



ESCOLA NAVAL

la sainte & bien faire



Ângela Beatriz Ramos Rodrigues de Castro

Desenvolvimento de um sistema de aprendizagem automática (Machine Learning) para previsão de condições de agitação marítima

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



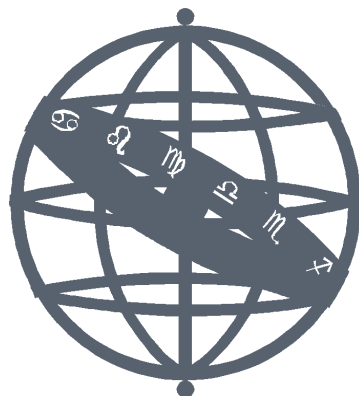
Alfeite

2024



ESCOLA NAVAL

talento & bem-fazer



Ângela Beatriz Ramos Rodrigues de Castro

Desenvolvimento de um sistema de aprendizagem automática (Machine Learning) para previsão de condições de agitação marítima

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

Orientação de : CMG RES Carlos Miguel Reis Silva de Oliveira e Lemos

Co-orientação de : CFR TSN-AMB Luís Quaresma dos Santos

O Aluno Mestrando ,

Assinado por: **ÂNGELA BEATRIZ RAMOS RODRIGUES DE CASTRO**

Num. de Identificação: BI14652027

Data: 2024.09.26 22.10.32 GMT Daylight time



Ângela Castro

O Orientador ,

Assinado por: **Carlos Miguel Reis Silva de Oliveira e Lemos**

Num. de Identificação: 05327698

Data: 2024.09.26 22:39:51+01'00'



Carlos Lemos

Alfeite

2024

Agradecimentos

Este trabalho representa o culminar do meu percurso na Escola Naval, um marco que só foi possível graças ao apoio inestimável daqueles a quem dedico os seguintes agradecimentos:

Aos meus orientadores, CMG Oliveira e Lemos e CFR Quaresma dos Santos, manifesto a minha profunda gratidão pela oportunidade de desenvolver este trabalho sob a vossa orientação. A vossa disponibilidade, dedicação e rigor académico foram essenciais para a concretização deste projeto. Agradeço, em particular, os valiosos contributos e a paciência que demonstraram sempre ao longo deste processo.

À equipa do CGEOMETOC, pelo acolhimento caloroso, pela vossa ajuda e por criarem um ambiente propício à realização deste trabalho. A vossa colaboração foi crucial para o sucesso deste projeto.

À Escola Naval, em particular, aos meus camaradas de curso, Fernão de Magalhães, por durante estes cinco anos me proporcionarem momentos que recordarei sempre, e por me ensinarem o verdadeiro significado da camaradagem.

Aos meus amigos, agradeço a compreensão e o apoio constantes, mesmo durante as longas ausências que a minha carreira naval exigiu. A vossa amizade foi um pilar fundamental ao longo desta jornada.

Aos meus pais e ao meu irmão, expresso o meu mais profundo agradecimento pelos valores que me transmitiram, pelo amor incondicional e pelo apoio em todas as fases da minha vida. Sem vocês, nada disto teria sido possível.

Por último, agradeço a todos os que, direta ou indiretamente, contribuíram para a realização desta tese ou para a minha formação ao longo destes anos.

Resumo

O conhecimento das condições de agitação marítima é essencial para o planeamento e condução das operações navais e anfíbias, bem como para as atividades da Autoridade Marítima Nacional, da Proteção Civil, e de outros serviços públicos. A previsão do estado do mar é atualmente obtida através de modelos computacionais que integram equações matemáticas capazes de descrever os processos de geração, propagação e transformação das ondas de vento. Esses modelos requerem a utilização de meios computacionais poderosos, operados continuamente por equipas especializadas, e a disseminação dos resultados desses modelos pode ser comprometida por ciberataques ou ações de Guerra Eletrónica. Neste trabalho, é apresentado um método de previsão baseado em modelos de aprendizagem automática (*Machine Learning*) supervisionada. O treino dos modelos foi efetuado a partir de registos dos modelos matemáticos de previsão e agitação marítima operados no Centro de Excelência Geoespacial, Meteorológico e Oceanográfico (CGEOMETOC) da NATO e dos dados da boia de Sines operada pelo Instituto Hidrográfico (IH). Estes modelos apresentam as vantagens de, uma vez treinados de forma adequada, permitirem obter previsões das condições de agitação marítima com um esforço computacional muito inferior ao dos modelos tradicionais, e de concederem uma maior autonomia às forças e unidades navais nas suas operações. O presente trabalho apresenta um passo significativo na tentativa de reduzir o esforço computacional despendidos pelo CGEOMETOC na produção de informação disponibilizada ao público sobre as condições de agitação marítima nas costas de Portugal, e de contribuir para uma maior autonomia e resiliência da elaboração das previsões meteo-oceanográficas por parte das forças e unidades navais empenhadas em operações. Representa também um contributo para o problema geral da previsão das condições de agitação marítima através de modelos de aprendizagem automática, que é de momento uma área de investigação ativa.

Palavras-chave: Agitação Marítima, Modelos de Previsão, Aprendizagem Automática, Redes Neurais

Abstract

Forecasting the sea state conditions is essential for planning and conducting naval and amphibious operations, as well as the activities of the Nation's Maritime Authority, Civil Protection and other public services. At present, sea state forecasting is mostly done with computational models based on mathematical equations that describe the processes of generation, propagation and transformation of wind waves. These models require the use of powerful computational resources, operated continuously by specialized teams. Cyberattacks or Electronic Warfare actions may compromise the dissemination of the results of these models to naval forces. In this work, a prediction method based on supervised Machine Learning models is presented, in which the data for the models training and testing were obtained from the numerical solutions provided by the mathematical models operated in the NATO Maritime Geospacial, Meteorological, and Oceanographic Center of Excellence (CGEOMETOC), and from the wave records from the Sines wave buoy operated by the Instituto Hidrográfico (IH) in Lisbon, Portugal. With adequate training, these models offer the advantages of providing sea state forecasts at selected points using a fraction of the computer resources presently used for this purpose, and a greater autonomy of naval forces and units in forecasting the sea state. The present work represents a step toward a significant reduction of the human and computational resources expended by CGEOMETOC in producing information made available to the public about the sea state conditions along the Portuguese coast. It also seeks to contribute to greater autonomy and resilience in the preparation of meteorological and oceanographic forecasts by naval units engaged in operations. In addition, it also represents a contribution to the broader scientific problem of predicting sea state conditions, as it is an active area of research.

Keywords: Sea Forecasting, Predictions Models, Machine learning, Neural Networks

Índice

Introdução	1
1 Revisão de literatura sobre a dinâmica das ondas de vento	5
1.1 Introdução	5
1.2 Dinâmica das ondas de vento	7
1.2.1 Generalidades	7
1.2.2 Teoria linear das ondas de superfície	8
1.2.3 Descrição estatística da agitação marítima	16
1.2.4 Geração, propagação e transformação das ondas de vento	25
Processo de geração de ondas	27
Dispersão angular e em frequência	32
Processos de transformação de ondas	32
1.3 Modelos físico-matemáticos de previsão de agitação marítima	39
2 Modelos estatísticos e modelos de aprendizagem automática	43
2.1 Séries temporais. Modelos descritivos	44
2.2 Modelo autoregressivo (AR)	48
2.3 Modelo médias móveis (MA)	48
2.4 Modelo ARIMA	49
2.5 Modelos de Regressão Linear	52
2.5.1 Modelo de Regressão Linear Simples	52
2.5.2 Modelo de Regressão Linear Múltipla	52
2.6 Modelos de Aprendizagem Automática: Redes Neurais	54
2.6.1 Processo de Aprendizagem	57
2.6.2 Redes Neurais	58
2.6.3 Funções de ativação	59
Função logística	59
Função tangente hiperbólica	60
Função unidade linear retificada (ReLU)	60
2.6.4 Redes neurais com várias camadas: Aprendizagem profunda	61

2.6.5	<i>Multilayer Perceptron</i>	62
2.6.6	Redes neuronais recorrentes (<i>Recurrent Neural Network</i>)	63
2.7	Modelos propostos para previsão das condições de agitação marítima baseados em Aprendizagem Automática	67
3	Metodologia de desenvolvimento	69
3.1	Conceito de desenvolvimento	69
3.2	Parâmetros Característicos da Agitação Marítima e Elementos Meteorológicos	73
3.3	Descrição, Preparação Prévia e Análise Exploratória das Séries Temporais	74
3.3.1	Descrição das séries temporais em estudo	74
3.3.2	Preparação prévia das séries temporais em estudo	77
3.3.3	Análise exploratória das séries temporais em estudo	81
3.4	Modelos estatísticos e de Aprendizagem Automática	87
3.4.1	Modelo ARIMA	87
3.4.2	Modelo de regressão linear	89
3.4.3	Modelos Aprendizagem Automática: Redes neuronais	90
	Estrutura da rede neuronal	92
	Hiperparâmetros	92
	Fontes de dados	94
4	Resultados	97
4.1	Modelo ARIMA(2,1,5)	97
4.2	Modelo de Regressão Linear com variáveis desfasadas (<i>lagged variables</i>)	102
4.3	Modelo de Aprendizagem Automática: Redes neuronais	105
5	Discussão dos resultados	125
5.1	Seleção dos dados/séries temporais	126
5.2	Seleção de variáveis com base no conhecimento de processos físicos	127
5.3	Estrutura dos modelos	128
5.4	Hiperparâmetros	128
5.5	Contribuições para o problema geral de previsão da agitação marítima	129
5.6	Limitações do estudo	131
	Conclusão	132
	Bibliografia	135

Apêndices	141
A Modelos de Regressão Linear aplicados	141

Lista de Figuras

1	Representação da área de interesse para o presente trabalho: limites geográficos, batimetria, linha de costa, localização das boias ondógrafo e multiparamétrica de recolha de dados de agitação marítima operadas pelo IH. Adaptado de: EMODnet [2024]	3
1.1	Distribuição da densidade de energia das ondas de superfície em função do período e classificação dos diversos sistemas de ondas em termos das forças perturbadoras e de restituição. Adaptado de Pope e Lockhart [2002]	6
1.2	Representação da agitação irregular como a soma de infinitas ondas sinusoidais com altura, período, fase e direção de propagação distintas. Fonte : Holthuijsen [2007]	9
1.3	Parâmetros característicos de uma onda simples (Adaptado :Sousa [2023]).	10
1.4	Representação das condições de propagação definidas pela teoria linear das ondas progressivas: águas profundas (em cima), profundidade intermédia (a meio), fundos baixos (em baixo)	13
1.5	Representação de uma onda sinusoidal e da variação de energia potencial	15
1.6	Ilustração de um registo da elevação de superfície (fonte :World Meteorological Organization [2023])	17
1.7	Representação dos dois critérios de classificação de ondas: zeros descendentes e zeros ascendentes. (Adaptado de: Holthuijsen [2007])	18
1.8	Representação da função de densidade de probabilidade associada à distribuição de Rayleigh	20
1.9	Representação dos espectros de amplitude, variância, densidade de variância contínuo e descontínuo. Fonte : Holthuijsen [2007]	22
1.10	Exemplo de um espectro direcional na costa holandesa, resultante de ondulação gerada por uma tempestade gerada pelo mar da Noruega e vaga gerada localmente. Fonte: Holthuijsen [2007]	24

1.11	Representação da geração e propagação das ondas de vento, adaptado de S. James e Stull [2024]	27
1.12	Representação do conceito de <i>fetch</i> equivalente sobre condições ideais. Adaptado de: Holthuijsen [2007]	29
1.13	Evolução da altura (no topo) e períodos (em baixo) significativos, em função da dimensão do fetch e gama de profundidades. Adaptado de: World Meteorological Organization [2023]	30
1.14	Método manual de previsão de agitação marítima. Adaptado de: World Meteorological Organization [2023]	31
1.15	Representação da propagação de ondas num plano bidimensional e do conceito de ortogonal.	34
1.16	Representação do fluxo de energia entre duas ortogonais	35
1.17	Exemplo da refração que ocorre num local de batimetria retilínea. Adaptado de Pope e Lockhart [2002]	37
1.18	Ilustração do tipo de rebentação. Fonte: Fleit [s.d.]	38
2.1	Representação de um neurónio. Fonte: I. d. Silva, Spatti, Flauzino, Liboni e Reis Alves [2016]	58
2.2	Representação da função logística $\text{sigmoid}(z)$	60
2.3	Representação da função tangente hiperbólica $\text{tanh}(z)$	60
2.4	Representação da função $\text{ReLU}(z)$	61
2.5	Representação da arquitetura de redes <i>Multilayer Perceptron</i> . Neste caso, a camada de saída contém m neurónios, sendo a saída (<i>output</i>) um vetor de comprimento m . Fonte: I. d. Silva, Spatti, Flauzino, Liboni e Reis Alves [2016]	63
2.6	Representação da arquitetura de redes neuronais recorrentes (<i>Recurrent Neural Network</i>). Fonte: I. d. Silva, Spatti, Flauzino, Liboni e Reis Alves [2016]	64
2.7	Representação da arquitetura de um neurónio de uma rede neuronal LSTM (<i>Long Short-Term Memory</i>). Fonte: Goodfellow, Bengio e Courville [2016]	66
3.1	Histograma da variável T_p	70
3.2	Distribuição do período de pico em função da direção principal de propagação da ondulação, agregada em setores. Registos referentes à bóia multiparamétrica de Sines.	71
3.3	Representação da localização do grupo de ondas associado aos períodos de pico 9, 12 e 14 segundos para o período de 24 e 48 horas.	72

3.4	Distribuição das séries temporais da altura significativa espectral H_{m0} (gráfico à esquerda) e do período de pico T_p (gráfico à direita) extraídas da boia multiparamétrica de Sines	75
3.5	Distribuição das séries temporais da altura significativa H_s (gráfico à esquerda) e do período de pico T_p (gráfico à direita) extraídas da malha geográfica do modelo <i>WaveWatch III</i> ao largo do arquipélago dos Açores	76
3.6	Distribuição das séries temporais da altura significativa H_s (gráfico à esquerda) e do período de pico T_p (gráfico à direita) extraídas da malha geográfica do modelo <i>WaveWatch III</i> ao largo da Ilha da Madeira	77
3.7	Representação de um ficheiro Netcdf. Adaptado de: Hoyer e Hamman 2017	78
3.8	Série temporal (1) da altura significativa espectral com lacunas temporais	80
3.9	Série temporal (1) da altura significativa espectral após a interpolação	80
3.10	Série temporal (2) da altura significativa espectral	80
3.11	Série temporal (3) da altura significativa espectral	81
3.12	Função autocorrelação da série temporal da altura significativa (topo), função de autocorrelação parcial da série temporal da altura significativa (segundo gráfico), função autocorrelação da série temporal da altura significativa diferenciada uma vez (terceiro gráfico) e respetiva função de autocorrelação parcial (último gráfico)	82
3.13	Matriz de correlação referente à variável H_{m0} considerando o ponto extraído do modelo WWIII ao largo dos Açores e ao largo da Madeira	83
3.14	Correlação cruzada referente às séries temporais de T_p na boia de Sines e no arquipélago dos Açores.	84
3.15	Correlação cruzada referente às séries temporais de H_{m0} na boia de Sines e no arquipélago dos Açores.	85
3.16	Correlação cruzada referente às séries temporais de H_{m0} na boia de Sines e na Ilha da Madeira.	85
3.17	Correlação cruzada referente às séries temporais de T_p na boia de Sines e na Ilha da Madeira.	86
3.18	Estratégia de exploração dos modelos AA de redes neurais	90
3.19	Divisão do conjunto de dados em conjunto de treino, validação e de teste.	91
4.1	Função autocorrelação dos resíduos do modelo (lado esquerdo) e função de autocorrelação parcial do modelo (lado direito)	98

4.2	Distribuição do valor ao longo do tempo (em cima) e histograma dos resíduos (em baixo) do modelo ARIMA(2,1,5)	99
4.3	Valores observados de H_{m0} na boia de Sines e previsões de 12 horas do modelo ARIMA(2,1,5), com uma janela deslizando de 12 horas.	100
4.4	Valores observados de H_{m0} na boia de Sines e previsões de 24 horas do modelo ARIMA(2,1,5), com uma janela deslizando de 12 horas.	101
4.5	Valores observados de H_{m0} na boia de Sines e previsões de 48 horas do modelo ARIMA(2,1,5), com uma janela deslizando de 12 horas.	101
4.6	Modelo de regressão linear para prever a variável H_{m0} no período de 12 horas, considerando os precursores significativos.	103
4.7	Modelo de regressão linear para prever a variável H_{m0} no período de 12 horas.	104
4.8	Resultado da aplicação do modelo de regressão linear ao conjunto de teste	105
4.9	Resultados da exploração manual da estrutura das redes neurais	107
4.10	Exemplo de um bom ajuste e de <i>overfitting</i> no treino de redes neu- ronais: curvas de aprendizagem da rede (4 X 2 X 1) à esquerda e a rede (4 X 2 X 1) à direita.	108
4.11	Resultados da rede (4 X 64 X 8 X 1) com treino de 250 épocas e com interrupção de treino provocada pela função <i>EarlyStopping</i> .	109
4.12	Resultados da variação do <i>batch size</i> na rede (4 X 5 X 1) com treino de 250 épocas e com interrupção de treino provocada pela função <i>EarlyStopping</i> .	110
4.13	Resultados da variação da taxa de aprendizagem na rede (4 X 5 X 1) com treino de 250 épocas e com interrupção de treino provocada pela função <i>EarlyStopping</i> .	111
4.14	Curvas de aprendizagem da rede (4 X 5 X 1) com taxa de aprendiza- gem 0,001 e 0,0001	112
4.15	Resultados da variação do fator de regularização da rede (4 X 5 X 1) com treino de 250 épocas e com interrupção de treino provocada pela função <i>EarlyStopping</i> .	113
4.16	Resultados da variação das épocas de treino na rede (4 X 5 X 1).	114
4.17	Valores observados e valores previstos pela rede (4 X 5 X 1).	115
4.18	Gráfico do valor de perda da função de custo em função das épocas de treino (à esquerda) e gráfico do valor de MSE da rede (4 X 5 X 1) (à direita).	116

4.19	Resultados da performance da rede neuronal (4 X 5 X 1) e hiperparâmetros selecionados pelo algoritmo <i>Random Search</i> com várias épocas de treino e de acordo com a função <i>EarlyStopping</i> .	117
4.20	Resultados da performance da rede neuronal (4 X 5 X 1) e hiperparâmetros selecionados pelo algoritmo <i>Random Search</i> com várias épocas de treino e de acordo com a função <i>EarlyStopping</i> .	118
4.21	Resultados da performance da rede neuronal (4 X 5 X 1) e hiperparâmetros selecionados pelo algoritmo <i>Random Search</i> com várias épocas de treino e de acordo com a função <i>EarlyStopping</i> .	119
4.22	Resultados da performance da rede neuronal (4 X 5 X 1) e (10 X 21 X 1) com 250 épocas de treino e de acordo com a função <i>EarlyStopping</i> .	121
4.23	Resultados da performance da rede neuronal (4 X 5 X 1) e (17 X 35 X 1) com 250 épocas de treino e de acordo com a função <i>EarlyStopping</i> .	121
4.24	Resultados da performance da rede neuronal (4 X 5 X 1) e (24 X 49 X 1) com 250 épocas de treino e de acordo com a função <i>EarlyStopping</i> .	122
4.25	Resultados da performance da rede neuronal (8 X 17 X 1) com 250 épocas de treino e de acordo com a função <i>EarlyStopping</i> .	123
4.26	Curvas de aprendizagem da rede neuronal (8 X 17 X 1) com 250 épocas de treino.	124
4.27	Valores observados e valores previstos pela rede neuronal (8 X 17 X 1) com 250 épocas de treino.	124

Lista de Tabelas

1.1	Descrição das ondas em diferentes profundidades	15
3.1	Variáveis selecionadas a partir dos registos da boia multiparamétrica.	74
3.2	Variáveis selecionadas a partir dos registos do modelo <i>Wave Watch III</i> .	74
4.1	Coeficientes do Modelo ARIMA com Intervalos de Confiança a 95%	97
4.2	Métricas MAE e RMSE do modelo ARIMA(2,1,5) para janelas desli- zantes de 72, 60, 48, 24 e 12 horas e previsões de 12, 24 e 48 horas.	99
4.3	Valores dos hiperparâmetros da rede neuronal inicial.	106
4.4	Valores possíveis para os hiperparâmetros no algoritmo <i>Random Search</i>	109
4.5	Hiperparâmetros selecionados para a rede (4 X 5 X 1)	114
4.6	Métricas de performance da rede neuronal (4 X 5 X 1) e hiperparâ- metros explorados manualmente.	114
4.7	Valores dos hiperparâmetros devolvidos pelo algoritmo <i>Random Se- arch</i> para uma rede 4 X 5 X 1	116

Lista de Abreviaturas

AA	<i>Aprendizagem Automática</i>
AIC	<i>Akaike Information Criterion</i>
BIC	<i>Bayesian Information Criterion</i>
CGEOMETOC	<i>Centro de Excelência Geospacial Meteorológico e Oceanográfico</i>
IH	<i>Instituto Hidrográfico</i>
MLP	<i>Multilayer Perceptron</i>
MSE	<i>Mean Squared Error</i>
NOAA	<i>National and Oceanic and Atmospheric Administration</i>
RMSE	<i>Root Mean Square Error</i>
RNN	<i>Recurrent Neural Networks</i>
SAR	<i>Search And Rescue</i>
WWIII	<i>Wave Watch III</i>

Lista de Símbolos

a : Amplitude de onda	m
C : Velocidade de propagação	m s^{-1}
C_{∞} : Velocidade de propagação em águas profundas	m s^{-1}
C_g : Velocidade de grupo	m s^{-1}
D : Profundidade	m
$\tilde{d}$: Variável adimensional da profundidade	
E : Densidade de energia	J m^{-2}
Ec : Energia cinética	J
Ep : Energia potencial	J
F : Fluxo de energia	W m^{-1}
$\tilde{F}$: Variável adimensional do <i>fetch</i>	
F_{eq} : <i>Fetch</i> equivalente	m
f : Frequência	Hz
f_p : Frequência de pico	Hz
g : Aceleração da gravidade	m s^{-2}
H : Altura de onda	m
H_{m0} : Altura significativa espectral	m
H_{rms} : Altura quadrática média	m
H_s : Altura significativa	m
$\tilde{H}$: Variável adimensional da altura significativa	
K_r : Coeficiente de refração	
K_s : Coeficiente de empolamento	
k : Número de onda	rad m^{-1}
κ_b : Índice de rebentação	
λ : Comprimento de onda	m
λ_{∞} : Comprimento de onda em águas profundas	m
μ : Valor esperado	
ω : Frequência angular	rad s^{-1}
ρ : Densidade	kg m^{-3}
σ : Desvio padrão	
T : Período	s

T_p :	Período de pico	s
$\tilde{T}$	Variável adimensional do período de pico	
w_t :	Ruído branco	
$X \sim \mathcal{N}(\mu, \sigma^2)$:	Distribuição normal	
$\overline{H}$:	Altura média	m
$\overline{x}_i$:	Média dos valores observados de uma série temporal	
$\vec{\phi}$:	Vetor de parâmetros	
$\vec{w}$:	Vetor de pesos sinápticos numa rede neuronal	
$\vec{x}$:	Vetor de variáveis de entrada (previsores)	
$\vec{y}$:	Vetor de variáveis de saída (ou de resposta)	
x_i :	Valores observados de uma série temporal	
y_i :	Valores previstos de uma série temporal	

Introdução

Motivação e âmbito

O estado do mar condiciona a execução de várias atividades que se desenvolvem nos oceanos, mares e águas costeiras. A previsão das condições de agitação marítima é essencial em muitas dessas atividades, nomeadamente para a segurança da navegação, a atividade piscatória, o projeto, operação e manutenção de obras marítimas e portuárias, as atividades desportivas e recreativas, entre outras.

No âmbito da Marinha Portuguesa e da Autoridade Marítima Nacional, o estado do mar condiciona a execução das operações navais, principalmente dos meios de superfície e dos meios aéreos orgânicos das unidades navais, a realização de ações de fiscalização e de Busca e Salvamento (SAR), a execução de operações anfíbias, o desempenho das armas e sensores, etc..

O primeiro modelo de previsão das condições de agitação marítima foi desenvolvido por Sverdrup e Munk (Sverdrup e Munk [1947](#)), perante a necessidade de planear e apoiar a decisão de realizar operações anfíbias no Norte de África em 1942 durante a segunda Guerra Mundial (Pinardi et al. [2017](#)). O trabalho de Sverdrup e Munk manteve-se classificado devido à importância das grandes operações anfíbias para o desfecho da guerra, só tendo sido publicado em 1947 (Sverdrup e Munk [1947](#)).

Desde o trabalho pioneiro de Sverdrup e Munk, foram desenvolvidos métodos de previsão de agitação marítima sucessivamente mais aperfeiçoados. Os primeiros métodos, orientados para a utilização manual através de ábacos, fórmulas simples e regras empíricas, permitem prever parâmetros característicos da agitação marítima, tais como a altura significativa e o período de pico (cujas definições serão apresentadas mais adiante). Esses métodos continuam a ter valor e interesse prático, por serem uma forma de detetar possíveis enviesamentos dos modelos numéricos, com o objetivo de os ajustar posteriormente (World Meteorological Organization [2023](#)). Contudo, a evolução do conhecimento teórico e empírico sobre os processos de geração, propagação e transformação das ondas de vento, bem como o aumento

das capacidades dos meios de cálculo, permitiu o desenvolvimento de gerações sucessivas de modelos computacionais caracterizados por representações cada vez mais refinadas dos termos da equação que descreve a evolução do espectro de energia das ondas (cuja definição também será apresentada mais adiante) em grandes áreas oceânicas. O modelo *WaveWatch III*, operado no CGEOMETOC, é um modelo de terceira geração (Holthuijsen [2007]), que permite obter previsões das condições de agitação com grande fiabilidade nos espaços marítimos de interesse nacional. A operação contínua deste modelo requer meios computacionais de grande capacidade, bem como o acompanhamento constante por parte de uma equipa especializada.

Recentemente têm sido propostos novos métodos recorrendo a modelos estatísticos e de Aprendizagem Automática (*Machine Learning*). Estes modelos podem ser treinados quer a partir de informação recolhida por boias ondógrafo (Oleg Makarynsky [2005]; Oleg Makarynsky [2005]; Demetriou, Michailides, Papanastasiou e Onoufriou [2021]), quer a partir dos resultados (soluções numéricas) dos modelos computacionais (Ma et al. [2020]; O'Donncha, Y. Zhang, B. Chen e S. C. James [2018]), ou a partir da combinação entre ambos (Minuzzi e Farina [2023]; S. C. James, Y. Zhang e O'Donncha [2018]). Uma vez treinados de forma correta, estes modelos permitem obter previsões fiáveis numa fração do tempo necessário à obtenção de soluções numéricas, e com meios de cálculo muito mais modestos. Por este motivo, a utilização de modelos de aprendizagem estatística (*data driven*) na previsão de condições ambientais é uma área de investigação muito ativa (Singh, Chaturvedi e Akhter [2019]; R. Zhang, Z.-Y. Chen, Xu e Ou [2019]; Scher e Messori [2018]).

Objetivos e questões de investigação

O objetivo deste trabalho é o desenvolvimento de modelos de aprendizagem automática para a previsão das condições de agitação marítima, em particular a altura significativa (Holthuijsen [2007]) e a sua aplicação à área entre o cabo Espichel e o Cabo de Sines ilustrada na figura [1]. A escolha desta área foi motivada pelo interesse operacional da área de Pinheiro da Cruz, por ser uma área muito ativa na realização de exercícios militares de projeção de forças (operações anfíbias), na qual a batimetria é aproximadamente retilínea e uniforme, o que possibilita a utilização de modelos simplificados para prever as condições de agitação marítima perto de costa, a partir dos parâmetros característicos da agitação ao largo. Por outro lado, estão disponíveis séries de dados recolhidos por uma boia multiparamétrica fundeada nesta área (Figura [1]) (Instituto Hidrográfico [2015]). Essas séries são essenciais para o treino e validação dos modelos de aprendizagem automática.

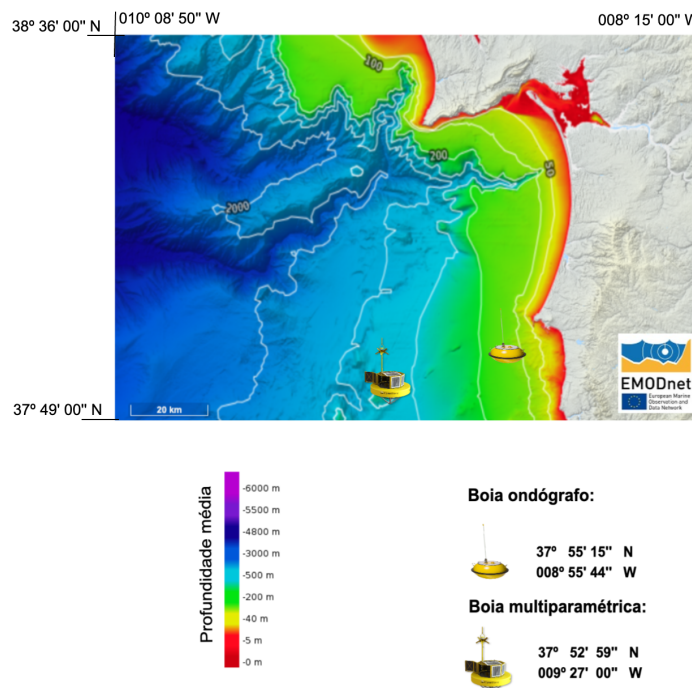


FIGURA 1: Representação da área de interesse para o presente trabalho: limites geográficos, batimetria, linha de costa, localização das boias ondógrafo e multiparamétrica de recolha de dados de agitação marítima operadas pelo IH. Adaptado de: EMODnet [2024](#)

Para além dos registos históricos das boias representadas na figura [1](#), o CGEOMETOC dispõe de registos dos resultados do modelo numérico *WaveWatch III* (WWIII), para toda a área do Atlântico Norte. Recentemente, Sousa [2023](#) apresentou uma descrição da climatologia de agitação marítima no Atlântico Norte, a partir da qual identificou e localizou zonas típicas de geração de agitação. Os resultados do modelo WWIII e o trabalho de Sousa [2023](#) serão utilizados no presente trabalho, por forma a determinar pontos do modelo cujo registos permitem treinar os modelos de aprendizagem automática de modo a deles obter previsões fiáveis das condições de agitação marítima.

No presente trabalho pretende-se analisar os resultados de previsão dos modelos de Aprendizagem Automática (*Machine Learning*) com os valores observados da boia de Sines. O conceito deste trabalho divide-se na exploração de três dimensões: a seleção da estrutura e hiperparâmetros da rede neuronal e na variação das fontes de dados. Quanto à variação das fontes de dados, há a considerar a escolha das variáveis que constituem os previsores mais importantes com base em explicações

fiáveis sobre os processos de geração e propagação das ondas. Por fim, a utilização de algoritmos com grau crescente de complexidade permite avaliar a qualidade relativa das previsões para cada parâmetro de agitação marítima a prever.

Assim, a questão de investigação central para o presente trabalho é a seguinte:

Os métodos de Aprendizagem Automática permitem prever de forma fiável as condições de agitação marítima num determinado ponto?

Da questão central, foram deduzidas as seguintes questões derivadas:

1. Como selecionar entre os dados disponíveis (bóia multiparamétrica e modelo numérico), as séries temporais que permitem e garantem o treino dos modelos e previsões fiáveis para os parâmetros de agitação marítima (altura significativa e período de pico) na área entre o cabo Espichel e o Cabo de Sines?
2. Como selecionar as variáveis que permitem treinar, testar e validar modelos estatísticos de Aprendizagem Automática com base no conhecimento geral sobre os processos de geração, propagação e transformação de ondas de vento?
3. Que tipos, estrutura e modelos garantem maior precisão na previsão da agitação marítima na área em estudo?
4. Que modelos e estratégia de exploração de hiperparâmetros permitem a melhoria da qualidade dos modelos de previsão?

Capítulo 1

Revisão de literatura sobre a dinâmica das ondas de vento

1.1 Introdução

O oceano é um sistema dinâmico influenciado pelas suas características físicas (morfologia do fundo, temperatura, densidade e salinidade da água do mar) e pelas forças externas que nele atuam (forças de volume, como a força da gravidade, a força geradora de maré e a força de Coriolis, assim como por forças de superfície como a tensão do vento, as flutuações da pressão atmosférica e a força da tensão superficial) (World Meteorological Organization [2023](#)). Nos oceanos existem três tipos de movimento: as correntes, as ondas e a turbulência.

As ondas de superfície, que se propagam na interface ar-água, manifestam-se pela variação da elevação da superfície do oceano, e propagam-se com periodicidades espacial e temporal, com comprimento de onda e período, dependentes das forças de perturbação e de restituição. As ondas podem ser classificadas pelo seu período, descrevendo assim um sistema de ondas.

A figura [1.1](#) ilustra a classificação dos sistemas de ondas de superfície em função do período, indicando também as forças perturbadoras e de restituição respectivas. Esta figura representa de forma qualitativa a distribuição de densidade de energia no domínio da frequência (que é o inverso do período), e mostra que as ondas de vento concentram a maior fração de energia das ondas de superfície. Como veremos mais adiante, o conhecimento do espectro de energia é fundamental na descrição matemática da dinâmica da agitação marítima.

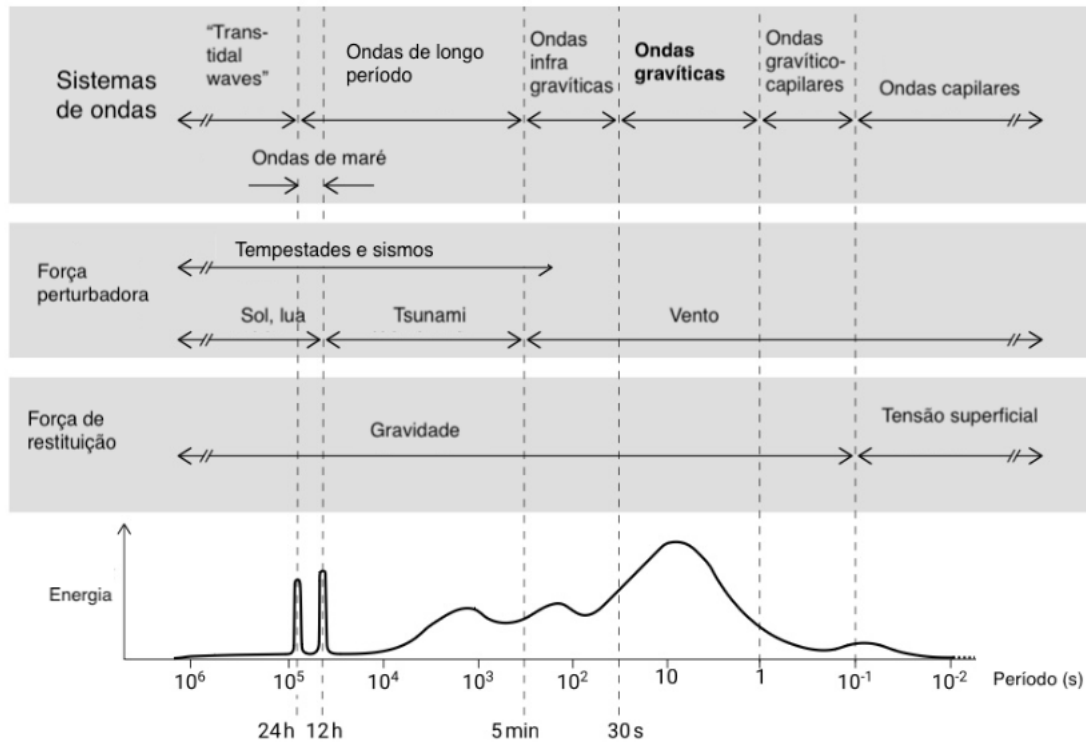


FIGURA 1.1: Distribuição da densidade de energia das ondas de superfície em função do período e classificação dos diversos sistemas de ondas em termos das forças perturbadoras e de restituição. Adaptado de Pope e Lockhart [2002](#)

O movimento irregular da agitação marítima pode ser descrito como a soma de infinitas ondas de amplitude pequena, com períodos, fases e direções de propagação diferentes. Desta forma, para descrever a dinâmica da agitação irregular é necessário analisar primeiro a dinâmica de uma onda simples monocromática, com período/ frequência únicos, propagando-se numa só direção, com cristas paralelas e uniformes (World Meteorological Organization [2023](#)). As teorias de onda desenvolvidas no âmbito da mecânica de fluidos teórica permitem descrever ondas com estas características. A mais simples destas teorias é a teoria linear, na qual se admite que a altura de onda é muito inferior ao comprimento de onda e à profundidade. Admite-se também que a superfície livre tem uma forma sinusoidal. A partir dos resultados das teorias de onda e do conhecimento das características de agitação irregular, em particular do espectro direcional (distribuições de densidade de energia em função da frequência e da direção de propagação) e dos processos de geração, propagação e transformação das ondas, é possível obter descrições matemáticas do estado do mar

e formular modelos para a sua previsão. Assim, este capítulo encontra-se dividido em três secções.

Na primeira secção são descritos os resultados mais importantes da teoria linear das ondas progressivas. Na segunda secção é abordada a descrição estatística do estado do mar, por forma a estabelecer parâmetros (altura significativa, altura máxima, período, período de pico, etc..) que descrevem ondas reais correspondentes aos parâmetros utilizados para definir ondas simples monocromáticas. Será também apresentado o conceito de espectro de energia ou variância. Este conceito é essencial porque permite estabelecer uma relação entre os resultados da teoria linear, os parâmetros característicos da agitação marítima e posteriormente com os fenómenos de geração, propagação e transformação de ondas (também abordados na primeira secção do capítulo).

1.2 Dinâmica das ondas de vento

1.2.1 Generalidades

O movimento da atmosfera e do oceano é descrito por três princípios de conservação (massa, quantidade de movimento e energia) e ainda pela equação de estado, que difere nos termos que por esta são considerados quando é calculada para a atmosfera ou para o oceano. Os princípios de conservação são descritos pelas equações da continuidade, da quantidade de movimento e de energia, respetivamente. As equações do movimento são não lineares, e não são conhecidas soluções analíticas. Por outro lado, os diversos fenómenos físicos na atmosfera e nos oceanos ocorrem em escalas de tempo muito distintas. Deste modo, para descrever matematicamente um determinado fenómeno, neste caso as ondas de vento, é necessário efetuar uma análise de escalas dos termos das equações do movimento, e introduzir hipóteses simplificadoras (Pope e Lockhart 2002).

No desenvolvimento da teoria linear, são assumidas as simplificações seguintes:

- Os efeitos da tensão superficial e o efeito de Coriolis podem ser desprezados
- A pressão na superfície livre é constante e uniforme
- As cristas das ondas são longas e paralelas (o movimento é bi-dimensional no plano xz)

As características físicas da água possibilitam que se assumam as seguintes simplificações (Stewart 2008):

- O fluido é incompressível;
- O fluido é de densidade constante;
- O fluido é ideal (não apresenta viscosidade);

A última simplificação mencionada é válida exceto nas camadas-limite junto à superfície (e ao fundo) e após a rebentação, que é caracterizada pela produção de turbulência e consequentemente vorticidade e dissipação de energia (Holthuijsen 2007).

Além disso, pode ainda considerar-se que:

- O escoamento é irrotacional;

Fora da zona de rebentação o movimento é quase irrotacional, excepto nas camadas-limite, isto é, camadas muito finas juntas à superfície ou ao fundo do oceano. Por isso, para o estudo das teorias de onda, considera-se que em águas profundas, a profundidade não afeta a estrutura do movimento e também se admite que o escoamento é irrotacional.

Ainda que a agitação marítima tenha um padrão irregular, pode ser descrita como a sobreposição de várias ondas sinusoidais de amplitude pequena, com altura, período, fase e direção de propagação diferentes, como é graficamente representado na figura 1.2 (World Meteorological Organization 2023).

Na secção seguinte, é apresentado um resumo dos resultados mais importantes da teoria linear, que descreve a dinâmica de ondas monocromáticas, de cristas paralelas e amplitude (ou altura) pequena (muito menor do que o comprimento de onda e a profundidade).

1.2.2 Teoria linear das ondas de superfície

A teoria linear das ondas de superfície admite, para além das hipóteses anteriormente mencionadas, que (Pope e Lockhart 2002):

- A superfície livre tem a forma de uma sinusóide;
- A amplitude da onda é muito menor do que o comprimento de onda e a profundidade ($\frac{H}{D} \ll 1$, $\frac{H}{\lambda} \ll 1$).

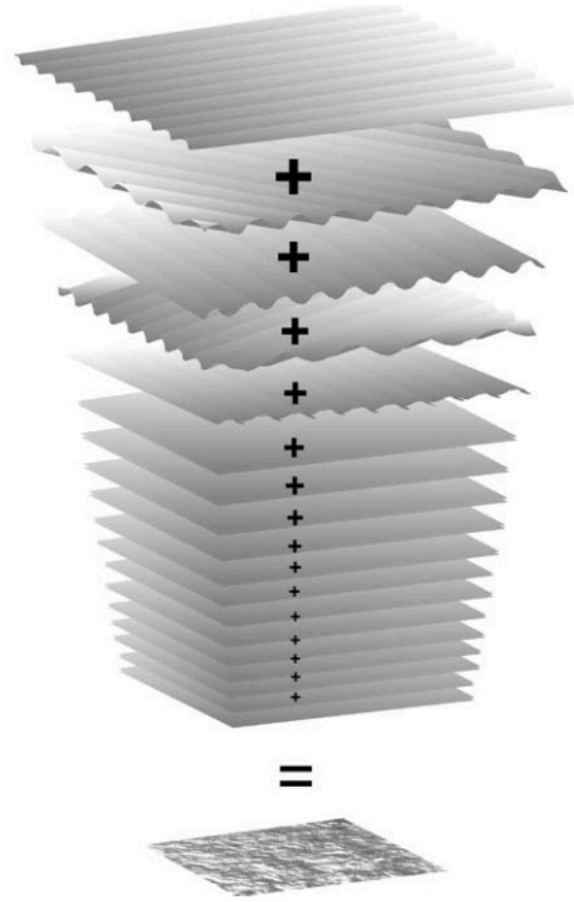


FIGURA 1.2: Representação da agitação irregular como a soma de infinitas ondas sinusoidais com altura, período, fase e direção de propagação distintas. Fonte : Holthuijsen [2007](#)

Admitindo que as ondas sinusoidais se propagam na direção do eixo dos xx (figura [1.3](#)), têm periodicidade espacial (λ) e temporal (T), então estas são descritas pelos parâmetros seguintes (World Meteorological Organization [2023](#)):

- Amplitude da onda, a : Diferença entre a elevação máxima e o nível de repouso;
- Altura de onda, H : Diferença entre a elevação das cristas e das cavas. Para uma onda sinusoidal (figura [1.3](#)) : $H = 2a$;
- Comprimento de onda, λ : Distância entre duas cristas/cavas sucessivas;
- Período, T : Intervalo de tempo entre a passagem de duas cristas/cavas sucessivas num ponto;
- Profundidade, D : Distância entre o nível médio e o fundo;

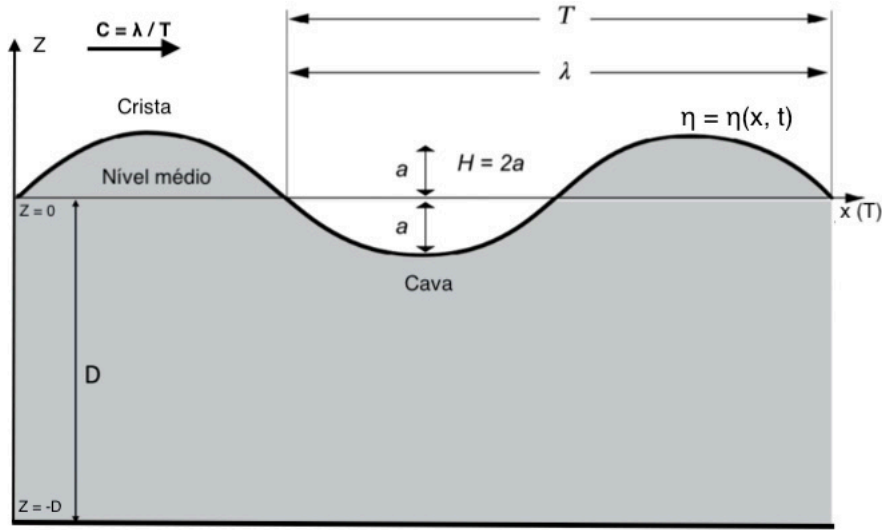


FIGURA 1.3: Parâmetros característicos de uma onda simples (Adaptado :Sousa [2023](#)).

- Velocidade de propagação, C : Quociente entre o comprimento de onda e o período;

$$C = \frac{\lambda}{T} \quad (1.1)$$

- Elevação da superfície, η : deslocamento vertical da superfície, relativamente ao nível de repouso, num determinado instante, podendo por isso ter valores positivos ou negativos. Admitindo, sem perda de generalidade, que no instante $t = 0$ a superfície livre tem a sua configuração indicada na figura [1.3](#), a superfície livre é descrita pela fórmula seguinte:

$$\eta(x, t) = a \sin(\omega t - kx) \quad (1.2)$$

Consideram-se ainda dois parâmetros úteis para o desenvolvimento analítico das teorias de onda:

- Frequência angular, ω :

$$\omega = \frac{2\pi}{T} \quad (1.3)$$

- Número de onda, κ :

$$k = \frac{2\pi}{\lambda} \quad (1.4)$$

A relação de dispersão permite relacionar o comprimento de onda (ou o número de onda k), com o período (ou a frequência angular ω) e a profundidade D .

Esta relação obtém-se a partir de duas condições de fronteira da superfície livre: a pressão atmosférica deve ser constante (condição dinâmica), e as partículas que se encontram na superfície livre devem permanecer na mesma (condição cinemática). A forma geral da equação da dispersão é a seguinte:

$$\omega^2 = gk \tanh(kD) \quad (1.5)$$

Esta relação pode ainda ser expressa nas formas equivalentes:

$$\lambda = \frac{gT^2}{2\pi} \tanh(kD) \quad (1.6)$$

$$C = \frac{gT}{2\pi} \tanh(kD) \quad (1.7)$$

A relação de dispersão é uma equação não-linear, e por isso só pode ser resolvida com recurso a métodos aproximados, como por exemplo, a aproximação de Eckard (1952), a aproximação de Beji (2013) ou ainda através de métodos iterativos. A forma da equação da dispersão mostra que existem três condições de propagação, entre as quais duas condições limite ($kD \rightarrow \infty$ e $kD \rightarrow 0$):

A condição de águas profundas corresponde ao limite $kD \rightarrow \infty$. Nesta condição, o comprimento de onda e velocidade de propagação são expressos respetivamente por:

$$\lambda_\infty = \frac{gT^2}{2\pi} \quad (1.8)$$

$$C_\infty = \frac{gT}{2\pi} \quad (1.9)$$

Portanto, na condição de águas profundas, o comprimento e a velocidade de propagação não dependem da profundidade D .

A condição de fundos baixos corresponde ao limite $kD \rightarrow 0$. Nesta condição, é válida a aproximação $\tanh(kD) \approx kD$. Na condição de fundos baixos, a velocidade de propagação não depende do período e é determinada pela profundidade. Neste caso, o comprimento de onda e a velocidade de propagação das ondas são dados pelas equações seguintes:

$$\lambda = T\sqrt{gD} \quad (1.10)$$

$$C = \sqrt{gD} \quad (1.11)$$

No caso geral, que se designa por condição de profundidade intermédia, o

comprimento de onda e velocidade de propagação dependem do período (T) e da profundidade (D) de acordo com as equações 1.6 e 1.7 ambas equivalentes à equação 1.5

As três condições de propagação acima descritas: águas profundas, profundidade intermédia e fundos baixos estão relacionados com a influência do fundo sobre a estrutura do movimento, em particular com a forma das órbitas das partículas. É usual estabelecer a condição de propagação de acordo com o critério seguinte:

- Águas Profundas: $\frac{D}{\lambda} \geq \frac{1}{2}$
- Profundidade Intermédia: $\frac{1}{25} \leq \frac{D}{\lambda} < \frac{1}{2}$
- Fundos Baixos: $\frac{1}{25} < \frac{D}{\lambda}$

Na condição de águas profundas, as órbitas das partículas são circunferências, pois o fundo não afeta a estrutura do movimento. O raio das trajetórias decresce exponencialmente ao longo da profundidade, como se demonstra analiticamente em Holthuijsen 2007 e é ilustrado graficamente no topo da figura 1.4.

No caso da condição de profundidade intermédia, as partículas descrevem órbitas elípticas, que no fundo degeneram em segmentos de recta horizontais conforme é representado no esquema do meio da figura 1.4.

Na condição de fundos baixos, tal como é apresentado no último exemplo da figura 1.4, as partículas também se movem em elipses. Porém, nesta condição a escala horizontal é muito superior à escala vertical, o que faz com que a componente vertical da velocidade seja muito inferior à componente horizontal da mesma e que esta seja constante em profundidade, ou seja, independente de z .

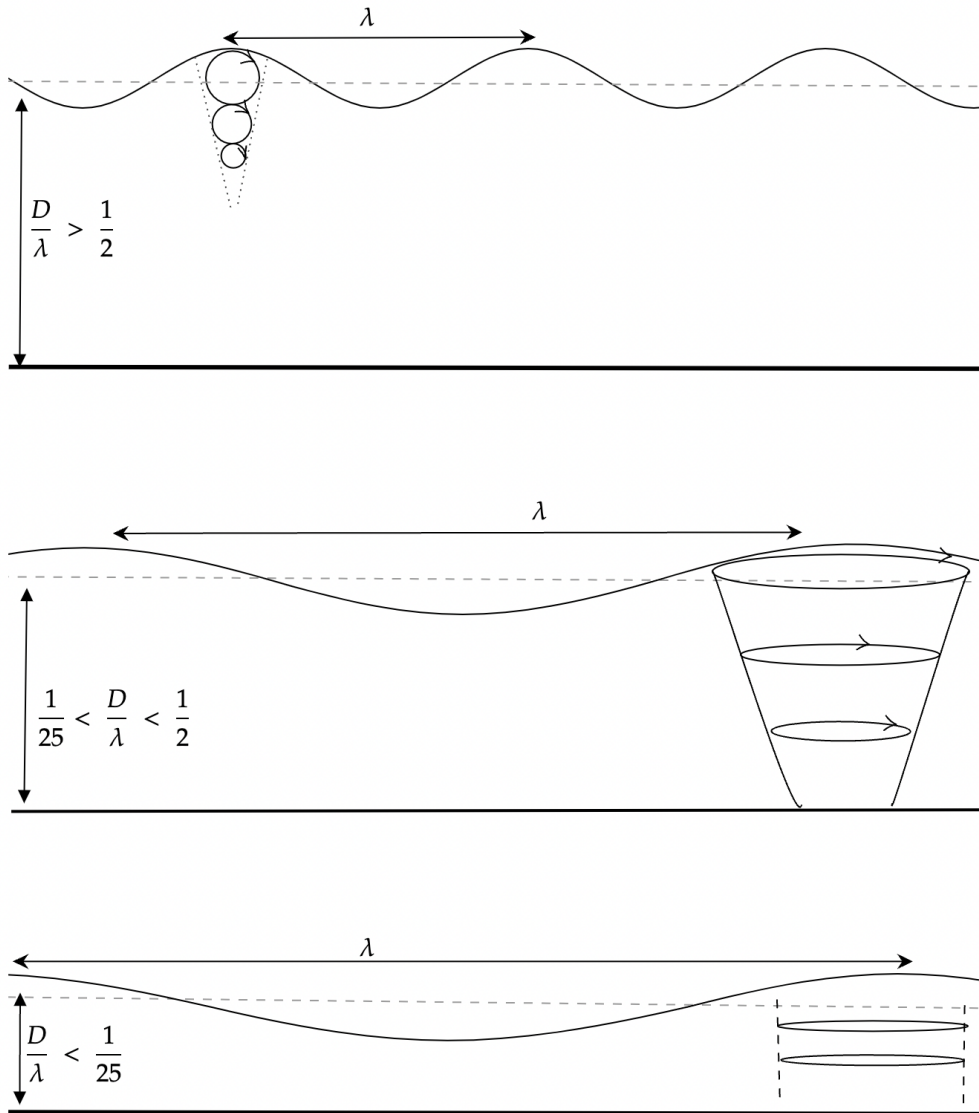


FIGURA 1.4: Representação das condições de propagação definidas pela teoria linear das ondas progressivas: águas profundas (em cima), profundidade intermédia (a meio), fundos baixos (em baixo)

Nas condições de águas profundas e de profundidade intermédia, a velocidade de propagação (C) depende do período (T). Nestas condições, as ondas dizem-se dispersivas, ou seja, ondas com períodos diferentes propagam-se com velocidades diferentes e a sua sobreposição forma grupos que se propagam com uma velocidade diferente das ondas individuais que os compõem. Esta velocidade é denominada velocidade de grupo.

É possível demonstrar que a energia das ondas se propaga com a velocidade de grupo, a qual tem a expressão (Stewart 2008):

$$C_g \equiv \frac{\partial \omega}{\partial k} \quad (1.12)$$

No caso das ondas de superfície, a velocidade de grupo deduz-se derivando a relação da dispersão, obtendo-se:

$$C_g = \frac{C}{2} \left(1 + \frac{2kD}{\sinh 2kD} \right) \quad (1.13)$$

Nas condições de águas profundas e de profundidade intermédia, a velocidade de grupo é inferior à velocidade de propagação das ondas. Nesta situação, existem ondas que entram e outras que saem do grupo, pelo que um grupo é composto por ondas diferentes em instantes diferentes (Holthuijsen 2007).

Na condição de águas profundas ($kD \rightarrow \infty$), a equação da velocidade de grupo pode ser simplificada ($\lim_{kD \rightarrow \infty} \frac{2kD}{\sinh(2kD)} = 0$), resultando na seguinte expressão:

$$(C_g)_\infty = \frac{1}{2}C = \frac{1}{2}C_\infty \quad (1.14)$$

Para a condição de fundos baixos ($kD \rightarrow 0$), $\lim_{kD \rightarrow 0} \frac{2kD}{\sinh(2kD)} = 1$, as ondas deixam de ser dispersivas, ou seja, a velocidade de propagação das ondas não depende do período (Holthuijsen 2007):

$$C_g = C \approx \sqrt{gD} \quad (1.15)$$

Conforme será abordado no capítulo 3, o cálculo da velocidade de grupo para certos períodos de onda, permitirá definir a localização dos pontos da malha geográfica utilizada no modelo *WaveWatch III* cujos registos serão usados para auxiliar o treino dos modelos de Aprendizagem Automática no presente trabalho.

A tabela seguinte resume as equações resultantes da teoria linear das ondas progressivas, para as diferentes condições de propagação:

	Águas profundas	Águas intermédias	Fundos baixos
Profundidade	$\frac{D}{\lambda} > \frac{1}{2}$	$\frac{1}{25} \leq \frac{D}{\lambda} < \frac{1}{2}$	$\frac{1}{25} < \frac{D}{\lambda}$
Comprimento de onda	$\lambda_{\infty} = \frac{gT^2}{2\pi}$	$\lambda = \frac{gT^2}{2\pi} \tanh(kD)$	$\lambda_{\infty} = T\sqrt{gD}$
Velocidade de propagação	$C_{\infty} = \frac{gT}{2\pi}$	$C = \sqrt{\frac{g}{k} \tanh(kD)}$	$C = \sqrt{gD}$
Velocidade de grupo	$C_g = \frac{1}{2}C$	$C_g = \frac{C}{2} \left(1 + \frac{2kD}{\sinh(2kD)}\right)$	$C_g = C$

TABELA 1.1: Descrição das ondas em diferentes profundidades

Uma propriedade importante das ondas de superfície é a densidade de energia mecânica, ou energia mecânica por unidade de área. A densidade de energia mecânica tem duas componentes, potencial e cinética, que no caso das ondas sinusoidais ou lineares são iguais.

A densidade de energia das ondas deduz-se integrando as energias potencial e cinética num comprimento de onda e para uma largura de crista arbitrária b , dividindo cada uma delas pela área $\lambda * b$ e somando no final os dois resultados obtidos (figura 1.5) (Holthuijsen 2007).

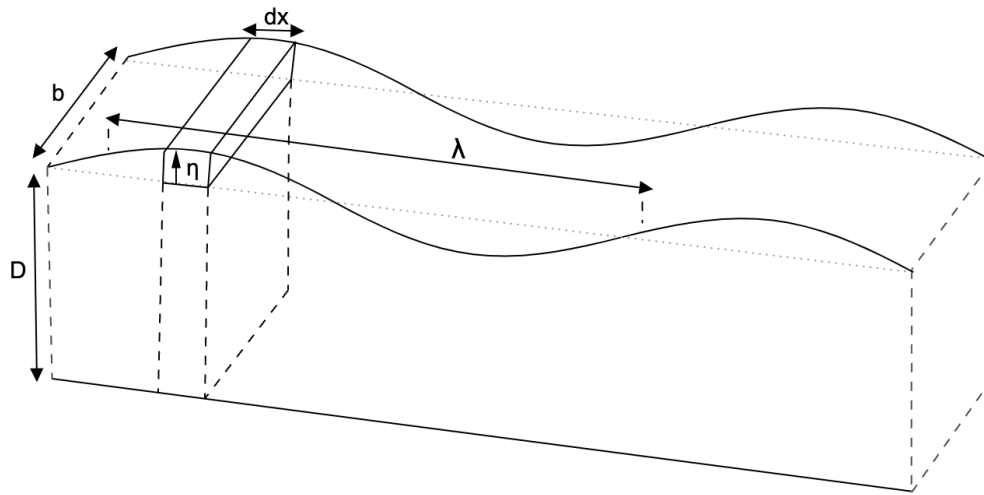


FIGURA 1.5: Representação de uma onda sinusoidal e da variação de energia potencial

Assim, a densidade de energia potencial é expressa da seguinte forma:

$$E_{potencial} = \frac{1}{4}\rho ga^2 \quad (1.16)$$

A equação que descreve a densidade de energia cinética é a seguinte:

$$E_{cinética} = \frac{1}{4}\rho ga^2 \quad (1.17)$$

pelo que a densidade de energia (E) nessa superfície é igual a:

$$E = E_{potencial} + E_{cinética} \quad (1.18)$$

$$E = \frac{1}{2}\rho ga^2 = \frac{1}{8}\rho gH^2 \quad (1.19)$$

O fluxo de energia é a grandeza associada ao transporte da energia por unidade de tempo e por largura de crista na direção perpendicular à direção de propagação. Na teoria linear, o fluxo de energia (P) associado à propagação das ondas é expresso pela equação (Holthuijsen [2007](#)):

$$P = EC_g \quad (1.20)$$

1.2.3 Descrição estatística da agitação marítima

Como referido anteriormente, a agitação marítima é um fenómeno irregular, mas pode ser caracterizada pela soma de um número infinito de ondas sinusoidais, propagando-se independentemente (World Meteorological Organization [2023](#)). A teoria linear permite obter relações teóricas importantes para a descrição da dinâmica das ondas de superfície. No caso da agitação irregular é necessário estabelecer primeiro um critério de definição de ondas individuais num registo de ondas e posteriormente os parâmetros estatísticos que correspondem aos parâmetros usados para descrever ondas regulares, como o período e a altura de onda.

Descrição no domínio do tempo

Uma das formas de medir a agitação marítima é a partir da utilização de boias ondógrafo. Uma boia ondógrafo é uma boia fundeada capaz de seguir livremente o movimento da superfície do oceano, dotada de um ou mais acelerómetros e um *link* rádio que permite transmitir dados para uma estação costeira. A integração

do sinal do acelerômetro permite obter o deslocamento vertical instantâneo (Soares, Rusu, A. A. P. Silva e Oleg Makarynskyy 2000). A figura 1.6 ilustra um registro típico da elevação da superfície livre obtido através de uma bóia ondógrafo.

A partir da análise da figura 1.6, reconhece-se a dificuldade em identificar ondas. Por essa razão, estabeleceu-se um critério de classificação de ondas: o critério dos zeros sucessivos.

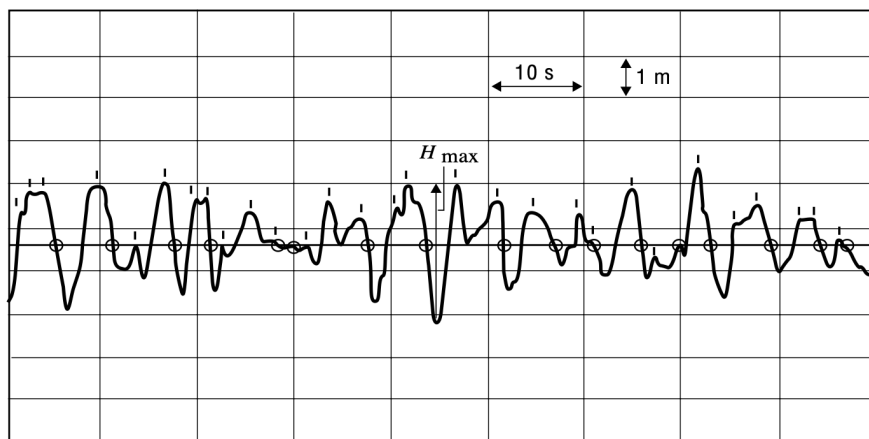


FIGURA 1.6: Ilustração de um registro da elevação de superfície (fonte :World Meteorological Organization 2023)

No critério dos zeros sucessivos, uma onda é definida entre dois instantes em que a elevação da superfície intersecta o nível médio do registro (zeros). Contudo, a contabilização dos zeros deve seguir, do princípio ao fim, apenas um de dois critérios que existem: o critério dos zeros descendentes ou o critério dos zeros ascendentes.

O critério dos zeros descendentes é a contagem de dois instantes sucessivos que a elevação superfície livre cruza o nível médio do registro ao alterar de um valor positivo para um valor negativo, conforme é representado no primeiro exemplo da figura 1.7. O inverso acontece no critério dos zeros ascendentes, ou seja, um zero ascendente é definido pelo instante em que a elevação da superfície cruza o nível médio do registro ao alterar de um valor negativo para um valor positivo conforme o segundo exemplo da figura 1.7.

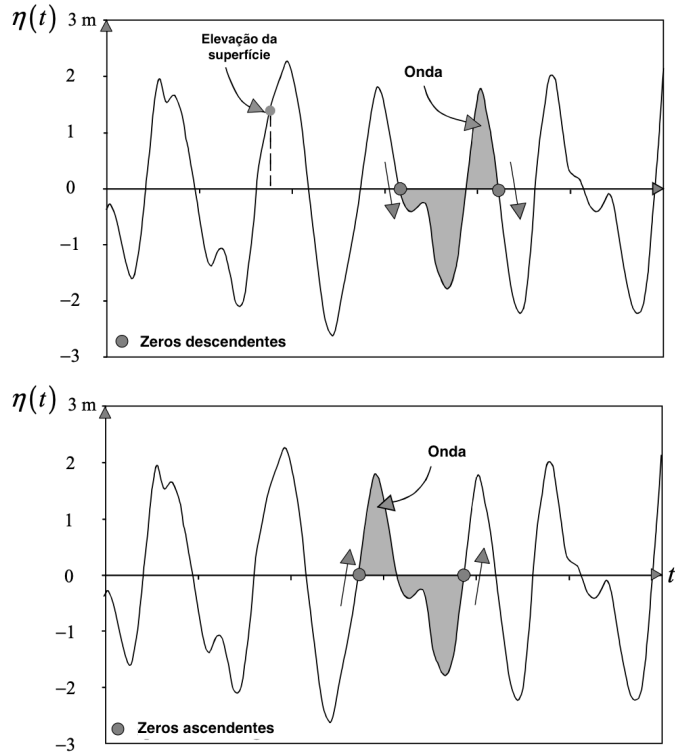


FIGURA 1.7: Representação dos dois critérios de classificação de ondas: zeros descendentes e zeros ascendentes. (Adaptado de: Holthuijsen 2007)

Parâmetros característicos

Embora as ondas num registo de agitação sejam irregulares, é possível definir, a partir de um registo, parâmetros estatísticos que tenham correspondência com os definidos nas teorias de onda, nomeadamente o período e a altura de onda (Pope e Lockhart 2002; Holthuijsen 2007).

Os parâmetros estatísticos mais importantes relacionados com a altura e o período definidos a partir da análise de registos no tempo são os seguintes:

- Altura média

$$\bar{H} = \frac{1}{N} \sum_{i=1}^N H_i \quad (1.21)$$

- Altura quadrática média (*root mean square*)

$$H_{\text{rms}} = \sqrt{\sum_{i=1}^N \frac{H_i^2}{N}} \quad (1.22)$$

- Altura significativa

$$H_s = H_{\frac{1}{3}} = \frac{1}{\frac{N}{3}} \sum_{j=1}^{\frac{N}{3}} H_j \quad (1.23)$$

Na equação [1.23](#), o índice j corresponde a uma ordenação decrescente das alturas de onda registadas. A altura significativa, que é definida como a média de $\frac{1}{3}$ das ondas mais altas num registo, é a que melhor aproxima as estimativas visuais (World Meteorological Organization [2023](#)).

- Altura máxima

$H_{\max}$ corresponde à altura máxima no intervalo considerado.

- Período médio

$$\bar{T} = \frac{1}{N} \sum_{i=1}^N T_i \quad (1.24)$$

- Período significativo

$$T_s = T_{\frac{1}{3}} = \frac{1}{\frac{N}{3}} \sum_{j=1}^{\frac{N}{3}} T_j \quad (1.25)$$

O período significativo é o período médio de $\frac{1}{3}$ das ondas mais altas de um registo. Nestas expressões, o índice i corresponde à numeração da onda conforme o registo e o índice j refere-se à ordenação decrescente das alturas de ondas registada.

Distribuição de Rayleigh

O matemático e oceanógrafo Longuet-Higgins ([1957](#)) provou que os registos da elevação da superfície livre do oceano seguem uma distribuição normal, e que os valores máximos da altura de onda seguem a distribuição de Rayleigh. A figura [1.8](#) representa a função de densidade de probabilidade associada à distribuição de Rayleigh, definida pela seguinte expressão:

$$f(H) = 2 \frac{H}{H_{rms}^2} e^{-\left(\frac{H}{H_{rms}}\right)^2} \quad (1.26)$$

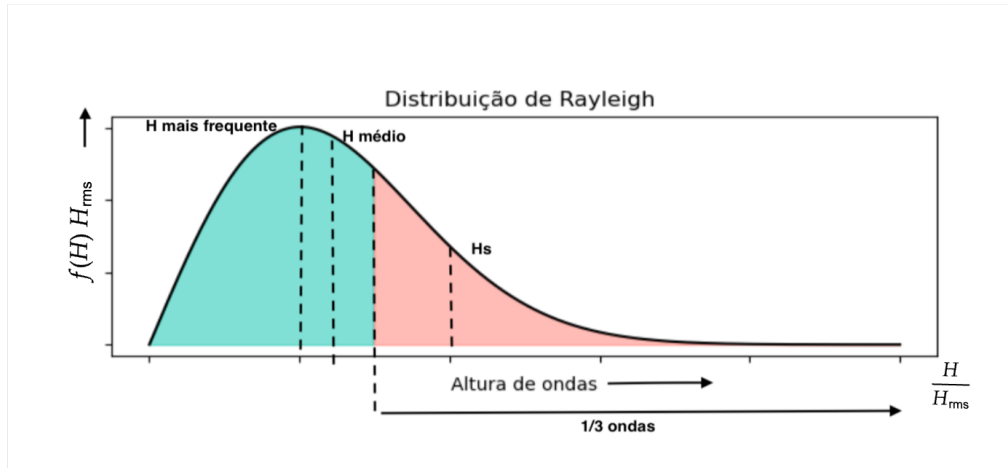


FIGURA 1.8: Representação da função de densidade de probabilidade associada à distribuição de Rayleigh

A distribuição de Rayleigh permite estabelecer as seguintes relações entre parâmetros:

$$H_{RMS} \approx \sqrt{2} H_{\text{mais provável}} \quad (1.27)$$

$$H_s = H_{\frac{1}{3}} \approx \sqrt{2} H_{RMS} \quad (1.28)$$

$$H_{\frac{1}{10}} \approx 1,27 H_s \quad (1.29)$$

$$H_{\text{max}} \approx 2,63 H_{RMS} \quad (1.30)$$

Espectro de densidade de energia e o espectro de densidade de variância

O objetivo de descrever as ondas com um espectro, ou através de parâmetros estatísticos característicos, não é descrever detalhadamente uma observação específica da elevação da superfície do oceano (um único registo de uma série temporal), mas sim descrever a superfície do mar como um processo estocástico¹, ou seja, caracterizar todas as observações possíveis de ocorrer sob as mesmas condições da observação ocorrida (amostra do registo temporal) (Holthuijsen [2007](#)).

Assumindo que um registo de elevação instantânea é um processo estacionário, ou seja as suas características estatísticas são independentes da variável tempo (conceito abordado no capítulo seguinte), então pode ser decomposto numa soma de um elevado número de funções sinusoidais, cada com amplitude dependente da

¹Um processo estocástico é a coleção de variáveis aleatórias organizadas temporalmente que podem tomar valores contínuos ou discretos. Em séries temporais estocásticas, os valores futuros não podem ser previstos com exatidão absoluta. No entanto, esses valores poderão ser estimados sobre a probabilidade da sua ocorrência, que é limitada pelo conhecimento de valores passados da própria série (Chatfield e Xing [2019](#)).

frequência e da fase aleatória (análise de Fourier), obtendo-se a representação do espectro de amplitude (primeiro exemplo da figura 1.9). Contudo, para que o registo da elevação da superfície seja completamente restituído na sua representação em função da frequência, é necessário que o critério de Nyquist seja respeitado. Este critério estabelece que a frequência da amostra do sinal deve ser pelo menos o dobro da frequência do sinal original, neste caso o registo, por forma a garantir que não existe perda de informação (Holthuijsen 2007).

Associado ao espectro de amplitude está o espectro de variância. Por definição, a variância da elevação da superfície é a média quadrática da elevação de superfície. No caso de uma onda sinusoidal com a amplitude a , a variância é:

$$Var(\eta(t)) = \frac{a^2}{2} \quad (1.31)$$

Demonstra-se que a variância da sobreposição de N ondas sinusoidais é igual à soma das variâncias de N ondas (segundo o exemplo da figura 1.9):

$$\text{Variância total} = \sum_{i=1}^N \frac{1}{2} a_i^2 \quad (1.32)$$

Na agitação irregular, o mínimo de ondas consideradas é tão grande que pode ser tomado em termos teóricos com infinito. Neste caso, a amplitude de uma onda com frequência f_i tende para zero, mas é possível redefinir uma densidade de variância, e associar uma variância a um intervalo Δf_i que contém f_i , conforme expresso matematicamente (e exemplificado no terceiro exemplo da figura 1.9):

$$E(f_i) = \frac{1}{\Delta f_i} E \left\{ \frac{1}{2} a_i^2 \right\} \quad (1.33)$$

Este espectro é particularmente importante para descrever a densidade de energia das ondas, obtido através da relação:

$$E = \rho g E_{\text{variância}} \quad (1.34)$$

Como na agitação irregular se assume que a densidade de energia das ondas varia de forma contínua com a frequência, a equação 1.33 permite definir o espectro de variância, conforme ilustrado no último exemplo da figura 1.9 (Holthuijsen 2007).

Obtém-se o espectro contínuo de densidade de variância a partir da seguinte relação:

$$E(f) = \lim_{\Delta f \rightarrow 0} \frac{1}{\Delta f} E \left\{ \frac{1}{2} a_f^2 \right\} \quad (1.35)$$

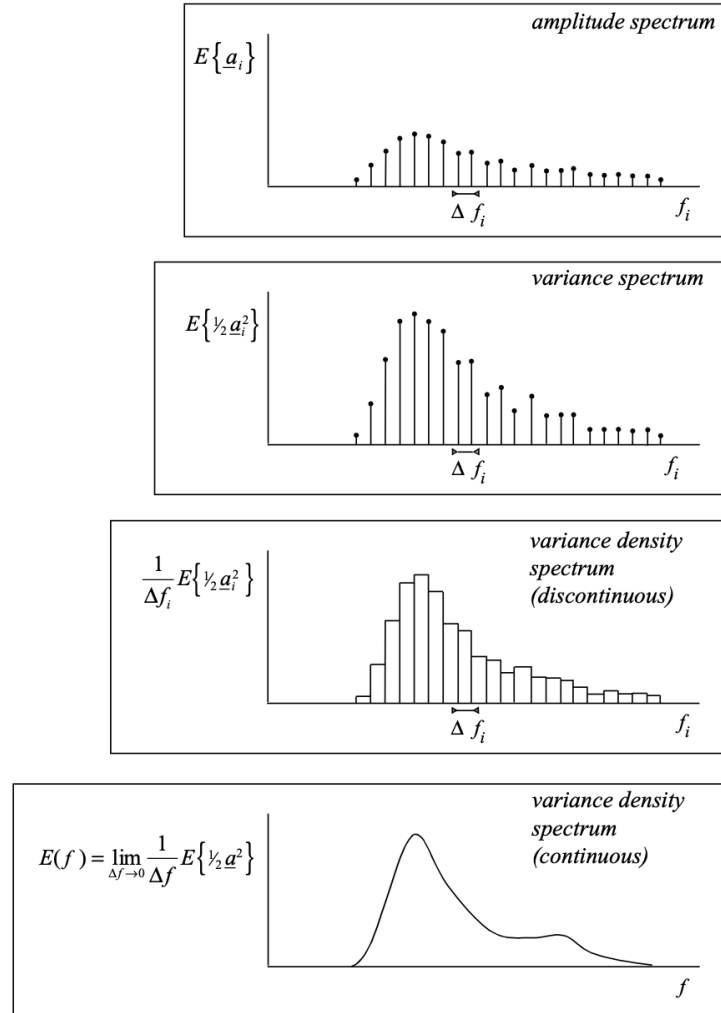


FIGURA 1.9: Representação dos espectros de amplitude, variância, densidade de variância contínuo e descontínuo. Fonte : Holthuijsen

2007

A partir do espectro de variância, é possível deduzir um conjunto de parâmetros característicos da agitação marítima, denominados parâmetros espectrais. Esses parâmetros são descritos em termos dos momentos do espectro de variância, são estes:

- Momento de ordem n

$$m_n = \int_0^{+\infty} f^n E(f) df \quad (1.36)$$

- Momento de ordem zero: corresponde ao valor da variância total

$$m_0 = \int_0^{+\infty} f^0 E(f) df \quad (1.37)$$

- Momento de primeira ordem:

$$m_1 = \int_0^{+\infty} f^1 E(f) df \quad (1.38)$$

- Momento de segunda ordem:

$$m_2 = \int_0^{+\infty} f^2 E(f) df \quad (1.39)$$

Os parâmetros espectrais mais importantes são os seguintes:

- A frequência de pico, f_p , corresponde à frequência com o valor máximo no espectro de variância
- O período de pico, T_p

$$T_p = \frac{1}{f_p} \quad (1.40)$$

- Altura significativa espectral, H_{m0} :

$$H_{m0} = 4\sqrt{m_0} \approx 1.05H_{\frac{1}{3}} \quad (1.41)$$

- Período da onda, T_{m01} correspondente à frequência média do espectro

$$T_{m01} = \frac{m_0}{m_1} \quad (1.42)$$

- Período da onda, T_{m02} correspondente ao segundo momento, teoricamente equivalente ao período médio

$$T_{m02} = \sqrt{\frac{m_0}{m_1}} \approx T_z \quad (1.43)$$

Alguns parâmetros espectrais, como por exemplo H_{m0} e T_{m02} , têm correspondência com os parâmetros característicos obtidos a partir da análise no tempo dos registos de agitação marítima (secção [1.2.3](#)).

Contudo, para descrever da maneira mais completa a agitação marítima, é necessário considerar a distribuição da variância não só em frequência, como também em direção (θ). Ao espectro de variância $E(f, \theta)$ que descreve a distribuição da variância em função da frequência e da direção dá-se o nome de espectro direcional, e é definido da seguinte forma (Holthuijsen 2007):

$$E(f, \theta) = \lim_{\Delta f \rightarrow 0} \lim_{\Delta \theta \rightarrow 0} \frac{1}{\Delta f \Delta \theta} E \left\{ \frac{1}{2} a^2 \right\} \quad (1.44)$$

O espectro direcional é importante porque descreve a direção dominante de agitação, a variação da densidade de energia em torno dessa direção e permite distinguir a vaga da ondulação. Consideremos um exemplo do espectro direcional obtido na costa holandesa, resultante da propagação de ondulação de uma tempestade gerada pelo Mar da Noruega e por vaga gerada pelo vento local, conforme a figura 1.10.

A ondulação geralmente é mais regular, e por isso tem um espectro mais estreito, e tem frequências mais baixas do que a vaga que é mais irregular (espectro largo), pelo que neste caso os dois sistemas de onda são facilmente distinguíveis no espectro direcional (e neste caso no espectro de variância também) (Holthuijsen 2007).

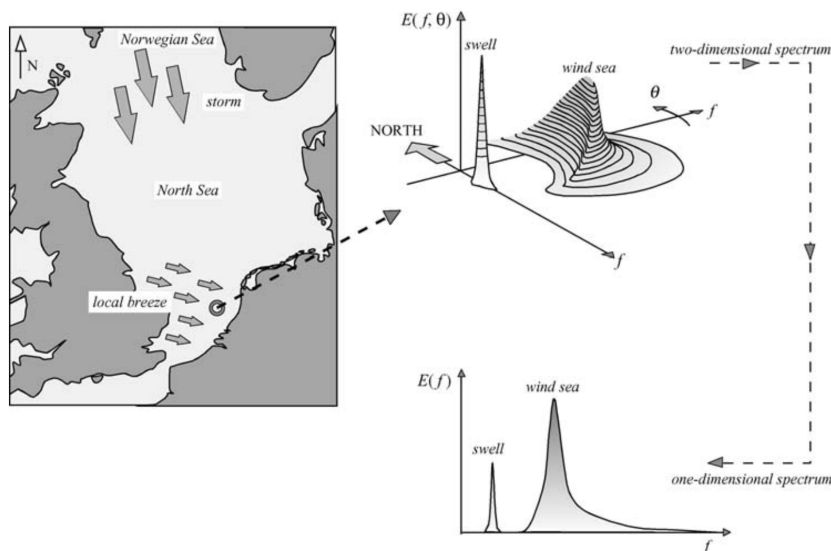


FIGURA 1.10: Exemplo de um espectro direcional na costa holandesa, resultante de ondulação gerada por uma tempestade gerada pelo mar da Noruega e vaga gerada localmente. Fonte: Holthuijsen 2007

1.2.4 Geração, propagação e transformação das ondas de vento

A primeira teoria com o propósito de explicar o processo de geração de ondas de vento foi desenvolvida por Kelvin e Helmholtz, que sugeriam que o processo de geração se iniciava quando a intensidade do vento excedia significativamente a velocidade de fase das ondas. Porém, foi comprovado empiricamente que se geravam ondas com uma velocidade de vento substancialmente mais baixa do que a prevista pela teoria de Kelvin e Helmholtz (Rabaud e Moisy [2020](#)).

Jeffreys ([1925](#)) propôs que o processo de geração de ondas era despoletado pela diferença de pressão entre os lados de barlavento e sotavento da onda, dependendo assim da declividade. À semelhança da teoria anterior, Jeffreys também determinou uma velocidade mínima do vento para a geração de ondas. Contudo, esta teoria tem duas limitações importantes: Primeiro, não descreve a formação inicial das ondas a partir da superfície livre plana, isto é, sem ondas. Além disso, constatou-se a partir de um conjunto de observações, que a velocidade do vento que estimara era muito superior à velocidade do vento necessária para a geração de ondas (Swart [2015](#)).

A teoria de Philips ([1957](#)) procura descrever o crescimento inicial das ondas de superfície. Segundo esta teoria, na fronteira entre a atmosfera e a superfície livre existe turbulência e remoinhos de diferentes dimensões, que são provocados pelo vento médio, que tem um perfil logarítmico. O vento médio propaga-se de forma aleatória ao longo da superfície livre, provocando flutuações de pressão ao longo da mesma (Swart [2015](#)). À semelhança da agitação marítima, a irregularidade da pressão induzida pelo vento, pode ser interpretada como a soma de várias componentes sinusoidais, porém ao contrário do caso da agitação marítima, cada componente propaga-se numa direção independente, mas normal à crista da onda. Neste sentido, algumas destas componentes têm velocidades, comprimentos de onda e direções semelhantes às ondas de superfície, o que permite, através do fenómeno de ressonância, que ocorra transferência de energia da pressão exercida pelo vento para a superfície livre, dando origem ao crescimento inicial das ondas de superfície. A teoria de Phillips limita-se a descrever o crescimento inicial das ondas, e por isso não explica o desenvolvimento das ondas até o estado do mar completamente desenvolvido (conceito abordado adiante na secção)(Holthuijsen [2007](#)).

Apesar da teoria de Philips explicar o crescimento inicial das ondas, a maior parte do desenvolvimento do espectro no processo de geração é explicado pela teoria

de Miles (1957). Esta teoria foi desenvolvida em paralelo da teoria de Phillips, ambas concordam que o vento é descrito por um perfil logarítmico, no entanto, a teoria de Miles assume que a pressão atmosférica induzida ao longo da superfície do oceano atinge um valor máximo a barlavento da crista e um valor mínimo a sotavento da crista de uma onda. De acordo com esta teoria, as ondas formam-se a partir da ressonância entre as flutuações de pressão e da tensão tangencial do vento. Por outras palavras, quando a pressão atmosférica é maior, as forças tangenciais do vento provocam a descida da superfície livre de forma a criar uma cava, no entanto, quando a pressão atmosférica é menor, esta combinada com a tensão tangencial do vento provoca a subida da superfície livre do oceano, formando-se uma crista. Neste processo ocorre transferência de energia do vento para a superfície livre através do fenómeno de ressonância. Há medida que a superfície livre é sujeita a este processo, mais energia será transferida para a mesma, originando-se ondas com maior altura. Assim, quanto maior for a altura de onda, maior será a eficácia do processo de transferência de energia, e por essa razão, gerar-se-ão ondas de maior amplitude (Holthuijsen 2007; Stewart 2008).

A geração de ondas de vento é provocada por tempestades locais em zonas distantes da costa, zonas de *fetch* (conceito abordado mais adiante), onde a intensidade e direção do vento são uniformes. Este fenómeno provoca grande irregularidade da superfície livre (vaga) e é limitado pelo comprimento do *fetch*, velocidade e duração do vento (Khandekar 2013).

Quando numa zona de *fetch*, a velocidade do vento diminui ou a direção varia mais do que 30° relativamente à direção inicial, o vento deixa de transferir energia para as ondas, pelo que estas se propagam para fora da zona de *fetch* sob a forma de ondulação. Durante a propagação da ondulação, estas sofrem o fenómeno de dispersão em frequência, e por isso dividem-se em grupos de ondas com período semelhante, propagando-se a uma velocidade conjunta (velocidade de grupo). Conforme será mencionado nesta secção do capítulo, ondas com maior período propagam-se mais rapidamente. Para além da dispersão em frequência, as ondas sofrem também dispersão angular.

Quando se aproximam de costa, as ondas são afetadas pelo fundo, correntes e (já muito próximo de costa) também por obstáculos, como por exemplo quebra-mares ou promontórios. Deste modo, as ondas sofrem diversos processos de transformação, os mais importantes para o presente trabalho são o empolamento, a refração, a reflexão e a rebentação (estes fenómenos serão abordados mais adiante no presente capítulo).

Processo de geração de ondas

Apesar de ainda não haver concordância científica a cerca do processo de geração de ondas, a velocidade do vento, (usualmente considerada a 10 metros acima da superfície), o *fetch*², e a duração que o vento exerce a sua influência no oceano são frequentemente considerados como fatores determinantes para a geração de ondas de vento (World Meteorological Organization [2023](#)).

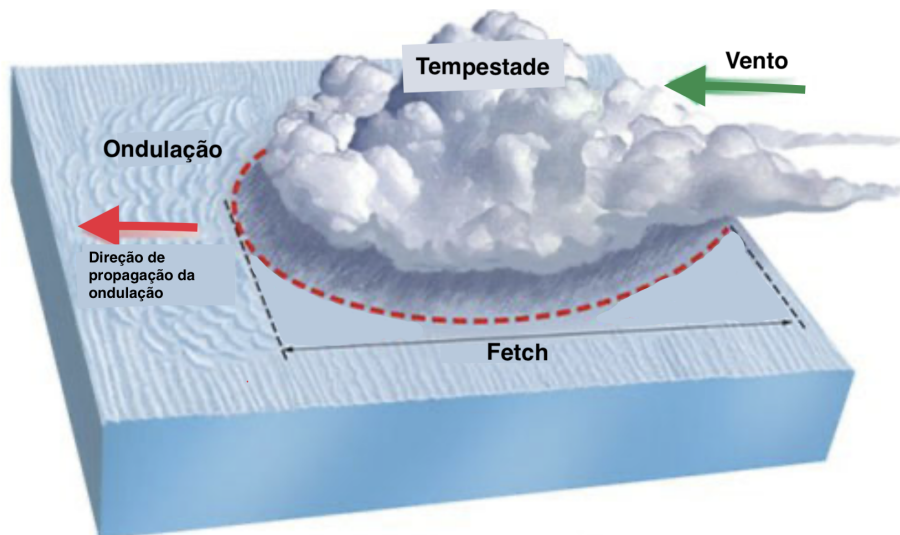


FIGURA 1.11: Representação da geração e propagação das ondas de vento, adaptado de S. James e Stull [2024](#)

Por vezes, reúnem-se as condições que garantem um *estado do mar completamente desenvolvido*³, nesta condição a altura das ondas não aumenta.

No entanto, o processo de geração de ondas pode ser limitado pelo comprimento do *fetch* ou pela duração que o vento exerce a sua influência sobre a superfície livre. Se a velocidade do vento se mantiver constante, o comprimento do *fetch* será um fator limitador, pois a duração em que o vento exerce a sua influência poderá ser insuficiente para atingir a condição de estado do mar completamente desenvolvido, e no caso de a duração ser suficiente, o comprimento do *fetch* também limitará as alturas de onda que podem ser atingidas. Conforme mencionado anteriormente,

²Área no oceano em que o vento exerce a sua influência mantendo a direção e intensidade uniformes. Não existe uma tradução geralmente aceite para este termo, pelo que é usual a utilização do termo original na literatura em português.

³Condição de equilíbrio entre a energia que é transmitida pelo vento às ondas e a energia que é dissipada e dispersada.

quando a velocidade do vento diminui ou a direção varia mais do que 30° relativamente à direção inicial, então deixa de haver transferência de energia para as ondas, e por esta razão, o fator limitador será a duração do vento.

Existem diversas fórmulas empíricas, baseadas em observações, que descrevem o processo de geração das ondas de vento. Estas fórmulas utilizam propriedades conhecidas do campo do vento (velocidade e intensidade do vento, *fetch* e duração), no entanto, de modo a simplificar as comparações as variáveis são expressas sob formas as variáveis adimensionais (World Meteorological Organization [2023](#)):

- Mínimo entre o *fetch* e o *fetch* equivalente ($\hat{F} = \min(F, F_{eq})$):

– *Fetch* equivalente (F_{eq}):

Considerando um ponto de previsão P a uma distância F da costa, onde uma componente de onda se dirige ao ponto P vinda de uma direção arbitrária θ no instante t , conforme representado na figura [1.12](#), percorrendo uma distância $C_g t$ desde que o vento começou a exercer sua influência (C_g é a velocidade de grupo da componente de onda considerada). Para simplificar, assume-se uma situação ideal onde o vento sopra continuamente na mesma direção, desde uma linha de “costa imaginária” até o ponto onde a onda é observada. A distância s entre a “costa imaginária” e o ponto de previsão é o *fetch* equivalente e é expressa como (Holthuijsen [2007](#)):

$$s = C_g t \cos \theta \quad (1.45)$$

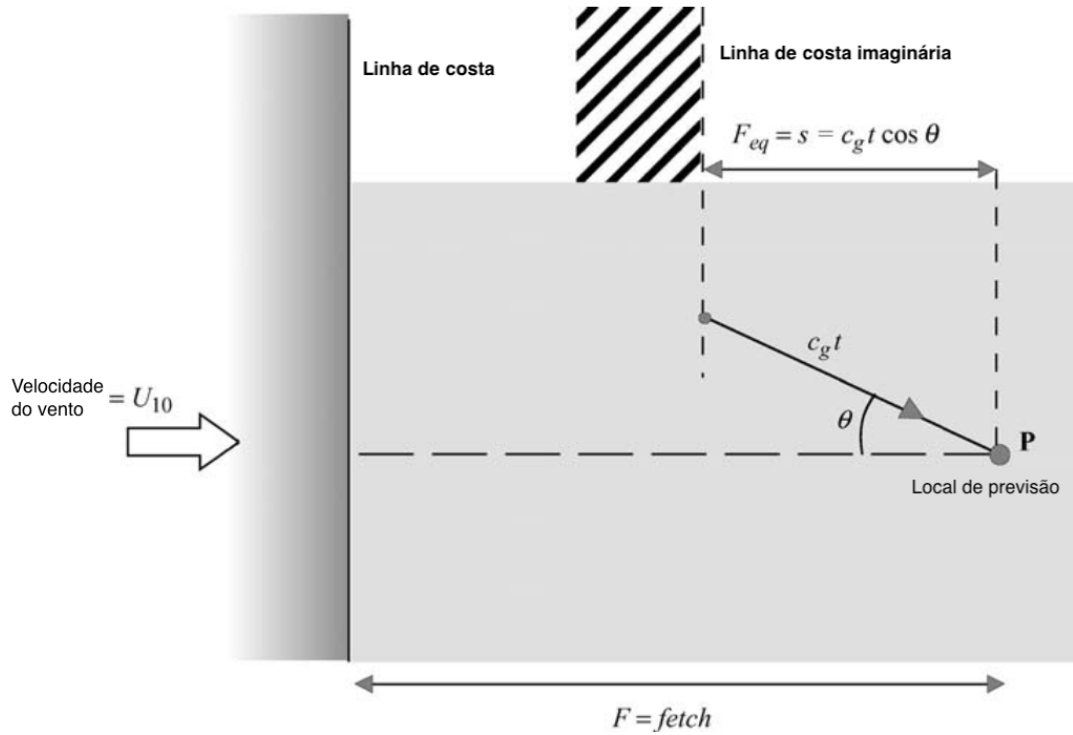


FIGURA 1.12: Representação do conceito de *fetch* equivalente sobre condições ideais. Adaptado de: Holthuijsen [2007](#)

O objetivo da relação entre o *fetch* e a duração do vento é garantir a equivalência nas situações em que o vento teve o mesmo tempo para transferir energia para a componente de onda considerada. Em outras palavras, independentemente de o vento manter a intensidade constante sobre uma grande distância (*fetch* maior) por um curto período de tempo ou sobre uma curta distância (*fetch* menor) por um longo período, o tempo total que o vento teve para transferir energia para a onda é considerado equivalente. As variáveis adimensionais são expressas da seguinte forma:

$$\tilde{F} = g \frac{\hat{F}}{U_{10}^2} \quad (1.46)$$

- Período de pico:

$$\tilde{T} = g \frac{T_p}{U_{10}} \quad (1.47)$$

- Altura significativa:

$$\tilde{H} = g \frac{H_{mo}}{U_{10}^2} \quad (1.48)$$

- Profundidade:

$$\tilde{d} = g \frac{D}{U_{10}^2} \quad (1.49)$$

A evolução da altura significativa e do período significativo são descritos, no caso da simplificação anteriormente considerada, pelas seguintes equações:

$$\tilde{H} = \tilde{H}_\infty \tanh(a_3 \tilde{d}^{b_3}) \tanh\left(\frac{a_1 \tilde{F}^{b_1}}{\tanh(a_3 \tilde{d}^{b_3})}\right) \quad (1.50)$$

$$\tilde{T} = \tilde{T}_\infty \tanh(a_4 \tilde{d}^{b_4}) \tanh\left(\frac{a_2 \tilde{F}^{b_2}}{\tanh(a_4 \tilde{d}^{b_4})}\right) \quad (1.51)$$

Os coeficientes $a_1, a_2, a_3, a_4, b_1, b_2, b_3, b_4$ são parâmetros determinados empiricamente (Holthuijsen 2007).

A evolução da altura significativa e do período significativo podem ser obtidos partir da análise dos gráficos da figura 1.13.

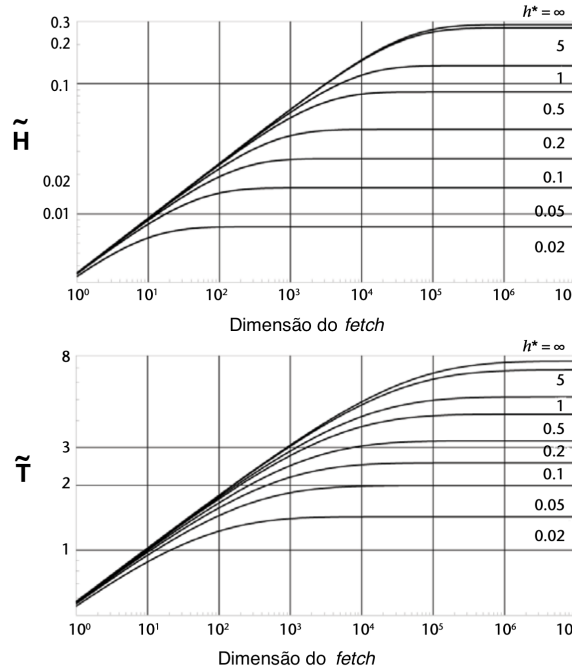


FIGURA 1.13: Evolução da altura (no topo) e períodos (em baixo) significativos, em função da dimensão do fetch e gama de profundidades. Adaptado de: World Meteorological Organization 2023

Conforme mencionado anteriormente Sverdrup e Munk desenvolveram o primeiro método manual para previsão de agitação marítima, com base num número

elevado de observações. No entanto, Breugem and Holthuijsen refinaram-no, desenvolvendo um novo conjunto de equações capaz de descrever a altura significativa e o período de pico. Este conjunto de equações é representado graficamente num diagrama na figura 1.14 que pode ser utilizado como um método manual de previsão.

Neste gráfico, as linhas vermelhas correspondem a uma velocidade constante do vento. A altura significativa H_{m0} é obtida a partir das linhas horizontais pretas e o período de pico T_p a partir das linhas azuis. As linhas verticais indicam a duração em horas na qual esse estágio de desenvolvimento será alcançado, a partir de uma altura de onda nula. Se a duração for um fator limitativo, as ondas não se desenvolverão ao longo das linhas escuras, independentemente do comprimento do *fetch*, pelo que se deve considerar o cruzamento das linhas que corresponde à duração das outras variáveis. No lado direito diagrama em questão, as curvas são aproximadamente horizontais, o que implica que para uma determinada velocidade do vento, quando a duração e o *fetch* são longos o suficiente, as ondas atingem a altura máxima, atingindo assim um estado de mar completamente desenvolvido (World Meteorological Organization 2023).

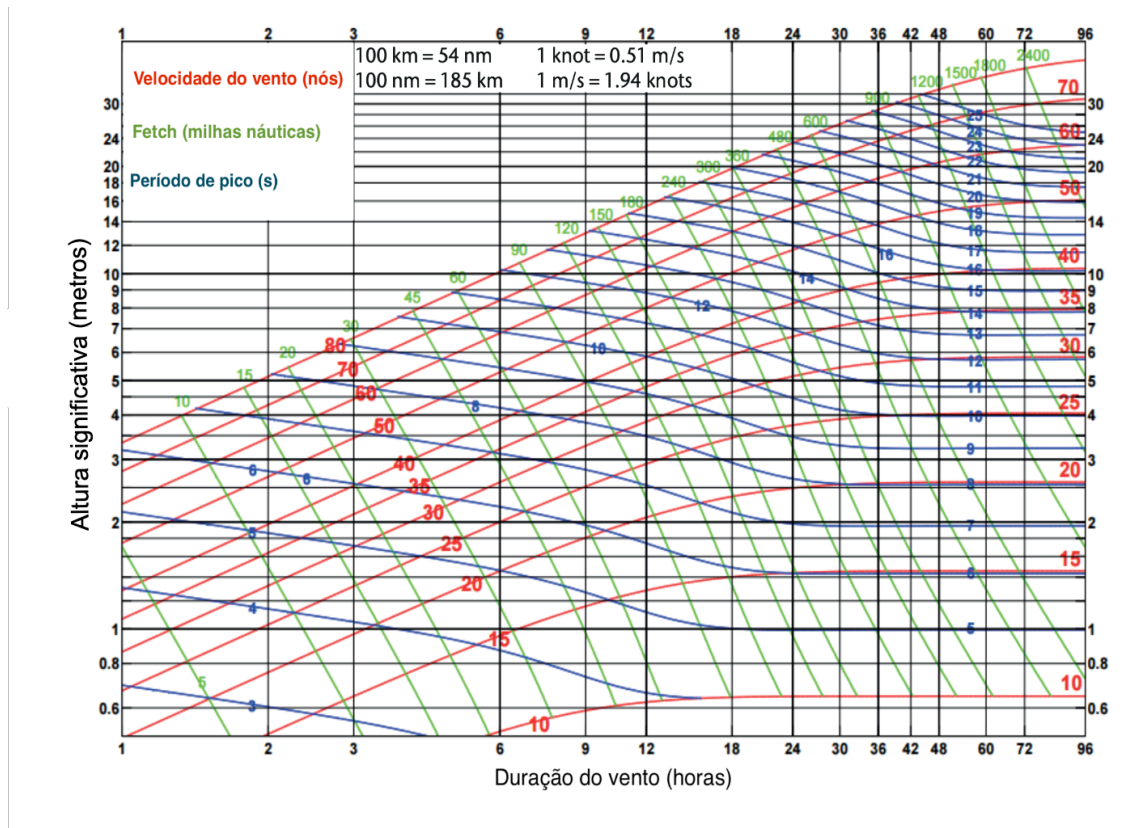


FIGURA 1.14: Método manual de previsão de agitação marítima.

Adaptado de: World Meteorological Organization 2023

Dispersão angular e em frequência

Conforme mencionado anteriormente, quando a velocidade do vento diminui significativamente ou a sua direção sofre uma alteração superior a 30° , o vento deixa de transferir energia para as ondas, pelo que estas se propagam para fora da zona de *fetch*.

A ondulação que se afasta da zona em que foi gerada (área de *fetch*) sofre dois tipos de dispersão: geométrica e em frequência. A primeira caracteriza-se pelo o espalhamento geométrico das ondas, ou seja, a maior parte da energia, e por isso as ondas com maior altura, deslocam-se na direção média do vento na zona de geração (*Fetch*). No entanto, em redor da direção do vento também se propagam ondas. Quanto maior o ângulo formado entre a direção de propagação da ondulação e a direção do vento na zona de *Fetch*, as ondas vão diminuindo em altura, por receberem menor quantidade de energia. A segunda resulta do agrupamento de ondas em função do seu período (ou frequência). Estes grupos de ondas propagam-se a uma velocidade conjunta (velocidade de grupo), por grandes distâncias sem que haja dissipação significativa de energia (Holthuijsen 2007). Existem ábacos para determinar manualmente os fatores de dispersão geométrica e em frequência, como é o caso dos ábacos desenvolvidos por Bretchneider em 1952 (World Meteorological Organization 2023)

Existem ainda interações onda-onda que serão mencionadas na secção 1.3 do capítulo, mas não são explicadas neste ponto do trabalho.

Processos de transformação de ondas

Antes da rebentação (dissipação máxima de energia) existe conservação de densidade de energia das ondas. No entanto, as ondas de vento podem transformar-se devido ao fenómeno do empolamento, processo no qual a altura de onda aumenta devido à variação da velocidade de grupo, causada pela redução de profundidade ou refração. O processo na qual as ondas alteram a sua direção de propagação, provocado também pela diminuição de profundidade designa-se por refração. A alteração da profundidade, e consequentemente velocidade ao longo de uma crista de onda, provoca alterações na direção de propagação da onda, forçando-as a propagarem-se consoante um ângulo de incidência, para uma onda harmónica. A lei de Snell permite descrever a nova direção de propagação (Holthuijsen 2007). Esta secção do capítulo tem como objetivo descrever estes fenómenos de transformação de ondas e ainda caracterizar o fenómeno e os critérios de rebentação. Seguidamente, serão resumidos os aspetos fundamentais destes três fenómenos.

Como definido previamente neste capítulo para o estudo da teoria linear, a elevação da superfície livre de cristas paralelas e uniformes (que se propaga na direção de x , é definida pela função):

$$\eta(x, y) = \frac{H}{2} \sin(kx - \omega t) \quad (1.52)$$

Para descrever os processos de transformação de ondas, consideremos uma onda sinusoidal que se propaga na direção de θ relativamente ao sistemas de eixos representados pela figura 1.15:

$$\eta(x, y, t) = \frac{H}{2} \sin(k_x x + k_y y - \omega t) \quad (1.53)$$

Como representado na figura 1.15 é possível reescrever a expressão 1.53 da seguinte forma:

$$\eta(x, y, t) = \frac{H}{2} \sin(k \cos(\theta)x + k \sin(\theta)y - \omega t) \quad (1.54)$$

Considerando κ , o número de onda, como definido na teoria linear (equação 1.4) a expressão 1.54 pode ser expressa nas suas formas equivalentes:

$$\eta(x, y, t) = \frac{H}{2} \sin\left(\frac{2\pi}{\lambda}(\cos(\theta)x + \sin(\theta)y) - \omega t\right) \quad (1.55)$$

$$\eta(x, y, t) = \frac{H}{2} \sin(\vec{k} \cdot \vec{x} - \omega t) \quad (1.56)$$

sendo $\vec{k} = (k \cos(\theta), k \sin(\theta))$ o vetor número de onda, cujo o módulo é obtido a partir da relação de dispersão para o caso anterior de uma onda que se propaga na direção de x sobre um fundo com profundidade constante.

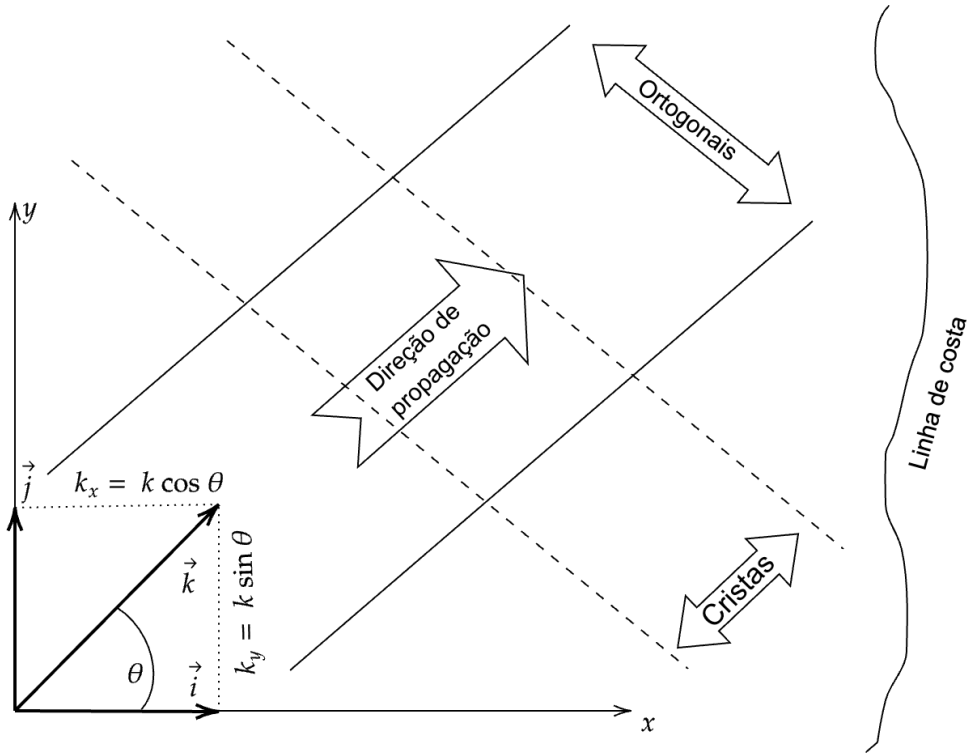


FIGURA 1.15: Representação da propagação de ondas num plano bidimensional e do conceito de ortogonal.

O diagrama da figura 1.15 será útil para consideração dos fenômenos de empolamento e da refração, que serão desenvolvidos seguidamente.

Empolamento: O fenômeno de empolamento caracteriza-se como a variação da altura de onda devido à variação da velocidade de grupo na aproximação das ondas a costa quando as ondas se propagam sobre uma batimetria retílinea e uniforme com incidência normal, isto é $\theta_\infty = 0$. Admitindo que não existe difração, reflexão ou dissipação de energia, então a energia numa área $b \times C_g \times t$ é constante (figura 1.16), pelo que se pode definir a seguinte igualdade:

$$E \times C_g \times t_1 \times b = E' \times C'_g \times t_2 \times b \quad (1.57)$$

Como b é constante, nas condições previamente descritas (batimetria retílinea e uniforme e $\theta_\infty = 0$) é possível concluir que o fluxo de energia é igualmente constante, desde a condição de propagação de águas profundas até à linha de rebentação. A expressão seguinte representa a igualdade do fluxo de energia entre a condição de águas profundas e a condição de propagação a uma profundidade D

(águas profundas ou fundos baixos até à linha de rebentação):

$$(E C_g)_\infty = (E C_g)_D \quad (1.58)$$

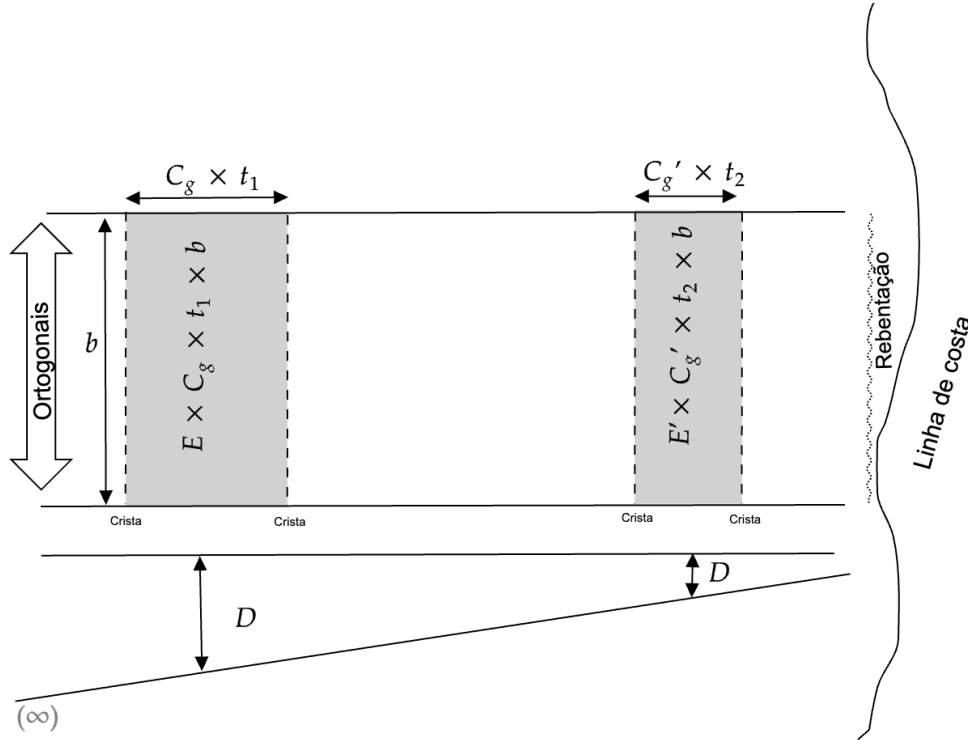


FIGURA 1.16: Representação do fluxo de energia entre duas ortogonais

A partir da igualdade definida pela equação 1.58 e substituindo-se E pela expressão que define a densidade de energia (equação 1.19), obtém-se a relação entre alturas de onda a diferentes profundidades expressa da seguinte forma:

$$H_D^2 = H_\infty^2 \frac{C_{g\infty}}{C_{gD}} = H_\infty^2 K_s \quad (1.59)$$

O parâmetro K_s é o coeficiente de empolamento, e pode ser expresso na seguinte forma:

$$K_s = \sqrt{\frac{\coth kD}{1 + \frac{2kD}{\sinh 2kD}}} \quad (1.60)$$

sendo k o número de onda conforme definido previamente na teoria linear e D a profundidade.

O efeito do empolamento, isoladamente dos efeitos de propagação, geração e dissipação, é caracterizado por uma diminuição inicial da altura de onda, porque a velocidade de grupo aumenta, mas à medida que a velocidade de grupo diminui, a altura de onda aumenta (Holthuijsen [2007]).

Refração: O fenómeno da refração ocorre quando o ângulo de incidência em águas profundas for diferente de zero ou a batimetria for irregular, que é o caso mais frequente. Durante a ocorrência deste fenómeno, zonas das frentes de onda com igual fase propagam-se em zonas de diferente profundidade, provocando uma alteração na sua velocidade, e consequentemente a alteração da direção de propagação (Pope e Lockhart [2002]).

Nesta situação as ortogonais não têm um espaçamento constante b . No entanto, desprezando os efeitos de difração e dissipação de energia, o princípio da conservação de energia mantém-se (Pope e Lockhart [2002]). Desta forma, considerando uma onda sinusoidal (equação 1.53), o fenómeno pode ser descrito pela lei de *Snell* (Holthuijsen [2007]):

$$\frac{\sin \theta_{\infty}}{C_{\infty}} = \frac{\sin \theta}{C} \quad (1.61)$$

Em que C e C_{∞} , correspondem às velocidades de propagação em águas profundas e de profundidade D , respetivamente.

Admitindo que não existe difração ou dissipação de energia, então o princípio de conservação de energia pode ser aplicado de modo a obter a altura de onda (ainda que b não seja constante):

$$(EC_g)_1 b_1 = (EC_g)_2 b_2 \quad (1.62)$$

$$H_2 = H_1 \sqrt{\frac{C_{g1}}{C_{g2}}} \sqrt{\frac{b_1}{b_2}} \quad (1.63)$$

Mesmo num local onde a batimetria é praticamente retilínea pode ocorrer refração, basta que a direção de propagação da onda seja tal que, pontos da frente de onda com a mesma fase se propaguem numa profundidade diferente, tal como é representado na figura 1.17 (Pope e Lockhart [2002]; World Meteorological Organization [2023]).

O coeficiente de refração pode ser determinado a partir da lei de Snell e é expresso por (Holthuijsen 2007):

$$K_r = \sqrt{\frac{b_\infty}{b_D}} = \sqrt{\frac{\cos \theta_\infty}{\cos \theta_D}} \quad (1.64)$$

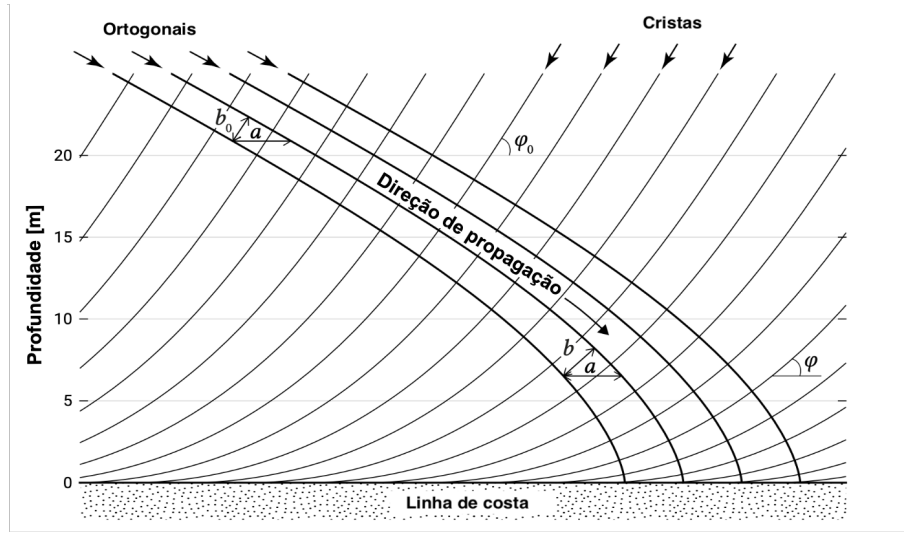


FIGURA 1.17: Exemplo da refração que ocorre num local de batimetria retilínea. Adaptado de Pope e Lockhart 2002

Rebentação: Com a aproximação da ondulação à costa, a profundidade vai diminuindo, as ondas deixam de ter uma forma sinusoidal passando a exibir cristas em pico. Inicialmente, a altura de onda começa por aumentar, devido ao efeito combinado do empolamento e da refração. Quando a velocidade das partículas da crista ultrapassa a velocidade de propagação da onda, esta torna-se instável o que faz com que se curve sobre si própria e rebente.

Existem diversos critérios de rebentação para determinar a altura de onda no fundo. Esses critérios estabelecem limites para a declividade, ou seja $\frac{H}{\lambda}$, como o critério de Miche 1944 ou para a profundidade relativa $\frac{H}{D}$, como por exemplo o critério de McCowan 1894 e Weggel 1972.

O critério de Miche 1944 define uma declividade limite nas condições de águas profundas e profundidade intermédia para que a onda rebente e é expressa por (Pope e Lockhart 2002):

$$\left(\frac{H}{\lambda}\right)_{max} = 0,142 \tanh(\kappa D) \quad (1.65)$$

O critério de McCowan [1894] define um valor limite (índice de rebentação κ_b) de profundidade relativa ($\frac{H}{D}$) para que a onda rebente, é este:

$$\left(\frac{H}{D}\right)_b = \kappa_b = 0,78 \quad (1.66)$$

O critério de Weggel [1972], é um critério mais rigoroso, que para além da profundidade relativa, considera também o declive do fundo (m). Pelo que índice de rebentação é calculado da seguinte forma:

$$\kappa_b = \frac{1,56}{1 - e^{-19,5m}} - 43,75(1 - e^{-19m}) \frac{H_b}{gT^2} \quad (1.67)$$

O tipo de rebentação é influenciado pela declividade do fundo $\frac{H}{\lambda}$, que pode ser descrita pelo número de Iribarren expresso nos seguintes termos:

$$Ir = \frac{\tan \theta}{\sqrt{\frac{H}{\lambda_\infty}}} \quad (1.68)$$

O critério de classificação do tipo de rebentação foi obtido através de um conjunto de observações (Holthuijsen [2007]), e respeita as seguintes condições (United States Army. Corps of Engineers e Coastal Engineering Research Center (U.S.) [1984]):

- *Spilling* $Ir < 0,4$
- *Plunging* $0,4 < Ir < 2,4$
- *Collapsing* $2,4 < Ir < 3,1$
- *Surging* $Ir > 3,1$

A figura 1.18 representa graficamente os tipos de rebentação definidos.

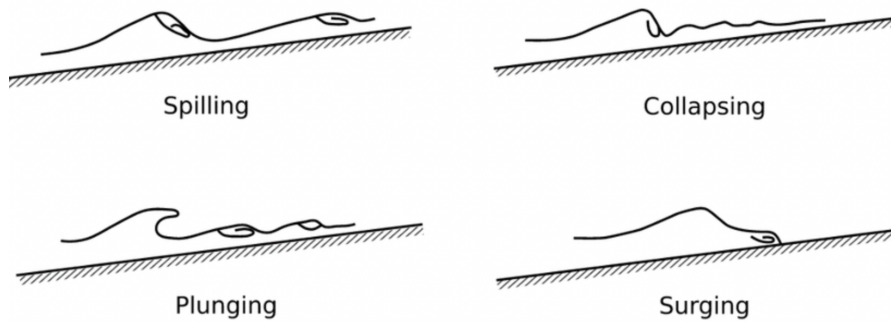


FIGURA 1.18: Ilustração do tipo de rebentação. Fonte: Fleit [s.d.]

1.3 Modelos físico-matemáticos de previsão de agitação marítima

O primeiro modelo de previsão da agitação marítima foi desenvolvido pelos cientistas Sverdrup e Munk, durante a segunda Guerra Mundial, como forma de apoiar o planeamento das operações dos Aliados na projeção de forças através de desembarques anfíbios (Sverdrup e Munk [1947]). Neste contexto, estes dois cientistas desenvolveram o seu modelo através de ábacos, regras empíricas e fórmulas simples para que pudessem ser facilmente utilizados no cenário operacional (Pinardi et al. [2017]). Os diagramas desenvolvidos permitiram a previsão da altura significativa e do período das ondas, com base no comprimento do *fetch*, na velocidade e duração do vento. Este modelo foi amplamente utilizado durante décadas como método primário de previsão até ao aparecimento dos primeiros modelos numéricos, corridos em computadores digitais.

A agitação marítima é irregular e por isso a sua descrição numérica requer a utilização de parâmetros estatísticos. O espectro de densidade de energia é bastante utilizado pelos modelos numéricos, visto que permite deduzir a maior parte dos parâmetros espectáveis de um modelo operacional, como por exemplo a altura significativa, o período e frequência de pico, o período e frequência máxima, o espectro direcional e a direção média das ondas (World Meteorological Organization [2023]).

Os modelos de agitação marítima calculam a evolução do espectro de energia para uma determinada área através da equação do balanço de energia, que pode ser expressa da seguinte forma (Holthuijsen [2007]):

$$\frac{\partial E(f, \theta, x, y, t)}{\partial t} + \frac{\partial C_{gx}}{\partial x} E(f, \theta, x, y, t) + \frac{\partial C_{gy}}{\partial y} E(f, \theta, x, y, t) = S_{in} + S_{nl} + S_{ds} \quad (1.69)$$

Os termos do membro esquerdo da equação representam a variação temporal da densidade de energia presente no sistema.

O termo S_{in} corresponde à energia inserida no sistema pelo vento. O termo S_{nl} corresponde à energia transferida pelas interações não lineares onda-onda. Finalmente, o termo S_{ds} corresponde à energia dissipada.

Até ao momento presente, foram desenvolvidos modelos de três gerações distintas (World Meteorological Organization [2023]; Holthuijsen [2007]). Este diferem

na forma como calculam o termo não-linear das interações onda-onda na equação do balanço de energia (S_{nl}).

Os modelos de primeira geração descreviam o processo de geração implementando as teorias de Jeffreys e Miles, mas não incluíam a representação matemática do processo de interação onda-onda (Pinardi et al. 2017).

Os modelos de segunda geração formulam o termo não-linear a partir da aplicação de um espectro de referência, como por exemplo o JONSWAP ou o espectro de Pierson-Moskowitz (World Meteorological Organization 2023). Estes modelos definem a agitação marítima como a soma de um grande número de ondas que se propagam a um período constante tal como na teoria linear.

Por fim, os modelos de terceira geração calculam o termo não-linear (S_{nl}), a partir da sua expressão teórica, ainda que seja necessário fazer aproximações numéricas e analíticas para tornar os cálculos mais expeditos (World Meteorological Organization 2023).

O modelo *WaveWatch III*, é um modelo de terceira geração desenvolvido pelo *US National Centers for Environmental Prediction* (NCEP) da agência *National Oceanic and Atmospheric Administration* (NOOA), que tem por base a equação de balanço de energia (Tolman et al. 2009). Este modelo é baseado na equação 1.69, mas utiliza mais termos para representar todos os processos descritos pelo membro direito dessa equação. No modelo *WaveWatch III*, o membro direito da equação do balanço de energia (equação 1.69) é composto nos termos seguintes:

$$S = (S_{ln} + S_{In} + S_{nl} + S_{ds}) + (S_{bot} + S_{db} + S_{tr} + S_{sc} + S_{ice} + S_{ref} + S_{user}) \quad (1.70)$$

A equação 1.70 é separada por dois conjuntos de termos. O primeiro refere-se aos termos que descrevem os fenómenos na condição de propagação de águas profundas, enquanto que o segundo conjunto de termos representam os fenómenos que ocorrem nas condições de propagação de águas intermédias e fundos baixos, os quais definem a evolução do espectro direcional de energia. Nesta equação, cada termo corresponde a energia inserida, redistribuída ou dissipada (Tolman et al. 2009), são estes:

- S_{ln} é responsável por modular a condição inicial do modelo (superfície livre na condição de repouso e sem consideração de forças externas)
- S_{In} corresponde à energia inserida no sistema pelo vento.

- S_{nl} representa a transferência energia provocada pelas interações não-lineares onda-onda.
- S_{ds} corresponde à dissipação de energia. Apesar deste termo não estar relacionado com o termo S_{In} , o balanço entre ambos descreve o crescimento integral da energia nas ondas.
- S_{bot} representa a interação das ondas com o fundo do oceano, como por exemplo o fenómeno da refração (pode ser expressa através da parametrização linear de JONSWAP)
- S_{db} caracteriza a energia dissipada através da rebentação que ocorre na condição de condição de fundos baixos
- S_{tr} representa as interações não lineares entre grupos de ondas
- S_{sc} , são as interações influenciadas pelas características do fundo, que de momento não está configurado na versão do modelo operada pelo IH.
- S_{ice} , representa interações entre as ondas e o gelo, também não está presente na configuração adotada pelo IH.
- S_{ref} é o parâmetro que descreve as condições de reflexão devido aos obstáculos que as ondas encontram quando se propagam em direção à linha de costa.
- S_{user} é um termo personalizável pelo utilizador, que pode acrescentar por exemplo, a influência das correntes oceânicas ou a variação da topografia no local de forma a refinar a previsão.

Capítulo 2

Modelos estatísticos e modelos de aprendizagem automática

O modelo *WaveWatch III* permite obter previsões com base em equações que descrevem os processos de geração, propagação e transformação das ondas de vento. Contudo, a previsão dos parâmetros caracterizadores da agitação marítima pode também ser interpretada como um problema de previsão de séries temporais. Recentemente, têm sido propostos métodos de previsão dos parâmetros de agitação marítima a partir de modelos estatísticos de Aprendizagem Automática, que uma vez treinados de forma correta, e com meios de cálculo modestos, permitem obter previsões fiáveis numa fração do tempo necessário para a obtenção de soluções numéricas com modelos físico-matemáticos (Oleg Makarynskyy [2005](#); Demetriou, Michailides, Papanastasiou e Onoufriou [2021](#)). A abordagem potencialmente mais promissora é a utilização de redes neuronais, que depois de treinadas funcionam como funções do tipo «caixa negra» (Demetriou, Michailides, Papanastasiou e Onoufriou [2021](#)). Neste capítulo, são descritos os conceitos e os modelos estatísticos e de Aprendizagem Automática que serão utilizados no presente trabalho.

Os modelos sugeridos pela Aprendizagem Automática são fundamentados pela estatística clássica (Bontempi [2021](#)), no âmbito da qual o problema de previsão das condições de parâmetros de agitação marítima deve ser interpretado como um problema estatístico. Além disso, a questão deve ainda ser considerada como um problema de previsão de séries temporais, conforme descrito seguidamente.

Este capítulo encontra-se organizado em sete secções. Nas primeiras cinco, são apresentados vários modelos estatísticos: modelos autorregressivos (AR), modelos de médias móveis (MA), modelos autorregressivos integrados de médias móveis (ARIMA), e modelos de regressão linear (simples e múltipla). Na secção [2.6](#) são descritos os modelos de Aprendizagem Automática. Na última secção é apresentado

um resumo dos trabalhos sobre a aplicação de técnicas de Aprendizagem Automática à previsão das condições de agitação marítima, com maior relevância para o presente trabalho.

A aplicação dos modelos AR, MA e ARIMA permite obter uma compreensão clara sobre como os registos de instantes passados influenciam as previsões atuais. Já os modelos de regressão linear (simples e múltipla) permitem identificar e quantificar as relações lineares entre as variável dependente (como altura significativa) e as restantes variáveis independentes (período de pico, altura máxima de onda, etc.).

2.1 Séries temporais. Modelos descritivos

Uma série temporal é uma coleção de variáveis $x_{t1}, x_{t2}, x_{t3}, x_{t4}, \dots, x_{tn}$, ordenadas ao longo do tempo, em que cada t_i é um instante de tempo observado e n é um número inteiro positivo que representa a quantidade de instantes observados (Shumway, Stoffer e Stoffer [2000](#)). Os registos dos parâmetros caracterizadores de agitação marítima são séries temporais, pois constituem um conjunto de observações que ocorrem sequencialmente ao longo do tempo.

Uma série temporal diz-se estocástica se for constituída por variáveis aleatórias, e por isso só poderão ser feitas previsões associadas a uma probabilidade de ocorrência, que é condicionada pelo conhecimento dos valores passados (Chatfield e Xing [2019](#)).

Numa série temporal é importante avaliar o seu mecanismo de geração, ou seja, o processo/fenómeno que a mesma representa. É também importante clarificar os objetivos da análise de exploração da série, definir as variáveis necessárias e o comprimento da série temporal, o intervalo entre observações, e determinar a(s) variável(eis) que deve(m) ser prevista(s).

Um outro aspeto fundamental na análise exploratória de séries temporais é a compreensão da estrutura estacionária ou não estacionária das mesmas.

Uma série temporal é dita estacionária se a distribuição conjunta de $x_{t_1}, x_{t_2}, x_{t_3}, \dots, x_{t_k}$ for igual à distribuição de $x_{t_1+\tau}, x_{t_2+\tau}, x_{t_3+\tau}, \dots, x_{t_k+\tau}$ para todos os valores de $t_1, t_2, t_3, \dots, t_k, \tau$. Por outras palavras, deslocar o instante de origem da série τ pontos no tempo não altera as distribuições conjuntas de uma série estacionária, que dependem apenas dos intervalos entre os valores $t_1, t_2, \dots, t_k$ (Chatfield e Xing [2019](#)).

Na análise exploratória de uma série temporal é também essencial caracterizar a dependência entre valores sucessivos da própria série temporal (autocorrelação). Essa dependência é descrita pelas funções de autocovariância, autocorrelação e autocorrelação parcial, que serão descritas em seguida.

A função de autocovariância caracteriza a dependência linear entre dois pontos na mesma série observados nos instantes s e t , e é expressa pela expressão:

$$\gamma_x(s, t) = \text{cov}(x_s, x_t) = E[(x_s - \mu_s)(x_t - \mu_t)] \quad (2.1)$$

A autocovariância é uma medida de similaridade linear entre valores da série temporal em instantes diferentes. No entanto, a sua escala depende da série temporal.

A função autocorrelação é a normalização da autocovariância pelo produto da variância, tornando os valores da mesma adimensionais no intervalo de valores entre -1 e 1 , o que permite caracterizar a dependência entre valores sucessivos independentes das unidades, facilitando também a comparação entre a dependência temporal de séries diferentes.

A função autocorrelação (ACF), mede a previsibilidade linear da série ao longo do tempo t . Se for possível prever perfeitamente x_t , a partir de x_s através de uma relação linear $x_t = \beta_0 + \beta_1 x_s$, então a correlação terá valor 1 quando $\beta_1 \geq 0$, e -1 quando $\beta_1 \leq 0$. Assim a função ACF permite avaliar a dependência temporal entre os valores de uma série. A função autocorrelação é expressa da seguinte forma (Shumway, Stoffer e Stoffer 2000):

$$\rho(s, t) = \frac{\gamma(s, t)}{\sqrt{\gamma(s, s)\gamma(t, t)}} \quad (2.2)$$

Para uma série estacionária terá a seguinte expressão:

$$\rho(h) = \frac{\gamma(t + h, t)}{\sqrt{\gamma(t + h, t + h)\gamma(t, t)}} = \frac{\gamma(h)}{\gamma(0)} \quad (2.3)$$

A função autocorrelação parcial (PACF) descreve a correlação entre diferentes instantes t num intervalo (atraso ou *lag*), eliminando a dependência linear entre x_{t+p} e os valores intermédios $x_{t+1}, \dots, x_{t+p-1}$, e é definida por (Shumway, Stoffer e Stoffer 2000):

$$\phi_{11} = \text{corr}(x_{t+1}, x_t) = \rho(1) \quad (2.4)$$

$$\phi_{hh} = \text{corr}(x_{t+h} - \hat{x}_{t+h}, x_t - \hat{x}_t), \quad h \geq 2. \quad (2.5)$$

Contudo, por vezes é necessário descrever a dependência linear entre uma série temporal y_t e outra série temporal x_s , pelo que partindo do pressuposto de que ambas as séries têm variâncias finitas, a covariância é definida pela expressão seguinte:

$$\rho_{xy}(s, t) = \frac{\gamma_{xy}(s, t)}{\sqrt{\gamma_x(s, s)\gamma_y(t, t)}} \quad (2.6)$$

sendo $\gamma_{xy}(s, t) = \text{cov}(x_s, y_t) = E[(x_s - \mu_s)(y_t - \mu_t)]$;

onde $\mu_s = E(x_s)$ e $\mu_t = E(y_t)$.

Para uma série temporal estacionária, a média μ , a variância σ^2 e autocovariância $\text{cov}(x_s, x_t)$ são constantes para todos os valores de t_i , o que facilita a análise da série, permitindo a previsão dos valores futuros da série com base nos valores passados da mesma, após a aplicação de modelos estatísticos de previsão, bem como a comparação entre diferentes séries. Como tal, é necessário transformar séries não estacionárias em séries estacionárias, de modo a estabilizar a variância, através da aplicação de uma transformação logarítmica ou, de forma mais geral, através da transformação de Box-Cox, expressa da seguinte forma (Chatfield e Xing [2019](#)):

$$y_t = \begin{cases} (x_t^\lambda - 1) / \lambda & \lambda \neq 0 \\ \log x_t & \lambda = 0 \end{cases} \quad (2.7)$$

Segundo Hyndman e Athanasopoulos [2014](#), efetuar previsões de uma série temporal a partir de uma abordagem estatística envolve cinco etapas distintas:

- Definir o problema
- Reunir informação
- Efetuar uma análise exploratória dos dados
- Escolher e ajustar modelos de previsão adequados
- Avaliar o modelo de previsão

A definição do problema consiste em determinar a variável (ou variáveis) que deve(m) ser previstas, o intervalo das previsões, o comprimento da série e os intervalos de amostragem entre registos. Depois de recolhidos os dados (amostra) a análise exploratória dos mesmos consiste na análise de padrões, deteção de lacunas

e *outliers*⁴ ou valores anómalos, na detecção de padrões e tendências na amostra, e ainda no estudo prévio das variáveis. A escolha do melhor modelo, depende muitas vezes dos dados históricos disponíveis e das relações entre a variável a prever e os previsores. Assim, o modelo selecionado corresponderá à descrição probabilística que melhor descreve a estrutura observada na série temporal. Os modelos podem classificar-se como (Chatfield e Xing 2019):

- Univariados, se a variável de previsão e os previsores pertencerem à mesma série temporal.
- Multivariados, se o valor da previsão de uma variável depende dos valores de uma ou mais séries temporais.

Assim que o modelo estiver selecionado e os parâmetros estimados, deve-se avaliar o desempenho do modelo e compará-lo se necessário com outros modelos, podendo esta ser feita através da análise dos resíduos, e da utilização de índices de ajustamento (ou métricas de desempenho).

O objetivo da previsão de séries temporais é minimizar o erro, isto é, a diferença entre as previsões e os valores observados, segundo critérios adequados avaliados em termos de índices de ajustamento. Alguns dos índices mais utilizados são o erro quadrático, o erro médio absoluto e o coeficiente de determinação R^2 (Wei 2006) conforme descrito seguidamente.

O erro médio quadrático (RMSE) é definido na forma seguinte:

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^m (x_i - y(x_i))^2} \quad (2.8)$$

em que n é o número de observações, x_i corresponde aos valores observados e $y(x_i)$ são os valores previstos. Embora o RMSE seja geralmente uma métrica mais utilizada para avaliar modelos de regressão, este índice é sensível à presença de valores anómalos (*outliers*). Um outro índice menos sensível é o erro médio absoluto, expresso da seguinte forma (Geron 2019):

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y(x_i) - x_i| \quad (2.9)$$

em que n é o número de observações, x_i corresponde aos valores observados e $y(x_i)$ corresponde aos valores previstos.

⁴Observação estatística atípica comparativamente às restantes observações da amostra considerada.

Em modelos de regressão, é muito comum utilizar o coeficiente de determinação R^2 como uma métrica de avaliação de desempenho do modelo. Este coeficiente mede a proporção de variância explicada pelo modelo de regressão e pode ser calculado através da correlação entre os valores observados x_i e os valores previstos $y(x_i)$, ou alternativamente da seguinte forma (Hyndman e Athanasopoulos 2014):

$$R^2 = \frac{\sum (x_i - \bar{x}_i)^2}{\sum (y(x_i) - \bar{x}_i)^2} \quad (2.10)$$

2.2 Modelo autoregressivo (AR)

Um dos modelos estatísticos mais utilizados para descrever processos estocásticos estacionários univariados cujas realizações são séries temporais x_t é o modelo autoregressivo. Um modelo autorregressivo de ordem p (abreviado por $AR(p)$), utiliza os valores de uma série temporal para prever os valores futuros de uma variável e é expresso da seguinte forma (Shumway, Stoffer e Stoffer 2000):

$$x_t = \alpha_1 x_{t-1} + \dots + \alpha_p x_{t-p} + w_t \quad (2.11)$$

em que x_t é uma série temporal estacionária e $\alpha_1, \dots, \alpha_p$ são constantes e diferentes de zero, e w_t é ruído branco, ou seja uma série aleatória estacionária, normalmente distribuída, com média e correlação zero e variância σ_w^2 . A ordem p dos modelos $AR(p)$ deve ser considerado após a análise da função correlação parcial (PACF). Esta corresponde ao número mais extremo de *lags* que difere de zero e é estatisticamente significativo. Após essa ordem os restantes *lags* não são considerados como previsores essenciais.

2.3 Modelo médias móveis (MA)

Um outro modelo utilizado para descrever séries temporais é o modelo de médias móveis de ordem q , $MA(q)$, o qual é expresso pela expressão seguinte:

$$x_t = \theta_1 w_{t-1} + \theta_2 w_{t-2} + \dots + \theta_q w_{t-q} + w_t \quad (2.12)$$

Neste modelo a ordem q simboliza o intervalo entre valores da média móvel, $\theta_1, \theta_2, \dots, \theta_q$ é diferente de 0 e w_t é uma série aleatória, com média zero e variância σ_w^2 . O valor de ordem q representa o número de termos dos resíduos que influenciam o valor atual x_t . Para a determinar, é necessário analisar a função autocorrelação

(ACF). Desta forma, a ordem q corresponde ao valor extremo de *lag* que difere de zero e é estatisticamente significativo.

Uma série estacionária pode ainda ser representada por um modelo ARMA(p, q) que contempla os dois modelos anteriores, sendo que p e q são parâmetros a determinar através das funções ACF e PACF bem como o resultado de testes estatísticos como o teste de Ljung-Box (explicado mais adiante na secção).

2.4 Modelo ARIMA

Os modelos AR e MA referidos anteriormente só são aplicáveis apenas para séries temporais estacionárias. Contudo, os modelos ARIMA (p, d, q) (*Autoregressive Integrated Moving Average Model*), permitem a transformação de séries não estacionárias em séries estacionárias e representam a série resultante como a combinação de um modelo autorregressivo como um modelo de médias móveis. Estes modelos têm três componentes:

- "AR"(p) : Componente autoregressiva.
- "I"(d) : Componente integrada, que torna séries não estacionárias em séries estacionárias. A ordem d é o número de diferenças necessárias para transformar a série numa série estacionária.
- "MA"(q): Componente de médias móveis.

No caso das séries temporais estacionárias poderá aplicar-se o modelo ARMA, semelhante ao modelo ARIMA, no entanto, a componente integrada é nula.

Apesar da flexibilidade e da facilidade de interpretar e detetar padrões de dependência temporal em séries temporais, estes modelos apresentam algumas limitações. São exemplos disso a dificuldade de identificação de padrões complexos, a sensibilidade do modelo quanto a valores omissos ou *outliers*, e a dificuldade para obter previsões a longo prazo. Estes modelos são mais adequados para previsão a curto prazo (com poucos intervalos de tempo no futuro). Contudo, quanto maior o intervalo de previsão, menos fiáveis serão as previsões obtidas (Stellwagen e Tashman [2013]). No caso da agitação marítima, o intervalo de previsão adequado para aplicação destes modelos deverá ser de poucas horas.

Para a aplicação de um modelo ARIMA, após a análise exploratória, sugere-se que sejam completados os seguintes passos (Hyndman e Athanasopoulos [2014]):

1. Avaliar a estacionaridade da série: Recomenda-se a visualização do gráfico dos registos em função do tempo. Uma série estacionária geralmente não apresenta padrões previsíveis a longo prazo.
2. Garantir a estacionaridade da série: Se a série não for estacionária, deve ser transformada numa estacionária, a partir por exemplo de diferenças simples (calculando a diferença de valores consecutivos da mesma série $y_t = x_t - x_{t-1}$). A estacionaridade pode ser garantida através de vários testes estatísticos, como por exemplo, o teste *Augmented Dickey-Fuller* (Dickey e Fuller 1979).
3. Selecionar a ordem p e q do modelo ARIMA: Após ter garantido a estacionaridade da série (e definido a componente integrada d) deverão ser determinadas as ordens p e q a partir da análise do gráfico da função de autocorrelação e da função autocorrelação parcial.
4. Testar modelos ARIMA de várias ordens: Considerar aplicar outros modelos ARIMA, selecionando o mais adequado com base no critério AIC (*Akaike's Information Criterion*) ou o critério BIC (*Bayesian Information Criterion*). Embora estes critérios não indiquem qual o modelo que descreve melhor a série temporal, a minimização de AIC ou BIC estabelece um balanço entre o ajustamento (minimização do erro) e a complexidade do modelo.

Para definir os critérios AIC e BIC é essencial compreender o conceito de máxima verossimilhança. Considerando uma amostra $X = (X_1, X_2, \dots, X_n)$ obtida através de uma distribuição com função de densidade de probabilidade $f(x; \vec{\theta})$ (em que $\vec{\theta}$ é o vetor de parâmetros da distribuição), a função de verossimilhança é expressa por:

$$L(\vec{\theta}; X) = \prod_{i=1}^n f(X_i; \vec{\theta}) \quad (2.13)$$

Na prática, é usual considerar o logaritmo da função de verossimilhança, por ser mais fácil de maximizar do que a própria função de verossimilhança. O logaritmo da função de verossimilhança tem a expressão seguinte: $\ell(\vec{\theta}; X) = \log L(\vec{\theta}; X)$, podendo ser expressa da seguinte forma:

$$\ell(\vec{\theta}; X) = \sum_{i=1}^n \log f(X_i; \vec{\theta}) \quad (2.14)$$

O critério AIC é um dos critérios mais utilizados e é expresso aproximadamente da seguinte forma (Chatfield e Xing [2019](#)):

$$\text{AIC} = -2 \ln(\text{máxima verossimilhança}) + 2r \quad (2.15)$$

em que r num modelo ARIMA é $r = p + q + 1$. Contudo, se a amostra for pequena, o critério a utilizar deverá ser o critério AICc, que acrescenta uma correção ao critério ao substituir o termo $2r$ por $\frac{2rn}{n-r-1}$, em que n é o número de observações.

O critério BIC é expresso aproximadamente como:

$$\text{BIC} = -2 \ln(\text{máxima verossimilhança}) + r \ln n \quad (2.16)$$

O termo $r \ln n$ aumenta conforme o número de parâmetros e de observações aumenta, penalizando mais os modelos complexos, especialmente quando a dimensão da amostra (neste caso, o comprimento da série) é um número elevado.

Ambos os critérios AIC e BIC penalizam modelos complexos favorecendo os modelos mais simples que melhor descrevem os dados. Contudo, o critério BIC para uma amostra com um número elevado de registos é mais penalizadora da complexidade do modelo do que o critério AIC. Desta forma, o modelo BIC é normalmente utilizado quando a amostra tem um conjunto elevado de registos e a simplicidade do modelo é essencial.

5. Verificar resíduos através da análise do gráfico dos resíduos em função do tempo: Nesta fase pretende-se identificar se a média dos resíduos é nula e a variância constante.
6. Aplicar o teste de correlação: os resíduos podem ser submetidos a testes de correlação, como por exemplo o teste de Ljung-Box expresso por (Chatfield e Xing [2019](#)) :

$$Q = N \sum_{k=1}^K r_k^2 \quad (2.17)$$

em que N é o número de termos na série diferenciada e K é o número de valores de resíduos que será avaliado conjuntamente a sua correlação, tipicamente este valor varia entre 15 a 30.

7. Calcular as previsões: As previsões devem ser efetuadas quando os resíduos tenham um gráfico semelhante a ruído branco, ou seja quando não apresentem correlação.

2.5 Modelos de Regressão Linear

Existem modelos de regressão linear simples, se a variável dependente for prevista a partir de uma variável independente, ou um modelo de regressão linear múltipla se a variável dependente for prevista por mais do que uma variável independente.

2.5.1 Modelo de Regressão Linear Simples

Admitindo que os resíduos w_t seguem uma distribuição normal com média igual a zero e variância σ_ε^2 então a regressão linear pode ser descrita como (Chatfield e Xing 2019):

$$y_t = \beta_0 + \beta_1 x_t + w_t \quad (2.18)$$

O modelo de regressão linear simples, apresenta uma reta em que β_0 simboliza o valor esperado quando $x = 0$ e β_1 é o declive da reta y_t .

2.5.2 Modelo de Regressão Linear Múltipla

Conforme referido anteriormente o modelo de regressão linear múltipla difere do modelo de regressão linear simples pelo número de variáveis independentes que contribuem para a previsão da variável dependente, podendo ser descritos como:

$$y_t = \beta_0 + \beta_1 x_{t1} + \beta_2 x_{t2} + \cdots + \beta_q x_{tq} + w_t \quad (2.19)$$

em que β_0 simboliza o valor esperado quando todos os x_{tq} são nulos. Os termos $\beta_1, \beta_2, \dots, \beta_q$ são os coeficientes de regressão na qual os parâmetros do modelo multiplicam cada variável independente x_{tq} . Estes modelos têm a vantagem de permitir analisar a importância relativa e o nível de significância dos coeficientes, e portanto dos vários previsores. No entanto, são limitados pela hipótese de que a série temporal de resposta y_t depende linearmente da série temporal x_t segundo a equação 2.19 e que os resíduos w_t são independentes e identicamente distribuídos com distribuição $X \sim \mathcal{N}(\mu, \sigma^2)$.

Uma outra vantagem destes modelos é que são fáceis de analisar e de implementar. No entanto, no caso geral baseiam-se em hipóteses restritivas: as relações entre as variáveis (dependente(s) e independente) são lineares, w_t tem média zero, não são correlacionados (independentes) e seguem uma distribuição normal. Estes modelos são sensíveis a *outliers* e não representam relações não lineares.

Após a análise exploratória dos dados, na implementação do modelo de regressão linear para previsão de uma série temporal deve completar-se as seguintes etapas (Hyndman e Athanasopoulos [2014](#)) :

1. Selecionar as variáveis independentes: Adicionar as várias variáveis e ir retirando uma a uma por forma a avaliar o modelo com base no critério R^2 .
2. Calcular os coeficientes do modelo: Os coeficientes do modelo são selecionados a partir do mínimo da soma dos quadrados dos erros:

$$\sum_{t=1}^K w_t^2 = \sum_{t=1}^K (y_t - \beta_0 - \beta_1 x_{1,t} - \beta_2 x_{2,t} - \dots)^2 \quad (2.20)$$

3. Avaliar o modelo: Conforme mencionado anteriormente no capítulo, o R^2 é bastante utilizado como uma métrica de avaliação do modelo. R^2 representa a variância explicada pelo modelo e os seus valores variam entre 0, se as previsões não estão correlacionadas com os dados, e 1 no caso inverso.
4. Efetuar as previsões: Uma vez determinados os coeficientes que minimiza, os erros. As previsões são obtidas ao calcular o termo y_t através da substituição dos coeficientes na equação [2.19](#).
5. Análise de resíduos: A análise de resíduos, à semelhança dos modelos ARIMA, deve ser feita através da análise do gráfico dos resíduos em função do tempo identificando se a variância dos resíduos é nula. Porém é essencial ser analisado o gráfico da função autocorrelação dos resíduos, como garantia que os mesmos não estão correlacionados.

O modelo descrito pela equação [2.19](#) pode ainda ser generalizado de modo a considerar valores da própria série e de outras séries temporais, generalizando assim, os modelos autoregressivos. Por exemplo, se y_t dependem de duas séries temporais x_1 e x_2 , um modelo de regressão que generaliza o modelo autoregressivo é expresso por:

$$y_t = \beta_0 + \sum_{i=1}^p \beta_i y_{t-i} + \sum_{j=0}^m \gamma_j x_{1(t-j)} + \sum_{k=0}^n \delta_k x_{2(t-k)} + w_t \quad (2.21)$$

Este modelo pressupõe que existe uma relação linear entre os valores sucessivos de y_t e das outras duas séries $x_{1,t}$ e $x_{2,t}$, e pode ser útil para descrever a importância relativa dos valores sucessivos de y_t e das outras duas séries temporais, mas não permite descrever relações complexas não lineares entre $y_t, x_{1,t}, x_{2,t}$ como

é o caso das séries temporais dos parâmetros de agitação marítima entre pontos diferentes, que dependem também da evolução dos elementos meteorológicos. Para descrever essas relações recomenda-se a utilização de redes neurais, que são uma das metodologias de aprendizagem automática potencialmente mais poderosas.

2.6 Modelos de Aprendizagem Automática: Redes Neurais

O termo "Aprendizagem Automática" (*Machine Learning*), pode ser entendido como uma área de inteligência artificial (IA) que se refere ao desenvolvimento de algoritmos que permitem aos computadores aprenderem a realizar tarefas ou fazer previsões a partir de dados (Prince 2023).

Os modelos estatísticos formalizam a relação entre variáveis na forma de equações matemáticas, permitindo assim a compreensão da importância relativa dos previsores. Estes modelos baseiam-se em algumas condições. No caso do modelo de regressão linear, assume-se que existe uma relação linear entre a variável dependente e a(s) variável(is) independente(s). No caso do modelo ARIMA, é essencial garantir a estacionaridade da série temporal. Além disso, estes modelos são apenas treinados com um conjunto de dados (conjunto de dados de treino), sem que os parâmetros do modelo variem de forma adaptativa.

Os modelos de Aprendizagem Automática (AA) representam relações entre variáveis diretamente a partir do conjunto de dados de treino sem assumir pré-condições do tipo dos modelos anteriores, privilegiando a precisão dos resultados em detrimento da interpretação da estrutura do modelo. Os modelos estatísticos são mais interpretáveis que os modelos de aprendizagem automática. No entanto, são geralmente mais limitados para descrever fenómenos complexos (Dangeti 2017).

Os modelos de AA podem ser descritos de forma genérica por uma relação matemática da seguinte forma:

$$\vec{y} = f[\vec{x}, \vec{\phi}] \quad (2.22)$$

As variáveis de entrada $\vec{x}$ (também denominadas como previsores ou *features*) são as variáveis independentes utilizadas como valores de entrada para treinar o modelo de AA. As variáveis de saída ou de resposta (*target*) $\vec{y}$ são as variáveis que o modelo de AA pretende prever. Numa base de dados referente aos valores de entrada, as colunas podem ser consideradas como *features* ou atributos. Numa

rede neuronal, o número de *features* determina por norma o número de neurónios utilizados na camada de entrada (Bhattacharjee [2017](#)) da rede.

Estes modelos podem resolver vários problemas. Os mais comuns são os de regressão, de classificação, de classificação de aglomerados (*clustering*) e de associação (Burkov [2019](#)):

- Problemas de regressão: Os problemas de regressão caracterizam-se pelas variáveis de resposta assumirem um valor contínuo.
- Problemas de classificação: Nos problemas de classificação, a variável de resposta é uma categoria ou classe.
- Problemas de classificação de aglomerados (*clustering*): Os problemas de classificação de aglomerados têm como objetivo agrupar registos em grupos com base em semelhanças.
- Problemas de associação: Nos problemas de associação o objetivo é encontrar regras que descrevem os registos de entrada.

A modelação a partir dos dados é um processo iterativo, com o objetivo de alcançar uma boa descrição do fenómeno em questão. Esta é constituída por duas fases distintas: a fase preliminar, que inclui a formulação do problema e o pré-processamento dos dados, que são comuns à aplicação dos modelos estatísticos tradicionais; e a fase de aprendizagem, na qual se pretende determinar o modelo cuja estrutura e parâmetros melhor descrevem a relação funcional entre os previsores (*features*) $\vec{x}$ e a resposta $\vec{y}$ (Bontempi [2021](#)). Na fase preliminar, a base dados deve ser dividida em dois conjuntos distintos: conjuntos de treino e de validação (utilizados no processo de aprendizagem). O conjunto de treino é normalmente de maiores dimensões, porque o treino do modelo deve considerar o máximo de condições representativas da relação entre variáveis, o que só é possível com uma dimensão da amostra ou comprimento da série temporal adequados.

O conjunto de validação utilizado para possíveis ajustes do algoritmo e o conjunto de teste, para a avaliação do modelo, têm normalmente dimensões menores. A divisão destes conjuntos pode respeitar diferentes critérios, como por exemplo, a validação cruzada, a regra 70%, 15%, 15% ou no caso de haver muitos dados disponíveis a separação entre 95%, 2, 5%, 2, 5% (Burkov [2019](#)). A segunda etapa do processo de modelação é o momento de aprendizagem que se inicia com a introdução dos dados de treino.

A natureza do problema (regressão, classificação, *clustering*, associação) permite dividir aprendizagem em três métodos: a aprendizagem supervisionada, não supervisionada e semi-supervisionada e aprendizagem por reforço:

- Na aprendizagem supervisionada, o conjunto de dados é a coleção de exemplos rotulados $\{(x_i, y_i)\}_{i=1}^N$, em que cada elemento x_i entre N é um vetor de *features*. Um vetor de *features* é um vetor no qual cada dimensão $j = 1, \dots, D$ contém um valor que descreve o exemplo. O objetivo deste tipo de aprendizagem é utilizar o conjunto de dados de treino para produzir um modelo que toma um vetor de *features* como entrada e calcula uma resposta que permite deduzir o rótulo para esse vetor. Esta forma de aprendizagem normalmente é aplicada a problemas de regressão e classificação.
- Na aprendizagem não supervisionada, ao contrário do que acontece na aprendizagem supervisionada, a base de dados é constituída por um conjunto de exemplos que não estão rotulados $(x_i)_{i=1}^N$. Esta forma de aprendizagem cria um modelo que utiliza o vetor das *features* $\vec{x}$ como variáveis de entrada e transforma num vetor ou num valor que pode ser utilizado como resposta ao problema, é normalmente aplicada a problemas de associação.
- A aprendizagem semi-supervisionada ocorre quando a base de dados, contém exemplos classificados e outros não classificados. O objetivo deste tipo de aprendizagem é o mesmo da aprendizagem supervisionada, contudo o facto de algumas das variáveis de entrada não estarem classificados poderá permitir que se obtenha um melhor modelo, por ser possível desta forma aproveitar a informação adicional disponível nos dados não rotulados, aumentando a quantidade e a variedade de dados utilizados no treino do modelo.
- Na aprendizagem por reforço existe um agente, como por exemplo um robô, que está em contacto com um ambiente dinâmico, na qual as decisões que efetua são penalizadas ou recompensadas com um *feedback*. O objetivo é que o agente maximize a recompensa cumulativa ao longo do tempo (Prince 2023). Este tipo de aprendizagem é frequentemente utilizado para otimização de sistemas, jogos ou robótica (Burkov 2019).

A previsão dos parâmetros de agitação marítima é um problema de regressão e por isso exige que a forma de aprendizagem seja supervisionada.

Os modelos de aprendizagem supervisionada permitem fazer previsões de uma ou mais variáveis de saída $\vec{y}$, a partir de uma ou várias variáveis de entrada $\vec{x}$.

2.6.1 Processo de Aprendizagem

A fase de aprendizagem é caracterizada pela seleção do modelo, cuja estrutura e parâmetros que representam a melhor relação funcional entre $\vec{x}$ e $\vec{y}$ (Bontempi 2021). O processo de treino de um modelo tem como objetivo identificar a estrutura e definir os parâmetros representados como componente do vector $\vec{\phi}$, que descrevem a melhor relação entre as variáveis de entrada e saída. A determinação dos valores ótimos destes parâmetros é obtida a partir da utilização de um conjunto de dados de treino de um par de variáveis de entrada e saída x_i, y_i , recorrendo a algoritmos especializados, descritos adiante no capítulo.

O grau de ajustamento no mapeamento é quantificado através de uma função custo L , a qual toma um valor escalar que resume o erro total das previsões do modelo relativamente aos valores correspondentes da resposta (Prince 2023).

O valor ótimo dos parâmetros, designado por $\vec{\phi}$, é obtido no mesmo processo de treino através da minimização da função de custo, ou seja:

$$\vec{\phi} = \arg \min_{\vec{\phi}} L[\vec{\phi}] \quad (2.23)$$

Se o custo for reduzido, então o modelo consegue prever com um erro reduzido (elevada precisão) os valores de saída $\vec{y}$ a partir dos valores de entrada $\vec{x}$ (Prince 2023).

O processo de aprendizagem de redes neuronais MLP é dividido em duas fases distintas. A primeira fase denominada de *feed forward propagation* é caracterizada pela passagem das variáveis de entrada nas várias camadas da rede desde a camada de entrada até à camada de saída, passando pelas camadas intermédias (processo descrito com detalhe mais adiante no capítulo) de modo a que a rede gere previsões. A segunda fase do processo de aprendizagem designada de *backpropagation* consiste na atualização dos pesos sinápticos das redes neuronais desde a camada de saída até à camada de entrada, em função dos erros obtidos das previsões, por forma a minimizar a função de custo. O algoritmo de *backpropagation* é baseado no gradiente descendente, com algumas modificações destinadas a evitar que os parâmetros da rede convirjam para um mínimo local (este processo será descrito de forma mais pormenorizada nas secções seguintes).

2.6.2 Redes Neurais

As redes neuronais são modelos computacionais inspirados nos neurónios biológicos (células do cérebro que processam e transmitem informação por meio de impulsos elétricos). Uma rede neuronal artificial é um modelo computacional que tenta replicar o processo biológico num ambiente computacional (I. d. Silva, Spatti, Flauzino, Liboni e Reis Alves [2016](#)).

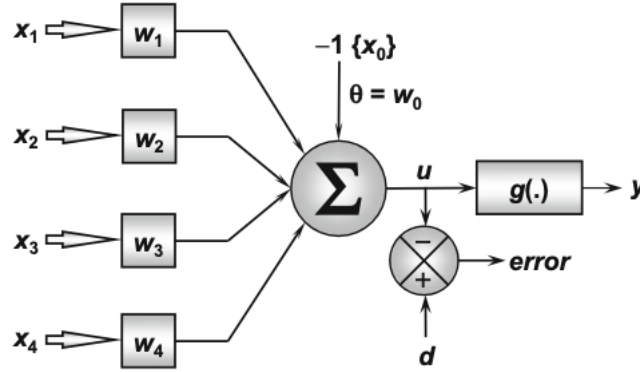


FIGURA 2.1: Representação de um neurónio. Fonte: I. d. Silva, Spatti, Flauzino, Liboni e Reis Alves [2016](#)

Uma rede neuronal é composta por neurónios artificiais, que são habitualmente denominados como nós. A figura [2.1](#) representa esquematicamente a estrutura de um neurónio. Cada neurónio é activado por um conjunto de sinais de entrada (*input*) ($x_1, x_2, x_3, \dots, x_n$), que são associados a pesos sinápticos ($w_1, w_2, w_3, \dots, w_n$), os quais quantificam a importância de cada sinal para a ativação do neurónio. Os sinais de entrada são agregados e sujeitos a uma soma ponderada, produzindo-se assim um potencial de ativação. O potencial de ativação é comparado com o limite de ativação (θ), produzindo-se o sinal de saída (u). De modo a garantir que o sinal de saída do neurónio permaneça dentro do limite de valores aceitáveis, este é submetido a uma função de ativação ($g(\cdot)$), que determina o comportamento não-linear do neurónio podendo converter o sinal final, por exemplo, num resultado contínuo ou binário. Por fim, o sinal de saída do neurónio pode ser utilizado como sinal de entrada para outros neurónios, estabelecendo-se assim uma rede neuronal. Os resultados produzidos por um neurónio artificial podem ser formalmente expressos como (I. d. Silva, Spatti, Flauzino, Liboni e Reis Alves [2016](#)):

$$y = g(u) \quad (2.24)$$

$$u = \sum_{i=1}^n w_i \cdot x_i - \theta \quad (2.25)$$

2.6.3 Funções de ativação

Da mesma forma que os neurónios no cérebro humano são ativados por sinais elétricos de outros neurónios, os neurónios artificiais são ativados por valores numéricos que dependem dos inputs e dos pesos associados aos mesmos. As funções de ativação são responsáveis por regular o grau de ativação do neurónio ao ajustar o valor da saída do neurónio de forma não linear permitindo que a rede represente relações complexas. Um modelo de redes neuronais pode ser definido com qualquer função de ativação g desde que esta seja diferenciável em todo o seu domínio com a eventual exceção de alguns pontos. Outras duas importantes propriedades das funções de ativação são:

- O cálculo da função da derivada deve ser eficiente;
- Deve conter um contradomínio restritivo para ser possível a regularização do nível de ativação do neurónio.

Algumas das funções mais comuns são a função logística (sigmoid), a tangente hiperbólica (tanh) e a unidade linear retificada (ReLU). Esta última é normalmente mais utilizada (Prince [2023](#)) devido à sua simplicidade computacional. Dependendo do problema poderão ser utilizadas diferentes funções de ativação.

Função logística

A função logística é uma função diferenciável em todo o seu domínio. Esta função é utilizada como função de ativação por transformar a soma ponderada dos valores de entrada do neurónio em outros valores de saída que podem ser interpretadas como uma probabilidade se for aplicada na última camada da rede, ou interpretada como uma ativação não-linear quando aplicada a camadas intermédias. O contradomínio da função logística é o intervalo $]0, 1[$. A função logística é diferenciável em todo o seu domínio $] - \infty, +\infty[$. Esta função é definida pela expressão:

$$\text{sigmoid}(z) = \frac{1}{1 + e^{-z}} \quad (2.26)$$

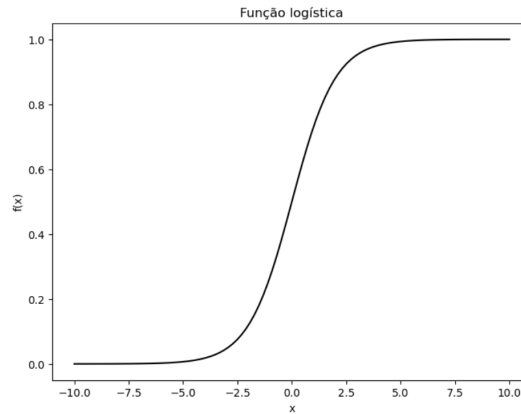


FIGURA 2.2: Representação da função logística $\text{sigmoid}(z)$

Função tangente hiperbólica

Esta função também é diferenciável em todo o seu domínio $]-\infty, +\infty[$ e o seu contradomínio é o intervalo $] -1, 1[$. A utilização desta função pode ser vantajoso em problemas de classificação binária. Esta função pode ser descrita pela seguinte expressão:

$$\tanh(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}} \quad (2.27)$$

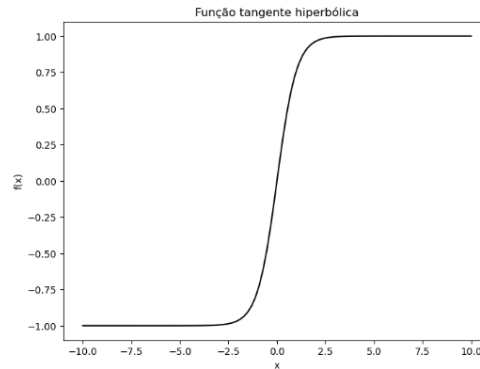


FIGURA 2.3: Representação da função tangente hiperbólica $\tanh(z)$

Função unidade linear retificada (ReLU)

Ao contrário das funções logística e tangente hiperbólica que têm o contradomínio limitado, a função ReLU assume o intervalo de valores $[0 + \infty[$, tendo vantagem na descrição dos valores de saída de variáveis como a altura significativa espectral e o período, que tomam apenas valores positivos, mas não são à partida

limitadas superiormente. O intervalo de valores das saídas dos neurónios é que influenciam a escolha das funções de ativação.

$$\text{ReLU}(z) = \begin{cases} 0 & \text{if } z < 0 \\ z & \text{otherwise} \end{cases} \quad (2.28)$$

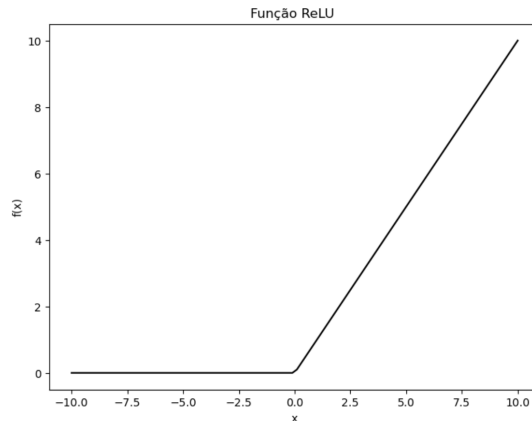


FIGURA 2.4: Representação da função $\text{ReLU}(z)$

2.6.4 Redes neurais com várias camadas: Aprendizagem profunda

O neurónio (nó), representado esquematicamente na figura [2.1](#), é o elemento essencial das redes neurais, as quais são constituída por várias camadas de neurónios, com os neurónios de cada camada ligadas aos da camada seguinte. Mais especificamente, a estrutura de uma rede neuronal compreende (I. d. Silva, Spatti, Flauzino, Liboni e Reis Alves [2016](#)):

- A camada de entrada: É a camada que recebe os valores das variáveis de entrada. Por norma estes valores são normalizados com o objetivo de tornar os cálculos mais eficientes, e o processamento mais rápido.
- Uma ou mais camadas escondidas ou ocultas (*hidden layers*): Camada entre a camada de entrada e a de saída. Pode existir mais que uma camada oculta.
- A camada de saída: À semelhança das camadas anteriores, a última camada é também composta por neurónios. Esta camada produz o resultado final que

resulta do processamento das camadas anteriores. Se a variável de resposta, é um escalar, a camada de saída tem um único neurónio.

Uma rede neuronal pode ser representada com apenas uma camada de entrada e de saída. Contudo, incluindo uma ou mais camadas escondidas, a rede neuronal torna-se capaz de descrever relações extremamente complexas e não lineares entre as variáveis de entrada e saída. No caso de uma rede neuronal ter mais que uma camada intermédia, o processo de treino e validação denomina-se "aprendizagem profunda" (*Deep Learning*) (Prince 2023).

A extensão da aplicação das redes neuronais a várias tipologias de problemas, como por exemplo reconhecimento de imagem, deteção de emails fraudulentos, reconhecimento facial, entre outros, motivou o desenvolvimento de diversas arquiteturas de redes neuronais adaptadas ao tipo de problema. Neste capítulo apenas serão apresentadas as duas arquiteturas de redes mais utilizadas para previsão dos parâmetros e agitação marítima: *Multilayer Perceptron* e *Recurrent Neural Network*. Conforme será descrito adiante, as redes *Multilayer Perceptron* são compostas por várias camadas de neurónios que não têm em consideração a estrutura sequencial e temporal dos valores de entrada, ao passo que as redes *Recurrent Neural Network* têm um processo de retroalimentação incluído que permite à rede processar sequências de dados.

2.6.5 *Multilayer Perceptron*

Uma das arquiteturas mais utilizadas em redes neuronais é a *Multilayer Perceptron* (MLP), representada na figura 2.5. É composta por duas ou mais camadas intermédias e n sinais de entrada. A rede MLP é conhecida pela aplicação sucessiva de duas etapas distintas. Na primeira etapa, designada de *feed forward propagation*, os sinais de entrada $(x_1, x_2, x_3, \dots, x_n)$ de um conjunto de treino são inseridos na camada de entrada da rede e "propagadas" através de camadas sucessivas até à produção do resultado na camada de saída. Nesta etapa, o resultado é obtido considerando os valores atuais dos pesos sinápticos. Os resultados produzidos pela rede são comparados com os correspondentes resultados da variável de resposta existentes no conjunto de treino (aprendizagem supervisionada) e calculado o erro respetivo, que servirá para ajustar os valores dos pesos sinápticos. Na segunda etapa, que ocorre devido à existência de erros na primeira etapa, é aplicado um algoritmo de *backpropagation*. Nesta fase, os pesos são ajustados em função do erro, através do método do gradiente descendente, na direção contrária à da etapa anterior, isto é, da última para a primeira camada, de modo a diminuir o valor da

função de custo (secção 2.6.1). As derivadas necessárias para o cálculo do gradiente descendente são obtidos a partir da regra da cadeia do cálculo diferencial (I. d. Silva, Spatti, Flauzino, Liboni e Reis Alves 2016; Prince 2023).

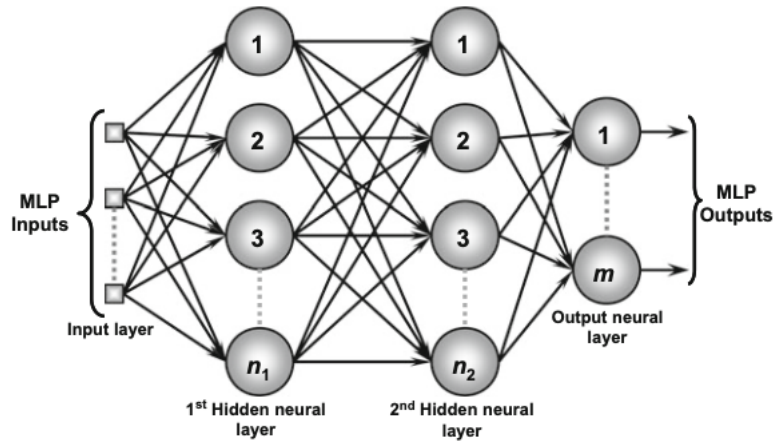


FIGURA 2.5: Representação da arquitetura de redes *Multilayer Perceptron*. Neste caso, a camada de saída contém m neurónios, sendo a saída (*output*) um vetor de comprimento m . Fonte: I. d. Silva, Spatti, Flauzino, Liboni e Reis Alves 2016

2.6.6 Redes neurais recorrentes (*Recurrent Neural Network*)

Neste tipo de rede, as saídas são combinadas com as entradas, de modo a produzir uma combinação entre a informação recente (das novas entradas) e o estado anterior do neurónio (retroalimentação). Numa rede recorrente, existem neurónios que recebem não apenas dados de entrada da camada anterior como também a sua própria saída. Este processo de retroalimentação permite à rede neuronal processar seqüências de dados. Por essa razão, a aplicação destas redes neurais é adequada para problemas que incluem séries temporais. A figura 2.6 representa um exemplo de uma rede recorrente.

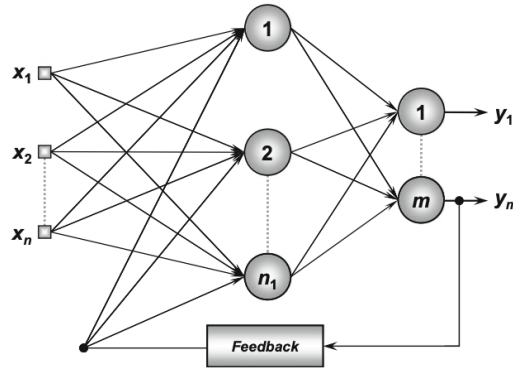


FIGURA 2.6: Representação da arquitetura de redes neurais recorrentes (*Recurrent Neural Network*). Fonte: I. d. Silva, Spatti, Flauzino, Liboni e Reis Alves [2016](#)

No treino de redes neurais profundas, e em particular em redes recorrentes é necessário ter em consideração que poderá ocorrer o problema do "gradiente desvanescente" ou "gradiente explosivo", que afetam o desempenho da rede neuronal ou podem impedir o seu treino (gerando paragens de execução por *overflow*) (Goodfellow, Bengio e Courville [2016](#)).

Durante o treino de uma rede neuronal, na qual se inclui o algoritmo de *backpropagation* ou o processo de retroalimentação de um neurónio, os pesos do neurónio são atualizados de modo a minimizar o valor da função de custo L . Esta atualização dos pesos sinápticos dos neurónios, desde a camada de saída até a camada de entrada, é feita através do cálculo do gradiente da função de custo L em relação a cada peso da rede w_{ij} , que é obtido através da regra da cadeia expresso na seguinte forma:

$$\frac{\partial L}{\partial w_{ij}} = \frac{\partial L}{\partial \vec{y}} \cdot \frac{\partial \vec{y}}{\partial \vec{u}} \cdot \frac{\partial \vec{u}}{\partial w_{ij}} \quad (2.29)$$

O problema do gradiente desvanescente (*vanishing gradient*), ocorre quando os gradientes das funções de ativação nas camadas iniciais da rede se tornam extremamente pequenos, isto é menores que 1, o que faz com que a atualização se torne mais lenta e o treino ineficaz. Por outro lado, quando os gradientes das funções de ativação nas camadas iniciais da rede se tornam muito elevados (superiores a 1), tornam instável o processo de treino, impedindo que este seja completado (Goodfellow, Bengio e Courville [2016](#); Burkov [2019](#)).

Quanto mais profunda for a rede, isto é, quanto maior for o número de camadas intermédias, mais exacerbado será o problema. As redes neurais profundas

são propensas a estes problemas, porque os valores das componentes dos gradientes serão propagados ao longo das várias camadas existentes, provocando ou uma diminuição ou um aumento exponencial dos pesos sinápticos.

As redes neuronais LSTM (*Long Short-Term Memory*) e GRU (*Gated Recurrent Unit*) são subtipos de redes neuronais recorrentes, capazes de processar sequências de dados, e que surgem como estratégias para evitar o problema do gradiente desvanescente, pela sua capacidade de regular o fluxo de informação (Prince 2023; Goodfellow, Bengio e Courville 2016).

As redes LSTM são caracterizadas por terem neurónios com dois “caminhos” possíveis para o fluxo de informação (figura 2.7), um correspondente a memória de curto prazo semelhante ao utilizado por redes neuronais recorrentes tradicionais em que o sinal é influenciado por pesos e atualizado a cada instante temporal, e outro caminho alternativo associado à memória de longo prazo, no qual funções sigmoideais ou tangentes hiperbólicas funcionam como interruptores no sentido em que determinam a percentagem de informação que é lembrada, atualizada e esquecida (Goodfellow, Bengio e Courville 2016). Uma célula LSTM é composta pelos seguintes módulos:

- *Forget gate*: Controla a percentagem de informação da memória a curto de prazo que deve ser eliminada;
- *Input gate*: Controla a informação que é adicionada à memória de longo prazo;
- *Output gate*: Determina a informação que será utilizada para atualizar a memória a curto prazo e gerar o valor de saída do neurónio.

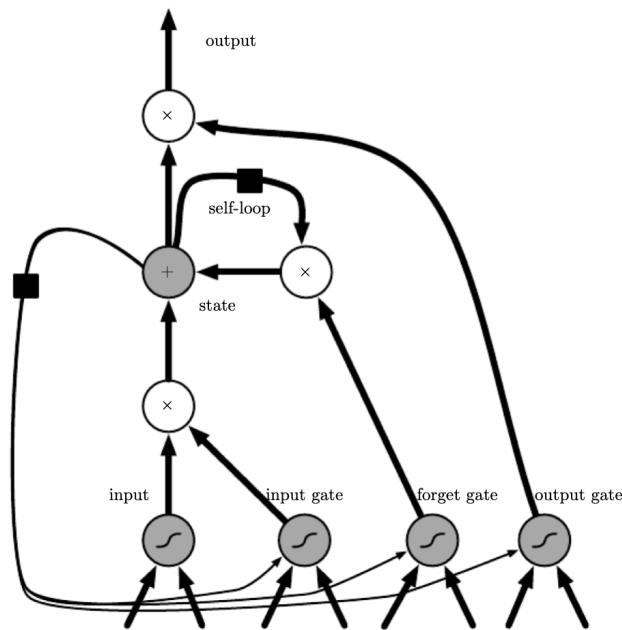


FIGURA 2.7: Representação da arquitetura de um neurônio de uma rede neuronal LSTM (*Long Short-Term Memory*). Fonte: Goodfellow, Bengio e Courville [2016](#)

As redes GRU são semelhantes às redes LSTM, mas, ao contrário dos neurónios das redes LSTM, os neurónios destas redes são mais simples, por terem apenas um interruptor que controla simultaneamente a percentagem de informação esquecida e a decisão de adicionar e atualizar a informação. Estes neurónios são constituídos por dois interruptores, são eles (Goodfellow, Bengio e Courville [2016](#)):

- *Reset gate*: Determina qual a percentagem de informação que será esquecida e simultaneamente, a quantidade de informação que servirá de atualização para o valor de saída do neurónio.
- *Actualization gate*: Controlam a informação que é transmitida com base no resultado obtido através do interruptor *Reset gate*. A informação pode ser copiada ou ignorada filtrando assim o resultado do valor de saída.

2.7 Modelos propostos para previsão das condições de agitação marítima baseados em Aprendizagem Automática

Recentemente têm sido propostos modelos de AA para a previsão de parâmetros de agitação marítima, com o objetivo de oferecer uma resposta com menor exigência computacional e com uma fiabilidade competitiva aos modelos físico-matemáticos. Na última década tem crescido a aplicação destes modelos em aplicações de engenharia costeira e oceanográfica, com particular foco na previsão da altura das ondas (Kumar, Savitha e Mamun [2018]). Estes modelos utilizam registos meteo-oceanográficos recolhidos a partir de boias ondógrafo e de sensores de pressão ou acústicos, como variáveis de entrada para treino e validação dos modelos.

As redes neuronais têm sido amplamente utilizadas para a previsão da altura significativa da onda (Deo, Jha, Chaphekar e Ravikant [2001]; Rao e Mandal [2005]; Tolman, Krasnopolsky e Chalikov [2005]; Browne, Castelle, Strauss, Tomlinson, Blumenstein e Lane [2007]), oferecendo uma boa alternativa à regressão estatística, à análises de séries temporais e à utilização de métodos numéricos (Javad Mahjoobi e Adeli [2009]).

Contudo, alguns investigadores concluíram que o fenómeno de *overfitting*⁵ é uma limitação destes métodos. Além disso, exigiam também muito tempo no treino e na exploração dos parâmetros das redes, nomeadamente no número de camada escondidas (Javad Mahjoobi e Adeli [2009]).

Foram também sugeridos, a utilização de redes neuronais com algoritmos de *feed-forward* (O. Makarynsky [2004]; Deo, Jha, Chaphekar e Ravikant [2001]), *recurrent* (Tayeb Sadeghifar e Mahdizadeh [2017]) e *cascade-correlation* (O. Makarynsky [2004]).

Contudo, estes estudos apontaram ainda algumas limitações. No estudo elaborado por Oleg Makarynsky [2005], foram utilizadas duas redes neuronais para previsão dos parâmetros de agitação H_s e T_p , para períodos de 3, 6, 12 e 24 horas, recorrendo a observações da boia ondógrafo de Sines no período compreendido entre 4-12-1999 00:00:00 e 31-01-2000 21:00:00. Estas redes, que recorrem apenas aos registos da boia de Sines intervalados em 3 horas, previram com grande precisão a

⁵Fenómeno que ocorre quando os modelos estatísticos ou de AA têm um bom desempenho na previsão dos dados de treino, mas um mau desempenho nos dados de validação ou de teste (Burkov [2019])

variável H_s para o período de 3 a 6 horas com RMSE= 0,20 m. Porém, nas previsões a 12 e 24 horas a rede neuronal teve uma performance inferior RMSE= 0,8 m. Além disso, a rede neuronal preditiva de T_{02} teve uma performance inferior à rede neuronal cuja variável de saída é H_s , com valores de RMSE aproximadamente de 1.

Tayeb Sadeghifar e Mahdizadeh [2017] utilizaram, no seu estudo, quatro bases de dados relativas ao Mar Cáspio, que correspondiam aos valores da altura significativas previstas intervaladas por períodos de 3, 6, 12, 24 horas no qual obteve previsões com grande precisão até os períodos de 12 horas (RMSE= 0,39; 0,44; 0,38; 0,32 m). Os resultados apresentados por este estudo apontam que as redes recorrentes aplicadas na previsão da altura significativa do Mar Cáspio são mais precisas a prever ondas subsequentes (até ao período de 12 horas de previsão).

Alguns estudos propõem a utilização de outros modelos de AA para o problema de previsão da altura da onda, nomeadamente na utilização de sistemas Fuzzy, de Support Vector Machines (SVM).

Javad Mahjoobi e Adeli [2009] apontam que os modelos SVM, apesar de necessitarem de menos parâmetros e esforço computacional do que os modelos de redes neurais, são mais complexos. Além disso, não permitem tratar dependências sequenciais, o que representa uma desvantagem relativamente a alguns tipos de redes neurais na aplicação destas redes a problemas de séries temporais.

J. Mahjoobi e Etemad-Shahidi [2008] e Demetriou, Michailides, Papanastasiou e Onoufriou [2021] sugerem a utilização de árvores de decisão como alternativa à utilização das redes neurais. O resultado da aplicação destes modelos resultou numa precisão próxima à obtida em trabalhos anteriores com a utilização de redes neurais. Contudo, Demetriou, Michailides, Papanastasiou e Onoufriou [2021] concluíram que o modelo de árvores de decisão, quando otimizado (isto é, quando eliminado os previsores menos relevantes), tornava-se mais sujeito a *overfitting* que as redes neurais.

O presente trabalho pretende desenvolver um modelo de AA com recurso a redes neurais MLP para previsão da altura significativa espectral H_{m0} a 12 horas, utilizando informação da boia multiparamétrica de Sines e de duas séries temporais extraídas da manha geográfica do modelo *Wave Watch III* numa localização próxima da Ilha da Madeira e do arquipélago dos Açores. A localização destes pontos foi estimada a partir do critério que será descrito no capítulo seguinte.

Capítulo 3

Metodologia de desenvolvimento

Este capítulo descreve a metodologia de desenvolvimento do presente trabalho subdividindo-se em quatro secções. A primeira secção diz respeito ao conceito de desenvolvimento do trabalho, na qual serão descritos os critérios para a seleção das fontes de dados para treino do modelo de AA. Na segunda secção serão descritos os parâmetros característicos de agitação marítima que foram considerados no presente trabalho. A terceira secção descreve as séries temporais em estudo, o pré-processamento necessário para a sua utilização nos modelos estatísticos e de AA, bem como uma breve análise exploratória das mesmas. Por fim, a última secção descreve os modelos estatísticos e de Aprendizagem Automática desenvolvidos.

3.1 Conceito de desenvolvimento

A escolha da zona de Sines foi motivada pelo interesse operacional da área de Pinheiro da Cruz para a realização de exercícios de operações anfíbias. Por outro lado, estão também disponíveis séries temporais recolhidas por uma boia multiparamétrica fundeada nesta área. Os registos desta boia permitiram a aplicação de modelos estatísticos tradicionais e também da aplicação de modelos de AA, a partir de processos de aprendizagem supervisionada, quando combinada com os registos obtidos do modelo *Wave Watch III*.

Conforme mencionado no capítulo [1](#), a geração das ondas de vento é provocada por tempestades locais em zonas distantes de costa (zonas de *fetch*) em forma de vaga. Quando o vento deixa de transferir energia para as ondas, estas dispersam-se em frequência e em direção, propagando-se à velocidade de grupo para fora da zona de geração como ondulação.

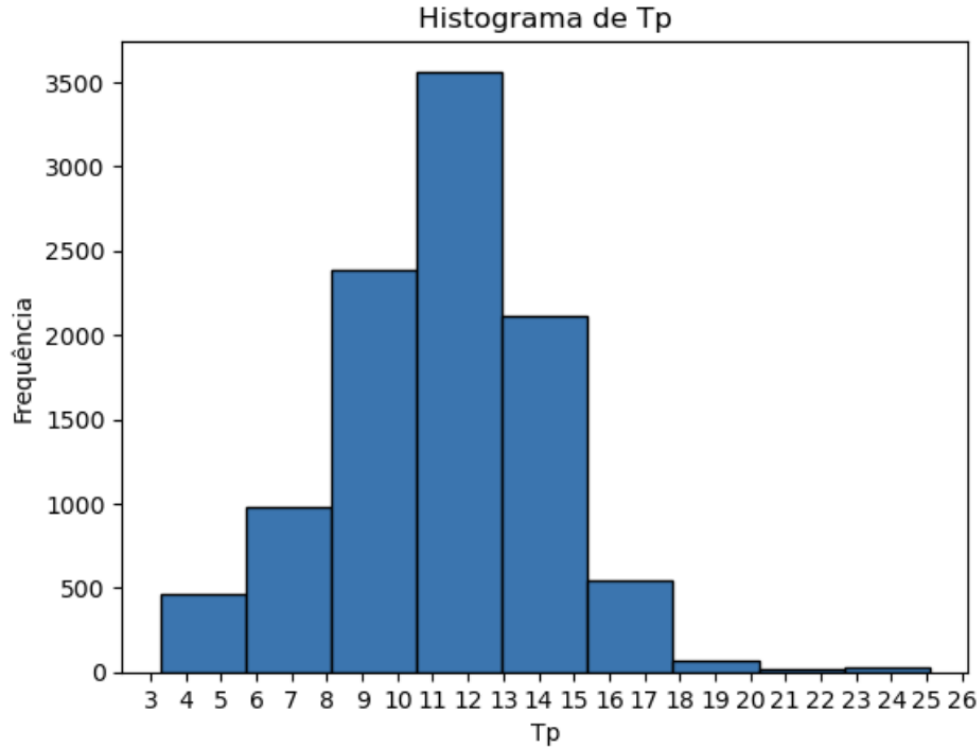


FIGURA 3.1: Histograma da variável T_p

A partir da análise dos resultados climatológicos nas áreas de interesse portuguesas desenvolvidas por Sousa [2023], identificou-se áreas de ondulação (figura 3.3), com o objetivo de calcular a localização de pontos a serem extraídos da malha geográfica do modelo *WaveWatch III*, nos quais pudessem ser extraídos séries temporais que servissem de treino para o modelo de regressão linear e o modelo de AA.

A partir da avaliação do histograma (figura 3.1) considerou-se que os três períodos de pico mais frequentes seriam $T_p = 9; 12; 14s$ que correspondem a velocidades de grupo e direção distintas. A direção de propagação destes grupos de ondas foi estimada com base na análise do gráfico de caixa e bigodes da direção principal das ondas em função do período de pico (figura 3.2). As três direções de propagação apresentadas neste gráfico (Sudoeste (SO), Oeste (O) e Noroeste (NO)) são as direções estimadas como zonas possíveis de geração de ondas pela análise dos resultados obtidos do trabalho de Sousa [2023]. O cálculo da velocidade de propagação do grupo de ondas descritas por estes períodos de pico ($T_p = 9; 12; 14s$) baseou-se na equação 1.14 expressa no capítulo 1.

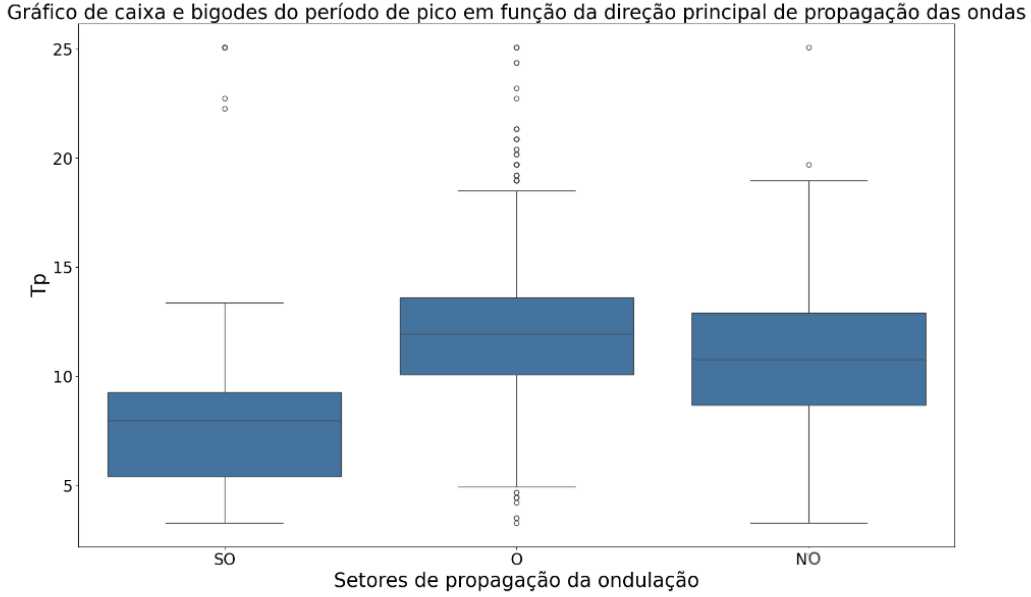


FIGURA 3.2: Distribuição do período de pico em função da direção principal de propagação da ondulação, agregada em setores. Registos referentes à bóia multiparamétrica de Sines.

As semicircunferências representadas na figura 3.3 caracterizam a localização estimada da ondulação para cada um dos períodos de pico selecionados a 24 e 48 horas antes de atingirem a localização da boia de Sines. A estimativa da localização dos grupos de ondas considerou não só a velocidade de grupo das mesmas como também a sua direção de propagação. De acordo com a figura 3.2, períodos de pico menores propagam-se até à boia de Sines provenientes de uma direção de Sudoeste, enquanto que períodos de pico maiores de uma direção de Oeste. Assim, estimou-se a localização dos grupos de onda associados aos períodos de pico ($T_p = 9; 12; 14s$) com base na distância conhecida entre os pontos Noroeste, Oeste e Sudoeste retirados da malha geográfica do modelo numérico e a localização da boia de Sines. Ao período 9 segundos associou-se a direção de Sudoeste, ao período de 12 segundos associou-se a direção de Noroeste e ao período de 14 segundos associou-se a direção de Oeste.

Conforme representado na legenda da figura 3.3, as semicircunferências a tracejado representam para os vários períodos de pico, a localização do grupo de ondas a um período de previsão de 48 horas ($t-48$), enquanto as de traço descontínuo representam a localização dos mesmos grupos de ondas para o período de previsão de 24 horas ($t-24$).

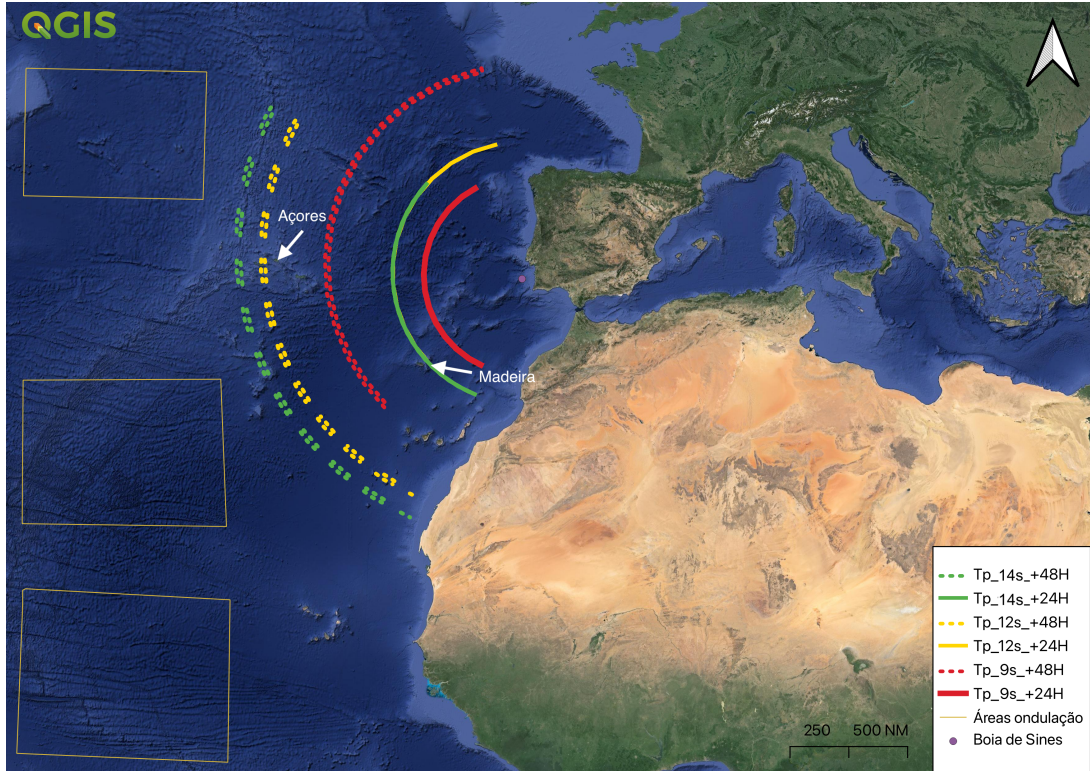


FIGURA 3.3: Representação da localização do grupo de ondas associado aos períodos de pico 9, 12 e 14 segundos para o período de 24 e 48 horas.

A análise da figura 3.3, indica que, próximo da Ilha da Madeira e do arquipélago dos Açores, encontram-se os grupos de ondas correspondentes aos períodos de pico de 12 e 14 segundos, para $t - 24$ no caso da Ilha da Madeira e de $t - 48$ no caso do arquipélago dos Açores.

As séries temporais extraídas ao largo da Ilha da Madeira correspondem aos vários parâmetros de agitação marítima fornecidos pelo modelo numérico WWIII (mencionados na tabela 3.2) para um período de 24H anterior aos registos da bóia multiparamétrica de Sines. Enquanto que as séries temporais extraídas do ponto localizado ao largo dos Açores, correspondem aos parâmetros de agitação marítima referentes ao período anterior de 48H em relação aos registos da boia multiparamétrica de Sines. Desta forma, foi possível construir uma base de dados para um modelo de AA treinado a partir de aprendizagem supervisionada, com informação relativa aos vários parâmetros de agitação marítima nos períodos correspondentes t , $t - 24H$ e $t - 48H$.

O presente trabalho tem como objetivo principal determinar se os modelos

de AA permitem prever de forma fiável as condições de agitação marítima, nomeadamente a altura significativa espectral H_{m0} . Para isso serão aplicados também modelos estatísticos tradicionais, nomeadamente o modelo ARIMA e o modelo de regressão linear, que permitirão interpretar o processo de previsão estatística destes parâmetros.

Inicialmente aplicar-se-á o modelo ARIMA, que é amplamente utilizado na análise de séries temporais, na medida em que a própria série temporal das variáveis H_{m0} é suficiente para prever a sua evolução. Além disso, os resultados obtidos por este modelo permitirão determinar a extensão temporal dos dados históricos que serão relevantes para previsões precisas.

Após esta análise, serão aplicados modelos de regressão linear com recurso às séries temporais de T_p e H_{m0} da boia e dos pontos extraídos do modelo WWII ao largo da Ilha da Madeira e do arquipélago dos Açores, que permitirão identificar os previsores adicionais (para além de H_{m0}) mais significativos na previsão da altura significativa espectral H_{m0} na boia de Sines.

O último modelo a ser aplicado será o modelo de AA com recurso a redes neurais MLP que será treinado com os previsores considerados mais significativos no desenvolvimento do modelo de regressão linear e considerando ainda a extensão temporal necessária dos dados valores anteriores da série temporal a prever sugerida pelos resultados do modelo ARIMA.

3.2 Parâmetros Característicos da Agitação Marítima e Elementos Meteorológicos

A boia multiparamétrica regista parâmetros de agitação marítima, como também alguns elementos meteorológicos. A tabela [3.1](#) descreve os parâmetros característicos de agitação marítima e os elementos meteorológicos extraídos da boia multiparamétrica de Sines.

Variável	Parâmetro
VTPK	Período de pico
VHM0	Altura significativa espectral ($\approx 1.05H_s$)
VTM02	Período da onda correspondente ao segundo momento espectral ($\approx T_m$)
VMDR	Direção média do vento
VPED	Direção principal das ondas referente no período de pico
WDIR	Direção do vento
WSPD	Velocidade do vento

TABELA 3.1: Variáveis selecionadas a partir dos registos da boia multiparamétrica.

A partir do modelo físico-matemático *WaveWatch III* é possível extrair um conjunto variado de parâmetros. Contudo, por forma a otimizar o processamento foram extraídos apenas os parâmetros essenciais que se apresentam na tabela 3.2.

Variável	Parâmetro
Tp	Período de pico
Hs	Altura significativa
Pdir	Direção principal das ondas no período de pico
Tm02	Período da onda correspondente ao segundo momento espectral ($\approx T_m$)
Uwnd	Componente meridional do vento (Oeste - Este)
Vwnd	Componente vertical do vento (Norte – Sul)
Tm	Período médio da onda

TABELA 3.2: Variáveis selecionadas a partir dos registos do modelo *WaveWatch III*.

3.3 Descrição, Preparação Prévia e Análise Exploratória das Séries Temporais

3.3.1 Descrição das séries temporais em estudo

A caracterização estatística dos dados é um passo fundamental para compreender as propriedades básicas das variáveis envolvidas no estudo. Esta secção inicia-se com a descrição da distribuição da amostra (da série temporal considerada) a partir da representação de histogramas e das estatísticas como a média, mediana e o desvio padrão de cada variável (figura 3.4).

A média e a mediana são estatísticas descritivas da tendência central de uma amostra, enquanto que o desvio padrão descreve a variabilidade da mesma.

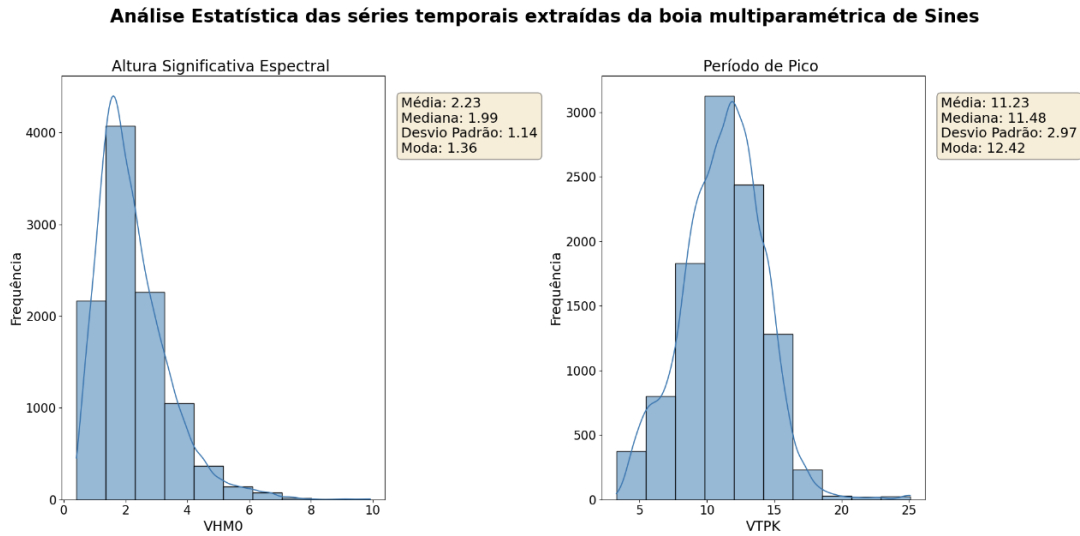


FIGURA 3.4: Distribuição das séries temporais da altura significativa espectral H_{m0} (gráfico à esquerda) e do período de pico T_p (gráfico à direita) extraídas da boia multiparamétrica de Sines

A amostra extraída da boia multiparamétrica de Sines é constituída por 10152 observações das variáveis anteriormente mencionadas na tabela 3.1. Esta amostra é constituída por três séries temporais distintas, que serão apresentadas na próxima secção deste capítulo. No presente trabalho foram utilizadas as séries temporais correspondentes à altura significativa espectral H_{m0} e ao período de pico T_p .

Na figura 3.4 observa-se que a distribuição da altura significativa espectral da onda segue uma distribuição assimétrica à direita da média = 2,23 m e centra-se entre as alturas de 1,5 a 3 m, com ocorrências pouco frequentes superiores a 5 m. O período de pico tem média de 11,23 s, mediana de 11,48 s e moda de 12,42 s. A maioria dos períodos de pico concentra-se entre 10 e 12 s. O desvio padrão é 2,97 s, o que sugere alguma variabilidade, refletindo a diversidade das condições de onda no local, onde se registaram ondas com o período de pico compreendido entre 5 e 20 s.

Análise Estatística das séries temporais extraídas do modelo WWIII ao largo do arquipélago dos Açores

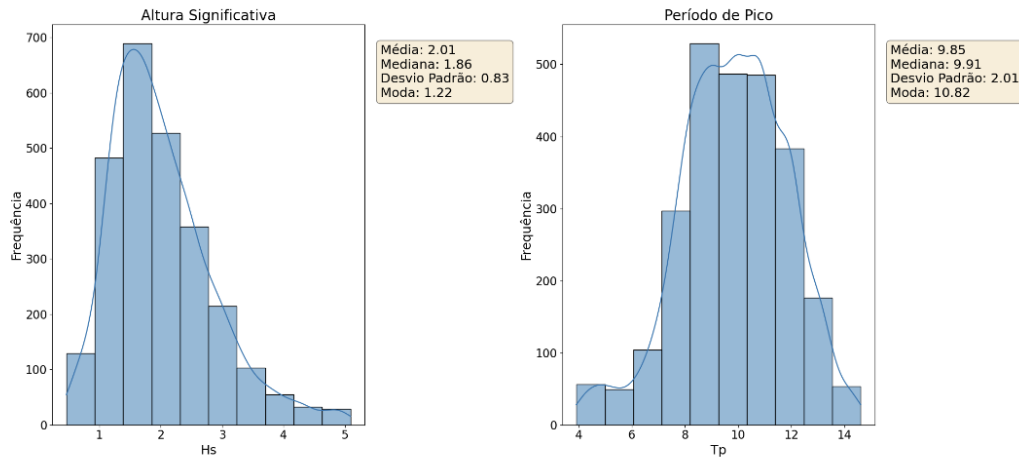


FIGURA 3.5: Distribuição das séries temporais da altura significativa H_s (gráfico à esquerda) e do período de pico T_p (gráfico à direita) extraídas da malha geográfica do modelo *WaveWatch III* ao largo do arquipélago dos Açores

A figura 3.5 representa a distribuição das séries temporais extraídas do modelo WWIII ao largo do arquipélago dos Açores. O gráfico da direita representa a distribuição da altura significativa H_s da onda que segue uma distribuição assimétrica à direita da média= 2,01 m e centra-se entre as alturas de 1 a 2,5 m, com ocorrências pouco frequentes superiores a 4 m. O período de pico das ondas apresenta uma média de 9,85 s, mediana de 9,91 s e moda de 10,82 s. A maioria dos períodos de pico concentra-se entre 8 e 9 s. O desvio padrão é 2 s, o que sugere alguma variabilidade, refletindo a diversidade das condições de onda no local, onde se registaram ondas com o período de pico compreendido entre 4 e 14 s.

Análise Estatística das séries temporais extraídas do modelo WWIII ao largo do arquipélago da Ilha da Madeira

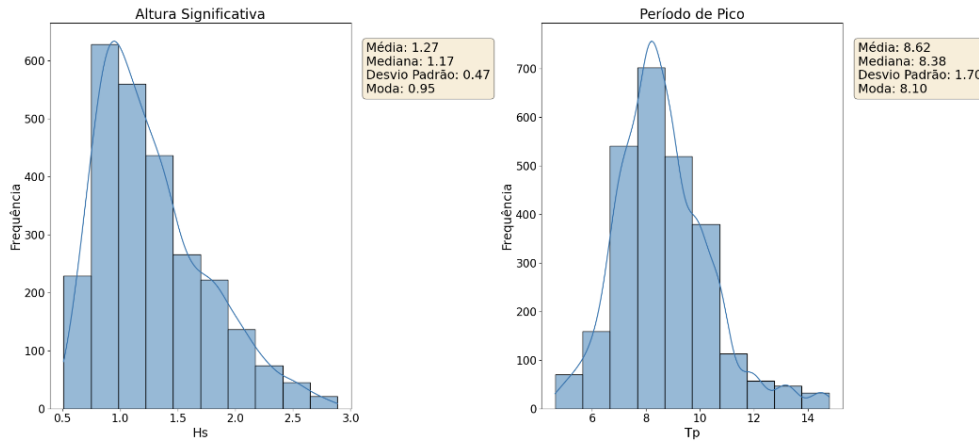


FIGURA 3.6: Distribuição das séries temporais da altura significativa H_s (gráfico à esquerda) e do período de pico T_p (gráfico à direita) extraídas da malha geográfica do modelo *WaveWatch III* ao largo da Ilha da Madeira

Por último, a figura 3.6 representa a distribuição estatística das séries temporais extraídas pela malha geográfica do modelo numérico WWIII ao largo da Ilha da Madeira. O gráfico da direita representa a distribuição da altura significativa H_s da onda segue uma distribuição assimétrica à direita da média= 1,27 m e centra-se entre as alturas de 1 a 2 m, com ocorrências pouco frequentes superiores a 2,5 m. O período de pico das ondas apresenta uma distribuição aproximadamente simétrica com uma média de 8,62 s, mediana de 8,38 s e moda de 8,10 s. A maioria dos períodos de pico concentra-se entre 8 e 9 s. O desvio padrão é de 1,7 s, o que sugere menor variabilidade que os casos anteriormente analisados.

3.3.2 Preparação prévia das séries temporais em estudo

Os dados da boia multiparamétrica foram extraídos a partir da plataforma *Copernicus Marine Service*, no formato Netcdf (*Network Common Data Form*). Nesta plataforma, para bóia multiparamétrica de Sines, na qual existiam registos que datavam desde maio de 2020 até janeiro de 2024, porém com alguns períodos intermédios sem registos.

O formato Netcdf é esquematicamente representado pela figura 3.7, e permite o manuseamento de uma grande quantidade de dados, pelo que se torna especialmente útil para armazenar dados extraídos por exemplo de modelos numéricos como *WaveWatch III*, por terem um conjunto de variáveis a serem representados

em matrizes que correspondem às áreas definidas pela superfície do oceano. Num ficheiro Netcdf, as variáveis ficam armazenadas em matrizes associadas em dimensões, neste caso: latitude, longitude e tempo. Uma das vantagens deste formato é que permite que seja selecionado pelo utilizador o intervalo de tempo.

No entanto, no caso da boia multiparamétrica e dos pontos extraídos pelo modelo numérico *WaveWatch III*, como se definiu coordenadas específicas para a localização da extração das séries temporais (a localização da bóia e dos pontos extraídos do modelo numérico), considerou-se transformar o formato do ficheiro para .csv com o intuito de otimizar o processamento dos dados.

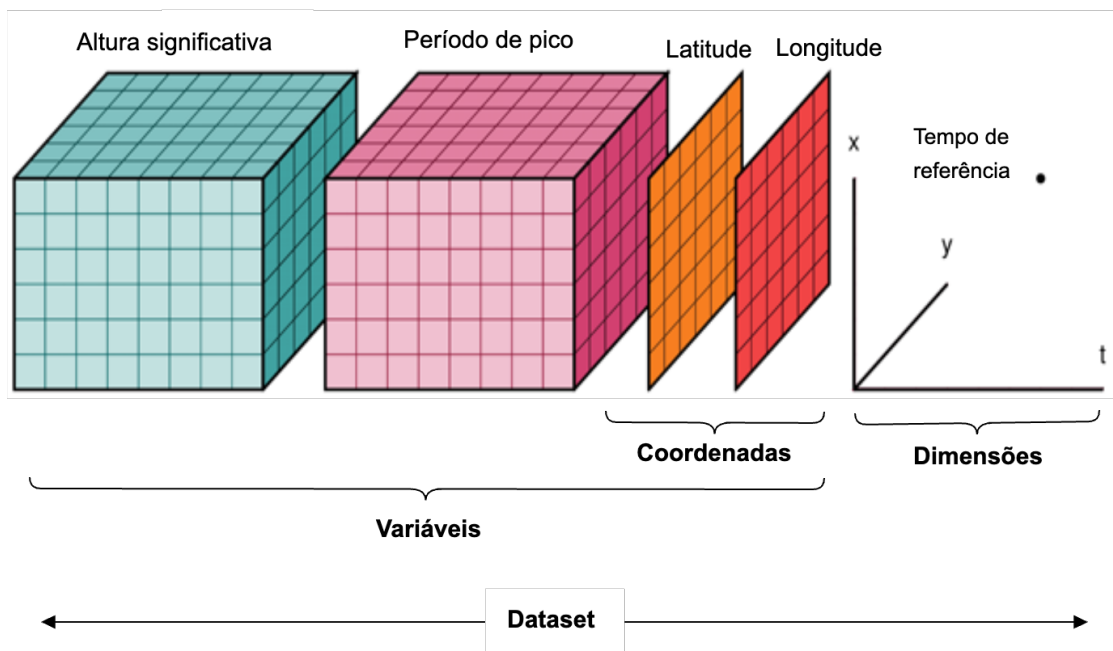


FIGURA 3.7: Representação de um ficheiro Netcdf. Adaptado de: Hoyer e Hamman [2017](#)

Após a extração de todos os registos da boia, foi essencial selecionar um período para as séries temporais que iriam ser utilizadas nos modelos estatísticos e de AA. O critério utilizado foi a definição de um intervalo em que a lacuna temporal máxima fosse igual ou inferior a 48 horas e simultaneamente tivesse a maior parte dos registos disponíveis, isto é, que não tivesse um grande número de valores omisso (Nan). Nos modelos estatísticos aplicados é importante garantir que as séries temporais são constantes. Desta forma, selecionou-se o período compreendido entre 08-11-2022 07:00:00 e 04-03-2023 22:00:00 como a série a ser utilizada pelos modelos

estatísticos. Nos modelos de AA de redes neurais MLP não é essencial garantir que as séries sejam contínuas, já que a continuidade dos dados é tratada pela preparação dos mesmos através de *lagged variables*, isto é, valores desfasados da série temporal original. Desta forma, e por os modelos de AA necessitarem de um volume considerável de dados para maior precisão, utilizou-se a três séries temporais:

- Série temporal 1: Compreendida entre 08-11-2022 07:00:00 e 04-03-2023 22:00:00, com três lacunas temporais de 23, 17 e 7 horas (figura 3.8) com 2800 registos (lacunas preenchidas);
- Série temporal 2: Compreendida entre 27-04-2023 10:00:00 e 10-08-2023 07:00:00, sem lacunas temporais (figura 3.10) com 2518 registos;
- Série temporal 3: Compreendida entre 12-08-2023 14:00:00 e 11-02-2024 16:00:00, sem lacunas temporais (figura 3.11) com 4395 registos.

No caso dos modelos estatísticos ARIMA e de regressão linear é importante garantir a continuidade da série temporal. Na utilização destes modelos considerou-se a série com mais registos mas que fosse simultaneamente mais representativa da agitação marítima de Sines. A série temporal 3 apesar de incluir mais registos, tem um pico elevado na ordem dos 10 metros, um valor atípico dos restantes, provavelmente devido a uma tempestade, tornando a variância da série maior. Por essa razão, optou-se por considerar a série temporal 1 para aplicação dos modelos estatísticos tradicionais.

Após a seleção do período inicial e final da série temporal, procurou-se imputar as lacunas temporais existentes na série temporal 1 e ainda verificar a existência de valores omissos nos registos de todas as séries, isto é, verificar se todos os parâmetros foram registados corretamente. Identificou-se que a série temporal 1 tinha dois valores de H_{m0} em falta que foram imputados pelo método linear assim como as lacunas temporais identificadas na mesma.

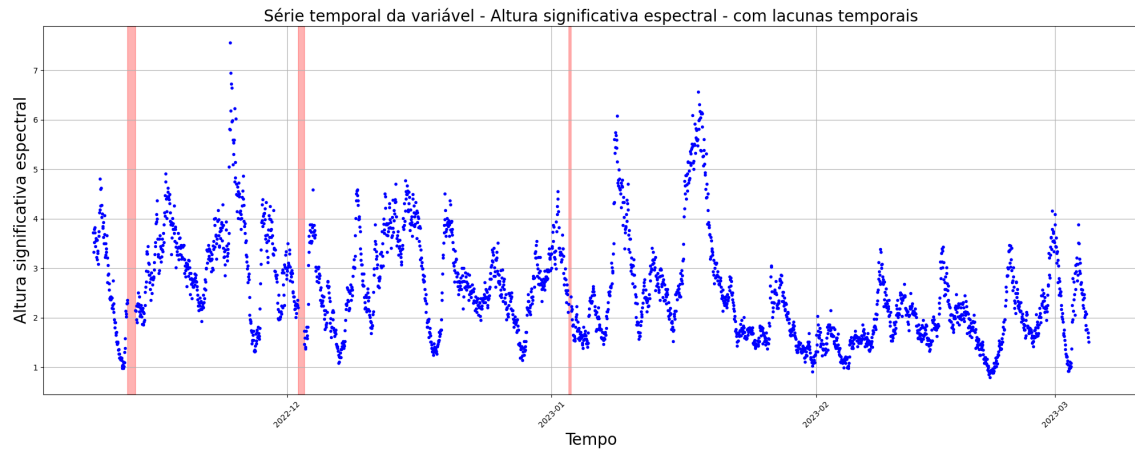


FIGURA 3.8: Série temporal (1) da altura significativa espectral com lacunas temporais

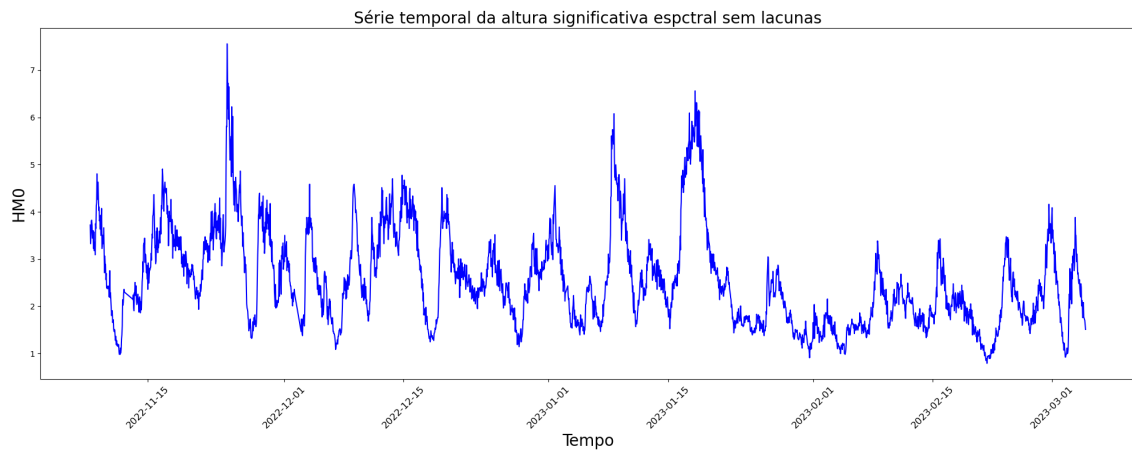


FIGURA 3.9: Série temporal (1) da altura significativa espectral após a interpolação

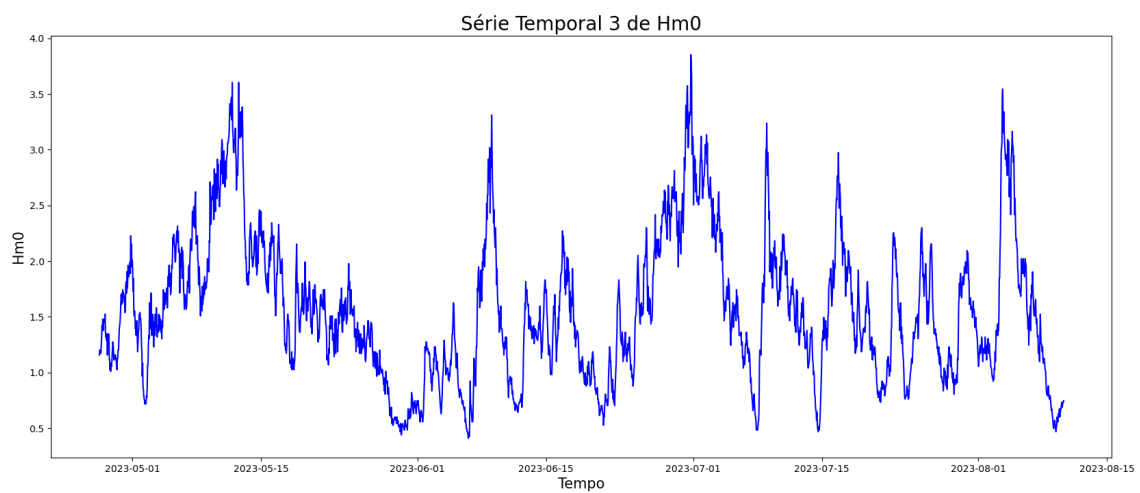


FIGURA 3.10: Série temporal (2) da altura significativa espectral

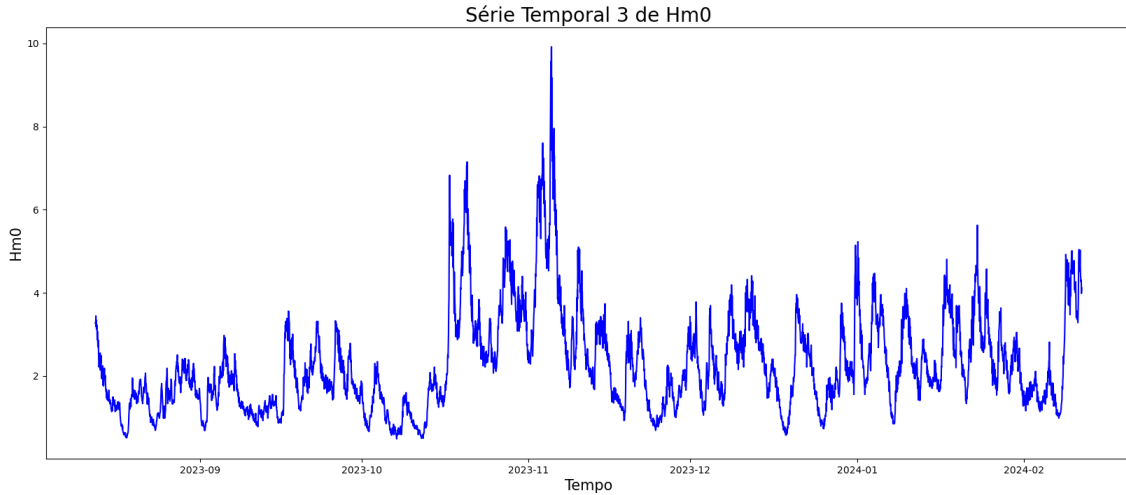


FIGURA 3.11: Série temporal (3) da altura significativa espectral

No pré-processamento que antecedeu a aplicação dos modelos de regressão linear e de AA, foi necessário selecionar apenas os registos das séries temporais extraídas da boia multiparamétrica de Sines, com as datas coincidentes com os registos extraídos do modelo numérico *WaveWatch III*. Este modelo calcula 4 registos diários para os períodos de 0, 6, 12, 18 horas. As séries temporais originalmente extraídas da boia multiparamétrica de Sines correspondiam a 10152 registos, porém após a seleção das datas coincidentes com os registos calculados pelo modelo numérico WWIII, a amostra diminuiu para 1677 registos.

3.3.3 Análise exploratória das séries temporais em estudo

A análise exploratória das séries temporais consistiu na autocorrelação da variável H_{m0} bem como na correlação da mesma com outras variáveis (tais como T_p), e ainda na correlação cruzada entre a variável H_{m0} e as séries temporais extraídas da malha geográfica do modelo numérico WWIII.

Inicialmente verificou-se a condição de estacionaridade da série, e sequencialmente procedeu-se à análise das funções de autocorrelação e autocorrelação parcial da série da altura significativa espectral H_{m0} de modo a ser possível a aplicação do modelo estatístico ARIMA. A análise da autocorrelação entre valores sucessivos da série foi a partir da função autocorrelação como demonstrado na figura [3.12](#). Neste gráfico é possível concluir-se que existe uma correlação muito forte entre valores sucessivos da amostra, o que é uma boa indicação que a série não é estacionária. Por forma a aplicar o modelo ARIMA, foi necessário aplicar uma transformação à série temporal, a qual consistiu no resultado da diferença dos seus valores sucessivos,

garantindo assim a estacionaridade. No primeiro gráfico da figura 3.12 verifica-se que a série apresenta autocorrelação significativa até aproximadamente 48 horas, sugerindo assim a presença de ciclos na mesma.

As ordens de p e q do modelo ARIMA inicialmente aplicado foram determinadas pela análise da função de autocorrelação e de autocorrelação parcial diferenciadas apresentadas na figura 3.12.

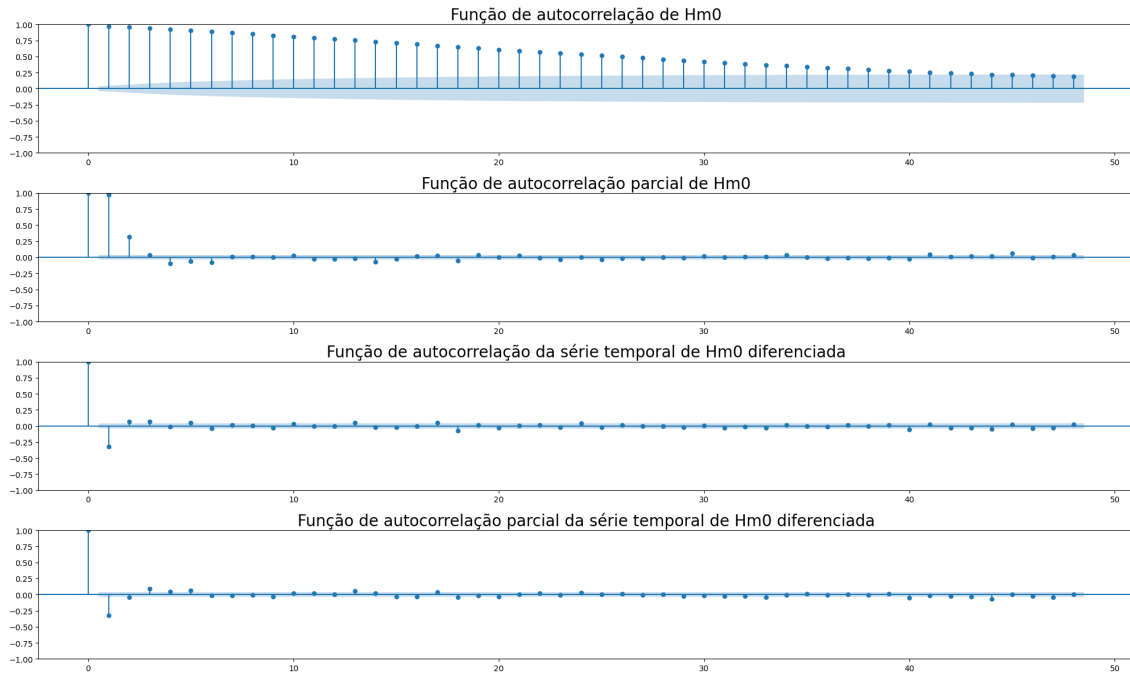


FIGURA 3.12: Função autocorrelação da série temporal da altura significativa (topo), função de autocorrelação parcial da série temporal da altura significativa (segundo gráfico), função autocorrelação da série temporal da altura significativa diferenciada uma vez (terceiro gráfico) e respetiva função de autocorrelação parcial (último gráfico)

A correlação de Pearson é uma ferramenta estatística que permite identificar relações lineares entre variáveis, esta é importante ser analisada, pois permite identificar as variáveis que devem ser consideradas no modelo de regressão linear.

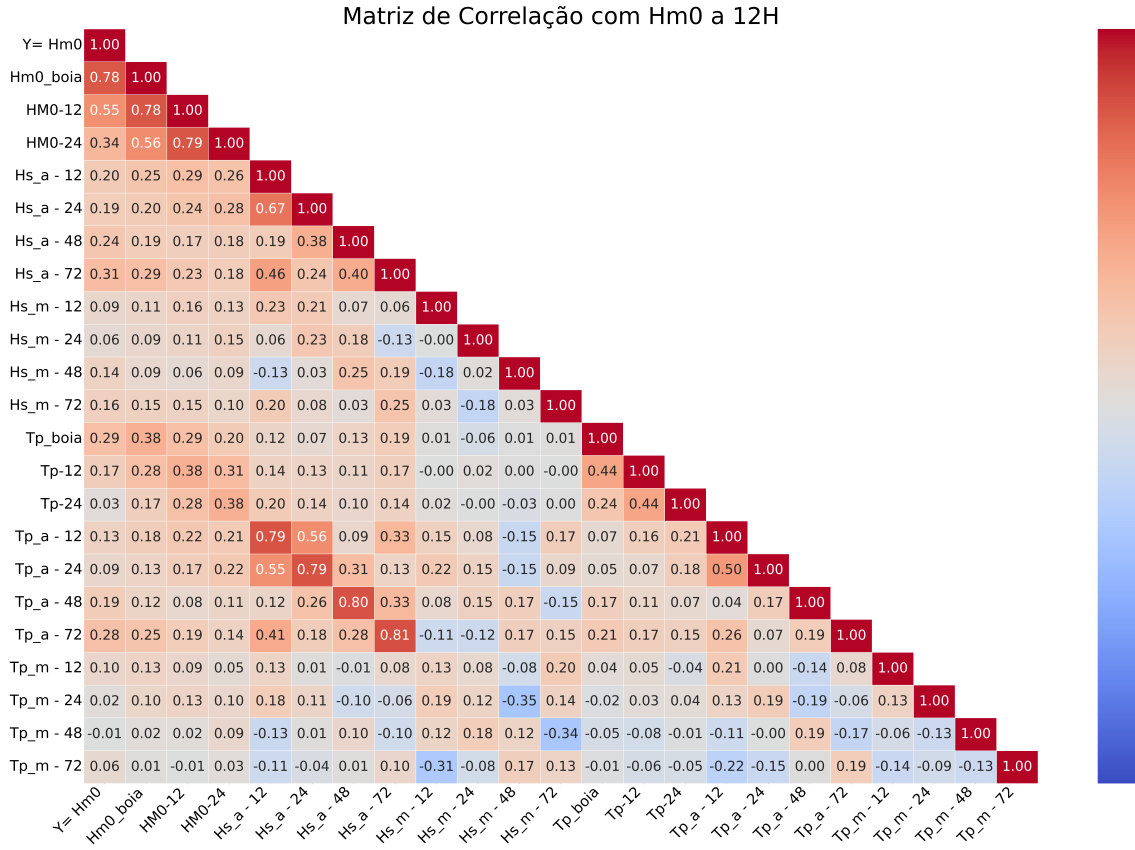


FIGURA 3.13: Matriz de correlação referente à variável H_{m0} considerando o ponto extraído do modelo WWIII ao largo dos Açores e ao largo da Madeira

A matriz de correlação apresentada na figura 3.13 representa a correlação entre a variável dependente $H_{m0} + 12$, representada por $Y = H_{m0}$, as variáveis H_{m0} e T_p nos instantes $t - 12, t - 24$ (instantes anteriores ao instante t) e ainda com as séries temporais de H_{m0} e T_p extraídas do modelo numérico nos *lags* de 12, 24, 48 e 72 horas.

A correlação cruzada é uma ferramenta útil para identificar a relação temporal entre duas séries temporais. No presente trabalho foi feita uma análise da correlação cruzada entre as séries extraídas pela boia multiparamétrica de Sines e as séries retiradas na malha geográfica do modelo WWIII ao largo da Ilha da Madeira e do arquipélago dos Açores.

A partir da análise de correlação cruzada entre as séries temporais de T_p da boia multiparamétrica de Sines e do ponto extraído da malha geográfica ao largo do arquipélago dos Açores representado na figura 3.14, conclui-se que as duas séries temporais apresentam uma correlação crescente a partir do *lag* de 20 horas e a sua correlação máxima corresponde aproximadamente ao de 72 horas ($t - 72$).

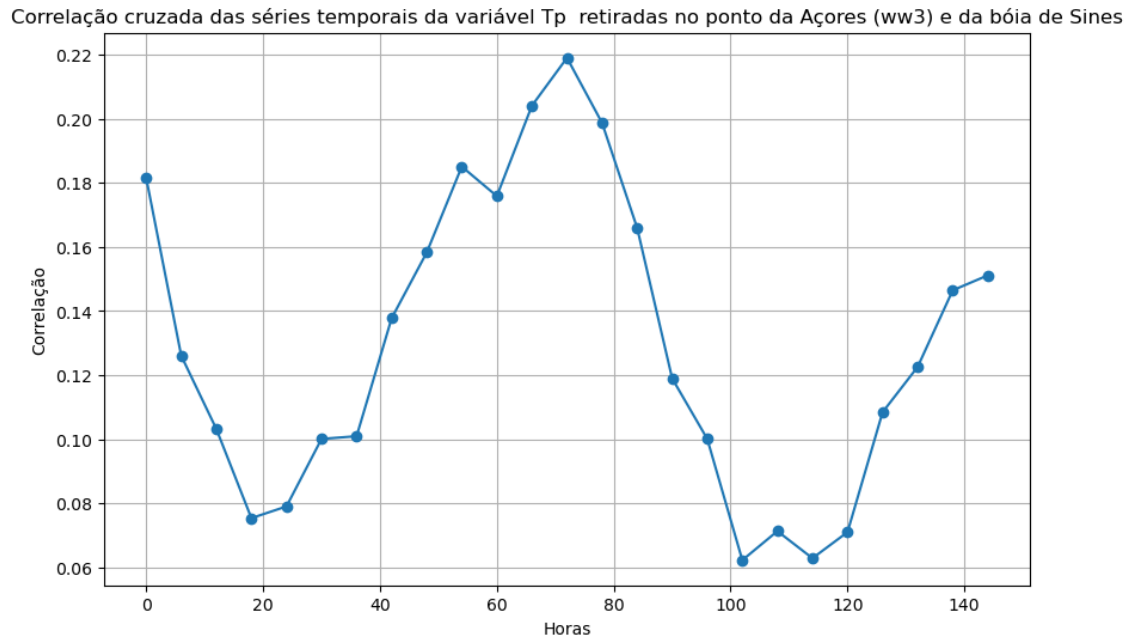


FIGURA 3.14: Correlação cruzada referente às séries temporais de T_p na boia de Sines e no arquipélago dos Açores.

A figura [3.15](#) representa o resultado da correlação entre a série temporal de H_{m0} extraída ao largo do arquipélago dos Açores e da boia de Sines, na qual é possível observar que existe uma forte correlação crescente entre ambas a partir do *lag* 40 horas ($t - 40$) e uma correlação máxima próxima das 80 horas ($t - 80$), o que sugere que existe um efeito de propagação considerável do efeito da agitação marítima da região dos Açores. Esta análise indica-nos que a série temporal de H_{m0} ao largo do arquipélago dos Açores, poderá ter valor significativo na previsão de H_{m0} em Sines para desfasamentos (*lags*) compreendidos entre 40 a 80 horas.

Correlação cruzada das séries temporais da variável H_{m0} retiradas no ponto da Açores (ww3) e da bóia de Sines

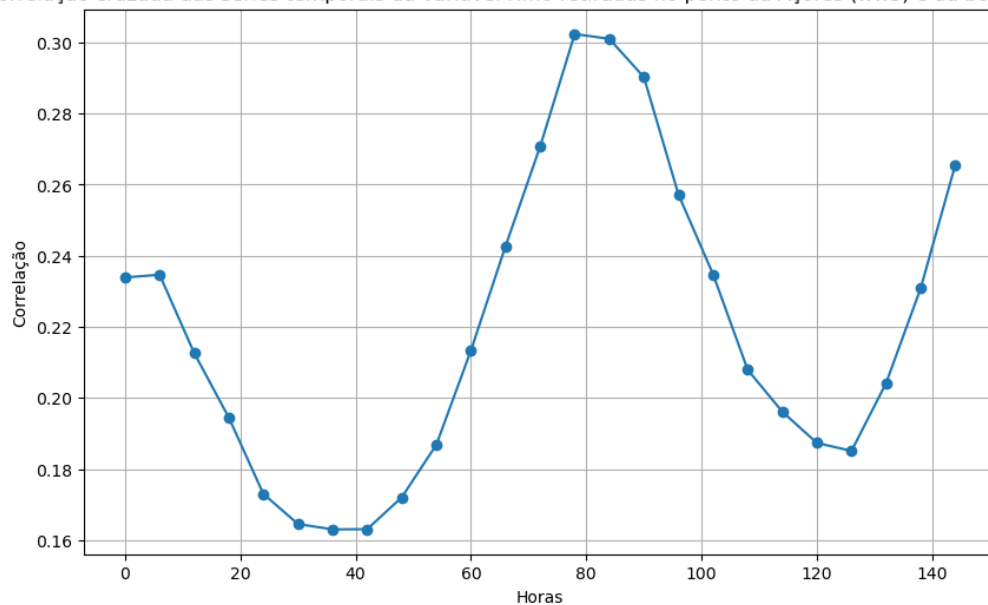


FIGURA 3.15: Correlação cruzada referente às séries temporais de H_{m0} na boia de Sines e no arquipélago dos Açores.

Correlação cruzada das séries temporais da variável H_{m0} retiradas no ponto da Madeira (ww3) e da bóia de Sines

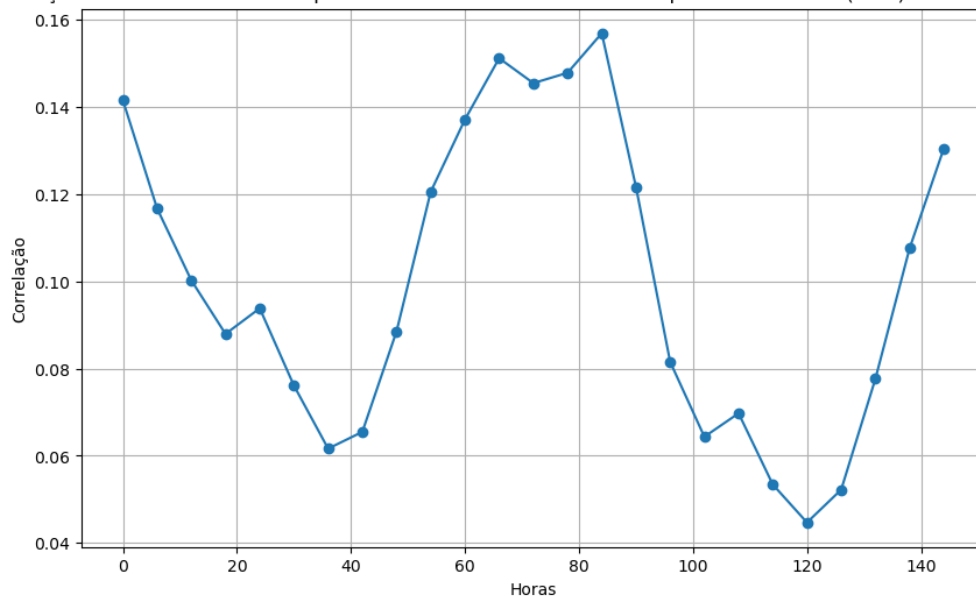


FIGURA 3.16: Correlação cruzada referente às séries temporais de H_{m0} na boia de Sines e na Ilha da Madeira.

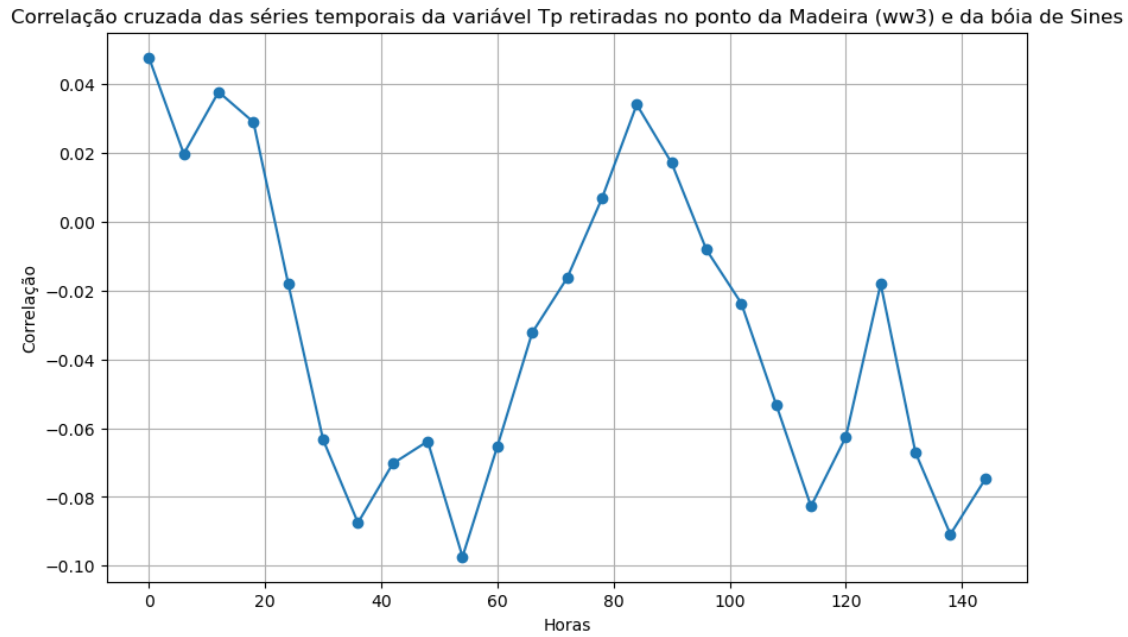


FIGURA 3.17: Correlação cruzada referente às séries temporais de T_p na boia de Sines e na Ilha da Madeira.

A figura 3.16 e a figura 3.17 representam a correlação cruzada entre as séries temporais registadas na boia de Sines e no modelo WWIII ao largo da Ilha da Madeira, referentes às variáveis H_{m0} e T_p . Relativamente à análise da primeira figura, verifica-se um aumento da correlação entre as duas séries temporais a partir do *lag* referente das 40 horas, com o máximo de correlação no *lag* próximo das 82 horas. Contudo, a correlação entre estas duas séries temporais apresentam uma menor correlação em comparação ao exemplo anterior (figura 3.15), podendo dever-se ao facto da Ilha dos Açores se localizar mais a Oeste estando mais próximos das zonas de geração de agitação marítima no Atlântico Norte, e por essa razão concentrar um maior rácio de energia das ondas que se propagam em direção à costa de Sines.

No caso da figura 3.17, o aumento de correlação entre as duas séries temporais (T_p em Sines e T_p ao largo da Ilha da Madeira) ocorre aproximadamente a partir do *lag* referente às 36 horas. Além disso, o pico de correlação entre as duas séries temporais ocorre aproximadamente a 72 horas ($t - 72$).

3.4 Modelos estatísticos e de Aprendizagem Automática

3.4.1 Modelo ARIMA

O modelo ARIMA aplicado tinha como objetivo a previsão da variável H_{m0} para períodos de 24 e 48 horas. Desta forma foram construídos dois modelos ARIMA, um para cada um destes intervalos de previsão. A partir do resultado das previsões do modelo ARIMA, foi possível analisar a influência dos valores anteriores da série temporal H_{m0} nas previsões e determinar a extensão temporal dos registos históricos a serem considerados na previsão da mesma.

Conforme descrito no capítulo 2, o modelo ARIMA é um modelo estatístico aplicável a séries temporais estacionárias. A condição de estacionariedade do modelo aplicado foi inicialmente verificada através da análise visual da série temporal da variável H_{m0} e garantida por meio da diferenciação da mesma, ou seja, através de uma transformação calculada pela diferença entre valores consecutivos da série temporal. A série temporal diferenciada foi submetida a testes estatísticos, nomeadamente ao teste de *Augmented Dickey-Fuller* (ADF) (Mushtaq 2011) e ainda *Kwiatkowski-Phillips-Schmidt-Shin* (KPSS) (Kwiatkowski, P. C. Phillips, Schmidt e Shin 1992), para confirmar a estacionariedade da série temporal diferenciada.

O teste de ADF determina a presença de uma raiz unitária (*unit root*), ou seja, verifica a presença de um coeficiente autorregressivo $\alpha = 1$. As hipóteses nulas e alternativas deste teste são:

- Hipótese nula (H_0): A série temporal tem uma raiz unitária⁶
- Hipótese alternativa (H_1): A série temporal é estacionária.

Ao rejeitar a hipótese nula, conclui-se que a série temporal é estacionária.

O teste de KPSS verifica também a estacionariedade da série, contudo verifica a ausência de uma raiz unitária em torno de uma média ou tendência. As hipóteses nulas e alternativas deste teste estatístico são:

- Hipótese nula (H_0): A série temporal é estacionária.

⁶As séries temporais descritas pelo modelo ARIMA podem ser representadas por polinómios nos quais a variável independente é o desfasamento (*lag*). A demonstração da relação entre a localização das raízes do polinómio que representa a componente AR e a estacionariedade da série não é simples e é demonstrada por Hamilton 1994. Em termos práticos, se as raízes estão fora do círculo unitário, ou seja, não têm uma raiz unitária, então a série temporal é estacionária.

- Hipótese alternativa (H_1): A série temporal não é estacionária.

Ao rejeitar a hipótese nula, conclui-se que a série temporal não é estacionária em torno de uma média ou numa tendência linear. O p-value definido como critério para rejeitar a hipótese nula no teste de ADF foi $p\text{-value} < 0,05$ e no teste de KPSS $p\text{-value} > 0,05$. Após uma diferenciação da série temporal, ambos os testes indicaram que a série era estacionária.

A determinação das componentes (p, q) do modelo foram a partir da análise das funções de autocorrelação e de autocorrelação parcial respetivamente. Testou-se vários modelos ARIMA com diferentes ordens das componentes p, q com base no critério AIC e BIC (descritos na secção 2.2), preferindo-se o critério BIC por dar ênfase à simplicidade do modelo.

Por forma a elaborar uma avaliação completa da performance do modelo estatístico, verificou-se a resposta do mesmo no intervalo correspondente ao conjunto de dados disponível, ao prever os vários valores de H_{m0} para esse intervalo.

Desta forma, desenvolveu-se um modelo ARIMA com base numa “janela deslizante” que percorreu os dados ao longo da série temporal. Inicialmente o modelo calculou os seus coeficientes e posteriormente foi aplicando as suas previsões ao longo da janela de teste deslizando. Definiram-se janelas deslizantes, que correspondiam a um conjunto de teste de 12, 24, 48, 60 e 72 horas. Testou-se a performance do modelo para previsões de 12, 24 e 48 horas para todas as janelas deslizantes descritas. Neste modelo, após terem sido calculados os coeficientes, a janela deslizando considerava um conjunto sucessivo de pontos, correspondente às 12, 24, 48, 60 ou 72 horas, para efetuar previsões a 12, 24 ou 48 horas. Após as previsões a janela deslocava-se uma hora para frente, considerando agora outro conjunto de 12, 24, 48, 60 ou 72 horas, de modo a fazer outras previsões, este processo é repetido iterativamente até a janela ter percorrido toda a série temporal.

Inicialmente testou-se a implementação do modelo ARIMA(2,1,1) na qual se determinou as ordens p e q com base na observação do gráfico da função de autocorrelação e da função autocorrelação parcial. Contudo, após se ter feito a análise dos resíduos do modelo, concluiu-se que este modelo não reconhecia os padrões da série temporal pelo que se procurou explorar outras ordens para o modelo.

No presente trabalho foi aplicado um modelo ARIMA (2,1,5) à série temporal H_{m0} , para prever valores nos intervalos de 12, 24 e 48 horas. No modelo

ARIMA(2,1,5), as previsões são expressas pela seguinte expressão:

$$y_i = c + \alpha_1 x_{t-1} + \alpha_2 x_{t-2} + \theta_1 w_{t-1} + \theta_2 w_{t-2} + w_t \quad (3.1)$$

Os coeficientes α e θ são calculados com base no método máxima verossimilhança (mencionado na secção 2.2). Após a determinação dos coeficientes o modelo aplica-os ao longo de uma janela de teste que considera um conjunto de valores correspondentes a 12, 24, 48, 60 ou 72 horas, para prever 12, 24 e 48 horas.

3.4.2 Modelo de regressão linear

O modelo de regressão linear aplicado no presente trabalho tinha como objetivo prever para o intervalo de 12 horas a variável H_{m0} , permitindo também identificar os previsores mais significativos no processo de previsão da série temporal em questão.

Conforme anteriormente mencionado no capítulo 2, o modelo de regressão linear tem como pressuposto que existem relações lineares entre as variáveis dependente(s) e independente.

A aplicação do modelo de regressão linear exigiu uma análise prévia de correlação, a qual está expressa na matriz de correlação apresentada na figura 3.13.

Para além disso, foi ainda essencial normalizar os dados de modo a transformar os valores numéricos dos dados numa escala próxima, com o objetivo de melhorar a eficiência e a precisão do modelo. Neste caso, aplicou-se um fator de escala de 2,5 metros para as variáveis correspondentes a alturas significativas e um fator de escala de 10 segundos para as variáveis correspondentes aos períodos de pico.

As variáveis independentes foram selecionadas com base na sua estatística $p - value$, e na performance do modelo relativamente ao R^2 . Após a seleção das variáveis, os registos da série temporal foram divididos em dados de treino e de teste, respeitando o critério 80%, 20% do conjunto de dados inicial. O modelo foi ajustado aos dados de treino, e posteriormente, verificado a sua performance através das previsões para o conjunto de dados de teste.

3.4.3 Modelos Aprendizagem Automática: Redes neuronais

O desenvolvimento dos modelos de AA apresentados neste trabalho seguiram uma estratégia de exploração iterativa dos três eixos fundamentais que sustentam o desempenho das redes neuronais: a estrutura, a variação de hiperparâmetros das redes neuronais, e as fontes de dados. Na análise da estrutura das redes neuronais importou conhecer o número de camadas intermédias e o número de neurónios essenciais para garantir o treino e a precisão das mesmas. A variação dos hiperparâmetros tais como o *batch size*, a taxa de aprendizagem, a regularização e o número de épocas de treino concorreu para compreensão dos fatores de otimização das mesmas. Por fim, a utilização de diferentes fontes de dados permitiu que os modelos fossem treinados com maior variedade de informação, decifrando a relevância dos diferentes previsores na estimativa da altura significativa espectral (figura 3.18).

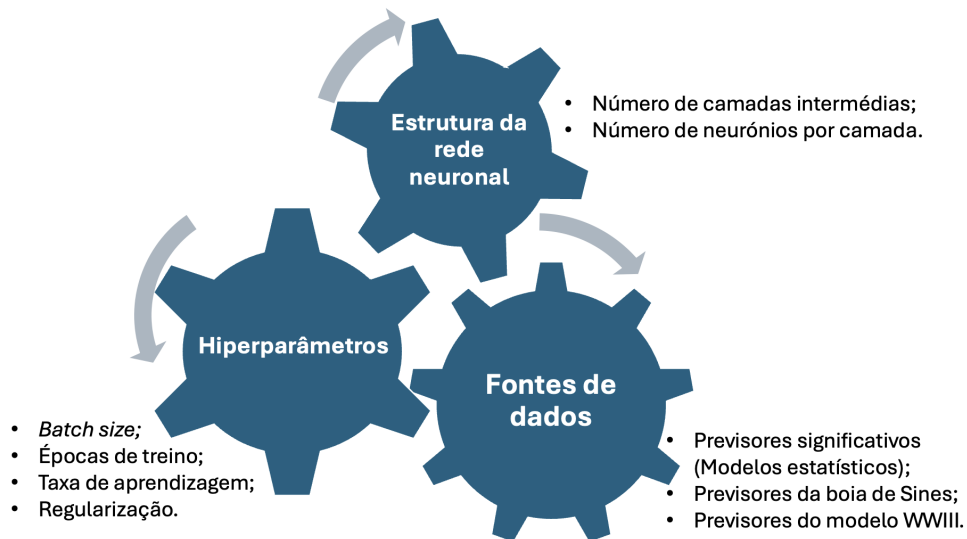


FIGURA 3.18: Estratégia de exploração dos modelos AA de redes neuronais

Conforme referido no capítulo anterior, as redes neuronais MLP não têm processos de retro-alimentação e tratam os seus valores de entrada como vetores independentes, por isso não é possível utilizar a variável tempo (variáveis *Timestamp*) como predictor. Para prever séries temporais, como é o caso da altura significativa espectral H_{m0} , com redes neuronais MLP é necessário criar *lagged variables*, ou seja, variáveis dos seus valores desfasadas no tempo. Desta forma, e porque os modelos de AA requerem um volume de dados significativo para garantir a sua precisão,

utilizou-se três séries temporais extraídas da boia multiparamétrica de Sines. A estas séries temporais ainda se fez corresponder as séries temporais extraídas do modelo *Wave Watch III* no período temporal considerado para cada uma das séries extraídas da boia de Sines, constituindo assim um conjunto de dados com 1677 registos, com intervalos de amostragem entre registos de 6 horas.

Os modelos de AA com algoritmos de redes neuronais MLP aplicados no presente trabalho utilizam uma das divisões recomendadas para conjuntos de dados de treino limitados (capítulo 1). Esta divisão é constituída por 70/15/15, isto é, 70% do conjunto de dados deve ser utilizada para treino, 15% dos dados deve constituir o conjunto de validação e 15% do conjunto de dados é aplicado como teste ao modelo (figura 3.19).

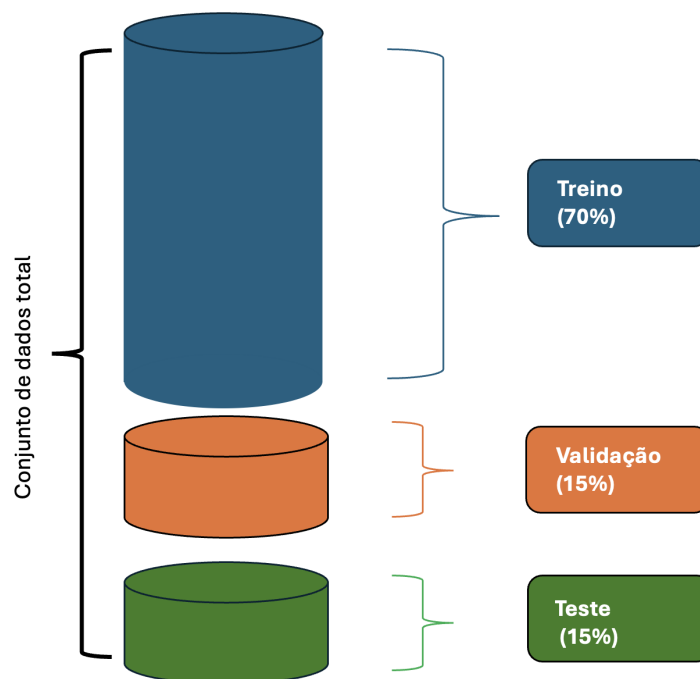


FIGURA 3.19: Divisão do conjunto de dados em conjunto de treino, validação e de teste.

Após a divisão do conjunto de dados em treino, validação e teste, sucedeu-se o processo de normalização. Os conjuntos de dados foram normalizados respeitando o critério adotado na aplicação do modelo de regressão linear, no qual foi aplicado o fator de escala de 2,5 metros para as variáveis que dizem respeito à altura significativa e de 10 metros para as variáveis que dizem respeito ao período de pico.

Estrutura da rede neuronal

As redes neuronais são constituídas por uma camada de entrada, uma ou mais camadas intermédias e uma camada de saída.

Os neurónios presentes na camada de entrada são responsáveis pelo processamento das variáveis de entrada (*features*) enquanto que os neurónio(s) que constituem a camada de saída produzem o resultado final que resulta do processamento das camadas anteriores. Como a variável de resposta pretendida é um escalar, a camada de saída deve ter apenas um único neurónio.

Na construção de estruturas de redes neuronais é uma boa prática definir o número de neurónios na camada de entrada igual ao número de *features* utilizadas para treino do modelo. Este critério foi considerado durante a exploração dos três eixos presentes neste trabalho.

Nos estudos publicados por Oleg Makarynsky [2005](#) e Minuzzi e Farina [2023](#) foram utilizadas redes neuronais MLP com três camadas (uma camada de entrada, uma camada intermédia e uma camada de saída) para prever a altura significativa espectral H_{m0} . Oleg Makarynsky [2005](#) selecionou o número de neurónios na camada intermédia H com base no seguinte critério (Huang e Foo [2002](#)):

$$H = 2Z + 1 \tag{3.2}$$

em que Z é o número de neurónios presentes na camada de entrada. Este critério tinha como objetivo evitar fenómenos de *overfitting*.

Deste modo, a estrutura da rede neuronal inicialmente aplicada neste trabalho, teve em consideração os dois critérios mencionados, desenvolvendo-se assim uma rede neuronal MLP constituída por três camadas (uma camada de entrada, uma camada intermédia e uma camada de saída), em que o número de neurónios da camada de entrada era o mesmo do número de variáveis de entrada, e o número de neurónios da camada intermédia respeitava o critério definido pela equação [3.2](#).

Hiperparâmetros

Os hiperparâmetros são variáveis nos algoritmos de redes neuronais, que afetam o seu comportamento no tempo e memória despendidos no processo de treino e na precisão do modelo (Goodfellow, Bengio e Courville [2016](#)). Existem duas abordagens para selecionar os hiperparâmetros: manualmente ou automaticamente. No presente trabalho foram testados ambas as abordagens, começando com o método

manual e aplicando posteriormente, o método automático através do algoritmo *Random Search*, que será mencionado adiante nesta secção.

O processo de seleção dos hiperparâmetros é um processo de otimização, em que o objetivo é encontrar a combinação de parâmetros para a qual o modelo é o mais simples, eficiente e preciso possível sem que ocorram fenómenos de *overfitting*. A análise das curvas de aprendizagem permite avaliar se o modelo apresenta sinais de *overfitting*. Estas curvas de aprendizagem referem-se aos valores de perda da função de custo ao longo das épocas de treino. Um modelo bem ajustado, é um modelo capaz de compreender os padrões presentes nos dados e de os generalizar. Modelos bem ajustados apresentam uma evolução das perdas da função de custo seguindo um padrão decrescente com valores relativamente baixos no final do treino do modelo. Os valores de perda do conjunto de validação e de treino, no final do processo de aprendizagem do modelo devem ser tão próximos quanto possível. A otimização de hiperparâmetros influencia bastante as curvas de aprendizagem das redes neuronais, regulando a aprendizagem do mesmo. Os hiperparâmetros considerados mais impactantes por Goodfellow, Bengio e Courville [2016] para o processo de aprendizagem das redes neuronais são:

- Épocas: Determina o número de ciclos de passagem do conjunto de treino pela rede;
- *Batch size*: Número de exemplos que são processados a cada época para atualização dos pesos sinápticos da rede neuronal;
- Taxa de aprendizagem: Controla a intensidade das atualizações feitas nos pesos da rede após cada iteração no processo de treino, determinando o fluxo de aprendizagem do modelo;
- Regularização: Modificação feita a um algoritmo de aprendizagem com o objetivo de reduzir os erros no conjunto de teste.

O algoritmo *Random Search* é uma estratégia de exploração de hiperparâmetros que otimizam o desempenho das redes neuronais, na qual se testam as várias combinações aleatórias de hiperparâmetros anteriormente mencionados (no qual também se pode incluir o número de camadas intermédias e neurónios por camadas). Contudo, este processo de otimização é computacionalmente dispendioso, o número de combinações possíveis a serem elaboradas pelo algoritmo varia conforme a capacidade de processamento do *hardware* de cada utilizador. Este algoritmo integra a função *EarlyStopping* da biblioteca do *Keras* do *software Python* (Chollet et al. [2015]) que interrompe o treino da rede neuronal se os valores de perda da

função de custo no conjunto de validação não diminuir significativamente ao longo de 10 camadas consecutivas. No final o algoritmo seleciona os hiperparâmetros que contribuíram para um menor erro definido pela métrica MAE.

No presente trabalho utilizou-se o algoritmo *Random Search* para determinar inicialmente a estrutura das redes neuronais, nomeadamente o número de camadas intermédias e o número de neurónios na(s) camada(s) intermédia(s). Este algoritmo foi também aplicado após de ter sido determinado a estrutura de rede mais eficaz, para determinar os hiperparâmetros como o *batch size*, taxa de aprendizagem e a regularização que otimizavam a aprendizagem da rede selecionada. Na implementação deste algoritmo foi apenas possível aplicar 36 combinações de hiperparâmetros.

Fontes de dados

No presente trabalho foram utilizadas séries temporais das variáveis da altura significativa espectral H_{m0} e do período de pico T_p registadas na boia multiparamétrica de Sines e calculadas pelo modelo numérico *WaveWatch III* para pontos ao largo da Ilha da Madeira e do arquipélago dos Açores.

As primeiras redes neuronais aplicadas utilizaram como valores de entrada (*features*) os previsores considerados significativos pelo modelo de regressão linear.

A variação das fontes dados regeu-se por um critério iterativo, no qual foram progressivamente adicionados aos previsores significativos o conjunto de previsores da boia de Sines, do ponto extraído da malha geográfica do modelo WWIII ao largo do arquipélago dos Açores e da Ilha da Madeira. A sequência de introdução dos previsores no modelo seguiu a ordem crescente correlação deste previsores com a série temporal de H_{m0} extraída da boia multiparamétrica de Sines.

No decorrer da exploração reconheceu-se que a quantidade de dados para treino das redes neuronais pudesse ter impacto nas soluções obtidas no presente trabalho. Por forma a fazer uma avaliação da quantidade de informação necessária para o treino das redes neuronais que preveem a altura significativa espectral, efetuou-se uma avaliação do desempenho de duas redes semelhantes em termos de estrutura e hiperparâmetros, diferindo apenas na quantidade de registos utilizada por cada uma destas redes. Ambas as redes consideram os mesmos previsores (séries temporais registadas pela boia de Sines), porém uma das redes considerou os registos utilizados no desenvolvimento das redes neuronais anteriores (1677 registos com intervalos entre registos de 6 horas), enquanto que a outra rede considerou todos

os registos da boia multiparamétrica de Sines (10152 registos com intervalos entre registos de 1 hora).

Capítulo 4

Resultados

Neste capítulo são apresentados os resultados obtidos decorrentes na investigação realizada que tinha como objetivo o desenvolvimento de modelos de aprendizagem automática para a previsão das condições de agitação marítima. Este capítulo encontra-se subdividido em três secções. A primeira secção diz respeito aos resultados do modelo ARIMA aplicado para previsões de 12, 24 e 48 horas. A segunda secção apresenta os resultados da aplicação do modelo de regressão linear. A última secção apresenta os resultados decorrentes da exploração dos três eixos de exploração do presente trabalho pela seguinte ordem: estrutura, hiperparâmetros da rede neuronal e fontes de dados utilizadas.

4.1 Modelo ARIMA(2,1,5)

No presente trabalho foi aplicado o modelo ARIMA(2,1,5), com o objetivo de obter previsões de 12, 24 e 48 horas para a série temporal H_{m0} registada na boia de Sines. Este modelo é descrito pelos coeficientes e respetivos intervalos de confiança descritos na tabela 4.1.

Parâmetro	Coefficiente	Limite inferior (95%)	Limite superior (95%)
α_1	-1.17	-1.22	-1.12
α_2	-0.94	-0.99	-0.90
θ_1	0.83	0.78	0.88
θ_2	0.64	0.59	0.69
θ_3	-0.12	-0.16	-0.09
θ_4	0.20	0.17	0.23
θ_5	0.11	0.08	0.14
w_t	0.05	0.05	0.06

TABELA 4.1: Coeficientes do Modelo ARIMA com Intervalos de Confiança a 95%

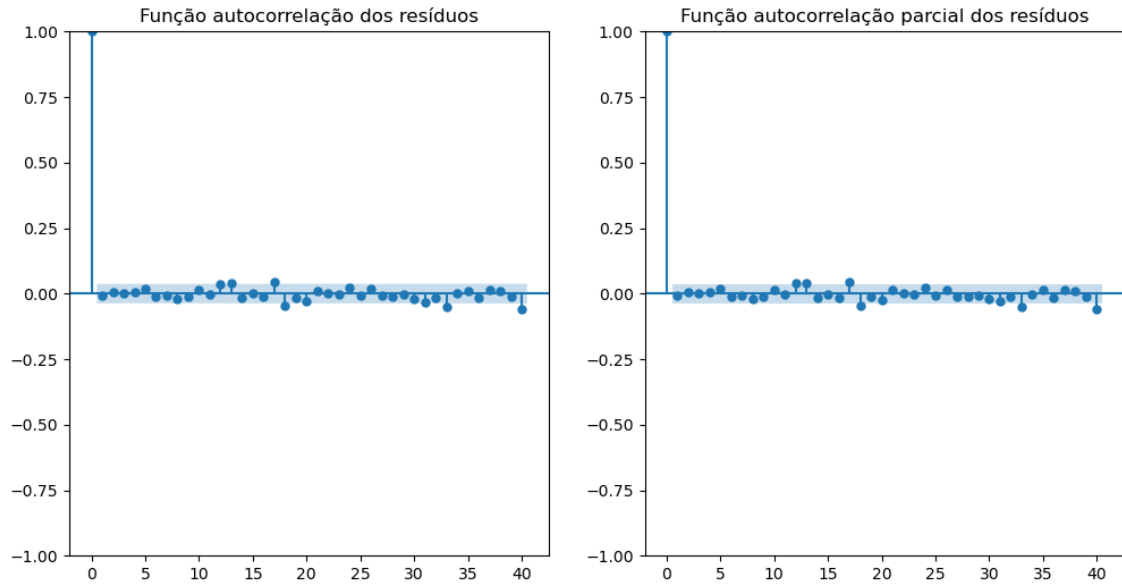


FIGURA 4.1: Função autocorrelação dos resíduos do modelo (lado esquerdo) e função de autocorrelação parcial do modelo (lado direito)

A figura 4.1 representa as funções de autocorrelação e autocorrelação parcial dos resíduos. Nesta figura, observa-se que para desfasamentos (*lags*) iguais ou superiores a 1, os valores encontram-se dentro do intervalo de confiança a 95% em torno de zero, o que indica que os resíduos não estão correlacionados.

O teste de *Ljung-Box* confirmou o resultado deduzido qualitativamente a partir da figura 4.1 (estatística de teste $Q = 4,29$, $p - value = 0,933 > 0,05$), pelo que se conclui que os resíduos não estão correlacionados a um nível de significância de 5%.

A figura 4.2 representa a distribuição dos resíduos do modelo em função do tempo (primeira imagem) e a distribuição estatística dos mesmos num histograma (imagem a baixo). Foi ainda aplicado o teste de *Shapiro-Wilk* com o objetivo de confirmar a normalidade dos resíduos. Conclui-se que existe evidência estatística com um nível de significância de 5% para rejeitar a distribuição normal dos resíduos ($W = 0,92$ com $p - value = 6.7 \times 10^{-37} \leq 0.05$).

A performance do modelo ARIMA(5,1,2) foi medida em termos das métricas de MAE e RMSE, conforme descrito na tabela 4.2, na qual se pode concluir que para intervalos de previsões maiores existe uma maior erro associado MAE e RMSE.

A aplicação do modelo ARIMA(5,1,2) oferece previsões razoáveis para 12 horas (RMSE = 0,7 , MAE= 0,52). Além disso, os resultados da aplicação deste

4.1. Modelo ARIMA(2,1,5)

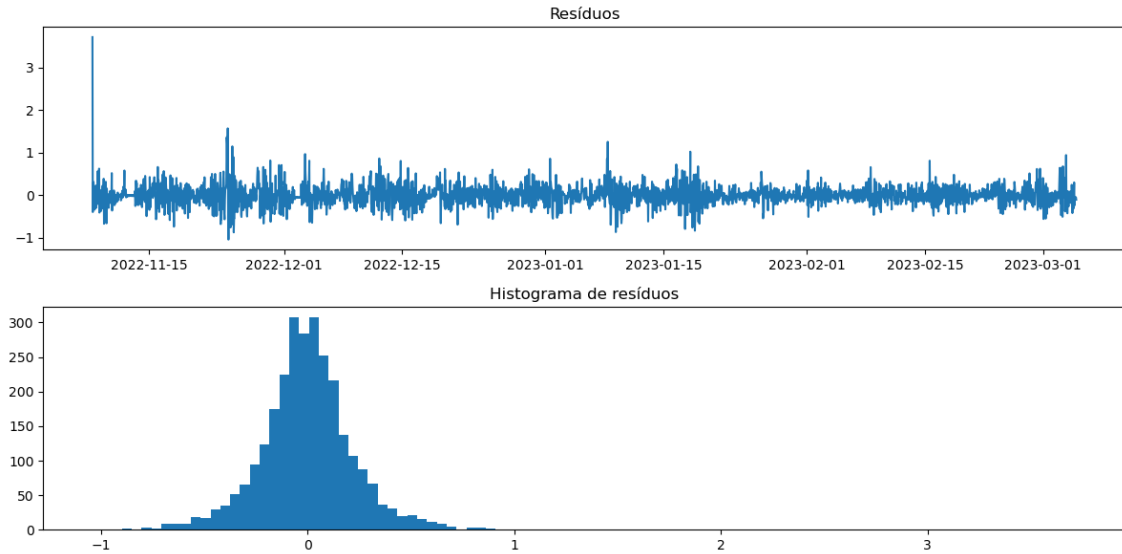


FIGURA 4.2: Distribuição do valor ao longo do tempo (em cima) e histograma dos resíduos (em baixo) do modelo ARIMA(2,1,5)

Previsões	Janela Deslizante	MAE	RMSE
t + 48 h	-72	0,92	1,21
	-60	1,06	1,32
	-48	1,10	1,39
	-24	0,96	1,27
	-12	1,06	1,32
t + 24 h	-72	0,78	0,98
	-60	0,75	0,97
	-48	0,78	0,98
	-24	0,79	1,00
	-12	0,75	0,97
t + 12 h	-72	0,52	0,69
	-60	0,52	0,69
	-48	0,52	0,69
	-24	0,52	0,69
	-12	0,52	0,70

TABELA 4.2: Métricas MAE e RMSE do modelo ARIMA(2,1,5) para janelas deslizantes de 72, 60, 48, 24 e 12 horas e previsões de 12, 24 e 48 horas.

modelo provam que a componente autoregressiva poderá ser significativa para previsões até 12 horas. O número de valores anteriores da série temporal da janela deslizante utilizados para as previsões, não contribui para melhores previsões.

No estudo publicado por Oleg Makarynsky [2005](#), utilizou-se redes neurais de diferentes arquiteturas para prever H_s a 12 e 24 horas recorrendo exclusivamente aos registos da boia ondógrafo de Sines. Nas previsões de 12 horas destas redes neuronais o valor de RMSE obtido foi entre 0,31 e 0,37 m, no caso das previsões de 24 horas, o erro é maior, sendo que o valor de RMSE obtido variou entre 0,53 e 0,77 m. Em comparação com as previsões das redes neuronais desenvolvidas por Oleg Makarynsky [2005](#), o modelo ARIMA(2,1,5) desenvolvido no presente trabalho teve uma performance inferior, obtendo um RMSE de aproximadamente 0,70 m para previsões de 12 horas e de 1 m para previsões de 24 horas.

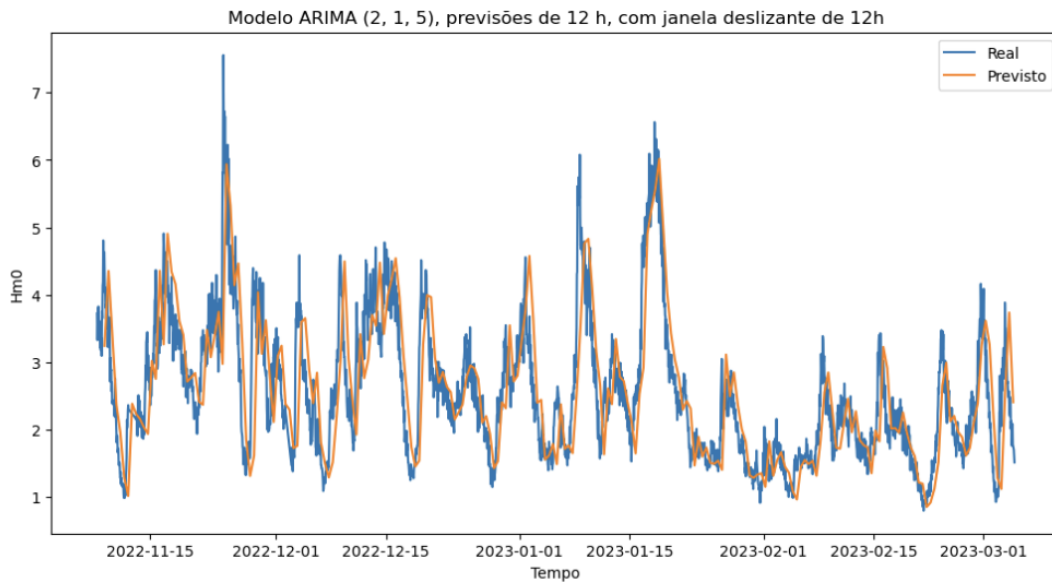


FIGURA 4.3: Valores observados de H_{m0} na boia de Sines e previsões de 12 horas do modelo ARIMA(2,1,5), com uma janela deslizante de 12 horas.

A figura [4.3](#) representa os valores observados pela boia de Sines e as previsões de 12 horas efetuadas pelo modelo ARIMA(5,1,2) utilizando uma janela deslizante correspondente a 12 horas. A análise desta figura permite concluir que o modelo se ajusta bem, porém tem dificuldades na previsão de valores extremos da altura significativa.

A figura [4.4](#) representa os valores observados pela boia de Sines e as previsões de 24 horas efetuadas pelo modelo ARIMA(5,1,2) utilizando uma janela deslizante correspondente a 12 horas. Em comparação com a figura anterior verifica-se que o

4.1. Modelo ARIMA(2,1,5)

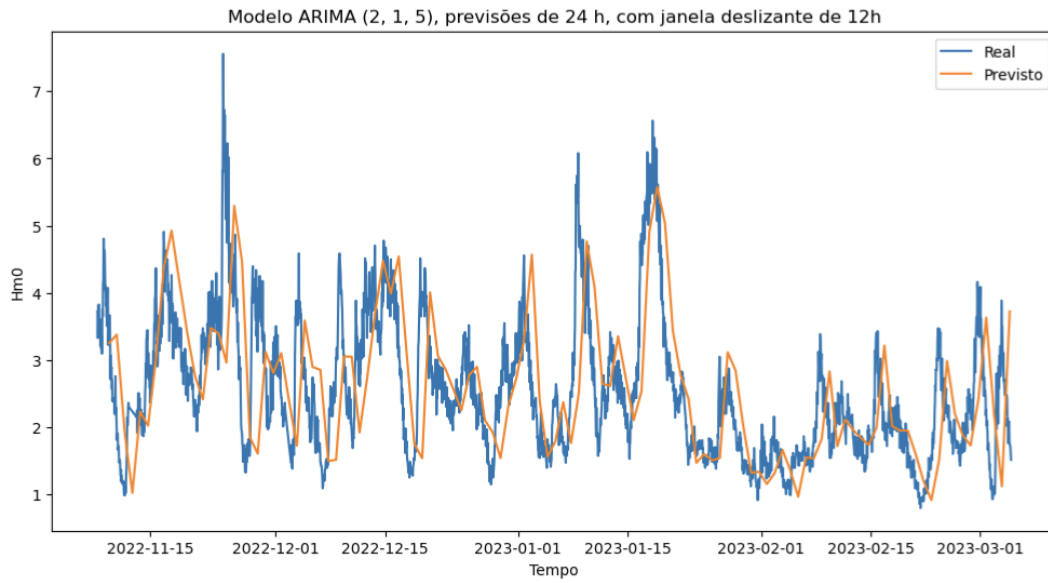


FIGURA 4.4: Valores observados de H_{m0} na boia de Sines e previsões de 24 horas do modelo ARIMA(2,1,5), com uma janela deslizante de 12 horas.

modelo tem uma performance inferior nas suas previsões, em especial na estimativa dos picos de H_{m0} .

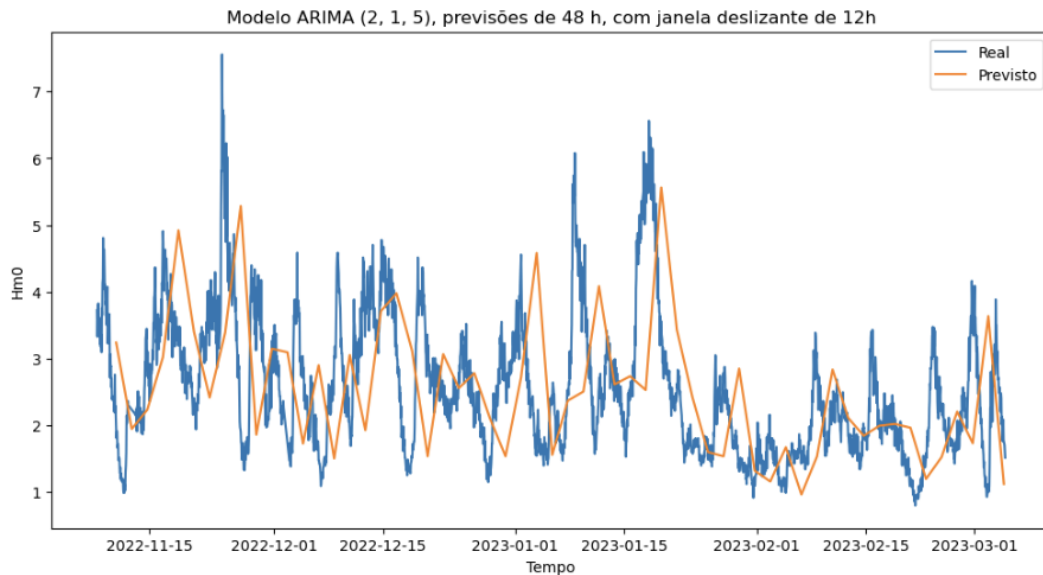


FIGURA 4.5: Valores observados de H_{m0} na boia de Sines e previsões de 48 horas do modelo ARIMA(2,1,5), com uma janela deslizante de 12 horas.

Na figura 4.5, onde é representada uma comparação entre os observados pela boia de Sines e as previsões de 48 horas efetuadas pelo modelo ARIMA(5,1,2) recorrendo a uma janela deslizante de 12 horas, verifica-se que os maiores valores das alturas de H_{m0} (picos) não são previstos com precisão aceitável e existe ainda uma discrepância entre os valores previstos e os valores observados, especialmente em períodos em que ocorrem rápidas variações da altura significativa espectral.

4.2 Modelo de Regressão Linear com variáveis desfasadas (*lagged variables*)

No presente trabalho desenvolveram-se vários modelos de regressão linear que tiveram como objetivos a determinação dos previsores mais significativos para a variável H_{m0} e ainda a avaliação da influência de outras séries temporais como previsores resultava num modelo com um desempenho superior ao dos modelos ARIMA.

Considerou-se como previsores as séries temporais de H_{m0} e T_p e as respectivas variáveis desfasadas de 12 e 24 horas.

Além disso, foi também considerado como previsores as séries temporais extraídas do modelo *WaveWatch III* ao largo do arquipélago dos Açores e da Ilha da Madeira, com os *lags* 12, 24, 48 e 72 horas. Estes previsores foram selecionados a partir da análise de correlação de cruzada entre as séries temporais de H_{m0} e T_p das séries temporais extraídas da malha geográfica do modelo numérico WWIII ao largo do arquipélago dos Açores e da Ilha da Madeira com a série temporal extraída da boia multiparamétrica de Sines, conforme mencionado no capítulo anterior.

A matriz de correlação (figura 3.13) permite avaliar a correlação entre todos os previsores e também com a variável dependente, esta avaliação é importante para garantir que não haja problemas de multicolinearidade.

Após a seleção inicial dos previsores, foi aplicado um factor de escala de normalização de 2,5 metros para as alturas de onda outro de 10 segundos para os períodos.

Inicialmente desenvolveram-se modelos de regressão linear definindo como previsores as séries temporais de H_{m0} e T_p extraídas da boia multiparamétrica de Sines como também as respectivas séries temporais desfasadas (*lagged variables*) a 24 e 48 horas. Posteriormente foram incluídas as séries temporais relativas à altura significativa e período de pico extraídas da malha geográfica do modelo WWIII ao largo do arquipélago dos Açores desfasadas de 72, 48, 24 e 12 horas por esta ordem.

Por fim, considerou-se como previsores dos modelos de regressão linear seguintes, as séries temporais de H_{m0} e T_p extraídas da boia multiparamétrica de Sines e as suas variáveis desfasadas (*lagged variables*) a 24 e 48 horas, e ainda as séries temporais extraídas do modelo numérico ao largo da ilha da Madeira desfasadas de 72, 48, 24 e 12 horas. No apêndice [A](#) é apresentado sequencialmente um sumário de todos os modelos de regressão linear desenvolvidos. Neste sumário inclui-se o valor dos coeficientes de cada predictor, bem como o valor de $p - value$. Durante o desenvolvimento de cada modelo foi feito uma análise do ajustamento do mesmo bem como das significância estatística dos previsores incluídos. Decorrente desta análise desenvolveu-se um modelo de regressão linear no qual se incluiu todos os previsores considerados significativos ($p - value < 0,05$) (figura [4.7](#)).

Modelo de Regressão Linear (Hm0+12) - Previsores significativos						
Dep. Variable:	Y= Hm0	R-squared (uncentered):	0.955			
Model:	OLS	Adj. R-squared (uncentered):	0.955			
Method:	Least Squares	F-statistic:	2178.			
Prob (F-statistic):	1.38e-205					
Log-Likelihood:	8.2003					
AIC:	-10.40					
BIC:	0.7896					
	coef	std err	t	P> t	[0.025	0.975]
Hm0_boia	0.9281	0.038	24.480	0.000	0.853	1.003
Hm0-24	-0.1452	0.037	-3.920	0.000	-0.218	-0.072
Tp_a - 48	0.2364	0.037	6.374	0.000	0.163	0.309
Cond. No.	6.36					

FIGURA 4.6: Modelo de regressão linear para prever a variável H_{m0} no período de 12 horas, considerando os previsores significativos.

Desenvolveu-se ainda uma função com o objetivo de selecionar os previsores mais significativos na previsão de H_{m0} na qual foi utilizada a biblioteca *statsmodels* do *software Python* (Seabold e Perktold [2010](#)).

A função desenvolvida tem como objetivo encontrar o melhor modelo de regressão linear com base no menor valor de AIC selecionando simultaneamente os previsores mais significativos $p - value < 0,05$. Esta função considera inicialmente um modelo sem variáveis independentes e simultaneamente $\beta_0 = 0$. Posteriormente são adicionados e removidos previsores de forma iterativa, avaliando simultaneamente o valor de AIC do modelo e o $p - value$ de cada predictor. Uma variável é removida se o valor $p - value > 0,10$ ou se a sua exclusão diminui o valor de AIC do modelo diminui, e é adicionada ao modelo se a sua adição contribuiu para um valor de AIC mais baixo ou se $p - value < 0,05$. Esta função determinou que os seguintes previsores como os significativos para a previsão de H_{m0} (figura [4.7](#)):

- H_{m0} boia: Corresponde à série temporal da altura significativa espectral registada pela boia de Sines;
- $H_{m0} - 12$: Série temporal que diz respeito à altura significativa espectral registada pela boia de Sines com um atraso de 12 horas;
- T_p a - 48: Série temporal do período de pico calculada pelo modelo WWIII desfasada de 48 horas;
- H_s m - 72: Série temporal da altura significativa espectral calculada pelo modelo WWIII desfasada de 72 horas.

Modelo de Regressão Linear (Hm0+12)						
Dep. Variable:	Y= Hm0	R-squared (uncentered):				0.953
Model:	OLS	Adj. R-squared (uncentered):				0.953
Method:	Least Squares	F-statistic:				1782.
Prob (F-statistic):	2.69e-230	No. Observations:				352
Log-Likelihood:	0.44698					
AIC:	7.106					
BIC:	22.56					
	coef	std err	t	P> t	[0.025	0.975]
Hm0_boia	0.8752	0.049	17.766	0.000	0.778	0.972
Tp_a - 48	0.1646	0.042	3.900	0.000	0.082	0.248
Hs_m - 72	0.4003	0.101	3.971	0.000	0.202	0.599
HM0-12	-0.0727	0.020	-3.694	0.000	-0.111	-0.034
Cond. No.	25.5					

FIGURA 4.7: Modelo de regressão linear para prever a variável H_{m0} no período de 12 horas.

O modelo de regressão linear de previsão de H_{m0} a 12 horas aplicado descreve cerca de 95,3 % da variância da série temporal de H_{m0} extraída da boia de Sines, com valores AIC, BIC e de condição relativamente pequenos, sugerindo que o modelo está bem ajustado e que não existem problemas de multicolinearidade. O valor de $F - statistic$ sugere que o modelo é significativo. Além disso, todos os coeficientes têm intervalos de confiança que não incluem o valor zero, reforçando a hipótese de que as variáveis são estatisticamente significativas (figura 4.7).

A figura 4.8 representa a comparação das previsões do conjunto de teste e dos valores observados ao longo do tempo (no topo à esquerda), o gráfico de dispersão dos valores observados e dos valores previstos relativamente ao conjunto de teste (no topo à direita), o histograma relativamente aos resíduos (em baixo à esquerda) e o gráfico de dispersão dos resíduos e dos valores previstos (em baixo à direita). Para o conjunto de teste, obteve-se como avaliação da performance do modelo, as métricas $MAE = 0,42$, $RMSE = 0,52$.

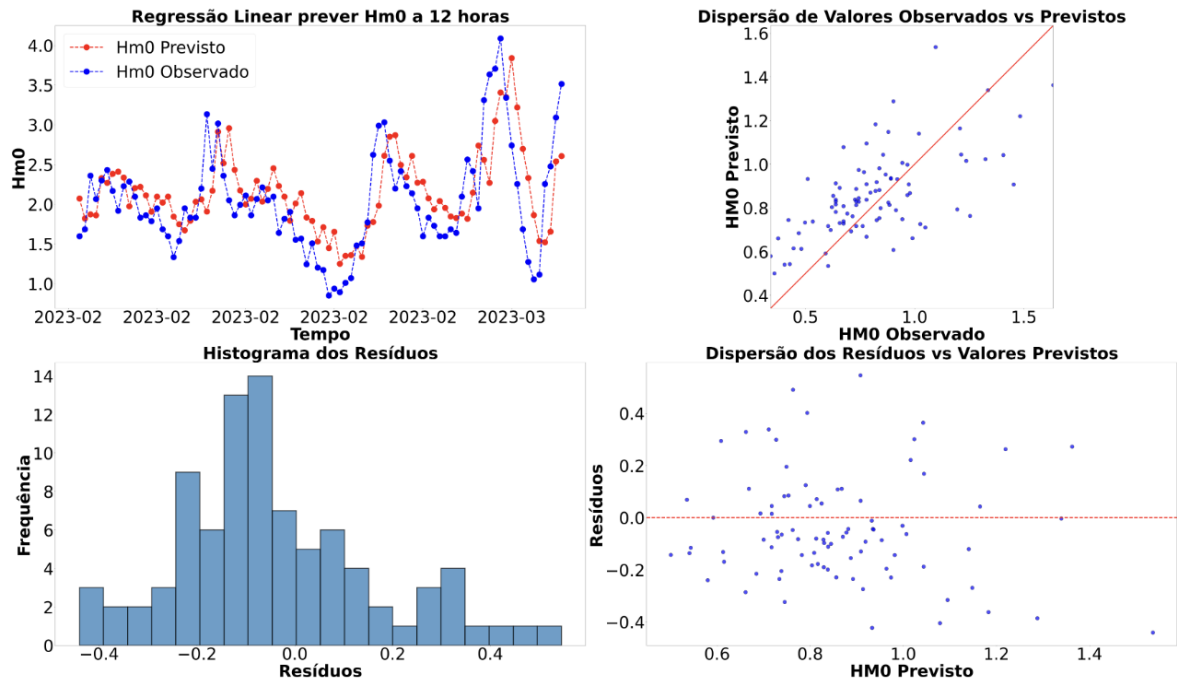


FIGURA 4.8: Resultado da aplicação do modelo de regressão linear ao conjunto de teste

4.3 Modelo de Aprendizagem Automática: Redes neurais

No presente trabalho desenvolveram-se várias redes neurais capazes de prever H_{m0} a 12 horas na área de Sines. O desenvolvimento destas redes seguiu uma abordagem de exploração tripartida, na qual se considerou a variação da estrutura (número de camadas intermédias e número de neurónios nessas camadas), dos hiperparâmetros das redes (*batch size*, regularização, taxa de aprendizagem e épocas de treino) e diversas escolhas de previsores (*features*). Para além da adição de previsores, desenvolveu-se ainda duas redes neurais com estrutura e hiperparâmetros semelhantes, diferindo apenas na quantidade de dados utilizado no processo de aprendizagem, com o objetivo de compreender a influência da quantidade de dados na aprendizagem das redes neurais para o contexto de previsão de parâmetros da agitação marítima.

A abordagem inicial consistiu na determinação da(s) estrutura(s) mais simples e que, simultaneamente, oferecessem menores erros nas previsões de H_{m0} . As estruturas de rede consideradas mantiveram o número de neurónios na camada de entrada igual ao número de variáveis de entrada (*features*) e apenas um neurónio na

camada de saída, por a variável de resposta ser um escalar ($Hm0$).

A determinação da estrutura de rede neuronal mais eficaz consistiu na variação do número de camadas intermédias e o número de neurónios nessas camadas. A primeira estrutura de rede a ser considerada foi a rede (4 X 9 X 1), por respeitar os critérios anteriormente referidos (capítulo 3). Durante a variação do número de camadas e neurónios, manteve-se constante os restantes hiperparâmetros das redes (*batch size*, taxa de aprendizagem, regularização e épocas de treino) bem como as *features* introduzidas na mesma, de modo a permitir comparações corretas entre o desempenho dos vários modelos. A tabela 4.3 ilustra os valores dos hiperparâmetros usados nesta fase da exploração. Os valores de entrada considerados durante a primeira fase de exploração foram os previsores considerados significativos pelo modelo de regressão linear (figura 4.7).

Hiperparâmetro	Valor
<i>Batch size</i>	32
Taxa de aprendizagem	0.001
Regularização L2	0.001
Épocas de treino	250

TABELA 4.3: Valores dos hiperparâmetros da rede neuronal inicial.

Durante a execução das várias experiências das redes neuronais desenvolvidas foi essencial garantir que o valor dos pesos de inicialização fossem os mesmos em todas as experiências. Este processo foi importante para procurar eliminar a influência das condições iniciais na comparação dos diferentes modelos de redes neuronais desenvolvidos.

No desenvolvimento de uma rede neuronal, adicionar camadas intermédias ou neurónios origina um aumento da complexidade da rede, podendo permitir que a rede capte melhor os padrões existentes nos dados ou, por outro lado, provocar problemas de *overfitting*, pelo que é importante analisar tanto as métricas de performance do modelo de AA como as suas curvas de aprendizagem.

A figura 4.9 ilustra os resultados dos 26 testes mais significativos sob a forma de tabela. Os resultados apresentados definem os valores limites de neurónios para a ocorrência de *overfitting*. Nestes testes começou por se testar o número de neurónios essenciais para uma camada intermédia. Posteriormente, adicionou-se de forma progressiva uma camada intermédia e fez-se variar o número de neurónios

nessa camada, determinando o limite de camadas e neurónios para qual ocorria *overfitting*.

Nº camadas intermédias	Estrutura da rede neuronal			
	Estrutura	RMSE teste	MAE teste	Treino na validação
1	4 X 2 X 1	0,74	0,54	Overfitting
	4 X 3 X 1	0,74	0,54	Overfitting
	4 X 4 X 1	0,73	0,54	Overfitting
	4 X 5 X 1	0,74	0,54	Bom ajuste
	4 X 9 X 1	0,73	0,54	Bom ajuste
	4 X 18 X 1	0,73	0,54	Bom ajuste
	4 X 19 X 1	0,73	0,54	Overfitting
2	4 X 2 X 9 X 1	0,74	0,54	Overfitting
	4 X 4 X 9 X 1	0,74	0,54	Overfitting
	4 X 9 X 9 X 1	0,73	0,54	Overfitting
	4 X 10 X 9 X 1	0,73	0,54	Overfitting
	4 X 18 X 9 X 1	0,73	0,54	Overfitting
	4 X 9 X 18 X 1	0,73	0,54	Bom ajuste
	4 X 9 X 25 X 1	0,73	0,54	Overfitting
	4 X 9 X 2 X 1	0,74	0,54	Bom ajuste
	4 X 9 X 1 X 1	0,74	0,54	Bom ajuste
3	4 X 4 X 9 X 2 X 1	0,73	0,54	Overfitting
	4 X 4 X 9 X 3 X 1	0,74	0,54	Bom ajuste
	4 X 9 X 9 X 3 X 1	0,73	0,54	Overfitting
	4 X 9 X 9 X 4 X 1	0,73	0,54	Overfitting
	4 X 4 X 9 X 4 X 1	0,73	0,54	Overfitting
	4 X 9 X 9 X 9 X 1	0,73	0,54	Bom ajuste
	4 X 9 X 9 X 10 X 1	0,73	0,54	Bom ajuste
	4 X 9 X 9 X 17 X 1	0,73	0,54	Bom ajuste
	4 X 9 X 9 X 18 X 1	0,73	0,54	Overfitting
	4 X 10 X 9 X 3 X 1	0,73	0,54	Overfitting

FIGURA 4.9: Resultados da exploração manual da estrutura das redes neuronais

A partir da análise da tabela anterior, conclui-se que a adição e remoção de camadas intermédias e/ou de neurónios nessas mesmas camadas não apresenta uma melhoria significativa nas métricas de performance do modelo RMSE e MAE. Paralelamente, foram analisadas as curvas de aprendizagem para cada uma das redes, na qual se identificou que para algumas dessas redes ocorrem fenómenos de *overfitting*. Apesar de existirem estruturas de rede compostas por três camadas intermédias que podem ser treinadas sem *overfitting*, verificou-se que a curva de aprendizagem do conjunto de validação apresentava algum ruído, sugerindo assim que a rede neuronal se estivesse a tornar excessivamente complexa para o problema em causa.

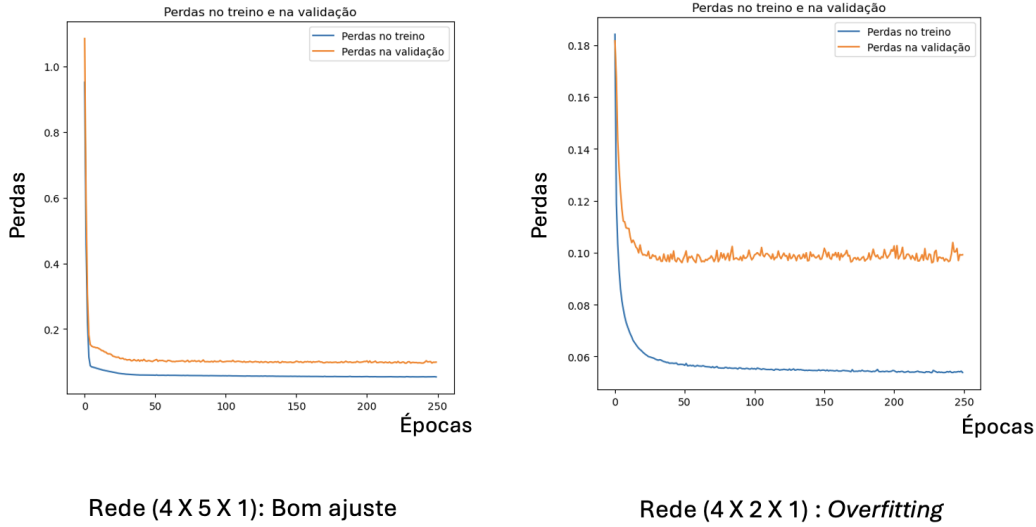


FIGURA 4.10: Exemplo de um bom ajuste e de *overfitting* no treino de redes neuronais: curvas de aprendizagem da rede (4 X 2 X 1) à esquerda e a rede (4 X 5 X 1) à direita.

A figura 4.10 apresenta dois exemplos de fenómenos diferentes das curvas de aprendizagem. Quando as perdas da função de custo no conjunto de validação são pequenas, tal como é apresentado no gráfico da esquerda da figura 4.10 então o modelo generaliza bem os resultados obtidos, pelo que será expectável que tenha uma boa performance no conjunto de teste. No caso de as perdas da função de custo no conjunto de validação forem muito mais elevadas, tal como é apresentado no gráfico da direita da figura 4.10, então tem *overfitting*.

Após a exploração manual da estrutura da rede neuronal, utilizou-se o algoritmo *Random Search* para determinar a estrutura de redes neuronais. Testar todas as combinações de hiperparâmetros possíveis seria impraticável, por isso foram fixados os valores dos hiperparâmetros descritos na tabela 4.4, bem como um número máximo de iterações. Definiu-se assim, que o algoritmo testaria 5% das combinações de hiperparâmetros de um modo aleatório, o que corresponde a 36 de 720 testes dos hiperparâmetros possíveis a partir da tabela 4.4.

Hiperparâmetros	Valores possíveis
Número de camadas intermédias	2; 3; 4; 5
Número de neurónios nas camadas intermédias	4; 8; 16; 32; 64
<i>Batch size</i>	16; 32; 64; 128
Taxa de aprendizagem	0,001; 0,01; 0,1
Regularização L2	0,001; 0,01; 0,1

TABELA 4.4: Valores possíveis para os hiperparâmetros no algoritmo *Random Search*

O resultado da aplicação do algoritmo *Random Search* foi uma estrutura de rede composta por 2 camadas intermédias com 64 e 8 neurónios, respetivamente, com os seguintes hiperparâmetros:

- *Batch size* = 128
- Taxa de Aprendizagem = 0,1
- Regularização = 0,001

A rede neuronal obtida pela aplicação do algoritmo *Random Search* foi testada com o mesmo número de épocas que as redes neuronais testadas inicialmente e, posteriormente, com o número de épocas definido pela função *EarlyStopping*. A tabela seguinte apresenta os resultados das métricas RMSE e MAE nos conjuntos de teste da rede (4 X 64 X 8 X 1) com treino de 250 épocas e com a interrupção de treino provocada pela função *EarlyStopping*.

RANDOM SEARCH	RMSE	MAE	Treino na validação
4 X 64 X 8 X 1 [épocas = 250]	1,3	0,97	Bom ajuste
4 X 64 X 8 X 1 [EarlyStopping]	1,5	1,1	Bom ajuste

FIGURA 4.11: Resultados da rede (4 X 64 X 8 X 1) com treino de 250 épocas e com interrupção de treino provocada pela função *EarlyStopping*.

A rede considerada adequada pelo algoritmo *Random Search* além de mais complexa que as redes exploradas manualmente não é mais precisa, pelo o valor de RMSE e MAE ser mais elevado em comparação com as redes anteriormente testadas. Considerou-se a rede (4 X 5 X 1) a estrutura mais simples e que simultaneamente oferecesse maior precisão no conjunto de teste, sem *overfitting*. Por essa razão, optou-se

pelo uso desta estrutura para explorar os hiperparâmetros que otimizavam a aprendizagem da rede, levando assim ao segundo eixo de exploração: Hiperparâmetros da rede neuronal (*batch size*, taxa de aprendizagem, regularização).

Na exploração dos hiperparâmetros da rede neuronal (4 X 5 X 1) considerou-se a variação do *batch size*, da taxa de aprendizagem, da regularização e das épocas de treino isoladamente.

O primeiro hiperparâmetro analisado foi o *batch size*, mantendo a taxa de aprendizagem, a regularização e as épocas de treino fixos conforme a tabela 4.3. O valor das épocas de treino pode influenciar a performance de todos os outros hiperparâmetros da rede. Neste ponto da investigação, o valor otimizado de épocas ainda não fora determinado. Por esta razão, foram efetuadas duas experiências: a primeira com 250 épocas de treino e a segunda com as épocas de treino interrompidas pela função *EarlyStopping*. A tabela seguinte apresenta os resultados obtidos.

<i>Batch size</i>					
Métricas	8	16	32	64	128
RMSE (Teste)	0,74	0,74	0,74	0,74	0,77
MAE (Teste)	0,55	0,54	0,54	0,55	0,52
RMSE (Validação)	0,8	0,78	0,78	0,74	0,74
MAE (Validação)	0,51	0,51	0,51	0,52	0,55
Observações	Bom ajuste	Bom ajuste	Bom ajuste	Bom ajuste	Bom ajuste
<i>Resultados com a função EarlyStopping</i>					
RMSE (Teste)	0,73	0,73	0,73	0,75	0,75
MAE (Teste)	0,56	0,56	0,56	0,57	0,57
RMSE (Validação)	0,76	0,77	0,75	0,78	0,78
MAE (Validação)	0,53	0,53	0,52	0,54	0,53
Observações	Bom ajuste	Bom ajuste	Bom ajuste	Bom ajuste	Bom ajuste

FIGURA 4.12: Resultados da variação do *batch size* na rede (4 X 5 X 1) com treino de 250 épocas e com interrupção de treino provocada pela função *EarlyStopping*.

A partir da análise dos resultados obtidos conclui-se que a variação do *batch size* para uma rede neuronal (4 X 5 X 1) com os hiperparâmetros definidos na tabela 4.4 e as *features* consideradas como significativas pelo modelo de regressão linear desenvolvido no presente trabalho, não altera significativamente os valores das métricas RMSE e MAE quer no conjunto de teste como de validação. Ao considerar-se *batch size* de menores (8, 16, 32) o modelo ajustará mais vezes os

pesos sinápticos da rede, o que poderá permitir ao modelo captar mais facilmente os padrões nos dados. Por outro lado, utilizar um valor superior para o *batch size*, como por exemplo 64 ou 128, poderá diminuir o tempo de treino da rede neuronal. Adotou-se o valor de *batch size* = 32 porque a curva de aprendizagem não apresentou melhorias significativas para valores de *batch size* superiores, pelo que se considerou que este valor seria adequado para que o modelo conseguisse captar padrões nos dados.

Taxa de Aprendizagem			
Métricas	0,01	0,001	0,0001
RMSE (Teste)	0,74	0,73	0,74
MAE (Teste)	0,54	0,54	0,55
RMSE (Validação)	0,77	0,77	0,77
MAE (Validação)	0,5	0,5	0,52
Obsrrvações	Overfitting	Bom ajuste	Bom ajuste
Resultados com a função <i>EarlyStopping</i>			
RMSE (Teste)	0,72	0,73	0,74
MAE (Teste)	0,56	0,56	0,56
RMSE (Validação)	3,89	0,53	0,53
MAE (Validação)	4,64	0,76	0,78
Observações	Overfitting	Bom ajuste	Bom ajuste

FIGURA 4.13: Resultados da variação da taxa de aprendizagem na rede (4 X 5 X 1) com treino de 250 épocas e com interrupção de treino provocada pela função *EarlyStopping*.

O segundo hiperparâmetro explorado foi a taxa de aprendizagem, mantendo os restantes parâmetros constantes. A figura [4.13](#) apresenta os resultados obtidos com a variação da taxa de aprendizagem.

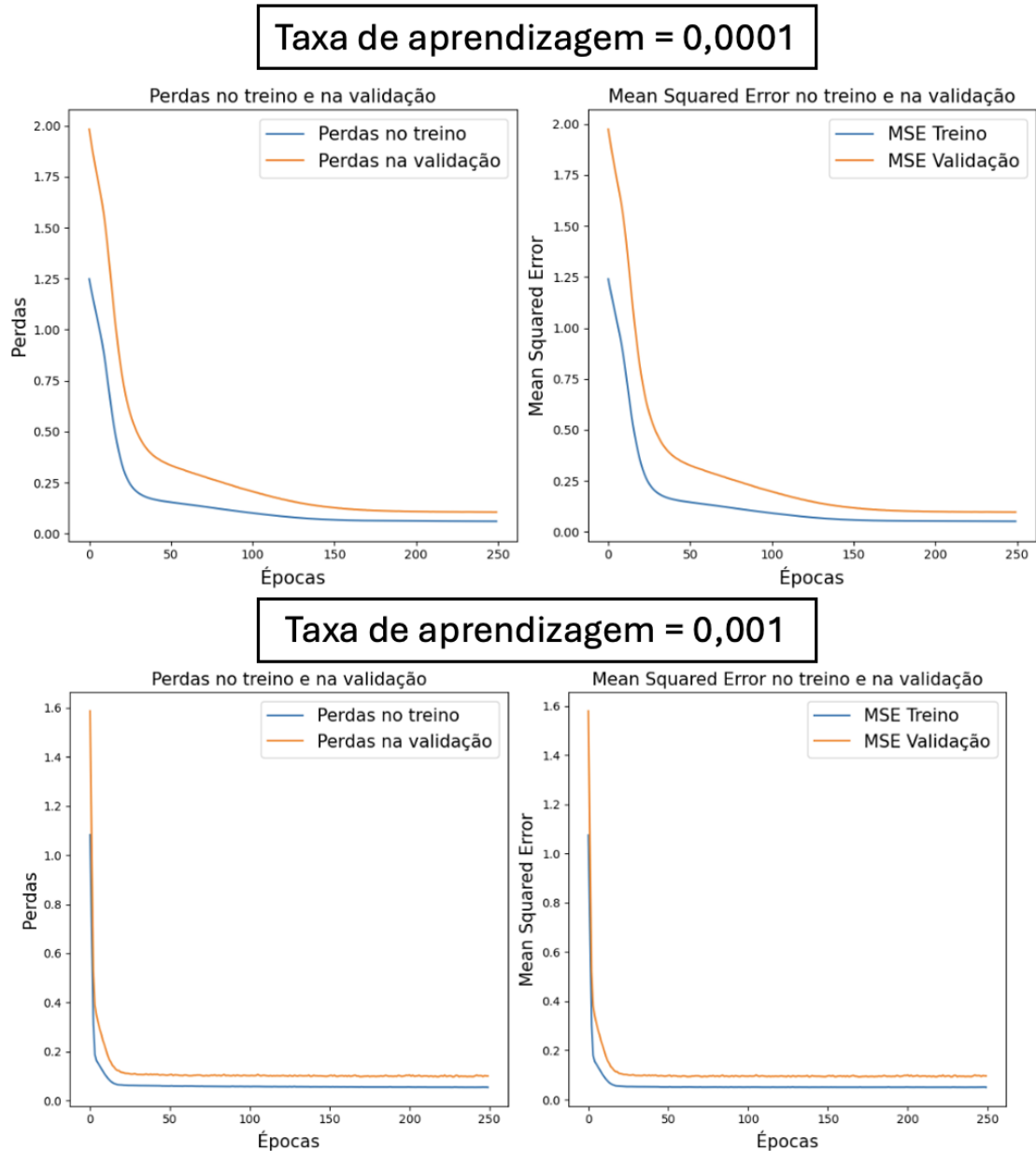


FIGURA 4.14: Curvas de aprendizagem da rede (4 X 5 X 1) com taxa de aprendizagem 0,001 e 0,0001

Os valores das métricas de MAE e RMSE são semelhantes tanto para o conjunto de teste e de validação para as taxas de aprendizagem de 0,001 e 0,0001. Porém, a curva de aprendizagem no conjunto de validação do modelo com a taxa de aprendizagem de 0,001 em comparação com a curva do modelo com a taxa de aprendizagem de 0,0001, tem valores de perdas menores durante as 250 épocas de treino, o que sugere que este modelo tem um desempenho melhor, e por esta razão selecionou-se o valor de 0,001.

Foram ainda testados valores de regularização L2 de modo a otimizar a aprendizagem da rede (4 X 5 X 1). A tabela apresentada na figura 4.15 apresenta os resultados obtidos da variação do fator de regularização, para a rede (4 X 5 X 1) na qual se verificou um bom ajuste para valores de regularização L2 entre 0,1 e 0,0001.

Métricas	Regularização L2			
	0,1	0,01	0,001	0,0001
RMSE (Teste)	0,8	0,74	0,74	0,73
MAE (Teste)	0,6	0,54	0,55	0,54
RMSE (Validação)	0,83	0,76	0,77	0,78
MAE (Validação)	0,55	0,5	0,51	0,52
Observações	Bom ajuste	Bom ajuste	Bom ajuste	Bom ajuste
<i>Resultados com a função EarlyStopping</i>				
RMSE (Teste)	0,81	0,73	0,73	0,73
MAE (Teste)	0,61	0,55	0,56	0,56
RMSE (Validação)	0,84	0,75	0,76	0,76
MAE (Validação)	0,56	0,52	0,53	0,52
Observações	Bom ajuste	Bom ajuste	Bom ajuste	Bom ajuste

FIGURA 4.15: Resultados da variação do fator de regularização da rede (4 X 5 X 1) com treino de 250 épocas e com interrupção de treino provocada pela função *EarlyStopping*.

Na análise das curvas de aprendizagem dos vários modelos desenvolvidos, não se verificou diferenças significativas entre os modelos desenvolvidos com valor de regularização L2 entre 0,01 e 0,0001, pelo que se optou por seleccionar o valor de 0,01.

Além da regularização, do *batch size* e da taxa de aprendizagem testou-se também a performance das previsões da rede neuronal (4 X 5 X 1) ao variar as épocas de treino.

Métricas	Épocas de treino					
	10	50	100	250	1000	10000
RMSE (Teste)	1,5	0,74	0,74	0,74	0,74	0,74
MAE (Teste)	1,4	0,56	0,55	0,55	0,55	0,55
RMSE (Validação)	1,27	0,78	0,78	0,76	0,71	0,71
MAE (Validação)	0,87	0,52	0,52	0,51	0,51	0,51
Observações	Underfitting	Bom ajuste	Bom ajuste	Bom ajuste	Bom ajuste	Bom ajuste

FIGURA 4.16: Resultados da variação das épocas de treino na rede (4 X 5 X 1).

A tabela ilustrada na figura 4.16 apresenta os resultados obtidos, na qual se obteve valores de RMSE e MAE bastante semelhantes tanto no conjunto de validação como no conjunto de teste. Os resultados obtidos mostram que o modelo tem uma boa performance quando treinado com apenas 50 épocas.

A tabela 4.5 apresenta os valores dos hiperparâmetros para uma rede neuronal (4 X 5 X 1) com menor erro associado, e a tabela 4.6 apresenta as métricas de desempenho obtidas por este modelo no conjunto de teste.

Hiperparâmetros	Valores possíveis
Batch size	32
Taxa de aprendizagem	0,001
Regularização L2	0,01

TABELA 4.5: Hiperparâmetros selecionados para a rede (4 X 5 X 1)

Métricas	Performance no conjunto de teste
RMSE	0,73
MAE	0,54

TABELA 4.6: Métricas de performance da rede neuronal (4 X 5 X 1) e hiperparâmetros explorados manualmente.

A figura seguinte apresenta uma comparação entre os valores observados e os valores previstos da série de teste pela rede neuronal (4 X 5 X 1) com os hiperparâmetros da tabela 4.5. A série de teste é constituída por registos entre 28-12-2023 00:00:00 a 29-02-2024 06:00:00 (253 exemplos com intervalos de amostragem entre registos de 6 horas) o que corresponde a aproximadamente 63 dias de previsões da variável H_{m0} .

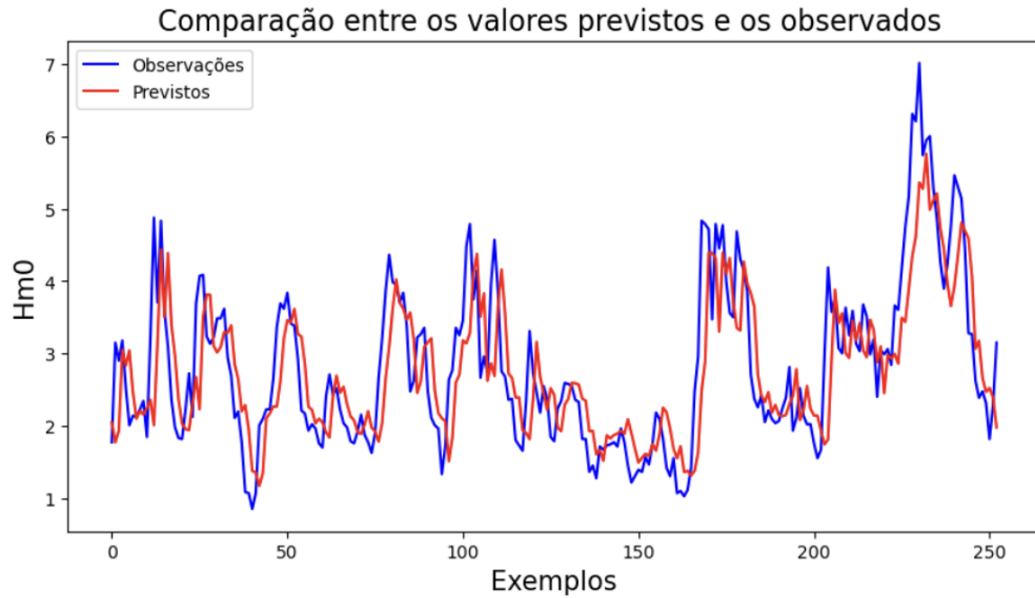


FIGURA 4.17: Valores observados e valores previstos pela rede (4 X 5 X 1).

As curvas de aprendizagem da rede neuronal anteriormente descritas estão representadas na figura [4.18](#). O gráfico da esquerda representa as perdas da função de custo ao longo das épocas no conjunto de treino e de validação e o gráfico da direita representa o MSE ao longo das épocas no conjunto de treino e de validação. A partir da análise da figura [4.17](#) é possível concluir que a o modelo prevê H_{m0} com algum desfasamento associado, e que tem menor erro nas previsões de valores extremos ("picos").

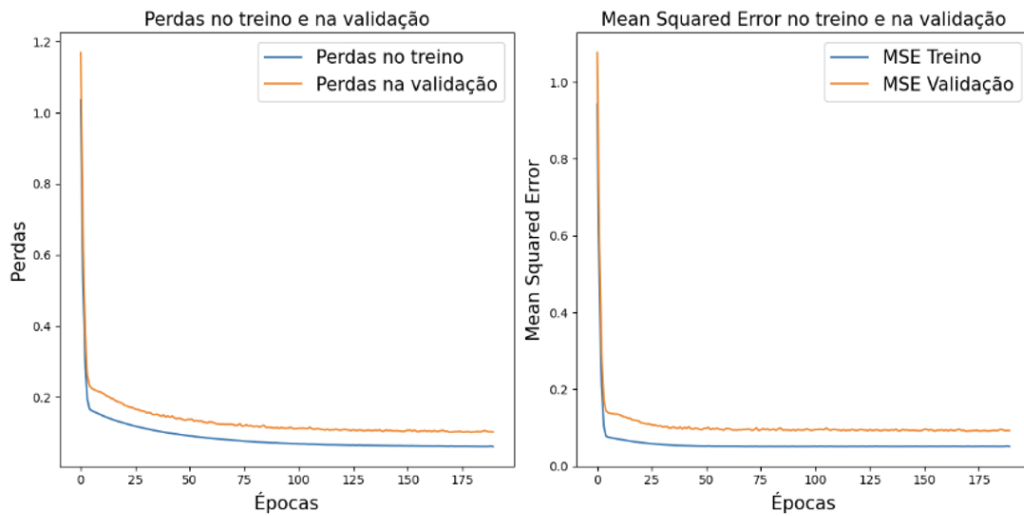


FIGURA 4.18: Gráfico do valor de perda da função de custo em função das épocas de treino (à esquerda) e gráfico do valor de MSE da rede (4 X 5 X 1) (à direita).

A análise da figura 4.18 permite concluir que o modelo não apresenta fenômenos de *overfitting*.

O algoritmo *Random Search* foi novamente utilizado para determinar, através de combinações aleatórias, os hiperparâmetros que otimizam a aprendizagem da rede (4 X 5 X 1). O algoritmo utilizou 36 combinações aleatórias dos valores dos hiperparâmetros: *batch size*, taxa de aprendizagem e regularização apresentados na tabela 4.4. A solução oferecida pelo algoritmo apresenta-se na tabela 4.7.

Hiperparâmetros	Valores possíveis
Batch size	128
Taxa de aprendizagem	0,01
Regularização L2	0,001

TABELA 4.7: Valores dos hiperparâmetros devolvidos pelo algoritmo *Random Search* para uma rede 4 X 5 X 1

A solução do algoritmo *Random Search* permitiu testar a aprendizagem da rede (4 X 5 X 1) com os hiperparâmetros por este determinados, para um intervalo compreendido entre 10 a 1000 épocas de treino, bem com as épocas de treino determinadas pela função *EarlyStopping*. A tabela da figura 4.19 apresenta os resultados obtidos.

Métricas	Épocas de treino - Random Search				
	10	50	100	250	1000
RMSE (Teste)	0,73	0,73	0,73	0,74	0,79
MAE (Teste)	0,54	0,57	0,54	0,55	0,57
RMSE (Validação)	0,77	0,76	0,76	0,76	0,82
MAE (Validação)	0,5	0,55	0,5	0,51	0,51
Observações	Bom ajuste	Bom ajuste	Overfitting	Overfitting	Overfitting
Resultados com a função EarlyStopping					
RMSE (Teste)	0,73				
MAE (Teste)	0,56				
RMSE (Validação)	0,75				
MAE (Validação)	0,53				
Observações	Bom ajuste				

FIGURA 4.19: Resultados da performance da rede neuronal (4 X 5 X 1) e hiperparâmetros selecionados pelo algoritmo *Random Search* com várias épocas de treino e de acordo com a função *EarlyStopping*.

A figura 4.20 apresenta uma comparação entre os registos dos valores observados e dos valores previstos da série de teste da rede neuronal (4 X 5 X 1), com as épocas de treino definidas pela função *EarlyStopping*. Na qual é possível observar que o desempenho nas previsões de H_{m0} parecem ser muito semelhantes ao modelo anteriormente analisado.

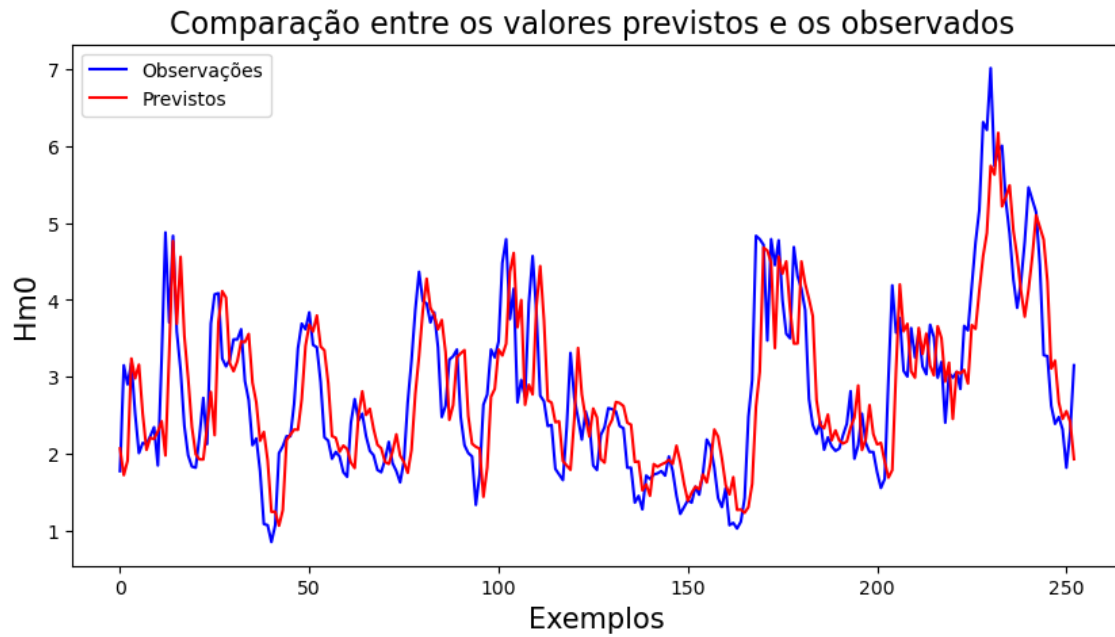


FIGURA 4.20: Resultados da performance da rede neuronal (4 X 5 X 1) e hiperparâmetros selecionados pelo algoritmo *Random Search* com várias épocas de treino e de acordo com a função *EarlyStopping*.

As curvas de aprendizagem deste modelo são apresentadas na figura [4.21](#). Na curva do treino do conjunto de validação deste modelo é possível observar que o modelo parecer ser demasiado complexo para o problema em questão, pelo facto de se observar a ocorrência de *overfitting*.

Comparativamente com os resultados obtidos anteriormente no desenvolvimento da rede neuronal (4 X 5 X 1) decorrente da exploração manual, conclui-se que a precisão dos resultados nas previsões de H_{m0} são bastante semelhantes, apesar de que a rede neuronal (4 X 5 X 1) com os hiperparâmetros selecionados pelo algoritmo *Random Search* tenham provocado fenómenos de *overfitting*.

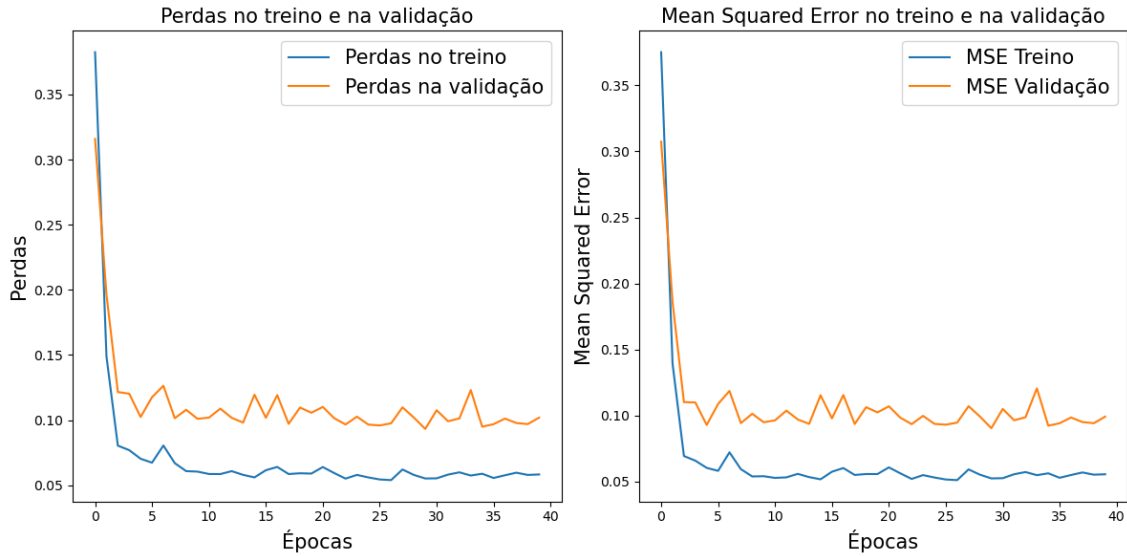


FIGURA 4.21: Resultados da performance da rede neuronal (4 X 5 X 1) e hiperparâmetros seleccionados pelo algoritmo *Random Search* com várias épocas de treino e de acordo com a função *EarlyStopping*.

A curva de aprendizagem no conjunto de validação desta rede neuronal em comparação com as curvas de aprendizagem da rede neuronal anterior (4 X 5 X 1) com os hiperparâmetros definidos pela tabela 4.5, sugere que o este modelo é sujeito ao fenómeno de *overfitting*.

Na terceira fase da exploração, procurou-se averiguar a influência do número e tipo de *features* no desempenho das redes neuronais. Nesta fase, foram considerados não apenas os previsores (*features*) identificados como significativos nos modelos de regressão linear, mas também outros, uma vez que estes poderiam mostrar-se significativos ao considerar relações não lineares.

A variação dos conjuntos de previsores considerados nas redes neuronais teve como objetivo identificar quais as variáveis que garantem previsões fiáveis da altura significativa, tendo estas sido implementadas seguindo um método incremental, que consistiu na adição progressiva de outras séries temporais provenientes da boia de Sines e dos dois pontos extraídos da malha geográfica do modelo *WaveWatch III*. Inicialmente, foram considerados os previsores relativos às séries temporais da boia de Sines. De seguida, adicionou-se as séries temporais da altura significativa e período de pico extraídas da malha do modelo *WaveWatch III* ao largo do arquipélago dos Açores. Por último, as redes neuronais incluíram também as séries temporais da altura significativa e período de pico extraídas deste modelo ao largo da ilha da Madeira.

No final, procedeu-se à comparação entre duas redes neuronais semelhantes em estrutura, hiperparâmetros e previsores, diferindo apenas na quantidade de informação processada, com o objetivo de compreender a quantidade de registos necessário para o treino adequado dos modelos. Na primeira rede neuronal desenvolvida foi utilizada a mesma quantidade de registos que as redes anteriormente (1677 registos com intervalos entre registos de 6 horas) desenvolvidas, enquanto que na segunda rede neuronal se utilizou todos os registos da boia multiparamétrica de Sines (10152 registos com intervalos entre registos de 1 hora).

Na exploração das fontes de dados, considerou-se duas estruturas de rede: a rede (4 X 5 X 1) aplicada anteriormente, e outra rede neuronal na qual as camadas de entrada correspondem ao número de *features* e o número de camadas intermédias respeita o critério anteriormente definido na secção 3.4.3. Considerou-se ainda o treino das várias redes neuronais para 250 épocas de treino como também com a função *EarlyStopping*. Os hiperparâmetros selecionados para otimizar a aprendizagem das redes foram os definidos pela exploração manual (tabela 4.5) por terem atingido uma performance ligeiramente melhor que o resultado obtido pela exploração de hiperparâmetros automática. Os previsores significativos incluíam as séries temporais de H_{m0} e $H_{m0} - 12$ relativas à boia de Sines. Pelo que a esses previsores bastou adicionar-se as restantes séries temporais: $H_{m0} - 24$, $H_{m0} - 48$, $T_p - 12$, $T_p - 24$ e $T_p - 48$. A performance do modelo de AA para estas fontes de dados, apresenta-se na tabela 4.22.

Os segundos previsores a incluir, foram as séries temporais relativas ao ponto extraído ao largo do arquipélago dos Açores, extraído da malha geográfica do modelo WWIII. As redes anteriormente desenvolvidas já incluíam a variável T_p a - 48. Assim, bastou-se adicionar aos previsores anteriormente considerados as variáveis: H_s a - 12, H_s a - 24, H_s a - 48, T_p a - 12, T_p a - 24 e T_p a - 72. A tabela 4.23 apresenta os resultados das métricas de performance nas duas redes neuronais consideradas.

Boia de Sines		
Arquitetura	4 X 5 X 1	10 X 21 X 1
RMSE (Teste)	0,76	0,75
MAE (Teste)	0,51	0,55
RMSE (Validação)	0,74	0,76
MAE (Validação)	0,55	0,51
Observações	Bom ajuste	Bom ajuste
Com a função <i>EarlyStopping</i>		
Arquitetura	4 X 5 X 1	10 X 21 X 1
RMSE (Teste)	0,72	0,76
MAE (Teste)	0,55	0,53
RMSE (Validação)	0,75	0,73
MAE (Validação)	0,52	0,56
Observações	Bom ajuste	Bom ajuste

FIGURA 4.22: Resultados da performance da rede neuronal (4 X 5 X 1) e (10 X 21 X 1) com 250 épocas de treino e de acordo com a função *EarlyStopping*.

Boia de Sines e Açores		
Arquitetura	4 X 5 X 1	17 X 35 X 1
RMSE (Teste)	0,73	0,74
MAE (Teste)	0,55	0,55
RMSE (Validação)	0,77	0,77
MAE (Validação)	0,52	0,52
Observações	Bom ajuste	Overfitting
Com a função <i>EarlyStopping</i>		
Arquitetura	4 X 5 X 1	17 X 35 X 1
RMSE (Teste)	0,73	0,76
MAE (Teste)	0,55	0,56
RMSE (Validação)	0,77	0,77
MAE (Validação)	0,52	0,54
Observações	Bom ajuste	Overfitting

FIGURA 4.23: Resultados da performance da rede neuronal (4 X 5 X 1) e (17 X 35 X 1) com 250 épocas de treino e de acordo com a função *EarlyStopping*.

Os últimos previsores a adicionar foram as séries temporais extraídas da malha geográfica do modelo WWIII ao largo da ilha da Madeira. Assim, esta rede neuronal considerou os previsores anteriormente mencionados e ainda: H_s m - 12, H_s m - 24, H_s m - 48, H_s m - 72, T_p m - 12, T_p m - 24, T_p m - 48 e T_p m - 72.

Boia de Sines , Açores e Madeira		
Arquitetura	4 X 5 X 1	24 X 49 X 1
RMSE (Teste)	0,74	0,73
MAE (Teste)	0,55	0,52
RMSE (Validação)	0,77	0,74
MAE (Validação)	0,51	0,55
Observações	Bom ajuste	Overfitting
Com a função <i>EarlyStopping</i>		
Arquitetura	4 X 5 X 1	24 X 49 X 1
RMSE (Teste)	0,73	0,73
MAE (Teste)	0,56	0,55
RMSE (Validação)	0,76	0,76
MAE (Validação)	0,53	0,53
Observações	Bom ajuste	Overfitting

FIGURA 4.24: Resultados da performance da rede neuronal (4 X 5 X 1) e (24 X 49 X 1) com 250 épocas de treino e de acordo com a função *EarlyStopping*.

Por fim, procedeu-se à comparação entre duas redes neuronais semelhantes em estrutura, hiperparâmetros e previsores, diferindo apenas na quantidade de informação processada. Ambas as redes neuronais desenvolvidas eram constituídas por 8 neurónios na camada de entrada, por uma só camada intermédia com 17 neurónios (respeitando o critério anteriormente mencionado) e 1 neurónio na camada de saída. Os hiperparâmetros adotados por estas redes foram os hiperparâmetros definidos inicialmente que constam na tabela 4.3. Uma das redes neuronais aplicadas processou o mesmo número de registos utilizados pelas redes neuronais anteriormente exploradas (1677 registos para treino, validação e teste da rede), enquanto que a outra rede neuronal desenvolvida processou os registos horários extraídos da boia multiparamétrica de Sines (10152 registos para treino, validação e teste da rede). A razão pela qual as redes aplicadas anteriormente têm menos registos disponíveis é devido aos resultados do modelo numérico WWIII serem extraídos apenas para os

períodos de 0, 6, 12 e 18 horas, pelo que foi necessário considerar apenas os registos temporais da boia multiparamétrica de Sines coincidentes com estes períodos.

A rede neuronal que processou apenas os registos da série temporal extraída da boia de Sines utilizou 10152 registos, de frequência horária.

A tabela da figura 4.25 apresenta os resultados das métricas obtidas das duas redes que processou os registos originalmente extraídos da boia multiparamétrica de Sines (10152 registos). Nesta imagem é possível verificar que a rede que processou mais informação teve uma performance superior.

Rede 8 X 17 X 1 (Registos da boia de Sines)		
Informação processada	Registos com intervalos de 6 horas	Registos horários [originais]
RMSE (Teste)	0,73	0,19
MAE (Teste)	0,55	0,14
RMSE (Validação)	0,73	0,16
MAE (Validação)	0,51	0,12
Observações	Bom ajuste	Bom ajuste
Com a função <i>EarlyStopping</i>		
Informação processada	Registos com intervalos de 6 horas	Registos horários [originais]
RMSE (Teste)	0,73	0,20
MAE (Teste)	0,56	0,15
RMSE (Validação)	0,75	0,17
MAE (Validação)	0,53	0,14
Observações	Bom ajuste	Bom ajuste

FIGURA 4.25: Resultados da performance da rede neuronal (8 X 17 X 1) com 250 épocas de treino e de acordo com a função *EarlyStopping*.

A figura 4.26 representa as curvas de aprendizagem da rede neuronal (8 X 17 X 1) na qual é possível observar um ajuste superior do modelo em comparação com as redes neuronais anteriormente apresentadas.

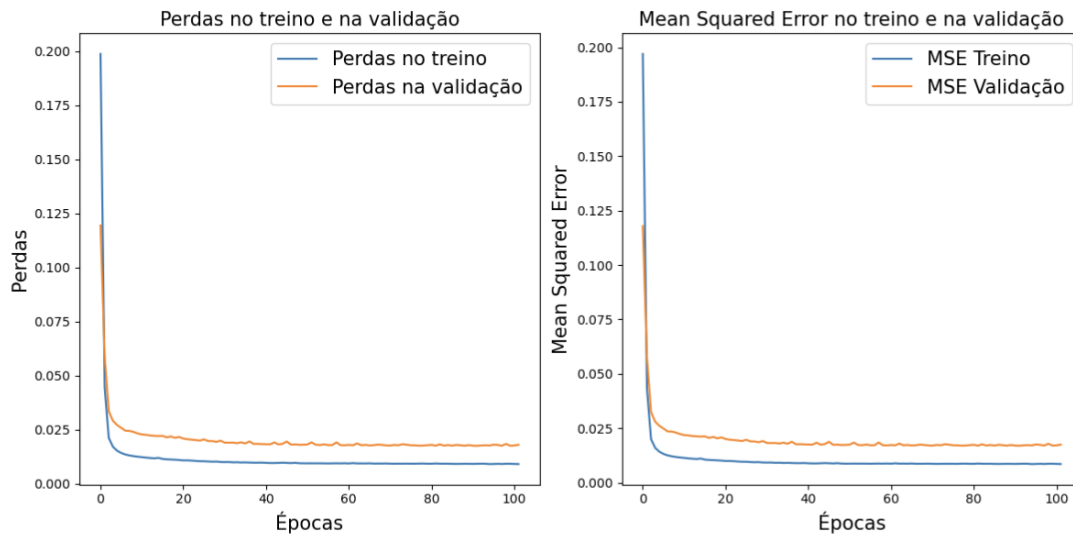


FIGURA 4.26: Curvas de aprendizagem da rede neuronal (8 X 17 X 1) com 250 épocas de treino.

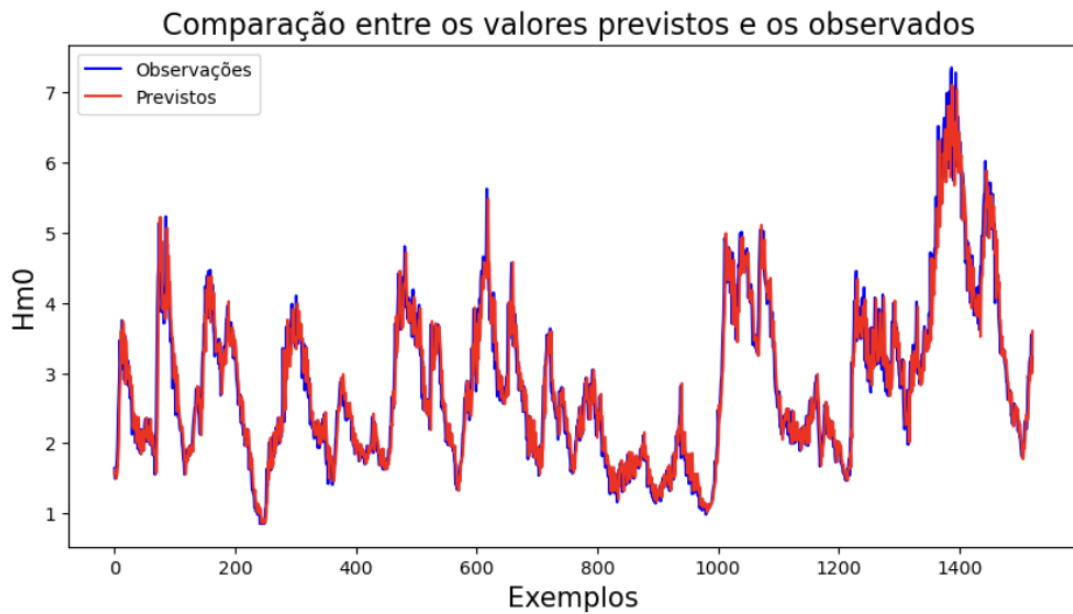


FIGURA 4.27: Valores observados e valores previstos pela rede neuronal (8 X 17 X 1) com 250 épocas de treino.

A figura 4.27 mostra uma comparação dos valores observados e dos valores previstos pela rede neuronal (8 X 17 X 1) no conjunto de teste, que corresponde a 1522 registos no período de 28-12-2023 14:00:00 até 29-02-2024 23:00:00. A análise desta figura permite concluir que o modelo prevê com grande precisão os valores de H_{m0} , mesmo os valores extremos da variável.

Capítulo 5

Discussão dos resultados

Neste capítulo, é discutida a eficácia dos métodos de AA na previsão da altura significativa na área compreendida entre o Cabo Espichel e o Cabo de Sines, procurando responder à questão central: “Os métodos de Aprendizagem Automática permitem prever de forma fiável as condições de agitação marítima num determinado ponto?”, através da análise dos resultados obtidos e das respostas às questões derivadas que orientaram a investigação.

Como resposta à questão central, os resultados obtidos demonstram que os métodos de AA têm potencial para prever a altura H_{m0} , que é um parâmetro importante da agitação marítima de uma forma fiável. Contudo, ao contrário do expectável, a performance das redes neuronais MLP aplicadas no presente trabalho para a previsão a 12 horas da altura significativa espectral não é significativamente superior aos modelos estatísticos aplicados: modelo ARIMA(2,1,5) e o modelo de regressão linear. Estes resultados podem ser explicados devido à implementação exclusiva de algoritmos de redes neuronais MLP. Estas redes, ao contrário das redes neuronais recorrentes, não têm processos de retroalimentação, e por isso não são tão eficazes a detetar padrões temporais das séries, processando-as como vetores independentes. Para além disso, os registos disponíveis para treino, validação e teste das redes neuronais desenvolvidas reduziram-se a um comprimento de 1677 registos, correspondentes a um período 1 ano 59 dias e 6 horas, com intervalos de 6 horas entre registos. Estas redes consideram apenas os registos com datas coincidentes aos registos do modelo *WaveWatch III*, que calcula 4 registos diários nos períodos de 0, 6, 12, 18 horas.

Os resultados obtidos na aplicação das duas redes neuronais semelhantes em termos de estrutura, hiperparâmetros e previsores, mas com a dimensão da amostra diferente, comprova que a segunda rede desenvolvida, que processou 10152 registos, tem uma melhor performance na utilização de segmentos de séries temporais com

maior comprimento. Por essa razão, a utilização de um maior conjunto de dados poderá otimizar a precisão dos modelos de redes neuronais anteriormente explorados.

Este capítulo apresenta-se subdividido em seis secções. Nas primeiras quatro secções serão discutidos os resultados da presente dissertação e procurar-se-á dar resposta às quatro questões derivantes da questão central. As últimas duas secções estão reservadas à discussão da contribuição do presente trabalho para o problema de previsão de agitação marítima, e para as limitações decorrentes do presente trabalho.

5.1 Seleção dos dados/séries temporais

Nesta secção, procurar-se-á responder à primeira questão derivada: "Como seleccionar entre os dados disponíveis (boia multiparamétrica e modelo numérico), as séries temporais que permitem e garantem o treino dos modelos e previsões fiáveis para a altura significativa?".

Para o desenvolvimento dos modelos estatísticos e de AA deste trabalho utilizou-se registos da boia multiparamétrica de Sines e de séries temporais extraídas da malha geográfica do modelo numérico *WaveWatch III*. O critério de seleção das séries temporais baseou-se nos resultados do modelo ARIMA, determinando a extensão temporal dos dados históricos que serão relevantes para previsões precisas, e ainda no modelo de regressão linear que permitiu identificar os previsores que mais contribuem para a previsão a 12 horas da altura significativa espectral.

Os resultados obtidos da aplicação do modelo ARIMA, permitiram concluir que os registos anteriores da própria série (H_{m0}) contribuem para a previsão da altura significativa espectral H_{m0} a 12 horas até desfasamentos de 72 horas. Além disso, poderá ainda concluir-se que a utilização exclusiva da série temporal da altura significativa espectral H_{m0} produz previsões aceitáveis e bastante semelhantes às obtidas com um modelo de redes neuronais treinado com poucos dados.

A aplicação do modelo de regressão linear permitiu identificar os previsores teoricamente mais significativos para a previsão de H_{m0} , contudo seria necessário recorrer a uma maior amostra e testar várias redes neuronais, para poder aferir conclusões relativas aos critérios de seleção de previsores adotados.

5.2 Seleção de variáveis com base no conhecimento de processos físicos

A segunda questão derivada do presente trabalho é a seguinte: "Como selecionar as variáveis que permitem treinar, testar e validar modelos de AA com base no conhecimento geral sobre os processos de geração, propagação e transformação das ondas de vento?".

No presente trabalho foram analisados os resultados climatológicos nas áreas marítimas de interesse nacional identificadas por Sousa [2023](#) a partir dos resultados numéricos do modelo *WaveWatch III*. Na análise deste trabalho identificou-se áreas de agitação de geração marítima a partir das quais se propaga ondulação que atinge a costa de Portugal Continental. Foram também analisados os períodos de pico registados com maior frequência na boia de Sines, tal como as direções de propagação para os quais esses períodos correspondiam. A partir do conhecimento dos processos de geração, propagação e transformação das ondas de vento, foram deduzidas duas localizações a partir das quais a ondulação que se propaga em Sines tivesse uma correlação forte a 24 e 48 horas. O resultado desta estimativa permitiu identificar a área marítima ao largo da ilha da Madeira e do arquipélago dos Açores como os pontos a serem extraídos da malha geográfica do modelo numérico WWIII.

Os resultados obtidos pelo modelo de regressão linear permitiram identificar que as séries temporais relativas ao período de pico ao largo dos Açores no período de 48 horas, bem como a altura significativa ao largo da ilha da Madeira 72 antes de se propagar até à boia de Sines, são previsores relevantes para previsões de 12 horas da altura significativa espectral em Sines.

A implementação destes previsores nos modelos de AA permitiram que as redes neuronais desenvolvidas obtivessem previsões de H_{m0} a 12 horas com RMSE inferior a 0,80 m e MAE inferior a 0,60 m. Contudo, não foi possível obter resultados conclusivos sobre a influência da inclusão destes previsores como variáveis de entrada (*features*) na performance dos modelos de AA. Para obter conclusões sobre a influência destas *features* na previsão de H_{m0} recomenda-se fazer uma nova investigação com a metodologia proposta pelo presente trabalho recorrendo a mais registos do modelo numérico *WaveWatch III*.

5.3 Estrutura dos modelos

A presente secção pretende dar resposta à terceira questão derivada: "Que tipos, estrutura de modelos garantem maior precisão na previsão da agitação marítima na área em estudo?".

No presente trabalho foram testadas várias estruturas de redes neuronais na quais se fez variar o número de camadas intermédias bem como o número de neurónios presentes nessas mesmas camadas. Apesar de ser possível obter um bom ajuste e métricas de desempenho aceitáveis ($RMSE = 0,74$ e $MAE = 0,54$) com redes neuronais de com apenas 1 camada intermédia e com redes neuronais constituídas por três camadas intermédias, adotou-se inicialmente a estrutura de rede mais simples ($4 \times 5 \times 1$). No desenvolvimento do presente trabalho foi variado o número de variáveis de entrada, pelo que nesses casos, se desenvolveram redes neuronais com o mesmo número de neurónios na camada de entrada que o número de previsores considerados.

Apesar de se ter obtido métricas de performance semelhantes na aplicação do modelo ARIMA, do modelo de regressão linear e ainda do modelo de redes neuronais, a última experiência efetuada, na qual foram comparadas redes neuronais semelhantes, com diferentes volumes de dados processados, sugere que para responder a esta questão derivada seria necessário efetuar mais experiências garantindo um maior volume de dados do modelo numérico. Para além disso, deve-se ainda considerar a aplicação das redes recorrentes, pela sua arquitetura tornar o processo de aprendizagem de séries temporais mais eficaz.

5.4 Hiperparâmetros

Nesta secção, procurar-se-á dar resposta à quarta questão derivada: "Que modelos e estratégia de exploração de hiperparâmetros permitem a melhoria da qualidade dos modelos de previsão?"

No presente trabalho foram aplicadas duas estratégias de exploração de hiperparâmetros: exploração manual, que consistiu na variação dos hiperparâmetros efetuando diversas iterações e exploração automática que consistiu na aplicação de um algoritmo (*Random Search*) que executou 36 combinações aleatórias dos valores de alguns hiperparâmetros, nomeadamente os do *batch size*, taxa de aprendizagem e de regularização, devolvendo os valores deste hiperparâmetros para os quais o modelo apresentava um menor MAE.

Os resultados obtidos permitem concluir que o algoritmo *Random Search* não apresentou valores de hiperparâmetros que tornassem a aprendizagem mais eficiente. Contudo, o comprimento relativamente reduzido dos segmentos das séries temporais utilizadas poderá ter constituído uma limitação para o resultado do algoritmo.

5.5 Contribuições para o problema geral de previsão da agitação marítima

Esta secção descreve as contribuições deste trabalho para o problema geral de previsão da agitação marítima, com foco na aplicação de modelos de Aprendizagem Automática (AA) para prever a altura significativa das ondas H_{m0} .

Este trabalho, em comparação com estudos científicos anteriormente publicados, adota uma metodologia distinta no desenvolvimento de redes neuronais capazes de prever H_{m0} , na qual integra a aplicação do modelo ARIMA por forma a identificar a extensão temporal dos dados históricos da série de H_{m0} mais relevantes para a previsões desta variável a 12 horas. Esta metodologia considerou ainda a aplicação de modelos de regressão linear por forma a fazer uma seleção das variáveis de entrada que poderiam contribuir para refinar a precisão dos modelos de AA na previsão de H_{m0} .

No presente trabalho foi realizada uma análise das direções de propagação dos períodos de pico T_p mais frequentes em Sines, de modo estimar-se, com base no conhecimento dos processos de geração, propagação e transformações de ondas, a localização deste grupos de ondas a 24 e 48 horas, para que posteriormente fossem extraídas da malha geográfica do modelo WWIII séries temporais de H_s e T_p nestes pontos. A inclusão destas séries temporais como previsores de H_s teve como objetivo maximizar o potencial preditivo das redes temporais desenvolvidas. Os resultados obtidos da aplicação do modelo de regressão linear permitiram concluir que os previsores que mais contribuem para a previsão de H_{m0} a 12 horas são as séries temporais de H_{m0} , $H_{m0} - 24$, o período de pico na região dos açores 48 antes da previsão ($T_{pa} - 48$) e ainda a altura significativa da ondulação na região da madeira 72 horas antes da previsão ($H_{sm} - 72$).

No estudo desenvolvido por Oleg Makarynskyy [2005] é proposto a utilização de registos da boia multiparamétrica de Sines desfasados de 3 horas ($H_s - 3$, $H_s - 6$, $H_s - 9$, $H_s - 12$, $H_s - 15$, $H_s - 18$, $H_s - 21$, $H_s - 24$) como variáveis de entrada nos

modelos de AA para previsões de H_s a 3, 6, 12 e 24 horas. Estes modelos recorreram a registos extraídos da boia ondógrafo de Sines registados com intervalos temporais de 3 horas no período de 04-12-1999 00:00:00 a 31-01-2000 21:00:00, constituindo assim uma amostra de 37 dias e 303 registos. Neste estudo obteve-se como métricas de avaliação de desempenho da rede neuronal aplicada a previsões de H_s a 12 horas valores de RMSE entre 0,31 e 0,37 m.

No presente trabalho foram utilizadas três séries temporais extraídas da boia multiparamétrica de Sines compreendidas entre 08-11-2022 07:00:00 e 04-03-2023 22:00:00, 27-04-2023 10:00:00 e 10-08-2027 07:00:00, 12-08-2023 14:00:00 e 11-02-2024, constituindo assim uma amostra de 1677 registos, ou seja, 5,5 vezes maior que a amostra considerada por Oleg Makarynskyy [2005]. Para comparar os resultados obtidos pelo presente trabalho e por Oleg Makarynskyy [2005] é necessário considerar que a agitação marítima em zonas costeiras é um fenómeno que poderá ser influenciada por padrões sazonais (Pope e Lockhart [2002]). Oleg Makarynskyy [2005] considerou um período de 2 meses para treinar, validar e testar redes neurais MLP, o poderá corresponder a um período relativamente curto para prever o comportamento de H_{m0} para outras alturas do ano, especialmente nos meses onde a altura significativa da onda é em média mais baixa, como por exemplo nos meses entre maio e setembro (figura 3.10 do capítulo 3).

O presente trabalho desenvolveu redes neuronais utilizando registos de H_{m0} com uma extensão temporal de aproximadamente 1 ano e dois meses, com uma intervalos de 6 horas entre registos. Estas redes previam H_{m0} para 12 horas obtendo como métrica de desempenho no conjunto de teste RMSE = 0,77 m. Na aplicação de uma rede neuronal (8 X 17 X 1) na qual se incluiu como variáveis de entrada as séries temporais de H_{m0} e T_p e respetivas *lagged variables* de 12, 24, 48 horas extraídas exclusivamente da boia multiparamétrica de Sines (10152 registos processados) obteve-se como métricas de performance do modelo no conjunto de teste RMSE = 0,20 m. Estes resultados apresentam uma melhoria de 0,11 m, aproximadamente 35% das previsões obtidas por Oleg Makarynskyy [2005]. Pelo que se pode concluir que para obter resultados mais precisos na previsão de H_{m0} a 12 horas, com redes neuronais MLP treinadas com registos de um aproximadamente 1 ano, é necessário que o intervalo de amostragem entre seja inferior a 6 horas.

5.6 Limitações do estudo

Esta secção descreve as principais limitações do presente estudo, que devem ser consideradas na interpretação dos resultados.

Uma das limitações mais relevantes dos modelos de AA desenvolvidos neste trabalho está relacionada com o comprimento relativamente curto dos segmentos das séries temporais utilizadas. A seleção de períodos de 6 horas entre registos afetou o processo de aprendizagem das redes neuronais, limitando potencialmente a capacidade preditiva dos modelos. Uma possibilidade para trabalhos futuros seria a extrapolação dos valores de H_s e T_p obtidos através do modelo numérico *WaveWatch III*, de modo a aumentar o intervalo de amostragem entre registos para uma base horária, permitindo que o modelo tivesse mais registos para treinar, validar e testar as suas previsões.

Outra limitação foi o presente estudo não considerar certos parâmetros de agitação marítima que podem ter um impacto significativo no processo de geração, propagação e transformação das ondas de vento, tais como a velocidade e direção do vento. Estudos anteriores, como o de Deo, Jha, Chaphekar e Ravikant [2001], sugerem a inclusão destes parâmetros por forma a melhorar a precisão dos modelos. A inclusão de previsores (*features*) como a direção e intensidade do vento, bem como a direção de propagação das ondas, poderá ser objeto de estudo para trabalhos futuros na área.

Apesar da relevância das redes neuronais recorrentes, em particular, as células LSTM (*Long Short-Term Memory*) e GRU (*Gated Recurrent Units*), para a previsão de séries temporais, destacadas nos estudos de Tayeb Sadeghifar e Mahdizadeh [2017] e Minuzzi e Farina [2023], estas arquiteturas não foram implementadas no presente estudo. A escolha por redes mais simples foi motivada por restrições de tempo e recursos, porém reconhece-se que as redes neuronais LSTM e GRU têm uma performance melhor que as redes MLP na previsão de séries temporais. Pelo que a implementação de redes neuronais recorrentes LSTM e GRU constituem uma linha de investigação promissora para estudos futuros, na qual poderá ser explorada a capacidade destas redes neuronais para processar as séries temporais extraídas da boia multiparamétrica de Sines e do modelo numérico WWIII para prever H_{m0} , tal como sugerido pela metodologia do presente trabalho.

Por fim, a aplicabilidade dos modelos desenvolvidos neste estudo está geograficamente limitada, uma vez que estes foram especificamente desenvolvidos com

base nas condições de propagação da agitação marítima na área entre o Cabo Espichel e o Cabo de Sines. Esta restrição geográfica sugere que os resultados obtidos podem não ser diretamente generalizáveis para outras regiões costeiras sem uma adaptação prévia dos modelos às condições locais.

Conclusão

O objetivo principal deste trabalho consiste na avaliação da precisão dos resultados dos modelos de Aprendizagem Automática (AA) na previsão dos parâmetros de agitação marítima num determinado ponto (boia multiparamétrica de Sines) na área entre o Cabo Espichel e o Cabo de Sines, que corresponde a uma zona de importância estratégica para a segurança marítima e atividades costeiras em Portugal.

O conceito do presente trabalho envolveu a avaliação dos resultados das previsões dos modelos de AA para H_{m0} a 12 horas em Sines, utilizando como previsores os registos de H_{m0} e T_p da boia multiparamétrica de Sines bem como as séries temporais de H_s e T_p extraídas de dois pontos da malha geográfica do modelo numérico *WaveWatch III*. Os pontos extraídos do modelo numérico foram selecionados a partir da análise de elementos da teoria linear das ondas de superfície e dos resultados climatológicos nas áreas de interesse portuguesas desenvolvidas por Sousa [2023](#).

A partir da aplicação de diferentes modelos, incluindo redes neuronais MLP, modelos estatísticos ARIMA e de regressão linear, foram obtidos resultados que confirmam o potencial dos métodos de AA para realizar previsões fiáveis da altura significativa espectral H_{m0} , embora a performance das redes MLP com poucos registos (1677 registos) não tenha superado significativamente os modelos estatísticos mais simples (ARIMA e regressão linear com variáveis desfasadas).

O estudo demonstrou que a seleção dos previsores e da estrutura dos modelos são importantes para a precisão das previsões de H_{m0} . Em particular, a inclusão de variáveis relacionadas com a geração e propagação das ondas de vento, como a altura significativa e o período de pico em regiões específicas, mostrou-se eficaz, porém o desempenho das redes neuronais alimentadas com estes previsores foi limitada pelo comprimento das séries utilizadas no seu processo de treino. Conclui-se que a utilização de uma amostra correspondente a período temporal de 1 ano e 2 meses para treino, validação e teste de redes com intervalos de amostragem de 6 horas tem um pior desempenho (RMSE= 0,73 m e MAE= 0,55 m) que redes que consideram a mesma extensão temporal, com intervalos de amostragem de 1 hora (RMSE=

0,19 m e MAE= 0, 14 m). Este trabalho também destacou as limitações das redes neuronais MLP, sugerindo que a aplicação rede recorrentes baseadas em elementos LSTM ou GRU no lugar de neurónios simples, poderá permitir obter previsões mais fiáveis para H_{m0} .

Além disso, a análise comparativa com estudos anteriores evidenciou que a utilização de intervalos de amostragem menores (inferiores a 6 horas) e o prolongamento do período de amostragem (superior a 1 ano) podem contribuir para previsões mais precisas. Neste contexto, a metodologia desenvolvida mostrou-se promissora, no entanto pode ser ainda refinada, nomeadamente em termos da extensão da base de dados utilizada e da inclusão de novos parâmetros de agitação marítima, como a direção e velocidade do vento e a direção dominante de propagação das ondas.

Os resultados obtidos permitiram concluir que os registos anteriores de H_{m0} até 72 horas, bem como os registos das séries temporais de H_s ao largo da Ilha da Madeira desfasados de 72 horas e de T_p ao largo do arquipélago dos Açores desfasados a 12 horas contribuem para prever H_{m0} a 12 horas. Além disso, foi possível ainda concluir que o algoritmo *Random Search* não contribui para determinar a estrutura nem os hiperparâmetros capazes de otimizar e melhorar a aprendizagem das redes neuronais aplicadas. Contudo, no presente estudo não foi possível concluir qual a estrutura de modelos que garantem maior precisão na previsão de H_{m0} .

As principais limitações deste estudo incluem o comprimento reduzido das séries temporais utilizadas, a exclusão de certos parâmetros relevantes para a caracterização completa da agitação marítima, e a aplicação dos modelos a áreas geográficas específicas. Apesar destas restrições, os resultados obtidos constituem um avanço significativo na aplicação de AA para previsões de H_{m0} e estabelecem uma base sólida para investigações futuras.

Em suma, este estudo avaliou a precisão dos modelos de AA na previsão de H_{m0} a 12 horas em Sines, destacando o impacto da seleção de previsores e da estrutura dos modelos na precisão das previsões. Embora a implementação de redes neuronais MLP tenham potencial para prever H_s , o seu desempenho foi limitado pela utilização de intervalos de amostragem (6 horas). Recomenda-se que futuramente que seja explorada a implementação de redes neuronais recorrentes com elementos LSTM ou GRU, que se utilizem séries temporais com intervalos de registos inferiores a 6 horas e ainda que se integrem mais previsores por forma a melhorar a precisão dos modelos de AA.

Bibliografia

- Beji, S. (2013). «Improved explicit approximation of linear dispersion relationship for gravity waves». Em: *Coastal Engineering* 73, pp. 11–12. ISSN: 0378-3839.
- Bhattacharjee, Joydeep (out. de 2017). «Some Key Machine Learning Definitions». Em: *Technology at Nineleaps*. URL: <https://medium.com/nineleaps/technology-at-nineleaps-3b0ff3d1b4d0>.
- Bontempi, Gianluca (fev. de 2021). *"Statistical foundations of machine learning: the handbook"*.
- Browne, Matthew, Bruno Castelle, Darrell Strauss, Rodger Tomlinson, Michael Blumenstein e Chris Lane (2007). «Near-shore swell estimation from a global wind-wave model: Spectral process, linear, and artificial neural network models». Em: *Coastal Engineering* 54.5, pp. 445–460. ISSN: 0378-3839.
- Burkov, Andriy (2019). *The hundred-page machine learning book*. Vol. 1. Andriy Burkov Quebec City, QC, Canada.
- Chatfield, C. e H. Xing (2019). *The Analysis of Time Series: An Introduction with R*. Chapman and Hall/CRC Texts in Statistical Science Series. CRC Press, Taylor & Francis Group.
- Chollet, François et al. (2015). *Keras*. <https://keras.io>.
- Dangeti, P. (2017). *Statistics for Machine Learning*. Packt Publishing. ISBN: 9781788291224. URL: <https://books.google.pt/books?id=C-dDDwAAQBAJ>.
- Demetriou, Demetris, Constantine Michailides, George Papanastasiou e Toulia Onoufriou (2021). «Coastal zone significant wave height prediction by supervised machine learning classification algorithms». Em: *Ocean Engineering*.
- Deo, M.C., A. Jha, A.S. Chaphekar e K. Ravikant (2001). «Neural networks for wave forecasting». Em: *Ocean Engineering* 28.7, pp. 889–898. ISSN: 0029-8018.
- Dickey, David A e Wayne A Fuller (1979). «Distribution of the estimators for autoregressive time series with a unit root». Em: *Journal of the American statistical association* 74.366a, pp. 427–431.
- Eckart, Carl (1952). «The propagation of gravity waves from deep to shallow water». Em: *National Bureau of Standards*.
- EMODnet (2024). *EMODnet GeoViewer*. <https://emodnet.ec.europa.eu/geoviewer/>. Acedido: 2024-05-24.

- Fleit, Gábor (s.d.). *Numerical Modeling of Complex Free Surface Flows*. https://www.researchgate.net/figure/Breaking-wave-types-FHWA-2008-The-type-of-wave-breaking-is-determined-by-the_fig20_324656508[acedido17Jun.2024]. Scientific Figure on ResearchGate.
- Geron, Aurelien (2019). *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*. 2nd. O'Reilly Media, Inc. ISBN: 1492032646.
- Goodfellow, Ian, Yoshua Bengio e Aaron Courville (2016). *Deep Learning*. MIT Press.
- Hamilton, James D. (1994). *Time Series Analysis*. Princeton University Press.
- Holthuijsen, Leo H. (2007). *Waves in Oceanic and Coastal Waters*. Cambridge University Press.
- Hoyer, Stephan e Joseph Hamman (2017). «xarray: N-D labeled Arrays and Datasets in Python». Em: *Journal of Open Research Software* 5. DOI: [10.5334/jors.148](https://doi.org/10.5334/jors.148).
- Huang, Wenrui e Simon Foo (2002). «Neural network modeling of salinity variation in Apalachicola River». Em: *Water Research* 36.1, pp. 356–362. ISSN: 0043-1354. DOI: [https://doi.org/10.1016/S0043-1354\(01\)00195-6](https://doi.org/10.1016/S0043-1354(01)00195-6). URL: <https://www.sciencedirect.com/science/article/pii/S0043135401001956>.
- Hyndman, R.J. e G. Athanasopoulos (2014). *Forecasting: Principles and Practice*. OTexts.
- Instituto Hidrográfico (2015). «MONIZEE : SISTEMA DE MONITORIZAÇÃO E PREVISÃO OPERACIONAL PARA O OCEANO COSTEIRO PORTUGUÊS». Em.
- James, Samantha e Roland Stull (2024). *Learning Goal 8a: Describe wind velocity, fetch, duration, and drag, and how they relate to ocean-wave creation*. UBC ATSC 113 - Weather for Sailing, Flying & Snow Sports. URL: https://www.eoas.ubc.ca/courses/atsc113/sailing/met_concepts/08-met-waves/8a-wave-formation/index.html (acedido em 13/05/2024).
- James, Scott C, Yushan Zhang e Fearghal O'Donncha (2018). «A machine learning framework to forecast wave conditions». Em: *Coastal Engineering* 137, pp. 1–10.
- Jeffreys, Harold (1925). «On the formation of water waves by wind». Em: *Proceedings of the royal society of London. Series A, containing papers of a mathematical and physical character* 107.742, pp. 189–206.

- Khandekar, Madhav L (2013). *Operational analysis and prediction of ocean wind waves*. Vol. 33. Springer Science & Business Media.
- Kumar, N. Krishna, Ramasamy Savitha e Abdullah Al Mamun (2018). «Ocean wave height prediction using ensemble of Extreme Learning Machine». Em: *Neurocomputing* 277, pp. 12–20.
- Kwiatkowski, Denis, Peter CB Phillips, Peter Schmidt e Yongcheol Shin (1992). «Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root?» Em: *Journal of Econometrics* 54.1-3, pp. 159–178.
- Longuet-Higgins, Michael Selwyn (1957). «The statistical analysis of a random, moving surface». Em: *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences* 249.966, pp. 321–387.
- Ma, Yu et al. (2020). «Machine learning in ocean applications: wave prediction for advanced controls of renewable energy and modeling nonlinear viscous hydrodynamics». Tese de doutoramento. Massachusetts Institute of Technology.
- Mahjoobi, J. e A. Etemad-Shahidi (2008). «An alternative approach for the prediction of significant wave heights based on classification and regression trees». Em: *Applied Ocean Research* 30.3, pp. 172–177. ISSN: 0141-1187. DOI: <https://doi.org/10.1016/j.apor.2008.11.001>. URL: <https://www.sciencedirect.com/science/article/pii/S0141118708000680>.
- Mahjoobi, Javad e Ehsan Adeli (abr. de 2009). «Prediction of significant wave height using regressive support vector machines». Em: *Ocean Engineering* 36, pp. 339–347. DOI: [10.1016/j.oceaneng.2009.01.001](https://doi.org/10.1016/j.oceaneng.2009.01.001).
- Makarynsky, O. (2004). «Improving wave predictions with artificial neural networks». Em: *Ocean Engineering* 31.5, pp. 709–724. ISSN: 0029-8018.
- Makarynsky, Oleg (jan. de 2005). «Neural pattern recognition and prediction for wind wave data assimilation». Em: *Pacific Oceanography* 3, pp. 76–85.
- McCowan, J. (1894). «VII. On the solitary wave». Em: *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 32.194, pp. 45–58.
- Miche, M. (1944). «Mouvements ondulatoires de la mer en profondeur constante ou décroissante». Em.
- Miles, John W (1957). «On the generation of surface waves by shear flows». Em: *Journal of Fluid Mechanics* 3.2, pp. 185–204.

- Minuzzi, Felipe C. e Leandro Farina (2023). «A deep learning approach to predict significant wave height using long short-term memory». Em: *Ocean Modelling* 181.
- Mushtaq, R. (2011). *Augmented Dickey Fuller Test*. SSRN.
- O Donncha, Fearghal, Yushan Zhang, Bei Chen e Scott C James (2018). «An integrated framework that combines machine learning and numerical models to improve wave-condition forecasts». Em: *Journal of Marine Systems* 186, pp. 29–36.
- Oleg Makarynskyy António A. Pires-Silva, Dina Makarynska (2005). «Artificial neural networks in wave predictions at the west coast of Portugal». Em: *Computers and Geosciences*, pp. 415–424.
- Phillips, O. M. (1957). «On the generation of waves by turbulent wind». Em: *Journal of Fluid Mechanics* 2, pp. 417–445.
- Pinardi, Nadia, L. Cavaleri, G. Coppini, P. De Mey, C. Fratianni, J. Huthnance, P. F. J. Lermusiaux, A. Navarra, R. Preller e S. Tibaldi (2017). «From weather to ocean predictions: an historical viewpoint.» Em: *Journal of Marine Research* 75.3, pp. 103–159.
- Pope, Joan e John H. Lockhart (2002). *Coastal Engineering Manual*.
- Prince, Simon J.D. (2023). *Understanding Deep Learning*. MIT Press.
- Rabaud, Marc e Frédéric Moisy (2020). «The Kelvin–Helmholtz instability, a useful model for wind-wave generation?» Em: *Comptes Rendus. Mécanique* 348.6–7, pp. 489–500.
- Rao, Subba e S. Mandal (2005). «Hindcasting of storm waves using neural networks». Em: *Ocean Engineering* 32.5, pp. 667–684. ISSN: 0029-8018. DOI: <https://doi.org/10.1016/j.oceaneng.2004.09.003>. URL: <https://www.sciencedirect.com/science/article/pii/S0029801804002057>.
- Scher, Sebastian e Gabriele Messori (2018). «Predicting weather forecast uncertainty with machine learning». Em: *Quarterly Journal of the Royal Meteorological Society* 144.717, pp. 2830–2841.
- Seabold, Skipper e Josef Perktold (2010). «statsmodels: Econometric and statistical modeling with python». Em: *9th Python in Science Conference*.
- Shumway, Robert H, David S Stoffer e David S Stoffer (2000). *Time series analysis and its applications*. Vol. 3. Springer.
- Silva, I.N. da, D.H. Spatti, R.A. Flauzino, L.H.B. Liboni e S.F. dos Reis Alves (2016). *Artificial Neural Networks: A Practical Course*. Springer International Publishing. ISBN: 9783319431628.

- Singh, Nitin, Saurabh Chaturvedi e Shamim Akhter (2019). «Weather forecasting using machine learning algorithm». Em: *2019 International Conference on Signal Processing and Communication (ICSC)*. IEEE, pp. 171–174.
- Soares, Carlos Ventura, Eugen Rusu, António Alberto Pires Silva e Oleg Makarynskyy (2000). *Técnicas de medição de parâmetros de agitação marítima : intercomparação e validação de modelos*. Edições Salamandra, Lda.
- Sousa, Miguel Serrão Bon (2023). «Construção da Climatologia da Agitação Marítima nas Áreas de Interesse Nacional». Tese de mestrado. Escola Naval.
- Stellwagen, Eric e Len Tashman (2013). «ARIMA: The Models of Box and Jenkins.». Em: *Foresight: The International Journal of Applied Forecasting*.
- Stewart, Robert H (2008). *Introduction to physical oceanography*. Robert H. Stewart.
- Sverdrup, H. U. e Walter H. Munk (1947). *Wind, sea, and swell : Theory of relations for forecasting*.
- Swart, H.E. (2015). *OCEAN WAVES*. Lecture notes, course presented as part of the Master's Programme 'Meteorology, Physical Oceanography and Climate', Utrecht University, The Netherlands.
- Tayeb Sadeghifar Maryam Nouri Motlagh, Massoud Torabi Azad e Mahdi Mohammad Mahdizadeh (2017). «Coastal Wave Height Prediction using Recurrent Neural Networks (RNNs) in the South Caspian Sea». Em: *Marine Geodesy* 40.6, pp. 454–465.
- Tolman, Hendrik L. et al. (2009). *User manual and system documentation of WAVWATCH III TM version 3.14*.
- Tolman, Hendrik L., Vladimir M. Krasnopolsky e Dmitry V. Chalikov (2005). «Neural network approximations for nonlinear interactions in wind wave spectra: direct mapping for wind seas in deep water». Em: *Ocean Modelling* 8.3, pp. 253–278. ISSN: 1463-5003. DOI: <https://doi.org/10.1016/j.ocemod.2003.12.008>. URL: <https://www.sciencedirect.com/science/article/pii/S1463500304000034>.
- United States. Army. Corps of Engineers e Coastal Engineering Research Center (U.S.) (1984). *Shore Protection Manual*. Shore Protection Manual. Department of the Army, Waterways Experiment Station, Corps of Engineers, Coastal Engineering Research Center.
- Weggel, J. R. (1972). «Maximum Breaker Height». Em: *J. Waterw., Harbours and Coastal Eng. Div.* 98, pp. 529–548.
- Wei, William (jan. de 2006). *Time Series Analysis: Univariate and Multivariate Methods*. 2nd. Boston: Addison-Wesley.

- World Meteorological Organization (2023). *Guide for Analysis and Forecasting*. 3rd. Geneva: World Meteorological Organization.
- Zhang, Rong, Zhao-Yue Chen, Li-Jun Xu e Chun-Quan Ou (2019). «Meteorological drought forecasting based on a statistical model with machine learning techniques in Shaanxi province, China». Em: *Science of the Total Environment* 665, pp. 338–346.

Apêndice A - Modelos de Regressão Linear aplicados

Modelo 1 [Prever Y= Hm0_boia_+12H]

Dep. Variable:	Y= Hm0	R-squared (uncentered):	0.955
Model:	OLS	Adj. R-squared (uncentered):	0.954
Method:	Least Squares	F-statistic:	1062.
Prob (F-statistic):	1.01e-199		
Log-Likelihood:	5.9796		
AIC:	0.04077		
BIC:	22.42		
No. Observations:	308		

	coef	std err	t	P> t	[0.025	0.975]
Hm0_boia	0.9532	0.042	22.771	0.000	0.871	1.036
Hm0-24	-0.1810	0.049	-3.716	0.000	-0.277	-0.085
Hm0-48	0.0462	0.044	1.043	0.298	-0.041	0.133
Tp_boia	0.0441	0.063	0.702	0.483	-0.080	0.168
Tp-24	-0.0569	0.063	-0.904	0.367	-0.181	0.067
Tp-48	0.1604	0.066	2.446	0.015	0.031	0.289

Cond. No. 17.1

Modelo 2[Prever Y = Hm0_boia_+12H]

Dep. Variable:	Y= Hm0	R-squared (uncentered):	0.955
Model:	OLS	Adj. R-squared (uncentered):	0.954
Method:	Least Squares	F-statistic:	797.2
Prob (F-statistic):	3.80e-197		
Log-Likelihood:	7.0421		
AIC:	1.916		
BIC:	31.76		

	coef	std err	t	P> t	[0.025	0.975]
Hm0_boia	0.9424	0.043	21.814	0.000	0.857	1.027
Hm0-24	-0.1708	0.049	-3.471	0.001	-0.268	-0.074
Hm0-48	0.0424	0.044	0.954	0.341	-0.045	0.130
Tp_boia	0.0149	0.066	0.225	0.822	-0.115	0.145
Tp-24	-0.0855	0.066	-1.294	0.197	-0.216	0.045
Tp-48	0.1419	0.067	2.124	0.035	0.010	0.273
Hs_a - 72	-0.0277	0.052	-0.532	0.595	-0.130	0.075
Tp_a - 72	0.1313	0.100	1.319	0.188	-0.065	0.327

Cond. No. 25.4

Modelo 3[Prever Y= Hm0_boia_+12H]

Dep. Variable:	Y= Hm0	R-squared (uncentered):	0.956
Model:	OLS	Adj. R-squared (uncentered):	0.955
Method:	Least Squares	F-statistic:	651.0
Prob (F-statistic):	6.37e-19620		
Log-Likelihood:	11.040		
AIC:	-2.080		
BIC:	35.22		
No. Observations:	308		

	coef	std err	t	P> t	[0.025	0.975]
Hm0_boia	0.9350	0.043	21.726	0.000	0.850	1.020
Hm0-24	-0.1575	0.049	-3.215	0.001	-0.254	-0.061
Hm0-48	0.0301	0.044	0.678	0.498	-0.057	0.118
Tp_boia	-0.0255	0.067	-0.379	0.705	-0.158	0.107
Tp-24	-0.1014	0.066	-1.541	0.124	-0.231	0.028
Tp-48	0.0834	0.070	1.190	0.235	-0.054	0.221
Hs_a - 72	-0.0378	0.053	-0.708	0.479	-0.143	0.067
Tp_a - 72	0.1130	0.100	1.128	0.260	-0.084	0.310
Hs_a - 48	-0.0064	0.051	-0.126	0.900	-0.107	0.094
Tp_a - 48	0.1941	0.096	2.027	0.044	0.006	0.382

Cond. No. 28.6

Modelo 4[Prever Y= Hm0_boia_+12H]

```

=====
Dep. Variable:          Y= Hm0      R-squared (uncentered):          0.956
Model:                  OLS         Adj. R-squared (uncentered):      0.955
Method:                 Least Squares   F-statistic:                650.0
Prob (F-statistic):     7.90e-196
Log-Likelihood:         10.817
AIC:                    -1.633
BIC:                    35.67
=====

```

	coef	std err	t	P> t	[0.025	0.975]
Hm0_boia	0.9408	0.042	22.276	0.000	0.858	1.024
HM0-24	-0.1730	0.049	-3.498	0.001	-0.270	-0.076
HM0-48	0.0339	0.045	0.762	0.447	-0.054	0.122
Tp_boia	-0.0032	0.065	-0.050	0.960	-0.131	0.125
Tp-24	-0.0728	0.068	-1.075	0.283	-0.206	0.060
Tp-48	0.0921	0.072	1.282	0.201	-0.049	0.233
Hs_a - 24	0.0443	0.051	0.868	0.386	-0.056	0.145
Tp_a - 24	-0.0390	0.100	-0.389	0.698	-0.236	0.158
Hs_a - 48	-0.0281	0.052	-0.538	0.591	-0.131	0.075
Tp_a - 48	0.2205	0.096	2.309	0.022	0.033	0.408

Cond. No. 28.2

Modelo 5[Prever Y= Hm0_boia_+12H]

```

=====
Dep. Variable:          Y= Hm0      R-squared (uncentered):          0.955
Model:                  OLS         Adj. R-squared (uncentered):      0.954
Method:                 Least Squares   F-statistic:                636.1
Prob (F-statistic):     1.71e-194
Log-Likelihood:         7.6365
AIC:                    4.727
BIC:                    42.03
=====

```

	coef	std err	t	P> t	[0.025	0.975]
Hm0_boia	0.9435	0.043	22.139	0.000	0.860	1.027
HM0-24	-0.1888	0.049	-3.818	0.000	-0.286	-0.091
HM0-48	0.0460	0.044	1.036	0.301	-0.041	0.133
Tp_boia	0.0376	0.064	0.592	0.555	-0.088	0.163
Tp-24	-0.0760	0.070	-1.092	0.276	-0.213	0.061
Tp-48	0.1248	0.071	1.748	0.081	-0.016	0.265
Hs_a - 24	0.0485	0.058	0.830	0.407	-0.066	0.164
Tp_a - 24	-0.0396	0.108	-0.368	0.713	-0.252	0.172
Hs_a - 12	-0.0323	0.058	-0.560	0.576	-0.146	0.081
Tp_a - 12	0.1262	0.108	1.167	0.244	-0.087	0.339

Cond. No. 33.0

Modelo 6[Prever Y= Hm0_boia_+12H]

```

=====
Dep. Variable:          Y= Hm0      R-squared (uncentered):          0.954
Model:                  OLS         Adj. R-squared (uncentered):      0.953
Method:                 Least Squares   F-statistic:                1250.
Prob (F-statistic):     7.59e-200
Log-Likelihood:         2.5768
AIC:                    4.846
BIC:                    23.50
=====

```

	coef	std err	t	P> t	[0.025	0.975]
Hm0_boia	0.9429	0.039	24.218	0.000	0.866	1.020
HM0-24	-0.1507	0.042	-3.621	0.000	-0.233	-0.069
Tp-24	0.0054	0.055	0.099	0.922	-0.102	0.113
Hs_a - 12	-0.0447	0.048	-0.926	0.355	-0.140	0.050
Tp_a - 12	0.2520	0.089	2.843	0.005	0.078	0.426

Cond. No. 17.7

Modelo 7[Prever Y= Hm0_boia_+12H]

Dep. Variable:	Y= Hm0	R-squared (uncentered):	0.956			
Model:	OLS	Adj. R-squared (uncentered):	0.955			
Method:	Least Squares	F-statistic:	813.9			
Prob (F-statistic):	1.97e-198					
Log-Likelihood:	10.085					
AIC:	-4.171					
BIC:	25.67					
=====						
	coef	std err	t	P> t	[0.025	0.975]
Hm0_boia	0.9326	0.042	22.109	0.000	0.850	1.016
HM0-24	-0.1667	0.049	-3.437	0.001	-0.262	-0.071
HM0-48	0.0301	0.044	0.681	0.497	-0.057	0.117
Tp_boia	-0.0016	0.065	-0.025	0.980	-0.129	0.126
Tp-24	-0.0967	0.064	-1.514	0.131	-0.222	0.029
Tp-48	0.1142	0.067	1.696	0.091	-0.018	0.247
Hs_m - 72	0.2038	0.116	1.758	0.080	-0.024	0.432
Tp_m - 72	0.1338	0.082	1.627	0.105	-0.028	0.296
=====						
Cond. No.	26.6					

Modelo 8[Prever Y= Hm0_boia_+12H]

Dep. Variable:	Y= Hm0	R-squared (uncentered):	0.956			
Model:	OLS	Adj. R-squared (uncentered):	0.955			
Method:	Least Squares	F-statistic:	652.3			
Prob (F-statistic):	4.77e-196					
Log-Likelihood:	11.340					
AIC:	-2.680					
BIC:	34.62					
=====						
	coef	std err	t	P> t	[0.025	0.975]

Hm0_boia	0.9232	0.043	21.680	0.000	0.839	1.007
HM0-24	-0.1660	0.048	-3.426	0.001	-0.261	-0.071
HM0-48	0.0277	0.044	0.626	0.532	-0.059	0.115
Tp_boia	-0.0109	0.065	-0.167	0.867	-0.139	0.117
Tp-24	-0.1066	0.066	-1.618	0.107	-0.236	0.023
Tp-48	0.0830	0.071	1.175	0.241	-0.056	0.222
Hs_m - 48	0.1374	0.134	1.023	0.307	-0.127	0.402
Tp_m - 48	0.0721	0.085	0.849	0.396	-0.095	0.239
Hs_m - 72	0.2118	0.118	1.797	0.073	-0.020	0.444
Tp_m - 72	0.0831	0.089	0.937	0.350	-0.091	0.258
=====						
Cond. No.	32.8					

Modelo 9[Prever Y= Hm0_boia_+12H]

```
=====
Dep. Variable:          Y= Hm0      R-squared (uncentered):          0.956
Model:                  OLS         Adj. R-squared (uncentered):      0.954
Method:                  Least Squares      F-statistic:                640.5
Prob (F-statistic):      6.40e-195
Log-Likelihood:          8.6511
AIC:                    2.698
BIC:                    40.00
=====
```

	coef	std err	t	P> t	[0.025	0.975]
Hm0_boia	0.9384	0.043	21.990	0.000	0.854	1.022
Hm0-24	-0.1765	0.049	-3.616	0.000	-0.273	-0.080
Hm0-48	0.0361	0.045	0.804	0.422	-0.052	0.125
TP_boia	0.0185	0.065	0.287	0.775	-0.109	0.146
TP-24	-0.0675	0.068	-0.985	0.325	-0.202	0.067
TP-48	0.1108	0.071	1.569	0.118	-0.028	0.250
Hs_m - 24	0.0398	0.123	0.324	0.746	-0.202	0.282
TP_m - 24	-0.0250	0.088	-0.286	0.775	-0.197	0.147
Hs_m - 48	0.2249	0.127	1.764	0.079	-0.026	0.476
TP_m - 48	0.0540	0.090	0.597	0.551	-0.124	0.232

Cond. No. 30.8

Modelo 10[Prever Y= Hm0_boia_+12H]

```
=====
Dep. Variable:          Y= Hm0      R-squared (uncentered):          0.956
Model:                  OLS         Adj. R-squared (uncentered):      0.954
Method:                  Least Squares      F-statistic:                640.6
Prob (F-statistic):      6.26e-195
Log-Likelihood:          8.6744
AIC:                    2.651
BIC:                    39.95
=====
```

	coef	std err	t	P> t	[0.025	0.975]
Hm0_boia	0.9375	0.043	21.869	0.000	0.853	1.022
Hm0-24	-0.1717	0.049	-3.484	0.001	-0.269	-0.075
Hm0-48	0.0358	0.045	0.799	0.425	-0.052	0.124
TP_boia	0.0189	0.065	0.292	0.770	-0.109	0.147
TP-24	-0.0807	0.069	-1.173	0.242	-0.216	0.055
TP-48	0.1418	0.067	2.115	0.035	0.010	0.274
Hs_m - 12	0.1847	0.126	1.471	0.142	-0.062	0.432
TP_m - 12	0.1115	0.091	1.223	0.222	-0.068	0.291
Hs_m - 24	0.0760	0.118	0.645	0.519	-0.156	0.308
TP_m - 24	-0.1136	0.097	-1.172	0.242	-0.304	0.077

Cond. No. 30.2

Modelo 11[Prever Y= Hm0_boia_+12H]

```
=====
Dep. Variable:          Y= Hm0      R-squared (uncentered):          0.955
Model:                  OLS         Adj. R-squared (uncentered):      0.954
Method:                  Least Squares      F-statistic:                801.8
Prob (F-statistic):      1.67e-197
Log-Likelihood:          7.8846
AIC:                    0.2308
BIC:                    30.07
=====
```

	coef	std err	t	P> t	[0.025	0.975]
Hm0_boia	0.9371	0.043	21.903	0.000	0.853	1.021
Hm0-24	-0.1711	0.049	-3.474	0.001	-0.268	-0.074
Hm0-48	0.0391	0.044	0.879	0.380	-0.048	0.127
TP_boia	0.0169	0.065	0.262	0.793	-0.110	0.144
TP-24	-0.0980	0.067	-1.470	0.143	-0.229	0.033
TP-48	0.1381	0.066	2.081	0.038	0.007	0.269
Hs_m - 12	0.1503	0.121	1.240	0.216	-0.088	0.389
TP_m - 12	0.0846	0.083	1.019	0.309	-0.079	0.248

Cond. No. 27.6