como muitos outros de menor significado, são ainda pouco conhecidos.

A Tabela 1.1 ilustra, de forma resumida, os principais fenômenos que desencadearam alterações no NMM ao longo dos últimos 96 mil anos.

PERIODICIDADE	MECANISMOS OU FENÓMENO	NATUREZA DO FENÓMENO
Período pós-Revolução Industrial	Alterações climáticas	Interação Oceano-Atmosfera
Dezenas a centenas de anos	Mal conhecidos	Mal conhecido ou desconhecido
1 a 3 mil anos	Fenômenos eustáticos e isostáticos	Interação Criosfera-Oceano-Atmosfera
7 a 17 mil anos		
21 mil anos	Precessão Equinocial	Variações orbitais
42 mil anos	Obliquidade da Elíptica	
96 mil anos	Excentricidade da órbita terrestre	

TABELA 1.1: Fenômenos que originaram variações no NMM ao longo dos últimos 96 mil anos (adaptado de Da Silva et al., 2014)

Nas subseções seguintes serão descritos, em detalhe, os principais fenômenos que influenciam a variação do NMM.

1.2.1 Variações Orbitais

Nos últimos 200 milhões de anos, os movimentos das placas tectônicas originaram a desintegração do supercontinente Pangeia. Estes movimentos provocaram uma grande assimetria da distribuição das massas continentais entre os dois hemisférios (Nott, 2017), com cerca de 67% das massas continentais localizadas no Hemisfério Norte (Eldredge, 2002). Os movimentos das placas também originaram a formação de cadeias montanhosas e alteraram a configuração das bacias oceânicas. Consequentemente, produziram também alterações dos padrões das correntes oceânicas. Paralelamente, as variações milenares de parâmetros da órbita terrestre, tal como a excentricidade, originaram diferenças na insolação recebida por cada hemisfério no verão e no inverno, o que provoca variações cíclicas do clima. Estes fatores

provocaram um arrefecimento gradual do clima nos últimos 60 milhões de anos. O aumento mais acentuado da superfície coberta por gelo começou há, aproximadamente, 2,6 milhões de anos, no início do período Quaternário. Entre 2,6 milhões de anos e 800 mil anos atrás, estima-se que a cada 41 mil anos ocorra um período Glacial; no entanto, desde há 800 mil anos, passaram a ocorrer a cada 100 mil anos (Nott, 2017).

Ao longo do período Quaternário, as mudanças climáticas foram fortemente afetadas por ligeiras variações na radiação solar incidente na superfície da Terra. Antes destes períodos não é possível reconstituir as alterações climáticas de forma fiável mas pensa-se que as variações climáticas também estejam relacionadas com variações de radiação. Estas alterações são uma consequência direta da posição da Terra em relação ao Sol, que depende de três fatores: a excentricidade da órbita terrestre, a obliquidade da elíptica e a precessão equinocial (NASA, 2020; Webb, 2021). A Figura 1.2 apresenta os três fatores mencionados.

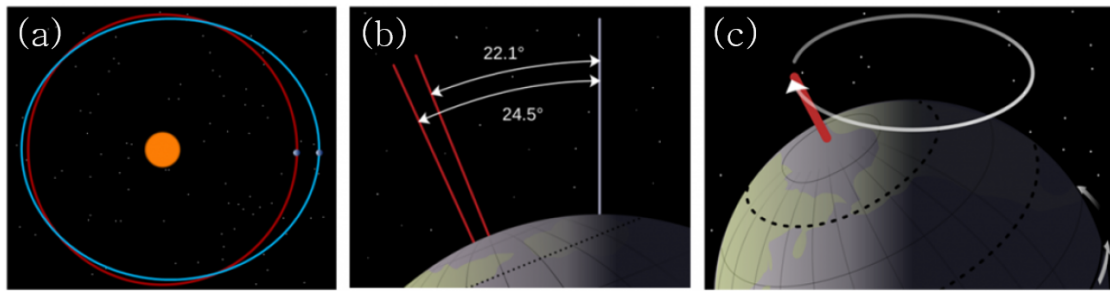


FIGURA 1.2: Exemplificação dos três ciclos de Milankovitch: (a) excentricidade da órbita terrestre; (b) obliquidade da elíptica; (c) precessão equinocial (adaptado de Webb, 2021)

A importância destes três fatores nas variações cíclicas do clima, conhecidos por ciclos de Milankovitch, foi uma hipótese apresentada pelo Jugoslavo Milutin Milankovitch no início do século XX. Milankovitch percebeu que apesar das variações nos ciclos orbitais não afetarem a quantidade de radiação recebida pela Terra, afetam a sua distribuição pela superfície terrestre entre os dois Hemisférios e entre o verão e o inverno (Webb, 2021).

De seguida, efetua-se uma descrição de cada um dos ciclos de Milankovitch e os respetivos períodos associados.

Em primeiro lugar, a órbita da Terra em torno do Sol é muito aproximadamente circular (atualmente, é uma elipse com excentricidade 0,0167), mas a forma da órbita terrestre sofre oscilações devido às perturbações gravitacionais exercidas

pelos dois maiores planetas do Sistema Solar, particularmente pelo planeta Júpiter. Como consequência, a distância entre a Terra e o Sol não é constante, sendo que, atualmente, é cerca de 3,4% superior no momento em que os dois corpos se encontram mais afastados relativamente ao momento que estão mais próximos. Esta situação faz com que a Terra receba aproximadamente 6,8% mais radiação no periélio do que no afélio, respetivamente, a designação dos períodos em que o Sol e a Terra se encontram mais próximos e mais afastados (Figura 1.3).

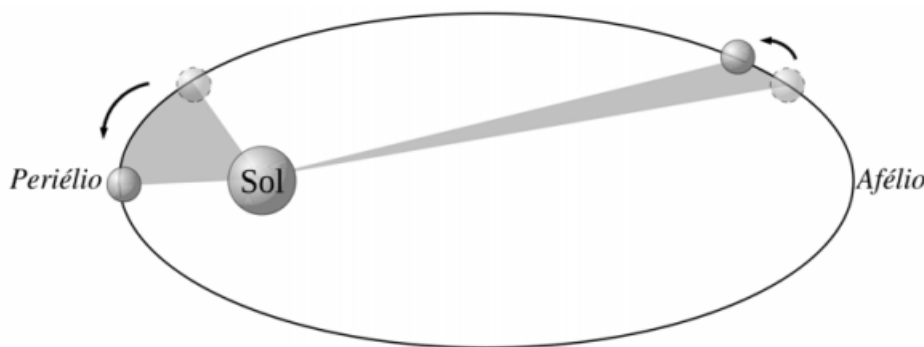


FIGURA 1.3: Afélio e periélio (Moura, 2011)

Além disso, as estações do ano não têm a mesma duração sendo que o verão no Hemisfério Norte é 4,5 dias maior do que o inverno e a primavera cerca de 3 dias maior que o outono (Webb, 2021).

A excentricidade da órbita terrestre varia ao longo de um ciclo de cerca de 100 mil anos. Atualmente, descreve uma órbita quase circular, devido à reduzida excentricidade. O atual valor de excentricidade encontra-se a reduzir gradualmente, o que significa que em alguns milhares de anos, a diferença na duração das estações do ano e radiação solar entre o periélio e afélio, como consequência da excentricidade, será ainda mais reduzida. No entanto, no momento em que a órbita da Terra atingir o valor máximo de excentricidade, a quantidade de radiação incidente na Terra, será 23% superior no periélio do que no afélio (Webb, 2021).

A Terra descreve uma rotação diária sobre o seu próprio eixo. Este eixo polar, forma um ângulo de $23,4^\circ$ com o plano orbital da Terra (que, de uma perspetiva terrestre, é o plano que o Sol descreve em torno da Terra). Este ângulo é designado de obliquidade. A obliquidade é o motivo pelo qual a Terra tem estações. Nos últimos milhões de anos, o valor da obliquidade, variou entre $22,1^\circ$ e $24,5^\circ$. Quanto maior for a obliquidade, mais extremas serão as estações do ano, com verões mais quentes e invernos mais frios. Durante o verão, ângulos de inclinação

maiores favorecem períodos de degelo. A obliquidade tem vindo a diminuir muito lentamente num ciclo que se estende por um período de 41 mil anos (Webb, 2021). À medida que a obliquidade diminui, as estações do ano vão ficando gradualmente mais amenas, resultando em invernos cada vez mais quentes e verões mais frios que, ao longo do tempo, permitem a formação de grandes camadas de neve e gelo nas latitudes elevadas. À medida que se dá o aumento das camadas de gelo, haverá aumento no albedo², potenciando ainda mais o arrefecimento global.

O eixo sobre o qual a Terra descreve uma rotação diária aponta numa determinada direção que varia numa escala de tempo de 20 mil anos. Esta variação, conhecida como precessão, significa que embora o Polo Norte esteja atualmente apontado para a estrela Polar, dentro de 10 mil anos apontará para a estrela Vega (NASA, 2020).

Os ciclos de Milankovitch desencadeiam pequenas alterações que a longo prazo levam a mudanças significativas do clima na Terra. Para Milankovitch, pequenas variações da radiação em determinadas latitudes e estações do ano têm grande influência na fusão e formação de camadas de gelo. Além disso, considera que obliquidade é o ciclo com maior influência no clima, por afetar com maior preponderância as latitudes elevadas, precisamente onde ocorre a generalidade da formação e fusão de gelo (NASA, 2020).

Atualmente, é amplamente aceite que os ciclos de Milankovitch foram um dos principais mecanismos naturais forçadores das variações climáticas, talvez mesmo o principal (NASA, 2020). Crowley e Kim (1994) mostraram que, muito provavelmente, o NMM entre 130 mil e 132 mil anos atrás está, diretamente, relacionado com os ciclos de Milankovitch, assim como noutros momentos do Quaternário. No entanto, ainda persistem dúvidas se os ciclos de Milankovitch podem ser considerados como fatores dominantes na evolução do clima durante a generalidade do período Quaternário. Alguns autores, tal como Grotzinger (1986), admitem que ainda não existem evidências suficientes que comprovem esta teoria.

Segundo Berger (1988), os ciclos de Milankovitch não explicam todas as alterações climáticas durante o período Quaternário. Em particular, não podem explicar as variações climáticas durante o período Holoceno. Os ciclos de Milankovitch fazem-se sentir em períodos de longa duração, variando de dezenas de milhares a centenas de milhares de anos. Em contraste, o aquecimento atual da Terra ocorreu

²Refletividade da superfície de um planeta; poderá ser interpretado como sendo o rácio entre a quantidade de luz refletida pela superfície e a quantidade de luz incidente no planeta.

numa escala de tempo muito mais curta. Vários estudos identificam as ações antrópicas como sendo as principais causas do rápido aquecimento das últimas décadas (Jevrejeva et al., 2010, Cazenave e Remy, 2011, Cazenave e Nerem, 2004).

1.2.2 Fenómenos Eustáticos

As mudanças eustáticas do NMM que acontecem à escala global, são provocadas por alterações na morfologia e volume das bacias oceânicas ou mesmo pela quantidade de água nelas contida. Entre os processos com maior relevância, podem referir-se os glácio-eustáticos, que ocorrem mediante as variações de volume das águas oceânicas em consequência dos processos cíclicos de fusão ou crescimento de calotes polares e de glaciares (Rovere et al., 2016); (Da Silva et al., 2014). Além disso, a expansão ou contração térmica também têm o seu contributo durante as variações cíclicas do clima, havendo ligeiras reduções do volume em períodos glaciares e ocorrendo o oposto em períodos interglaciares. As variações do volume de água no oceano sofrem ainda interferência dos processos hidro-eustáticos, através da redistribuição de água entre diferentes reservatórios (humidade do solo, águas subterrâneas) (Rovere et al., 2016).

Processos tectónicos-eustáticos, que estão associados a deformações tectónicas que modificam a capacidade das bacias oceânicas, também originam variações no NMM. A expansão e contração dos fundos oceânicos estão associadas à elevação e afundamento de cadeias meso-oceânicas, sendo este fenômenos geológicos relativamente rápidos (Vieira, 1981).

Além deste processo, a capacidade das bacias oceânicas também é modificada devido à acumulação de sedimentos. Este processo designa-se por sedimentoeustático. A acumulação de sedimento faz com que, numa primeira fase, se dê uma elevação do fundo marinho e, portanto, elevação do NMM. No entanto, ao longo do tempo, inicia-se a compactação dos sedimentos, o que modifica novamente o volume da bacia oceânica.

Por fim, mas não menos importante, os processos geóide-eustáticos também influenciam o NMM. Este último processo refere-se às variações do nível de água devidas a flutuações no relevo do geóide, sendo influenciado por diversos fatores tais como a pressão atmosférica, densidade da água do mar e temperatura, batimetria e anomalias gravíticas (Dias e Taborda, 1988).

Todos estes processos são muito complexos e difíceis de quantificar. Por exemplo, os processos glácio-eustáticos, fortemente dependentes da temperatura e da

velocidade de fusão ou formação do gelo, são muito lentos e, como tal, a quantificação das suas variações é extremamente complicada. Por este motivo, também é difícil determinar fenómenos isostáticos visto que estes derivam dos fenómenos eustáticos.

Segundo o Painel Internacional para as Alterações Climáticas da Organização das Nações Unidas (Griggs e Noguer, 2002), com o recente aumento das temperaturas atmosféricas, como consequência das atividades humanas, é lícito esperar um agravamento da atual subida do NMM motivada por processos eustáticos. Contudo, esta ideia ainda não está claramente provada cientificamente, muito por causa da sua ocorrência ser lenta relativamente à escala humana. No entanto, estes processos também têm algum significado em escalas de tempo mais curtas, embora tenham uma magnitude mais baixa. Atualmente, estima-se que haja subidas de cerca de 1,5 mm/ano do NMM em consequência de processos glácio-eustáticos (Vieira, 1981).

1.2.3 Fenómenos Isostáticos

Como já foi referido, a problemática referente às variações do NMM envolve inúmeros processos difíceis de quantificar, sendo amplificado pela interligação entre eles. Por exemplo, o aquecimento global provoca a fusão generalizada do gelo das calotas polares e glaciares, provocando a redução do albedo da Terra, levando a uma situação em que a Terra passa a absorver ainda mais radiação solar e, consequentemente, o aquecimento global é amplificado. Assim como neste exemplo, também os processos isostáticos acontecem em consequência das variações eustáticas do NMM.

Os fenómenos isostáticos são compostos por dois grandes processos, designados por glácio-isostáticos e hidro-isostáticos. Os processos glácio-isostáticos ocorrem devido a ajustamentos litosféricos causados por variação das massas glaciárias, ou seja, sempre que um território passa a estar coberto de gelo, terá lugar a uma ligeira subsidência, pelo contrário, quando se dá a fusão do gelo, o território emerge. Este processo é verificado, por exemplo, nos países escandinavos, que emergem por ação da composição isostática pós-glacial (Da Silva et al., 2014). Sempre que existe uma variação do NMM, ocorre ou a inundação de áreas terrestres ou a sua imersão, ou aumento/diminuição do peso da coluna de água no solo marinho. Como resultado destas alterações, iniciam-se também ajustamentos litosféricos devido às variações de espessura da coluna de água e, consequentemente, da força que esta exerce no leito marinho. Este último processo é designado de hidro-isostático (Dias e Taborda, 1988).

Apesar da interligação existente, os fenómenos isostáticos desencadeiam variações do NMM independentes, em termos de valores numéricos, das variações

eustáticas. Em várias regiões do planeta, as variações locais do NMM diferem das variações globais. Por exemplo, a elevação de um determinado local (ou de uma determinada zona) pode ser interpretada como uma diminuição do NMM global, se esta subida for superior às subidas associadas aos fenómenos eustáticos. Desta forma, considera-se que os fenómenos isostáticos estão associados a variações do nível médio relativo, ocorrendo a nível regional/local. Estes fenómenos ainda não estão totalmente explicados, mas deverão estar relacionados com flutuações na quantidade de calor armazenado na circulação e transporte de calor (COI, 1985) .

O conhecimento dos fenómenos isostáticos é fundamental tanto para estudos locais como globais sobre o NMM, e não devem ser esquecidos pois podem resultar em conclusões erradas.

1.3 Causas da Atual Subida do Nível Médio do Mar

No último século, verificou-se um aumento da tendência de subida do NMM, com particular incidência nas últimas décadas (Picoto et al., 2022a). Esta subida pode ser atribuída em grande parte à expansão térmica dos oceanos, com maior significado nas camadas superiores, à fusão dos glaciares alpinos e calotes de gelo polares e ainda às mudanças no ciclo hidrológico da água, além de outros fatores menos significativos (Cazenave e Remy, 2011; Church et al., 2013).

Os dois primeiros fatores aludidos podem ser entendidos como uma resposta da atmosfera terrestre ao aquecimento global. O aquecimento da atmosfera e a subida do NMM podem ser, em parte, relacionados com o final da última pequena oscilação climática, que se iniciou no século XVI e se prolongou até meados dos século XIX, caracterizada por temperaturas ligeiramente mais frias e NMM estável ou com ligeira redução, este período é comumente designado por “Pequena Idade do Gelo” (Mann, 2002).

Não obstante esta relação, a recente tendência de subida do NMM não deixa de ter uma forte relação com as atividades humanas, em especial, ao nível atmosférico, devido ao aumento da quantidade de gases que potencializam o chamado “efeito de estufa”. Por exemplo, entre 1958 e o final da década de setenta, verificou-se um aumento de 8% na concentração de CO_2 na atmosfera, entre 1970 e 1980 o metano teve um acréscimo de 1% a 2% (Dias e Taborda, 1988). O aumento da temperatura em consequência do incremento da concentração de gases como o CO_2 , CH_4 , NO_2 ,

potenciam a evaporação das massas de água. O vapor de água contribui decisivamente para o “efeito de estufa”, potenciando ainda mais o aumento generalizado da temperatura atmosférica.

Embora ainda existam questões em aberto sobre o aquecimento global e os respetivos mecanismos autorreguladores da Terra, há consenso que o aumento da temperatura média é uma consequência inevitável das ações antrópicas. Com efeito, desde os primórdios da revolução industrial, em meados do século XIX, a temperatura média da Terra subiu cerca de 1°C, valor esse que descreve uma subida de, aproximadamente, 0,2°C por década (NASA, s.d.).

Existem ainda incertezas no que concerne à relação entre o aquecimento global e o aumento do NMM devido à expansão térmica e ao aumento da massa de água presente nos oceanos, como resultado da fusão de gelo. Contudo, as medições de gravidade terrestre realizadas por satélites a partir de meados da década de noventa vieram, em grande parte, facilitar a compreensão das variações de gelo nos glaciares e calotes polares. Cazenave e Remy (2011) mostraram que, a fusão de glaciares e das calotes polares contribuíram, cada um, em 30% da subida do NMM entre 1993 e 2010.

A ambiguidade persiste na expansão térmica devido ao desfasamento temporal na resposta dos oceanos às alterações climáticas. Segundo Dias e Taborda (1988), as oscilações do NMM em consequência da variação da temperatura global dos oceanos acontecem com um desfasamento de 18 anos, com uma elevação de 0,16 m por cada grau. Apesar destas incertezas, a maioria dos autores referem que estes dois fatores são responsáveis por grande parte da atual tendência de subida. Church et al. (2013) mostraram que, entre 1971 e 2010, as contribuições da expansão térmica e da transferência de grandes quantidades de água das camadas de gelo, excluindo as da Antártida, para os oceanos explicam cerca de 75% da atual tendência de subida do NMM.

As atividades humanas introduziram alterações não só a nível atmosférico, mas também no ciclo hidrológico. Cazenave e Nerem (2004) descrevem as principais ações antrópicas que induzem alterações no balanço hidrológico da água. Para estes autores, processos como a captação de grandes massas de água em albufeiras através de barragens, irrigação, urbanização, extração de água subterrânea e desflorestação, provocam a alteração da distribuição da água a nível global e, como tal, provocam alterações no NMM. Destes processos, a grande maioria favorece o aumento do NMM, no entanto, a grande quantidade de água retida nas albufeiras bem como a água que é utilizada na irrigação provocam um efeito contrário, de facto, estes dois

fatores provocam uma descida generalizada de, respetivamente, $-1\pm0,2$ mm/ano e $-0,56\pm0,1$ mm/ano (Cazenave e Nerem, 2004).

Por outro lado, a urbanização apresenta um efeito positivo na subida do NMM. Devido à ocupação de zonas de vegetação por superfícies impermeáveis dá-se o aumento do escoamento superficial, redução da infiltração no solo, e do caudal dos lençóis freáticos. Os dados obtidos apontam para uma subida com um valor associado de $0,34\pm0,04$ mm/ano como resultado da urbanização. Entre os fatores com maiores contribuições no aumento do NMM, além do mencionado anteriormente, está a desflorestação e a extração de água subterrânea, contribuindo respetivamente em $0,09$ mm/ano e $0,2\pm0,1$ mm/ano (Cazenave e Nerem, 2004).

Todas as alterações antrópicas no ciclo hidrológico da água refletem-se numa ligeira redução do NMM. Cazenave e Nerem (2004) mostraram que os distúrbios do ciclo hidrológico contribuem para uma redução de $-0,91\pm0,45$ mm/ano. Contudo, este valor está associado a uma elevada incerteza, aproximadamente 50%, devido à falta de evidências. Estas alterações não são, no entanto, suficientes para contrariar o aumento do NMM provocado pelo aquecimento global.

Em termos regionais, algumas atividades humanas podem também resultar em fenómenos de subsidência, intensificando o aumento do nível médio local. Este fenómeno resulta de atividades como drenagem (para fins de desenvolvimento urbano), a extração de recursos tais como água subterrânea, petróleo ou gás e a realização de obras ao longo do litoral. A título exemplificativo, a extração de recursos no subsolo nas cidades costeiras da Indonésia resulta em tendências de subsidência que em algumas regiões pode chegar a 22 cm/ano; na Itália, a extração de água subterrânea entre 1950 e 1970, resultou na subsidência de até 3 metros em parte do delta do rio Pó (Rovere et al., 2016).

1.4 Variações Sazonais

A variação intraanual do NMM apresenta um padrão de variabilidade, designado por variação sazonal. As variações sazonais são frequentemente ignoradas quando se estudam tendências seculares, sendo consideradas como uma fonte de ruído. Apesar de descartadas, as variações sazonais alteram a variância dos dados e aumentam a incerteza das estimativas das tendências de subida do NMM, (Antunes, 2011).

Portanto, o estudo das variações sazonais é muito relevante para a compreensão e aumento da nossa capacidade de previsão tanto da variabilidade das grandes

correntes oceânicas como das interações oceano/atmosfera. As variações sazonais quando acompanhadas por períodos de tempestades extremas, elevam os riscos de erosão e inundações costeiras e, como tal, a análise destas é de extrema relevância.

Na costa portuguesa, as variações apresentam valores mínimos no período de maio a julho, apresentando depois valores mais altos entre outubro a dezembro, comportamento este que é normal para o Hemisfério Norte (Antunes, 2016). Entre as causas mais prováveis para este tipo de comportamento encontram-se a temperatura da água do mar, a pressão atmosférica, o vento, fatores astronómicos, e dependendo do local em questão, as descargas fluviais (Taborda e Dias, 1988). De acordo com Taborda e Dias (1988), cada um destes fatores influencia o NMM da seguinte forma:

1. Temperatura da água do mar: a variação da temperatura do oceano provoca a expansão térmica da água. No meio do Oceano Pacífico, por exemplo, a expansão térmica é responsável por variações que podem chegar a 1 metro.

2. Pressão atmosférica: as variações da pressão atmosférica causam oscilações na pressão exercida sobre a superfície do mar. O oceano atua como se fosse um barómetro invertido, ou seja, o nível do mar sofre uma redução de 1 cm por cada acréscimo de um milibar na pressão atmosférica.

3. Vento: o vento é um fator de relevo nas variações sazonais. No caso particular da costa portuguesa, entre os meses de maio e julho, os ventos são predominantemente do quadrante de Noroeste, originando o transporte da camada superficial para o largo, verificando-se afloramento costeiro, denominado por “*upwelling*”. Durante o inverno, os ventos fazem-se sentir dos quadrantes de Sul/Sudeste, com elevada intensidade durante as tempestades, provocando rápidas subidas do nível do mar.

4. Descarga fluvial: em estuários ou regiões adjacentes, os regimes de descargas fluviais são de extrema importância, podendo provocar oscilações com uma magnitude que pode ser superior a um metro. Em algumas regiões as variações provocadas por este fator podem ser decisivas nas flutuações anuais do NMM. Na costa atlântica dos Estados Unidos da América, por exemplo, entre 7 a 21% das variações ao longo do ano podem ser atribuídas ao regime de descargas fluviais.

Além dos fatores apresentados, também há a considerar fatores astronómicos, por exemplo, na costa portuguesa, as marés vivas podem ultrapassar os 3 metros, valor superior à média da maré astronómica, aproximadamente de 2,1 metros (Taborda e Dias, 1988).

1.5 Métodos de Observação do Nível Médio do Mar

As mudanças no NMM podem ser observadas em diferentes escalas de tempo com recurso a diferentes técnicas. Independentemente da técnica usada, nenhuma observação consegue inferir apenas variações eustáticas do nível médio. Atualmente, as medições do NMM são realizadas a partir de séries de dados obtidos, tanto por satélites como *in situ* através de marégrafos. Antes da década de noventa apenas era possível utilizar dados maregráficos para inferir o NMM e estudar as respetivas variações (Cazenave e Nerem, 2004). Em escalas de tempo superiores, desde algumas centenas até milhões de anos, a reconstituição do NMM e das suas variações é realizada a partir de diferentes indicadores, os principais são: geológicos, geomorfológicos, biológicos, arqueológicos.

1.5.1 Marégrafos e Registos Maregráficos

O estudo de séries de registos maregráficos constitui-se como o método mais comum para avaliar alterações do NMM em escalas de tempo de décadas ou eventualmente séculos (apenas possível na totalidade no século XX). O primeiro marégrafo de registo contínuo foi instalado no rio Tamisa em 1830 (Bradshaw et al., 2016). Nos anos seguintes, vários marégrafos foram instalados em portos europeus, nos Estados Unidos da América, e na Austrália para assegurar a segurança da navegação. Na década de oitenta do século XIX, já se encontravam operacionais cerca de 70 marégrafos no continente europeu, um deles foi o marégrafo de Cascais, o qual segundo Boski (2013) era, provavelmente, um dos menos afetados por processos de subsidência.

Os marégrafos são instrumentos que permitem medir as variações do NMM em relação a uma marca de nivelamento. Ao longo do tempo, o estudo de séries maregráficas tornou-se a principal ferramenta para o estudo das flutuações do NMM, de curta duração, provocadas por fenómenos atmosféricos, por exemplo, e tendências de mudanças ao longo de escalas temporais de séculos/décadas (Boski, 2013). A maioria dos marégrafos modernos já se encontram equipados com estações GPS, por forma a registar fenómenos de elevação ou subsidência provocados por processos isostáticos. Estes movimentos verticais da crosta terrestre podem influenciar decisivamente os registos do NMM e, portanto, a análise das séries maregráficas envolve um cuidado especial no processamento destes.

A utilização de registos maregráficos, em estudos do NMM, apresenta três grandes desvantagens: em primeiro lugar, podem ser influenciados por fenómenos regionais, tais como fenómenos isostáticos, em segundo, as estações maregráficas encontram-se desigualmente distribuídas pelo globo, com registos suficientemente longos para poderem ser utilizados em estudos sobre as variações do NMM localizados predominantemente no Hemisfério Norte (Figura 1.4). Esta distribuição espacial assimétrica dificulta a caracterização das tendências do NMM à escala global. Por último, os registos maregráficos são, por vezes, caracterizados pela descontinuidade dos dados. Apesar destes aspetos, os registos maregráficos são essenciais para inferir as variações do NMM nas zonas costeiras. Além disso, são a única ferramenta disponível para realizar reconstruções do NMM desde finais do século XIX até ao aparecimento das medições altimétricas realizadas por satélites.

A maioria dos estudos atuais sobre o NMM a nível global utilizam séries de dados maregráficos provenientes do *Permanent Service for Mean Sea Level* (*PSMSL*), sendo este o maior repositório de dados maregráficos. Disponibiliza dados desde o final do século XIX, provenientes de cerca de 2000 marégrafos, permitindo realizar estudos das alterações no NMM a longo prazo. A maioria das séries de registos disponíveis na *PSMSL* são referenciadas a um *datum* comum, designado por *Revised Local Reference* (*RLR*). O *datum RLR* foi definido de maneira a estar, aproximadamente, 7000 mm abaixo do NMM, com esta escolha, feita há muitos anos, para evitar valores médios mensais e anuais negativos (*PSMSL*, s.d.).

Segundo Cazenave e Remy (2011), muitos dos registos disponíveis na *PSMSL* não apresentam homogeneidade entre si em termos de qualidade e distribuição temporal. Para estudos a longo prazo da evolução de NMM, apenas 10% deste conjunto de dados podem ser utilizados, podendo apesar disso permitir inferências quanto às tendências seculares.

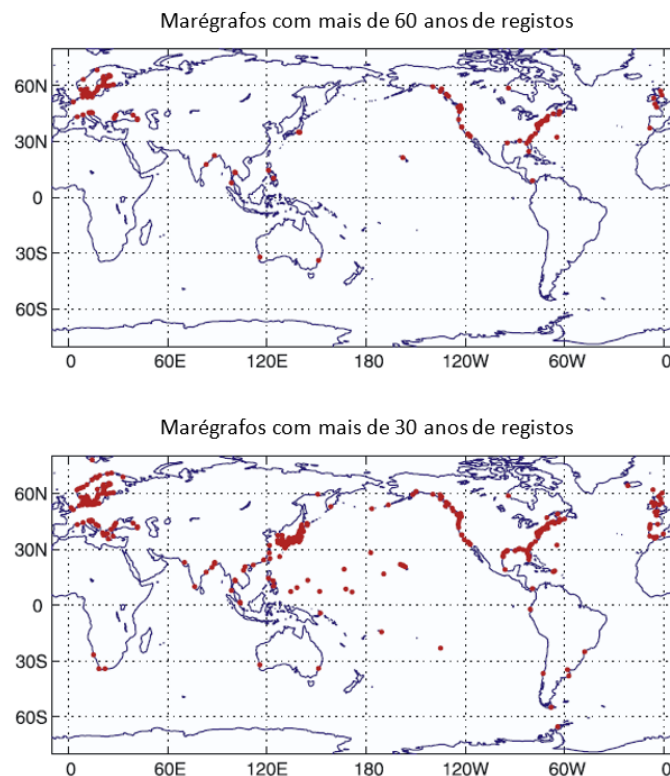


FIGURA 1.4: Distribuição de registros maregráficos da base de dados da *PSMSL*. Marégrafos com séries de dados maregráficos superiores a 60 anos (gráfico superior), marégrafos com séries de dados maregráficos inferiores a 30 anos (gráfico inferior) (adaptado de Cazenave e Nerem, 2004)

1.5.2 Altimetria por Satélite

As medições da elevação da superfície oceânica realizadas com recursos a levantamentos altimétricos a partir do espaço tiveram o seu início em 1978, quando o satélite norte-americano SEASAT entrou em operação. Na década de oitenta, os dados recolhidos dos levantamentos topográficos oceânicos contavam já com uma resolução vertical de 5 cm. Em 1992, com o lançamento do satélite franco-americano Topex/Poseidon (T/P), vocacionado para a investigação de questões oceanográficas e climáticas, iniciou-se um novo paradigma, em grande parte devido aos avanços tecnológicos. Este satélite, que esteve operacional até 2006, permitia a determinação da altura da superfície oceânica em relação ao elipsoide de referência da órbita do satélite. Os avanços na precisão foram significativos, com este satélite a realizar medições com uma precisão de 2-3 cm (Cazenave e Nerem, 2004). A entrada em operação de novos satélites, nomeadamente do Jason 1 e Jason 2, possibilitou o aumento da precisão altimétrica (Figura 1.5). Além disso, assim como o T/P,

também estes dois satélites possibilitam uma cobertura orbital quase completa com uma periodicidade de 10 dias (Boski, 2013).

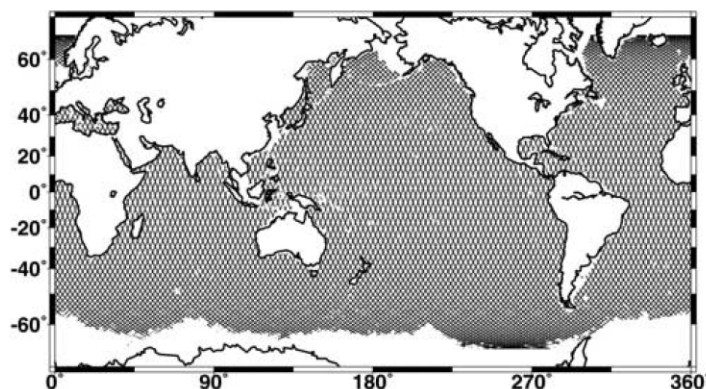


FIGURA 1.5: Distribuição dos dados altimétricos recolhidos ao longo de um ciclo de 10 dias pelo satélite Topex/Poseidon (extraído de Cazenave e Nerem, 2004)

A medição altimétrica da superfície oceânica é realizada com recurso a ondas eletromagnéticas. A altitude do satélite relativamente à superfície instantânea do mar é obtida a partir do cálculo do tempo que o sinal de frequência na banda das micro-ondas demora a percorrer o percurso emissor/recetor, após ser refletido pela superfície do mar. Uma vez que a altura de órbita do satélite é conhecida, pode-se então determinar a altura da superfície do mar em relação ao elipsoide de referência da órbita do satélite. Por fim, são aplicadas correções à medição que visam eliminar problemas como a variação da velocidade, estado do mar, entre outros (Dalazoana et al., 2005).

Ao longo das últimas três décadas, apesar da complexidade do processo de compilação dos dados das 3 missões realizadas, podemos afirmar o seguinte sobre as variações interanuais do NMM registadas pelo processo de medição altimétrico, (Boski, 2013):

- São uma consequência de eventos climáticos, nomeadamente do fenómeno no Pacífico Sul conhecido por El Niño (Hamlington et al., 2015, Wang et al., 2018);
- A tendência de subida do NMM medida com recurso à altimetria por satélite situa-se no intervalo 3,0-3,2 mm/ano, muito superior à determinada a partir dos registos maregráficos (1,5-2,1 mm/ano (Church e White, 2011)). Cazenave e Nerem (2004) refere que grande parte desta subida tem origem na expansão térmica dos oceanos;

- As mudanças do NMM não são uniformes, com regiões a registarem tendências de subida na ordem de +10 mm/ano enquanto noutras regiões se verificam de -4 a -8 mm/ano;

A utilização de satélites permite ainda a medição de gravidade. As mudanças de massa causada pela fusão de gelo das calotas polares e dos glaciares podem ser detetadas pelas medições de gravidade. Deste modo, as observações realizadas pelos satélites já referidos ganharam ainda mais relevância, sob a forma de mapas de gravidade. Desta forma, é possível quantificar a subida eustática que se deve exclusivamente às mudanças de massa resultantes da fusão das calotas polares e glaciares de montanha.

As medições altimétricas realizadas por satélites vieram possibilitar o mapeamento da distribuição geográfica da mudança do NMM. Assim como a maioria dos estudos das séries maregráficas, que sugerem uma subida do NMM não uniforme ao longo do globo, também as medições altimétricas sugerem uma grande variabilidade regional do NMM. A subida do NMM no Pacífico Ocidental e Índico Oriental atinge a maior magnitude, com valores várias vezes superiores à média global. No Atlântico verifica-se uma subida generalizada, enquanto em regiões como Pacífico Oriental e Índico Ocidental se tem verificado uma descida do NMM (Cazenave e Nerem, 2004) (Figura 1.6). Os registos altimétricos parecem indicar que o aumento do NMM nas últimas décadas está a ser superior ao aumento registado pelas séries de dados maregráficos ao longo de todo o século XX.

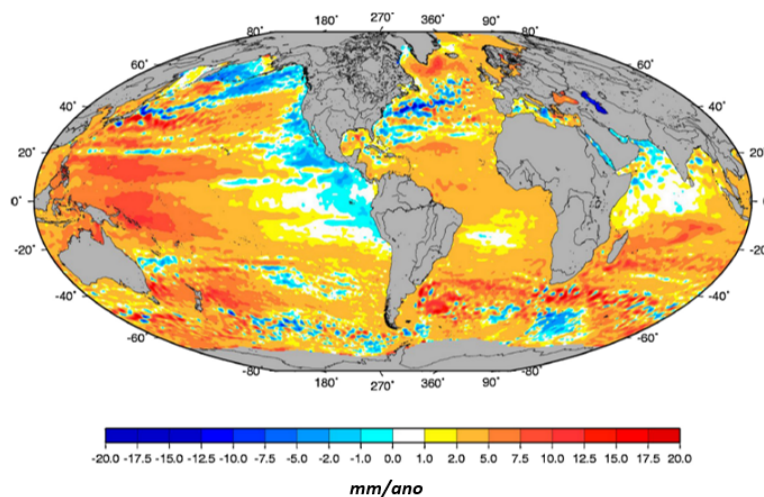


FIGURA 1.6: Distribuição geográfica das tendências do NMM entre 1992-2008, obtidas através de altimetria pelos satélites T/P e Jason 1 (extraído de Cazenave e Nerem, 2004)

Dada a relevância das medições altimétricas na monitorização da subida do NMM, iniciou-se em 2020, a missão Sentinel-6, que vem dar continuidade ao legado das missões T/P e Jason pelo menos até 2030. A missão Sentinel-6, composta atualmente pelo satélite Sentinel-6 Michael Freilich (lançado em novembro de 2020) tem como objetivo específico o estudo dos oceanos. O satélite é equipado com um radar altímetro orientado para medir, quase na sua globalidade e com elevada precisão, a altura da superfície do oceano e, desta forma, continuar o trabalho já desenvolvido pelos seus antecessores na monitorização do NMM (ESA, s.d.).

1.5.3 Outros Métodos de Reconstrução do Nível Médio do Mar

Os atuais registos instrumentais das mudanças do NMM, séries de dados maregráficos e medições altimétricas, são muito precisos. No entanto, abrangem um intervalo de tempo muito curto relativamente à escala temporal das variações cíclicas do nível de água.

Ao longo dos últimos 20 milénios, a Terra presenciou a transição de um período glaciário para um período interglaciário. Esta transição teve como consequência a transferência de grandes massas de água que se encontravam retidas sob a forma de gelo sobre as calotas polares e glaciares nas massas continentais para os oceanos. Este processo teve como consequência uma subida do NMM de cerca de 120 metros (Boski, 2013). Assim como neste pequeno intervalo de tempo, ao longo de todo o período geológico da Terra sucederam-se grandes oscilações do NMM. A questão é, como foi possível inferir esta realidade?

Existem várias evidências que provam que o NMM já esteve muito abaixo do atual. É possível inferir esta subida através de evidências como deltas e prolongamentos de sistemas de vales fluviais para o largo, através da plataforma continental. A subida do NMM também é confirmada pelo estudo de planos intertidais, sapais, plataformas de abrasão marinha, sistemas de grutas em arribas, entre outros (COI, 1985). As reconstruções do NMM são conseguidas através da análise de indicadores. Podemos definir um indicador como sendo qualquer evidência cuja relação entre a localização onde é encontrada e a sua datação permitem perceber onde se situava o NMM num determinado período. Estes indicadores podem ser divididos nas seguintes categorias: arqueológicas, geológicas e geomorfológicas, biológicas.

No passado, tal como na atualidade, as sociedades procuraram estabelecer-se junto às zonas costeiras. Ao longo do período Holoceno, com a subida do NMM,

em consequência da transição do período glaciário para um período interglaciário, deu-se a inundação de áreas que dantes continham povoaamentos prósperos, deixando imersas várias evidências que se conservaram ao longo do tempo, tais como cerâmica, tanques e portos. No Mediterrâneo, onde as amplitudes da maré são pequenas, as variações do NMM podem ser, em muitos casos, avaliadas a partir de levantamentos arqueológicos de portos antigos. O Mediterrâneo é muito abundante em evidências arqueológicas graças às várias culturas clássicas que prosperaram nas suas costas, possibilitando reconstruções muito satisfatórias do nível do mar ao longo dos últimos 2,5 mil anos (Rovere et al., 2016). Os sambaquis ou concheiros, locais onde eram depositadas essencialmente grandes quantidades de conchas e restos de animais, constituem-se também como excelentes auxiliares na constatação de antigos níveis marinhos quaternários, devido à sua localização costeira.

Existem também outros indicadores de cariz geomorfológico e geológico, como por exemplo: existência de lagunas isoladas na parte emersa da plataforma continental, níveis de erosão marinhos, paleo-lagunas ou vales fluviais ou existência de antigas linhas de costas submersas, são também bons indicadores, permitindo realizar reconstruções relativamente precisas devido, à natural proximidade espacial entre esses indicadores e o NMM passado (Vieira, 1981).

As evidências biológicas ou orgânicas são apontadas, na maioria dos estudos, como sendo a forma mais credível para reconstruir o NMM, sendo por isso as mais utilizadas. A partir deste tipo de evidências é possível reconstruir o NMM com recurso à datação de fósseis de corais ou outros seres vivos, cuja profundidade do habitat é conhecida. Por exemplo, Vieira (1981) refere que os recifes de corais não se formam em profundidades muito superiores a cerca de cinco metros da linha de água. Como tal é possível estabelecer uma relação clara entre a idade, determinada por processos de datação, e a posição vertical das espécies que permitem indicar o NMM (Boski, 2013).

A reconstrução do nível do mar através dos processos aludidos está naturalmente sujeita a várias incertezas. Os intervalos associados a essas incertezas não são uniformes, podendo ser da ordem de alguns centímetros até cerca de dois mil anos atrás, e aumentando progressivamente até vários metros para o Paleoceno (Rovere et al., 2016).

1.6 A Evolução do Nível Médio do Mar

Conforme já ficou patente na secção anterior, o NMM esteve sempre em constante mudança. Ao longo do período Quaternário, a Terra passou por vários episódios de expansão e contração massiva das camadas de gelo. À medida que o clima da Terra alternou entre estes dois períodos, o NMM sofreu grandes alterações. Durante os períodos glaciários, o nível do mar desceu em consequência da grande quantidade de água aprisionada nas calotes polares e glaciares. Em períodos interglaciários, à medida que o volume de gelo retido sobre os continentes diminuiu, o nível médio global dos oceanos aumentou. Os avanços e recuos fizeram-se sentir com particular incidência nos últimos dois milhões de anos. Durante este período, existem evidências da existência de até dezassete ciclos glaciação-interglaciação. Contudo na história geológica total estes períodos de grandes oscilações são anormais, sendo que apenas foi possível identificar mais quatro oscilações nos últimos 900 milhões de anos (COI, 1985).

A reconstrução do NMM antes do período Holoceno é muito complicada essencialmente porque, na maioria das vezes, existem apenas referências de valores mínimos e máximos ocorridos ao longo dos ciclos de glaciação-interglaciação. Além disso, a compreensão dos fenómenos que provocaram a alternância entre períodos glaciários e interglaciários suscita muitas dúvidas, essencialmente porque a natureza dos fenómenos é diferente dos últimos milhares de anos.

Por este motivo, o foco das seguintes subsecções será a evolução do NMM ao longo do atual período, o Holoceno, onde existem evidências que possibilitam a sua reconstrução. Além disso, a natureza dos fenómenos que provocaram as oscilações do NMM permanece relativamente constante ao longo de todo o Holoceno, apesar de haver grande aceitação que no período pós-Revolução Industrial as ações antrópicas tenham sido o principal fator de variação do clima e consequentemente do NMM.

1.6.1 Período Holoceno

Na escala de tempo geológico, o Holoceno é a época atual do período quaternário. O Holoceno iniciou-se há cerca de 11,7 mil anos, durante a transição de um período mais frio, para um período geralmente mais quente, com oscilações térmicas relativamente estáveis, permitindo um rápido crescimento e desenvolvimento da sociedade (Wanner et al., 2011).

A reconstrução do NMM durante o período Holoceno é muito complicada devido à ausência de evidências. No estudo realizado por Smith et al. (2011), referente à evolução do NMM na costa atlântica dos Estados Unidos da América, 93% de todos os dados disponíveis, no período do Holoceno, são referentes aos últimos 6 mil anos. A existência de apenas 7% de dados desde os primórdios até meados do Holoceno apresenta-se como um enorme desafio na determinação do NMM durante este período. Assim como este estudo, a maioria dos estudos ficam limitados a reconstruções do NMM a partir de poucos dados e que, por norma, são muito dispersos globalmente. Além disso, em várias regiões do globo, não existem dados que permitam uma reconstrução contínua, apresentando em muitos casos, grande dispersão temporal entre dados consecutivos que, por vezes, podem ser da ordem de várias centenas de anos ou até mesmo de milénios, dificultando muito a reconstrução fiável do NMM e em muitos casos, impedindo a reconstrução em grande parte do período Holoceno.

As reconstruções do NMM na costa atlântica dos Estados Unidos da América, apresentam tendências de subida do NMM ao longo de todo o período Holoceno (Engelhart e Horton, 2012). No entanto, supõe-se que na primeira metade deste período tenha havido uma subida mais significativa. O número reduzido de dados disponíveis entre 9 e 6 mil a.C. impossibilita uma reconstrução contínua da evolução do NMM neste período. No entanto, estima-se que tenha subido pelo menos 26 metros (Engelhart e Horton, 2012).

Posteriormente, entre 6 mil e 2 mil a.C., o NMM continuou a subir, embora com uma tendência menos acentuada do que durante o período anterior. Este período é marcado por elevada incerteza relativamente à tendência de subida do nível médio devido à escassez de dados, com níveis de incerteza a variar entre $\pm 0,4$ a $\pm 2,0$ mm/ano. A tendência de subida do NMM entre este período varia entre 0,5 e 4,5 mm/ano (Engelhart e Horton, 2012). Em comparação, entre 2 mil a.C. e 1900 d.C. a tendência foi significativamente inferior, a variar entre 0,6 e 1,8 mm/ano. Para Engelhart e Horton (2012), a redução da tendência de subida foi uma consequência da diminuição da influência de processos isostáticos e do volume de água que chega aos oceanos através da fusão de gelo. Este evento resultou por um lado do desaparecimento da camada de gelo Laurentide em 5 mil a.C., que cobria parte da América do Norte, e por outro lado da redução da camada de gelo sobre a Gronelândia, que atingiu um mínimo histórico no Holoceno por volta de 6 mil a.C.

Garcia-Artola et al. (2018) determinaram tendências de subida do NMM mais elevadas durante o período inicial do Holoceno, a variar entre $6,3 \pm 0,8$ mm/ano

e $6,8 \pm 0,5 \text{ mm/ano}$, na orla costeira localizada entre a Flandres Francesa e o Algarve, ocorrendo uma diminuição notável da tendência por volta de 5 mil a.C. A diminuição da tendência de subida é progressiva ao longo do restante período do Holoceno. No Holoceno médio, entre 5 mil e 2 mil a.C., a tendência variou entre $1,8 \pm 0,5 \text{ mm/ano}$ e $0,5 \pm 0,4 \text{ mm/ano}$, acentuando-se a redução no Holoceno tardio, de 2 mil a.C. até ao presente, para valores entre $0,9 \pm 0,4 \text{ mm/ano}$ e $0,1 \pm 0,5 \text{ mm/ano}$. Assim como na costa atlântica da América do Norte, também se assume que a tendência de subida do NMM tenha diminuído ao longo do Holoceno em consequência da redução do degelo e da diminuição da influência de processos isostáticos.

Smith et al. (2011), apresentaram uma reconstrução do NMM no Taiti, Oceano Pacífico, que segundo estes, é indicativa das tendências do NMM em grande parte do globo. Smith et al. (2011) estima que nos primeiros milénios do Holoceno, entre 9,6 a.C. e 7,9 a.C., o NMM tenha subido entre 50 e 60 metros, representando uma tendência de subida na ordem de 12-13 mm/ano. Estes valores são claramente superiores aos apresentados por Engelhart e Horton (2012), que se fixaram na ordem dos 26 metros entre 9 e 6 mil a.C. Já no período posterior a 5 a.C., o NMM subiu entre 3 e 5 metros, representando uma tendência de subida de 0,4 a 0,7 mm/ano. De acordo com Engelhart e Horton (2012), o estudo do período entre o início e meados do Holoceno pode ser crucial para compreender a evolução de futuras subidas do NMM devido às semelhanças com as tendências atuais.

Os estudos apresentam reconstruções significativamente diferentes, no entanto, ambos apontam para um aumento do NMM ao longo de todo o Holoceno com valores muito elevados na fase inicial, em particular no Taiti. Apresentam também uma desaceleração da tendência de subida por volta de 6 mil a.C., que se acentuou até ao final do século XIX.

1.6.2 Os Últimos Dois Mil Anos

Com base na análise de sedimentos de sapais costeiros, da costa atlântica dos Estados Unidos da América, foi possível reconstruir o NMM nos últimos 2100 anos; os dados do estado americano da Carolina do Norte, após realizadas correções para os fenómenos isostáticos, revelam quatro fases distintas de alteração (Kemp et al., 2011). Inicialmente, o NMM permaneceu estável entre 100 a.C. até 950 d.C., seguindo-se um período de subida na ordem de 0,6 mm/ano, por um período de 400 anos. Posteriormente, verificou-se que o NMM permaneceu estável ou, eventualmente, com uma ligeira descida, até ao século XIX. Desde então, o NMM descreveu uma tendência de subida sem precedentes, com uma taxa de aumento médio de

2,1mm/ano. O estudo concluiu que essas variações não excedem 0,25 m até ao início do século XIX, aquando do início de uma nova fase de subida. A atual tendência de subida está a ser maior do que qualquer tendência dos últimos 2100 anos. A Figura 1.7 apresenta a representação gráfica da curva de evolução do NMM no período analisado juntamente com as correções isostáticas realizadas.

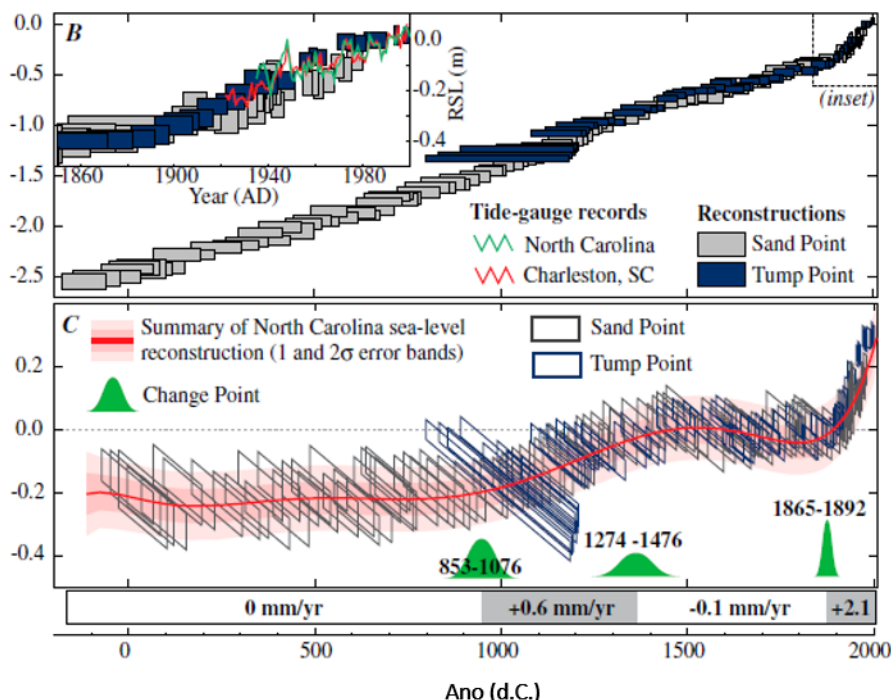


FIGURA 1.7: (B) reconstrução do NMM relativo a partir de duas localizações diferentes, Sand Point e Tump Point, no estado norte-americano da Carolina do Norte desde 100 a.C. (C) reconstrução do NMM após aplicadas as correções isostáticas nas duas localizações. (extraído de Kemp et al., 2011)

Outros estudos foram desenvolvidos para as demais regiões do planeta, tendo-se obtido resultados semelhantes. Em Massachusetts, também localizado na costa atlântica dos Estados Unidos da América, foi desenvolvido um modelo de alta resolução da evolução do NMM nos últimos 1500 anos, tendo por base os microfósseis de plantas existentes em sapais costeiros. Os resultados obtidos estão em concordância com a reconstrução realizada para o estado da Carolina do Norte, com exceção do período entre 700 e 1000 d.C. onde se verificou uma tendência de subida (Kemp et al., 2011). No Golfo do México, os dados mostram que o NMM permaneceu estável até 1000 d.C., seguindo-se uma tendência de subida positiva, cujo pico terá sido atingido por volta de 1200 d.C.. No estado do Connecticut, estima-se que tenha ocorrido um rápido aumento do NMM até 1000 d.C. No entanto, os dados obtidos da época em questão podem estar comprometidos pela erosão causada pela

influência de furacões. No período entre 1300 e 1800 d.C., os dados mostram que o nível permaneceu estável, correspondendo com o estudo feito no estado da Carolina do Norte.

Na Islândia, os dados recolhidos de sapais costeiros mostram que o NMM teve uma tendência de subida entre 100 A.C e 500 d.C., seguindo-se um período de descida entre 500 e 1800 d.C., possivelmente, como consequência da contração térmica dos oceanos (Gehrels et al., 2006). Estima-se que entre 1800 e 1840, se tenha iniciado uma nova fase de tendência, verificando-se uma rápida subida do mesmo, que permanece até ao presente (Gehrels et al., 2006). Na América do Norte, com exceção do Connecticut, existe concordância nos valores apresentados para a maior parte dos dois últimos milénios (Gehrels et al., 2006). Além disso, estudos mostram que o NMM manteve-se constante, ou com ligeiras reduções, entre 1400 e 1900 d.C. tanto na América do Norte como no Golfo do México e Islândia, sendo que foi neste período em que se verificou a “Pequena Idade do Gelo”.

No Mediterrâneo, as estruturas e artefactos de culturas clássicas fornecem elementos quantitativos importantes para os estudos. Com base em evidências arqueológicas romanas, demonstrou-se que o atual NMM, em Itália, está cerca de 0,13 metros acima daquilo que se estima que estaria há 2000 anos. Estes dados foram obtidos através da análise de tanques de piscicultura da época romana. No entanto, este tipo de estruturas deixou de ser edificado por volta de 50 d.C., como tal, posteriormente a esta data, deixou de ser possível reconstruir o NMM através deste tipo de dados (Lambeck et al., 2004).

Em Israel, no Mediterrâneo Oriental, a relação entre lençóis freáticos e o NMM foi utilizada para recrear níveis passados na antiga cidade de Caesarea, povoada até 1300 d.C.. Evidências arqueológicas provenientes de poços localizados próximos de costa (a menos de 150 metros) mostram que entre 100 e 900 d.C. o NMM se manteve estável ou, eventualmente, com ligeira descida. De notar que, durante os períodos de ocupação bizantina e árabe, o NMM foi muito semelhante ao atual e que entre 300 e 700 a.C. esteve, muito possivelmente, acima do atual (Sivan et al., 2004).

No Pacífico Ocidental, dados provenientes da Nova Zelândia, mostram que o NMM se manteve estável ou com ligeira descida entre 1400 e 1900 d.C. (Gehrels et al., 2008). Nas ilhas Cook, os dados sugerem que houve uma descida do nível do mar ao longo dos 2 milénios (Goodwin e Harvey, 2008).

Na Figura 1.8, estão representadas as reconstruções do NMM passado nas

diferentes regiões do estudo. A curva cor-de-rosa mostra a evolução do NMM na Carolina do Norte, permitindo a comparação visual entre os dados obtidos para as demais regiões com a reconstrução da Carolina do Norte (Kemp et al., 2011).

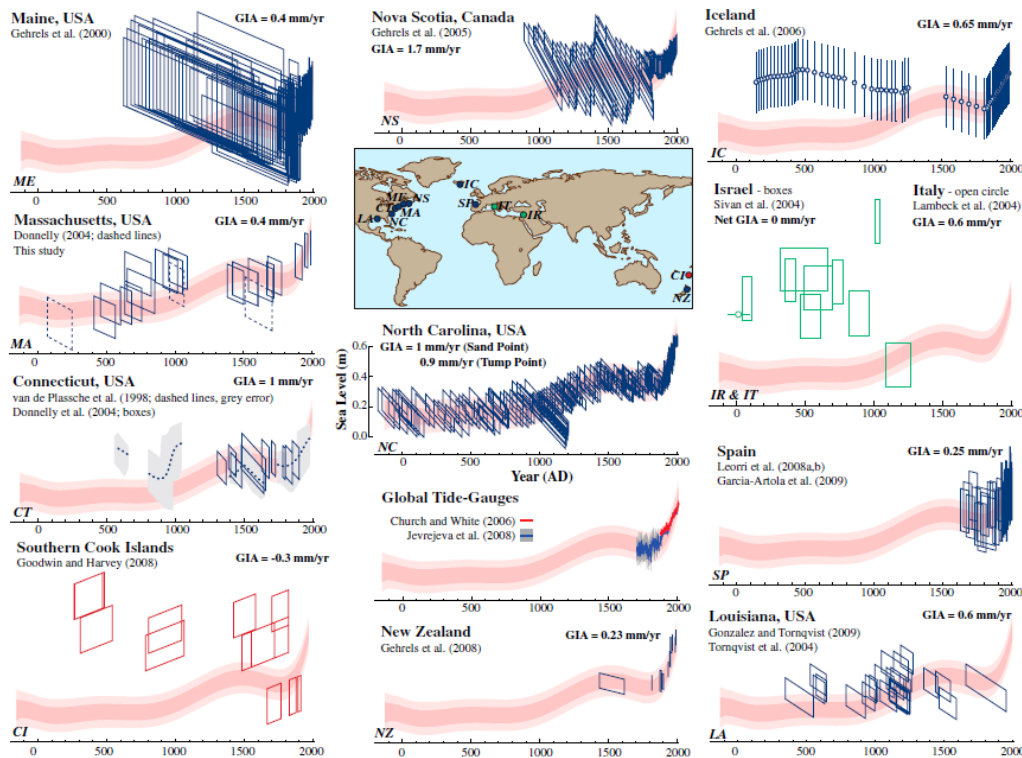


FIGURA 1.8: Comparação entre reconstruções do NMM durante o Holoceno tardio após aplicadas correções isostáticas na Carolina do Norte e evidências de NMM noutras regiões do planeta (extraído de Kemp et al., 2011).

Os estudos que permitiram realizar a reconstrução do NMM do mar do Oceano Atlântico estão em consonância em grande parte do período dos dois últimos milénios, em especial, no que se refere à última tendência de subida, que ainda hoje persiste. As reconstruções realizadas no Atlântico e no Pacífico permitem limitar o início da recente subida do nível do mar ao intervalo de tempo entre 1880 e 1920 (Kemp et al., 2011). No entanto, os registos recolhidos na Islândia, sugerem que este acontecimento tenha começado anteriormente, possivelmente entre 1800 e 1840, com os dados a indicar que apenas neste período, houve uma elevação do NMM na ordem dos 0,4 metros na costa ocidental da ilha (Gehrels et al., 2006).

1.6.3 Um Novo Paradigma no Século XX

Nas últimas décadas surgiram vários estudos referentes à evolução do NMM, tendo por base registos maregráficos e altimétricos, sendo que no geral existe consenso que efetivamente houve aumento acentuado do NMM e que a sua evolução não se realizou de forma linear ao longo do século XX (Cazenave e Remy, 2011).

A Figura 1.9, apresenta o resultado do estudo de Church e White (2011), onde se verificam dois períodos com tendências distintas de subida do NMM. O período entre 1880 e 1935 com valores $1,1 \pm 0,7$ mm/ano, e o outro entre 1936 e 2009 com uma subida de $1,8 \pm 0,3$ mm/ano.

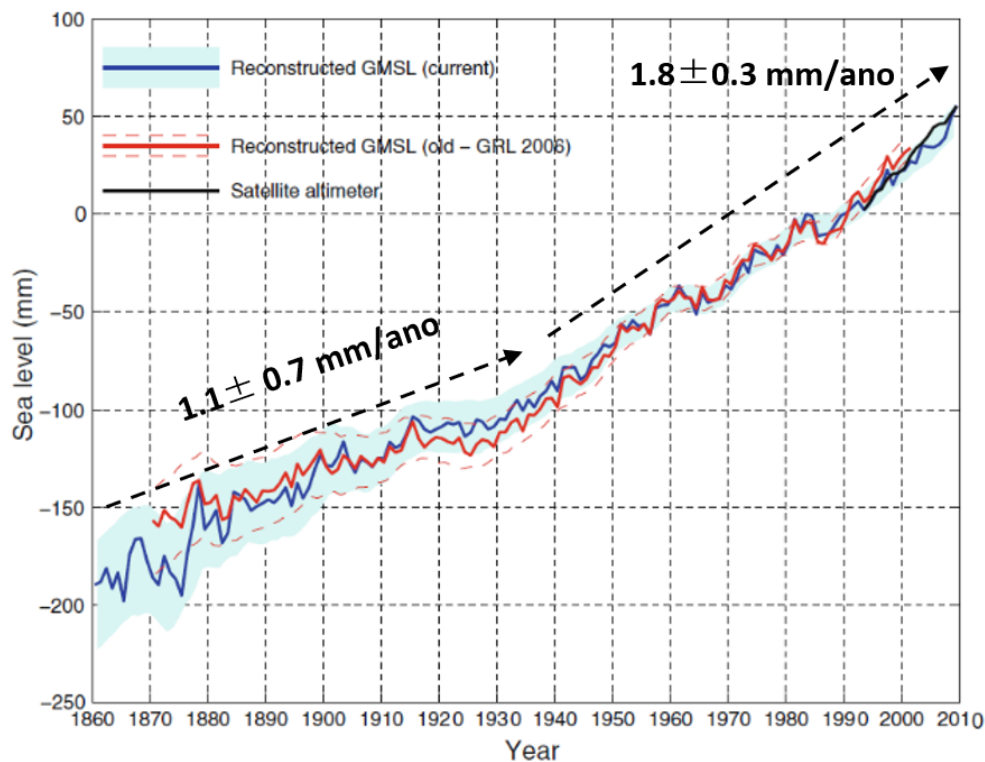


FIGURA 1.9: Evolução do NMM global entre 1860 a 2009. A linha vermelha apresenta uma reconstrução realizada por Church e White, 2006. Linha azul e linha preta representam reconstruções realizadas, respetivamente, através da análise de registos maregráficos e medições altimétricas realizadas por (adaptado de Church e White, 2011)

No segundo período, entre 1936 e 2009, verificou-se uma descida abrupta após 1962. O NMM desceu cerca de 10 mm entre 1962 e 1967. Após esta inesperada descida, o NMM voltou novamente a subir. Posteriormente a 1967, o NMM passou a registar subidas na ordem de 2,4 mm/ano. Por fim, logo após 1993, a tendência de subida acentuou-se, passando a ser de $2,8 \pm 0,8$ mm/ano (Church e White, 2011).

Em 1993, iniciou-se a medição do NMM com recurso a satélites altimétricos. Para o mesmo período, o aumento fixou-se em $3,2 \pm 0,4$ mm/ano através deste método. Este valor é coerente com outros estudos, segundo Cazenave e Remy (2011) entre 1993 e 2011 houve um aumento de aproximadamente $3,3 \pm 0,4$ mm/ano. As duas tendências, uma com base nos registos de séries de dados maregráficos e outra com base em altimetria por satélite, apresentam valores relativamente próximos. Até 2009, as maiores variações anuais entre as duas tendências não ultrapassaram os 5 mm, como se pode ver na Figura 1.10.

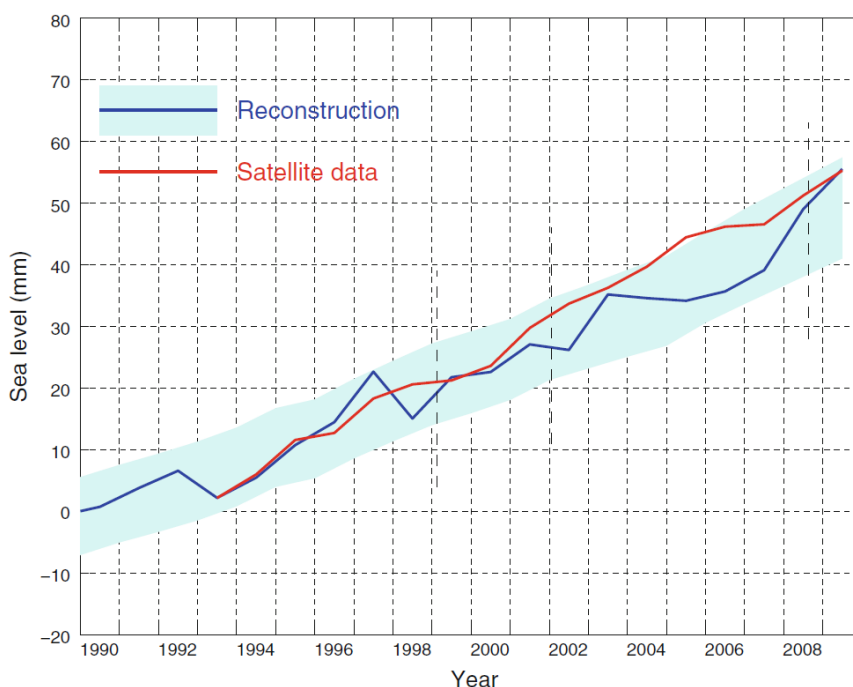


FIGURA 1.10: Variação do NMM global entre 1990 e 2009. Linha Vermelha representa a reconstrução feita através de dados altimétricos; Linha Azul representa a reconstrução feita com séries de dados mareógrafos (extraído de Church e White, 2011)

O estudo de Church e White (2011) conclui que entre 1880 e 2009 o NMM global subiu cerca de 210 mm. No entanto, as últimas décadas têm um grande significado no valor final, com Church e White (2006) a identificarem uma subida de 195 mm apenas entre 1970 e 2006.

Ao longo do século XX e inícios do século XXI verificou-se um gradual aumento da tendência de subida, superior a qualquer período dos últimos dois milénios. O NMM está a subir a um ritmo muito elevado e assim deverá continuar nas próximas décadas, devido ao esperado aumento do aquecimento global. No entanto, o valor exacto (ou os limites) da tendência de subida do NMM até ao final deste século

é ainda uma questão em aberto. Segundo Jevrejeva et al. (2008), se se mantiver a tendência atual, pode verificar-se uma elevação do NMM em cerca de 34 cm durante o século XXI.

1.6.4 O Século XX em Portugal

Em Portugal existem várias séries de registos maregráficos provenientes de marégrafos dispersos ao longo do seu litoral (Picoto et al., 2022a). Segundo Dias e Taborda (1988), todas as séries temporais das estações maregráficas consideradas no seu estudo apresentavam tendências de subida do NMM ao longo do século XX (até 1982). No entanto, apenas as estações maregráficas de Cascais e Lagos permitem chegar a conclusões fiáveis, por possuírem registos longos que possibilitam a determinação de tendências seculares. Todas as outras estações têm registos claramente insuficientes.

Como já foi apresentado em Picoto et al. (2022a), tanto os registos maregráficos de Cascais como de Lagos evidenciam uma tendência de subida do NMM, no entanto, as análises das séries de dados maregráficos sugerem que essa subida não tem uma tendência consistente ao longo do tempo. A análise dos dados indica dois períodos temporais distintos. Desde final do século XIX até por volta de 1915-1920, o NMM teve inclusive uma tendência negativa, $-5 \pm 0,4$ mm/ano em Cascais (Ferreira et al., 2008). Só posteriormente a este período é que se nota uma tendência de subida (Picoto et al., 2022a).

No segundo período, verificou-se um aumento de $1,7 \pm 0,2$ mm/ano em Cascais e $1,2 \pm 0,2$ mm/ano em Lagos. Dias e Taborda (1988) mostraram que a natureza destes números aponta para uma subida predominantemente eustática, considerando que esta se deve à expansão térmica do oceano. Uma tendência idêntica foi obtida para Cascais, no período ligeiramente mais alargado de 1920-2000, onde se registou uma tendência de 1,9 mm/ano (Antunes e Taborda, 2009). Na Figura 1.11, encontra-se apresentada a reconstrução do NMM em Cascais no período considerado.

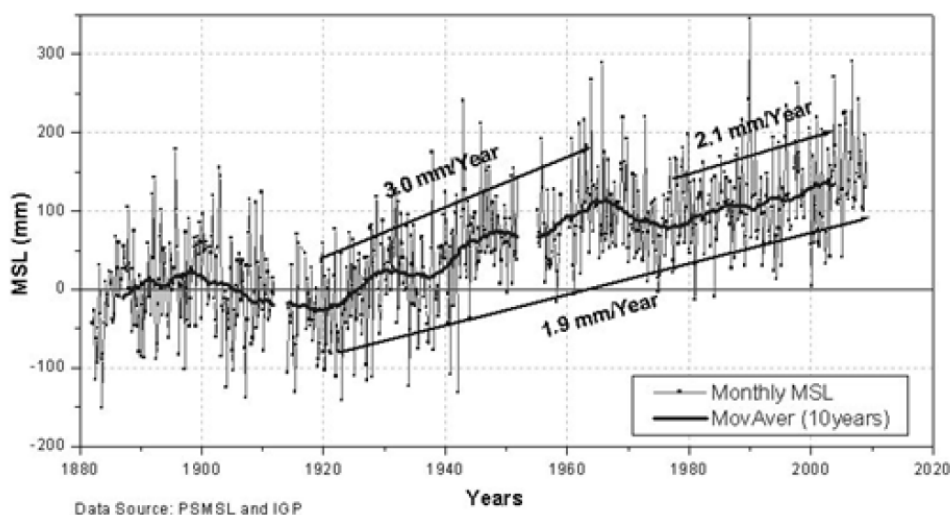


FIGURA 1.11: NMM no marégrafo de Cascais (1882-2008). Reconstituição realizada a partir de séries temporais da média mensal do NMM com sobreposição da média móvel de 10 anos (extraído de Antunes e Taborda, 2009)

Estes valores são consistentes com as estimativas globais, por exemplo, Church e White (2006), referem uma tendência de $1,8 \pm 0,3$ mm/ano. Para Antunes e Taborda (2009), a coerência entre valores regionais, de Cascais, e valores globais é justificado pelo facto de Cascais verificar movimento verticais pouco significantes. No estudo realizado por Mendes et al. (2020), tendo por base a utilização de registos *GPS*, concluiu-se que os movimentos verticais na Península Ibérica são, na sua maioria caracterizados por tendências de subsidência lentas/moderadas, com um valor médio de 0,6 mm/ano, sendo que em Cascais verificou-se uma tendência de 0,4 mm/ano (Picoto et al., 2022a). Como já referido no trabalho de Picoto et al. (2022a) isto significa que as séries maregráficas do marégrafo de Cascais são fidedignas de serem usadas em estudos globais sobre o NMM e que podem ser usadas para caracterizar a tendência de subida do NMM nas últimas décadas.

Capítulo 2

Fundamentação Teórica

2.1 Definições

Neste capítulo pretende-se explicar a metodologia utilizada no estudo das séries de dados maregráficos. Estes dados apresentam-se na forma de série temporal, como tal, este capítulo foca-se nos fundamentos teóricos de análise de uma série temporal, com a finalidade de analisar o seu comportamento e fazer previsões.

As séries temporais decompõem-se em três componentes: 1) tendência; 2) sazonalidade; e 3) componente puramente aleatória (Morettin e Toloi, 2006). Como tal, numa primeira fase, serão abordados os métodos aplicáveis à obtenção de estimativas da tendência e da sazonalidade. Numa segunda fase será descrita a metodologia de Box e Jenkins, por forma a identificar o melhor modelo que se ajuste ao comportamento geral de uma dada série temporal e que, a partir deste, se possam fazer as previsões da série.

Iniciaremos este capítulo por abordar o princípio dos testes de hipóteses que serão amplamente utilizados ao longo da metodologia utilizada.

Toda esta metodologia será aplicada no capítulo seguinte com recurso ao *software* estatístico **R** e respetivas bibliotecas necessárias à obtenção dos resultados das técnicas que se descrevem em seguida.

2.2 Teste de Hipótese

Nesta secção apresenta-se, de forma genérica, a metodologia para a construção e análise de um teste de hipótese (para mais detalhe, podem ser consultados os livros de Murteira (1990) e Velosa e Pestana (2010)).

O objetivo de um teste de hipóteses, consiste em validar, ou não, uma dada suposição sobre uma ou mais populações.

As hipóteses a testar são:

- H_0 : Hipótese nula (é sempre a hipótese mais restritiva) – algo que se pretende manter;
- H_1 : Hipótese alternativa (contém: $>$, $<$ ou $\neq$)

Uma hipótese nula pode ser falsa ou verdadeira, no entanto, existe sempre um grau de incerteza na decisão a tomar (Murteira, 1990).

De acordo com Nunes e Afonso (2005), para se realizar um teste de hipóteses, deve-se percorrer as seguintes etapas:

1.^a etapa: Formular as hipóteses (H_0 e H_1), o teste diz-se unilateral quando H_1 contiver uma das desigualdades $<$ ou $>$; o teste diz-se bilateral quando contiver a não igualdade ($\neq$);

2.^a etapa: Escolher a estatística de teste e calcular o seu valor;

3.^a etapa: Decidir pela rejeição, ou não, de H_0 .

A Figura 2.1 ilustra o intervalo de rejeição de H_0 para um teste bilateral.

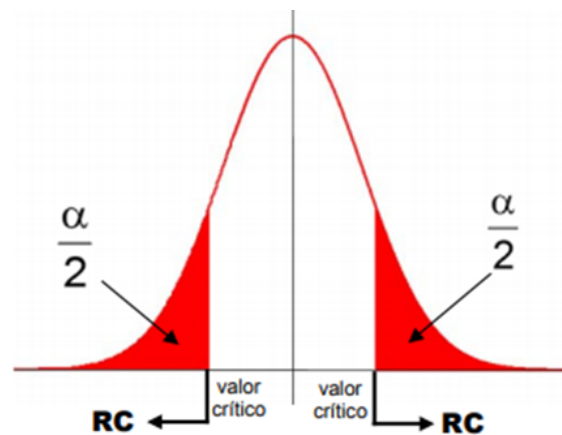


FIGURA 2.1: Intervalo de rejeição de H_0 para um teste bilateral quando se assume a normalidade da variável aleatória (adaptado de: shorturl.at/fjnS0, consultado em 19 de maio de 2022)

Nota: quando se rejeita H_0 podemos concluir que esta é falsa; no entanto, se não se rejeitar H_0 conclui-se que não existe evidência (estatística) que suporte a rejeição de H_0 (Maroco, 2007).

Ao decidir sobre aceitação, ou rejeição, de H_0 , podem cometer-se dois tipos de erros; H_0 é verdadeira e é rejeitada, trata-se de um erro do tipo I; se H_0 é falsa e é aceite, comete-se um erro do tipo II (Nunes e Afonso, 2005).

A Tabela 2.1 esquematiza os erros associados à tomada de decisão (Maroco, 2007):

Verdade	Aceitar H_0	Rejeitar H_0
H_0 Verdadeiro	Decisão Correta ($1 - \alpha$)	Erro Tipo I (α)
H_0 Falso	Erro Tipo II β	Decisão Correta ($1 - \beta$)

TABELA 2.1: Erros associados à aceitação, ou rejeição, de H_0

O fundamental dos testes de hipóteses, prende-se com a escolha da região de rejeição que permite minimizar cada um dos tipos de erro (Nunes e Afonso, 2005).

O nível de significância habitual é $\alpha = 0,05$ (5%) mas podem ser usados outros níveis de significância, tais como 0,1 e 0,01 (Maroco, 2007).

O $p - value$ é o menor valor de significância, α , a partir do qual se rejeita H_0 (Nunes e Afonso, 2005).

Assim, fixado à partida, o erro de tipo I aceitável e conhecida a distribuição amostral da estatística de teste, é possível construir uma região de rejeição que, sob a validade de H_0 , tem uma probabilidade ($1 - \alpha$) de conter o valor observado para a estatística de teste (Maroco, 2007). Conhecidas as distribuições amostrais sob a validade de H_0 e de H_1 é possível determinar α e β (Figura 2.2):

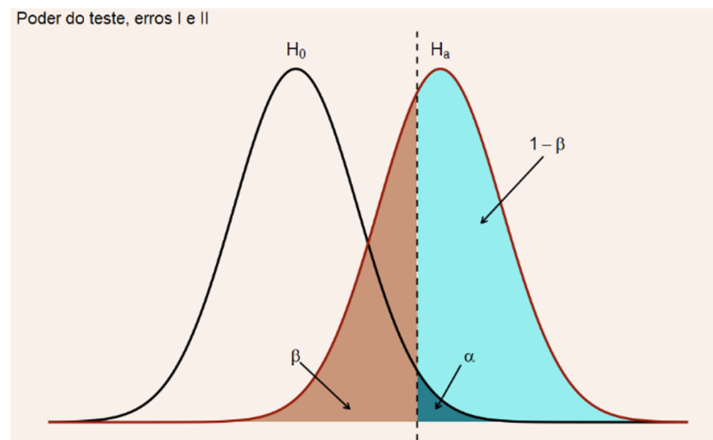


FIGURA 2.2: Determinação de α e β (adaptado de: shorturl.at/gHOS1, consultado em 20 de maio de 2022)

Atualmente, a maioria dos *softwares* calcula apenas o $p - value$ e, para um determinado nível de significância ou probabilidade de erro de tipo I (α), deve-se rejeitar H_0 se $p - value \leq \alpha$ (Fisher e Bischel, 1973).

2.3 Série Temporal

Entende-se por série temporal, qualquer sucessão de observações ordenadas no tempo. São exemplo de séries temporais: valores diários de temperatura de Lisboa, precipitação atmosférica mensal na cidade de Coimbra, alturas horárias de marés registadas no marégrafo de Cascais, entre outras.

As séries temporais podem ser contínuas ou discretas. No entanto, para que possam ser analisadas, as séries temporais têm obrigatoriamente de ser discretas, isto é, o conjunto de observações que a constitui é apresentado em intervalos de tempo sucessivos (por exemplo, de hora em hora). Portanto, sempre que estamos perante uma série temporal contínua, obtida no período $[0, T]$, torna-se necessário convertê-la numa série discreta com n observações, onde $n = \frac{T}{\Delta t}$ (Morettin e Toloi, 2006). Nestas situações, dizemos que a série foi discretizada.

O estudo de séries temporais tem essencialmente três objetivos: 1) identificar e compreender o mecanismo gerador da série temporal; 2) caracterizar o comportamento da série, ou seja, identificação de tendências, de sazonalidade; e 3) efetuar previsões de valores futuros (Morettin e Toloi, 2006).

2.3.1 Componentes de Uma Série Temporal

Considerando o conjunto de observação $X_t : t = 1, \dots, n$ de uma série temporal. Morettin e Toloi (2006) mostram que X_t pode ser escrito como a soma de três componentes não-observáveis,

$$X_t = T_t + S_t + \epsilon_t, \quad (2.1)$$

onde T_t e S_t representam, respetivamente, a tendência e a sazonalidade, enquanto ϵ_t é uma componente puramente aleatória, de média igual a zero e variância constante σ_ϵ^2 . Uma série temporal diz-se puramente aleatória se cumprir com as condições de estacionariedade, que serão apresentadas na secção seguinte.

Nas duas próximas subsecções, serão apresentados dois métodos de regressão distintos que, quando combinados, permitem caracterizar o comportamento geral da série, tanto em termos de tendência como de sazonalidade.

2.3.2 Tendência

Supondo a ausência da sazonalidade, X_t é dada por:

$$X_t = T_t + \epsilon_t, \quad (2.2)$$

onde ϵ_t é ruído branco³, com variância σ_ϵ^2 .

Existem vários métodos possíveis para estimar T_t . De acordo com Morettin e Toloi (2006), um dos métodos mais utilizados para a determinação da tendência consiste em ajustar uma função aos valores da série temporal. Entre as funções possíveis de serem utilizadas encontram-se as funções exponencial, logarítmica e polinomial. No entanto, vamos nos limitar a descrever sucintamente o ajuste polinomial visto que, assim como nos estudos que foram abordados no primeiro capítulo, também neste trabalho, se fará uso apenas deste tipo de função.

Ajuste de Uma Função Polinomial

Supondo que pretendemos representar a série temporal através da soma de uma função linear (polinômio de grau um) e de uma componente puramente aleatória. Nessas condições, temos que:

$$X_t = \beta_0 + \beta_1 t + \epsilon_t, \quad (2.3)$$

onde β_0 e β_1 são parâmetros desconhecidos e ϵ_t representa a componente puramente aleatória (ou erro). Conforme Morettin e Toloi (2006), para estimar estes dois parâmetros, minimizamos a soma dos quadrados dos resíduos,

$$f(\beta_0, \beta_1) = \sum_{t=1}^n (X_t - \beta_0 - \beta_1 t)^2, \quad (2.4)$$

onde β_0 e β_1 são as soluções que minimizam (2.4),

$$\hat{\beta}_1 = \frac{\sum_{t=1}^n (X_t - \bar{X})(t - \bar{t})}{\sum_{t=1}^n (t - \bar{t})^2}, \quad (2.5)$$

³Ver página 47.

$$\hat{\beta}_0 = \bar{X} - \hat{\beta}_1 \bar{t},$$

onde $\bar{X} = \frac{1}{n} \sum_{t=1}^n X_t$ e $\bar{t} = (n+1)/2$.

Na Figura 3.3, encontra-se representado um exemplo do ajuste de um polinómio de grau um a uma série temporal, cujo declive corresponde à tendência geral da série.

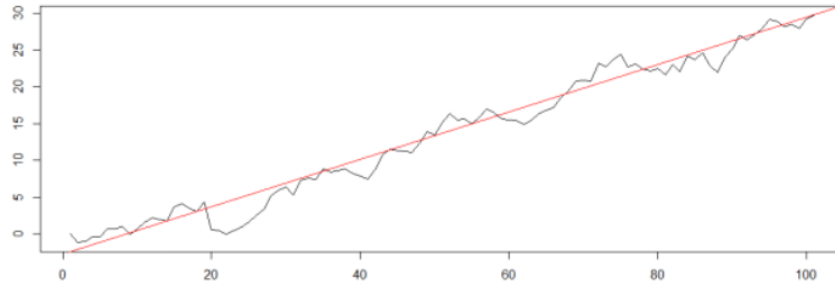


FIGURA 2.3: Ajuste de uma função polinomial de grau um (adaptado de: shorturl.at/lvwZ0, consultado em 26 de maio de 2022)

Suponhamos agora que se pretende ajustar um polinómio de grau superior a um. Nesse caso, o modelo seria o seguinte (Morettin e Toloi, 2006):

$$X_t = \beta_0 + \beta_1 t + \dots + \beta_m t^m + \epsilon_t, \quad (2.6)$$

onde m representa o grau do polinómio. Para se estimar os valores $(\beta_i, i = 0, 1, 2, \dots, m)$ utiliza-se o método dos mínimos quadrados, que de acordo com Morettin e Toloi (2006), segue o seguinte formalismo,

$$f(\beta_0, \beta_1, \dots, \beta_m) = \sum_{t=1}^n (X_t - \beta_0 - \beta_1 t - \dots - \beta_m t^m - \epsilon_t)^2, \quad (2.7)$$

obtendo-se assim, os estimadores de mínimos quadrados $(\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_m)$.

Importa realçar que o ajuste polinomial embora possa ajustar-se bem aos valores da série temporal, pode originar extrapolações muito diferentes dos valores reais (Morettin e Toloi, 2006).

2.3.3 Sazonalidade

Supondo a ausência da tendência, X_t é dada por:

$$X_t = S_t + \epsilon_t. \quad (2.8)$$

Assim como na determinação da tendência, existem várias formas de se obter a sazonalidade. No presente trabalho, optou-se pelo ajuste de funções periódicas/tendência cíclica, uma vez que, as oscilações sazonais NMM apresentam frequências conhecidas, facilitando assim, o ajuste deste tipo de função.

Ajuste de Uma Função de Tendência Cíclica

Num modelo de tendência cíclica supõem-se que o comportamento geral da série é explicado por um número de sinusoides e por uma variável aleatória. Um modelo de tendência cíclica, conforme Murteira et al. (2000), é expresso por,

$$X_t = \sum_{p=0}^{n/2} [a_p \cos(\omega_p t) + b_p \sin(\omega_p t)] + \epsilon_t, \quad (2.9)$$

onde ω é a frequência angular, em radianos, (a_p, b_p) são coeficientes e se admite que ϵ_t é ruído branco (Murteira et al., 2000).

Se o investigador conhece *a priori* as frequências das sinusoides, em número de j , que contribuem significativamente para a variância do processo observado e supondo que $\sum_{t=1}^j X_t = 0$, o modelo passa a ter a forma,

$$X_t = \sum_{p=1}^j [a_p \cos(\omega_p^* t) + b_p \sin(\omega_p^* t)] + \epsilon_t, \quad (2.10)$$

onde $\omega_p^*, p = 1, 2, \dots, j$ representam as frequências das sinusoides, que contribuem, significativamente.

A estimação dos coeficientes $(a_p, b_p), p = 1, 2, \dots, j$ e σ_ϵ^2 pode fazer-se pelo método dos mínimos quadrados (Murteira et al., 2000).

2.4 Estacionariedade

Sempre que se tenta modelar uma série temporal é necessário saber se esta é ou não temporalmente invariante do ponto de vista probabilístico. Se as características probabilísticas variam ao longo do tempo, torna-se muito difícil representar, através de modelos algébricos, as observações passadas e futuras.

Se a série se desenvolve ao longo do tempo em torno de um único nível médio, apresentando equilíbrio estatístico, podemos dizer que a estrutura probabilística

(conjunto de todas as propriedades estatísticas) é invariante relativamente a alterações da origem do tempo (Murteira et al., 2000). Nesta situação, torna-se possível representar a série temporal através de um modelo algébrico cujos coeficientes são estimados tendo por base as observações em instantes anteriores (Santos, 2011).

Dada uma série temporal,

$$X_t : t = \dots, -1, 0, 1, \dots,$$

de acordo com Murteira et al. (2000) o processo diz-se:

- estacionário de 1.^a ordem em distribuição se,

$$F^4(x_1; t_1) = F(x_1; t_{1+k}), \quad \forall t_1, k \in Z \quad (2.11)$$

- estacionário de ordem superior a um em distribuição se,

$$F(x_1, \dots, x_n; t_1, \dots, t_n) = F(x_1, \dots, x_n; t_{1+k}, \dots, t_{n+k}), \quad \forall t_1, k \in Z \quad (2.12)$$

Um processo diz-se estritamente (fortemente ou completamente) estacionário quando verifica (2.12) para qualquer ordem (Murteira et al., 2000).

A estacionariedade no sentido estrito é uma propriedade difícil de verificar. Desta forma, na maioria das aplicações práticas é usual considerar a estacionariedade dos momentos até à 2.^a ordem (valor esperado e autocovariância⁵). A este tipo menos restritivo de estacionariedade dá-se o nome de estacionariedade em sentido lato (Murteira et al., 2000).

A estacionariedade até 2.^a ordem verifica-se quando a média e a variância da série são constantes. Ou seja, caso o valor esperado exista, $(E(|X_t|^2) < \infty)$, cumpre os seguinte pressupostos apresentados por Santos (2011):

$$\begin{aligned} E(X_t) &= \mu, \forall t \in Z, \\ Var(X_t) &= \sigma^2, \forall t \in Z, \\ Cov(X_t, X_s) &= Cov(X_{t+k}, X_{s+k}) = \gamma(|t-s|), \forall t, s \in Z, \end{aligned} \quad (2.13)$$

⁴Função distribuição.

⁵Ver página 45.

onde μ e σ^2 representam, respetivamente, a média e variância de um conjunto de dados e γ representa a autocovariância.

Isto significa que a covariância entre X_t e X_s é constante para qualquer valor de t e s , dependendo apenas do espaçamento temporal entre as variáveis, $|t - s|$ (Santos, 2011).

A maioria das séries apresentam alguma forma de não estacionariedade. A Figura 2.4 apresenta um exemplo de uma série estacionária, representada pelo gráfico da esquerda, e uma série nitidamente não estacionária, gráfico da direita, caracterizada por tendência e oscilações sazonais com amplitude crescente ao longo do tempo.

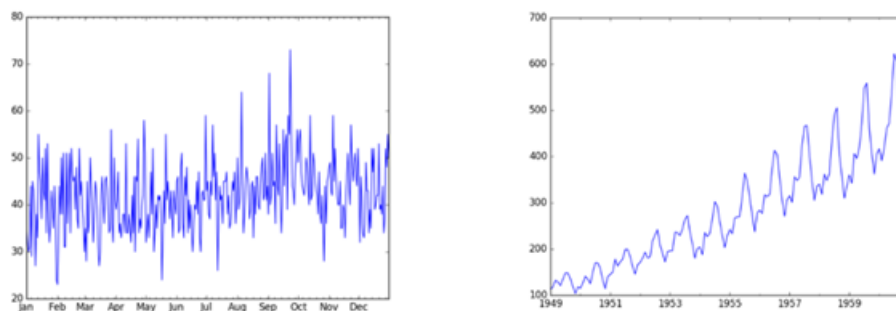


FIGURA 2.4: Gráfico da esquerda: Série estacionária. Gráfico da direita: Série não estacionária (adaptado de: <https://directions.pt/inteligencia/machine-learning-na-previsao-de-series-temporais>, consultado em 15 de abril de 2022)

Como a maioria das séries temporais não são estacionárias, torna-se necessário aplicar transformações por forma a estabilizar tanto a média como a variância. Só depois de se atingir a estacionariedade é possível aplicar modelos algébricos que descrevam o comportamento geral da série e que possibilitem a realização de previsões.

NOTA: No presente trabalho, passará a ser usada a designação de “série estacionária” para referir séries estacionárias em sentido lato.

2.5 Funções de Autocovariância, Autocorrelação e Autocorrelação Parcial

Assumindo uma série X_t estacionária, podemos definir a ordem k da função de autocovariância, γ_k , de acordo com Murteira et al. (2000), dada por,

$$\gamma_k = E \{ (X_t - \mu)(X_{t+k} - \mu) \} \quad (2.14)$$

A função γ_k , mede a intensidade com que covariam pares de valores do processo separados por um intervalo (“lag”) de amplitude $k \times \Delta t$ (Murteira et al., 2000). Segundo o mesmo autor, a autocovariância pode ser normalizada e apresentada como uma função de autocorrelação, ρ_k ,

$$\rho_k = \frac{\gamma_k}{\gamma_0}, \quad (2.15)$$

A função de autocorrelação, também designada de FAC, permite identificar a correlação entre pares de valores separados por um intervalo k , ou seja, é uma medida de dependência entre uma observação e a mesma deslocada k unidades de tempo (Murteira et al., 2000).

De acordo com Murteira et al. (2000), a representação gráfica de ρ_k denomina-se por correlograma; este é um dos principais instrumentos de análise de séries temporais. Por norma, à medida que se dá o aumento de k , verifica-se uma diminuição de γ_k e ρ_k . A forma como se verifica o decaimento de ρ_k pode ser interpretada como uma medida da memória do processo ou dos processos geradores da série temporal.

A função de autocorrelação parcial, também designada de FACP, permite quantificar a correlação entre pares de valor separados por um intervalo k (X_t e X_{t+k}), depois de eliminar o efeito sobre eles das variáveis intermédias (Murteira et al., 2000).

A função de autocorrelação parcial, ϕ_{kk} , de acordo com Murteira et al. (2000), é dada por,

$$\phi_{kk} = \begin{vmatrix} 1 & \rho_1 & \dots & \rho_{k-2} & \rho_1 \\ \rho_1 & 1 & \dots & \rho_{k-3} & \rho_2 \\ \dots & \dots & \dots & \dots & \dots \\ \rho_{k-1} & \rho_{k-2} & \dots & \rho_1 & \rho_k \end{vmatrix} \div \begin{vmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{k-1} \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{k-2} \\ \dots & \dots & \dots & \dots & \dots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \dots & 1 \end{vmatrix} \quad (2.16)$$

2.6 Ruído Branco

Um processo diz-se que constitui ruído branco, ϵ_t , quando é constituído por uma sequência de variáveis aleatórias com a mesma distribuição e cujas características probabilísticas, de acordo com Murteira et al. (2000), são tais que,

$$\begin{aligned} E(\epsilon_t) &= \mu_\epsilon \text{ (usualmente } \mu_\epsilon = 0) & Var(\epsilon_t) &= \sigma_\epsilon^2 \\ \gamma_k &= \sigma_\epsilon^2, \text{ se } k = 0 & \rho_k &= 1, \text{ se } k = 0 \\ \gamma_k &= 0, \text{ se } k \neq 0 & \rho_k &= 0, \text{ se } k \neq 0 \end{aligned}$$

2.7 Estimação de Parâmetros Característicos dos Processos Estacionários

Estimar os parâmetros que caracterizam o processo estacionário é extremamente relevante na modelação de séries temporais. Os parâmetros a estimar são: a média μ , a variância σ^2 , as autocorrelações ρ_k e as autocorrelações parciais ϕ_{kk} .

A Tabela 2.2 descreve os estimadores destes parâmetros, obtidos a partir de uma amostra de uma série temporal, de acordo com Santos (2011):

parâmetro	Estimador
Média (μ)	$\bar{X} = \frac{1}{n} \sum_{t=1}^n X_t$
Variância (σ^2)	$\hat{\gamma} = \frac{1}{n} \sum_{t=1}^{n-k} (X_t - \bar{X})(X_{t+k} - \bar{X}), \text{ com } 0 \leq k \leq n-1$
Autocorrelações (ρ_k)	$\hat{\rho}_K = \frac{\hat{\gamma}_k}{\hat{\gamma}_0} = \frac{\sum_{t=1}^{n-k} (X_t - \bar{X})(X_{t+k} - \bar{X})}{\sum_{t=1}^n (X_t - \bar{X})^2}$
Autocorrelações parciais (ϕ_{kk})	$\hat{\phi}_{kk} = \frac{\hat{\rho}_k - \sum_{j=1}^{k-1} \hat{\phi}_{k-1} \hat{\rho}_{k-j}}{1 - \hat{\rho}_k - \sum_{j=1}^{k-1} \hat{\phi}_{k-1} \hat{\rho}_j}$

TABELA 2.2: Estimadores de parâmetros característicos

2.8 Processos Estacionários e Lineares

2.8.1 Processo Autoregressivo de Ordem p

O processo X_t diz-se autoregressivo de ordem p , $AR(p)$, de média nula, quando satisfaz a equação apresentada por Murteira et al. (2000),

$$X_t - \phi_1 X_{t-1} - \phi_2 X_{t-2} - \dots - \phi_p X_{t-p} = \epsilon_t, \quad \text{ou} \quad \phi_p(B)X_t = \epsilon_t, \quad (2.17)$$

$$\phi_p(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p,$$

sendo B o operador de atraso, de maneira que $BX_t = X_{t-1}$.

A metodologia utilizada para a obtenção dos parâmetros autoregressivos e dos respetivos valores das funções FAC e FACP pode ser encontrada em Murteira et al. (2000) (páginas 56-59).

Na análise da FACP deve ser tido em conta que, por norma, o processo $AR(p)$ descreve um decaimento muito rápido para zero a partir de um determinado intervalo (“lag”). Por exemplo, para um processo $AR(2)$, a FACP é nula depois do “lag” 2. No que diz respeito à FAC, esta apresenta, por norma, um decaimento gradual para zero (Murteira et al., 2000). A Figura 2.5 pretende exemplificar o comportamento das funções FAC e FACP de um processo $AR(2)$.

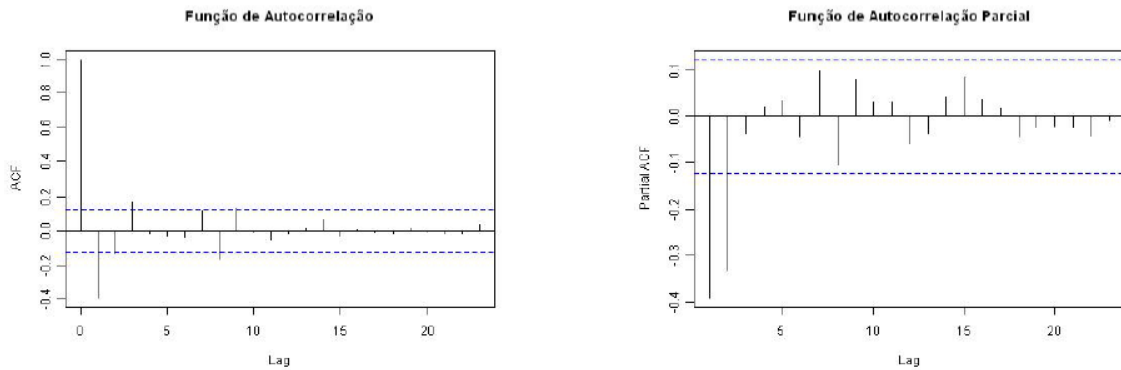


FIGURA 2.5: correlograma das funções FAC e FACP de um processo $AR(2)$, (adaptado de Santos, 2011)

2.8.2 Processos de Médias Móveis de Ordem q

De acordo com Murteira et al. (2000), diz-se que uma série temporal X_t resulta de um processo de média móvel se for descrita por uma equação da forma:

$$X_t = \epsilon_t - \theta_1 \epsilon_{t-1} - \dots - \theta_q \epsilon_{t-q} \quad \text{ou} \quad X_t = \theta_q(B) \epsilon_t, \quad (2.18)$$

$$\theta_q(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q,$$

sendo $\theta_q(B)$ o polinómio de média móvel de ordem q .

Estes processos resultam da ideia de exprimir o processo de X_t em termos de um processo mais simples ϵ_t . Ao contrário daquilo que se sucede com os processos autoregressivos, no modelo de médias móveis os efeitos causados por valores passados de X_t apenas perduram por um período de tempo finito (Murteira et al., 2000).

A metodologia utilizada para a obtenção dos parâmetros de médias móveis e dos respetivos valores das funções FAC e FACP pode ser encontrada no trabalho de Murteira et al. (2000) (páginas 59-66).

Graficamente, a FACP de um processo MA(q) decai gradualmente para zero, por valores positivos ou negativos, ou valores alternados, dependendo do sinal do parâmetro. A FACP de um processo MA(q) apresenta um comportamento semelhante ao de uma FAC de um modelo AR(p) (Murteira et al., 2000).

A Figura 2.6 pretende exemplificar o comportamento das funções FAC e FACP de um processo AM(2).

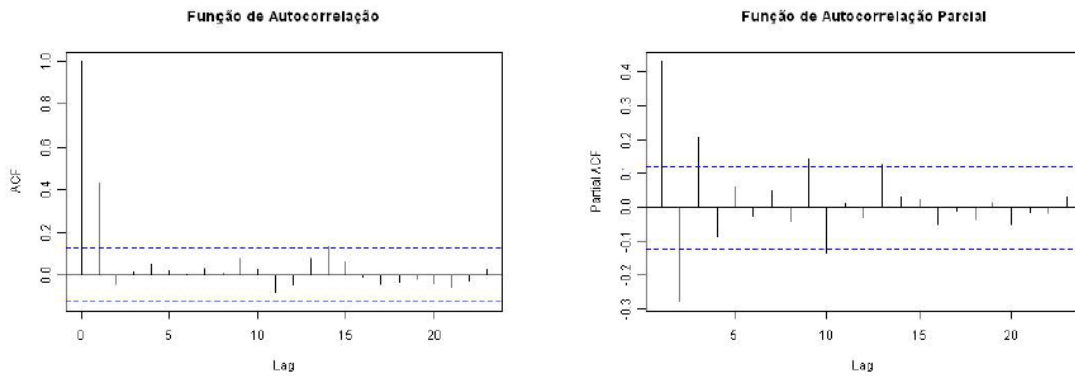


FIGURA 2.6: correlograma das funções FAC e FACP de um processo AM(2), (adaptado de Santos, 2011)

2.8.3 Processos Mistos Autoregressivos e de Médias Móveis

Pode acontecer que o processo gerador de série temporal tenha uma estrutura que resulte da combinação dos dois processos anteriores. Nesta situação, qualquer ensaio para utilizar uma aproximação exclusivamente com processo AR(p) ou MA(q), resulta num número elevado de parâmetros. Os modelos mistos surgem deste modo como modelos com menos parâmetros (Murteira et al., 2000).

O processo X_t , de média nula, diz-se ARMA(p,q) quando satisfaz o formalismo de Murteira et al. (2000), dado por,

$$X_t - \phi_1 X_{t-1} - \dots - \phi_p X_{t-p} = \epsilon_t - \phi_1 \epsilon_{t-1} - \dots - \phi_q \epsilon_{t-q} \quad (2.19)$$

ou,

$$\phi(B) = \phi_p(B)X_t = \theta_p(B)\epsilon_t,$$

com,

$$\begin{aligned} \phi(B) &= 1 - \sum_{i=1}^p \phi_i B^i, \\ \theta(B) &= 1 - \sum_{i=1}^p \theta_i B^i, \end{aligned}$$

sendo que os polinómios $\phi(B)$ e $\theta(B)$ não podem ter fatores comuns.

Em Murteira et al. (2000) (páginas 66-70) poderá ser consultada a metodologia utilizada para a obtenção dos parâmetros autorregressivos e de médias móveis, bem como da obtenção dos respetivos valores das funções FAC e FACP.

Assim como num processo autoregressivo AR(p), a FAC de um processo misto apresenta um decaimento exponencial e/ou sinusóide para zero. Paralelamente, a FACP de um processo misto comporta-se como a FACP de um processo de médias móveis MA(q), com um decaimento exponencial e/ou sinusóide amortecida⁶, (Murteira et al., 2000).

A Figura 2.7 pretende exemplificar o comportamento das funções FAC e FACP de um processo ARMA(1,1).

⁶sinusóide cuja amplitude decresce lentamente até se anular

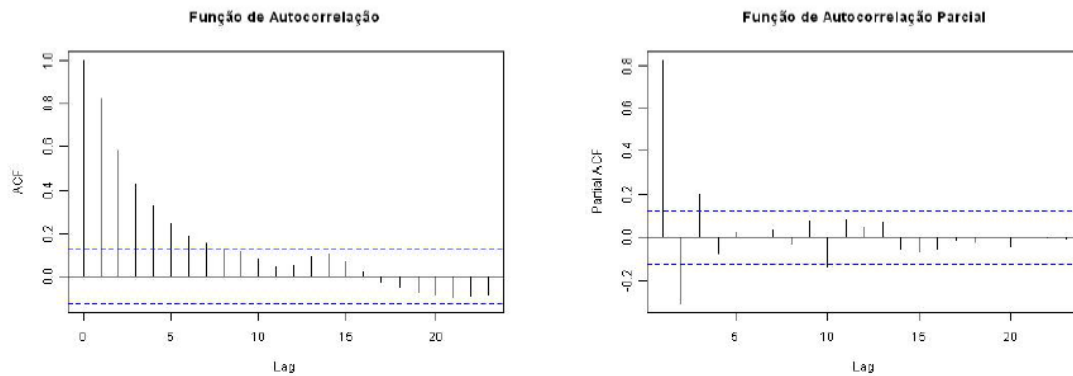


FIGURA 2.7: correlograma das funções FAC e FACP de um processo ARMA(1,1), (adaptado de Santos, 2011)

A Tabela 2.3, adaptada de Murteira et al. (2000), pretende sintetizar o comportamento das funções FAC e FACP dos três processos descritos.

	AR(p)	MA(q)	ARMA(p, q)
Modelo em termos dos valores anteriores de X_t	$\phi_p(B)X_t = \epsilon_t$ Série finita em X_t	$[\theta(B)]^{-1}X_t = \epsilon_t$ Série finita em X_t	$[\theta(B)]^{-1}\phi_p(B)X_t = \epsilon_t$, Série finita em X_t
FAC	Descida a tender para zero (assumindo a forma de uma exponencial/sinusoide)	Rápida descida a tender para zero a partir de $K = q + 1$	Descida a tender para zero (assumindo a forma de uma exponencial/sinusoide)
FACP	Rápida descida a tender para zero a partir de $K = p + 1$	Descida a tender para zero (assumindo a forma de uma exponencial/sinusoide)	Descida a tender para zero (assumindo a forma de uma exponencial/sinusoide)

TABELA 2.3: Processos estacionários lineares (adaptado de Murteira et al., 2000)

2.9 Processos Não Estacionários e Lineares

Como foi referido no início deste capítulo, a maioria das séries temporais não são estacionárias. A não estacionariedade pode apresentar-se de várias formas

e convém analisar cada uma delas antes de descrever os processos adequados para a conversão em série estacionária. De acordo com Murteira et al. (2000), podemos dizer que um série não é estacionária quando apresenta algum dos seguintes comportamentos,

- tendência linear,

$$X_t = a + bt + Z_t, \quad (2.20)$$

onde Z_t é um processo estacionário;

- componente sazonal de periodicidade S ,

$$X_t = A \cos(2\pi t/S) + Z_t; \quad (2.21)$$

- tendência linear e componente sazonal na forma aditiva,

$$X_t = a + bt + A \cos(2\pi t/S) + Z_t; \quad (2.22)$$

- série com mudança no nível médio,

$$\begin{aligned} X_t &= Z_t, & t \leq t_0, \\ X_t &= Z_t + x, & t > t_0, \end{aligned} \quad (2.23)$$

- estacionariedade em média, mas não em variância,

$$\begin{aligned} X_t &\sim N(0, \sigma_0^2), & \sigma_0^2 &= x, & t &\leq t_0, \\ X_t &\sim N(0, \sigma_1^2), & \sigma_1^2 &\neq \sigma_0^2, & t &> t_0 \end{aligned}$$

Além disso, uma série temporal que não seja estacionária em média, também não será em variância e covariância (Murteira et al., 2000).

Como tal, quando se pretende estudar uma série temporal não estacionária, como é o caso dos registos de dados maregráficos, recorre-se a transformações que estabilizam a média e/ou variância, ou seja, transformando uma série não estacionária em estacionária (Murteira et al., 2000). Além destas duas transformações pode ainda ser necessária estabilizar a série temporal em sazonalidade. De seguida, apresentam-se os operadores que realizam as referidas estabilizações.

2.9.1 Estabilização de Média e Sazonalidade

A estacionariedade em média e sazonalidade é conseguida através dos seguintes métodos,

- Modelos paramétricos;
- Filtragem.

A construção de modelos paramétricos é mais propícia a caracterizar as diferentes componentes do que identificar o processo gerador, aspeto esse, mais importante. Por exemplo, as equações (2.20), (2.21) e (2.22) são exemplos de modelos paramétricos onde se nota que o principal propósito é a decomposição de cada uma das componentes (Murteira et al., 2000).

Os modelos paramétricos possuem essencialmente dois inconvenientes. Em primeiro lugar, admitem que a tendência é uma função do tempo, determinística, que permanece constante para qualquer valor de t , e como tal, as equações de regressão, podem ser pouco fiáveis para fazerem extrapolações. Em segundo lugar, não fazem uso da estrutura de dependência entre as observações em diferentes intervalos de tempo e, por isso, atribuem o mesmo peso a todas as observações na construção do modelo ou nas previsões (Murteira et al., 2000).

Devido a estas desvantagens, os modelos paramétricos têm sido substituídos por modelos não determinísticos através dos quais se atribui maior flexibilidade à função média para que se possa dar resposta às variações que a série temporal sofre ao longo do tempo. Desta forma, é possível aplicar uma transformação que estabilize a média, sem exigir compromissos com quaisquer modelos, motivo pelo qual, iremos apenas focar neste último método.

A estabilização da média é conseguida por meio de transformações que fazem parte dos chamados filtros lineares. Um filtro linear, de acordo com Murteira et al. (2000), tem a forma,

$$Y_t = \sum_{l=-\nu}^{\kappa} a_l X_{t-l}, \quad (2.24)$$

onde X_t é a série original, Y_t a série transformada e a_l os pesos. A escolha específica de diferentes pesos a_l e diferentes ν, κ é feita em função dos objetivos da operação e resulta em diferentes tipos de filtragem.

De seguida, apresentam-se dois operadores que permitem realizar dois tipos de filtragem, o operador diferença e o operador diferença sazonal.

Operador Diferença

Uma das formas mais usuais de estabilizar a média consiste em diferenciar a série até que ela se torne estacionária em média. A primeira diferença (ou diferença de ordem um) de X_t , de acordo com Murteira et al. (2000), é expressa por:

$$\nabla X_t = X_t - X_{t-1} = (1 - B)X_t \quad (2.25)$$

Poderá ser necessária aplicar diferenças de ordem superior a um. De modo geral, a diferença de ordem d de X_t é definida por:

$$\nabla^d X_t = (1 - B)^d X_t \quad (2.26)$$

A aplicação de diferenças de ordem um ou dois é geralmente suficiente para induzir a estacionariedade da série. De notar que, se a série transformada $\nabla^{d_0} X_t$ já está no estado estacionário então, para qualquer $d \geq d_0$, a série $\nabla^d X_t$ permanece estacionária, mas a variância aumenta com a ordem de grandeza da diferenciação. Como tal, deve evitar-se a sobrediferenciação para não se introduzir desnecessariamente variância excessiva na série temporal transformada (Murteira et al., 2000).

Operador Diferença Sazonal

De um modo análogo, a primeira diferença sazonal, ∇_S , conforme Murteira et al. (2000), define-se por:

$$\nabla_S X_t = X_t - X_{t-S} = (1 - B^S)X_t, \quad (2.27)$$

onde S corresponde à periodicidade da componente sazonal determinística.

Com a diferença de ordem D definida por:

$$\nabla_s^D X_t = (1 - B^s)^D X_t, \quad (2.28)$$

2.9.2 Estabilização da Variância

A diferenciação não elimina todos os casos de não estacionariedade. De facto, uma parte significativa das séries temporais, independentemente de serem ou não estacionárias em média, não o são em variância. Este problema é solucionado através da aplicação da transformação de *Box-Cox* que, segundo Murteira et al. (2000), segue o formalismo,

$$\begin{cases} T(X_t) = X_t^{(\lambda)} = \frac{X_t^\lambda - 1}{\lambda}, & \text{se } \lambda \neq 0, \\ T(X_t) = \ln X_t, & \text{se } \lambda = 0, \end{cases} \quad (2.29)$$

onde o parâmetro λ é escolhido dentro do intervalo $[-1,1]$. Os valores mais usuais na prática da análise de séries temporais correspondem às transformações indicadas na Tabela 2.4

Valores de λ	Transformações
-1.0	$1/X_t$
-0.5	$1/\sqrt{X_t}$
0.0	$\ln X_t$
0.5	$\sqrt{X_t}$
1.0	X_t

TABELA 2.4: Parâmetros de λ (Murteira et al., 2000)

Conforme Murteira et al. (2000), a correspondência entre as transformações de *Box-Cox* e logarítmica, resulta do limite seguinte:

$$\lim_{\lambda \rightarrow 0} \frac{X_t^\lambda - 1}{\lambda} = \ln X_t$$

A transformação da variância deve ser a primeira a ser aplicada visto que, a determinação de diferenças pode induzir variação nas observações o que dificultaria ainda mais a escolha da transformação adequada para estabilizar a variância. Só depois desta transformação, e caso a série transformada apresente tendência e/ou sazonalidade, se deve proceder à diferenciação e à diferenciação sazonal, aplicando a diferenciação até à ordem que se julgue necessário (Murteira et al., 2000).

A transformação só pode ser aplicada a séries positivas. Caso não se verifique esta condição, deverá ser adicionada uma constante positiva para evitar este

problema. A aplicação da constante não altera a estrutura variância-covariância (Murteira et al., 2000).

A informação contida na tabela 2.5 resume as várias transformações que podem ser aplicadas às séries temporais em situação de não estacionariedade. É de notar que a aplicação de cada uma das estabilizações apenas deve ser realizada se necessário:

Transformação em:	Expressão matemática
Variância	$\begin{cases} T(X_t) = X_t^{(\lambda)} = \frac{X_t^\lambda - 1}{\lambda}, & \text{se } \lambda \neq 0 \\ T(X_t) = \ln X_t, & \text{se } \lambda = 0 \end{cases}$
Média	$\nabla^d X_t = \nabla[\nabla^{d-1} X_t] = (1 - B)^d X_t$
Sazonalidade	$\nabla_s^D X_t = (1 - B^s)^D X_t$

TABELA 2.5: Transformações a aplicar a série não estacionária

2.9.3 Modelos Integrados Mistos Não Sazonais e Sazonais (ARIMA e SARIMA)

Depois de se verificar que séries não estacionárias podem ser convertidas, através da diferenciação, em estacionárias, convém introduzir os modelos ARIMA. Estes modelos resultam da combinação de três componentes: componente autoregressiva (AR), diferenciação e componente médias móveis (MA).

De acordo com Murteira et al. (2000), um processo X_t diz-se um ARIMA(p,d,q) se,

$$\phi_p(B)(1 - B)^d X_t = \theta_0 + \theta_q(B)\epsilon_t, \quad (2.30)$$

onde, $\phi_p(B) = 1 - \phi_1 B - \dots - \phi_p B^p$ é o polinómio autoregressivo de ordem p e $\theta_q(B) = 1 - \theta_1 B - \dots - \theta_q B^q$ é o polinómio de médias móveis de ordem q .

Se $d = 0$, θ_0 está relacionado com a média, $\theta_0 = \mu(1 - \phi_1 - \dots - \phi_p)$. Quando $d \geq 1$, representa o coeficiente da tendência determinística (Murteira et al., 2000).

O processo X_t , que numa primeira fase é não estacionário, depois de diferenciado d vezes converte-se num processo estacionário e invertível ARMA(p,q) (Murteira et al., 2000).

Os modelos ARIMA são de grande utilidade para modelar valores vizinhos na série, $X_{t-p}, X_{t-p+1}, \dots, X_t$. Contudo, se houver uma componente cíclica S , com periodicidade sazonal, é de esperar que $X_{t-2S}, X_{t-S}, \dots, X_t$ estejam também relacionados. Nestas situações, a escolha adequada dos operadores de diferenciação ∇^d e ∇_S^D é fundamental para assegurar que a série se converta num processo estacionário. Como tal, precisamos de um novo processo que tenha em conta os dois operadores referidos, surgindo assim os processos SARIMA.

Murteira et al. (2000) define que um processo X_t comporta-se como um SARIMA (p,d,q) $\times$ (P,D,Q) $_S$ se,

$$\phi_p(B)\Phi_P(B^S)(1-B)^d(1-B^S)^D X_t = \theta_q(B)\Theta_Q(B^S)\epsilon_t, \quad (2.31)$$

2.10 Modelação ARIMA de Séries Temporais

Segundo Murteira et al. (2000), para se obter um “bom modelo” de *Box Jenkins* deve-se adotar a seguinte metodologia constituída por três etapas:

Identificação $\rightarrow$ Estimação $\rightarrow$ Avaliação do diagnóstico

2.10.1 Identificação

Como, em geral, as séries temporais não se encontram no estado estacionário e os modelos ARMA(p,q) correspondem a processos estacionários, torna-se necessário realizar operações sobre as séries de modo que as séries transformadas possam ser descritas por tais modelos. Assim sendo, esta primeira fase, visa realizar transformações, para estabilização da variância e eliminação da tendência, assim como retirar o efeito da sazonalidade. Além disso, após atingida a estacionariedade da série, procede-se à escolha dos valores adequados para d , S , D e para p , q (Santos, 2011).

As várias técnicas usadas nesta fase são essencialmente exploratórias e dependem na sua maioria de métodos gráficos, comparação dos correlogramas das FAC e FACP. Consequentemente, torna-se mais relevante eliminar modelos implausíveis

do que escolher o melhor modelo (Murteira et al., 2000). Na fase da identificação é muito importante a experiência do analista para a determinação do modelo adequado.

Na identificação do modelo devem ser tidos em conta os seguintes passos (Murteira et al., 2000; Santos, 2011):

Passo 1 – Calcular e analisar as FAC e FACP por forma a proceder à confirmação da necessidade de aplicar operações de modo a tornar a série estacionária. Assim, apresentam-se algumas regras que devem ser consideradas:

1. Se a FAC descreve um lento decaimento para zero e a FACP se anula depois de “lag 1”, a série deve ser diferenciada para eliminar a tendência;
2. Numa situação em que estamos perante uma série (constituída por valores mensais) que não é estacionária em média, devido à existência de uma tendência, com uma aproximação suave para zero sobre os “lag’s” múltiplos de 12 e onde se note ainda a existência de movimentos periódicos, sugere-se a aplicação do operador diferença sazonalidade de ordem um;
3. Por norma, o primeiro ou segundo grau de diferenciação são suficientes para conseguir a estacionariedade da série.

Passo 2 - Garantir a estacionariedade da série, através de sucessivas transformações, até que se consiga estabilizar a variância, e eliminar a tendência e a sazonalidade. Como já definidas na secção anterior, as transformações mais usuais são:

1. A transformação *Box-Cox*;
2. O operador diferença simples de potência d . Normalmente, $d \leq 2$ é suficiente;
3. O operador diferença sazonal de potência D . Normalmente $D = 1$ é suficiente.

Passo 3 – Após efetuadas as transformações necessárias à estacionariedade da série, são calculadas e examinadas as FAC e FACP das séries transformadas por forma a identificar os operadores p e q .

De notar que a identificação de modelos AR e MA é relativamente fácil visto que estão associados, respetivamente, a um decaimento muito rápido para zero das FACP e da FAC. No entanto, nos modelos ARMA, devido ao decaimento lento para zero em ambas as funções, torna-se difícil a identificação das ordens p e q .

Na identificação de modelos deve-se apenas prestar atenção ao comportamento geral da série, uma vez que os modelos podem ser melhorados numa fase posterior.

2.10.2 Estimação

Após identificados os modelos procede-se à estimação dos parâmetros, $(\phi_1 \dots \phi_p)$, $(\theta_1 \dots \theta_q)$, $\sigma_\epsilon^2 = v(\epsilon_t)$ do modelo ARMA(p, q),

$$X_t^* - \phi_1 X_{t-1}^* \dots - \phi_p X_{t-p}^* = \epsilon_t - \theta_1 \epsilon_{t-1} \dots - \theta_q \epsilon_{t-q},$$

onde $X_t^* = X_t - \mu$ e X_t são observações correspondentes à série temporal já regularizada (Murteira et al., 2000).

De acordo com Murteira et al. (2000), tratando-se de sistemas de equações não lineares, são utilizados métodos de análise numérica para a sua resolução; os estimadores são a solução desse sistema de equação não lineares.

Devido à existência de pacotes informáticos que fornecem, depois de identificados os modelos, as estimativas dos respetivos parâmetros, esta etapa, não exige grande intervenção por parte de quem está a analisar as séries. Os métodos mais utilizados para a estimação dos modelos são o método da máxima verosimilhança, o método dos mínimos quadrados, e o método dos momentos (Murteira et al., 2000).

A metodologia utilizada para a obtenção dos parâmetros para os processos já abordados pode ser encontrada em Murteira et al. (2000), páginas 115-120.

2.10.3 Métodos de Diagnóstico dos Modelos

Após a identificação de um modelo e a determinação dos respetivos parâmetros chega-se à etapa em que se procura analisar a qualidade estatística do modelo e do ajustamento. Se o modelo selecionado não estiver de acordo com os pressupostos que serão apresentados de seguida, deve repetir-se o ciclo, retrocedendo à fase inicial, numa nova tentativa de identificar um novo modelo que descreva a série em causa de forma satisfatória.

Os modelos de regressão e modelos que combinam tendência linear e tendência cíclica também devem verificar as qualidades estatísticas descritas a seguir. Caso os modelos não estejam em conformidade, também devem ser excluídos do estudo.

Avaliação da Qualidade Estatística dos Modelos

Existem vários métodos possíveis para avaliar a qualidade estatística do modelo, no entanto, neste trabalho, vamos apenas apresentar os dois mais conhecidos.

1. **Significância estatística dos parâmetros estimados.** Procurando um modelo equilibrado, devem ser eliminados os parâmetros considerados desnecessários, ou seja, eliminar todos aqueles que sejam significativamente iguais a zero.

Em geral, modelos com menor número de parâmetros produzem melhores previsões (Murteira et al., 2000).

Tem, por isso, interesse, aplicar um teste univariado para aferir significância estatística de cada coeficiente do modelo, onde se consideram as seguintes hipóteses,

$$H_0 : \beta_i = 0 \text{ vs } H_1 : \beta_i \neq 0,$$

onde β_i representa cada um dos parâmetros $(\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q)$ do modelo.

De acordo com Santos (2011) a estatística de teste é dada por,

$$T = \sqrt{n} \frac{\hat{B}_i}{\sqrt{\hat{Var}[\hat{B}_i]}} \sim t_{n-p-q}$$

isto é, T segue uma distribuição $t - Student$ com $n - p - q$ graus de liberdade. Rejeita-se H_0 considerando um nível de significância α sempre que $[T] > t_{n-q-p, \frac{\alpha}{2}}$, com o $[T] > t_{n-q-p, \frac{\alpha}{2}}$ o quantil de probabilidade $1 - \frac{\alpha}{2}$ da distribuição de $t - Student$, com $n-q-p$ graus de liberdade. Quando rejeitamos H_0 , o parâmetro mantém-se no modelo (Murteira et al., 2000).

Em alternativa ao teste de hipótese, pode ser utilizado o intervalo de confiança para avaliar a significância do parâmetro. Para cada parâmetro, são definidas duas estatísticas, o limite inferior (LI) e o limite superior (LS), que delimitam, com um determinado nível de confiança $(1 - \alpha)100\%$, o intervalo onde o verdadeiro valor do parâmetro se encontra (Murteira e Antunes, 2012; Murteira, 1990). Conforme Murteira (1990), o intervalo de um determinado parâmetro (ζ), a $(1 - \alpha)100\%$ de confiança, é dado por,

$$LI \leq \zeta \leq LS.$$

Se um número grande de intervalos for construído, cada um resultante de um conjunto de dados, é expectável que $(1 - \alpha)100\%$ deles contenha o verdadeiro valor de ζ (Murteira, 1990). Para construir um intervalo de confiança é necessário definir um estimador do parâmetro, conhecer a respetiva distribuição e determinar uma estimativa pontual do parâmetro (Z. Pereira e Requeijo, 2012).

As distribuições amostrais anteriormente apresentadas desempenham aqui um papel fundamental. Para detalhar a estimação por intervalos de confiança, podem ser consultados os trabalhos de Murteira (1990) e Velosa e Pestana (2010).

2. Estacionariedade e invertibilidade do modelo

A Tabela 2.6, adaptada de Murteira et al. (2000), apresenta as condições que os modelos devem verificar:

	AR(p)	MA(q)	ARMA(p, q)
Condições de Estacionariedade	Raízes de $\phi_p(B) = 0$ fora do círculo unitário	Sempre estacionários	Raízes de $\phi_p(B) = 0$ fora do círculo unitário
Condições de invertibilidade	Sempre invertíveis	Raízes de $\theta_q(B) = 0$ fora do círculo unitário	Raízes de $\theta_q(B) = 0$ fora do círculo unitário

TABELA 2.6: Condições de Estacionariedade e invertibilidade (adaptado de Murteira et al. (2000))

Avaliação da Qualidade de Ajustamento

A qualidade do ajustamento de um modelo estimado pode ser medida através da análise dos seus resíduos (Murteira et al., 2000). Considere-se o modelo ARMA(q, p) no qual as ordens q e p foram identificadas,

$$\phi(B)X_t = \theta(B)\epsilon_t,$$

e se $\hat{\phi} = (\hat{\phi}_1, \dots, \hat{\phi}_p)$ e $\hat{\theta} = (\hat{\theta}_1, \dots, \hat{\theta}_q)$ representam os estimadores de máxima verossimilhança, podemos definir os resíduos ϵ_t por,

$$\hat{\epsilon}_t = \hat{\theta}^{-1}(B)\hat{\phi}(B)(X_t - \hat{\mu}),$$

onde $\hat{\theta}(B) = (1 - \hat{\theta}_1 B - \dots - \hat{\theta}_q B^q)$ e $\hat{\phi}(B) = (1 - \hat{\phi}_1 B - \dots - \hat{\phi}_p B^p)$ são, respetivamente, os valores estimados dos coeficientes dos modelo.

Segundo Murteira et al. (2000), nestas condições o modelo estimado descreve bem a série temporal caso se verifique que à medida que o número de valores da série aumenta, o comportamento dos resíduos aproxima-se do comportamento do ruído branco. Como tal, a análise dos resíduos pode ser interpretada como uma medida de avaliação da adequação do modelo. Pode admitir-se que, modelos cujos resíduos tenham um comportamento análogo a um ruído branco são adequados para descrever a série em estudo (Murteira et al., 2000).

O comportamento dos resíduos pode ser avaliado através dos respetivos valores estimados das funções FAC e FACP. Designamos por $\hat{\rho}_k(\hat{\epsilon})$ e $\hat{\phi}_k(\hat{\epsilon})$, respetivamente, as FAC e FACP dos resíduos.

De acordo com Murteira et al. (2000), o teste estatístico de Ljung e Box permite avaliar a qualidade do ajustamento tendo por base as $\hat{\rho}_k(\hat{\epsilon})$ e $\hat{\phi}_k(\hat{\epsilon})$. Para y “lag” considerados, $\hat{\rho}_k, k = 1, \dots, y$, testa a hipótese,

$$H_0 : \rho_1(\hat{\epsilon}) = \rho_2(\hat{\epsilon}) = \dots = \rho_y(\hat{\epsilon}) = 0.$$

O teste de Ljung-Box, segundo Murteira et al. (2000), é baseado na estatística,

$$Q = N(N+2) \sum_{k=1}^y \frac{1}{N-k} \hat{\rho}_k^2(\hat{\epsilon}), \quad N = n - d, \quad (2.32)$$

que converge também, mas mais rapidamente, para uma distribuição Qui-quadrado (χ_{m-p-q}^2). Este teste torna-se conveniente quando a amostra não é muito grande (Murteira et al., 2000).

A rejeição da hipótese nula indica que as autocorrelações são significativamente diferentes de zero, o que leva a concluir que os resíduos não são aleatórios e independentes ao longo do tempo. Portanto, a rejeição da hipótese nula sugere que o modelo estimado não é apropriado para descrever a série temporal (Murteira et al., 2000).

2.11 Critérios de Seleção de Modelos

Na prática, o analista encontra vários modelos satisfatórios, sejam eles modelos ARIMA, polinomiais ou de tendência cíclica. Nestas situações deve ser usado

um conjunto de critérios que permitem identificar o melhor modelo que descreve o comportamento dos dados em análise. Neste trabalho adotou-se a seguinte metodologia de seleção do modelo:

- Comparar dois ou mais modelos encaixados (modelos com diferentes quantitativos de parâmetros), para saber se o modelo com mais parâmetros é estatisticamente superior aos modelos com menor número de parâmetros. Se não for, adotamos o mais simples. Os testes de Análise de variância (ANOVA) e de *Wilcoxon* podem ser utilizados para realizar esta inferência estatística;
- Caso se pretenda selecionar o melhor modelo entre modelos ARIMA, deve recorrer aos critérios AIC e BIC, como auxiliares de análise. Estes critérios permitem estabelecer um equilíbrio entre ajustamento e número de parâmetros.
- Por último, recorre-se ao coeficiente de determinação (R^2) e ao *Root Mean Square Error* (*RMSE*) para avaliar se os modelos encontrados são adequados.

2.11.1 Análise de Variância (ANOVA) e Teste de *Wilcoxon*

A ANOVA é uma técnica estatística que permite verificar a igualdade estatística entre médias populacionais, ou seja, a análise de variância compara médias de diferentes populações para verificar se essas populações possuem médias iguais ou diferentes (Nunes e Afonso, 2005). Esta técnica baseia-se na comparação da variabilidade em torno de cada uma das médias amostrais e a variabilidade entre as médias amostrais (Nunes e Afonso, 2005).

Face a outras técnicas de verificação da igualdade estatística entre diferentes médias populacionais, a ANOVA tem a vantagem de permitir verificar igualdades em mais do que duas médias populacionais ao mesmo tempo.

Segundo Nunes e Afonso (2005) o teste ANOVA pode ser utilizado quando se verificam os seguintes pressupostos:

- As amostras devem ser aleatórias e independentes;
- As populações seguem a distribuição Normal;
- As populações apresentam a mesma variância.

As hipóteses a testar, de acordo com Nunes e Afonso (2005), são,

$$H_0 : \mu_1 = \mu_2 = \dots = \mu_a \quad vs. \quad H_1 : \exists_{i \neq j} : \mu_i \neq \mu_j \quad para \quad i, j = 1, 2, \dots, a$$

ou equivalentemente

$$H_0 : \alpha_1 = \alpha_2 = \dots = \alpha_\alpha = 0 \quad vs. \quad H_1 : \text{Existe pelo menos um } \alpha_i \neq 0$$

Conforme Nunes e Afonso (2005), a estatística de teste é dada por,

$$F = \frac{MQA}{MQE} = \frac{\frac{SQA}{\alpha-1}}{\frac{SQE}{n-\alpha}} \sim F_{\alpha-1; n-\alpha}, \quad (2.33)$$

onde,

$SQA = \sum_{i=1}^a n_i (\bar{X}_i - \bar{\bar{X}})^2$, sendo $\bar{\bar{X}}$ a média global e $\bar{X}_i$ a média do i ;

$SQE = \sum_{i=1}^a (n_i - 1) S_i^2$, sendo S_i a variância de nível i ;

$SQT = (N - 1) S^2$ sendo S_i a variância global.

Rejeita-se H_0 para o nível de significância α sempre que $p\text{-value} < \alpha$. Para rejeitar H_0 é suficiente que duas médias sejam diferentes. Ao rejeitar H_0 , apenas podemos concluir que não existe igualdade entre as médias populacionais.

Quando não se verificam os pressupostos do teste de ANOVA, o teste não paramétrico de *Wilcoxon* constitui-se uma alternativa. Este teste tem como objetivo avaliar se duas amostras têm a mesma mediana.

Segundo Nunes e Afonso (2005), para cada par (X_i, Y_i) de observações da amostra, é calculada a diferença $D_i = X_i - Y_i$ retendo-se o resultado final da operação (+ ou -). De notar que, quando a diferença é nula, ou seja, quando X_i for igual a Y_i , deve-se eliminar o valor da diferença e corrigir a dimensão da amostra.

Seguidamente, os valores $|D_1|, |D_2|, \dots, |D_n|$ são ordenados por ordem crescente, associando-se o sinal + ou - se D_i for, respetivamente, positivo ou negativo (Nunes e Afonso, 2005).

As hipóteses a testar, de acordo com Nunes e Afonso (2005), são,

$$H_0 : \tilde{\mu}_1 = \tilde{\mu}_2 \quad vs. \quad H_1 : \tilde{\mu}_1 \neq \tilde{\mu}_2$$

A estatística de teste segue a seguinte forma (Nunes e Afonso, 2005),

- Para amostras pequenas ($n \leq 30$),

$$W = \text{Soma das ordens dos } |D_i| \text{ com sinal } +, \quad (2.34)$$

onde os pontos críticos podem ser consultados em tabela própria (consultar em Afonso e Nunes (2019), página 395).

- Para amostras grandes ($n > 30$),

$$Z = \frac{W - \frac{n'(n'+1)}{4}}{\sqrt{\frac{n'(n'+1)(2n'+1)}{24}}} \sim N(0; 1), \quad (2.35)$$

onde n' representa a dimensão da amostra corrigida, ou seja, $n' = n - n^\circ$ de diferenças nulas.

2.11.2 Critérios AIC e BIC

Estes critérios baseiam-se em estatísticas obtidas a partir dos resíduos correspondentes ao ajustamento de um modelo, ou a partir dos desvios nas suas previsões (Murteira et al., 2000).

Critério AIC

Akaike (1974), introduziu o critério *Akaike Information Criterion*, AIC, para medir a qualidade dos ajustamentos. Conforme Murteira et al. (2000), o critério AIC, penaliza os modelos com maior variância e maior número de parâmetros, ou seja, representa um compromisso entre qualidade do ajustamento e parcimónia. O modelo mais adequado está associado ao melhor compromisso entre o ajustamento conseguido e o número de coeficientes usados para especificar o modelo. Como regra de decisão, deve ser escolhido o modelo que apresente o menor valor AIC (Murteira et al., 2000).

De notar que este critério apenas pode ser utilizado para comparar modelos que tenham sido ajustados à mesma série estacionarizada, isto é, apenas se pode aplicar a modelos que apliquem as mesmas transformações à serie (Murteira et al., 2000).

O formalismo matemático, inerente ao critério AIC, pode ser consultado em Murteira et al. (2000), páginas 133-135.

Critério BIC

O mesmo autor do critério anterior, desenvolveu mais tarde o critério *Bayesian Information Criterion*, BIC. Este critério consiste numa extensão do critério

AIC que, conforme Murteira et al. (2000), permite aumentar a penalização ao número elevado de parâmetros do modelo, de maneira que, a minimização de BIC resulte, em geral, na seleção de modelos de menor ordem relativamente aos obtidos pela minimização do AIC (Murteira et al., 2000).

O formalismo matemático, inerente ao critério BIC, pode ser consultado em Murteira et al. (2000), página 135.

2.11.3 Comparação de Resíduos

Coefficiente de Determinação

O coeficiente de determinação, R^2 , é utilizado na avaliação do ajustamento. O R^2 toma valores entre 0 e 1 e representa a proporção da variância que é explicada pelo modelo (Maroco, 2007). Se o R^2 for próximo de um, o modelo explica quase toda a variância da série temporal. Se for próximo de zero, o modelo não é adequado para explicar a variância da série. O coeficiente de determinação é dado, de acordo com Cryer e Chan (2008) e Quinino et al. (1991), pela equação,

$$R^2 = \frac{\sum_{i=1}^n (\hat{X}_i - \bar{X})^2}{\sum_{i=1}^n (X_i - \bar{X})^2} = 1 - \frac{\sum_{i=1}^n e_i^2}{\sum_{i=1}^n (X_i - \bar{X})^2}, \quad (2.36)$$

onde $\hat{X}_i$ corresponde ao polinómio estimado e $e_i = X_i - \hat{X}_i$ é o resíduo de ordem i .

Root Mean Square Error (RMSE)

O *RMSE* representa o desvio médio entre observações e previsões. Caracteriza-se por possuir dimensão igual à dimensão dos valores observados e previstos (Willmott, 1981). De acordo com Willmott (1981) o *RMSE* é dado por:

$$RMSE = \left[\frac{\sum_{i=1}^n (O_i - P_i)^2}{n} \right]^{\frac{1}{2}}, \quad (2.37)$$

onde O_i e P_i são, respetivamente, o valor observado e valor previsto de ordem i .

O *RMSE* é uma importante estatística de validação uma vez que dá indicação sobre o tamanho real do erro produzido, contrariando o que se sucede com o coeficiente de determinação, onde erros de elevada magnitude podem ser encobertos por um desvio padrão elevado ou pequenos erros podem tornar-se significativos quando o desvio padrão é reduzido (Willmott, 1981).

Capítulo 3

Análise das Séries de Dados Maregráficos

Pretende-se estudar a recente evolução do NMM na costa portuguesa, porém, é impraticável realizar o estudo para todos os portos nacionais. Para mitigar este problema, optou-se pela seleção de três portos de elevada relevância estratégica e económica e, ao mesmo tempo, geograficamente distribuídos pelo território nacional: porto de Sines (região sul); porto de Viana do Castelo (região norte); e porto do Funchal (Região Autónoma da Madeira).

A seleção do porto de Sines deve-se à sua importância a nível nacional e internacional. O porto de Sines apresenta-se como sendo o grande *hub*⁷ da fachada Euro-Atlântica, com condições únicas na costa portuguesa que lhe permitem acolher qualquer tipo de navio. É considerado líder nacional na quantidade de mercadorias movimentadas, sendo a principal porta de abastecimento energético do país (Cacho, 2020). A nível económico, o porto de Sines tem um peso de 1,5% na economia nacional, podendo-se considerar um importante ativo (Cacho, 2020). Além disso, apresenta um elevado valor estratégico tendo em conta a atual situação geopolítica da Europa.

Atualmente, o porto de Viana do Castelo está dotado de excelentes infraestruturas, distinguindo-se a nível nacional pelo sua localização estratégica, localizado numa das regiões mais industrializadas do país. Simultaneamente, apresenta-se como um importante *hub* na área da arquitetura naval a nível nacional e ibérico.

O porto do Funchal, líder nacional no ano 2020 no que toca ao movimento de passageiros, tem-se apresentado como sendo um dos principais motores económicos

⁷ponto central de uma atividade.

da Região Autónoma da Madeira (S. Pereira, 2020). Paralelamente, o estudo da evolução do NMM no porto do Funchal é fundamental para compreender a volatilidade da tendência de subida do NMM entre regiões oceânicas e a orla costeira.

A caracterização da atual subida é fundamental para que as entidades competentes consigam tomar as medidas necessárias para que num futuro a médio e longo prazo se minimize, ou mesmo, evite os impactos negativos desta problemática, por isso, o estudo detalhado da evolução do NMM neste três portos está revestido de elevada importância.

Parte dos resultados que a seguir se apresentam, já fizeram parte dos trabalhos apresentados em Picoto et al. (2022a) e Picoto et al. (2022b). O estudo foi realizado tendo por base a aplicação das metodologias apresentadas no capítulo anterior através de bibliotecas do *software* estatístico **R** (R Core Team, 2021). No presente trabalho, recorreram-se às seguintes bibliotecas: “**Forecast**” (Hyndman et al., 2022a), “**Astsa**” (Hyndman et al., 2022b), “**DescTools**” (Andri et mult. al., 2021), “**Xts**” (Ryan e Ulrich, 2020), “**Lubridate**” (Grolemund e Wickham, 2011), “**TSA**” (Chan e Ripley, 2020), “**Ggplot2**” (Wickham, 2016), “**Metrics**” (Hamner e Frasco, 2018), “**Openxlsx**” (Schauberger e Walker, 2021). Os dados, utilizados nesta dissertação, foram disponibilizados pela orientadora que desenvolve atividade na Divisão de Oceanografia do IH.

3.1 Análise dos Registos Maregráficos de Sines

3.1.1 Análise Descritiva dos Dados

A série dos dados é constituída por 288 observações do nível médio mensal (MNM), obtidas a partir das alturas horárias registadas pela estação maregráfica do porto de Sines, Figura 3.1. Os dados analisados referem-se ao período compreendido entre 1997 e 2020.

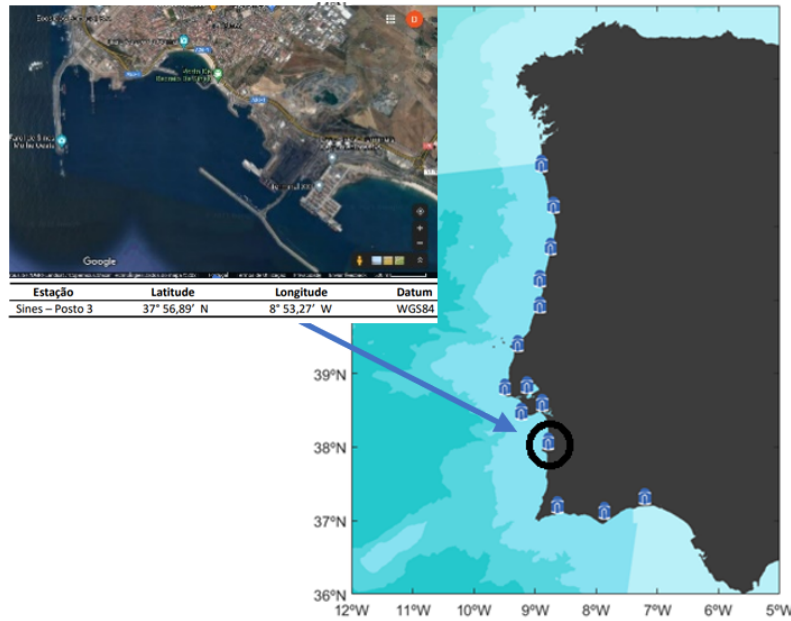


FIGURA 3.1: Localização da estação maregráfica do porto de Sines ($\varphi = 37^{\circ}56,89'N$ e $\lambda = 008^{\circ}53,27'W$) (fonte: IH, 2020)

O histograma da Figura 3.2, apresenta uma ligeira assimetria negativa (coeficiente de assimetria igual a -0,070, aproximadamente) demonstrando que a distribuição dos NMm's se concentra no lado direito da média = 2,198 m com Intervalo de Confiança (95%) igual a (2,191; 2,206) com uma ligeira cauda para o seu lado esquerdo (Picoto et al., 2022a e Picoto et al., 2022b).

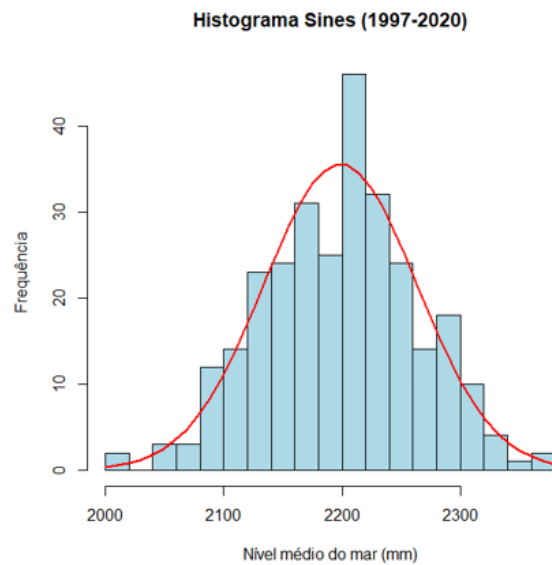


FIGURA 3.2: Histograma das frequências dos NMm em Sines no período 1997 - 2020 (Picoto et al., 2022a e Picoto et al., 2022b)

A hipótese de normalidade dos dados foi testada através do teste de *Shapiro-Wilk*, onde as hipóteses são dadas por (Shapiro e Wilk, 1965):

$$\begin{cases} H_0 = \text{os NMm's seguem a distribuição normal} \\ H_1 = \text{os NMm's não seguem a distribuição normal} \end{cases}$$

Obteve-se o valor da estatística de teste $W = 0,9976$ com um valor associado de $p - \text{value} = 0,9516 > \alpha = 0,05$, concluindo-se que existe evidência estatística de que os NMm's seguem uma distribuição normal

A Tabela 3.1 apresenta as principais estatísticas descritivas dos NMm's (1997-2020) (Picoto et al., 2022a e Picoto et al., 2022b):

<i>NMm's (mm)</i>	
Média	2198
Mediana	2205
Desvio-padrão	64,88
Assimetria	-0,07
Mínimo	2007
Máximo	2379
Dimensão da amostra	288

TABELA 3.1: Estatísticas descritivas dos NMm's (mm) de Sines

3.1.2 Avaliação da Tendência e Sazonalidades

A Figura 3.3, apresenta a série temporal dos NMm's observados em Sines. Foi ajustado um modelo de regressão polinomial de grau um ($X_t = \beta_0 + \beta_1 t$) aos NMm's, de tendência β_1 .

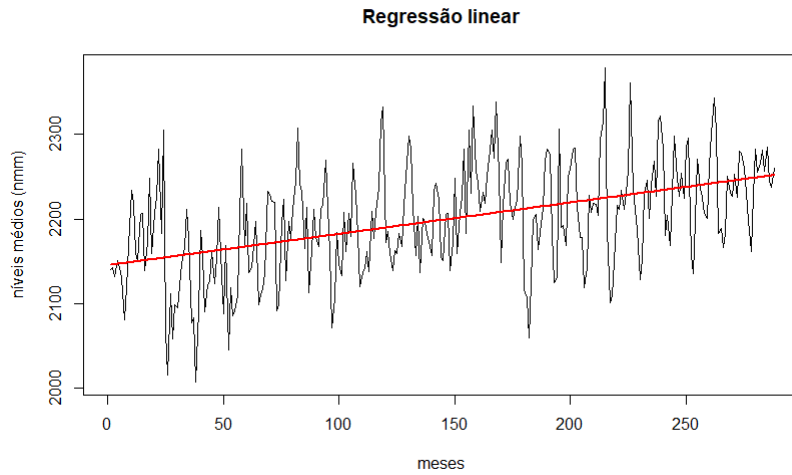


FIGURA 3.3: Linha preta: evolução dos NMm's em Sines no período em análise; linha vermelha: modelo ajustado (Picoto et al., 2022a e Picoto et al., 2022b)

A Figura 3.4 apresenta os coeficientes e os respetivos p -value, obtidos pelo *software R*.

```

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  2.145e+03  6.731e+00  318.671  <2e-16 ***
X             3.699e-01  4.038e-02   9.161  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

FIGURA 3.4: Coeficientes do modelo polinomial de grau um estimado

A hipótese de que os coeficientes são significativos foi confirmada através do teste de significância estatística, onde as hipóteses são dadas por:

$$\begin{cases} H_0 = \text{coeficientes nulos} \\ H_1 = \text{coeficientes não nulos} \end{cases}$$

Obteve-se $p\text{-value} \leq 0,05$, pelo que rejeitamos H_0 a favor de H_1 e assim, conclui-se que os coeficiente são significativos.

Face ao declive positivo, $\beta_1 = 3,699 \times 10^{-1}$, pode-se concluir que o NMM está a aumentar.

Ajustou-se um polinómio de grau dois ($X_t = \beta_0 + \beta_1 t + \beta_2 t^2$), com vista a determinar se a tendência sofreu aceleração (Figura 3.5).

```

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  2.137e+03  1.014e+01 210.779 < 2e-16 ***
X            5.419e-01  1.620e-01   3.345 0.000932 ***
I(X^2)       -5.951e-04  5.428e-04  -1.096 0.273853
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

FIGURA 3.5: Coeficientes do modelo polinomial de grau dois estimado

Obteve-se $p - value = 0,2738 > 0,05$, conclui-se que o coeficiente β_2 não é significativo e como tal, não se verifica aceleração da tendência da subida no período de 1997 a 2020 (Figura 3.5).

A hipótese de que os resíduos dos dois modelos de tendência polinomial seguem a distribuição normal foi confirmada através do teste de *Shapiro-Wilk*; a hipótese de normalidade não foi rejeitada ao nível de significância de 5% ($W = 0,995$ com $p - value = 0,628 > 0,05$ e $W = 0,996$ com $p - value = 0,812 > 0,05$, respetivamente, para o modelo polinomial de grau um e grau dois). A normalidade dos dois modelos indica-nos que o teste de ANOVA pode ser aplicado.

Através do teste de ANOVA procura-se comparar entre duas médias populacionais; as hipóteses a testar são:

$$\begin{cases} H_0 = \text{os resíduos possuem médias iguais} \\ H_1 = \text{os resíduos não possuem médias iguais} \end{cases}$$

Verifica-se que os modelos são semelhantes ($F = 1,202$ com um valor de $p - value = 0,273 > 0,05$); não rejeitamos H_0 . Como tal, o modelo com mais parâmetros (polinómio de grau dois) não acrescenta valor relativamente ao polinómio mais simples.

O facto de o coeficiente β_2 não ser significativo e de os dois modelos não apresentarem diferenças, estatisticamente, significativas permite concluir que não se evidenciou aceleração da tendência do NMM.

Doravante, passaremos a referir-nos ao modelo polinomial de grau um e grau dois, respetivamente, por modelo 1 e modelo 2.

A Figura 3.6, apresenta as distribuições do NMm's registados pelo marégrafo de Sines entre 1997 e 2020:

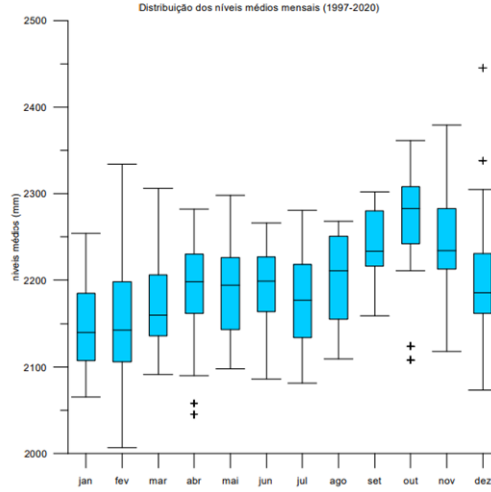


FIGURA 3.6: Comportamento dos níveis médios ao longo do ano em Sines no período 1997 a 2020 (Picoto et al., 2022a e Picoto et al., 2022b)

Da análise da figura anterior, a variação do NMM aparenta ter uma componente sazonal muito forte ao longo do ano. Por este motivo, recorreu-se a modelos que combinam tendência polinomial com tendência cíclica, onde se assume que existe um número de sinusoides que descrevem as variações sazonais (Picoto et al., 2022a e Picoto et al., 2022b). Surgem assim duas hipóteses:

- um modelo que considera apenas sazonalidade anual (uma senoide), com período igual a 12 meses, que segue a forma:

$$X_t = \beta_0 + \beta_1 t + a_1 \cos(\omega_1 t) + b_1 \sin(\omega_1 t),$$

onde $\omega_1 = \frac{2\pi}{p_1}$, sendo o período $p_1 = 12$.

Doravante, passaremos a referir-nos a este modelo por modelo 3.

- um modelo que considera sazonalidade anual e semianual (duas sinusoides), respetivamente, com períodos iguais a 12 e 6 meses, que segue a forma:

$$X_t = \beta_0 + \beta_1 t + a_1 \cos(\omega_1 t) + b_1 \sin(\omega_1 t) + a_2 \cos(\omega_2 t) + b_2 \sin(\omega_2 t),$$

onde $\omega_1 = \frac{2\pi}{p_1}$, com $p_1 = 12$; e $\omega_2 = \frac{2\pi}{p_2}$, com $p_2 = 6$.

Doravante, passaremos a referir-nos a este modelo por modelo 4.

A Figura 3.7 apresenta o modelo 3 ajustado à série de dados,

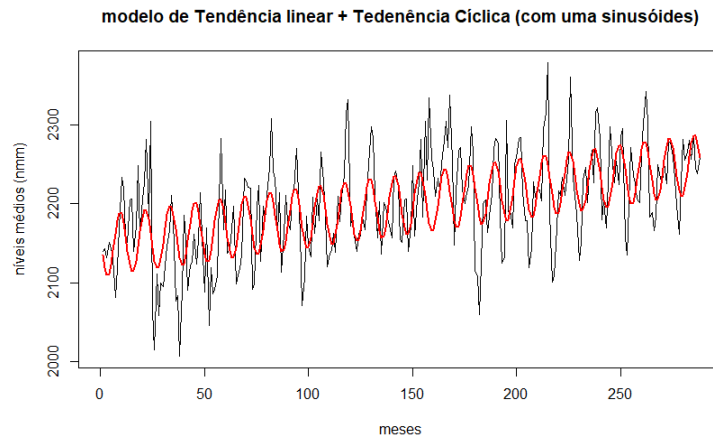


FIGURA 3.7: Linha preta: série de dados em análise; linha vermelha: modelos ajustado (modelo 3)

A Figura 3.8, apresenta os coeficientes e os respetivos p -values encontrados,

```

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    2146.6719     5.9171  362.789  <2e-16 ***
X               0.3591      0.0355   10.115  <2e-16 ***
I(cos(2 * pi/12 * X))  8.3258     4.1716    1.996  0.0469 *
I(sin(2 * pi/12 * X)) -37.8987     4.1736   -9.081  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
    
```

FIGURA 3.8: Coeficientes do modelo do modelo 3

Para os coeficientes encontrados, $p - value \leq 0,05$, pelo que rejeitamos H_0 , e assim, conclui-se que os coeficiente são significativos.

Foi aplicado o teste de *Shapiro-Wilk* para averiguar a normalidade dos resíduos. Verifica-se que não existe evidência estatística para rejeitar a normalidade dos resíduos ao nível de significância de 5% ($W = 0,994$ com um valor de $p - value = 0,36 > 0,05$).

A Figura, 3.9 apresenta o modelo 4 ajustado à série de dados,

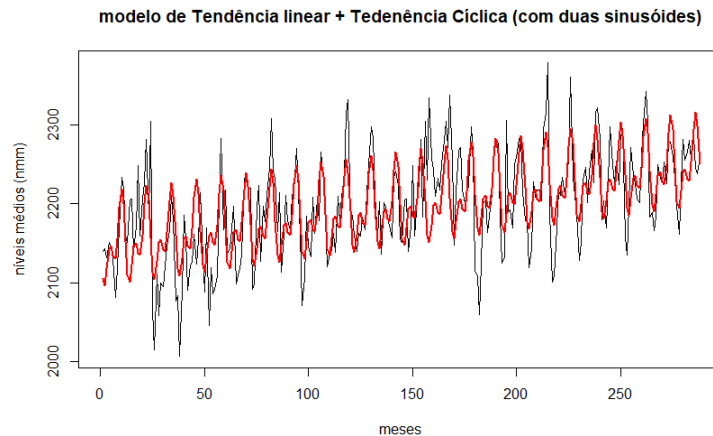


FIGURA 3.9: Linha preta: série de dados em análise; linha vermelha: modelos ajustado (modelo 4)

A Figura 3.10, apresenta os coeficientes e os respetivos p – values encontrados,

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	2147.15970	5.28098	406.583	< 2e-16	***
X	0.35568	0.03168	11.226	< 2e-16	***
I(cos(2 * pi/12 * X))	8.32921	3.72272	2.237	0.026	*
I(sin(2 * pi/12 * X))	-37.91129	3.72446	-10.179	< 2e-16	***
I(cos(4 * pi/12 * X))	-7.53276	3.72272	-2.023	0.044	*
I(sin(4 * pi/12 * X))	-31.26450	3.72299	-8.398	2.25e-15	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

FIGURA 3.10: Coeficientes do modelo do modelo 4

Para os coeficientes encontrados, p – value $\leq 0,05$, pelo que rejeitamos H_0 , e assim, conclui-se que os coeficiente são significativos.

Foi aplicado o teste de *Shapiro-Wilk* para averiguar a normalidade dos resíduos. Verifica-se que existe evidência estatística para rejeitar a normalidade dos resíduos ao nível de significância de 5% ($W = 0,984$ com p – value $= 0,0028 \leq 0,05$).

A normalidade da amostra é um pressuposto do teste de ANOVA e, tendo-se verificado que o modelo 4 não cumpre com este pressuposto, optou-se por se utilizar o teste de *Wilcoxon*.

O teste de *Wilcoxon* foi utilizado para testar a igualdade entre as medianas dos resíduos; as hipóteses a testar são:

$$\begin{cases} H_0 = \text{os resíduos possuem medianas iguais} \\ H_1 = \text{os resíduos não possuem medianas iguais} \end{cases}$$

Através do teste de *Wilcoxon*, não rejeitamos H_0 , assim, pode-se concluir que não existe evidência estatística de que os modelos 3 e 4 sejam diferentes ($V = 562$ com $p - \text{value} = 0,617 > 0,05$)

Através do teste de ANOVA, realizou-se a comparação dos modelos 1 e 3. Pela sua aplicação, mostra-se que os dois modelos são estatisticamente diferentes ($F = 43,216$ com valor de $p - \text{value} = 2,2 \times 10^{-16} \leq 0,05$). Com base no teste de *Wilcoxon* procedeu-se à comparação dos modelos 1 e 4. O teste de *Wilcoxon* mostra que os modelos são estatisticamente equivalentes ($V = 8150$ com valor de $p - \text{value} = 0,691 > 0,05$).

O melhor ajustamento é verificado com a obtenção do maior R^2 e menor $RMSE$. A Tabela 3.2, apresenta os respectivos valores para cada modelo:

modelo	R^2	$RMSE$
modelo 1	0,23	56,77
modelo 2	0,23	56,65
modelo 3	0,41	49,70
modelo 4	0,53	44,20

TABELA 3.2: R^2 e $RMSE$ dos quatro modelos em análise

O modelo 4 apresenta menor $RMSE$ e, além disso, é o que melhor explica a variabilidade dos dados, uma vez que apresenta o maior R^2 . Como tal, conclui-se que o modelo 4 reflete o melhor ajustamento.

Deste modo, o ajustamento de um modelo que combina tendência polinomial de grau um com tendência cíclica, vem confirmar que a sazonalidade é uma componente com grande preponderância nos valores observados do NMM.

A Figura 3.11, apresenta os correlogramas das funções FAC e FACP residuais do modelo 4.

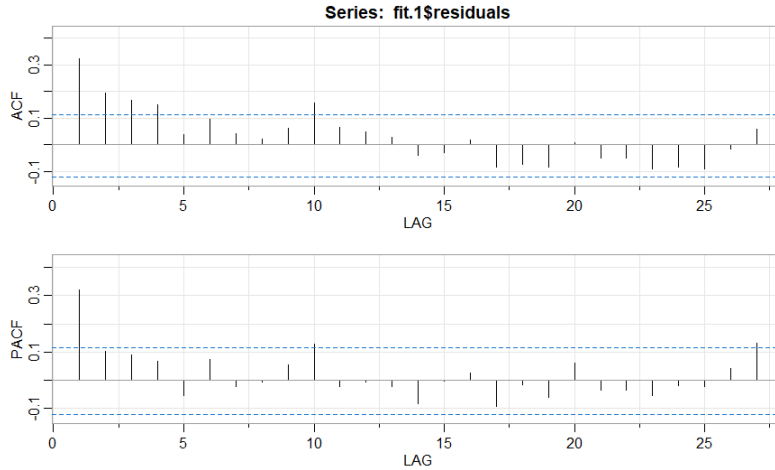


FIGURA 3.11: correlogramas da FAC e FACP residual do modelo selecionado

As duas funções apresentam um lento decaimento para zero, com vários valores fora dos intervalos de confiança no caso da FAC. O comportamento da FAC e FACP mostra que os resíduos estão correlacionados.

Esta ideia é ainda confirmada pelo teste de *Ljung e Box*; onde as hipóteses são dadas por:

$$\begin{cases} H_0 = \text{existe independência entre os resíduos} \\ H_1 = \text{não existe independência entre os resíduos} \end{cases}$$

Através da sua aplicação, rejeitamos H_0 a favor de H_1 ; pode-se concluir que não existe evidência estatística de que os resíduos são independentes ao nível de significância 5% ($Q = 42,876$ com $p\text{-value} = 5,233 \times 10^{-6} \leq 0,05$).

A tendência obtida, $T = 4,27$ mm/ano, dado que $\beta_1 = 0,3557$, é coerente com o trabalho de Mendes et al. (2020) que obteve T entre 3,96 e 5,38 mm/ano.

O modelo 4, mostra que: 1) existe uma oscilação periódica anual com amplitude de 38,8 mm/ano ($A_1 = \sqrt{(a_1)^2 + (b_1)^2}$); 2) existem duas oscilações periódicas semianuais com amplitudes de 32,2 mm/ano ($A_2 = \sqrt{(a_2)^2 + (b_2)^2}$).

Foi adicionado o coeficiente β_2 ao modelo 4, para verificar se, assim como no modelo 2, a tendência de subida do NMM não sofre aceleração. Para o coeficiente β_2 ($p\text{-value} = 0,1635 > 0,05$), pelo que não rejeitamos H_0 e assim, também neste modelo, prova-se que a aceleração não é significativa (Figura 3.12).

```

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  2.139e+03  7.940e+00 269.366 < 2e-16 ***
X            5.273e-01  1.269e-01   4.157 4.29e-05 ***
I(X^2)       -5.938e-04  4.251e-04  -1.397  0.1635
I(cos(2 * pi/12 * X)) 8.337e+00  3.716e+00  2.243  0.0257 *
I(sin(2 * pi/12 * X)) -3.791e+01  3.718e+00 -10.196 < 2e-16 ***
I(cos(4 * pi/12 * X)) -7.531e+00  3.716e+00  -2.026  0.0437 *
I(sin(4 * pi/12 * X)) -3.126e+01  3.717e+00  -8.412 2.07e-15 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

FIGURA 3.12: Coeficientes do modelo 4 com a adição do coeficiente β_2 .

Foram ainda testados modelos com oscilações trimestrais e quadrimestrais. A Figura 3.13 apresenta os resultados dos coeficientes e os respetivos p – values. Obteve-se coeficientes com (p – value $> 0,05$), pelo que não rejeitamos H_0 ; conclui-se que os coeficientes não são significativos e como tal, não se verifica a existência de oscilações trimestrais e quadrimestrais.

Coefficients:					Coefficients:				
	Estimate	Std. Error	t value	Pr(> t)		Estimate	Std. Error	t value	Pr(> t)
(Intercept)	2147.18525	5.27987	406.674	< 2e-16 ***	(Intercept)	2147.19920	5.29036	405.870	< 2e-16 ***
X	0.35550	0.03168	11.223	< 2e-16 ***	X	0.35541	0.03174	11.197	< 2e-16 ***
I(cos(2 * pi/12 * X))	8.32939	3.72173	2.238	0.0260 *	I(cos(2 * pi/12 * X))	8.32949	3.72899	2.234	0.0263 *
I(sin(2 * pi/12 * X))	-37.91195	3.72347	-10.182	< 2e-16 ***	I(sin(2 * pi/12 * X))	-37.91231	3.73073	-10.162	< 2e-16 ***
I(cos(4 * pi/12 * X))	-7.53259	3.72173	-2.024	0.0439 *	I(cos(4 * pi/12 * X))	-7.53249	3.72899	-2.020	0.0443 *
I(sin(4 * pi/12 * X))	-31.26481	3.72200	-8.400	2.27e-15 ***	I(sin(4 * pi/12 * X))	-31.26498	3.72926	-8.384	2.6e-15 ***
I(cos(6 * pi/12 * X))	-2.43884	3.72173	-0.655	0.5128	I(cos(6 * pi/12 * X))	-2.43874	3.72899	-0.654	0.5137
I(sin(6 * pi/12 * X))	-4.88061	3.72173	-1.311	0.1908	I(sin(6 * pi/12 * X))	-4.88070	3.72899	-1.309	0.1917
---					I(cos(8 * pi/12 * X))	2.68279	3.72899	0.719	0.4725
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					I(sin(8 * pi/12 * X))	2.34019	3.72890	0.628	0.5308

					Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1				

(A)

(B)

FIGURA 3.13: Coeficientes dos modelos com sazonalidade trimestral e quadrimestral: (A) modelo com 3 sinusoides; (B) modelo com 4 sinusoides.

3.1.3 Avaliação da aplicação de modelos ARIMA

Com o intuito de identificar o modelo ARIMA com melhor ajustamento e estimar os seus coeficientes, consideramos apenas os primeiros 230 meses (19) anos, os restantes 58 meses (aproximadamente 5 anos) serão usados para aferir a qualidade da previsão dos modelos identificados. A Figura 3.14 apresenta a representação gráfica dos primeiros 230 meses,

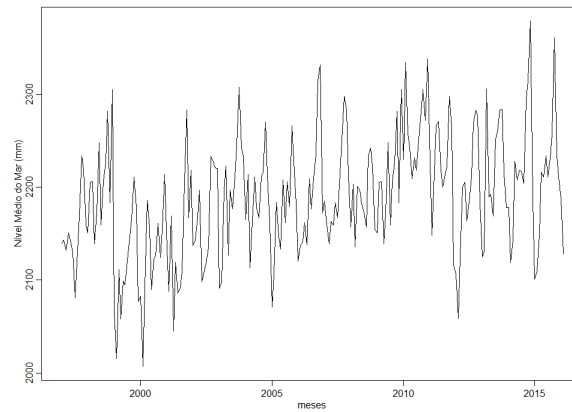


FIGURA 3.14: Registos maregráficos de Sines, primeiros 230 meses.

A trajetória da série apresenta uma tendência crescente, assim será aconselhável considerar a série das diferenças. A não estacionaridade é verificada através dos correlogramas (Figura 3.15). A FAC descreve movimentos periódicos e um decaimento lento para zero sobre os “lag’s” múltiplos de 12. O facto das auto-correlações decrescerem muito lentamente, confirma a necessidade de aplicar uma diferença simples para estabilizar os dados.

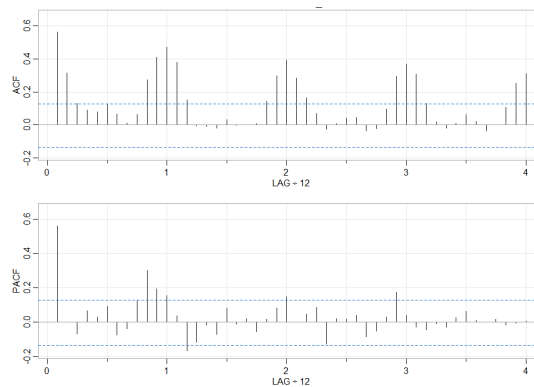


FIGURA 3.15: Correlogramas da série de Sines: ACF (gráfico superior) e ACFP (gráfico inferior)

A série foi sujeita a uma diferenciação simples e a uma diferenciação sazonal, Figura 3.16.

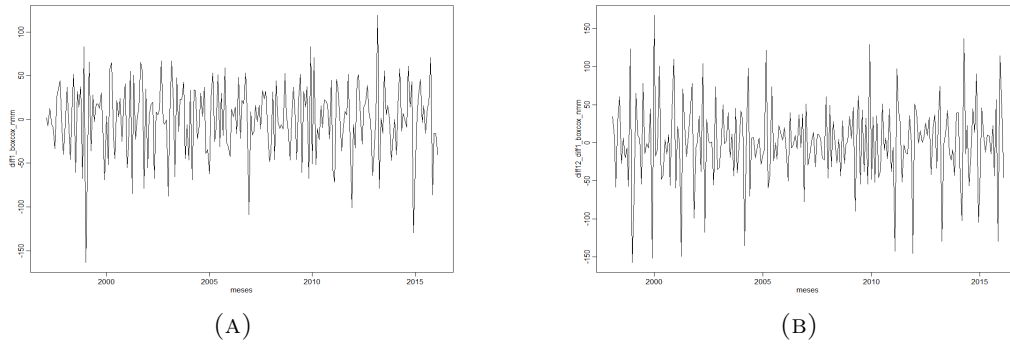


FIGURA 3.16: (A) Diferenciação simples; (B) Diferenciação sazonal

Considerando a série estacionária, a Figura 3.17, apresenta os correlogramas dos dados em análise.

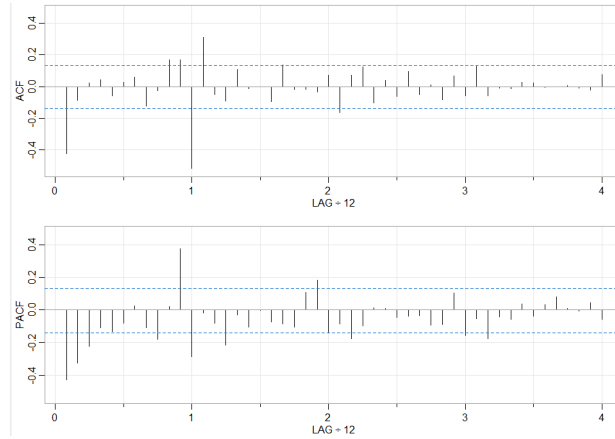


FIGURA 3.17: Correlogramas da série estacionária - FAC (gráfico superior) e FACP (gráfico inferior)

Nesta situação nota-se que tanto a FAC como a FACP da série transformada, descrevem dois tipos distinto de comportamento, um sobre os “lag’s” $K = 1, 2, 3, \dots$ e outro para os “lag’s” múltiplos de 12, espectável devido à oscilação anual do NMM. Portanto, estamos perante um modelo $SARIMA(p, d, q) \times (P, D, Q)_{12}$, onde o comportamento dos “lag’s” $K = 1, 2, 3, \dots$ ajuda a identificar (p, d, q) e o comportamento dos “lag’s” múltiplos de 12, ajuda a identificar (P, D, Q) . Nos “lag’s” $K = 1, 2, 3, \dots$, a FAC descreve um decaimento brusco para zero depois do “lag” 1 enquanto a FACP descreve um decaimento exponencial para zero, sugerindo um ajustamento $p = 0$, $d = 1$ (diferenciação simples) e $q = 1$.

Nos “lag’s” múltiplos de 12, tanto a FAC como a FACP descrevem um rápido decaimento exponencial para zero. Como tal, não dão uma clara indicação do modelo que dever ser escolhido, surgindo várias hipóteses nomeadamente, $(P, D, Q) = (0, 1, 1)$ ou $(1, 1, 0)$ ou $(1, 1, 1)$.

De seguida, apresenta-se a aplicação da metodologia de avaliação da qualidade estatística e do ajustamento ao modelo $(p, d, q)(P, D, Q) = (1, 1, 1)(0, 1, 1)_{12}$; a aplicação desta metodologia, aos restantes modelos considerados, pode ser consultada no Apêndice A.

Considerou-se a série diferenciada; a análise da Figura 3.17 mostra que as autocorrelações desta nova série não são significativas, isto é, estamos na presença de modelo $SARIMA = (1, 1, 1)(0, 1, 1)_{12}$.

A Tabela 3.3, apresenta os coeficientes encontrados para o modelo,

Modelo	θ_1	ϕ_1	Φ_1
$SARIMA(1, 1, 1)(0, 1, 1)_{12}$	0,2103	-0,8734	-0,9992

TABELA 3.3: Coeficientes do modelos $SARIMA(1, 1, 1)(0, 1, 1)_{12}$

O modelo $SARIMA(1, 1, 1)(0, 1, 1)_{12}$ é dado por:

$$(1 - 0,2103B)(1 - B)(1 - B^{12})X_t = (1 + 0,8734B)(1 + 0,9992B^{12})\epsilon_t,$$

O modelo identificado verifica os seguintes pressupostos:

1. Coeficientes significativamente diferentes de zero. A Tabela 3.4 apresenta os intervalos de confiança de cada um dos coeficientes estimados. Visto que nenhum dos intervalos contém o valor zero, ou seja, os intervalos variaram apenas entre valores negativos ou positivos, todos eles são considerados significativos.

	2,5%	97,5%
θ_1	0,02	0,4006
ϕ_1	-0,9980	-0,7488
Φ_1	-1,2243	-0,7740

TABELA 3.4: Intervalos de confiança dos coeficientes do modelo

2. Os correlogramas da FAC e FACP dos resíduos, apresentam valores que, na generalidade, estão compreendidos entre os limites de confiança (Figura 3.18);

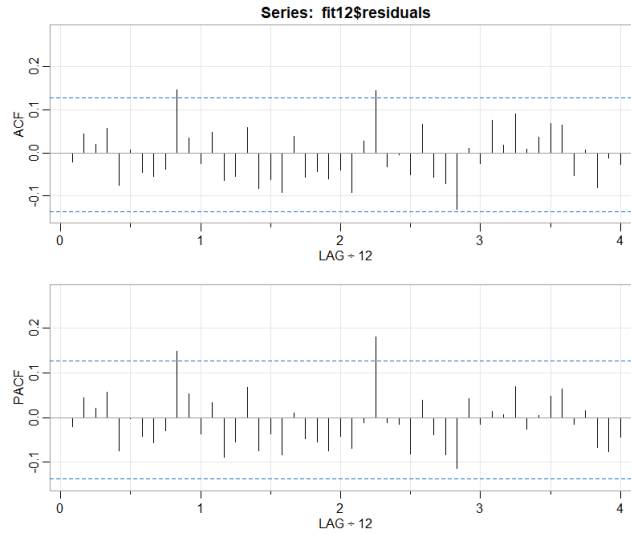


FIGURA 3.18: Correlogramas dos resíduos do modelo $SARIMA(1, 1, 1)(0, 1, 1)_{12}$: FAC (gráfico superior) e FACP (gráfico inferior)

- Os resíduos não são correlacionados. Através do teste de *Ljung e Box*, não rejeitamos H_0 , assim pode-se concluir que existe evidência estatística de que os resíduos não estão correlacionados nível de significância 5% ($Q = 20,984$ com $p - value = 0,4599 > 0,05$);

A Figura 3.19, apresenta o histograma dos resíduos do modelo. A média dos resíduos é 0,3514 m com intervalo de confiança a 95% (-3,4607 ; 4,1635).

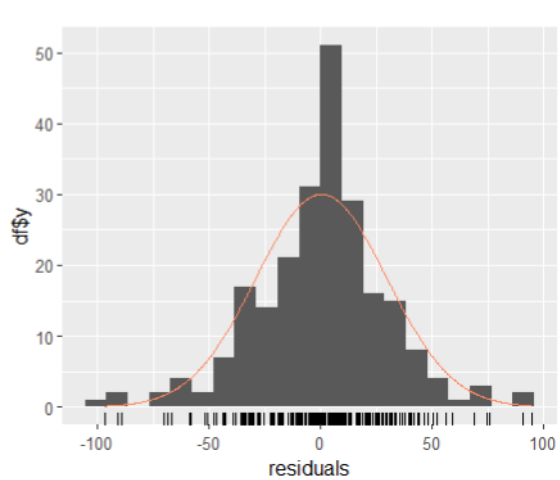


FIGURA 3.19: Histograma da frequências dos resíduos do modelo

Foi aplicado o teste de *Shapiro-Wilk* para averiguar a normalidade dos resíduos. Verifica-se que existe evidência estatística para rejeitar a normalidade dos

resíduos ao nível de significância de 5% ($W = 0,9791$ com um valor associado de $p - value = 0,0017 \leq 0,05$).

Na avaliação dos diversos modelos identificados mostrou-se que, os únicos que asseguram, simultaneamente, a qualidade estatística e a qualidade do ajustamento são os modelos $SARIMA(1, 1, 1)(0, 1, 1)_{12}$ e $SARIMA(0, 1, 1)(0, 1, 1)_{12}$. A avaliação do modelo $SARIMA(0, 1, 1)(0, 1, 1)_{12}$ pode ser consultada no Apêndice A.

Depois de identificados os modelos que descrevem, de forma satisfatória, os dados em análise, passaremos agora, à escolha do modelo que melhor descreva a série em questão, tendo em conta os critérios AIC e BIC, o $RMSE$ e o ajustamento gráfico das previsões à série temporal completa.

Modelo	AIC	BIC	$RMSE$
$SARIMA(0, 1, 1)(0, 1, 1)_{12}$	2141,11	2151,25	43,49
$SARIMA(1, 1, 1)(0, 1, 1)_{12}$	2138,69	2152,21	43,01

TABELA 3.5: Critérios AIC e BIC; $RMSE$

A análise da Tabela 3.5 evidencia que os três critérios apresentam valores relativamente semelhantes para os dois modelos. Não obstante, o modelo $SARIMA(1, 1, 1)(0, 1, 1)_{12}$ apresenta melhor desempenho, na generalidade, pelo que deve ser escolhido como sendo o modelo que melhor se ajusta à série em estudo.

Por último, comparamos graficamente as previsões com os meses da série que não foram considerados na obtenção do modelo (os últimos 58 meses). A Figura 3.20 mostra que, apesar de alguma variabilidade, as previsões apresentam um comportamento semelhante aos valores reais.

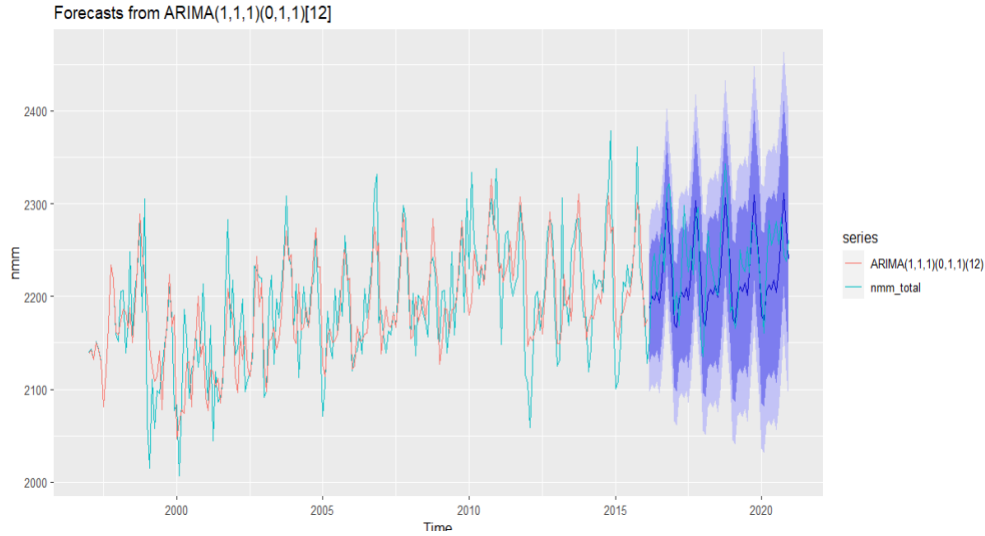


FIGURA 3.20: Ajustamento das previsões aos meses não considerados na análise para construção dos modelos.

Procedeu-se à aplicação do teste de *Shapiro-Wilk* para averiguar a normalidade dos resíduos. Com um nível de significância de 5%, podemos concluir que os resíduos do modelo não seguem uma distribuição normal ($W = 0,979$ com um valor de $p - value = 0,0017 \leq 0,05$).

3.1.4 Seleção e Previsão

Foram identificados vários modelos que podem explicar o comportamento do NMM. Nesta fase do trabalho, vamos avaliar se o modelo do tipo ARIMA garante melhor ajustamento que o modelo de tendência polinomial e tendência cíclica (modelo 4). Para isso importa analisar:

- Comparação da distribuição dos resíduos;
- Analisar o $RMSE$ dos modelos.

Para avaliar se os resíduos, obtidos por modelos distintos são provenientes da mesma população, recorreu-se ao teste de *Wilcoxon*, visto que uma das amostras não segue uma distribuição normal ($SARIMA(1,1,1)(0,1,1)_{12}$). A hipótese de igualdade das medianas dos resíduos não foi rejeitada, para um nível de significância de 5% ($V = 6714$ com $p - value = 0,984 > 0,05$), isto é, existe evidência estatística para afirmar que os dois modelos são equivalentes.

A Tabela 3.18 apresenta o $RMSE$ e o número de coeficientes de cada modelo.

Modelo	<i>RMSE</i>	nº de coeficientes
<i>modelo 4</i>	44,20	6
<i>SARIMA</i> (1, 1, 1)(0, 1, 1) ₁₂	43,01	3

TABELA 3.6: *RMSE* e nº de coeficientes dos modelos 4 e *SARIMA*(1, 1, 1)(0, 1, 1)₁₂

No que diz respeito ao *RMSE*, o modelo *SARIMA*(1, 1, 1)(0, 1, 1)₁₂ apresenta o menor valor, além disso, é o modelo mais parcimonioso. Desta forma, conclui-se que o modelo *SARIMA*(1, 1, 1)(0, 1, 1)₁₂ apresenta o melhor ajustamento.

Pode-se agora analisar as previsões que este modelo forneceria para um período de 10 anos, isto é, de 2020 a 2030. Na Figura 3.21 apresentam-se as previsões obtidas e os correspondentes limites dos intervalos de confiança a 95%. Os valores destas previsões estão próximos de 2200 mm, com uma ligeira tendência de crescimento.

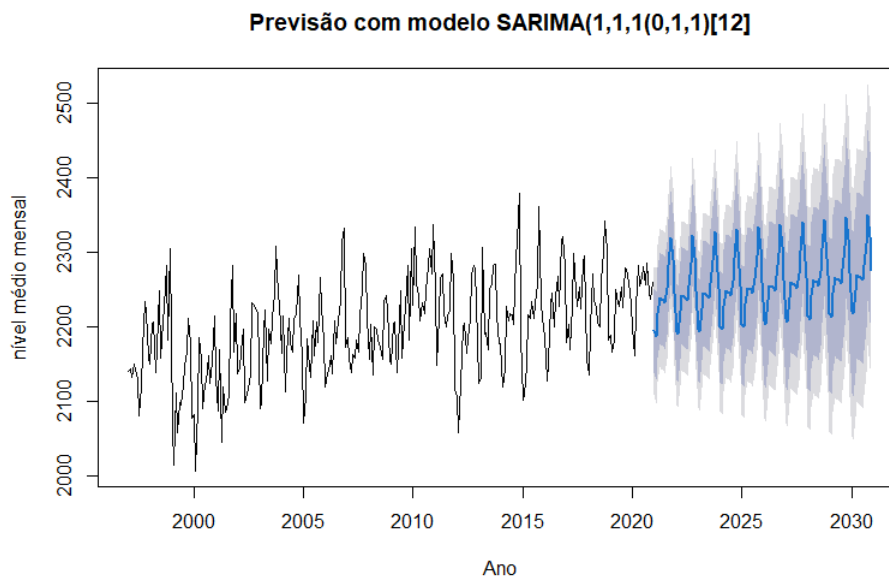


FIGURA 3.21: Previsão dos NMm's para os próximos 10 anos

Quantificou-se a tendência do NMM para os valores previstos, e podemos esperar um aumento de 44 mm até 2030, o que se traduz numa tendência de +4,36 mm/ano.

3.2 Análise dos Registos Maregráficos de Viana do Castelo

3.2.1 Análise Descritiva dos Dados

A série dos dados alisados é constituída por 384 observações do MNm obtidas a partir das alturas horárias registadas pela estação maregráfica do porto Viana do Castelo, Figura 3.22. Os dados analisados referem-se ao período compreendido entre 1984 a 2015.

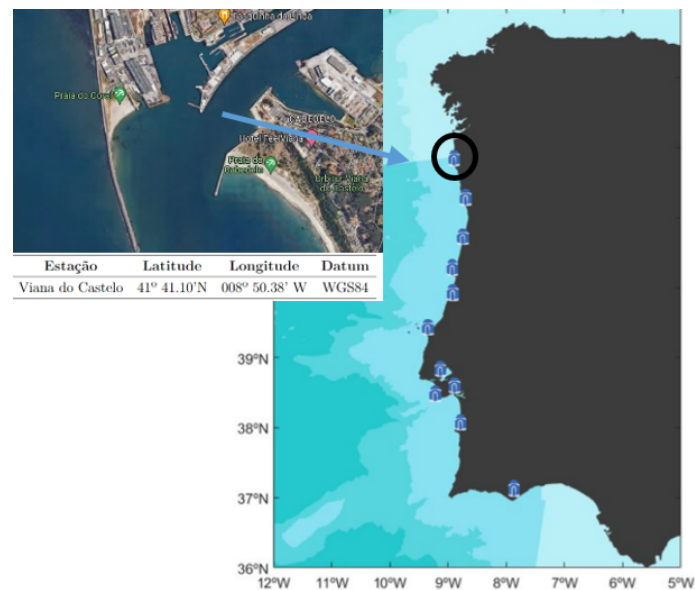


FIGURA 3.22: Localização da estação maregráfica do porto de Viana do Castelo ($\varphi = 41^{\circ}41,1'N$ e $\lambda = 008^{\circ}50,28'W$) (fonte: IH, 2020)

A Tabela 3.7 apresenta as principais estatísticas descritivas dos NMm's (1984-2015):

<i>NMm's (mm)</i>	
Média	2165
Mediana	2148
Desvio-padrão	80,69
Assimetria	0,83
Mínimo	1982
Máximo	2440
Dimensão da amostra	384

TABELA 3.7: Estatísticas descritivas dos NMm's (mm) de Viana do Castelo

3.2.2 Avaliação da Tendência e Sazonalidade

A Figura 3.23 apresenta o ajustamento um modelo de regressão polinomial de grau um ($X_t = \beta_0 + \beta_1 t$) aos NMm's, de tendência β_1 .

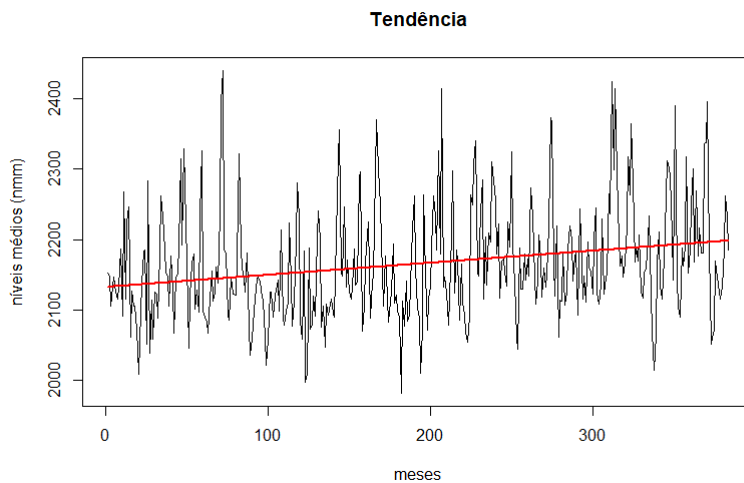


FIGURA 3.23: Linha preta: evolução dos NMm's em Viana do Castelo no período em análise; linha vermelha: modelo ajustado

A Figura 3.24 apresenta os coeficientes e os respetivos $p - value$ obtidos pelo *software R*.

```

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) 2.132e+03   8.030e+00 265.556 < 2e-16 ***
X            1.713e-01   3.615e-02   4.739 3.03e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

FIGURA 3.24: Coeficientes do modelo polinomial de grau um estimado para Viana do Castelo

Aplicou-se o teste de significância estatística para verificar se os coeficientes são significativos. Obteve-se ($p - value \leq 0,05$); rejeitamos H_0 a favor de H_1 , assim conclui-se que os coeficientes são significativos.

Face ao declive positivo, $\beta_1 = 1,713 \times 10^{-1}$, pode-se concluir que o NMM está a aumentar.

Foi aplicado o teste de *Shapiro-Wilk* para averiguar a normalidade dos resíduos. Verifica-se que existe evidência estatística para rejeitar H_0 ao nível de significância de 5% ($W = 0,9537$ com $p = 1,31 \times 10^{-09} \leq 0,05$), pelo que os resíduos não seguem uma distribuição normal.

Doravante, passaremos a referir-mo-nos ao modelo de regressão polinomial de grau um por modelo 1.

Ajustou-se um polinómio de grau dois ($X_t = \beta_0 + \beta_1 t + \beta_2 t^2$), com vista a determinar se a tendência sofreu aceleração (Figura 3.25).

```

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) 2.139e+03   1.209e+01 176.941 <2e-16 ***
X            6.331e-02   1.450e-01   0.436  0.663
I(X^2)       2.806e-04   3.648e-04   0.769  0.442
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

FIGURA 3.25: Coeficientes do modelo polinomial de grau dois estimado para a série de registos de Viana do Castelo

Obteve-se $p - value > 0,05$, pelo que concluímos que os coeficientes β_1 e β_2 não são significativos e como tal, não se verifica aceleração da tendência da subida de 1984 a 2015 (Figura 3.25).

Através da aplicação do teste de *Shapiro-Wilk*, verifica-se que existe evidência estatística para rejeitar H_0 ao nível de significância de 5% ($W = 0,9537$ com $p = 1,015 \times 10^{-09} \leq 0,05$), pelo que os resíduos não seguem um distribuição normal.

Doravante, passaremos a referir-mo-nos ao modelo polinomial de grau dois por modelo 2.

A Figura 3.26 apresenta as distribuições do NMm's registados pelo marégrafo de Viana do Castelo entre 1984 e 2015:

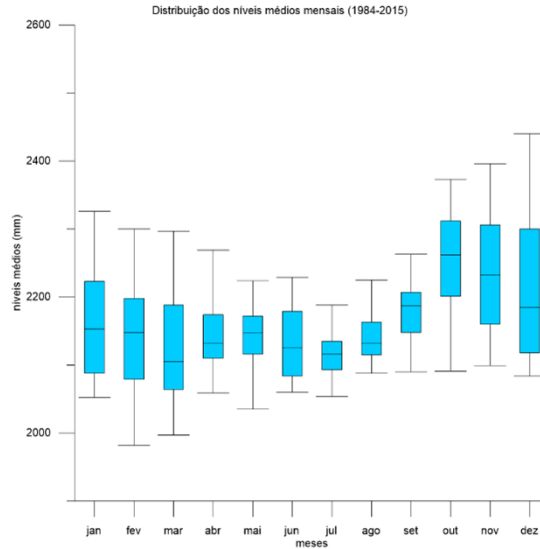


FIGURA 3.26: Comportamento dos níveis médios ao longo do ano, em Viana do Castelo no período 1984 a 2015

Assim como nos registos maregráficos de Sines, também nestes se verifica que a variação do nível médio aparenta ter uma componente sazonal. Verificando-se níveis mais elevados nos meses de outubro e novembro e níveis mais reduzidos nos meses de março e julho.

O modelo de tendência linear e tendência cíclica que proporcionou o melhor ajustamento é dado pela expressão,

$$X_t = \beta_0 + \beta_1 t + a_1 \cos(\omega_1 t) + b_1 \sin(\omega_1 t) + b_2 \sin(\omega_2 t),$$

onde $\omega_1 = \frac{2\pi}{p_1}$ com $p_1 = 12$; e $\omega_2 = \frac{2\pi}{p_2}$ com $p_2 = 6$.

onde os coeficientes assumem os valores apresentados pela Figura 3.27.

```

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  2133.83498    6.93638  307.630 < 2e-16 ***
X              0.16403     0.03123   5.253 2.51e-07 ***
I(cos(2 * pi/12 * X))  41.86940    4.89394   8.555 2.95e-16 ***
I(sin(2 * pi/12 * X)) -22.70148    4.89523  -4.637 4.86e-06 ***
I(sin(4 * pi/12 * X)) -30.39664    4.89414  -6.211 1.39e-09 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

FIGURA 3.27: Coeficientes do modelo de tendência linear e cíclica

Aplicou-se o teste de *Shapiro-Wilk* para averiguar a normalidade dos resíduos. Verifica-se que existe evidência estatística para rejeitar a normalidade dos resíduos ao nível de significância de 5% ($W = 0,9722$ com um valor associado de $p - value = 1,047 \times 10^{-06} \leq 0,05$).

Doravante, passaremos a referirmo-nos a este modelo por modelo 3.

O teste de *Wilcoxon* foi utilizado para comparar o modelo 3 com os modelos 1 e 2. Através da sua aplicação, mostra-se que as diferenças entre os modelos 1 e 3 e os modelos 2 e 3 são estatisticamente equivalentes ($V = 134856$ com valor de $p - value = 0,0375 \leq 0,05$ e $V = 134760$ com valor de $p - value = 0,03612 \leq 0,05$, respetivamente, entre os modelos 1 e 3 e entre os modelos 2 e 3).

A Tabela 3.8 apresenta o R^2 e $RMSE$ para cada modelo.

modelo	R^2	$RMSE$
modelo 1	0,06	78,32
modelo 2	0,06	78,26
modelo 3	0,30	67,36

TABELA 3.8: R^2 e $RMSE$ dos modelo em análise

O modelo 3 apresenta-se como sendo o modelo que melhor explica a variabilidade dos dados (maior R^2) e que possui o menor $RMSE$. Como tal, conclui-se que o modelo 3 reflete o melhor ajustamento.

A Figura 3.28 apresenta o ajustamento do modelo 3 à série de dados de Viana do Castelo.

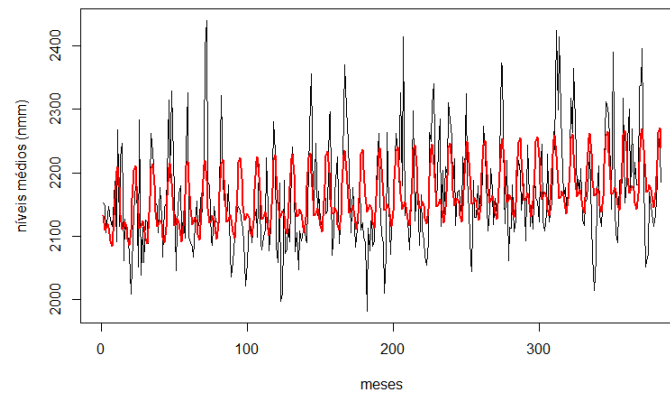


FIGURA 3.28: Linha preta: série de dados em análise (Viana do Castelo); linha vermelha: modelo ajustado (modelo 3)

A Figura 3.29 apresenta os correlogramas das funções FAC e FACP residuais do modelo 3.

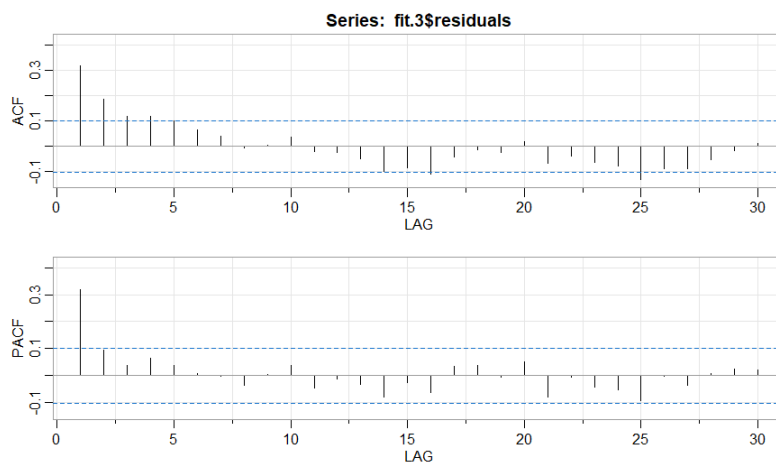


FIGURA 3.29: Correlogramas dos resíduos do modelo: FAC (gráfico superior) e FACP (gráfico inferior)

As duas funções apresentam um lento decaimento para zero, com vários valores fora dos intervalos de confiança no caso da FAC. O comportamento da FAC mostra que os resíduos estão correlacionados.

Esta ideia é ainda confirmada pelo teste de *Ljung e Box*. Através da sua aplicação, rejeitamos H_0 a favor de H_1 , assim pode-se concluir que existe evidência estatística de que os resíduos não são independentes ao nível de significância 5% ($Q = 45,309$ com $p - value = 1,913 \times 10^{-6} \leq 0,05$).

A tendência obtida $T = 1,97$ mm/ano, dado que $\beta_1 = 0,164$, é coerente com o trabalho de Mendes et al. (2020) que obteve T entre 1,97 e 3,75 mm/ano.

O modelo selecionado, mostra que: 1) existe uma oscilação periódica anual com amplitude de 47,6 mm/ano ($A_1 = \sqrt{(a_1)^2 + (b_1)^2}$); 2) existem duas oscilações periódicas semianuais com amplitudes de 32,4 mm/ano ($A_2 = |b_2|$).

Foi adicionado o coeficiente β_2 ao modelo 3, para verificar se, assim como no modelo 2, a tendência de subida do NMM não sofre aceleração. A Figura 3.30 apresenta os resultados dos coeficientes e os respectivos p – values. Obteve-se um valor de p – value $> 0,05$, pelo que concluímos que os coeficientes β_1 , β_2 e $b_2\omega_2$ não são significativos. Como tal, também neste modelo, existe evidência estatística de que não se verificou aceleração.

Coefficients:				
	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	2.141e+03	1.045e+01	204.895	< 2e-16 ***
X	5.661e-02	1.253e-01	0.452	0.652
I(X^2)	2.787e-04	3.152e-04	0.884	0.377
I(cos(2 * pi/12 * X))	4.187e+01	4.899e+00	8.546	3.22e-16 ***
I(sin(2 * pi/12 * X))	-2.270e+01	4.900e+00	-4.633	4.97e-06 ***
I(cos(4 * pi/12 * X))	3.163e+00	4.899e+00	0.646	0.519
I(sin(4 * pi/12 * X))	-3.040e+01	4.899e+00	-6.204	1.45e-09 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1				

FIGURA 3.30: Coeficientes do modelo 3 com a adição do coeficiente β_2

Foram ainda testados modelos com oscilações trimestrais e quadrimestrais. A Figura 3.31 apresenta os resultados dos coeficientes e os respectivos p – values. Obteve-se coeficientes com (p – value $> 0,05$), pelo que não rejeitamos H_0 ; conclui-se que os coeficientes não são significativos e como tal, não se verifica a existência de oscilações trimestrais e quadrimestrais.

Coefficients:									
	Estimate	Std. Error	t value	Pr(> t)		Estimate	Std. Error	t value	Pr(> t)
(Intercept)	2133.81654	6.94672	307.169	< 2e-16 ***	(Intercept)	2133.79833	6.96369	306.418	< 2e-16 ***
X	0.16413	0.03128	5.248	2.58e-07 ***	X	0.16422	0.03135	5.238	2.72e-07 ***
I(cos(2 * pi/12 * X))	41.86930	4.90101	8.543	3.30e-16 ***	I(cos(2 * pi/12 * X))	41.86921	4.91289	8.522	3.90e-16 ***
I(sin(2 * pi/12 * X))	-22.70112	4.90230	-4.631	5.03e-06 ***	I(sin(2 * pi/12 * X))	-22.70077	4.91418	-4.619	5.30e-06 ***
I(cos(4 * pi/12 * X))	3.16400	4.90101	0.646	0.519	I(cos(4 * pi/12 * X))	3.16390	4.91289	0.644	0.520
I(sin(4 * pi/12 * X))	-30.39647	4.90121	-6.202	1.47e-09 ***	I(sin(4 * pi/12 * X))	-30.39631	4.91309	-6.187	1.61e-09 ***
I(cos(6 * pi/12 * X))	-5.96621	4.90101	-1.217	0.224	I(cos(6 * pi/12 * X))	-5.96631	4.91289	-1.214	0.225
I(sin(6 * pi/12 * X))	-0.45045	4.90101	-0.092	0.927	I(sin(6 * pi/12 * X))	-0.45036	4.91289	-0.092	0.927
					I(cos(8 * pi/12 * X))	-2.06006	4.91289	-0.419	0.675
					I(sin(8 * pi/12 * X))	0.45566	4.91282	0.093	0.926
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1				

(A)

(B)

FIGURA 3.31: Coeficientes dos modelos com sazonalidade trimestral e quadrimestral: (A) modelo com 3 sinusoides; (B) modelo com 4 sinusoides.

3.2.3 Avaliação da aplicação de modelos ARIMA

O modelo $SARIMA(1, 0, 0)(0, 1, 1)_{12}$ foi identificado como sendo o modelo que melhor se ajusta ao comportamento geral dos registos maregráficos de Viana do Castelo. A avaliação dos restantes modelos identificados pode ser consultada no Apêndice B.

A Tabela 3.9, apresenta os coeficientes encontrados no ajustamento do modelo $SARIMA(1, 0, 0)(0, 1, 1)_{12}$ aos dados em análise,

Modelo	ϕ_1	Θ_1
$SARIMA(1, 0, 0)(0, 1, 1)_{12}$	0,2814	-0,9089

TABELA 3.9: Coeficientes do modelos $SARIMA(1, 0, 0)(0, 1, 1)_{12}$

O modelo $SARIMA(1, 0, 0)(0, 1, 1)_{12}$ é dado por:

$$(1 - 0,2814B)(1 - B^{12})X_t = (1 + 0,9089B^{12})\epsilon_t.$$

O modelo identificado encontrando-se em conformidade com os seguintes pressupostos:

- Coeficientes são significativamente diferentes de zero. A Tabela 3.10 apresentam os coeficientes e os respetivos intervalos de confiança.

	2,5%	97,5%
ϕ_1	0,1701	0,3928
Θ_1	-1,2367	-0,8083

TABELA 3.10: Intervalos de confiança do modelo identificado

- FAC e FACP residuais têm os seus valores compreendidos entre os intervalos de confiança (Figura 3.32);

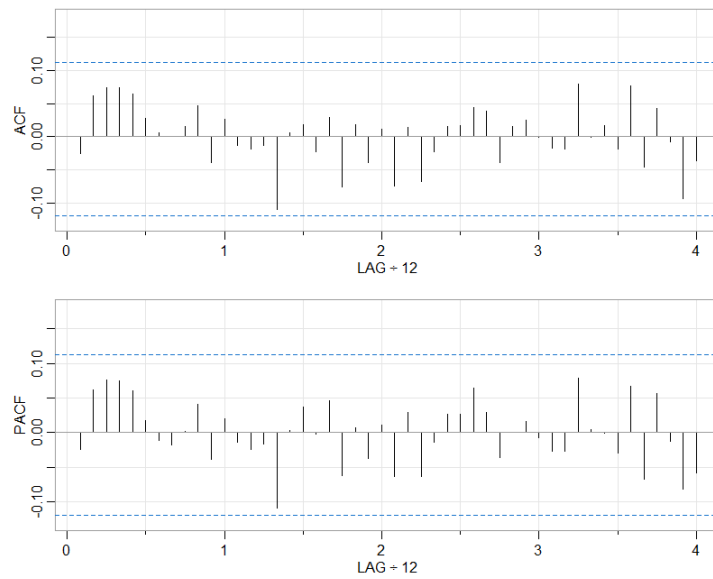


FIGURA 3.32: Correlogramas dos resíduos do modelo $SARIMA(1,0,0)(0,1,1)_{12}$: FAC (gráfico superior) e FACP (gráfico inferior)

- A hipótese de independência dos resíduos foi testada através do teste de *Ljung e Box*; não rejeitamos H_0 , assim pode-se concluir que existe evidência estatística de que os resíduos não estão correlacionados nível de significância 5% ($Q = 14,679$ com $p - value = 0,8387 > 0,05$).

A Figura 3.33, apresenta o histograma dos resíduos do modelo. A média dos resíduos é -0,2337 m com intervalo de confiança a 95% (-7,4617 ; 6,9942).

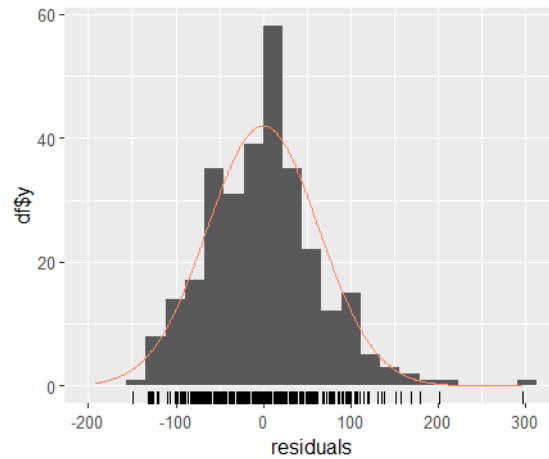


FIGURA 3.33: Histograma das frequências dos resíduos do modelo escolhido

Foi aplicado o teste de Shapiro-Wilk para averiguar a normalidade dos resíduos. Verifica-se que existe evidência estatística para rejeitar a normalidade dos resíduos ao nível de significância de 5% ($W = 0,9787$ com $pvalue = 0,00019 \leq 0,05$).

Na fase de avaliação dos modelos identificados, mostrou-se que os modelos $SARIMA(1, 0, 0)(0, 1, 1)_{12}$, $SARIMA(0, 1, 1)(0, 1, 1)_{12}$ e $SARIMA(1, 0, 0)(2, 0, 0)_{12}$, descrevem, de forma satisfatória, a série em análise, cumprindo com os pressupostos definidos. A avaliação dos dois últimos modelos, $SARIMA(0, 1, 1)(0, 1, 1)_{12}$ e $SARIMA(1, 0, 0)(2, 0, 0)_{12}$, pode ser consultada no Apêndice B.

Passaremos agora à escolha do modelo que melhor descreva a série em questão, tendo em conta os critérios AIC, BIC e $RMSE$.

Modelo	AIC	BIC	$RMSE$
$SARIMA(1, 0, 0)(0, 1, 1)_{12}$	3249,13	3263,78	63,5
$SARIMA(0, 1, 1)(0, 1, 1)_{12}$	3252,21	3263,19	65,2
$SARIMA(1, 0, 0)(2, 0, 0)_{12}$	3293,61	3311,93	70,2

TABELA 3.11: Critérios AIC e BIC; $RMSE$

A análise da Tabela 3.11, mostra que o modelo $SARIMA(1, 0, 0)(0, 1, 1)_{12}$ apresenta melhor desempenho, na generalidade, pelo que deve ser escolhido como sendo o modelo que melhor se ajusta à série em estudo.

A Figura 3.34, apresenta o ajustamento gráfico das previsões aos meses da série que não foram considerados na obtenção do modelo (os últimos 84 meses). A avaliação gráfica confirma a adequabilidade do modelo.

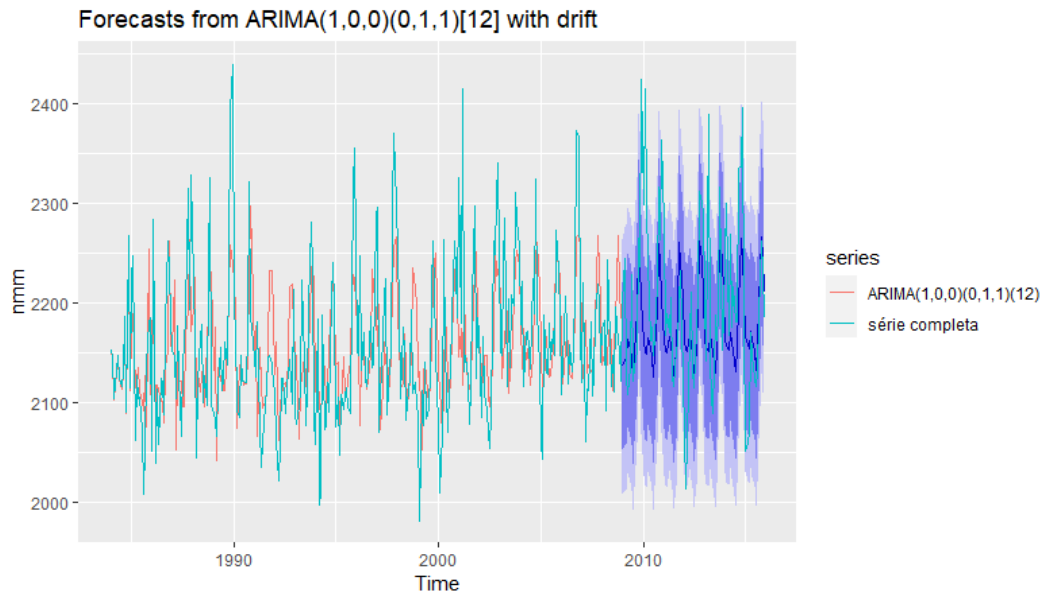


FIGURA 3.34: Ajustamento das previsões aos meses não considerados na análise para construção dos modelos para Viana do Castelo

3.2.4 Seleção e Previsão

Recorreu-se ao teste de *Wilcoxon* para avaliar se os resíduos, obtidos por métodos distintos (modelo 3 e $SARIMA(1, 0, 0)(0, 1, 1)_{12}$), são provenientes da mesma população. A hipótese de igualdade das medianas dos resíduos não foi rejeitada, para um nível de significância de 5% ($V = 11148$ com um valor de $p - value = 0,274 > 0,05$), isto é, existe evidência estatística para afirmar que os dois modelos são equivalentes.

A Tabela 3.12 apresenta o *RMSE* e o número de coeficientes de cada modelo.

Modelo	<i>RMSE</i>	nº de coeficientes
<i>modelo 3</i>	67,36	5
$SARIMA(1, 0, 0)(0, 1, 1)_{12}$	63,50	2

TABELA 3.12: *RMSE* e nº de coeficientes dos modelos 3 e $SARIMA(1, 0, 0)(0, 1, 1)_{12}$

O modelo $SARIMA(1, 0, 0)(0, 1, 1)_{12}$ apresenta o menor valor de *RMSE* e é o modelo mais parcimonioso. Desta forma, conclui-se que o melhor ajustamento é assegurado pelo modelo $SARIMA(1, 0, 0)(0, 1, 1)_{12}$.

Pode-se agora analisar as previsões que este modelo forneceria para um período de 10 anos, isto é de 2015 a 2025. Na Figura 3.35 apresentam-se as previsões obtidas e os correspondentes limites dos intervalos de confiança a 95%. Os valores destas previsões estão próximos de 2200 mm, com uma ligeira tendência de crescimento.

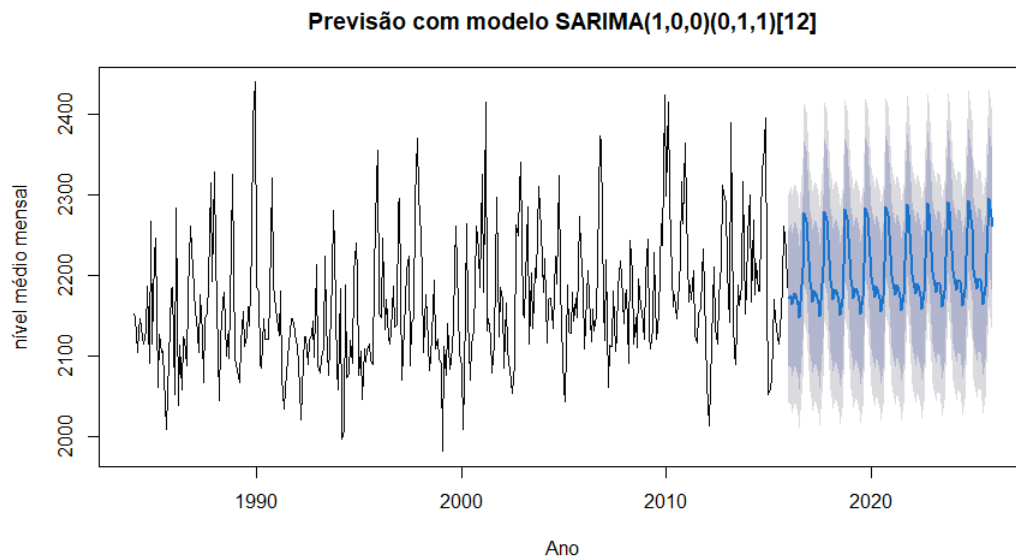


FIGURA 3.35: Previsão dos NMM's para os próximos 10 anos para Viana do Castelo

Quantificou-se a tendência do NMM para os valores previstos, e podemos esperar um aumento de 30 mm até 2025, o que se traduz numa tendência de, aproximadamente, +3 mm/ano.

A diferença entre a tendência verificada ao longo dos 31 anos e a tendência dos valores previstos é de aproximadamente 1 mm/ano. Nos modelos de médias móveis, os efeitos causados por valores passados de X_t , apenas perduram por um período finito. Isto significa que, a memória do processo fará com que os valores que tenham acontecido nos últimos instantes da série tenham maior influência nos valores previstos (Murteira et al., 2000). Como tal, a diferença poderá ser o resultado da utilização de um coeficiente de médias móveis e de uma tendência ligeiramente superior nos últimos registos. De facto, quando analisamos apenas os últimos 6,5 anos (aproximadamente 20% dos dados), verificamos uma tendência de + 5,1 mm/ano, $\beta_1 = 0,4249$ (Figura 3.36).

```

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) 2173.7073    21.5729 100.761  <2e-16 ***
X              0.4249     0.6151   0.691   0.492
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

FIGURA 3.36: Ajustamento de um modelo polinomial de grau um aos últimos 6,5 anos de registos em Viana do Castelo

3.3 Análise dos Registos Maregráficos do Funchal

3.3.1 Análise Descritiva dos Dados

A série dos dados alisados é constituída por 192 observações do NMm obtidas a partir das alturas horárias registadas pela estação maregráfica do porto do Funchal, Figura 3.37. Os dados analisados referem-se ao período compreendido entre 2005 a 2020.

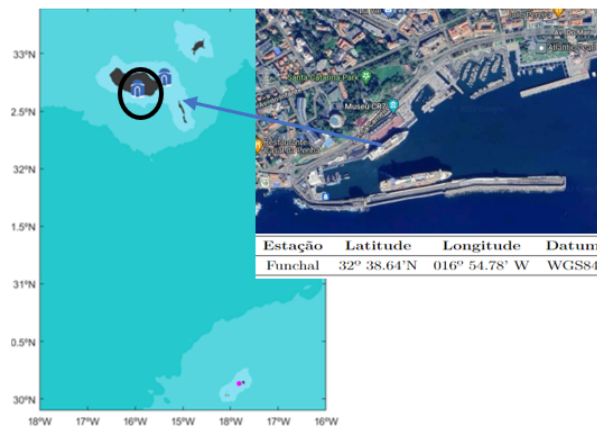


FIGURA 3.37: Localização da estação maregráfica do porto do Funchal ($\varphi = 32^{\circ}38,64'N$ e $\lambda = 016^{\circ}54,78'W$) (fonte: IH, 2020)

A Tabela 3.13 apresenta as principais estatísticas descritivas dos NMm's (2005-2020):

<i>NMm's (mm)</i>	
Média	1613
Mediana	1612
Desvio-padrão	70,14
Assimetria	0,15
Mínimo	1481
Máximo	1782
Dimensão da amostra	192

TABELA 3.13: Estatísticas descritivas dos NMm's (mm) do Funchal

3.3.2 Avaliação da Tendência e Sazonalidade

A Figura 3.38 apresenta o ajustamento um modelo de regressão polinomial de grau um ($X_t = \beta_0 + \beta_1 t$) aos NMm's, de tendência β_1

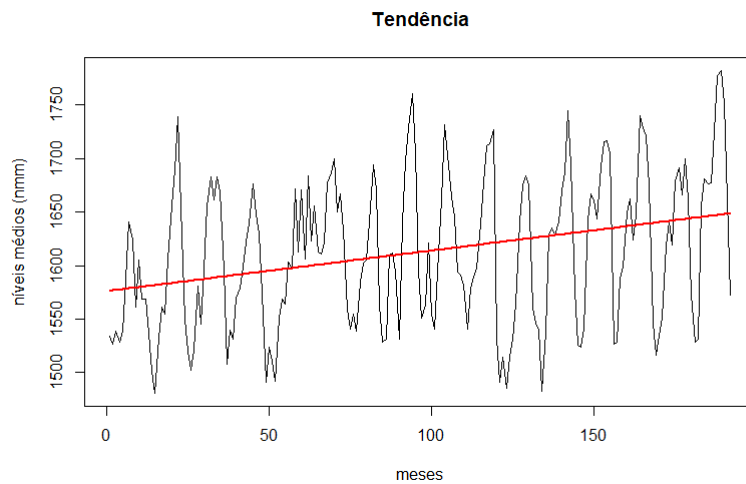


FIGURA 3.38: Linha preta: evolução dos NMm's no Funchal no período em análise; linha vermelha: modelo ajustado

A Figura 3.39 apresenta os coeficientes e os respetivos $p - value$ obtidos pelo *software R*.

```

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) 1.576e+03   9.720e+00 162.132 < 2e-16 ***
X            3.787e-01   8.735e-02   4.336 2.36e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

FIGURA 3.39: Coeficientes do modelo de polinomial de grau um estimado para o Funchal

Dado que $p - value \leq 0,05$, rejeitamos H_0 a favor de H_1 , assim conclui-se que os coeficientes são significativos.

Face ao declive positivo, $\beta_1 = 3,787 \times 10^{-1}$, pode-se concluir que o NMM está a aumentar.

Foi aplicado o teste de *Shapiro-Wilk* para averiguar a normalidade dos resíduos. Verifica-se que existe evidência estatística para rejeitar H_0 ao nível de significância de 5% ($W = 0,984$ com $p = 0,0278 \leq 0,05$), pelo que os resíduos não seguem uma distribuição normal.

Doravante, passaremos a referir-mo-nos ao modelo de regressão polinomial de grau um por modelo 1.

Ajustou-se um polinómio de grau dois ($X_t = \beta_0 + \beta_1 t + \beta_2 t^2$), com vista a determinar se a tendência sofreu aceleração 3.40.

```

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) 1.577e+03   1.471e+01 107.147 <2e-16 ***
X            3.578e-01   3.520e-01   1.016  0.311
I(X^2)       1.085e-04   1.767e-03   0.061  0.951
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

FIGURA 3.40: Coeficientes do modelo polinomial de grau dois estimado para a série de registos do Funchal

Obteve-se $p - value > 0,05$, pelo que concluímos que os coeficientes β_1 e β_2 não são significativos e como tal, não se verifica aceleração da tendência da subida de 2005 a 2015 (Figura 3.40).

Através do teste de *Shapiro-Wilk*, verifica-se que existe evidência estatística para rejeitar H_0 ao nível de significância de 5% ($W = 0,984$ com um valor de $p - value = 0,0272 \leq 0,05$), pelo que os resíduos não seguem uma distribuição normal.

A Figura 3.41 apresenta as distribuições do NMm's registados pelo marégrafo do Funchal entre 2005 e 2020

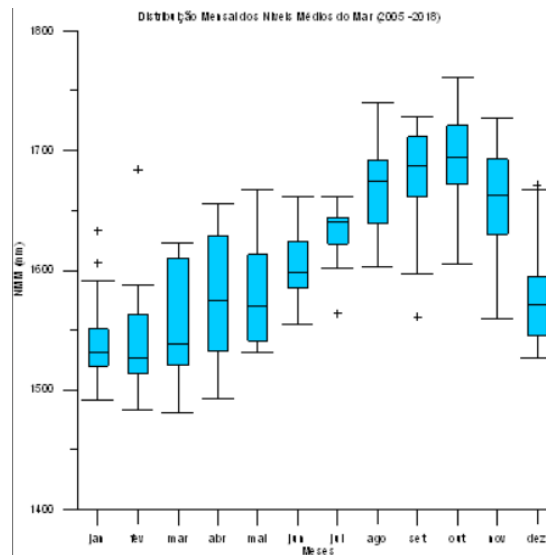


FIGURA 3.41: Comportamento dos níveis médios ao longo do ano, no Funchal no período de 2005 a 2020

Verificando-se níveis mais elevados no período compreendido entre os meses de agosto e novembro e níveis mais reduzidos entre os meses de janeiro e março.

O modelo de tendência polinomial e tendência cíclica que proporcionou o melhor ajustamento é dado pela expressão,

$$X_1 = \beta_0 + \beta_1 t + a_1 \cos(\omega_1 t) + b_1 \sin(\omega_1 t) + a_2 \cos(\omega_2 t) + b_2 \sin(\omega_2 t),$$

onde $\omega_1 = \frac{2\pi}{p_1}$ com $p_1 = 12$; e $\omega_2 = \frac{2\pi}{p_2}$ com $p_2 = 6$.

onde os coeficientes assumem os valores apresentados pela Figura 3.42.

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	1580.09090	6.04911	261.211	< 2e-16	***
X	0.33595	0.05438	6.178	4.00e-09	***
I(cos(2 * pi/12 * X))	-11.19496	4.25607	-2.630	0.009244	**
I(sin(2 * pi/12 * X))	-69.15214	4.26056	-16.231	< 2e-16	***
I(cos(4 * pi/12 * X))	-16.90887	4.25607	-3.973	0.000101	***
I(sin(4 * pi/12 * X))	-18.84956	4.25677	-4.428	1.62e-05	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

FIGURA 3.42: Coeficientes do modelo de tendência linear e tendência cíclica

Aplicou-se o teste de *Shapiro-Wilk* para averiguar a normalidade dos resíduos. Verifica-se que não existe evidência estatística para rejeitar H_0 ao nível de significância de 5% ($W = 0,9920$ com $p - value = 0,377 > 0,05$), pelo que os resíduos não seguem uma distribuição normal.

Doravante, passaremos a referir-nos a este modelo por modelo 3.

Utilizou-se o teste de *Wilcoxon* para comparar o modelo 3 com os modelos 1 e 2. Pela sua aplicação, mostra-se que as diferenças entre os modelos 1 e 3 e os modelos 2 e 3 são estatisticamente equivalentes ($W = 18208$ com um valor de $p - value = 0,837 > 0,05$ e $W = 18206$ com um valor de $p - value = 0,835 > 0,05$, respetivamente, entre os modelos 1 e 3 e entre os modelos 2 e 3).

A Tabela 3.14 apresenta o R^2 e $RMSE$ de cada uma dos modelos.

modelo	R^2	$RMSE$
modelo 1	0,09	66,73
modelo 2	0,09	66,73
modelo 3	0,66	41,40

TABELA 3.14: R^2 e $RMSE$ dos modelo em análise

O modelo 3 apresenta-se como sendo o modelo que melhor explica a variabilidade dos dados (maior R^2) e que possui o menor $RMSE$. Como tal, conclui-se que o modelo 3 reflete o melhor ajustamento.

A Figura 3.43 apresenta o ajustamento do modelo 3 à série de dados do Funchal.

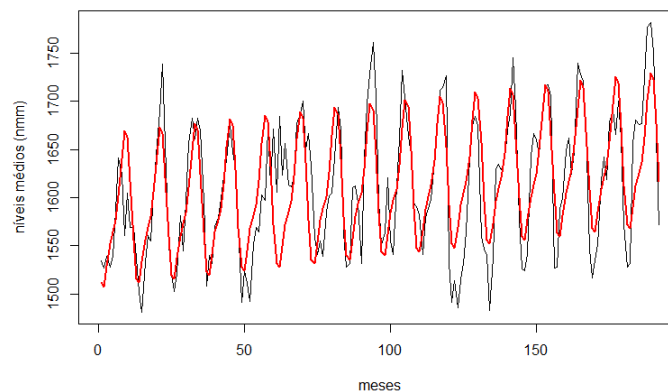


FIGURA 3.43: Linha preta: série de dados em análise (Funchal);
linha vermelha: modelo ajustado (modelo 3)

A Figura 3.44 apresenta os correlogramas das funções FAC e FACP residuais do modelo 3.

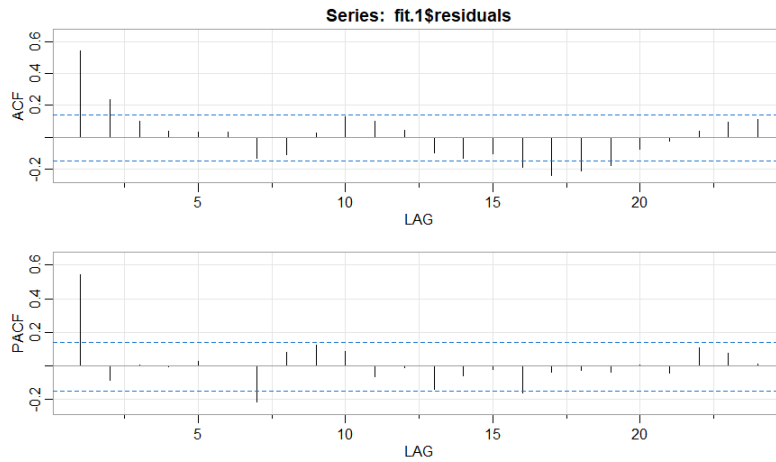


FIGURA 3.44: Correlogramas dos resíduos do modelo: FAC (gráfico superior) e FACP (gráfico inferior)

A função FAC apresenta um lento decaimento para zero e vários valores fora do intervalo de confiança a partir do “lag” 15. O comportamento da FAC mostra que os resíduos estão correlacionados.

Esta ideia é ainda confirmada pelo teste de *Ljung e Box*. Através da sua aplicação, rejeitamos H_0 a favor de H_1 , assim pode-se concluir que não existe evidência estatística de que os resíduos são independentes ao nível de significância 5% ($Q = 67,958$ com $p - value = 1,097 \times 10^{-10} \leq 0,05$).

A tendência obtida $T = 4,03$ mm/ano, dado que $\beta_1 = 0,3359$, é coerente com o trabalho de Mendes et al. (2017) que obteve T entre 2,13 e 5,77 mm/ano.

O modelo selecionado, mostra que: 1) existe uma oscilação periódica anual com amplitude de 70 mm/ano ($A_1 = \sqrt{(a_1)^2 + (b_1)^2}$); 2) existem duas oscilações periódicas semianuais com amplitudes de 25 mm/ano ($A_2 = \sqrt{(a_2)^2 + (b_2)^2}$).

Foi adicionado o coeficiente β_2 ao modelo 3, para verificar se, assim como no modelo 2, a tendência de subida do NMM não sofre aceleração. A Figura 3.45 apresenta os resultados dos coeficientes e os respetivos $p - values$. Obteve-se um valor de $p - value > 0,05$, pelo que concluímos que os coeficientes β_1 e β_2 não são significativos. Como tal, também neste modelo, existe evidência estatística de que não se verificou aceleração.

```

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.581e+03  9.930e+00 159.183  <2e-16 ***
X            3.123e-01  2.375e-01   1.315   0.1901
I(X^2)       1.359e-04  1.192e-03   0.114   0.9093
I(cos(2 * pi/12 * X)) -1.120e+01  4.631e+00  -2.418   0.0165 *
I(sin(2 * pi/12 * X)) -6.914e+01  4.636e+00 -14.915  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

FIGURA 3.45: Coeficientes do modelo 3 com a adição do coeficiente β_2

Foram ainda testados modelos com oscilações trimestrais e quadrimestrais. A Figura 3.46 apresenta os resultados dos coeficientes e os respetivos p – values. Obteve-se (p – value $> 0,05$), pelo que não se rejeitar H_0 ; conclui-se que os coeficientes não são significativos e como tal, não se verifica a existência de oscilações trimestrais e quadrimestrais.

<pre> Coefficients: Estimate Std. Error t value Pr(> t) (Intercept) 1580.1765 6.0514 261.127 < 2e-16 *** X 0.3351 0.0544 6.159 4.49e-09 *** I(cos(2 * pi/12 * X)) -11.1941 4.2571 -2.629 0.009275 ** I(sin(2 * pi/12 * X)) -69.1555 4.2616 -16.227 < 2e-16 *** I(cos(4 * pi/12 * X)) -16.9080 4.2571 -3.972 0.000102 *** I(sin(4 * pi/12 * X)) -18.8511 4.2578 -4.427 1.63e-05 *** I(cos(6 * pi/12 * X)) -0.4288 4.2571 -0.101 0.919875 I(sin(6 * pi/12 * X)) -5.8628 4.2571 -1.377 0.170129 --- Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 </pre>	<pre> Coefficients: Estimate Std. Error t value Pr(> t) (Intercept) 1580.04401 6.01867 262.524 < 2e-16 *** X 0.33644 0.05411 6.218 3.36e-09 *** I(cos(2 * pi/12 * X)) -11.19545 4.23379 -2.644 0.0089 ** I(sin(2 * pi/12 * X)) -69.15033 4.23826 -16.316 < 2e-16 *** I(cos(4 * pi/12 * X)) -16.90936 4.23379 -3.994 9.42e-05 *** I(sin(4 * pi/12 * X)) -18.84871 4.23448 -4.451 1.49e-05 *** I(cos(6 * pi/12 * X)) -0.43019 4.23379 -0.102 0.9192 I(sin(6 * pi/12 * X)) -5.86148 4.23379 -1.384 0.1679 I(cos(8 * pi/12 * X)) -8.50311 4.23379 -2.008 0.0461 * I(sin(8 * pi/12 * X)) -0.16660 4.23356 -0.039 0.9687 --- Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 </pre>
--	--

(A)

(B)

FIGURA 3.46: Coeficientes dos modelos com sazonalidade trimestral e quadrimestral: (A) modelo com 3 sinusoides; (B) modelo com 4 sinusoides.

3.3.3 Avaliação da aplicação de modelos ARIMA

O modelo $SARIMA(1, 0, 0)(0, 1, 1)_{12}$ foi identificado como sendo o modelo que melhor se ajusta ao comportamento geral dos registos maregráficos do Funchal. A avaliação dos restantes modelos identificados, pode ser consultada no Apêndice C.

A Tabela 3.15, apresenta os coeficientes encontrados no ajustamento do modelo $SARIMA(1, 0, 0)(0, 1, 1)_{12}$ aos dados em análise,

Modelo	θ_1	Φ_1
$SARIMA(1, 0, 0)(0, 1, 1)_{12}$	0,5619	-0,9999

TABELA 3.15: Coeficientes do modelos $SARIMA(1, 0, 0)(0, 1, 1)_{12}$

O modelo $SARIMA(1, 0, 0)(0, 1, 1)_{12}$ é dado por:

$$(1 - 0.5619B)(1 - B^{12})X_t = (1 + 0,999B^{12})\epsilon_t.$$

O modelo identificado encontrando-se em conformidade com os seguintes pressupostos:

- Coeficientes significativamente diferentes de zero. A Tabela 3.16 apresentam os intervalos de confiança de cada um dos coeficientes.

	2,5%	97,5%
θ_1	0,4219	0,7020
Φ_1	-1,2607	-0,7391

TABELA 3.16: Intervalos de confiança do modelo identificado

- Os valores das funções FAC e FACP residuais estão compreendidos entre os intervalos de confiança (Figura 3.47), exceto no “lag” 10 da FAC e nos “lag’s” 10 e 41 da FACP;

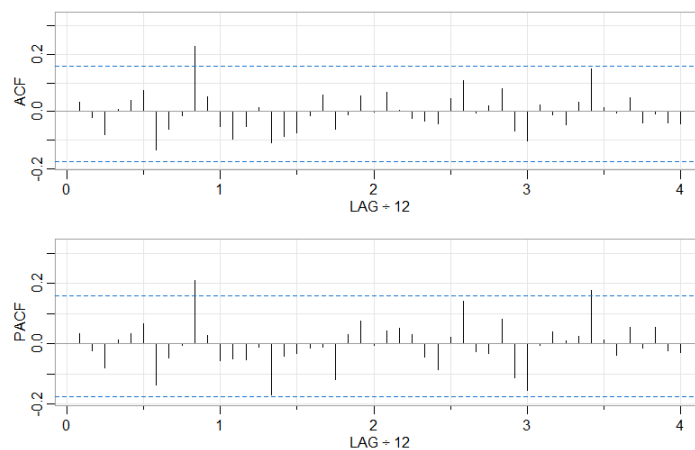


FIGURA 3.47: Correlogramas dos resíduos do modelo $SARIMA(1, 0, 0)(0, 1, 1)_{12}$: FAC (gráfico superior) e FACP (gráfico inferior)

- A hipótese de independência dos resíduos foi testada através do teste de Ljung e Box; pelo que não rejeitamos H_0 , assim pode-se concluir que existe evidência estatística de que os resíduos não estão correlacionados nível de significância 5% ($Q = 22,51$ com $p - value = 0,37 > 0,05$).

A Figura 3.48, apresenta o histograma dos resíduos do modelo. A média dos resíduos é 0,0054 m com intervalo de confiança a 95% (-0,0093 ; 0,02028).

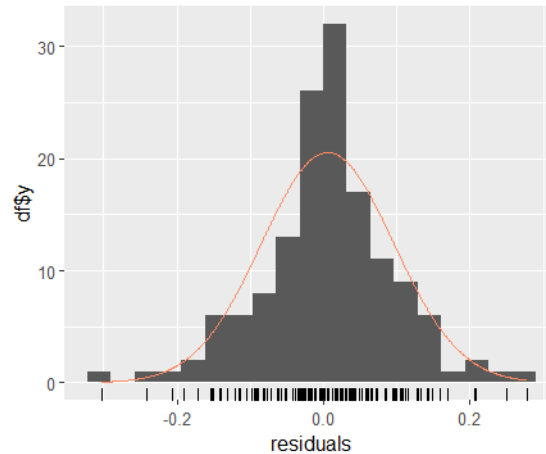


FIGURA 3.48: Histograma dos resíduos do modelo escolhido

Foi aplicado o teste de Shapiro-Wilk para averiguar a normalidade dos resíduos. Verifica-se que existe evidência estatística para rejeitar a normalidade dos resíduos ao nível de significância de 5% ($W = 0,9796$ com $pvalue = 0,0304 \leq 0,05$).

Na fase de avaliação dos modelos identificados, mostrou-se que os modelos $SARIMA(1, 0, 0)(0, 1, 1)_{12}$, $SARIMA(1, 0, 0)(2, 1, 0)_{12}$ e $SARIMA(0, 0, 2)(0, 1, 1)_{12}$, descrevem, de forma satisfatória, a série em análise, cumprindo com os pressupostos definidos. A avaliação dos dois últimos modelos, $SARIMA(1, 0, 0)(2, 1, 0)_{12}$ e $SARIMA(0, 0, 2)(0, 1, 1)_{12}$, pode ser consultada no Apêndice C.

Passaremos agora à escolha do modelo que melhor descreva a série em questão, tendo em conta os critérios AIC, BIC e $RMSE$.

Modelo	AIC	BIC	$RMSE$
$SARIMA(1, 0, 0)(0, 1, 1)_{12}$	-212,04	-200,51	32,98
$SARIMA(1, 0, 0)(2, 1, 0)_{12}$	-189,54	-175,13	39,00
$SARIMA(0, 0, 2)(0, 1, 1)_{12}$	-209,57	-195,16	33,00

TABELA 3.17: Critérios AIC e BIC;RMSE

A análise da Tabela 3.17 mostra que o modelo $SARIMA(1, 0, 0)(0, 1, 1)_{12}$ apresenta melhor desempenho nos três critérios, pelo que deve ser considerado o modelo que melhor se ajusta à série em estudo.

A Figura 3.49, ajustamento gráfico das previsões aos meses da série que não foram considerados na obtenção do modelo (os últimos 48 meses). A avaliação gráfica confirma a adequabilidade do modelo.

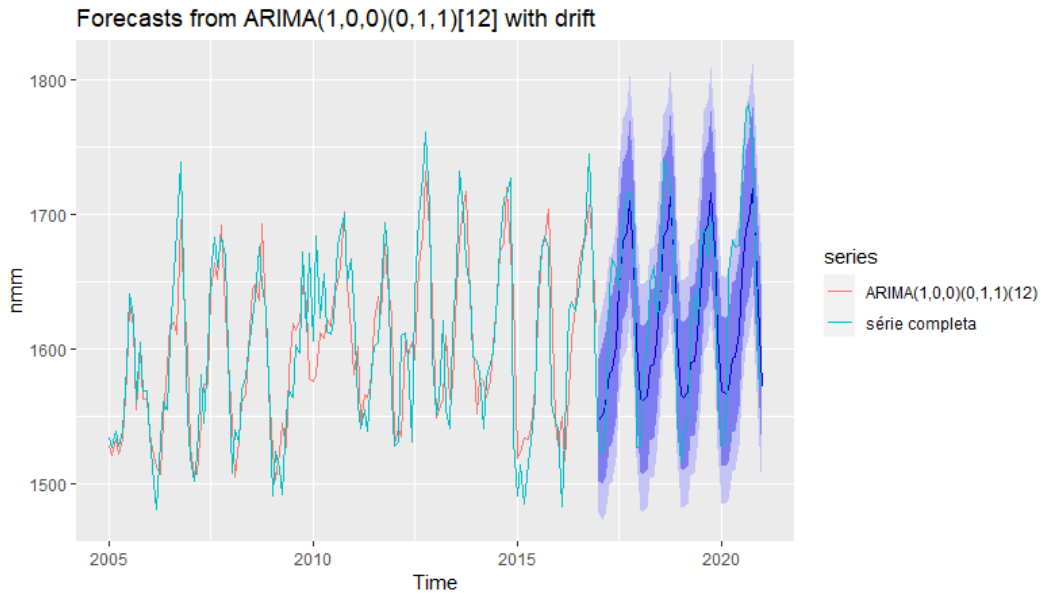


FIGURA 3.49: Ajustamento das previsões aos meses não considerados na análise para construção dos modelos para o Funchal

3.3.4 Seleção e Previsão

Recorreu-se ao teste de *Wilcoxon* para avaliar se os resíduos, obtidos por metodologias distintas (modelo 3 e $SARIMA(1, 0, 0)(0, 1, 1)_{12}$), são provenientes da mesma população. A hipótese de igualdade das medianas dos resíduos não foi rejeitada, para um nível de significância de 5% ($V = 28848$ com $p - value = 0,762 > 0,05$), isto é, existe evidência estatística para afirmar que os dois modelos são equivalentes.

A Tabela 3.18 apresenta o *RMSE* e o número de coeficientes de cada modelo.

Modelo	<i>RMSE</i>	nº de coeficientes
modelo 3	41,40	5
$SARIMA(1, 0, 0)(0, 1, 1)_{12}$	32,98	2

TABELA 3.18: *RMSE* e nº de coeficientes dos modelos 3 e $SARIMA(1, 0, 0)(0, 1, 1)_{12}$

O modelo $SARIMA(1, 0, 0)(0, 1, 1)_{12}$ apresenta o menor valor de $RMSE$ e é o modelo mais parcimonioso. Desta forma, conclui-se que o melhor ajustamento é conseguido com o modelo $SARIMA(1, 0, 0)(0, 1, 1)_{12}$.

Pode-se agora analisar as previsões que este modelo forneceria para um período de 10 anos, isto é de 2020 a 2030. Na Figura 3.50 apresentam-se as previsões obtidas e os correspondentes limites dos intervalos de confiança a 95%. Os valores destas previsões estão próximos de 1650 mm com uma ligeira tendência de crescimento.

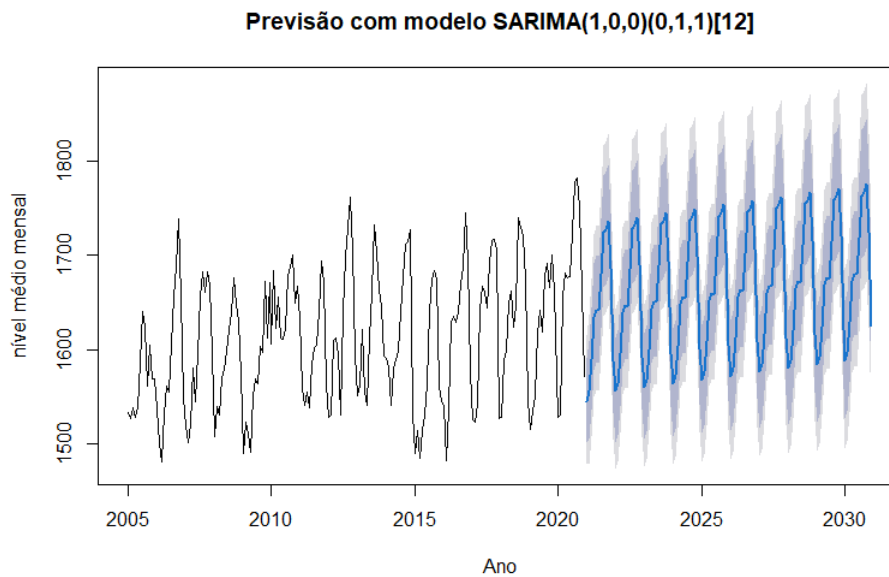


FIGURA 3.50: Previsão dos NMm's para os próximos 10 anos para o Funchal

Quantificou-se a tendência do NMM para os valores previstos, e podemos esperar um aumento de 55 mm até 2025, o que se traduz numa tendência de, aproximadamente, +5,52 mm/ano.

Assim como no caso dos registos de Viana do Castelo, também aqui, se identifica uma diferença entre a tendência observada na série e a tendência dos valores previstos. Esta situação deve-se, novamente, à utilização de coeficientes de médias móveis e da existência de uma tendência superior nos últimos registos. De facto, quando analisamos apenas os últimos 3 anos (aproximadamente 20% dos dados), verificamos uma tendência de 13,3 mm/ano, $\beta_1 = 1,112$ (Figura 3.51).

```

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) 1602.873    18.974   84.477  <2e-16 ***
X              1.112     0.541    2.056  0.0443 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

FIGURA 3.51: Ajustamento de um modelo polinomial de grau um aos últimos 3 anos de registos da estação maregráfica do Funchal

3.4 Semelhança Estatística entre a Sazonalidade nos Diferentes Registos Maregráficos Estudados

Como foi demonstrado até agora, as séries de registos maregráficos são fortemente influenciadas pelas variações anuais e semianuais. Nesta secção procura-se analisar se a sazonalidade apresenta um comportamento semelhante ou se difere entre os três portos estudados. Para isso, procura-se verifica a hipótese de que os níveis médios mensais característicos, entre os três portos analisados, apresentam diferenças estatisticamente significativas ao longo do ano.

A Figura 3.52 apresenta a evolução dos níveis médios mensais característicos das três séries de dados maregráficos. Salienta-se que foram acrescentadas cotas aos níveis médios característicos de Sines e Funchal para evitar a sobreposição.

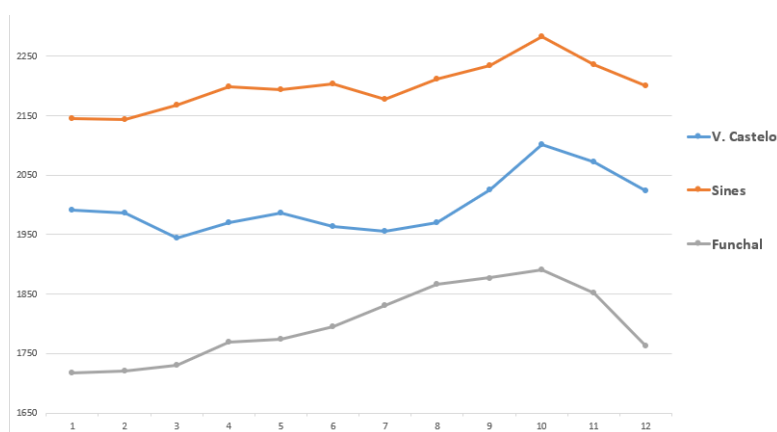


FIGURA 3.52: Níveis médios mensais característicos dos níveis médios das três séries de registos maregráficos

Foi aplicado o teste de *Shapiro-Wilk* para averiguar a normalidade dos níveis médios mensais característicos. Verifica-se que não existe evidência estatística para rejeitar a normalidade dos níveis médios mensais característicos ao nível

de significância de 5% ($W = 0,956$ com $p - value = 0,726 > 0,05$, para Sines; $W = 0,890$ com $p - value = 0,119 > 0,05$, para Viana do Castelo; e $W = 0,917$ com $p - value = 0,267 > 0,05$ para o Funchal).

Assegurado o pressuposto de normalidade das amostras, podemos agora, através do teste de ANOVA, testar a igualdade entre as médias amostrais. Através da sua aplicação, verificamos que,

- Rejeitamos H_0 a favor de H_1 ; existe evidência de que a diferença entre os níveis médios mensais característicos de Viana do Castelo e Sines é estatisticamente significativa ($F = 12,47$ com $p - value = 0,00543 \leq 0,05$);
- Existe evidência de que a diferença entre os níveis médios mensais característicos de Sines e Funchal é estatisticamente significativa ($F = 27,47$ com $p - value = 0,000378 \leq 0,05$);
- Existe evidência de que a diferença entre os níveis médios mensais característicos de Viana do Castelo e Funchal não é estatisticamente significativa ($F = 3,972$ com $p - value = 0,0742 > 0,05$);

Uma vez que o nível de referência, o ZH, não é igual nas três estações maregráficas ⁸, procedeu-se ao mesmo ensaio com os os níveis médios mensais característicos referenciados ao mesmo *datum*, o *RLR*. Na Tabela 3.19 são apresentados os valores das cotas adicionas aos registos, por forma a homogeneizar o nível de referência.

Estação maregráfica	cota (mm)
Viana do Castelo	4890
Sines	4930
Funchal	5382

TABELA 3.19: Cota adicionada aos registos maregráficos (PSMSL, s.d.)

Através da aplicação do Teste de ANOVA verificamos que,

- Rejeitamos H_0 a favor de H_1 ; existe evidência de que a diferença entre os níveis médios mensais característicos de Viana do Castelo e Sines é estatisticamente significativa ($F = 12,56$ com $p - value = 0,00532 \leq 0,05$);

⁸O ZH em Sines e Viana do Castelo está estabelecido 2,00 metros abaixo do NMA. No Arquipélagos da Madeira, o ZH situa-se 1,40 metros abaixo do NMA (IH, 2020).

- Existe evidência de que a diferença entre os níveis médios mensais característicos de Sines e Funchal é estatisticamente significativa ($F = 27,08$ com $p - value = 0,000399 \leq 0,05$);
- Existe evidência de que a diferença entre os níveis médios mensais característicos de Viana do Castelo e Funchal não é estatisticamente significativa ($F = 3,972$ com $p - value = 0,0742 > 0,05$);

A aplicação do teste de ANOVA apresenta os mesmos resultados para os dois ensaios. Podemos assim concluir que as variações anuais e semianuais de Viana do Castelo e Funchal apresentam semelhanças. No entanto, não fica evidente a existência de semelhança entre Sines e as duas restantes estações maregráficas analisadas.

Os resultados obtidos são consistentes com trabalhos já efetuados no IH (Figura 3.53):

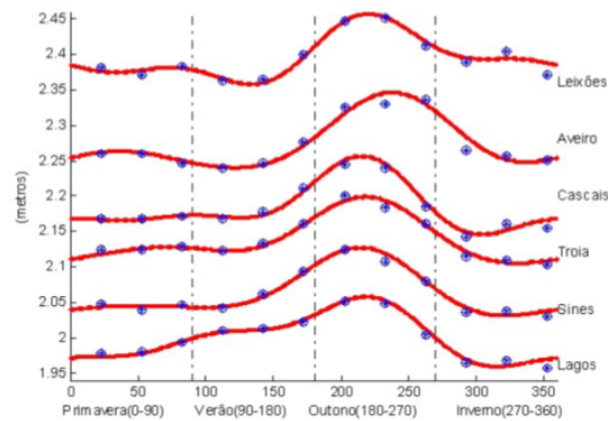


FIGURA 3.53: Níveis médios mensais característicos (pontos azuis) com ajustamento dos níveis médios diários (vermelho), (fonte: Carinhas et al., 2021)

A análise da figura mostra que, ao longo da costa portuguesa, se verifica um comportamento sazonal semelhante.

Conclusão

Neste trabalho, procurou-se contribuir para o estudo das variações do NMM na costa portuguesa a partir dos dados recolhidos pelas estações maregráficas de Sines, Viana do Castelo e Funchal. Os resultados apresentados neste trabalho, mostram que as três estações maregráficas manifestam uma tendência de subida do NMM. No entanto, a subida não é uniforme ao longo da costa portuguesa, notando-se uma elevada divergência entre Viana do Castelo e as duas restantes estações estudadas. A série de registos de Viana do Castelo, a mais longa estudada, permitiu apurar que o NMM está a subir com uma tendência de 1,97 mm/ano. Por sua vez, o NMM em Sines e no Funchal, apresenta, respetivamente, uma tendência de 4,27 mm/ano e 4,03 mm/ano, sensivelmente o dobro da tendência verificada em Viana do Castelo. A Península Ibérica apresenta movimento de subsidência com um valor médio de, aproximadamente, 0,6 mm/ano, pelo que a ordem de grandeza das três tendências aponta para uma subida predominantemente eustática. Os resultados apresentam uma clara concordância com estudos já realizados para as três estações maregráficas. Não obstante, verificam-se diferenças significativas a nível regional. As tendências estimadas para Sines e Funchal apresentam-se superiores às tendências estimadas, em trabalhos anteriores, para Cascais ($1,7 \pm 0,2$ mm/ano) e Lagos ($1,2 \pm 0,2$ mm/ano). Não se identificou, em nenhuma das estações maregráficas, aceleração da tendência de subida do NMM.

Mostrou-se que o nível do mar não é constante ao longo do ano. As variações do nível do mar são uma consequência de dois movimentos: 1) uma oscilação periódica anual; 2) duas oscilações periódicas semianuais. O primeiro movimento é muito mais expressivo e caracteriza-se por aumento gradual do nível do mar entre o primeiro trimestre do ano e o mês de outubro, apresentado, de seguida, uma redução acentuada. O segundo movimento, decompõe-se em duas oscilações com periodicidade semianual que, se materializam através da existência de um mínimo e máximo relativo que ocorrem ao longo do segundo e terceiro trimestre do ano. Verificou-se que, no caso do Funchal, a variação do nível do mar resulta, quase na totalidade, da oscilação periódica anual, pelo que os mínimos e máximos relativos têm muito pouco significado.

O comportamento do nível do mar ao longo do ano apresenta concordância com estudos já realizados para o porto de Sines. Todavia, devido à inexistência de estudos referentes à sazonalidade para as duas restantes estações maregráficas, a comparação apenas é possível com estudos da costa portuguesa que, em grande parte, são produzidos com séries de dados maregráficos de Cascais e Lagos. Sendo que estes, revelam um comportamento ligeiramente diferente, com os valores mais reduzidos a ocorrerem no segundo trimestre do ano. Apresentando, no entanto, concordância no período em que ocorrem os valores mais elevados.

Os modelos utilizados mostram que as amplitudes das variações anuais do nível do mar são consideravelmente superiores à tendência verificada. Esta realidade vem confirmar a necessidade de se utilizar registos maregráficos, com janelas temporais suficientemente representativas, em detrimento do aumento da incerteza das estimativas das tendências, provocado pelas oscilações anuais e semianuais.

A análise das variações do NMM ao longo do ano conduziu a resultados inesperados, com as estações maregráficas de Viana do Castelo e Funchal a apresentarem comportamento semelhante, porém, não fica evidente a existência de semelhança entre Sines e as duas restantes estações maregráficas.

O ensaio de diferentes tipos de modelos, às séries maregráficas das três estações, revelou que os modelos da classe ARIMA proporcionam o melhor ajustamento. Os modelos encontrados são:

1. $(1 - 0,2103B)(1 - B)(1 - B^{12})X_t = (1 + 0,8734B)(1 + 0,9992B^{12})\epsilon_t$, para Sines;
2. $(1 - 0,2814B)(1 - B^{12})X_t = (1 + 0,9089B^{12})\epsilon_t$, para Viana do Castelo;
3. $(1 - 0,5619B)(1 - B^{12})X_t = (1 + 0,999B^{12})\epsilon_t$, para o Funchal.

Sendo B o operador de atraso e ϵ_t o ruído branco presente em cada uma das séries.

Foram feitas previsões para um horizonte de 10 anos através dos modelos ARIMA encontrados. As tendências dos valores previstos são: +3 mm/ano (Viana do Castelo), + 4,36 mm/ano (Sines) e + 5,52 mm/ano (Funchal). Os valores estimados acompanham o comportamento geral das séries, contudo, as tendências dos valores previstos, tanto em Viana do Castelo como no Funchal, são substancialmente superiores às tendências das séries dos registos maregráficos. As duas tendências diferem 1 mm/ano e 1,5 mm/ano, respetivamente, para Viana do Castelo e Funchal.

Devido à volatilidade da interação climático-oceânica é muito provável que, tanto as estimativas como as previsões, sofram alterações em escalas de tempo de

curto e médio prazo. Não obstante, se as tendências estimadas não sofrerem alterações, podemos antever que, até ao final do século, o NMM esteja, aproximadamente, 43 cm acima do atual, na perspetiva mais pessimista dos três cenários.

3.5 Recomendações Futuras

Devemos continuar a melhorar a previsão, especialmente a precisão das previsões de séries temporais; uma tarefa importante, embora muitas vezes difícil, que os decisores enfrentam em muitas áreas e a oceanografia não é exceção. Assim, com base neste trabalho, surgem as seguintes perspetivas:

- Desenvolver, testar e comparar com os resultados obtidos neste trabalho, metodologias que apliquem redes neuronais no estudo da evolução do NMM;
- Aplicar a metodologia utilizada neste trabalho a outros portos nacionais, em especial na Região Autónoma dos Açores;
- Quantificar as causas da atual subida do NMM em Portugal;
- Incorporar as tendências de movimentos isostáticos nas várias tendências estimadas, obtendo assim, a evolução do NMM absoluto para as três estações analisadas;
- Utilizar séries de dados, com uma janela temporal suficientemente representativa de todos os fenómenos que provocam flutuações no NMM. Recomenda-se a análise de um período superior à duração do ciclo nodal lunar; o fenómeno com maior duração, que se expande por um período aproximado de 18,6 anos. No caso dos registos maregráficos do Funchal, a janela temporal utilizada, 16 anos, é claramente insuficiente para inibir o efeito do ciclo nodal lunar nas variações do NMM. Pelo que se aconselha a aplicação da mesma metodologia ao porto do Funchal, utilizando séries de registos maregráficos com duração superior a 18,6 anos.

Apesar dos inúmeros modelos de séries temporais disponíveis, a pesquisa para melhorar a eficácia dos modelos de previsão nunca para.

Bibliografia

- Afonso, A. & Nunes, C. (2019). Probabilidades e Estatística. Aplicações e Soluções em SPSS. Versão revista e aumentada.
- Akaike, H. (1974). A new look at the statistical model identification. *IEEE transactions on automatic control*, 19(6), 716–723.
- Andri et mult. al., S. (2021). *DescTools: Tools for Descriptive Statistics* [R package version 0.99.44].
- Antunes, C. (2011). Variação actual do NMM em Cascais. *VII Conferência Nacional de Cartografia e Geodesia*.
- Antunes, C. (2016). Subida do Nivel Médio do Mar em Cascais, revisão da taxa actual. *Actas das*, 4.
- Antunes, C. & Taborda, R. (2009). Sea level at Cascais tide gauge: data, analysis and results. *Journal of Coastal Research*, 218–222.
- Berger, A. (1988). Milankovitch theory and climate. *Reviews of geophysics*, 26(4), 624–657.
- Boski, T. (2013). O nível do mar - Historia e perspectiva de investigação. *Interações Homem-Meio nas Zonas Costeiras Brasil/Portugal*, 235–242.
- Bradshaw, E., Woodworth, P., Hibbert, A., Bradley, L., Pugh, D., Fane, C. & Bingley, R. (2016). A century of sea level measurements at Newlyn, Southwest England. *Marine Geodesy*, 39(2), 115–140.
- Cacho, J. (2020). As potencialidades do Porto de Sines [acedido em 10 de maio de 2022, disponível em: <https://adfersit.pt/noticia/439/as-potencialidades-do-porto-de-sines>]. *ADFERSIT-Associação Portuguesa para o Desenvolvimento dos Sistemas Integrados de Transportes*.
- Carinhas, D., Infante, P., Martinho, A. & Vasquez, F. (2021). As marés ao serviço da variação do nível médio do mar. *Apresentado no XXV Congresso da Sociedade Portuguesa de Estatística – SPE 2021, Évora*.
- Cazenave, A. & Nerem, R. S. (2004). Present-day sea level change: Observations and causes. *Reviews of Geophysics*, 42(3).
- Cazenave, A. & Remy, F. (2011). Sea level and climate: measurements and causes of changes. *Wiley Interdisciplinary Reviews: Climate Change*, 2(5), 647–662.

- Chan, K.-S. & Ripley, B. (2020). *TSA: Time Series Analysis* [R package version 1.3].
- Church, J. A., Clark, P. U., Cazenave, A., Gregory, J. M., Jevrejeva, S., Levermann, A., Merrifield, M. A., Milne, G. A., Nerem, R. S., Nunn, P. D. et al. (2013). *Sea level change* (rel. téc.). PM Cambridge University Press.
- Church, J. A. & White, N. J. (2006). A 20th century acceleration in global sea-level rise. *Geophysical research letters*, 33(1).
- Church, J. A. & White, N. J. (2011). Sea-level rise from the late 19th to the early 21st century. *Surveys in geophysics*, 32(4), 585–602.
- COI. (1985). *Manual de Medição e Interpretação do Nível do Mar* (rel. téc.). COI-Comissão Oceanográfica Internacional, Manuais e Guias No.14, UNESCO.
- Crowley, T. J. & Kim, K.-Y. (1994). Milankovitch forcing of the last interglacial sea level. *Science*, 265(5178), 1566–1568.
- Cryer, J. D. & Chan, K.-S. (2008). *Time series analysis: with applications in R* (Vol. 2). Springer.
- Da Silva, A. A. A., Freire, E. & Crisóstomo, G. (2014). Variações do nível médio anual do mar em Cascais: características e tendências. *Estudos do Quaternário/Quaternary Studies*, (5).
- Dalazoana, R., Luz, R. T. & Freitas, S. (2005). Estudos do NMM a partir de séries temporais maregráficas e de Altimetria por satélites visando a integração da rede vertical brasileira ao SIRGAS. *Rev. Bras. Cartogr*, 57, 140–153.
- Dias, J. & Taborda, R. (1988). Evolução recente do nível médio do mar em Portugal. *Anais do Instituto hidrográfico*, 9, 83–97.
- Eldredge, N. (2002). *Life on Earth: AG* (Vol. 1). ABC-CLIO.
- Engelhart, S. E. & Horton, B. P. (2012). Holocene sea level database for the Atlantic coast of the United States. *Quaternary Science Reviews*, 54, 12–25.
- ESA. (s.d.). Copernicus Sentinel-6 Michael Freilich [acedido em 10 de maio de 2022, disponível em shorturl.at/hkxy4].
- Ferreira, Ó., Dias, J. A. & Taborda, R. (2008). Implications of sea-level rise for continental Portugal. *Journal of Coastal Research*, 24(2), 317–324.
- Fisher, R. A. & Bischel, W. (1973). The role of linear dispersion in plane-wave self-phase modulation. *Applied Physics Letters*, 23(12), 661–663.
- Garcia-Artola, A., Stéphan, P., Cearreta, A., Kopp, R. E., Khan, N. S. & Horton, B. P. (2018). Holocene sea-level database from the Atlantic coast of Europe. *Quaternary Science Reviews*, 196, 177–192.

- Gehrels, W. R., Hayward, B. W., Newnham, R. M. & Southall, K. E. (2008). A 20th century acceleration of sea-level rise in New Zealand. *Geophysical Research Letters*, 35(2).
- Gehrels, W. R., Marshall, W. A., Gehrels, M. J., Larsen, G., Kirby, J. R., Eiriksson, J., Heinemeier, J. & Shimmield, T. (2006). Rapid sea-level rise in the North Atlantic Ocean since the first half of the nineteenth century. *The Holocene*, 16(7), 949–965.
- Goodwin, I. D. & Harvey, N. (2008). Subtropical sea-level history from coral microatolls in the Southern Cook Islands, since 300 AD. *Marine Geology*, 253(1-2), 14–25.
- Griggs, D. J. & Noguer, M. (2002). Climate change 2001: the scientific basis. Contribution of working group I to the third assessment report of the intergovernmental panel on climate change. *Weather*, 57(8), 267–269.
- Grolemund, G. & Wickham, H. (2011). Dates and Times Made Easy with lubridate. *Journal of Statistical Software*, 40(3), 1–25.
- Grotzinger, J. (1986). Upward shallowing platform cycles: A response to 2.2 billion years of low-amplitude, high-frequency (Milankovitch band) sea level oscillations. *Paleoceanography*, 1(4), 403–416.
- Hamlington, B., Leben, R., Kim, K.-Y., Nerem, R., Atkinson, L. & Thompson, P. (2015). The effect of the El Niño-Southern Oscillation on US regional and coastal sea level. *Journal of Geophysical Research: Oceans*, 120(6), 3970–3986.
- Hamner, B. & Frasco, M. (2018). *Metrics: Evaluation Metrics for Machine Learning* [R package version 0.1.4].
- Hansel, A. K. & McKay III, E. D. (2010). Quaternary period. *Geology of Illinois*, 216–247.
- Hyndman, R., Athanasopoulos, G., Bergmeir, C., Caceres, G., Chhay, L., O’Hara-Wild, M., Petropoulos, F., Razbash, S., Wang, E. & Yasmeen, F. (2022a). *forecast: Forecasting functions for time series and linear models* [R package version 8.16].
- Hyndman, R., Athanasopoulos, G., Bergmeir, C., Caceres, G., Chhay, L., O’Hara-Wild, M., Petropoulos, F., Razbash, S., Wang, E. & Yasmeen, F. (2022b). *forecast: Forecasting functions for time series and linear models* [R package version 8.16].
- IH. (2020). *Tabela de Marés volume I - Portugal 2021* (rel. téc.). IH - Instituto Hidrográfico, Lisboa-Portugal.

- IH. (2021). O Instituto Hidrográfico [acedido em 07 de setembro de 2022, disponível em <https://www.hidrografico.pt/op/1>].
- IOC. (2002). *Manual on sea level measurement and interpretation. Volume III- Reappraisals and Recommendations as of the year 2000*. (rel. téc.). IOC - Intergovernmental Oceanographic Commission, Manual e Guides No.14, UNESCO.
- Jevrejeva, S., Moore, J. C., Grinsted, A. & Woodworth, P. L. (2008). Recent global sea level acceleration started over 200 years ago? *Geophysical Research Letters*, 35(8).
- Jevrejeva, S., Moore, J. C. & Grinsted, A. (2010). How will sea level respond to changes in natural and anthropogenic forcings by 2100? *Geophysical research letters*, 37(7).
- Kemp, A. C., Horton, B. P., Donnelly, J. P., Mann, M. E., Vermeer, M. & Rahmstorf, S. (2011). Climate related sea-level variations over the past two millennia. *Proceedings of the National Academy of Sciences*, 108(27), 11017–11022.
- Lambeck, K., Anzidei, M., Antonioli, F., Benini, A. & Esposito, A. (2004). Sea level in Roman time in the Central Mediterranean and implications for recent change. *Earth and Planetary Science Letters*, 224(3-4), 563–575.
- Mann, M. E. (2002). Little ice age. *Encyclopedia of global environmental change*, 1(504), e509.
- Maroco, J. (2007). *Análise Estatística-Com Utilização do SPSS* (Edições Sílabo, Lda).
- Martins, L. & Reis, J. (2007). Nível médio, maré e análise harmónica. *5^a Jornadas Portuguesas de Engenharia Costeira e Portuária*.
- Mendes, V., Barbosa, S. & Carinhas, D. (2020). Vertical land motion in the Iberian Atlantic coast and its implications for sea level change evaluation. *Journal of Applied Geodesy*, 14(3), 361–378.
- Mendes, V., Barbosa, S., Romero, I., Madeira, J. & Brum da Silveira, A. (2017). Vertical land motion and sea level change in Macaronesia. *Geophysical Journal International*, 210(2), 1264–1280.
- Morettin, P. A. & Toloi, C. (2006). Análise de séries temporais. *Análise de séries temporais* (pp. 538–538).
- Moura, C. S. (2011). Física para o ensino médio: gravitação, eletromagnetismo e física moderna.
- Murteira, B. & Antunes, M. (2012). *Probabilidades e Estatística”, Volume II*. Livraria Escolar Editora.
- Murteira, B. (1990). *Probabilidades e estatística*.

- Murteira, B., Müller, D. & Turkman, K. F. (2000). *Análise de sucessões cronológicas*. NASA. (s.d.). Overview: Weather, Global Warming and Climate Change [acedido em 8 de abril de 2022, disponível em shorturl.at/AGQX5].
- NASA. (2020). Milankovitch (Orbital) Cycles and Their Role in Earth's Climate [acedido em 20 de junho de 2022, disponível em shorturl.at/bdEZ8].
- Nott, J. (2017). How climate and sea level have changed over the long-term past.
- Nunes, C. & Afonso, A. (2005). Apontamentos de Introdução às Probabilidades e à Estatística - Volume II.
- Pereira, S. (2020). Porto do Funchal recebe quase 600 mil passageiros e destrona Lisboa da liderança [acedido em 10 de maio de 2022, disponível em: shorturl.at/DHNPS].
- Pereira, Z. & Requeijo, J. (2012). Qualidade: Planeamento e Controlo Estatístico de Processos (Quality: Statistical Process Control and Planning). *FCT/UNL Foundation Editor, Lisboa*.
- Picoto, M., Carinhas, D. & Lemos, C. (2022a). Caracterização da tendência de subida do nível médio em Portugal. *7.ª Jornadas de Engenharia Hidrográfica, 2.ª Jornadas Luso-Espanhola de Hidrografia*, 161–164.
- Picoto, M., Carinhas, D. & Lemos, C. (2022b). Caracterização da tendência de subida do nível médio em Portugal. *Jornadas do Mar 22. Marín-Espanha*.
- PSMSL. (s.d.). Revised Local Reference (RLR) Definition [acedido em 27 de julho de 2022, disponível em <https://www.psmsl.org/data/obtaining/rlr.php>].
- Quinino, R. C., Reis, E. A. & Bessegato, L. F. (1991). O coeficiente de determinação R^2 como instrumento didático para avaliar a utilidade de um modelo de regressão linear múltipla. *Belo Horizonte: UFMG*.
- R Core Team. (2021). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria.
- Rocha, C., Antunes, C. & Catita, C. (2018). Estudo da vulnerabilidade costeira à subida do nível médio do mar para Portugal Continental como consequência das alterações climáticas. *Atas das 5as Jornadas de Engenharia Hidrográfica*, 428–431.
- Rovere, A., Stocchi, P. & Vacchi, M. (2016). Eustatic and relative sea level changes. *Current Climate Change Reports*, 2(4), 221–231.
- Ryan, J. A. & Ulrich, J. M. (2020). *xts: eXtensible Time Series* [R package version 0.12.1].
- Santos, J. L. C. d. (2011). *Modelação de séries temporais em R-Séries com suporte em números reais e números inteiros* (tese de mestrado). Universidade de Évora.

- Schauberger, P. & Walker, A. (2021). *openxlsx: Read, Write and Edit xlsx Files* [R package version 4.2.5].
- Shapiro, S. S. & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, 52(3/4), 591–611.
- Sivan, D., Lambeck, K., Toueg, R., Raban, A., Porath, Y. & Shirman, B. (2004). Ancient coastal wells of Caesarea Maritima, Israel, an indicator for relative sea level changes during the last 2000 years. *Earth and Planetary Science Letters*, 222(1), 315–330.
- Smith, D., Harrison, S., Firth, C. R. & Jordan, J. T. (2011). The early Holocene sea level rise. *Quaternary Science Reviews*, 30(15-16), 1846–1860.
- Taborda, R. & Dias, A. (1988). *O nível do mar em Portugal. Tendências Sazonais e Seculares* (rel. téc.).
- Velosa, S. & Pestana, D. (2010). Introdução à Probabilidade e à Estatística. *Lisboa: Fundação Calouste Gulbenkian*.
- Vieira, P. C. (1981). Variações do nível marinho. Alterações eustáticas no Quaternário. *Revista do Instituto Geológico*, 2(1), 39–57.
- Wang, H., Liu, K., Wang, A., Feng, J., Fan, W., Liu, Q., Xu, Y. & Zhang, Z. (2018). Regional characteristics of the effects of the El Niño-Southern Oscillation on the sea level in the China Sea. *Ocean Dynamics*, 68(4), 485–495.
- Wanner, H., Solomina, O., Grosjean, M., Ritz, S. P. & Jetel, M. (2011). Structure and origin of Holocene cold events. *Quaternary Science Reviews*, 30(21-22), 3109–3123.
- Webb, P. (2021). *Introduction to oceanography*. Roger Williams University.
- Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.
- Willmott, C. J. (1981). On the validation of models. *Physical geography*, 2(2), 184–194.

Apêndice A - Avaliação dos Modelos Identificados - Registos Maregráficos de Sines

Considerando a série estacionária, apresenta-se os modelos que aparentam descrever, satisfatoriamente, o comportamento geral da série, Tabela A.1:

$SARIMA(0, 1, 1)(0, 1, 1)_{12}$
$SARIMA(0, 1, 1)(1, 1, 0)_{12}$
$SARIMA(0, 1, 1)(1, 1, 1)_{12}$
$SARIMA(11, 1, 0)_{12}$
$SARIMA(1, 1, 1)(0, 1, 1)_{12}$

TABELA A.1: modelos identificados

Após se ter identificado os modelos, que numa primeira fase, essencialmente exploratória, garantem um bom ajustamento aos dados em análise, procura-se agora, analisar a qualidade estatística e do ajustamento de cada modelo. De salientar que, um modelo deve satisfazer, simultaneamente, os pressupostos da qualidade estatística e do ajustamento para que, seja considerado um modelo adequado.

A.1 $SARIMA(0, 1, 1)(1, 1, 0)_{12}$

A Tabela A.2, apresenta os coeficientes encontrados para o modelo,

Modelo	ϕ_1	Θ_1
$SARIMA(0, 1, 1)(1, 1, 0)_{12}$	-0,7289	-0,5194

TABELA A.2: Coeficientes do modelos $SARIMA(0, 1, 1)(1, 1, 0)_{12}$

O modelo $SARIMA(0, 1, 1)(1, 1, 0)_{12}$ é dado por:

$$(1 + 0,5194B^{12})(1 - B)(1 - B^{12})X_t = (1 + 0,7289B)\epsilon_t,$$

Análise dos pressupostos que garantem a qualidade estatística e do ajustamento:

1. Coeficientes significativamente diferentes de zero. A Tabela A.3 apresenta os intervalos de confiança de cada um dos coeficientes estimados. Visto que

nenhum dos intervalos contém o valor zero, ou seja, os intervalos variaram apenas entre valores negativos, todos eles são considerados significativos.

	2,5%	97,5%
ϕ_1	-0,8451	-0,6128
Θ_1	-0.6380	-0,4009

TABELA A.3: Intervalos de confiança dos coeficientes

- Os cronogramas da FAC e FACP dos resíduos, apresentam valores que, na generalidade, estão compreendidos entre os limites de confiança. No entanto, os “Lag’s” que se encontram fora do intervalo de confiança, apresentam valores de FAC e FACP muito elevados, nomeadamente, nos “Lag’s” 10 e 24 (Figura A.1);

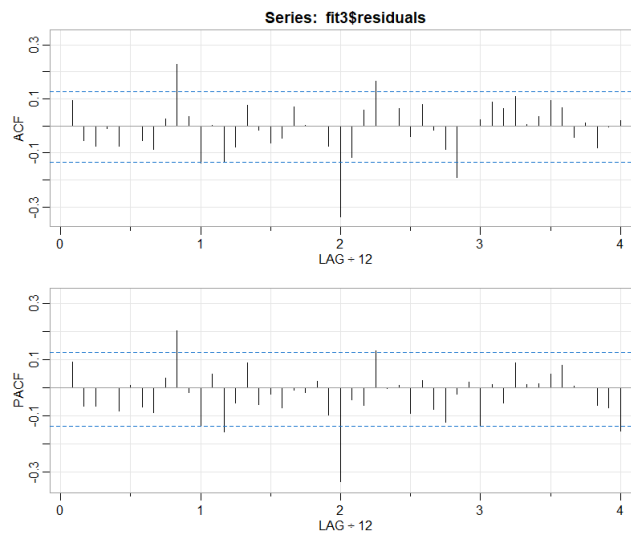


FIGURA A.1: Correlogramas dos resíduos do modelo $SARIMA = (0, 1, 1)(1, 1, 0)_{12}$: FAC (gráfico superior) e FACP (gráfico inferior)

- Os resíduos estão correlacionados. Através do teste de *Ljung e Box*, rejeitamos H_0 a favor de H_1 , assim pode-se concluir que existe evidência estatística de que os resíduos estão correlacionados nível de significância 5% ($Q = 66,73$ com $p - value = 2,141 \times 10^{-06} \leq 0,05$);

A Figura A.2, apresenta o histograma dos resíduos do modelo. A média dos resíduos é 0,1645 m com intervalo de confiança a 95% (-4,8537 ; 4,6797).

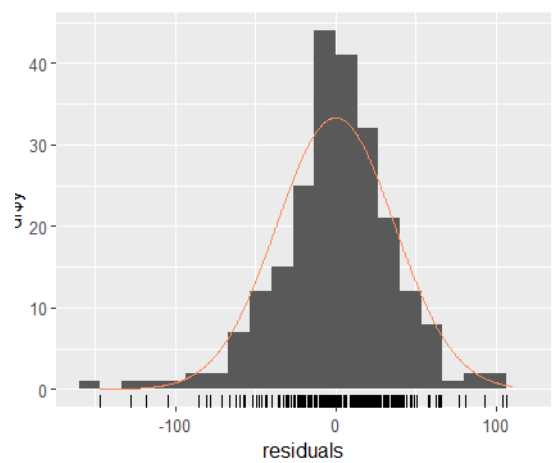


FIGURA A.2: Histograma das frequências dos resíduos do modelo

Foi aplicado o teste de *Shapiro-Wilk* para averiguar a normalidade dos resíduos. Verifica-se que existe evidência estatística para rejeitar a normalidade dos resíduos ao nível de significância de 5% ($W = 0,9835$ com $p\text{-value} = 0,0089 \leq 0,05$).

Devido à estrutura de dependência dos resíduos, consideramos que o modelo $SARIMA(0, 1, 1)(1, 1, 0)_{12}$, não cumpre com os pressupostos da qualidade do ajustamento, assim conclui-se que o modelo não é apropriado para descrever a série de dados maregráficos de Sines.

A.2 $SARIMA(0, 1, 1)(0, 1, 1)_{12}$

A Tabela A.4, apresenta os coeficientes encontrados para o modelo,

Modelo	ϕ_1	Φ_1
$SARIMA(0, 1, 1)(0, 1, 1)_{12}$	-0,7517	-0,9999

TABELA A.4: Coeficientes do modelos $SARIMA(0, 1, 1)(0, 1, 1)_{12}$

O modelo $SARIMA(0, 1, 1)(0, 1, 1)_{12}$ é dado por:

$$(1 - B)(1 - B^{12})X_t = (1 + 0,7517B)(1 + 0,9999B^{12})\epsilon_t,$$

Análise dos pressupostos que garantem a qualidade estatística e do ajustamento:

1. Coeficientes significativamente diferentes de zero. A Tabela A.5 apresenta os intervalos de confiança de cada um dos coeficientes estimados. Visto que nenhum dos intervalos contém o valor zero, ou seja, os intervalos variaram apenas entre valores negativos ou positivos, todos eles são considerados significativos.

	2,5%	97,5%
ϕ_1	-0,8733	-0,6301
Φ_1	-1,2610	-0,7389

TABELA A.5: Intervalos de confiança dos coeficientes

2. Os cronogramas da FAC e FACP dos resíduos, apresentam alguns valores fora dos intervalos de confiança; (Figura A.3);

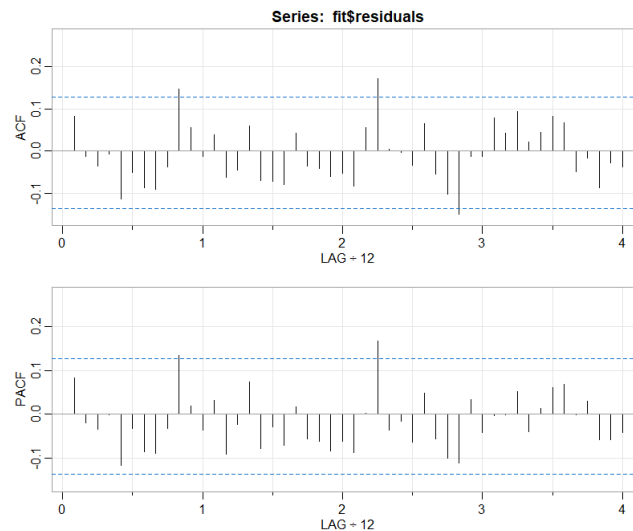


FIGURA A.3: Correlogramas dos resíduos do modelo $SARIMA(0, 1, 1)(0, 1, 1)_{12}$: FAC (gráfico superior) e FACP (gráfico inferior)

3. Os resíduos não são correlacionados. Através do teste de *Ljung e Box*, não rejeitamos H_0 , assim pode-se concluir que existe evidência estatística de que os resíduos não estão correlacionados nível de significância 5% ($Q = 25,465$ com $p - value = 0,275 > 0,05$);

A Figura A.4, apresenta o histograma dos resíduos do modelo. A média dos resíduos é 0,1645 m com intervalo de confiança a 95% (-3,6906 ; 4,0197).

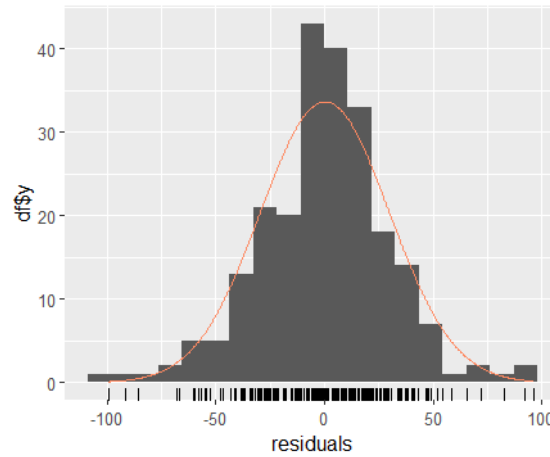


FIGURA A.4: Histograma de frequências dos resíduos do modelo

Foi aplicado o teste de *Shapiro-Wilk* para averiguar a normalidade dos resíduos. Verifica-se que existe evidência estatística para rejeitar a normalidade dos resíduos ao nível de significância de 5% ($W = 0,98351$ com $p\text{-value} = 0,0089 \leq 0,05$).

O modelo $SARIMA(0, 1, 1)(0, 1, 1)_{12}$ demonstra ser um modelo que descreve, de forma satisfatória, o comportamento geral da série, assim deve ser comparado com o modelo $SARIMA(1, 1, 1)(0, 1, 1)_{12}$.

A.3 $SARIMA(0, 1, 1)(1, 1, 1)_{12}$

A Tabela A.6, apresenta os coeficientes encontrados para o modelo,

Modelo	ϕ_1	Θ_1	Φ_1
$SARIMA(0, 1, 1)(1, 1, 1)_{12}$	-0,7518	-0,0094	-0,9998

TABELA A.6: Coeficientes do modelos $SARIMA(0, 1, 1)(1, 1, 1)_{12}$

O modelo $SARIMA(0, 1, 1)(1, 1, 1)_{12}$ é dado por:

$$(1 + 0,0094B^{12})(1 - B)(1 - B^{12})X_t = (1 + 0,7518B)(1 + 0,9998B^{12})\epsilon_t,$$

Análise dos pressupostos que garantem a qualidade estatística e do ajustamento:

1. Os Coeficientes não são significativamente diferentes de zero. A Tabela A.7 apresenta os intervalos de confiança de cada um dos coeficientes estimados.

O intervalo de confiança do coeficiente Θ_1 , oscila entre valores negativos e positivos, assim conclui-se que este coeficiente não é significativo.

	2,5%	97,5%
ϕ_1	-0,8737	-0,6299
Θ_1	-0,1471	0,1282
Φ_1	-1,2909	-0,7086

TABELA A.7: Intervalos de confiança dos coeficientes

- Os cronogramas da FAC e FACP dos resíduos, apresentam valores que, na generalidade, estão compreendidos entre os limites de confiança (Figura A.5);

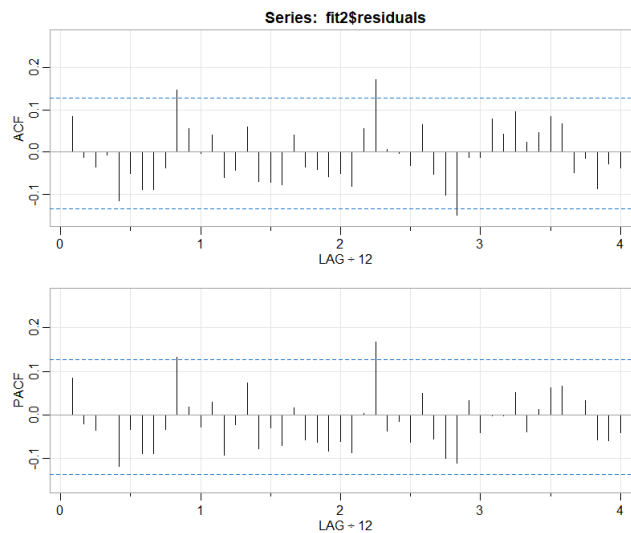


FIGURA A.5: Correlogramas dos resíduos do modelo $SARIMA = (0, 1, 1)(1, 1, 1)_{12}$: FAC (gráfico superior) e FACP (gráfico inferior)

- Os resíduos não são correlacionados. Através do teste de *Ljung e Box*, não rejeitamos H_0 , assim pode-se concluir que existe evidência estatística de que os resíduos não estão correlacionados nível de significância 5% ($Q = 25,646$ com $p - value = 0,2203 > 0,05$);

A Figura A.6, apresenta o histograma dos resíduos do modelo. A média dos resíduos é 0,1819 m m com intervalo de confiança a 95% (-3,6714 ; 4,0354).

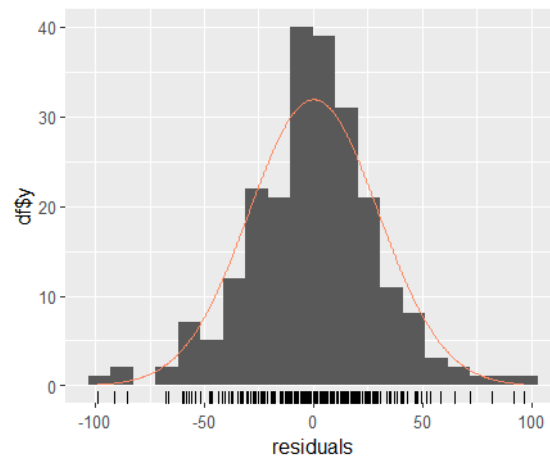


FIGURA A.6: Histograma das frequências dos resíduos do modelo

Foi aplicado o teste de *Shapiro-Wilk* para averiguar a normalidade dos resíduos. Verifica-se que existe evidência estatística para rejeitar a normalidade dos resíduos ao nível de significância de 5% ($W = 0,9837$ com $p - value = 0,0098 \leq 0,05$).

Devido à existência de coeficientes não significativos, consideramos que o modelo $SARIMA(0, 1, 1)(1, 1, 1)_{12}$, não é apropriado para descrever a série de dados maregráficos de Sines.

A.4 $SARIMA(11, 1, 0)$

A Tabela A.8, apresenta os coeficientes encontrados para o modelo,

	<i>SARIMA</i> (11, 1, 0)
θ_1	-0,5541
θ_2	-0,4918
θ_3	-0,5346
θ_4	-0,4991
θ_5	-0,5380
θ_6	-0,4074
θ_7	-0,4032
θ_8	-0,4949
θ_9	-0,5302
θ_{10}	-0,3305
θ_{11}	-0,1894
β_0	0,2953

TABELA A.8: Coeficientes do modelos *SARIMA*(11, 1, 0)

O modelo *SARIMA*(11, 1, 0) é dado por:

$$\phi_{11}(B)(1 - B)X_t = 0,2953 + \epsilon_t.$$

Análise dos pressupostos que garantem a qualidade estatística e do ajustamento:

1. Os Coeficientes não são significativamente diferentes de zero. A Tabela A.9 apresenta os intervalos de confiança de cada um dos coeficientes estimados. O intervalo de confiança da ordenada na origem, oscila entre valores negativos e positivos, assim conclui-se que não é significativo.

	2,5%	97,5%
θ_1	-0,6816	-0,4267
θ_2	-0,6318	-0,3519
θ_3	-0,6719	-0,3973
θ_4	-0,6401	-0,3580
θ_5	-0,6847	-0,3914
θ_6	-0,5602	-0,2547
θ_7	-0,5483	-0,2581
θ_8	-0,6346	-0,3551
θ_9	-0,6676	-0,392
θ_{10}	-0,4709	-0,1901
θ_{11}	-0,3172	-0,0616
β_0	-0,7710	1,3616

TABELA A.9: Intervalos de confiança dos coeficientes

- Os cronogramas da FAC e FACP dos resíduos, apresentam valores estão dentro do confiança (Figura A.7);

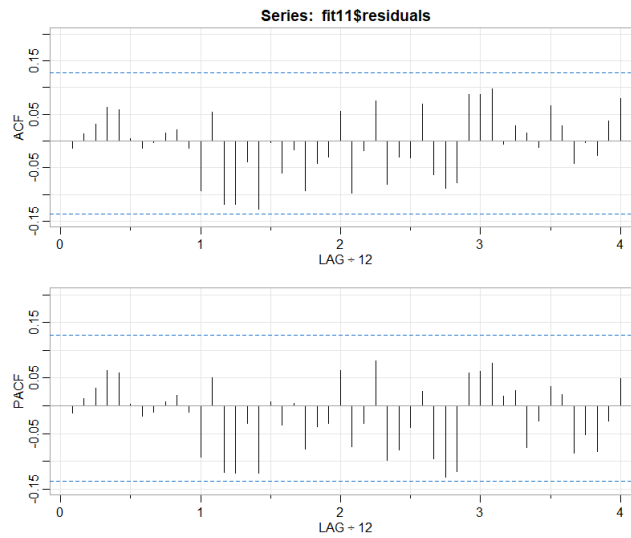


FIGURA A.7: Correlogramas dos resíduos do modelo $SARIMA = (11, 1, 1)$: FAC (gráfico superior) e FACP (gráfico inferior)

- Os resíduos estão correlacionados. Através do teste de *Ljung e Box*, rejeitamos H_0 a favor de H_1 , assim pode-se concluir que existe evidência estatística de que os resíduos estão correlacionados nível de significância 5% ($Q = 21,224$ com $p - value = 0,0472 \leq 0,05$);

A Figura A.8, apresenta o histograma dos resíduos do modelo. A média dos resíduos é 0,1587 m m com intervalo de confiança a 95% (-6,1010 ; 6,4185).

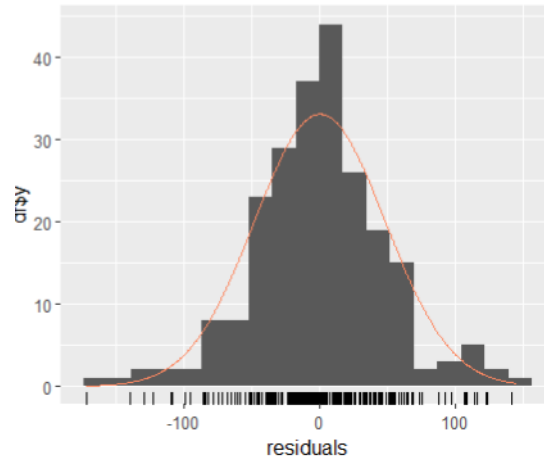


FIGURA A.8: Histograma das frequências dos resíduos do modelo

Foi aplicado o teste de *Shapiro-Wilk* para averiguar a normalidade dos resíduos. Verifica-se que existe evidência estatística para rejeitar a normalidade dos resíduos ao nível de de significância de 5% ($W = 0,9851$ com $p - value = 0,0169 \leq 0,05$).

Apesar dos resíduos do modelo apresentarem todos os seus valores da FAC e FAP dentro do intervalo de confiança, o teste de *Ljung e Box* mostra que, os resíduos estão correlacionados para um nível de significância de 5%. Como tal, o modelo *SARIMA*(11,1,0), não é apropriado para descrever a série de dados maregráficos de Sines.

Apêndice B - Avaliação dos Modelos Identificados - Registos Maregráficos de Viana do Castelo

Considerando a série estacionária, apresenta-se os modelos que aparentam descreve, satisfatoriamente, o comportamento geral da série, Tabela B.1

$SARIMA(1, 0, 0)(0, 1, 1)_{12}$
$SARIMA(0, 1, 1)(0, 1, 1)_{12}$
$SARIMA(1, 0, 0)(2, 0, 0)_{12}$

TABELA B.1: modelos identificados

Após se ter identificado os modelos, que numa primeira fase, essencialmente exploratória, garantem um bom ajustamento aos dados em análise, procura-se agora, analisar a qualidade estatística e do ajustamento de cada modelo. De salientar que, um modelo deve satisfazer, simultaneamente, os pressupostos da qualidade estatística e do ajustamento para que, seja considerado um modelo adequado.

B.1 $SARIMA(0, 1, 1)(0, 1, 1)_{12}$

A Tabela B.2, apresenta os coeficientes encontrados para o modelo,

Modelo	ϕ_1	Φ_1
$SARIMA(0, 1, 1)(0, 1, 1)_{12}$	-0,8005	-0,9083

TABELA B.2: Coeficientes do modelos $SARIMA(0, 1, 1)(0, 1, 1)_{12}$

O modelo $SARIMA(0, 1, 1)(0, 1, 1)_{12}$ é dado por:

$$(1 - B)(1 - B^{12})X_t = (1 + 0,8005B)(1 + 0,9083B^{12})\epsilon_t,$$

Análise dos pressupostos que garantem a qualidade estatística e do ajustamento:

1. Coeficientes significativamente diferentes de zero. A Tabela B.3 apresenta os intervalos de confiança de cada um dos coeficientes estimados. Visto que nenhum dos intervalos contém o valor zero, ou seja, os intervalos variaram apenas entre valores negativos, os coeficientes são considerados significativos.

	2,5%	97,5%
ϕ_1	-0,9038	-0,6973
Φ_1	-1,0023	-0,8143

TABELA B.3: Intervalos de confiança dos coeficientes

2. Os cronogramas da FAC e FACP dos resíduos, apresentam valores que, na generalidade, estão compreendidos entre os limites de confiança (Figura B.1):

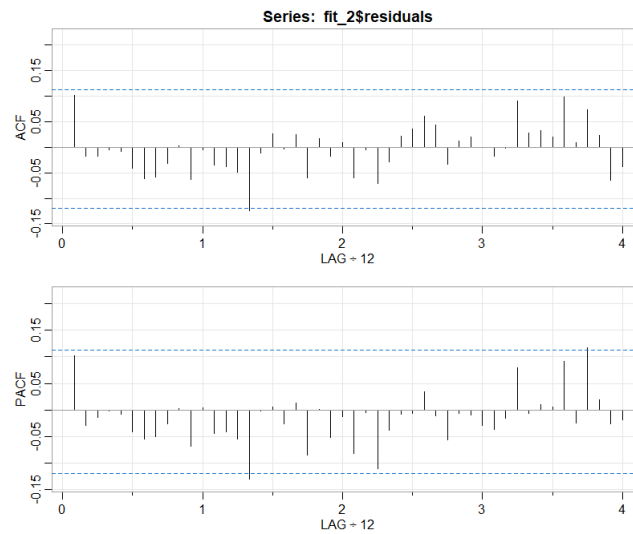


FIGURA B.1: Correlogramas dos resíduos do modelo $SARIMA = (0, 1, 1)(0, 1, 0)_{12}$: FAC (gráfico superior) e FACP (gráfico inferior)

3. Os resíduos não estão correlacionados. Através do teste de *Ljung e Box*, não rejeitamos H_0 , assim pode-se concluir que não existe evidência estatística de que os resíduos estão correlacionados ao nível de significância 5% ($Q = 16,08$ com $p - value = 0,8119 > 0,05$);

A Figura B.2, apresenta o histograma dos resíduos do modelo. A média dos resíduos é -0,932 m com intervalo de confiança a 95% (-8,3412 ; 6,4771).

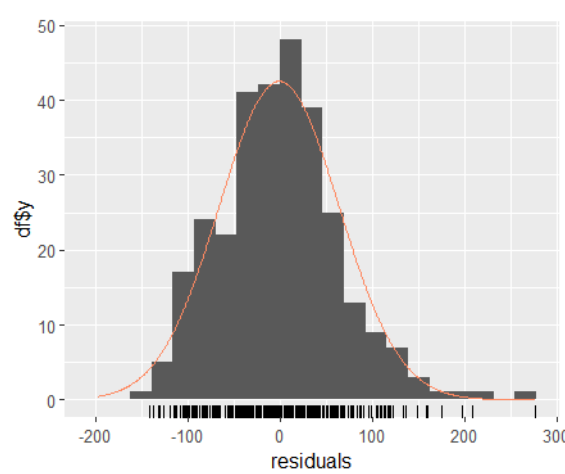


FIGURA B.2: Histograma dos resíduos do modelo

Foi aplicado o teste de *Shapiro-Wilk* para averiguar a normalidade dos resíduos. Verifica-se que existe evidência estatística para rejeitar a normalidade dos resíduos ao nível de significância de 5% ($W = 0,98142$ com $p\text{-value} = 0,0006 \leq 0,05$).

O modelo $SARIMA(0, 1, 1)(0, 1, 1)_{12}$ cumpriu com todos os pressupostos e como tal, demonstra ser um modelo que descreve, de forma satisfatória, o comportamento geral da série.

B.2 $SARIMA(1, 0, 0)(2, 0, 0)_{12}$

A Tabela B.4, apresenta os coeficientes encontrados para o modelo,

Modelo	θ_1	Θ_1	Θ_2	β_0
$SARIMA(1, 0, 0) \times (2, 0, 0)_{12}$	0,3541	0,1983	0,1747	2155,5200

TABELA B.4: Coeficientes do modelos $SARIMA(1, 0, 0)(2, 0, 0)_{12}$

O modelo $SARIMA(1, 0, 0)(2, 0, 0)_{12}$ é dado por:

$$(1 - 0,3541B)X_t = 2155.52 + (1 - 0,1983B)(1 - 0,1747B)\epsilon_t,$$

Análise dos pressupostos que garantem a qualidade estatística e do ajustamento:

1. Coeficientes significativamente diferentes de zero. A Tabela B.5 apresenta os intervalos de confiança de cada um dos coeficientes estimados. Visto que nenhum dos intervalos contém o valor zero, são considerados significativos.

	2,5%	97,5%
θ_1	0,24675	0,4615
Θ_1	0,0864	0,3102
Θ_2	0,0618	0,2876
β_0	$2,137 \times 10^3$	2173,5036

TABELA B.5: Intervalos de confiança dos coeficientes

2. Os cronogramas da FAC e FACP dos resíduos, apresentam valores que, na generalidade, estão compreendidos entre os limites de confiança (Figura B.3):

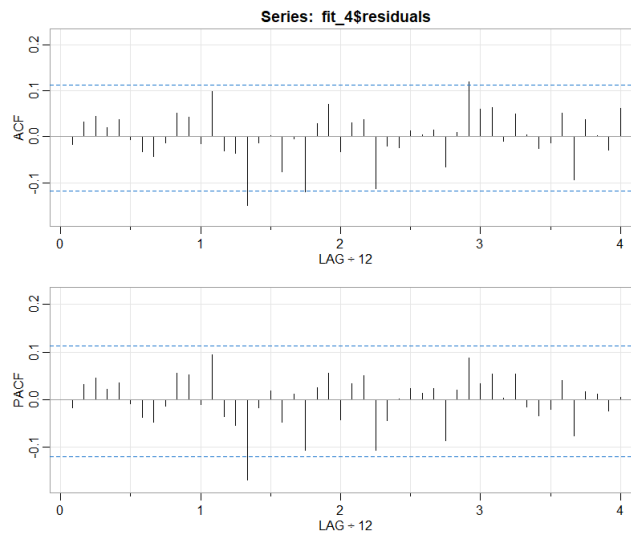


FIGURA B.3: Correlogramas dos resíduos do modelo $SARIMA = (1, 0, 0)(2, 0, 0)_{12}$: FAC (gráfico superior) e FACP (gráfico inferior)

3. Os resíduos não estão correlacionados. Através do teste de *Ljung e Box*, não rejeitamos H_0 , assim pode-se concluir que não existe evidência estatística de que os resíduos estão correlacionados ao nível de significância 5% ($Q = 23,723$ com $p - value = 0,2547 > 0,05$);

A Figura B.4, apresenta o histograma dos resíduos do modelo. A média dos resíduos é 0,2908 m com intervalo de confiança a 95% (-7,2948 ; 7,8765).

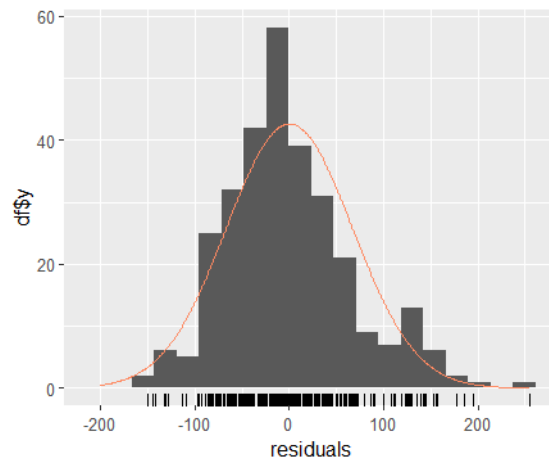


FIGURA B.4: Histograma dos resíduos do modelo

Foi aplicado o teste de *Shapiro-Wilk* para averiguar a normalidade dos resíduos. Verifica-se que existe evidência estatística para rejeitar a normalidade dos resíduos ao nível de significância de 5% ($W = 0,9728$ com $p\text{-value} = 1,857 \times 10^{-55} \leq 0,05$).

O modelo $SARIMA(1, 0, 0)(2, 0, 0)_{12}$ cumpriu com todos os pressupostos e como tal, demonstra ser um modelo que descreve, de forma satisfatória, o comportamento geral da série.

Apêndice C - Avaliação dos Modelos Identificados - Registos Maregráficos do Funchal

Considerando a série estacionária, apresenta-se os modelos que aparentam descreve, satisfatoriamente, o comportamento geral da série, Tabela C.1:

$SARIMA(1, 0, 0)(0, 1, 1)_{12}$
$SARIMA(1, 0, 0)(2, 1, 0)_{12}$
$SARIMA(0, 0, 2)(0, 1, 0)_{12}$

TABELA C.1: modelos identificados

Após se ter identificado os modelos, que numa primeira fase, essencialmente exploratória, garantem um bom ajustamento aos dados em análise, procura-se agora, analisar a qualidade estatística e do ajustamento de cada modelo. De salientar que, um modelo deve satisfazer, simultaneamente, os pressupostos da qualidade estatística e do ajustamento para que, seja considerado um modelo adequado.

C.1 $SARIMA(1, 0, 0)(2, 1, 0)_{12}$

A Tabela C.2, apresenta os coeficientes encontrados para o modelo,

Modelo	θ_1	Φ_1	Φ_2
$SARIMA(1, 0, 0)(2, 1, 0)_{12}$	0,5384	-0,6684	-0,2769

TABELA C.2: Coeficientes do modelos $SARIMA(1, 0, 0)(2, 1, 0)_{12}$

O modelo $SARIMA(1, 0, 0)(2, 1, 0)_{12}$ é dado por:

$$(1 - 0,5384B)(1 - B^{12})X_t = (1 + 0,6684B^{12})(1 + 0,2769B^{12})\epsilon_t,$$

Análise dos pressupostos que garantem a qualidade estatística e do ajustamento:

1. Coeficientes significativamente diferentes de zero. A Tabela C.3 apresenta os intervalos de confiança de cada um dos coeficientes estimados.

	2,5%	97,5%
θ_1	0,3951	0,6817
Φ_1	-0,8389	0,4979
Φ_2	-0,4521	-0,1018

TABELA C.3: Intervalos de confiança dos coeficientes

2. Os cronogramas da FAC e FACP dos resíduos, apresentam valores que, na generalidade, estão compreendidos entre os limites de confiança (Figura C.1):

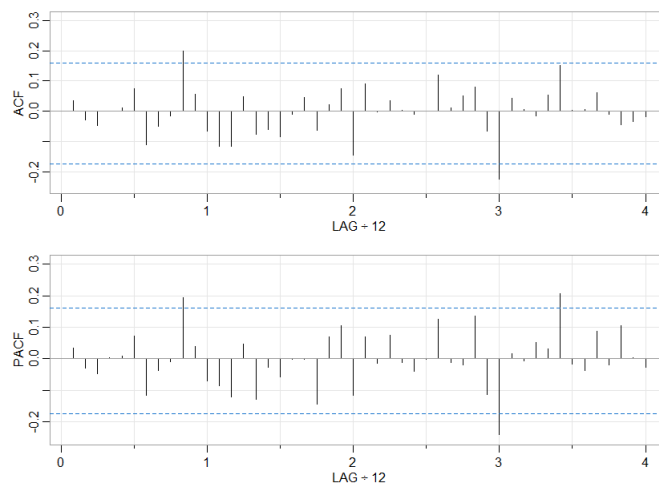


FIGURA C.1: Correlogramas dos resíduos do modelo $SARIMA = (1, 0, 0)(2, 1, 0)_{12}$: FAC (gráfico superior) e FACP (gráfico inferior)

3. Os resíduos não estão correlacionados. Através do teste de *Ljung e Box*, não rejeitamos H_0 , assim pode-se concluir que não existe evidência estatística de que os resíduos estão correlacionados ao nível de significância 5% ($Q = 24,353$ com $p - value = 0,2273 > 0,05$);

A Figura C.2, apresenta o histograma dos resíduos do modelo. A média dos resíduos é 0,0029 m com intervalo de confiança a 95% (-0,0146 ; 0,0204).

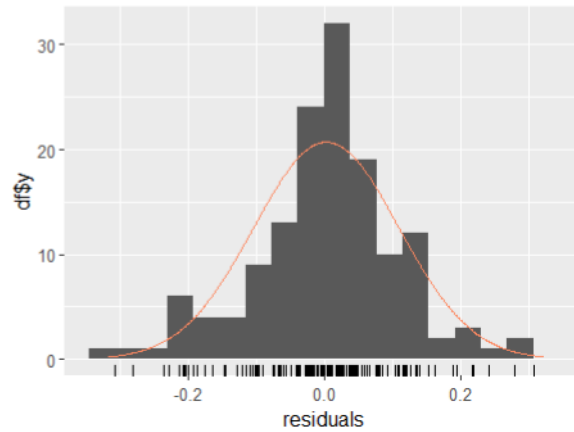


FIGURA C.2: Histograma dos resíduos do modelo

Foi aplicado o teste de *Shapiro-Wilk* para averiguar a normalidade dos resíduos. Verifica-se que existe evidência estatística para rejeitar a normalidade dos resíduos ao nível de significância de 5% ($W = 0,9779$ com $p\text{-value} = 0,02013 \leq 0,05$).

O modelo $SARIMA(0, 1, 1)(0, 1, 1)_{12}$ cumpre com todos os requisitos, pelo que foi aceite na fase de avaliação.

C.2 $SARIMA(0, 0, 2)(0, 1, 1)_{12}$

A Tabela C.4, apresenta os coeficientes encontrados para o modelo,

Modelo	ϕ_1	ϕ_2	Φ_1
$SARIMA(0, 0, 2)(0, 1, 1)_{12}$	0,5767	0,2599	-0,9999

TABELA C.4: Coeficientes do modelos $SARIMA(0, 0, 2) \times (0, 1, 1)_{12}$

O modelo $SARIMA(0, 0, 2)(0, 1, 1)_{12}$ é dado por:

$$(1 - B^{12})X_t = (1 - 0,5767B)(1 - 0,2599B)(1 + 0,9999B^{12})\epsilon_t,$$

Análise dos pressupostos que garantem a qualidade estatística e do ajustamento:

1. Coeficientes significativamente diferentes de zero. A Tabela C.5 apresenta os intervalos de confiança de cada um dos coeficientes estimados.

	2,5%	97,5%
ϕ_1	4,1444	0,7390
ϕ_2	1,0611	0,4137
Φ_1	-1,2389	-0,7609

TABELA C.5: Intervalos de confiança dos coeficientes

2. Os cronogramas da FAC e FACP dos resíduos, apresentam valores que, na generalidade, estão compreendidos entre os limites de confiança (Figura C.3):

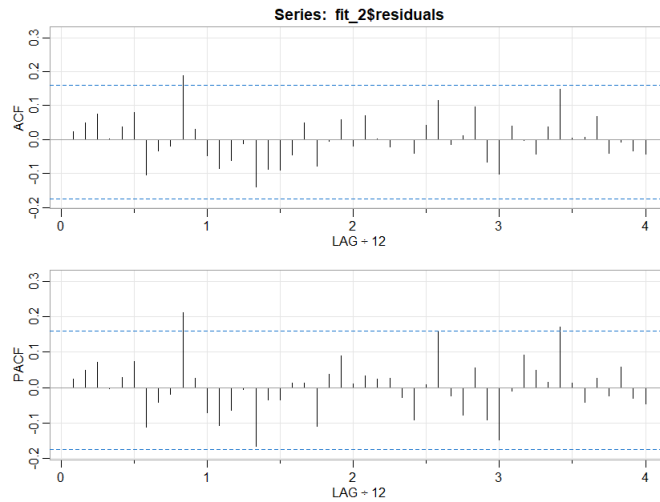


FIGURA C.3: Correlogramas dos resíduos do modelo $SARIMA = (0, 0, 2)(1, 1, 0)_{12}$: FAC (gráfico superior) e FACP (gráfico inferior)

3. Os resíduos não estão correlacionados. Através do teste de *Ljung e Box*, não rejeitamos H_0 , assim pode-se concluir que não existe evidência estatística de que os resíduos estão correlacionados ao nível de significância 5% ($Q = 20,303$ com $p - value = 0,4392 > 0,05$);

A Figura C.4, apresenta o histograma dos resíduos do modelo. A média dos resíduos é 0,0064 m com intervalo de confiança a 95% (-0,0084 ; 0,0212).

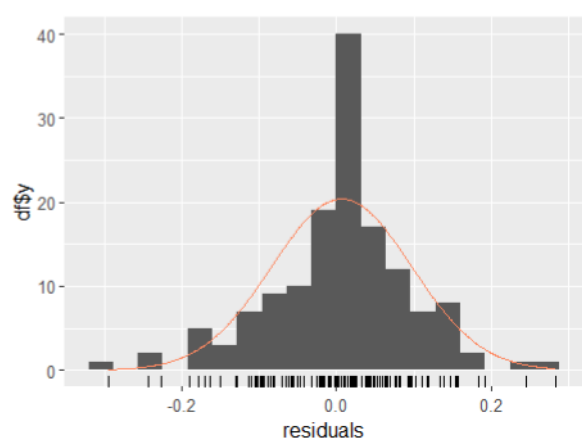


FIGURA C.4: Histograma dos resíduos do modelo

Foi aplicado o teste de *Shapiro-Wilk* para averiguar a normalidade dos resíduos. Verifica-se que existe evidência estatística para rejeitar a normalidade dos resíduos ao nível de significância de 5% ($W = 0,9746$ com $p\text{-value} = 0,0089 \leq 0,05$).

O modelo $SARIMA(0, 0, 2)(1, 1, 0)_{12}$ cumpre com todos os requisitos, pelo que foi aceite na fase de avaliação.

Apêndice D - Código Utilizado na Análise e Processamento das Séries Temporais

```
1
2 library(forecast)
3 library(astsa)
4 library(DescTools)
5 library(xts)
6 library(lubridate)
7 library(TSA)
8 library(ggplot2)
9 library(Metrics)
10 library(openxlsx)
11
12
13 # ler dados do ficheiro nmm_R.txt
14 sines=read.table("nmm.txt", header = TRUE)
15 str(sines) # estrutura de sines
16
17 #####
18 #
19 #   Análise Exploratória do NMM mensal em Sines
20 #
21 #####
22 # transformar sines$nmm num objeto do tipo série temporal (ts),
23 #   indicando a
24 #   frequência, neste caso mensal, e a data de início
25 nmmtotal=ts(sines$nmm,frequency=12,start=c(1997,1))
26 plot(nmmtotal, ylab= "log_N vel Médio do Mar (mm)", xlab="meses")
27 str(nmmtotal) # estrutura de nmmtotal
28
29 # alguma caracterização
30 mean(nmmtotal)
31 var(nmmtotal)
32 summary(nmmtotal)
33
34 # Outliers dos nmm
35 boxplot(nmmtotal, col="lightblue") #identificados 3 outliers
36
37 # Histograma dos nmm
38 hist(nmmtotal, breaks= 20, col="lightblue",
39      main="Histograma Sines (1997-2020)",
40      ylab= "Frequência", xlab="Nível médio do mar (mm)")
```

```

41
42 # Teste de normalidade
43 st = shapiro.test(nmmtotal)
44 st #dados seguem distr. normal
45
46 # FAC amostral
47 acf1(nmmtotal)
48 # FACP amostral
49 acf1(nmmtotal, pacf=TRUE)
50
51 # ou, usando o pacote astsa, as duas funções numa única figura
52 acf2(nmmtotal)
53
54 # Teste em que a hipótese nula indica a independência da série dos
    nmm
55 Box.test(nmmtotal, lag = 1, type = "Ljung-Box", fitdf = 0)
56 #p-value < 0,05 indica rejeição da hipótese nula
57
58
59 #####
60 #
61 #     ANÁLISE DE TENDÊNCIA E SAZONALIDADE
62 #
63 #####
64
65
66 #criar uma dataframe
67 df = data.frame(cbind(sines$nmmtotal, 1:nrow(sines)))
68 #alterar nome das colunas
69 names(df) = c("Y", "X")
70
71
72 #####
73     ### Tendência Polinomial###
74 #####
75 # ajustamento do polinômio de grau 1
76 fit = lm(Y ~ 1 + X, data = df,
77         singular.ok = FALSE)
78
79 #representação gráfica do dataframe
80 plot(df$Y, type = "l", main = "Regressão linear", xlab = "meses",
      ylab = "n vezes médios (nmmtotal)")
81 #representação gráfica do polinômio de grau 1
82 lines(fit$fitted, col = "red", lwd= 2)
83 #caracterização (coeficientes, p-value, R^2, etc.)

```



```

84 summary(fit)
85 #intervalos de confiança dos coeficientes estimados.
86 confint(fit)
87 # root mean squared error
88 rmse(sines$nm, fit$fitted.values)
89
90 # ajustamento do polinómio de ordem 2( ACELERAÇÃO DA TENDÊNCIA)
91 fit2 = lm(Y ~ 1 + X + I(X^2) , data = df,
92          singular.ok = FALSE)
93
94 plot(df$Y, type = "l",main = "polinómio de grau 2",xlab = "meses",
95       ylab = "n veis médios (nm)")
96 lines(fit2$fitted, col = "red", lwd= 2)
97
98 #caracterização
99 summary(fit2)
100 #intervalos de confiança dos coeficientes estimados.
101 confint(fit2)
102
103 # verificar se os res duos seguem a distribuição normal
104 shapiro.test(fit$residuals)
105 shapiro.test(fit2$residuals)
106
107 #verificar se as médias dos res duos dos dois modelos são
108 estaticamente diferentes
109 # usar se ambos os res duos seguem distribuição normal
110 anova(fit,fit2)
111 #verificar se as medianas dos res duos dos dois modelos são
112 estaticamente diferentes
113 # usar se ambos os res duos não seguem distribuição normal
114 amostras=matrix(c(fit$residuals,fit2$residuals),ncol=2)
115 wilcox.test(amostras)
116
117 #####
118 ##### Tendência C clica#####
119 #####
120
121 ##### modelo com DUAS sinusóides#####
122 fit.1 = lm(Y ~ 1 + X + I(cos(2*pi/12*X)) + I(sin(2*pi/12*X))+
123           I(cos(4*pi/12*X)) + I(sin(4*pi/12*X)), data = df,
124           singular.ok = FALSE)
125
126 summary(fit.1)

```

```

125 plot(df$Y, type = "l", main="modelo de Tendência linear + Tendência
      C clica (com duas sinusoides)" , xlab = "meses", ylab = "
      n veis médios (nmm)")
126 lines(fit.1$fitted, col = "red", lty = 1, lwd = 2)
127 rmse(sines$nmm, fit.1$fitted.values)
128 acf2(fit.1$residuals)
129 checkresiduals(fit.1)
130 shapiro.test(fit.1$residuals)
131
132 ##### modelo com DUAS sinusóides e coeficnete beta_{2}#####
133 fit.3 = lm(Y ~ 1 + X +I(X^2)+ I(cos(2*pi/12*X)) + I(sin(2*pi/12*X))
      +
134           I(cos(4*pi/12*X)) + I(sin(4*pi/12*X))#+
135           #I(cos(6*pi/12*X)) + I(sin(6*pi/12*X))#+
136           #I(cos(8*pi/12*X)) + I(sin(8*pi/12*X))
137           , data = df,
138           singular.ok = FALSE)
139
140 summary(fit.3)
141 plot(df$Y, type = "l", main="modelo de Tendência linear + Tendência
      C clica (com duas sinusoides)" , xlab = "meses", ylab = "
      n veis médios (nmm)")
142 lines(fit.3$fitted, col = "red", lty = 1, lwd = 2)
143 rmse(sines$nmm, fit.3$fitted.values)
144 acf2(fit.3$residuals)
145 checkresiduals(fit.3)
146 shapiro.test(fit.3$residuals)
147
148 # verificar se os modelos são estatisticamente semelhantes
149 nmm_sinusoi = matrix(c(fit$residuals,fit.1$residuals),ncol=2)
150 wilcox.test(nmm_sinusoi)
151 nmm_sinusoi2 = matrix(c(fit2$residuals,fit.1$residuals),ncol=2)
152 wilcox.test(nmm_sinusoi2)
153 nmm_sinusoi3 = matrix(c(fit.1$residuals,fit.3$residuals),ncol=2)
154 wilcox.test(nmm_sinusoi3)
155
156 #####
157
158 #ARIMA/////SARIMA#
159
160 #####
161
162 # Para efeitos de previsão, consideramos apenas os primeiros 230
      meses (19 anos),

```

```

163 # os restantes 58 meses (5 anos) serão usados para aferir,
    graficamente, a qualidade da previsão
164 nmm=ts(sines$nm[1:230],frequency=12,start=c(1997,1))
165 plot(nmm, ylab= "log_N vel Médio do Mar (mm)", xlab="meses")
166
167
168 #####
169
170 #Transformações ( VARIÂNCIA; MÉDIA; SAZONALIDADE)#
171
172 #####
173
174 # estabilizar a variância aplicando o logaritmo
175 log_nmm=log(nmm) #a variância não melhorou!
176 plot(log_nmm, ylab= "log_N vel Médio ", xlab="meses")
177
178 # estabilizar a variância aplicando a transf. de Box-Cox
179 (lambda = BoxCox.lambda(nmm,lower=-1,method = "loglik")) #
    determina o lambda ótimo'
180 boxcox_nmm=BoxCox(nmm, BoxCox.lambda(nmm,lower=-1,method = "loglik"
    )) # aplica a transformação
181 plot(boxcox_nmm)
182 acf2(boxcox_nmm)
183
184 #retirar a tendência:Diferença de 1 ordem
185 diff1_boxcox_nmm=diff(boxcox_nmm, lag=1)
186 plot(diff1_boxcox_nmm, xlab="meses")
187 acf2(diff1_boxcox_nmm)
188
189
190 # Agora a sazonalidade: Diferença sazonal, s=12
191 diff12_diff1_boxcox_nmm=diff(diff1_boxcox_nmm,lag=12)
192 plot(diff12_diff1_boxcox_nmm, main = "Série estacionária " , xlab =
    "meses")
193 acf2(diff12_diff1_boxcox_nmm)
194
195 #####
196 #
197 # Ajuste (e diagnóstico) de modelos SARIMA para o n vel médio
    mensal
198 # de Sines, PARA OS PRIMEIROS 230 VALORES.
199 #
200 #####
201
202

```

```

203 #aplicação do modelo #sarima(1,1,1)(0,1,1)(12) aos primeiros 230
    meses (19 anos)
204 fit = Arima(nmm,order=c(1,1,1),
205             seasonal=list(order=c(0,1,1),period=12),
206             #include.mean = TRUE,
207             #transform.pars = TRUE,
208             lambda=0.95,
209             include.constant=TRUE)
210 #verificar FAC e Histograma dos res duos
211 checkresiduals(fit)
212 #coeficientes estimados
213 coefficients(fit)
214 # FAC e FACP dos res duos
215 acf2(fit$residuals)
216 # representação gráfica das previs es
217 autoplot(forecast(fit))
218 # intervalos de confiança de cada um dos coeficientes
219 confint(fit)
220 # teste de normalidade ao res duos
221 shapiro.test(fit$residuals)
222 #média dos res duos
223 mean(fit$residuals)
224 #assimetria das frequências dos res duos
225 skewness(fit$residuals)
226 summary(fit)
227
228 # calcular intervalo de confiança da média dos res duos ao n vel
    de confiança de 95%
229
230     # média dos res duos
231     média = mean(fit$residuals)
232     # tamanho da amostra
233     n = length(fit$residuals)
234     # desvio padrão
235     desvio.padrão = sd(fit$residuals)
236     # standard error
237     standard_error = desvio.padrão / sqrt(n)
238     alpha = 0.05
239     degrees_of_freedom = n - 1
240     t_score = qt(p=alpha/2, df=degrees_of_freedom,lower.
tail=F)
241     margin_error = t_score * standard_error
242
243     # calcular o limite inferior e o limite superior
244     lower_bound <- média - margin_error

```

```

245         upper_bound <- média + margin_error
246         # intervalo de confiança
247         print(c(lower_bound, upper_bound))
248
249 # COMPARAR AS PREVIS ES [231,288] COM OS VALORES REAIS
250 # previs es para um horizonte de 58 meses
251 autoplot(forecast(fit, h = 58 )) +
252     #modelo encontrado
253     autolayer(fit$fitted, series="ARIMA(1,0,0)(0,1,1)(12)") +
254     # série completa
255     autolayer(nmmtotal)
256
257
258 #####
259
260 # Ajustar o modelo      série completa (288 valores de NMM) e fazer
    previs es  #
261 #####
262 #sarima(1,1,1)(0,1,1)(12)
263 fit13 = Arima(nmmtotal, order=c(1,1,1),
264               seasonal=list(order=c(0,1,1), period=12),
265               lambda=0.95,
266               include.constant=TRUE)
267 checkresiduals(fit13)
268 coefficients(fit13)
269 acf2(fit13$residuals)
270 autoplot(forecast(fit13, h=120), xlab = "Ano", ylab = "n vel médio
    mensal")
271 plot(forecast(fit13, h=120), main = "Previsão com modelo SARIMA
    (1,1,1(0,1,1)[12]", xlab = "Ano", ylab = "n vel médio mensal")
272 autoplot(fit13)
273 confint(fit13)
274 summary(fit13)
275
276
277
278 # COMPARAR AS PREVIS ES [231,288] COM OS VALORES REAIS
279 autoplot(forecast(fit12, h=58))+
280     autolayer(nmmtotal)+
281     autolayer(fit12$fitted, series="ARIMA(1,1,1)(0,1,1)(12)")
282
283 #valores da previsão
284 resultados = forecast(fit13, h=120)
285
286 #CRIAR FICHEIRO EXCELL COM OS VALORES PREVISTOS

```

```
287 wb <- createWorkbook()
288
289 addWorksheet(wb, "Forecasts")
290 writeData(wb, "Forecasts", resultados, rowNames = TRUE)
291 saveWorkbook(wb, "Forecasts.xlsx", overwrite = TRUE)
```

Apêndice E - Código Utilizado na Avaliação da Semelhança Estatística da Sazonalidade Entre as Três Estações Maregráficas

```
1
2 library(Metrics)
3
4 #####
5 ##### Semelhanças entre a sazonalidade #####
6 #####
7
8
9 # mediana do NMm's Sines ( criar dataset)
10 sines=c(2146, 2143, 2167, 2199, 2195, 2204, 2177, 2211, 2234, 2283,
11         2236, 2200)
12 #Teste de normalidade
13 shapiro.test(sines)
14
15 # mediana do NMm's Viana do Castelo ( criar dataset)
16 viana=c(2153, 2148, 2105, 2132, 2147, 2125, 2116, 2132, 2187, 2262,
17         2233, 2185)
18 shapiro.test(viana)
19
20 # mediana do NMm's Funchal ( criar dataset)
21 funchal=c(1528, 1531, 1541, 1580, 1584, 1605, 1641, 1677, 1687,
22         1700, 1662, 1572)
23 #Teste de normalidade
24 shapiro.test(funchal)
25
26 # as três medianas apresentam distribuição Normal
27 #Aplicar teste de ANOVA
28
29
30 #criar data frame com as três amostras em estudo
31 medianas_sazonalidade = data.frame(cbind(viana,sines,funchal))
32 str(medianas_sazonalidade)
33
34 # Teste ANOVA( Registos de Sines e Viana)
35 one.way = aov(sines ~ viana, data = medianas_sazonalidade)
36 summary(one.way)
37
38 # Teste ANOVA( Registos de Sines e Funchal)
39 one.way = aov(sines ~ funchal, data = medianas_sazonalidade)
40 summary(one.way)
```

```
37  
38 # Teste ANOVA( Registos de Viana e Funchal)  
39 one.way = aov(viana ~ funchal, data = medianas_sazonalidade)  
40 summary(one.way)
```


Anexo I - Participação em: 7.^a Jornadas de Engenharia Hidrográfica/ 2^aJornadas Luso-Espanholas de Engenharia Hidrográfica - Resumo Alargado

Caracterização da tendência de subida do nível médio em Portugal

M. Picoto (1), D. Carinhas (2), C. Lemos (3)

- (1) Escola Naval, Instituto Universitário Militar; migpicoto@gmail.com.
 (2) Instituto Hidrográfico, IIFA/Universidade de Évora.
 (3) CINAV, Escola Naval, Instituto Universitário Militar, Base Naval de Lisboa – Alfeite 2810-001, Almada.

Resumo: As alterações climáticas e as suas consequências são um dos temas dominantes na atualidade. Uma das consequências mais evidentes do aquecimento global é a tendência de subida do nível médio do mar, associada à fusão dos glaciares nas regiões polares e à expansão térmica da água dos oceanos. Esta tendência tem vindo a ser estudada desde há várias décadas a nível global, utilizando séries maregráficas longas e, mais recentemente, outros métodos de observação (por exemplo altimetria por satélite). Neste trabalho, pretende-se estender estudos já efetuados no Instituto Hidrográfico sobre a subida do nível médio do mar através da utilização de regressão linear e de técnicas de análise de séries temporais a registos maregráficos de portos nacionais. Por exemplo em Sines, o incremento foi de 106 milímetros nos 24 anos o que representa uma subida de 4.67 mm/ano, não se evidenciando nenhuma aceleração.

Palavras-chave: modelos de séries temporais, nível médio do mar, registos maregráficos, tendência do nível do mar.

1. INTRODUÇÃO

Atualmente, existe uma grande preocupação com a subida do nível médio do mar (NMM), porque este fenómeno tem, à escala global, um grande impacto na humanidade, nomeadamente nas atividades económicas e sociais.

Muitos são os fatores que influenciam o NMM e, em geral, fazem-se sentir no espaço e no tempo, com maior ou menor intensidade e interligação. A compreensão de muitos destes processos é, em muitos casos complexa, e apenas alguns fenómenos, como a maré astronómica, podem ser previstos de forma rigorosa. Em contrapartida, as variações inter-anuais são mais difíceis compreender e de prever. Em estudos de análise das variações do NMM, cujo objetivo é provar a relação entre estas e o atual aquecimento global, associado a atividades antrópicas, são muitas vezes omitidos ou subvalorizados os fenómenos a eles associados, os quais embora conhecidos não são quantificáveis de forma determinística (Dias e Taborda, 1988).

As variações na órbita terrestre têm sido apontadas como um dos principais fenómenos de variação do clima e consequentemente do NMM ao longo do período Quaternário, entre 2.7 milhões de anos e 11.7 mil anos atrás. Após o último período glaciário e essencialmente durante o Período Holoceno, que se iniciou há 11.7 mil anos e perdura até ao presente, existem essencialmente dois grandes fenómenos, a nível global, que desencadeiam variações de longas escalas de tempo no NMM devido às interações Criosfera-Oceano-Atmosfera, sendo eles, os fenómenos eustáticos e isostáticos (Da Silva et al.,

2014). Fenómenos eustáticos provocam variações no volume e na massa dos oceanos, enquanto os fenómenos isostáticos estão relacionados com a elevação ou subsidência da crosta terrestre, provocando uma variação relativa do NMM. As flutuações do NMM causadas por fenómenos glácio-eustáticos estão fortemente associadas ao aquecimento e arrefecimento global que, consequentemente, levam à fusão ou à formação das calotas polares e glaciares. Estes processos desenvolvem-se em escalas temporais de vários milénios. Durante os períodos glaciários, o nível do mar desce porque a humidade que é evaporada dos oceanos não é reposta, ficando aprisionada sob a forma de gelo nas calotas polares. Em períodos interglaciários, com temperaturas mais amenas, dá-se o recuo dos glaciares e, consequentemente, o aumento do NMM devido à reposição da água que estava retida nas calotas.

No último século verificou-se um aumento da tendência de subida do NMM, com particular incidência nas últimas décadas (Figura 1).

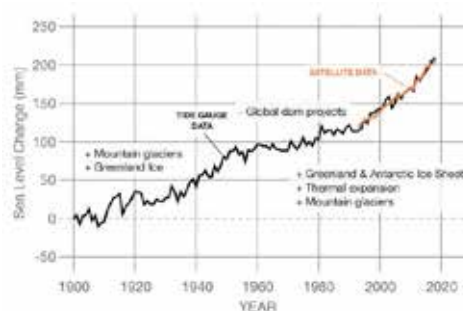


Fig. 1. Nível do mar entre 1900 e 2020: dados obtidos de observações de marés, em estações maregráficas (preto) e através de satélites (vermelho) (fonte: <https://climate.nasa.gov/>).

No entanto, essa subida do NMM é difícil de quantificar, porquanto:

(<https://tidesandcurrents.noaa.gov/sltrends/>):

1. Desconhece-se o comportamento geral desta subida, ou seja, desconhece-se o modelo (ou função) aplicável às recentes variações do NMM;
2. Os valores da tendência de subida apresentam diferenças notáveis de local para local (Figura 2);
3. O número de estações maregráficas, com séries de observações, suficientemente, significativas é bastante pequeno e a sua distribuição, pelo globo, é assimétrica.

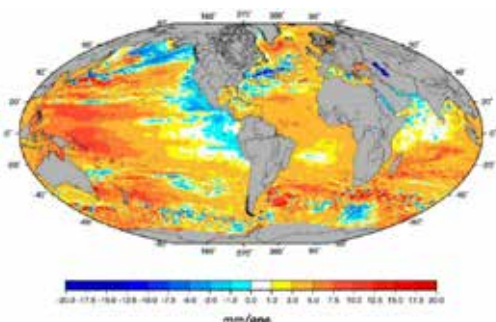


Fig. 2. Distribuição geográfica das tendências do nível do mar entre 1992-2008, obtidas através de altimetria de satélites (adaptado de Cazenave e Nerem, (2004)).

Em Portugal existem várias séries de registos maregráficos provenientes de marégrafos dispersos ao longo do seu litoral. Tanto os registos maregráficos de Cascais como de Lagos evidenciam uma tendência de aumento do NMM (Antunes e Taborda, 2009). No entanto, as análises das séries de dados maregráficos sugerem que o NMM não subiu com uma tendência consistente ao longo do tempo (Antunes e Taborda, 2009). Desde final do século XIX até por volta de 1915-1920, o NMM teve inclusive uma tendência negativa, -5 ± 0.4 mm/ano em Cascais (Ferreira et al., 2008). Só posteriormente a este período é que se nota uma tendência de subida do NMM. Mendes *et al.*, (2020), tendo por bases a utilização de registos GPS, concluíram que os movimentos verticais na Península Ibérica são, na sua maioria, caracterizados por tendências de subsidência lentas/moderadas, com um valor médio de 0.6mm/ano, sendo que em Cascais se verificou uma tendência de 0.4 mm/ano. Isto significa que as séries maregráficas da estação de Cascais são fidedignas para serem usadas em estudos globais sobre o NMM e que podem ser usadas para caracterizar a tendência de subida do NMM.

2. MATERIAIS E MÉTODOS

A estatística desempenha um papel fundamental no processamento de dados porque permite inferir parâmetros caracterizadores de fenómenos com uma componente aleatória a partir de amostras. O seu principal objetivo é a recolha, a análise dos dados,

inferência (determinação de parâmetros de distribuições ou relações funcionais) e interpretação dos resultados, de forma qualitativa e quantitativa. Quando se efetuam medições, é sempre expectável uma certa variabilidade da característica ou atributo que se está a medir. A variabilidade (também chamada de dispersão) refere-se à distribuição de um conjunto de dados. As quatro principais formas de descrever a variabilidade em um conjunto de dados são: Amplitude, Amplitude interquartil, Variância e Desvio padrão. Graficamente, a variabilidade pode ser avaliada através de gráficos *box-plot* e/ou histogramas.

Por definição, uma série temporal é uma sucessão de observações que evoluem no tempo, igualmente espaçadas (Murteira *et al.*, 2000).

Segundo o modelo clássico todas as séries temporais são compostas de quatro padrões:

- tendência (T), que é o comportamento de longo prazo da série, que pode ser causada pelo crescimento ou decrescimento que afete a variável de interesse ao longo do tempo;
- variações cíclicas ou ciclos (C), flutuações nos valores da variável com duração superior a um ano, e que se repetem com certa periodicidade, que podem ser resultado de variações de fenómenos climáticos como o El Niño (que se repete com periodicidade superior a um ano);
- variações sazonais ou sazonalidade (S), flutuações nos valores da variável com duração inferior a um ano, e que se repetem todos os anos, geralmente em função das estações do ano, se os dados forem registados anualmente não haverá influência da sazonalidade na série;
- variações irregulares (I), que são as flutuações aleatórias, não correlacionadas entre si (ao contrário das variações descritas anteriormente), resultando de fatores fortuitos e inesperados, como fenómenos extremos e catástrofes naturais.

É importante salientar que nem sempre uma série temporal, mesmo que o modelo clássico seja considerado apropriado para analisá-la, irá apresentar todos os componentes citados.

A tendência descreve o comportamento da variável retratada na série temporal no longo prazo. Um dos objetivos principais na sua identificação é avaliar o seu comportamento para utilizá-lo em previsões. A obtenção da tendência pode ser feita de três formas: através de um modelo de regressão (como o modelo linear - reta), através de médias móveis, ou através de ajuste exponencial (que não deixa de ser uma média móvel) (Morettin e Toloi, 2018). O procedimento para a obtenção da tendência através da reta de regressão é semelhante ao usado na regressão linear simples, mas agora a variável independente será sempre o tempo. Para uma série registada anualmente, por exemplo, de 2005 a 2014, a variável

independente assumiria os valores dos anos. Para uma série registada mensalmente, por exemplo, com 60 meses, a variável independente poderia assumir os valores de 1 a 60. A tendência pode ser obtida através da equação $T = bt + a$;

onde:

t é o valor do tempo, no caso linear b é o declive da reta (se positivo indica tendência crescente, se negativo a tendência é decrescente) e a é o coeficiente linear da reta. As equações dos coeficientes são expressas por:

$$b = \frac{n \times \sum_{i=1}^n (t_i \times Y_i) - \sum_{i=1}^n t_i \times \sum_{i=1}^n Y_i}{n \times \sum_{i=1}^n (t_i^2) - \left(\sum_{i=1}^n t_i \right)^2} \quad a = \frac{\sum_{i=1}^n Y_i - b \times \sum_{i=1}^n t_i}{n}$$

onde Y_i é um valor qualquer da variável registada na série temporal, t_i é o período associado a Y_i , e n é o número de dados da série. Para encontrar os coeficientes basta calcular os somatórios (tal como na análise de regressão linear simples).

3. RESULTADOS

Com vista à aplicação prática das metodologias apresentadas, a amostra em estudo compreende 288 NMM mensais, observados na estação maregráfica do porto de Sines (Figura 3), referentes ao período de 1997 a 2020. Estes dados, foram disponibilizados pelo Instituto Hidrográfico (IH).



Fig. 3. Identificação do porto estudado (fonte: Divisão de Oceanografia do Instituto Hidrográfico).

O histograma (Figura 4) apresenta uma ligeira assimetria negativa (coeficiente de assimetria igual a -0.07, aproximadamente) demonstrando que a distribuição dos níveis médios mensais se concentra no lado direito da média (2198 mm com Intervalo de Confiança (95%) igual a (2191; 2206)) com uma ligeira cauda para o seu lado esquerdo. Através da aplicação do teste de normalidade de *shapiro-wilk*, obteve-se o valor $p = 0.9516$ ($\alpha = 0.05$), pelo que não se rejeita a hipótese da normalidade dos dados.

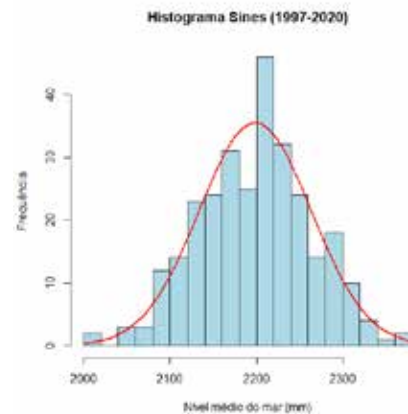


Fig. 4. Histograma de frequências dos níveis médios mensais em Sines no período 1997 - 2020.

A Tabela 1 apresenta as principais estatísticas descritivas dos NMM mensais (1997-2020):

Tabela 1. Estatísticas descritivas dos NMM mensais (mm) de Sines

NMM mensais (mm)	
Média	2198
Mediana	2205
Desvio-padrão	64,88
Assimetria	-0,07
Mínimo	2007
Máximo	2379
Dimensão da amostra	288

A Figura 5 apresenta a evolução dos níveis médios mensais e anuais em Sines até finais de 2020. Foi ajustada uma função linear aos níveis médios anuais, através de um polinómio de grau 1, com vista a representar a tendência dos níveis médios mensais.

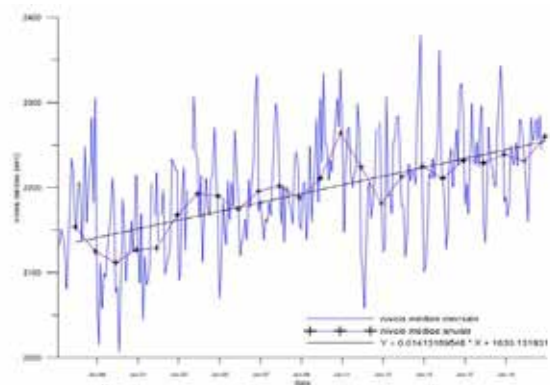


Fig. 5. Evolução dos níveis médios mensais e anuais, em Sines desde 1997 até finais de 2020.

Face ao declive positivo da reta de regressão obtida, pode-se concluir que o nível médio está a aumentar. O incremento foi de 106 milímetros nos 24 anos o que representa uma subida de 4.67 mm/ano, aproximadamente, não se evidenciando nenhuma aceleração. Durante o ano de 2020 o nível médio mensal máximo foi de 2441 mm registado em novembro e o nível médio mensal mínimo foi de 2160

mm registado em fevereiro. Dos 288 valores de nível médio mensal registados desde 1997, 95% situam-se no intervalo [2080; 2320] milímetros, sendo que 5% são superiores a 2300 mm e apenas 1% são inferiores a 2010 mm.

A Figura 6 apresenta as distribuições mensais dos níveis médios do mar registados no marégrafo de Sines 1997 e 2020:

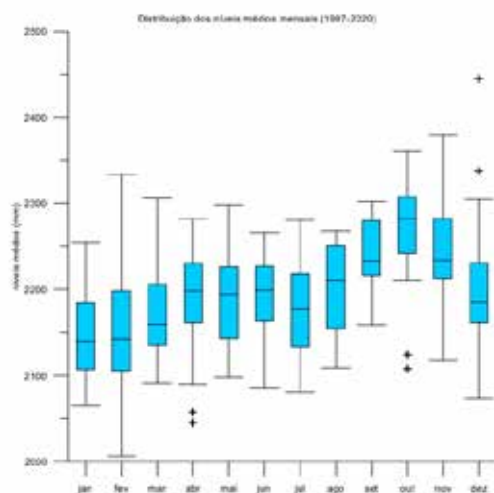


Fig. 6. Comportamento dos níveis médios ao longo do ano, em Sines no período 1997 a 2020.

É possível verificar que a variação do nível médio tem uma componente sazonal. Verificam-se acentuadas variações anuais e semi-anuais do NMM devido a variações sazonais da pressão atmosférica, da densidade da água e da circulação oceânica (Carinhas *et al.*, 2021).

Esta análise é consistente com estudos já efetuados no IH. A Figura 7 ilustra os registos dos níveis médios mensais característicos para os diferentes portos nacionais:

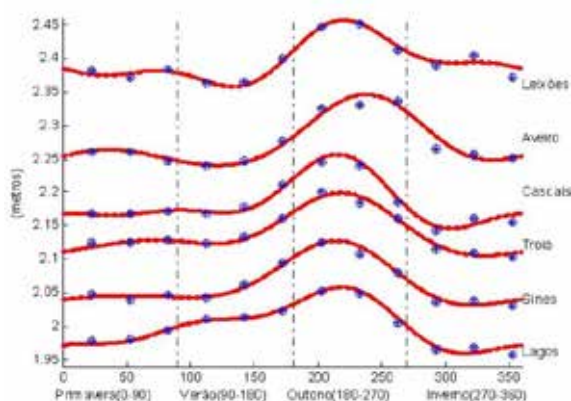


Fig. 7. Níveis médios mensais característicos (pontos azuis) com ajustamento dos níveis médios diários (vermelho) (fonte: Carinhas, *et al.*, 2021).

Durante os meses de verão tendem a predominar as variações da densidade da água e nos meses de inverno as variações de origem meteorológica (Carinhas *et al.*, 2021).

4. CONSIDERAÇÕES FINAIS

Neste trabalho, procurou-se contribuir para o estudo das variações do NMM na costa de Portugal Continental a partir dos dados recolhidos pelo marégrafo de Sines entre 1997 e 2020. A longa série de dados permitiu apurar que o NMM está a subir neste local com uma tendência de 4.67 ± 0.71 mm/ano. Foram ainda calculados os níveis médios mensais característicos, e verificada a existência de sazonalidade ao longo do ano. Observaram-se acentuadas variações anuais e semi-anuais do nível médio do mar devido a variações da pressão atmosférica, da densidade da água e da circulação oceânica. Em síntese, durante os meses de verão tendem a predominar as variações da densidade da água e nos meses de inverno as variações de origem meteorológica.

Como trabalho futuro, pretendem-se aplicar o modelo de média móvel sazonal integrada autorregressiva (SARIMA) (Murteira *et al.*, 2000), para tentar compreender melhor a estrutura de variabilidade do nível médio e analisar a possibilidade de efetuar previsões mais rigorosas do que é possível obter a partir de modelos de regressão linear.

REFERÊNCIAS

- Antunes, C., Taborda, R., (2009). Sea level at Cascais tide gauge: data, analysis and results. *Journal of Coastal Research*, 218-222
- Cazenave, A., Nerem, R. (2004). Present-day Sea Level Change: Observations and Causes. *Rev. Geophys.* 42. (doi:10.1029/2003RG000139)
- Carinhas, D., Infante, P., Martinho, A., Vasquez, F. As marés ao serviço da variação do nível médio do mar. Apresentado no XXV Congresso da Sociedade Portuguesa de Estatística – SPE 2021, Évora., Portugal, outubro de 2021.
- Da Silva, A. A. A., Freire, E., & Crisóstomo, G. (2014). Variações do nível médio anual do mar em Cascais: características e tendências. *Estudos do Quaternário/Quaternary Studies*, (5).
- Dias, J., Taborda, R. (1988). Evolução recente do nível médio do mar em Portugal, *Anais do Instituto hidrográfico* 9, 83-97
- Mendes, V.B., Barbosa, S.M. and Carinhas, D. (2020). "Vertical land motion in the Iberian Atlantic coast and its implications for sea level change evaluation" *Journal of Applied Geodesy*, vol. 14, no. 3, 2020, pp. 361-378. (doi:10.1515/jag-2020-0012)
- Morettin, P. A., & Toloi, C. M. (2018). *Análise de séries temporais: modelos lineares univariados*. Editora Blucher.
- Murteira, B., Müller, D., & Turkman, K. F. (2000). *Análise de sucessões cronológicas*.

Anexo II - Participação em: Jornadas do Mar'22 - Artigo Submetido

Caracterização da tendência de subida do nível médio em Portugal

M. Picoto (1), D. Carinhas (2) e C. Lemos (3)

(1) Escola Naval; migpicoto@gmail.com.

(2) Instituto Hidrográfico.

(3) Escola Naval

Resumo: As alterações climáticas e as suas consequências são um dos temas dominantes na atualidade. Uma das consequências mais evidentes do aquecimento global é a tendência de subida do nível médio do mar, associada à fusão dos glaciares nas regiões polares e à expansão térmica da água dos oceanos. Esta tendência tem vindo a ser estudada desde há várias décadas a nível global, utilizando séries maregráficas longas e, mais recentemente, outros métodos de observação (por exemplo altimetria por satélite). Neste trabalho, pretende-se estender estudos já efetuados no Instituto Hidrográfico sobre a subida do nível médio do mar através da utilização de regressão linear e de técnicas de análise de séries temporais a registos maregráficos de portos nacionais. Pretende-se ainda identificar fatores com periodicidade sazonal que se sobreponham à tendência geral de variação do nível médio, a partir do cálculo dos níveis médios mensais característicos.

Palavras-chave: modelos de séries temporais, nível médio do mar, registos maregráficos, tendência do nível do mar.

1. INTRODUÇÃO

Atualmente, existe uma grande preocupação com a subida do nível médio do mar (NMM), porque este fenómeno tem, à escala global, um grande impacto na humanidade, nomeadamente nas atividades económicas e sociais.

Muitos são os fatores que influenciam o NMM e, em geral, fazem-se sentir no espaço e no tempo, com maior ou menor intensidade e interligação. A compreensão de muitos destes processos é, em muitos casos complexa, e apenas alguns fenómenos, como a maré astronómica, podem ser previstos de forma rigorosa. Em contrapartida, as variações inter- anuais são mais difíceis compreender e de prever. Em estudos de análise das variações do NMM, cujo objetivo é provar a relação entre estas e o atual aquecimento global, associado a atividades antrópicas, são muitas vezes omitidos ou subvalorizados os fenómenos a eles associados, os quais embora conhecidos não são quantificáveis de forma determinística (Dias e Taborda, 1988).

As variações na órbita terrestre têm sido apontadas como um dos principais fenómenos de variação do clima e consequentemente do NMM ao longo do período Quaternário, entre 2,7 milhões de anos e 11,7 mil anos atrás. Após o último período glaciário e essencialmente durante o Período Holoceno, que se iniciou há 11,7 mil anos e perdura até ao presente, existem essencialmente dois grandes fenómenos, a nível global, que desencadeiam variações

de longas escalas de tempo no NMM devido às interações Criosfera-Oceano-Atmosfera, sendo eles, os fenómenos eustáticos e isostáticos (Da Silva et al., 2014). Fenómenos eustáticos provocam variações no volume e na massa dos oceanos, enquanto os fenómenos isostáticos estão relacionados com a elevação ou subsidência da crosta terrestre, provocando uma variação relativa do NMM. As flutuações do NMM causadas por fenómenos glácio-eustáticos estão fortemente associadas ao aquecimento e arrefecimento global que, consequentemente, levam à fusão ou à formação das calotas polares e glaciares. Estes processos desenvolvem-se em escalas temporais de vários milénios. Durante os períodos glaciários, o nível do mar desce porque a humidade que é evaporada dos oceanos não é reposta, ficando aprisionada sob a forma de gelo nas calotas polares. Em períodos interglaciários, com temperaturas mais amenas, dá-se o recuo dos glaciares e, consequentemente, o aumento do NMM devido à reposição da água que estava retida nas calotas.

No último século verificou-se um aumento da tendência de subida do NMM, com particular incidência nas últimas décadas (Figura 1).

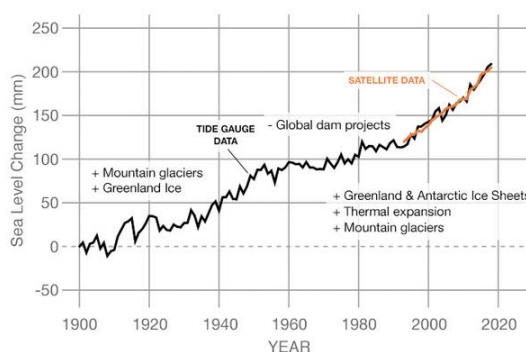


Fig. 1. Nível do mar entre 1900 e 2020: dados obtidos de observações de marés, em estações maregráficas (preto) e através de satélites (vermelho) (fonte: <https://climate.nasa.gov/>).

No entanto, essa subida do NMM é difícil de quantificar, porquanto: (<https://tidesandcurrents.noaa.gov/sltrends/>):

1. Desconhece-se o comportamento geral desta subida, ou seja, desconhece-se o modelo (ou função) aplicável às recentes variações do NMM;
2. Os valores da tendência de subida apresentam diferenças notáveis de local para local (Figura 2);
3. O número de estações maregráficas, com séries de observações, suficientemente, significativas é bastante pequeno e a sua distribuição, pelo globo, é assimétrica.

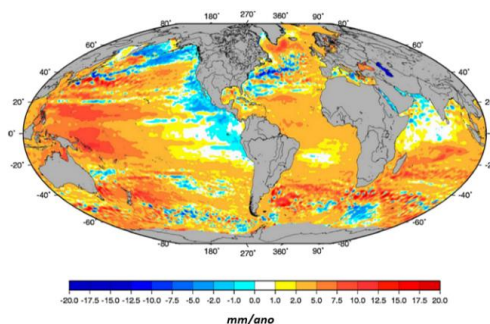


Fig. 2. Distribuição geográfica das tendências do nível do mar entre 1992-2008, obtidas através de altimetria de satélites (adaptado de Cazenave e Nerem, (2004)).

Em Portugal existem várias séries de registos maregráficos provenientes de marégrafos dispersos ao longo do seu litoral. Tanto os registos maregráficos de Cascais como de Lagos evidenciam uma tendência de aumento do NMM (Antunes e Taborda, 2009). No entanto, as análises das séries de dados maregráficos sugerem que o NMM não subiu com uma tendência consistente ao longo do tempo (Antunes e Taborda, 2009). Desde final do século XIX até por volta de 1915-1920, o NMM teve inclusive uma tendência negativa, -5 ± 0.4 mm/ano em Cascais (Ferreira et al., 2008). Só posteriormente a este período é que se nota uma tendência de subida do NMM. Mendes *et al.*, (2020), tendo por bases a utilização de registos GPS, concluíram que os movimentos verticais na Península Ibérica são, na sua maioria, caracterizados por tendências de subsidência lentas/moderadas, com um valor médio de 0.6mm/ano, sendo que em Cascais se verificou uma tendência de 0.4 mm/ano. Isto significa que as séries maregráficas da estação de Cascais são fidedignas para serem usadas em estudos globais sobre o NMM e que podem ser usadas para caracterizar a tendência de subida do NMM.

2. MATERIAIS E MÉTODOS

A estatística desempenha um papel fundamental no processamento de dados porque permite inferir parâmetros caracterizadores de fenómenos com uma componente aleatória a partir de amostras. O seu principal objetivo é a recolha, a análise dos dados, inferência (determinação de parâmetros de distribuições ou relações funcionais) e interpretação dos resultados, de forma qualitativa e quantitativa. Quando se efetuam medições, é sempre expectável uma certa variabilidade da característica ou atributo que se está a medir. A variabilidade (também chamada de dispersão) refere-se à distribuição de um conjunto de dados. As quatro principais formas de descrever a variabilidade em um conjunto de dados são: Amplitude, Amplitude

interquartil, Variância e Desvio padrão. Graficamente, a variabilidade pode ser avaliada através de gráficos *box-plot* e/ou histogramas.

Por definição, uma série temporal é uma sucessão de observações que evoluem no tempo, igualmente espaçadas (Murteira *et al.*, 2000),

Segundo o modelo clássico todas as séries temporais são compostas de quatro padrões:

- tendência (T), que é o comportamento de longo prazo da série, que pode ser causada pelo crescimento ou decrescimento que afete a variável de interesse ao longo do tempo;

- variações cíclicas ou ciclos (C), flutuações nos valores da variável com duração superior a um ano, e que se repetem com certa periodicidade, que podem ser resultado de variações de fenómenos climáticos como o El Niño (que se repete com periodicidade superior a um ano);

- variações sazonais ou sazonalidade (S), flutuações nos valores da variável com duração inferior a um ano, e que se repetem todos os anos, geralmente em função das estações do ano, se os dados forem registados anualmente não haverá influência da sazonalidade na série;

- variações irregulares (I), que são as flutuações aleatórias, não correlacionadas entre si (ao contrário das variações descritas anteriormente), resultando de fatores fortuitos e inesperados, como fenómenos extremos e catástrofes naturais.

É importante salientar que nem sempre uma série temporal, mesmo que o modelo clássico seja considerado apropriado para analisá-la, irá apresentar todos os componentes citados.

A tendência descreve o comportamento da variável retratada na série temporal no longo prazo. Um dos objetivos principais na sua identificação é avaliar o seu comportamento para utilizá-lo em previsões. A obtenção da tendência pode ser feita de três formas: através de um modelo de regressão (como o modelo linear - reta), através de médias móveis, ou através de ajuste exponencial (que não deixa de ser uma média móvel) (Morettin e Toloi, 2018). O procedimento para a obtenção da tendência através da reta de regressão é semelhante ao usado na regressão linear simples, mas agora a variável independente será sempre o tempo. Para uma série registada anualmente, por exemplo, de 2005 a 2014, a variável independente assumiria os valores dos anos. Para uma série registada mensalmente, por exemplo, com 60 meses, a variável independente poderia assumir os valores de 1 a 60. A tendência (T) pode ser obtida através da equação $T_t = b_1 t + b_0$;

onde:

t é o valor do tempo, no caso linear b_1 é o declive da reta (se positivo indica tendência crescente, se negativo a tendência é decrescente) e b_0 é o coeficiente linear da reta. As equações dos coeficientes são expressas por (Morettin e Toloi, 2018),

$$b = \frac{n \times \sum_{i=1}^n (t_i \times Y_i) - \sum_{i=1}^n t_i \times \sum_{i=1}^n Y_i}{n \times \sum_{i=1}^n (t_i^2) - \left(\sum_{i=1}^n t_i \right)^2} \quad a = \frac{\sum_{i=1}^n Y_i - b \times \sum_{i=1}^n t_i}{n}$$

onde Y_i é um valor qualquer da variável registada na série temporal, t_i é o período associado a Y_i , e n é o número de períodos da série. Para encontrar os coeficientes basta calcular os somatórios (tal como na análise de regressão linear simples). Caso se pretenda inferir a variabilidade da tendência deverá ser feito o ajuste de um polinómio de segundo grau, dado pela equação $T_t = b_2 t^2 + b_1 t + b_0$;

onde,

b_2 é a aceleração. O termo b_2 permite identificar se a tendência permaneceu ou não constante ao longo de t . A estimação dos coeficientes b_2, b_1 e b_0 pode ser feita através do método dos mínimos quadrados que é dado por (Murteira *et al.*, 2000),

$$f(\beta_0, \beta_1, \dots, \beta_m) = \sum_{t=1}^n (T_t - \beta_0 - \beta_1 t - \dots - \beta_m t^m)^2$$

Onde m representa o grau do polinómio.

No que diz respeito ao estudo do NMM enquanto série temporal, a tendência e a sazonalidade justificam grande parte do comportamento geral da série. Portanto, a utilização de modelos de tendência cíclica pode ser vista como uma abordagem interessante e adequada. Num modelo de tendência cíclica supõem-se que o comportamento geral da série é explicado por um número de sinusóides e por uma da variável aleatória. Um modelo de tendência cíclica, conforme Murteira *et al.*, (2000), é expresso por,

$$T_t = b + \sum_{p=1}^m [a_p \cos(\omega_p^* t) + b_p \sin(\omega_p^* t)],$$

Onde ω_p^* , $p = 1, 2, \dots, m$ representam as frequências das sinusóides, que contribuem, significativamente, para a variância da série observada e b representa a variável aleatória.

Importa agora identificar a função que melhor se ajusta aos valores da série. O coeficiente de determinação, R^2 , é utilizado na avaliação do ajuste. O R^2 varia entre 0 e 1 e o melhor ajustamento é verificado com a obtenção do R^2 mais elevado. O coeficiente de determinação é dado pela equação (Cryer e Chan, (2008); Quinino *et al.*, 1991),

$$r^2 = \frac{\sum_{i=1}^n (\hat{X}_i - \bar{X})^2}{\sum_{i=1}^n (X_i - \bar{X})^2} = 1 - \frac{\sum_{i=1}^n e_i^2}{\sum_{i=1}^n (X_i - \bar{X})^2},$$

Onde $\hat{X}_i$ e X_i são, respetivamente, os valores previstos e observados no instante i ; $\bar{X}$ é a média aritmética da série observada; e $e_i^2 = (\hat{X}_i - X_i)^2$.

A previsão dos valores futuros de uma série temporal é um dos objetivos mais importantes da sua análise e pode obter-se através de várias metodologias. Uma das metodologias mais populares, com vista à obtenção de previsões mais precisas, consiste em usar modelos de média móvel integrada autorregressiva (ARIMA). Segundo Murteira *et al.*, (2000), um modelo ARIMA(p,d,q) é dado pela expressão

$$\phi(B)(1-B)^d T_t = \theta_0 + \theta_q(B)\varepsilon_t,$$

Onde $\phi_p(B) = 1 - \phi_1 B - \dots - \phi_p B^p$ e $\theta_q(B) = 1 - \theta_1 B - \dots - \theta_q B^p$ são, respetivamente, os polinómios autorregressivos e de médias móveis e d é a ordem de integração, ou seja, o número de diferenças necessárias para tornar a série estacionária. Os modelos ARIMA apenas podem ser usados quando não há dados omissos. Estes modelos são sugeridos para trabalhar séries não estacionárias (Makriadakis *et al.*, 1997).

Quando os dados analisados apresentam autocorrelação, deve-se modelar diretamente a estrutura correlacional com um modelo apropriado de série temporal, ou seja, um modelo que remova a autocorrelação dos dados (Makriadakis *et al.*, 1997). Segundo Makriadakis *et al.*, (1997), a verificação da autocorrelação amostral (FAC) pode ser estimada pela equação,

$$\rho_h = Cor[Z_t, Z_{t+h}] = \frac{Cov[Z_t, Z_{t+h}]}{\sqrt{\sigma_t^2 \sigma_{t+h}^2}},$$

Onde h é um número inteiro e

$$\gamma_h = Cov[Z_t, Z_{t+h}] = E[(Z_t - \mu_t)(Z_{t+h} - \mu_{t+h})]$$

é a covariância. A função de autocorrelação parcial (FACP) permite quantificar a correlação entre pares de valor separados por um intervalo h , depois de eliminar o efeito sobre eles das variáveis intermédias. Ou seja, mede a correlação entre T_t e T_{t+h} sem ter em conta o efeito de $T_{t+1}, \dots, T_{t+h-1}$. (Morettin e Toloi, 2018).

Contudo, quando nos deparamos com uma série que apresente sazonalidade, necessitamos de um novo modelo que incorpore essa importante característica temporal, surgindo assim os modelos de média móvel integrada autorregressiva sazonais (SARIMA). Segundo Morettin e Toloi, (2018), o modelo SARIMA(p,d,q)(P,D,Q) é dado por:

$$\phi_p(B)\Phi_p(B^S)(1-B)^d(1-B^S)^D T_t = \theta_q(B)\Theta(B^S)\varepsilon_t,$$

Onde (p,d,q) são as ordens do modelo referentes à dinâmica já mencionada e (P,D,Q) são as ordens da parte sazonal; $\Phi_1, \dots, \Phi_p$ e $\Theta_1, \dots, \Theta_p$ são, respetivamente, os parâmetros autorregressivos sazonais e de médias móveis sazonais. Para observações mensais, com sazonalidade de 12 meses, temos que $S = 12$.

A modelação ARIMA deve seguir três passos básicos, antes de se proceder aos cálculos das previsões: a) identificação/seleção do modelo; b) estimação; c) verificação. O passo mais importante é o primeiro, em que se determinam os termos autorregressivos, o número de diferenças e o número de termos da média móvel.

Por vezes, após concluídos os passos da modelação ARIMA, podem existir vários modelos que descrevam de forma satisfatória a série temporal, ou seja, que tenham cumprido todos os critérios descritos anteriormente. Nestas situações deve-se recorrer a critérios de seleção por forma a identificar o modelo correto. Entre os critérios de seleção, destaca-se o AIC (*Akaike Information Criterion*) que é dado por (Murteira *et al.*, 2000),

$$AIC = -2\ln[L(\hat{\xi})] + 2k,$$

onde $L(\hat{\xi})$ é a função de verossimilhança; $\hat{\xi}$ são as estimativas de máxima verossimilhança do vetor paramétrico e K é o número de parâmetros do modelo.

Pela forma como o este critério foi concebido verifica-se que são penalizados os modelos com maior número de parâmetros e com maior variância, ou seja, o critério representa um compromisso entre qualidade do ajustamento e parsimónia. De acordo com este critério, o modelo correto é o que representa o melhor compromisso entre o ajustamento conseguido e o número de coeficientes usados para especificar o modelo. Como regra de decisão deve ser escolhido o modelo que apresente o menor valor de AIC (Murteira *et al.*, 2000).

Como forma de seleção do modelo pode-se ainda utilizar-se medidas de qualidade para avaliar a qualidade do ajustamento dos modelos à série. Entre essas medidas destacamos o erro quadrático médio (EQM), dado por (Bayer *et al.*, 2012),

$$EQM(\hat{\theta}) = \frac{1}{N} \sum_{i=1}^N (\hat{\theta}_i - \theta_i)^2$$

onde n corresponde ao número de observações; $\hat{\theta}_i$ e θ_i são, respetivamente, os valores previstos e observados no instante i .

3. RESULTADOS

3.1. Análise Descritiva dos Dados

Com vista à aplicação prática das metodologias apresentadas, a amostra em estudo compreende 288 NMM mensais, observados na estação maregráfica do porto de Sines (Figura 3), referentes ao período de 1997 a 2020. Estes dados, foram disponibilizados pelo Instituto Hidrográfico (IH).

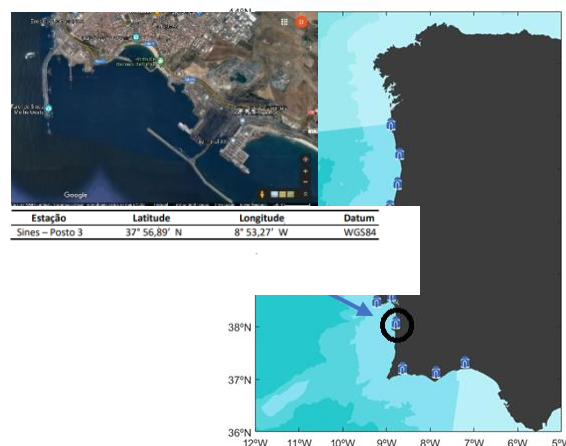


Fig. 3. Identificação do porto estudado (fonte: Divisão de Oceanografia do Instituto Hidrográfico).

O histograma (Figura 4) apresenta uma ligeira assimetria negativa (coeficiente de assimetria igual a -0,070, aproximadamente) demonstrando que a distribuição dos níveis médios mensais se concentra no lado direito da média (2,198 m com Intervalo de Confiança (95%) igual a (2,191; 2,206)) com uma ligeira cauda para o seu lado esquerdo. Através da aplicação do teste de normalidade de *Shapiro-Wilk* (Shapiro e Wilk, 1965), obteve-se o valor $p = 0.9516$ ($\alpha = 0.05$), pelo que não se rejeita a hipótese da normalidade dos dados.

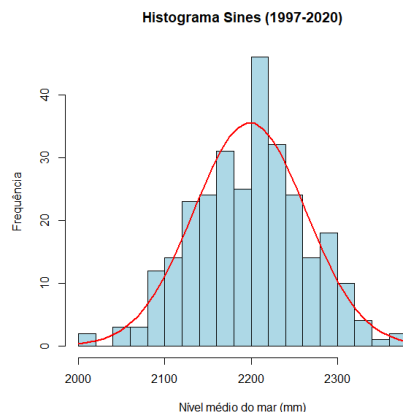


Fig. 4. Histograma de frequências dos níveis médios mensais em Sines no período 1997 - 2020.

A Tabela 1 apresenta as principais estatísticas descritivas dos NMM mensais (1997-2020):

Tabela 1. Estatísticas descritivas dos NMM mensais (mm) de Sines

NMM mensais (mm)	
Média	2198
Mediana	2205
Desvio-padrão	64,88
Assimetria	-0,07
Mínimo	2007
Máximo	2379
Dimensão da amostra	288

A Figura 5 apresenta a evolução dos níveis médios mensais e anuais em Sines até finais de 2020. Foi ajustada uma função linear aos níveis médios anuais, através de um polinómio de grau 1, com vista a representar a tendência dos níveis médios mensais.

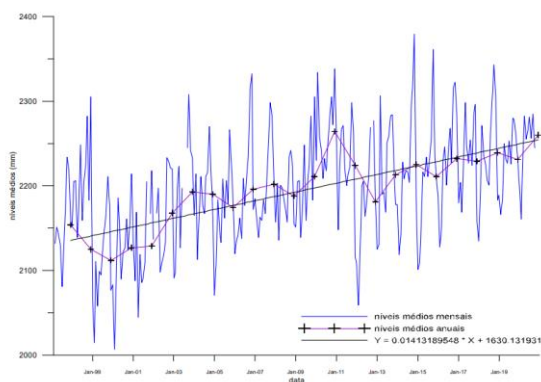


Fig. 5. Evolução dos níveis médios mensais e anuais, em Sines desde 1997 até finais de 2020.

Face ao declive positivo da reta de regressão obtida, pode-se concluir que o nível médio está a aumentar. O incremento foi de 106 milímetros nos 24 anos o que representa uma subida de +4,4 mm/ano.

Posteriormente, foi ajustado um polinómio de grau 2, com vista a determinar se a tendência sofreu alguma aceleração (Tabela 2)

Tabela 2 Coeficientes do modelo estimado

Coeficientes Estimados	<i>p</i> -valeu
$b_0=2137$	2×10^{-16}
$b_1= 0.5419$	0.0009
$b_2= -0.0005951$	0.2738

Através do teste de ANOVA, prova-se que não existe evidência estatística para rejeitar a hipótese de igualdade dos modelos pois, $p - value = 4.209 \times 10^{-15} < 0,05$. O facto de os dois polinómios não serem estatisticamente diferentes permite concluir que não se evidenciou nenhuma aceleração significativa da tendência.

Doravante passaremos a referir-mo-nos ao modelo ao polinómio de primeiro grau e de segundo grau, respetivamente, por modelo 1 e modelo 2.

Durante o ano de 2020 o nível médio mensal máximo foi de 2,441 metros registado em novembro e o nível médio mensal mínimo foi de 2,161 metros registado em fevereiro. Dos 288 valores de nível médio mensal registados desde 1997, 95% situam-se no intervalo [2,08; 2,32] metros, sendo que 5% são superiores a 2,30 metros e apenas 1% são inferiores a 2,01 metros.

A Figura 6 apresenta as distribuições mensais dos níveis médios do mar registados no marégrafo de Sines 1997 e 2020:

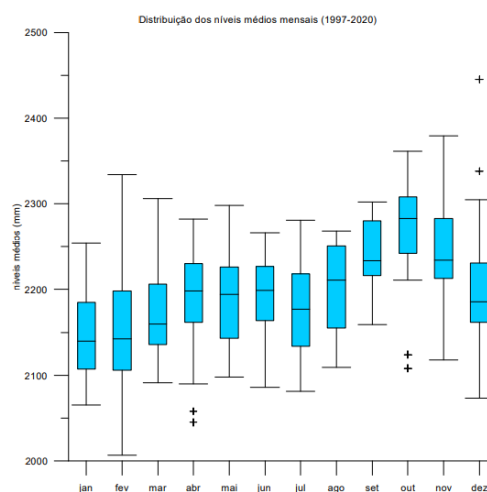


Fig. 6. Comportamento dos níveis médios ao longo do ano, em Sines no período 1997 a 2020.

É possível verificar que a variação do nível médio tem uma componente sazonal. Verificam-se acentuadas variações anuais e semi-anuais do NMM devido a variações sazonais da pressão atmosférica, da densidade da água e da circulação oceânica (Carinhas *et al.*, 2021). Esta análise é consistente com estudos já efetuados no IH. A Figura 7 ilustra os registos dos níveis médios mensais característicos para os diferentes portos nacionais:

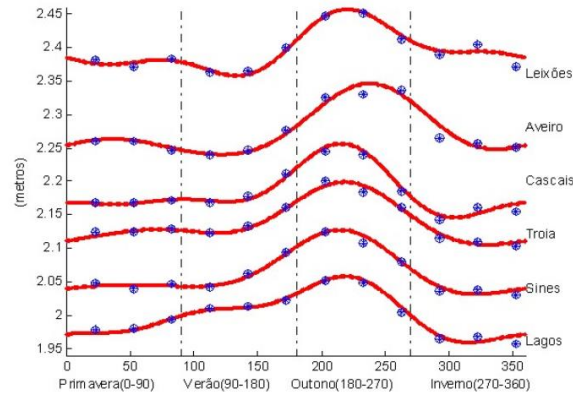


Fig. 7. Níveis médios mensais característicos (pontos azuis) com ajustamento dos níveis médios diários (vermelho) (fonte: Carinhas, *et al.*, 2021).

Durante os meses de verão tendem a predominar as variações da densidade da água e nos meses de inverno as variações de origem meteorológica (Carinhas *et al.*, 2021).

A Figura 7 é indicadora de que o NMM, durante o ano, tem sazonalidade. Recorreu-se a modelos de tendência cíclica onde se assume que existe um número de sinusóides que descrevem essas variações sazonais. Surgem assim duas hipóteses:

1. um modelo que considera apenas sazonalidade anual (uma sinusóide), com período igual a 12 meses, que segue a forma:

$$T_1 = b_0 + b_1 t + \alpha_{12} \cos\left(\frac{2\pi}{12}\right) + b_{12} \sin\left(\frac{2\pi}{12}\right)$$

Tabela 3 - coeficientes estimados do modelo 3

Coeficientes Estimados	<i>p</i> -valeus
$b_0 = 2146.672$	$< 2 \times 10^{-16}$
$b_1 = 0.359$	$< 2 \times 10^{-16}$
$\alpha_{12} = 8.326$	0.046
$b_{12} = -37.899$	$< 2 \times 10^{-16}$

Doravante passaremos a referir-mo-nos a este modelo por modelo 3.

2. um modelo que considera sazonalidade anual e semianual (duas sinusóides), respetivamente, com períodos iguais a 12 e 6 meses, que segue a forma:

$$T_2 = b_0 + b_1 t + \alpha_{12} \cos\left(\frac{2\pi}{12}\right) + b_{12} \sin\left(\frac{2\pi}{12}\right) + \alpha_6 \cos\left(\frac{4\pi}{12}\right) + b_6 \sin\left(\frac{4\pi}{12}\right)$$

Tabela 4 -coeficientes estimados do modelo 4

Coeficientes Estimados	<i>p-valeus</i>
$b_0 = 2147.159$	$< 2 \times 10^{-16}$
$b_1 = 0.356$	$< 2 \times 10^{-16}$
$\alpha_{12} = 8.329$	0.026
$b_{12} = 37.911$	$< 2 \times 10^{-16}$
$\alpha_6 = -7.532$	0.044
$b_6 = -31.264$	2.25×10^{-15}

Doravante passaremos a referir-mo-nos a este modelo por modelo 4.

Através do teste de ANOVA, prova-se que existe evidência estatística para rejeitar a hipótese de igualdade entre os modelos 3 e 4 pois obteve-se $p - value = 4.209 \times 10^{-15} < 0,05$. Prova-se também que existe evidência estatística para rejeitar a hipótese de igualdade entre os modelos 1, 3 e 4 $p - value = 4.209 \times 10^{-15} < 0,05$, ou seja, os modelos 1, 3 e 4 são estatisticamente.

Assim, entre as quatro soluções estatísticas apresentadas procedeu-se ao ensaio e seleção do modelo matemático que melhor se ajusta à série temporal. O melhor ajustamento é verificado com a obtenção do maior coeficiente de determinação (R^2). Obteve-se os seguintes coeficientes de determinação: modelo 1 ($R^2 = 0.23$), modelo 2 ($R^2 = 0.23$), modelo 3 ($R^2 = 0.41$) e modelo 4 ($R^2 = 0.53$). O modelo 4 apresenta correlações mais forte e, desta forma, reflete o melhor ajustamento, Figura 8.

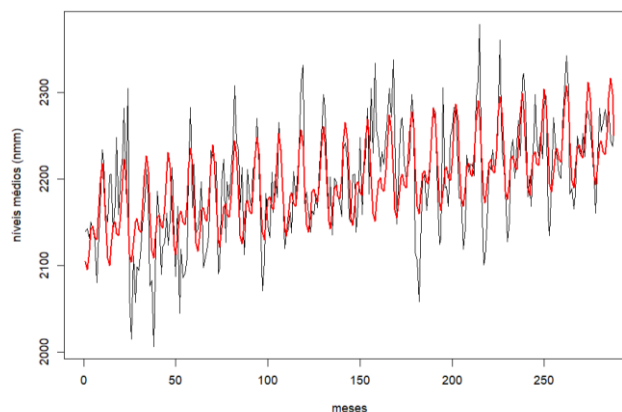


Figura 8 Ajustamento do modelo tendência cíclica 2

Deste modo, o ajustamento de um modelo de tendência cíclica vem confirmar que a sazonalidade é uma componente com grande preponderância nos valores observados do NMM.

3.2. Modelagem e Previsão

Com o objetivo de efetuar previsões da subida do NMM, vamos identificar o modelo com melhor ajustamento e estimar os seus coeficientes, para tal, consideramos apenas os primeiros 230 meses (19) anos, os restantes 58 meses (aproximadamente 5 anos) serão usados para aferir a qualidade da previsão dos modelos identificados. A Figura 9 apresenta a representação gráfica dos primeiros 230 meses.

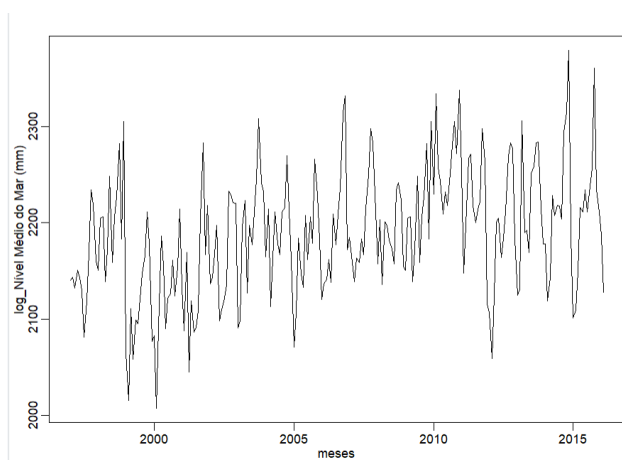


Figura 9 - Registos maregráficos de Sines- primeiros 230 meses

Na Figura 10 encontra-se a representação gráfica da série depois de se estabilizar a variância, através da logaritmização. No entanto, não aparenta ter tido diferenças muito significativas quando comparado com a série origina, Figura 9.

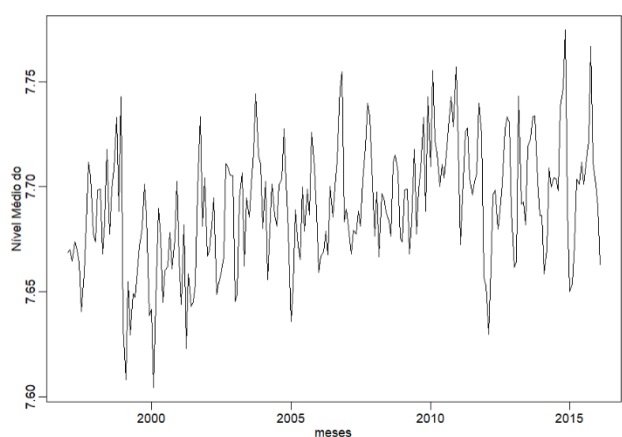


Figura 10 - Estabilização da variância

A série resultante é manifestamente não estacionária, notando-se uma ligeira tendência crescente e a existência de movimentos periódicos com grande significado.

Outro método visual de verificar a não estacionaridade é através da inspeção das FAC e FACP, (Figura 11), donde se verifica que a transformação não é suficiente para garantir a estacionaridade. Aliás, a FAC descreve movimentos periódicos e um decaimento lento para zero sobre os “lags” múltiplo de 12. Para além da necessidade de se estabilizar a sazonalidade, o facto da FAC decair lentamente para zero indica a necessidade de aplicar diferença simples para estabilizar a média.

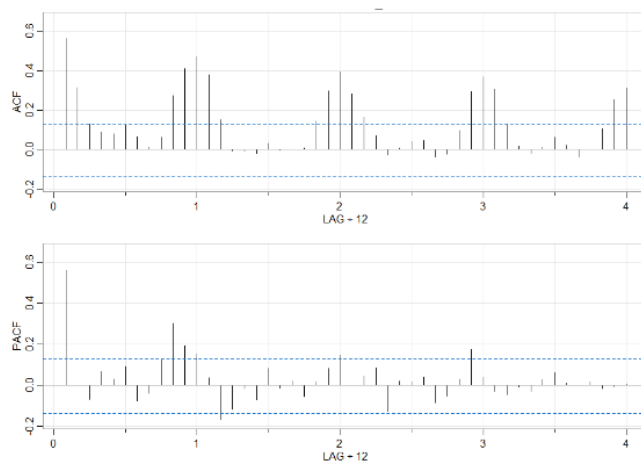


Figura 11 - FAC e FACP após estabilizada a variância

A série transformada foi sujeita a uma diferenciação simples e a uma diferenciação sazonal, Figura 12.

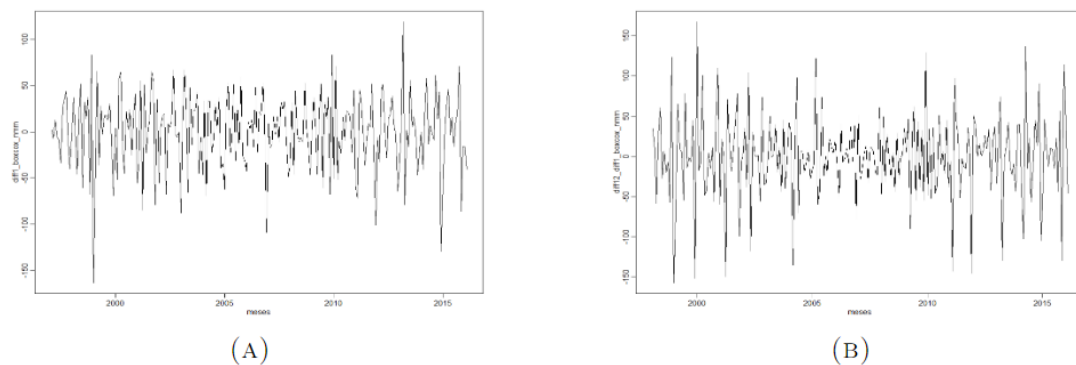


Figura 12 - Gráfico (A) - Diferenciação simples e Gráfico (B) - Diferenciação sazonal

A série está agora estacionária uma vez que os valores se encontram centrados em torno de zero. Após aplicadas os três tipos de transformações e com a série estacionária, vamos agora analisar a FAC e FACP geradas, por forma a identificar modelos que descrevam a série de forma satisfatória, (Figura 13).

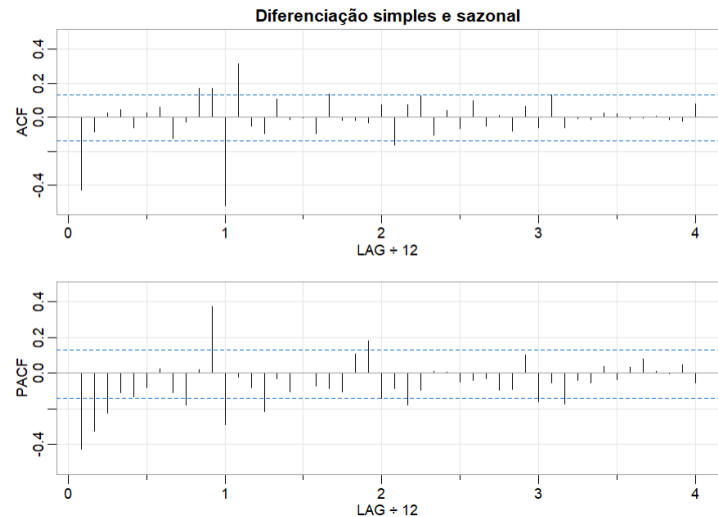


Figura 13 FAC e FACP da série estacionária

Nesta situação nota-se que tanto a FAC como a FACP da série transformada descrevem dois tipos distintos de comportamento, um sobre os “lags” $K = 1, 2, 3, \dots$ e outro para os “lags” K múltiplo de 12, expectável devido à oscilação anual do NMM. Portanto, estamos perante um modelo $SARIMA(p,d,q) \times (P,D,Q)_{12}$, onde a primeira parte ajuda a identificar (p,d,q) e a segunda ajuda a identificar (P,D,Q) . Na primeira parte, a FAC descreve um decaimento brusco para zero depois do “lags” 1 enquanto a FACP descreve um decaimento exponencial para zero, sugerindo um ajustamento $p = 0$, $d = 1$ (diferenciação simples) e $q = 1$.

Na segunda parte, tanto a FAC como a FACP descrevem um rápido decaimento exponencial para zero. Como tal, não dão uma clara indicação do modelo que dever ser escolhido, surgindo várias hipóteses nomeadamente, $(P, D, Q) = (0,1,1)$ ou $(1,1,0)$ ou $(1,1,1)$.

Desta forma, através das indicações dos correlogramas podemos equacionar os seguintes modelos:

$SARIMA(0, 1, 1) \times (0, 1, 1)_{12}$
$SARIMA(0, 1, 1) \times (1, 1, 0)_{12}$
$SARIMA(0, 1, 1) \times (1, 1, 1)_{12}$
$SARIMA(11, 1, 0)_{12}$
$SARIMA(1, 1, 1) \times (0, 1, 1)_{12}$

Após se ter determinado os modelos candidatos e estimados os respetivos coeficientes foi realizada a análise de diagnóstico, com o objetivo de verificar quais dos modelos cumprem

todas as condições para assegurar tanto a qualidade estatística como qualidade de ajustamento.

De todos os modelos, apenas os modelos $SARIMA(1,1,1) \times (0,1,1)_{12}$ e $SARIMA(11,1,0)_{12}$ se encontram em conformidade com os seguintes pressupostos:

1. Parâmetros significativamente diferentes de zero. As Figuras 14 e 15 apresentam os intervalos de confiança de cada um dos coeficientes estimados. Visto que nenhum dos intervalos de contém o valor zero, ou seja, os intervalos variaram apenas entre valores negativos ou positivos, todos eles são considerados significativos.

	2.5 %	97.5 %
ar1	-0.6816265	-0.42676408
ar2	-0.6318114	-0.35191355
ar3	-0.6719012	-0.39737165
ar4	-0.6401602	-0.35806417
ar5	-0.6847620	-0.39140151
ar6	-0.5602412	-0.25474246
ar7	-0.5483347	-0.25819370
ar8	-0.6346785	-0.35518530
ar9	-0.6676566	-0.39279522
ar10	-0.4709557	-0.19010412
ar11	-0.3172849	-0.06162307

Figura 14 - Intervalos de confiança dos coeficientes do modelo $SARIMA(11,1,0)_{12}$

	2.5 %	97.5 %
ar1	0.02008988	0.4006043
ma1	-0.99807812	-0.7488560
sma1	-1.22435882	-0.7740980

Figura 15 - Intervalos de confiança dos coeficientes do modelo $SARIMA(1,1,1) \times (0,1,1)_{12}$

2. FAC e FACP residuais têm os seus valores compreendidos entre os limites, Figuras 16 e 17. No caso do modelo $SARIMA(1,1,1) \times (0,1,1)_{12}$ nota-se a existência de dois valores fora do intervalo de confiança da FAC, mas pelo facto de cumprir com todos os outros critérios, também foi considerado;

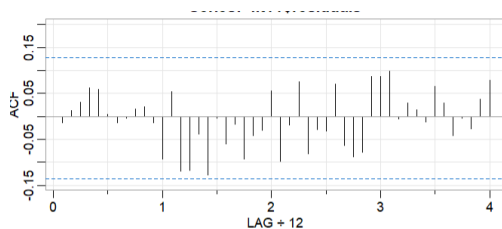


Figura 16 - FAC do modelo $SARIMA(11,1,0)_{12}$

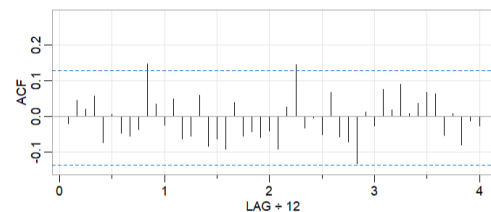


Figura 17 - FAC do modelo $SARIMA(1,1,1) \times (0,1,1)_{12}$

3. Os resíduos não são correlacionados. A aplicando o teste de *Ljung e Box [1978]* comprova-se que os resíduos são independentes, pois obteve-se os valores $p - value = 0.4599 > 0,04$ e $p - value = 0.0472 > 0,04$, respetivamente, para os modelos $SARIMA(1,1,1) \times (0,1,1)_{12}$ e $SARIMA(11,1,0)_{12}$.

A etapa de avaliação permitiu concluir que apenas dois modelos são adequados para descrever a série temporal. Passaremos agora à escolha do modelo que melhor descreva a série em questão, tendo por base o critério AIC, o erro quadrático médio e o ajustamento gráfico das previsões à série temporal completa.

Tabela 5 - Modelos selecionados

Modelo	AIC	Erro quadrático médio
SARIMA (11,1,1)(12)	2453.67	2311.29
SARIMA (1,1,1) (0,1,1)(12)	2138,69	1849.88

A análise da Tabela 5 permite concluir que o segundo modelo, apesar de menos parcimonioso, garante um melhor ajustamento. Com efeito, o segundo modelo apresenta o menor valor do critério AIC e de erro quadrático médio.

Por último, comparamos graficamente as previsões com os meses da série que não foram considerados na obtenção do modelo (os últimos 58 meses). A Figura 17 mostra que, apesar de alguma volatilidade, as previsões apresentam um comportamento semelhante aos valores reais.

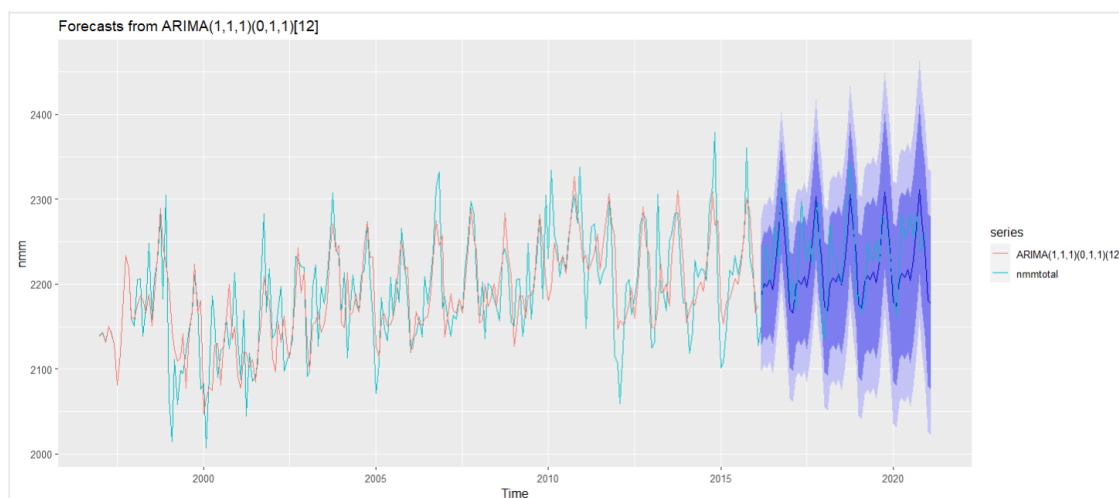


Figura 17 - Ajustamento das previsões aos meses não considerados

Podemos assim concluir que o modelo $SARIMA (1,1,1) \times (0,1,1)_{12}$ é aquele que descreve da melhor forma possível a série temporal em questão.

Escolhido o modelo, procedeu-se à previsão dos níveis médios para um horizonte de 10 anos (120 meses). Na Figura 18 apresenta-se a evolução das previsões e os respetivos intervalos de confiança a 95%.

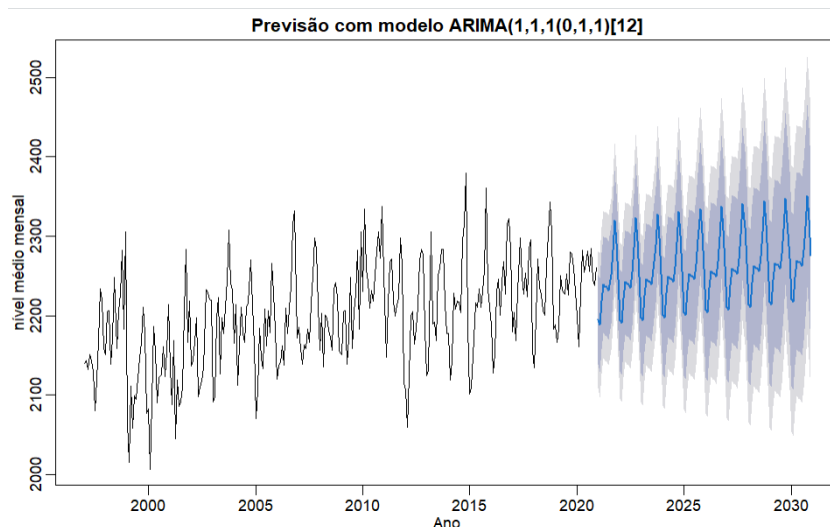


Figura 18 - Previsão dos níveis médios mensais para os próximos 10 anos

Verificam-se acentuadas variações anuais do NMM devido à componente sazonal. De notar a existência de tendência positiva, qua vai de encontro à tendência estimada (na ordem de +4.36 mm/ano) pelo ajustamento de um polinómio de grau 1.

4. CONSIDERAÇÕES FINAIS

Neste trabalho, procurou-se contribuir para o estudo das variações do NMM na costa de Portugal Continental a partir dos dados recolhidos pelo marégrafo de Sines entre 1997 e 2020. A longa série de dados permitiu apurar que o NMM está a subir neste local com uma tendência de 4,44 mm/ano. Foram ainda calculados os níveis médios mensais característicos, e verificada a existência de sazonalidade anual e durante o ano. Observaram-se acentuadas variações anuais e semi-anuais do NMM devido a variações da pressão atmosférica, da densidade da água e da circulação oceânica. Em síntese, durante os meses de verão tendem a predominar as variações da densidade da água e nos meses de inverno as variações de origem meteorológica. As variações anuais e semi-anuais foram ainda confirmadas através de ajustamento de modelos de tendência cíclica e pelas funções de autocorrelação e autocorrelação parcial.

Foi realizado o ajuste de um modelo de séries temporais da classe ARIMA a uma série de registos maregráficos. Visto que a série não era estacionária, aplicaram-se três

transformações com o intuito de estabilizar a variância, a média e a sazonalidade, permitindo, desta forma, modelar a correlação entre a série.

Dos modelos identificados, o selecionado foi o modelo selecionado foi o *SARIMA* $(1,1,1) \times (0,1,1)_{12}$. Na fase de diagnóstico, o modelo demonstrou ajustar-se bem aos dados, tendo cumprido com todos os requisitos necessários.

Foram feitas previsões para um horizonte de 10 anos. Os valores estimados acompanham o comportamento geral da série e apresentam uma tendência idêntica, aproximadamente 4,36 mm/ano. No entanto, devido à volatilidade da interação climático-oceânica é muito provável que esta estimativa sofra alterações em escalas de tempo de curto e médio prazo. Não obstante, se o valor estimado permanecer contante, podemos antever que em 2100 o NMM esteja aproximadamente 35 cm acima do atual.

Devemos continuar a melhorar a previsão, especialmente a precisão da previsão de séries temporais, é uma tarefa importante, embora muitas vezes difícil, que os decisores enfrentam em muitas áreas e a oceanografia não é exceção. Apesar dos inúmeros modelos de séries temporais disponíveis, a pesquisa para melhorar a eficácia dos modelos de previsão nunca para.

5. REFERÊNCIAS

- Antunes, C., Taborda, R., (2009). Sea level at Cascais tide gauge: data, analysis and results. *Journal of Coastal Research*, 218-222.
- Bayer, D. M., Castro, N. M. D. R., & Bayer, F. M. (2012). Modelagem e previsão de vazões médias mensais do rio Potiribu utilizando modelos de séries temporais. *Rbrh: revista brasileira de recursos hídricos*. Porto Alegre, RS. Vol. 17, n. 2 (abr./jun. 2012), p. 229-239.
- Cazenave, A., Nerem, R. (2004). Present-day Sea Level Change: Observations and Causes. *Rev. Geophys.* 42. (doi:10.1029/2003RG000139).
- Carinhas, D., Infante, P., Martinho, A., Vasquez, F. As marés ao serviço da variação do nível médio do mar. Apresentado no XXV Congresso da Sociedade Portuguesa de Estatística – SPE 2021, Évora., Portugal, outubro de 2021.
- Cryer, J. D., & Chan, K. S. (2008). *Time series analysis: with applications in R* (Vol. 2). New York: Springer.

- Da Silva, A. A. A., Freire, E., & Crisóstomo, G. (2014). Variações do nível médio anual do mar em Cascais: características e tendências. *Estudos do Quaternário/Quaternary Studies*, (5).
- Dias, J., Taborda, R. (1988). Evolução recente do nível médio do mar em Portugal, *Anais do Instituto hidrográfico* 9, 83-97.
- Mendes, V. B., Barbosa, S. M. and Carinhas, D. (2020). "Vertical land motion in the Iberian Atlantic coast and its implications for sea level change evaluation" *Journal of Applied Geodesy*, vol. 14, no. 3, 2020, pp. 361-378. ([doi:10.1515/jag-2020-0012](https://doi.org/10.1515/jag-2020-0012)).
- Makridakis, S., Wheelwright, S., Hyndman, R. (1997). *Forecasting methods and applications*, 3 ed. Wiley.
- Morettin, P. A., & Toloi, C. M. (2018). *Análise de séries temporais: modelos lineares univariados*. Editora Blucher.
- Murteira, B., Müller, D., & Turkman, K. F. (2000). *Análise de sucessões cronológicas*.
- Quinino, R. C., Reis, E. A., & Bessegato, L. F. (1991). O coeficiente de determinação R^2 como instrumento didático para avaliar a utilidade de um modelo de regressão linear múltipla. Belo Horizonte: UFMG.
- Shapiro, S., Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, 52, p.591-611.