

Automação Avançada

Previsão de Ondas com Dados Oceânicos e Meteorológicos

Autores:

Daniel Ferreira daniel.resina.ferreira@tecnico.ulisboa.pt

Margarida Ferreira margaridanferreira@tecnico.ulisboa.pt

Miguel Mateus miguelcmateus@tecnico.ulisboa.pt

1º Semestre - 2025/2026

1. Introdução

A previsão das propriedades das ondas desempenha um papel fundamental em diversas aplicações, como navegação marítima, gestão costeira e operações offshore. A melhoria da capacidade preditiva depende não só da escolha dos modelos, mas também da qualidade e relevância das variáveis utilizadas. Neste contexto, este projeto tem como objetivo desenvolver modelos de previsão das características das ondas e comparar o desempenho de duas fontes de dados amplamente utilizadas: **ERA5** e **Copernicus**.

A análise proposta visa responder a questões essenciais para a modelação oceânica:

- Existem diferenças significativas na capacidade preditiva entre os conjuntos de variáveis fornecidos por ERA5 e Copernicus?
- Quais os parâmetros das ondas que apresentam maior previsibilidade e quais são mais desafiantes?
- De que forma diferentes horizontes temporais influenciam a precisão das previsões?
- Qual é o impacto das variáveis de vento e correntes na previsão das características das ondas?

Para tal, serão explorados modelos que utilizam separadamente e em combinação as variáveis das duas fontes, permitindo avaliar sinergias e limitações. A comparação da qualidade das previsões será realizada para diferentes horizontes temporais, identificando padrões de desempenho e possíveis fontes de incerteza. Por fim, será analisada a influência das variáveis atmosféricas e oceânicas na previsão das ondas, contribuindo para uma melhor compreensão dos mecanismos que governam a dinâmica marítima. O código desenvolvido neste trabalho encontra-se disponível em: <https://github.com/margaridanferreira/LAB>

1.2 Metodologías Utilizadas

Para identificar as abordagens mais adequadas à previsão das características das ondas, foram explorados diferentes métodos de aprendizagem automática. Em particular, foram considerados métodos baseados em árvores, nomeadamente Random Forest e Tree Boosting, pela sua capacidade de modelar relações não lineares e de fornecer informação sobre a relevância das variáveis. Adicionalmente, foram utilizadas redes neurais, com o objetivo de capturar padrões mais complexos e melhorar a capacidade preditiva dos modelos.

Os modelos foram desenvolvidos para três cenários distintos: (i) utilizando exclusivamente variáveis da base ERA5, (ii) utilizando apenas variáveis da base Copernicus e (iii) combinando variáveis das duas fontes de dados.

Nos modelos baseados na base Copernicus e nos modelos híbridos ERA5 + Copernicus, a existência de resoluções temporais distintas entre as fontes resultou na presença de valores em falta após o alinhamento temporal dos dados. Para garantir a consistência e a integridade dos modelos, optou-se pela remoção dos registos incompletos, assegurando que o treino fosse realizado exclusivamente com amostras totalmente observadas e sincronizadas. A interpolação dos valores em falta não foi adotada, uma vez que implicaria a introdução de um número significativo de valores artificiais que não pertencem originalmente ao dataset, podendo comprometer a representatividade física dos dados e enviesar a avaliação comparativa entre as diferentes configurações de modelos.

Adicionalmente, foram criados modelos sem as variáveis de correntes, com o objetivo de analisar o impacto destas na previsão das características das ondas.

Para avaliar a precisão dos modelos e permitir uma comparação consistente entre as fontes ERA5 e Copernicus, foram utilizadas as métricas MAE e RMSE. O MAE fornece uma medida direta do erro médio nas unidades físicas das variáveis, permitindo uma interpretação intuitiva do desempenho esperado em condições operacionais. O RMSE, por penalizar fortemente erros de maior magnitude, foi incluído para avaliar a sensibilidade dos modelos a desvios extremos, sendo particularmente relevante em contextos de segurança marítima. No caso da Direção Média das Ondas (MWD), dada a sua natureza cíclica, a avaliação foi ajustada para utilizar o erro angular mínimo. Esta abordagem garante que a descontinuidade numérica entre 0° e 360° (onde uma diferença real de 2° poderia ser erroneamente interpretada como 358°) não comprometa o rigor estatístico dos resultados.

2. Neural Networks

As Redes Neurais Artificiais são sistemas de processamento de informação inspirados na estrutura do cérebro humano, caracterizando-se pela sua natureza massivamente paralela e capacidade de aprender através de "pesos sinápticos" (Haykin, 2009). Ao contrário dos métodos lineares, estas redes conseguem modelar a complexidade e a não-linearidade inerentes à dinâmica dos oceanos.

2.1 Abordagem para a análise

O estudo comparou as arquiteturas CNN 1D e LSTM para a previsão de estados de mar em horizontes de 1h a 24h. Dadas as características físicas distintas de cada parâmetro oceânico, foram desenvolvidos e otimizados modelos independentes para cada variável alvo (swl, mwp, pp1d e mwd). Embora a LSTM seja reconhecida pela sua memória sequencial, a CNN 1D superou-a consistentemente, demonstrando maior eficácia na extração de padrões locais e melhor generalização através da sua estrutura de partilha de pesos.

Com base nos resultados de treino e validação (divisão 70/20/10), a CNN 1D foi selecionada como o modelo base para todas as análises subsequentes. Esta decisão fundamenta-se na obtenção de métricas de erro (RMSE e MAE) mais baixas, garantindo a fiabilidade necessária para comparar as fontes de dados ERA5 e Copernicus e avaliar o impacto das variáveis de vento e corrente no sistema.

2.2 Otimização e Configuração da Arquitetura do Modelo

Esta análise foca-se na otimização da arquitetura CNN 1D, o objetivo consistiu em identificar a configuração que maximiza a capacidade preditiva do modelo.

2.2.1 Número de Camadas e Filtros

Foram testadas diversas profundidades de rede e variações na densidade de filtros para equilibrar a complexidade do modelo e a sua capacidade de generalização.

- Configuração Ótima: Embora tenham sido experimentadas arquiteturas mais profundas (com três ou mais camadas) e diferentes volumes de filtros, os melhores resultados foram obtidos com uma estrutura de duas camadas convolucionais.

- Densidade: A combinação de uma primeira camada com 64 filtros e uma segunda com 32 filtros revelou-se a mais eficaz. Esta progressão permite que a rede extraia características fundamentais dos dados brutos na primeira fase e refine os padrões temporais na segunda, evitando o *overfitting* que surgiu em modelos mais densos.

2.2.2 Tamanho do Lookback

Foi comparado um lookback de 24h e 48h. O de 48h melhorou previsões de longo prazo (24h) por captar dois ciclos diários, mas prejudicou o curto prazo (1h) ao introduzir ruído e reduzir a importância da autocorrelação imediata. Assim, optou-se por um lookback de 24h, garantindo melhor precisão operacional e menor custo computacional.

2.2.3 Avaliação de Variáveis Exógenas

Além da avaliação base das previsões, será realizada uma análise de sensibilidade para quantificar o contributo das variáveis exógenas, nomeadamente vento e correntes. Para isso, o desempenho do modelo será comparado entre duas configurações: uma utilizando apenas o histórico das ondas e outra integrando variáveis meteorológicas e oceânicas. Esta comparação permite identificar de forma clara o impacto do vento e das correntes na dinâmica do estado do mar e na capacidade preditiva do modelo.

2.3. Análise dos Resultados

Apesar do bom desempenho das CNN 1D em diversos problemas de séries temporais, os resultados obtidos neste estudo indicam limitações na sua aplicação à previsão de estados de mar. A operação de convolução pressupõe a existência de padrões locais e alguma invariância temporal, o que nem sempre é compatível com a dinâmica oceânica, caracterizada por dependências temporais de médio e longo prazo e por interações físicas complexas entre ondas, vento e correntes. Como consequência, a convolução pode induzir combinações locais das variáveis que não refletem diretamente os mecanismos físicos subjacentes, conduzindo a uma degradação da capacidade preditiva, sobretudo para horizontes temporais mais longos e para variáveis intrinsecamente mais caóticas, como a direção média e o período de pico.

2.3.1 ERA5

Os resultados obtidos com a base ERA5 evidenciam uma boa capacidade preditiva dos modelos, particularmente para a altura significativa das ondas (SWH), que se destaca consistentemente como o parâmetro mais previsível, apresentando os menores valores de RMSE e MAE para todos os horizontes analisados. Em contraste, a direção média das ondas (MWD) e os períodos, sobretudo o período de pico (PP1D), revelam maiores dificuldades de previsão, com erros significativamente superiores, conforme detalhado nas tabelas de métricas de desempenho apresentadas em Anexo (Tabelas 1 e 2).

Observa-se ainda uma degradação sistemática do desempenho com o aumento do horizonte de previsão, tanto em termos de RMSE como de MAE, refletindo a acumulação de incerteza inerente à previsão multi-passo. A análise da influência das variáveis de vento mostra que a remoção de u_{10} e v_{10} conduz, na maioria dos casos, a um aumento dos erros, especialmente para horizontes curtos, confirmando o papel dominante do vento na evolução imediata do estado do mar. Para horizontes mais longos, as diferenças entre os modelos completos e sem vento tornam-se menos pronunciadas, sugerindo que a persistência e a dinâmica interna das ondas passam a ter um peso relativamente maior na previsão.

2.1.2 Copernicus

A análise das métricas RMSE e MAE evidencia que os modelos baseados no Copernicus apresentam, de forma consistente, erros superiores aos obtidos com o ERA5. Este resultado está intrinsecamente associado à menor resolução temporal do dataset Copernicus (intervalos de 3 em 3 horas face à resolução horária do ERA5), o que limita a capacidade do modelo em capturar variações de curto prazo na dinâmica oceânica. Os dados detalhados que sustentam esta comparação encontram-se sistematizados nas Tabelas 3 e 4, apresentadas em Anexo.

Tal como observado anteriormente, a altura significativa das ondas (SWH) confirma-se como o parâmetro mais previsível em ambos os datasets. Em contrapartida, a direção média e os períodos de onda apresentam maior dificuldade de previsão, verificando-se uma degradação progressiva do desempenho à medida que o horizonte temporal aumenta. No entanto, é de notar que a inclusão das variáveis de corrente oceânica (específicas do dataset Copernicus) conduz a melhorias moderadas nas métricas de erro, sobretudo para horizontes mais longos. Este comportamento sugere que, embora a resolução temporal seja um fator penalizante, a informação sobre a circulação oceânica contribui positivamente para a previsão da evolução do estado do mar, ainda que de forma limitada quando comparada com o peso das variáveis meteorológicas.

2.1.3 ERA5 + Copernicus

A análise dos modelos que combinam variáveis ERA5 e Copernicus mostra que, embora a inclusão conjunta de vento e correntes conduza, em alguns casos, a melhorias face às configurações sem variáveis exógenas, os melhores desempenhos globais continuam a ser obtidos pelos modelos baseados exclusivamente no ERA5.

Tal como nas análises anteriores, a altura significativa das ondas (SWH) mantém-se como o parâmetro mais previsível, enquanto a direção média e os períodos apresentam maiores dificuldades de previsão, com degradação progressiva do desempenho à medida que o horizonte aumenta. Estes resultados sugerem que a maior resolução temporal e consistência do ERA5 são determinantes para a qualidade das previsões, não sendo totalmente compensadas pela adição das variáveis do Copernicus no modelo híbrido.

Os resultados detalhados destas métricas, comparando as configurações com e sem influências externas, encontram-se apresentados nos anexos através das seguintes tabelas:

- Tabela 5: Resultados do RMSE para modelos com dados COPERNICUS+ERA (Versão Completa).
- Tabela 6: Resultados do RMSE para modelos com dados COPERNICUS+ERA sem influências de vento e corrente.
- Tabela 7: Resultados do MAE para modelos com dados COPERNICUS+ERA (Versão Completa).
- Tabela 8: Resultados do MAE para modelos com dados COPERNICUS+ERA sem influências de vento e corrente.
- Estas tabelas permitem uma observação direta de como a exclusão das variáveis de vento e corrente impacta a precisão do modelo para cada um dos horizontes temporais analisados.

3. Tree-based method - Boosting

3.1 Motivação para a Escolha do Gradient Boosting

A escolha do Gradient Boosting para a previsão de características de ondas é fundamentada em várias vantagens intrínsecas. Em primeiro lugar, o algoritmo destaca-se pela Captura de Relações Não-Lineares, essencial para modelar as interações complexas entre os preditores meteorológicos e oceanográficos (do ERA5 e Copernicus) e os parâmetros de onda, como altura, período e direção. Em segundo, é particularmente eficaz para lidar com Dependências Temporais quando se integram lag features (características desfasadas), permitindo capturar os padrões autocorrelativos inerentes a séries temporais. A sua Robustez constitui outro fator decisivo, pois é menos sensível a outliers e valores extremos comparativamente a modelos lineares clássicos.

3.2 Arquitetura do algoritmo

Para prever múltiplas características de ondas de forma simultânea e coerente, nomeadamente a altura significativa (swh_cop), o período médio (mwp_cop), a direção média (mwd_cop_cos, mwd_cop_sin) e o período de pico (peak_period_cop), foi adotada a classe MultiOutputRegressor do scikit-learn. Esta estratégia encapsula um GradientBoostingRegressor para cada variável alvo, criando uma arquitetura paralela onde cada regressor tem a mesma configuração e hiperparâmetros, mas aprende de forma independente para o seu target específico. Esta abordagem permite efetuar previsões coordenadas de todos os parâmetros, avaliar o desempenho individualmente para cada um e implementar forecasting recursivo, onde as previsões de um passo são usadas para atualizar os lags de todos os targets nos passos seguintes.

A fim de modelar explicitamente a autocorrelação temporal nas séries de ondas, foram criadas features desfasadas (lag features). Para cada parâmetro, foram geradas variáveis representando os seus valores nas 48 horas anteriores. Este desfasamento temporal, definido por LAGS = 48, permite ao modelo aprender padrões fundamentais como a Persistência (o estado atual é influenciado pelo passado imediato), os Ciclos associados a marés e regimes de vento, e a Inércia Dinâmica das características oceânicas. Desta forma, o modelo não depende exclusivamente de preditores exógenos, mas também incorpora a memória intrínseca da série temporal.

Para garantir uma avaliação realista e evitar o data leakage, os dados foram divididos cronologicamente. A divisão alocou 70% dos dados mais antigos para Treino, 15% seguintes para Validação (para ajuste de hiperparâmetros e early stopping), e os 15% mais recentes para Teste.

O GradientBoostingRegressor foi configurado com hiperparâmetros que equilibram capacidade preditiva e generalização: n_estimators=300 fornece capacidade modelar suficiente; learning_rate=0.05 reduz o contributo de cada árvore, promovendo um aprendizado mais suave e robusto; max_depth=3 limita a complexidade individual das árvores, prevenindo o sobreajuste; validation_fraction=0.15 reserva parte dos dados de treino para validação interna; e n_iter_no_change=20 combinado com tol=1e-4 implementa early stopping automático, interrompendo o treino se não houver melhoria significativa na validação após 20 iterações.

Para gerar previsões para múltiplos passos no futuro (ex., 24 horas), foi implementada uma estratégia de forecasting recursivo. O algoritmo inicia usando os valores reais até ao instante t^* para prever o instante $t+1^*$. Em seguida, essa previsão é incorporada no vetor de features como o lag mais recente para prever $t+2^*$, enquanto os lags anteriores são deslocados. Este processo repete-se sequencialmente para todo o horizonte de previsão. A principal vantagem é simular um cenário de previsão realista, onde apenas informações passadas estão disponíveis. Contudo, o principal desafio é a propagação e acumulação de erros à medida que o horizonte se alonga, o que pode degradar a precisão das previsões para horizontes mais distantes (tipicamente além de 24-48 horas).

3.3 Análise de Resultados

3.3.1 ERA5

Usando apenas variáveis do ERA5, o modelo consegue explicar muito bem a variabilidade da altura significativa de onda: para `swh_cop` obtém-se $MAE \approx 0,086$ e $RMSE \approx 0,111$, com , indicando previsões muito próximas do observado no conjunto de teste. Para os períodos, o desempenho continua bom mas já com maior incerteza: em `mwp_cop` o $RMSE \approx 0,590$ e mostram que o ERA5 capta bem a dinâmica dominante, enquanto em `peak_period_cop` o $RMSE$ sobe para $\approx 1,63$, sugerindo que o pico do espectro é mais difícil de inferir apenas a partir do vento (maior sensibilidade a processos não representados diretamente). Na direção média (`mwd_deg`), a abordagem com seno/cosseno também apresenta resultados consistentes: $MAE \approx 7,1^\circ$ e $RMSE \approx 14,5^\circ$, o que é compatível com o facto de o vento ser um forte condicionante da direção de propagação das ondas. Em suma, o ERA5 por si só já fornece informação suficiente para previsões muito robustas de `swh_cop` e razoáveis de período/direção, sendo o `peak_period_cop` o alvo onde mais se nota a limitação de usar apenas vento.

3.3.2 Copernicus

Usando apenas variáveis do Copernicus como entradas, os resultados degradam-se de forma clara em todos os alvos, o que sugere que as correntes, por si só, não captam os principais mecanismos que controlam as características das ondas neste local/período. Para os períodos (`mwp_cop` e `peak_period_cop`) obtém-se $RMSE$ elevados ($\approx 1,54$ e $\approx 2,47$, respetivamente), indicando que o modelo baseado só em Copernicus chega a ser pior do que uma previsão trivial pela média no conjunto de teste. O mesmo acontece para `swh_cop`, onde apesar do alvo ser bem previsto com ERA5, com Copernicus o erro sobe muito ($RMSE \approx 1,36$), evidenciando fraca capacidade explicativa. Na direção (`mwd_deg`), o erro angular também é bastante maior ($MAE \approx 43^\circ$ e $RMSE \approx 54^\circ$), reforçando que as correntes isoladamente têm poder preditivo limitado para a direção e energia das ondas, quando comparadas com o a influência atmosférica.

3.3.3 Copernicus + ERA5

Quando se combina ERA5 + Copernicus, o desempenho mantém-se muito próximo do obtido com ERA5 sozinho, o que indica que o vento já contém a maior parte da informação relevante para prever as características de onda neste problema, e que as correntes acrescentam apenas um ganho marginal (ou nulo) nas métricas globais. Em `swh_cop`, por exemplo, continua excelente ($MAE \approx 0,090$; $RMSE \approx 0,116$), mas ligeiramente pior do que apenas ERA5 ($RMSE \approx 0,111$), sugerindo que as variáveis Copernicus não ajudam a

reduzir o erro médio e podem introduzir ruído. O mesmo padrão aparece em `mwp_cop` e `peak_period_cop`: os RMSE ($\approx 0,599$ e $\approx 1,694$) e os ($\approx 0,827$ e $\approx 0,467$) são semelhantes, porém um pouco inferiores aos de ERA5, reforçando que a contribuição adicional das correntes é pequena para estes alvos.

Para a direção `mwd_deg`, a combinação praticamente empata com ERA5 (RMSE angular $\approx 14,35^\circ$ vs $14,50^\circ$), mostrando uma melhoria muito ligeira. Em resumo, juntar Copernicus ao ERA5 não muda substancialmente a capacidade preditiva: o sinal dominante vem do ERA5 e o Copernicus atua, no máximo, como refinamento pequeno em direção, sem ganhos consistentes nos restantes alvos. Todos os dados podem ser analisados na tabela 18.

3.3.4 Influência dos Horizontes de Previsão

À medida que aumentamos o horizonte de previsão de 1h até 24h, observa-se o comportamento esperado de degradação progressiva da qualidade: as curvas de erro (MAE e RMSE) tendem a crescer com o horizonte, porque a incerteza aumenta e, no caso de previsão multi-passo/recursiva, os erros acumulam-se (cada passo usa previsões anteriores como entrada). Tipicamente, nas primeiras horas o modelo consegue capturar bem a persistência e a dinâmica local do estado do mar, mantendo erros relativamente baixos; porém, conforme nos aproximamos das 24h, a previsão passa a depender mais da capacidade do modelo em antecipar mudanças de regime (variações de vento/correntes e evolução do “sea state”), o que se traduz em maior dispersão e pior ajuste. Este efeito tende a ser mais visível em variáveis mais voláteis ou menos determinísticas (por exemplo, direção média das ondas, mesmo medida em graus via \sin/\cos), enquanto variáveis com maior inércia (como períodos/altura em situações estáveis) podem degradar de forma mais suave. Em resumo, os gráficos de evolução do erro mostram que existe um custo evidente em ir até 24h: a previsibilidade diminui e o modelo perde precisão, refletindo limites físicos e estatísticos do sistema e do próprio método multi-horizonte.

4. Tree-based method - Random Forest

4.1. Metodologia e Configuração do Modelo

Para superar as limitações dos modelos lineares na captura de relações complexas e não-lineares entre as variáveis atmosféricas e oceanográficas, implementou-se o algoritmo Random Forest Regressor. Este método de ensemble learning opera através da construção de múltiplas árvores de decisão durante o treino, agregando os seus resultados pela média para reduzir a variância e mitigar o risco de overfitting.

A estratégia de modelação adotada consistiu numa abordagem de previsão direta (Direct Multi-step Forecasting). Para cada horizonte de previsão definido (1, 3, 6, 12, 24 horas), foi treinado um modelo independente. A estrutura dos dados foi organizada da seguinte forma:

- **Variáveis de Entrada (X_t):** Conjunto de features meteorológicas e oceanográficas no instante atual. Foram testados três cenários de input: apenas ERA5, apenas Copernicus e o modelo Combinado.
- **Variável Alvo (Y_{t+h}):** O valor da propriedade da onda (ex: Altura Significativa) deslocado h horas no futuro.

- **Divisão Temporal:** Manteve-se a divisão cronológica de 80% para treino e 20% para teste, garantindo que o modelo não acede a informação futura (data leakage), simulando um cenário operacional real.

Para a variável Direção Média das Ondas (MWD), devido à sua natureza circular ($0^\circ - 360^\circ$), implementou-se uma decomposição vetorial em componentes seno e cosseno ($\sin(\theta), \cos(\theta)$), transformando o problema num espaço contínuo. A avaliação de desempenho para esta variável recorreu a métricas angulares específicas, nomeadamente o erro angular calculado sobre a menor distância no círculo trigonométrico.

4.2. Análise de Resultados: Comparação de Fontes

Os resultados detalhados para os diferentes modelos encontram-se sistematizados no anexo (Tabelas 9 a 17), apresentando os valores das métricas de avaliação para cada horizonte temporal e variável alvo.

A análise comparativa entre as reanálises atmosféricas (ERA5) e o modelo operacional oceânico (Copernicus), bem como a sua combinação, permitiu identificar padrões distintos de previsibilidade:

- **Altura Significativa das Ondas (SWH):** Esta revelou-se a variável com maior previsibilidade. Ambas as fontes individuais apresentaram excelente desempenho a curto prazo ($R^2 > 0.95$ para $h=1h$). Contudo, o Modelo Combinado demonstrou maior robustez nos horizontes longos (24h), confirmando que a fusão da informação meteorológica (vento do ERA5) com a inércia hidrodinâmica (ondas do Copernicus) minimiza o erro quadrático médio (RMSE).
- **Período de Pico (Peak Period):** Em contraste, o Período de Pico confirmou-se como a variável mais desafiante. Os valores de RMSE mantiveram-se elevados e o R^2 degrada rapidamente com o aumento do horizonte. Isto deve-se à natureza discreta e volátil deste parâmetro, que muda abruptamente com a chegada de grandes ondulações, eventos que os modelos baseados apenas em condições locais têm dificuldade em antecipar sem dados de fronteira global.
- **Direção Média das Ondas (MWD):** A previsão da direção beneficiou da abordagem de decomposição vetorial. Observou-se que o modelo Copernicus (e o Combinado) tende a apresentar erros angulares ligeiramente inferiores em comparação com o ERA5 isolado, sugerindo que a inclusão da dinâmica das correntes oceânicas contribui para uma melhor representação da refração e da direção de propagação das ondas.

4.3. Influência dos Horizontes de Previsão

A análise temporal evidenciou a esperada degradação da qualidade da previsão à medida que nos afastamos do instante inicial (t_0):

- **Curto Prazo (1h - 6h):** Os modelos exibem um comportamento quase persistente. A inércia do sistema físico é dominante, permitindo previsões de altíssima precisão com erros residuais mínimos.
- **Médio/Longo Prazo (12h - 24h):** A partir das 12 horas, a incerteza meteorológica começa a dominar. Para o horizonte de 24 horas, observa-se uma queda notória no coeficiente de determinação (R^2) e um aumento do RMSE. Este limiar indica o ponto onde o conhecimento do "estado atual" deixa de ser suficiente, sugerindo que previsões para além de um dia necessitam da incorporação de previsões meteorológicas numéricas (forecasts) e não apenas de observações passadas.

4.4. Análise de Sensibilidade


Para responder ao objetivo de isolar o impacto físico dos agentes externos, realizou-se um estudo de Feature Ablation (remoção de variáveis), cujos resultados se destacam:

- **Impacto do Vento (ERA5):** Ao comparar o modelo ERA5 completo com uma versão "Sem Vento" (removendo u_{10}, v_{10}), verificou-se um agravamento significativo do erro no horizonte de 24h. A visualização dos resultados (Heatmap de RMSE) mostrou que a ausência do vento penaliza severamente a previsão da altura das ondas a longo prazo. Isto valida a hipótese de que o vento local é essencial para modelar a geração de windsea e a manutenção da energia da onda ao longo do dia.
- **Impacto das Correntes (Copernicus):** A comparação entre modelos com e sem correntes (u_0, v_0) revelou um impacto assimétrico. Enquanto a previsão da Altura Significativa (SWH) foi pouco afetada pela remoção das correntes (sugerindo fraca interação onda-corrente na magnitude da onda nesta região), a previsão da Direção Média (MWD) beneficiou da inclusão destas variáveis. A redução do erro angular no modelo "Com Corrente" indicia a captura de fenómenos de refração, onde a direção de propagação da onda é modulada pelo campo de correntes, ainda que de forma subtil.
- **Impacto Global dos Agentes Externos (Modelo Completo vs. Só Ondas):** Por fim, realizou-se uma avaliação global para quantificar o ganho total de incluir agentes externos (Vento e Corrente) em comparação com um modelo baseado puramente na inércia das ondas. Comparou-se o **Modelo Completo** (Ondas + Vento + Corrente) contra o cenário "**Só Ondas**". Os resultados demonstram inequivocamente que a inclusão da física atmosférica e oceanográfica é determinante para a precisão a longo prazo. Enquanto o modelo "Só Ondas" degrada rapidamente a sua performance após as 6 horas, o Modelo Completo mantém erros significativamente mais baixos no horizonte de 24 horas, tanto para a altura (SWH) como para a direção (MWD), provando que a memória do estado do mar não é suficiente para prever a evolução das condições num ciclo diário completo.

5. Conclusão

Neste trabalho foi desenvolvido e avaliado um conjunto de modelos de aprendizagem automática para a previsão de características das ondas, recorrendo a dados meteorológicos do ERA5 e a dados oceânicos do Copernicus, analisados isoladamente e em combinação. O estudo permitiu avaliar não apenas o desempenho preditivo dos diferentes modelos, mas também o impacto das fontes de dados, das variáveis exógenas e do horizonte temporal na qualidade das previsões.

Os resultados obtidos demonstram de forma consistente que a altura significativa das ondas (SWH) é o parâmetro com maior previsibilidade em todos os cenários analisados. Em particular, os modelos baseados exclusivamente no ERA5 apresentaram erros muito reduzidos, tanto em termos de MAE como de RMSE, evidenciando que o forçamento atmosférico, nomeadamente o vento, contém a informação dominante para a previsão da energia das ondas no local em estudo. Em contraste, parâmetros como o período de pico e a



direção média das ondas revelaram-se mais difíceis de prever, apresentando erros significativamente superiores e uma maior sensibilidade ao aumento do horizonte temporal.

A comparação entre as diferentes fontes de dados mostrou que os modelos baseados apenas no Copernicus apresentam, de forma geral, um desempenho inferior aos baseados no ERA5. Esta limitação está fortemente associada à menor resolução temporal do Copernicus, que dificulta a captura de variações rápidas na dinâmica do estado do mar. Ainda assim, observou-se que a inclusão das correntes oceânicas pode contribuir positivamente para a previsão da direção média das ondas, sugerindo que estes dados ajudam a representar fenómenos de refração e modulação da propagação das ondas, embora com impacto limitado nas restantes variáveis.

A combinação de variáveis ERA5 e Copernicus não conduziu a melhorias substanciais face ao uso exclusivo do ERA5. Os resultados indicam que a informação adicional introduzida pelas correntes não é suficiente para compensar a menor consistência temporal do Copernicus, sendo o vento o principal agente controlador da evolução do estado do mar neste problema. Assim, o ERA5 revelou-se a fonte mais robusta e informativa para previsões de curto e médio prazo.

No que diz respeito aos métodos de modelação, as redes neuronais convolucionais 1D (CNN 1D) demonstraram limitações na previsão de estados de mar, sobretudo para horizontes mais longos e para variáveis mais caóticas. Em contraste, os métodos baseados em árvores, nomeadamente o Gradient Boosting e o Random Forest, apresentaram um desempenho superior e mais consistente, beneficiando da capacidade de modelar relações não lineares e de integrar eficazmente a informação temporal através de variáveis desfasadas. A abordagem de previsão recursiva permitiu simular um cenário operacional realista, embora com a penalização expectável da acumulação de erro à medida que o horizonte de previsão aumenta.

A análise da evolução dos erros com o horizonte confirmou uma degradação progressiva da previsibilidade entre 1h e 24h, refletindo limites físicos do sistema oceânico e incertezas inerentes à previsão multi-passo. Este efeito foi particularmente pronunciado para a direção média das ondas, mesmo quando avaliada através de métricas angulares apropriadas, enquanto variáveis com maior inércia apresentaram uma degradação mais suave.

Em síntese, este trabalho demonstra que modelos de aprendizagem automática, quando corretamente configurados e avaliados, podem fornecer previsões fiáveis das características das ondas a curto prazo, sobretudo quando alimentados por dados atmosféricos de elevada resolução como o ERA5. No entanto, os resultados também evidenciam limitações claras para horizontes mais longos e para variáveis mais complexas, sugerindo que futuras abordagens poderão beneficiar da integração de previsões meteorológicas numéricas, modelos híbridos físico-estatísticos ou arquiteturas que capturem dependências temporais de longo alcance. Estes desenvolvimentos constituem caminhos promissores para melhorar a robustez e aplicabilidade operacional de sistemas de previsão de estados de mar baseados em dados.



Anexos

--- COMPARAÇÃO DE RESULTADOS (RMSE) ---

target	mwd_era5		mwp_era5		pp1d_era5		swh_era5	
Versão	Completo	Sem Vento	Completo	Sem Vento	Completo	Sem Vento	Completo	Sem Vento
horizon								
1	11.2143	10.5047	0.5475	0.4440	0.9303	1.0707	0.3262	0.3416
3	16.3484	15.9435	0.5892	0.6439	1.2624	1.0466	0.4273	0.4322
6	23.1795	22.8936	0.8104	0.8131	1.4893	1.4019	0.5863	0.5727
12	32.4102	33.1582	1.2734	1.2615	1.7216	1.7717	0.7239	0.7560
24	43.7820	49.0745	1.3893	1.3954	2.2931	2.0990	1.0105	1.0268

Tabela 1 - Resultados do RMSE para modelos com dados ERA5

--- COMPARAÇÃO DE RESULTADOS (MAE) ---

target	mwd_era5		mwp_era5		pp1d_era5		swh_era5	
Versão	Completo	Sem Vento	Completo	Sem Vento	Completo	Sem Vento	Completo	Sem Vento
horizon								
1	7.3806	6.0418	0.4213	0.3332	0.6435	0.8233	0.2076	0.1930
3	10.3612	8.9179	0.4558	0.5068	0.9102	0.7101	0.2688	0.2652
6	14.4495	13.1883	0.6226	0.6086	1.1901	1.0691	0.3859	0.3706
12	21.0080	19.9000	1.0026	1.0238	1.3279	1.4114	0.5002	0.5369
24	30.0830	33.9387	1.1127	1.1383	1.8338	1.7234	0.7598	0.8209

Tabela 2 - Resultados do MAE para modelos com dados ERA5

target	mwd_cop		mwp_cop		peak_period_cop		swh_cop	
Versão	Completo	Sem corrente	Completo	Sem corrente	Completo	Sem corrente	Completo	Sem corrente
horizon								
1	34.5999	26.8674	2.0688	1.8464	3.3091	2.5593	0.7692	0.8714
2	40.0492	36.0642	2.1756	1.9725	3.2717	2.7821	1.0128	0.7962
4	46.4700	51.9725	2.0513	2.1205	3.3604	2.7824	0.8766	0.9290
8	57.1256	56.9347	2.1394	2.1450	3.4595	3.1220	1.1038	1.2641

Tabela 3 - Resultados do RMSE para modelos com dados COPERNICUS

--- COMPARAÇÃO DE RESULTADOS (MAE) ---

target	mwd_cop		mwp_cop		peak_period_cop		swl_cop	
Versão	Completo	Sem corrente	Completo	Sem corrente	Completo	Sem corrente	Completo	Sem corrente
horizon								
1	22.0554	19.0303	1.6010	1.4193	2.5577	2.0525	0.5816	0.6415
2	27.8078	23.6038	1.7278	1.5448	2.5749	2.1797	0.7905	0.5778
4	33.5179	36.4233	1.6263	1.6791	2.7653	2.1556	0.6818	0.6703
8	40.9121	43.2060	1.6869	1.6669	2.5795	2.4781	0.8998	0.8603

Tabela 4 - Resultados do MAE para modelos com dados COPERNICUS

===== MÉTRICAS: VERSÃO COMPLETA =====

--- RMSE ---

target	mwd_cop	mwd_era5	mwp_cop	mwp_era5	peak_period_cop	pp1d_era5	swl_cop	swl_era5
horizon								
1	33.4454	35.3029	2.0984	2.1369	3.5097	2.9980	0.8674	0.8894
2	41.0105	43.4656	2.1118	2.1899	3.2646	2.8841	0.8347	0.9726
4	53.3808	47.7136	2.1988	2.2455	3.5597	2.9646	0.9645	0.9552
8	58.7702	54.8706	1.8660	2.2326	3.2150	3.2715	0.9798	0.9304

Tabela 5 - Resultados do RMSE para modelos com dados COPERNICUS+ERA

===== MÉTRICAS: SEM VENTO/CORRENTE =====

--- RMSE ---

target	mwd_cop	mwd_era5	mwp_cop	mwp_era5	peak_period_cop	pp1d_era5	swl_cop	swl_era5
horizon								
1	36.5572	38.0006	2.0367	1.8285	2.7939	2.7459	0.9169	0.6524
2	35.1422	43.0122	1.8949	1.8374	3.1402	3.1778	0.7026	0.6279
4	49.0133	53.8515	2.0014	2.0605	3.3033	2.8674	0.7281	0.9835
8	61.4538	60.1788	2.0019	1.9388	3.0744	2.9237	0.9860	0.8444

Tabela 6 - Resultados do RMSE para modelos com dados COPERNICUS+ERA sem influências de vento e corrente

--- MAE ---

target	mwd_cop	mwd_era5	mwp_cop	mwp_era5	peak_period_cop	pp1d_era5	swh_cop	swh_era5
horizon								
1	21.1336	21.6317	1.6528	1.7158	2.7876	2.4083	0.6785	0.6747
2	24.8424	28.8106	1.6296	1.7548	2.6047	2.3536	0.6149	0.7930
4	34.8861	32.7832	1.7386	1.7613	2.8099	2.3545	0.7429	0.7391
8	41.2980	38.9116	1.5364	1.7723	2.5729	2.6166	0.7171	0.7254

Tabela 7 - Resultados do MAE para modelos com dados COPERNICUS+ERA

--- MAE ---

target	mwd_cop	mwd_era5	mwp_cop	mwp_era5	peak_period_cop	pp1d_era5	swh_cop	swh_era5
horizon								
1	24.7626	24.3279	1.6407	1.4670	2.2616	2.1604	0.6871	0.4942
2	23.5829	29.4931	1.5382	1.4838	2.4894	2.5786	0.5428	0.4492
4	34.6385	38.2826	1.5795	1.6622	2.6560	2.2970	0.5033	0.7199
8	45.7240	45.7473	1.5102	1.5469	2.4993	2.4309	0.7615	0.6136

Tabela 8 - Resultados do MAE para modelos com dados COPERNICUS+ERA sem influências de vento e corrente

>>> RMSE Completo (Escalares + Direção): Menor é melhor

Horizonte	1		3		6		12		24	
Input_Source	ERA5 (Com Vento)	ERA5 (Sem Vento)	ERA5 (Com Vento)	ERA5 (Sem Vento)	ERA5 (Com Vento)	ERA5 (Sem Vento)	ERA5 (Com Vento)	ERA5 (Sem Vento)	ERA5 (Com Vento)	ERA5 (Sem Vento)
Target										
Direção Média (Copernicus)	19.2828	19.8699	29.6405	32.2515	45.0927	49.8685	51.3315	56.4498	55.8620	55.8141
Direção Média (ERA5)	14.0360	13.8130	27.7934	30.6623	35.5089	48.4672	50.9837	58.9756	57.8515	58.5702
MWP (Copernicus)	0.6919	0.7796	0.9115	1.0454	1.4221	1.4645	1.5822	1.5861	1.6159	1.6932
MWP (ERA5)	0.2827	0.2941	0.7061	0.7765	1.3155	1.3310	1.5268	1.5657	1.5765	1.6487
Peak Period (Copernicus)	1.7882	1.7841	2.0373	2.1195	2.4100	2.5240	2.8902	2.8428	2.7366	2.8392
Peak Period (ERA5)	0.9471	0.9757	1.5153	1.5801	2.0741	2.1264	2.6060	2.5534	2.4714	2.5112
SWH (Copernicus)	0.1687	0.1708	0.3520	0.3561	0.5239	0.5718	0.8744	0.9579	0.9435	1.0890
SWH (ERA5)	0.1570	0.1596	0.3413	0.3509	0.5188	0.5635	0.8272	0.9443	0.8924	1.0072

Tabela 9- Resultados da métrica RMSE para os modelos baseados em dados ERA5.

>>> MAE Completo (Escalares + Direção): Menor é melhor

Horizonte	1		3		6		12		24	
Input_Source	ERA5 (Com Vento)	ERA5 (Sem Vento)	ERA5 (Com Vento)	ERA5 (Sem Vento)	ERA5 (Com Vento)	ERA5 (Sem Vento)	ERA5 (Com Vento)	ERA5 (Sem Vento)	ERA5 (Com Vento)	ERA5 (Sem Vento)
Target										
Direção Média (Copernicus)	9.7714	10.0877	16.0780	18.4459	27.7540	32.7936	37.5944	43.3045	43.0781	42.5213
Direção Média (ERA5)	5.6640	6.4388	14.3651	18.0529	22.0887	32.3965	37.3136	44.7111	43.5884	43.6873
MWP (Copernicus)	0.5154	0.5837	0.6950	0.7949	1.1447	1.1750	1.2352	1.2827	1.3143	1.3485
MWP (ERA5)	0.1941	0.2093	0.5201	0.5911	0.9919	1.0542	1.1184	1.1823	1.2208	1.2746
Peak Period (Copernicus)	1.0610	1.0429	1.3660	1.4359	1.7532	1.8304	2.1547	2.1110	2.1484	2.2405
Peak Period (ERA5)	0.4398	0.4678	0.9642	1.0272	1.4560	1.4895	1.9571	1.9723	1.8525	1.9038
SWH (Copernicus)	0.1091	0.1155	0.2016	0.2086	0.3341	0.3734	0.6145	0.6803	0.6730	0.7996
SWH (ERA5)	0.0680	0.0719	0.1783	0.1962	0.3172	0.3641	0.5942	0.6629	0.6347	0.7219

Tabela 10 – Resultados da métrica MAE para os modelos baseados em dados ERA5.

>>> R² Completo (Escalares + Direção): Maior é melhor

Horizonte	1		3		6		12		24	
Input_Source	ERA5 (Com Vento)	ERA5 (Sem Vento)	ERA5 (Com Vento)	ERA5 (Sem Vento)	ERA5 (Com Vento)	ERA5 (Sem Vento)	ERA5 (Com Vento)	ERA5 (Sem Vento)	ERA5 (Com Vento)	ERA5 (Sem Vento)
Target										
Direção Média (Copernicus)	0.8605	0.8518	0.6703	0.6097	0.2347	0.0641	0.0050	-0.2033	-0.1904	-0.1884
Direção Média (ERA5)	0.9306	0.9328	0.7279	0.6688	0.5549	0.1708	0.0802	-0.2308	-0.1929	-0.2228
MWP (Copernicus)	0.7421	0.6725	0.5523	0.4111	-0.0877	-0.1537	-0.3440	-0.3506	-0.3925	-0.5288
MWP (ERA5)	0.9611	0.9579	0.7575	0.7068	0.1595	0.1396	-0.1304	-0.1885	-0.1984	-0.3106
Peak Period (Copernicus)	0.4798	0.4821	0.3247	0.2691	0.0551	-0.0365	-0.3594	-0.3152	-0.2215	-0.3148
Peak Period (ERA5)	0.8274	0.8167	0.5580	0.5194	0.1716	0.1294	-0.3083	-0.2559	-0.1777	-0.2160
SWH (Copernicus)	0.9336	0.9319	0.7110	0.7042	0.3601	0.2379	-0.7812	-1.1374	-1.0716	-1.7595
SWH (ERA5)	0.9338	0.9316	0.6872	0.6694	0.2776	0.1476	-0.8361	-1.3925	-1.1348	-1.7195

Tabela 11 – Resultados da métrica R² para os modelos baseados em dados ERA5.

>>> RMSE Completo (Escalares + Direção): Impacto da Corrente (Menor é melhor)

Horizonte	1		3		6		12		24	
Input_Source	Copernicus (Com Corrente)	Copernicus (Sem Corrente)	Copernicus (Com Corrente)	Copernicus (Sem Corrente)	Copernicus (Com Corrente)	Copernicus (Sem Corrente)	Copernicus (Com Corrente)	Copernicus (Sem Corrente)	Copernicus (Com Corrente)	Copernicus (Sem Corrente)
Target										
Direção Média (Copernicus)	12.7903	12.5898	26.9910	26.7561	39.4163	40.2377	53.6306	55.4545	55.5178	58.7328
Direção Média (ERA5)	25.0704	25.4225	33.9055	33.5819	42.7800	42.9583	56.9748	58.5180	57.2248	57.5590
MWP (Copernicus)	0.3599	0.3649	0.8561	0.8892	1.3620	1.3928	1.6842	1.6914	1.5613	1.5771
MWP (ERA5)	0.5969	0.6023	0.9207	0.9053	1.2616	1.3039	1.4988	1.5360	1.5770	1.5933
Peak Period (Copernicus)	1.2260	1.2827	1.9081	1.9666	2.3352	2.4472	2.7868	2.8574	2.7521	2.8437
Peak Period (ERA5)	1.5493	1.5165	1.7047	1.7149	2.1492	2.1709	2.4397	2.3978	2.5016	2.5553
SWH (Copernicus)	0.1525	0.1497	0.4004	0.4091	0.6260	0.6353	0.9334	0.8680	0.9505	0.9699
SWH (ERA5)	0.1944	0.1913	0.4195	0.4176	0.6241	0.6316	0.8360	0.8331	0.8836	0.8989

Tabela 12 – Resultados da métrica RMSE para os modelos baseados em dados Copernicus.

>>> MAE Completo (Escalares + Direção): Menor é melhor

Horizonte	1			3		6		12		24	
Input_Source	Copernicus (Com Corrente)	Copernicus (Sem Corrente)	Copernicus (Com Corrente)	Copernicus (Sem Corrente)	Copernicus (Com Corrente)	Copernicus (Sem Corrente)	Copernicus (Com Corrente)	Copernicus (Sem Corrente)	Copernicus (Com Corrente)	Copernicus (Sem Corrente)	Copernicus (Com Corrente)
Target											
Direção Média (Copernicus)	5.9613	5.8342	15.8480	15.7007	26.5436	26.7349	40.1971	41.9919	43.5023	46.1100	
Direção Média (ERA5)	13.6095	13.6926	20.5955	20.3931	29.5941	29.0913	42.9321	44.6457	44.6949	43.7280	
MWP (Copernicus)	0.2442	0.2479	0.6268	0.6441	1.0585	1.0881	1.3734	1.3827	1.2598	1.2679	
MWP (ERA5)	0.4552	0.4541	0.6450	0.6255	0.9652	0.9766	1.1172	1.1553	1.1772	1.2016	
Peak Period (Copernicus)	0.6053	0.6215	1.2581	1.2799	1.6971	1.7889	2.1111	2.1679	2.1422	2.2742	
Peak Period (ERA5)	0.9172	0.8860	1.1032	1.0965	1.4950	1.5180	1.8275	1.8024	1.9220	2.0218	
SWH (Copernicus)	0.0887	0.0846	0.2344	0.2352	0.4139	0.4071	0.6858	0.6444	0.7136	0.7383	
SWH (ERA5)	0.1183	0.1170	0.2430	0.2427	0.4146	0.4041	0.6104	0.6033	0.6580	0.6648	

Tabela 13 – Resultados da métrica MAE para os modelos baseados em dados Copernicus.

>>> R² Completo (Escalares + Direção): Qualidade da Previsão (Maior é melhor)

Horizonte	1			3		6		12		24	
Input_Source	Copernicus (Com Corrente)	Copernicus (Sem Corrente)	Copernicus (Com Corrente)	Copernicus (Sem Corrente)	Copernicus (Com Corrente)	Copernicus (Sem Corrente)	Copernicus (Com Corrente)	Copernicus (Sem Corrente)	Copernicus (Com Corrente)	Copernicus (Sem Corrente)	Copernicus (Com Corrente)
Target											
Direção Média (Copernicus)	0.9386	0.9405	0.7266	0.7314	0.4153	0.3907	-0.0861	-0.1613	-0.1758	-0.3159	
Direção Média (ERA5)	0.7786	0.7724	0.5951	0.6028	0.3540	0.3486	-0.1487	-0.2117	-0.1672	-0.1809	
MWP (Copernicus)	0.9302	0.9283	0.6051	0.5740	0.0021	-0.0434	-0.5229	-0.5358	-0.2999	-0.3263	
MWP (ERA5)	0.8267	0.8236	0.5877	0.6014	0.2271	0.1744	-0.0892	-0.1440	-0.1991	-0.2241	
Peak Period (Copernicus)	0.7555	0.7323	0.4077	0.3708	0.1128	0.0257	-0.2638	-0.3286	-0.2354	-0.3189	
Peak Period (ERA5)	0.5380	0.5573	0.4406	0.4339	0.1105	0.0925	-0.1466	-0.1075	-0.2068	-0.2591	
SWH (Copernicus)	0.9457	0.9477	0.6260	0.6095	0.0866	0.0591	-1.0297	-0.7553	-1.1024	-1.1890	
SWH (ERA5)	0.8985	0.9017	0.5273	0.5316	-0.0454	-0.0709	-0.8753	-0.8621	-1.0930	-1.1659	

Tabela 14 – Resultados da métrica R^2 para os modelos baseados em dados Copernicus.

>>> RMSE Global: Só Ondas vs. Completo (Menor é melhor)

Horizonte	1		3		6		12		24	
Input_Source	Completo (Tudo)	Só Ondas (Inércia)	Completo (Tudo)	Só Ondas (Inércia)	Completo (Tudo)	Só Ondas (Inércia)	Completo (Tudo)	Só Ondas (Inércia)	Completo (Tudo)	Só Ondas (Inércia)
Target										
Direção Média (Copernicus)	11.7040	12.4831	22.6919	25.4866	35.1586	40.7074	51.4092	54.9640	55.2296	56.7972
Direção Média (ERA5)	14.0681	14.2886	28.6067	30.8803	34.3716	45.9998	51.0346	57.0850	56.8473	57.5825
MWP (Copernicus)	0.3470	0.3679	0.7767	0.8393	1.2640	1.3352	1.6315	1.6015	1.5862	1.6023
MWP (ERA5)	0.2806	0.2879	0.7015	0.7309	1.2837	1.2684	1.5153	1.5352	1.5515	1.5753
Peak Period (Copernicus)	1.1655	1.2339	1.7211	1.7372	2.3482	2.3703	2.8119	2.7194	2.6408	2.8636
Peak Period (ERA5)	1.0159	0.9868	1.4366	1.4272	2.0452	2.0588	2.5675	2.4891	2.4409	2.5112
SWH (Copernicus)	0.1445	0.1426	0.3595	0.3582	0.5424	0.5652	0.8722	0.9118	0.9057	0.9983
SWH (ERA5)	0.1580	0.1539	0.3515	0.3537	0.5277	0.5487	0.8353	0.9023	0.8381	0.9276

Tabela 15– Resultados da métrica RMSE para os modelos com dados Combinados (ERA5 + Copernicus).

>>> MAE Global: Só Ondas vs. Completo (Menor é melhor)

Horizonte	1		3		6		12		24	
Input_Source	Completo (Tudo)	Só Ondas (Inércia)	Completo (Tudo)	Só Ondas (Inércia)	Completo (Tudo)	Só Ondas (Inércia)	Completo (Tudo)	Só Ondas (Inércia)	Completo (Tudo)	Só Ondas (Inércia)
Target										
Direção Média (Copernicus)	4.9833	5.6084	12.4464	15.0534	21.6894	27.5825	38.0233	41.8220	42.1574	44.7744
Direção Média (ERA5)	5.6318	6.3625	14.8514	17.1499	21.6441	29.3451	37.8969	43.0510	42.4092	43.5538
MWP (Copernicus)	0.2327	0.2545	0.5595	0.6223	0.9769	1.0302	1.2794	1.2880	1.2819	1.2788
MWP (ERA5)	0.1942	0.1986	0.5138	0.5469	0.9429	0.9476	1.1286	1.1570	1.1706	1.1954
Peak Period (Copernicus)	0.5916	0.6220	1.1762	1.1867	1.6981	1.7123	2.1595	2.0659	2.0454	2.2637
Peak Period (ERA5)	0.5079	0.4942	0.9280	0.9318	1.4586	1.4646	1.9717	1.9195	1.8451	1.9595
SWH (Copernicus)	0.0796	0.0789	0.2017	0.2040	0.3438	0.3739	0.6120	0.6717	0.6633	0.7465
SWH (ERA5)	0.0689	0.0690	0.1847	0.1931	0.3296	0.3576	0.5974	0.6495	0.6102	0.6820

Tabela 16 – Resultados da métrica MAE para os modelos com dados Combinados (ERA5 + Copernicus).

>>> R² Global: Só Ondas vs. Completo (Maior é melhor)

Horizonte	1		3		6		12		24	
Input_Source	Completo (Tudo)	Só Ondas (Inércia)	Completo (Tudo)	Só Ondas (Inércia)	Completo (Tudo)	Só Ondas (Inércia)	Completo (Tudo)	Só Ondas (Inércia)	Completo (Tudo)	Só Ondas (Inércia)
Target										
Direção Média (Copernicus)	0.9486	0.9415	0.8068	0.7562	0.5348	0.3763	0.0020	-0.1408	-0.1636	-0.2306
Direção Média (ERA5)	0.9303	0.9281	0.7118	0.6641	0.5830	0.2531	0.0784	-0.1531	-0.1519	-0.1819
MWP (Copernicus)	0.9351	0.9271	0.6749	0.6205	0.1406	0.0411	-0.4290	-0.3770	-0.3417	-0.3692
MWP (ERA5)	0.9617	0.9597	0.7606	0.7402	0.1997	0.2187	-0.1133	-0.1427	-0.1607	-0.1965
Peak Period (Copernicus)	0.7790	0.7523	0.5181	0.5090	0.1029	0.0859	-0.2867	-0.2034	-0.1375	-0.3375
Peak Period (ERA5)	0.8014	0.8126	0.6028	0.6080	0.1946	0.1839	-0.2699	-0.1936	-0.1488	-0.2160
SWH (Copernicus)	0.9513	0.9525	0.6985	0.7007	0.3142	0.2552	-0.7721	-0.9371	-0.9091	-1.3191
SWH (ERA5)	0.9329	0.9364	0.6682	0.6639	0.2525	0.1919	-0.8722	-1.1845	-0.8830	-1.3068

Tabela 17 – Resultados da métrica R² para os modelos com dados Combinados (ERA5 + Copernicus).

	MAE	MAE2	MAE3	RMSE1	RMSE2	RMSE3
Cenário	COP	ERA5	ERA5+COP	COP	ERA5	ERA5+COP
mwp_cop	1.2819	0.428	0.4376	1.5426	0.5901	0.5992
peak_period_cop	1.86	0.8688	0.949	2.4651	1.6299	1.6936
swh_cop	1.1132	0.0858	0.0895	1.3572	0.1107	0.1162
med_deg	43.06	7.063	7.314	54.199	14.496	14.354

Tabela 18 - Resultados Resumidos para horizonte de 24h das várias métricas e parâmetros de estudo

Horizonte(h)	MAE_ sw_h_cop	MAE_ mwp_cop	MAE_peak_ period_cop	MAE_ mwd_deg	RMSE_ sw_h_cop	RMSE_ mwp_cop	RMSE_peak_ period_cop	RMSE_ mwd_deg
1	0.088	0.255	0.569	4.37	0.182	0.371	1.157	8.789
3	0.217	0.626	1.122	12.947	0.436	0.843	1.798	22.593
6	0.366	1.003	1.551	25.136	0.624	1.303	2.176	37.373
12	0.552	1.301	1.687	37.42	0.808	1.636	2.156	50.238
24	0.736	1.267	1.716	38.99	1.016	1.653	2.156	51.396

Tabela 19 - Valores dos erros ao longo do tempo