



UNIVERSIDADE
FEDERAL
DE PERNAMBUCO

UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE INFORMÁTICA
CURSO DE BACHARELADO EM ENGENHARIA DA COMPUTAÇÃO

JOÃO VITOR DA SILVA GOMES

**ESTRATÉGIAS PARA A PREVISÃO DE MÚLTIPLOS HORIZONTES
USANDO REDES NEURAIS ARTIFICIAIS: UMA APLICAÇÃO EM
SÉRIES TEMPORAIS DE POTÊNCIA EÓLICA**

RECIFE

2021

UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE INFORMÁTICA
CURSO DE BACHARELADO EM ENGENHARIA DA COMPUTAÇÃO

JOÃO VITOR DA SILVA GOMES

**ESTRATÉGIAS PARA A PREVISÃO DE MÚLTIPLOS HORIZONTES
USANDO REDES NEURAS ARTIFICIAIS: UMA APLICAÇÃO EM
SÉRIES TEMPORAIS DE POTÊNCIA EÓLICA**

Monografia apresentada ao Centro de Informática (CIN) da Universidade Federal de Pernambuco (UFPE) como requisito parcial para conclusão do Curso de Engenharia da Computação, orientada pelo professor Tsang Ing Ren.

RECIFE

2021

UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE INFORMÁTICA
CURSO DE BACHARELADO EM ENGENHARIA DA COMPUTAÇÃO

JOÃO VITOR DA SILVA GOMES

**ESTRATÉGIAS PARA A PREVISÃO DE MÚLTIPLOS HORIZONTES
USANDO REDES NEURAS ARTIFICIAIS: UMA APLICAÇÃO EM
SÉRIES TEMPORAIS DE POTÊNCIA EÓLICA**

Monografia submetida ao corpo docente da Universidade Federal de Pernambuco, defendida e aprovada em 24 de agosto de 2021.

Banca Examinadora:

Prof. Tsang Ing Ren

Doutor

Orientador

Prof. Paulo Salgado Gomes de Mattos Neto

Doutor

Examinador

AGRADECIMENTOS

Gostaria de agradecer primeiramente a minha mãe por me aconselhar a jamais desistir de meus sonhos. Também quero agradecer ao professor Tsang Ing Ren pela orientação nos projetos de pesquisa durante a graduação, além de toda a equipe de *Deep Learning* do Centro de Energias Renováveis (CER-UFPE). No mais, desejo muitíssimo obrigado a Emmanuel Carreira e Sérgio Victor, companheiros de turma com os quais construí ótimas memórias ao longo desta jornada.

*“It is sometimes a mistake to climb; it is always
a mistake never even to make the attempt.”*

Neil Gaiman

RESUMO

A modelagem do comportamento das massas de ar é uma atividade de grande importância para a integração da energia eólica na matriz energética. Alterações inoportunas na disponibilidade dos ventos tendem a gerar instabilidades no balanço de produção e consumo de eletricidade, e por isso precisam ser previstas a tempo de uma tomada de decisões. Algoritmos de aprendizado de máquina são modelos orientados a dados cuja aplicação em séries temporais de potência eólica ainda vem sendo explorada. Estes modelos são capazes de reconhecer padrões contidos nos dados e então realizar previsões de não somente um, mas de vários instantes no futuro (horizontes). Este trabalho compara o uso de diferentes estratégias para previsão de múltiplos horizontes (direta, recursiva e de múltiplas saídas) aliadas a redes neurais artificiais *Multilayer Perceptron* e *Long Short-Term Memory*. Foram utilizados dados públicos de 3 parques eólicos australianos com amostras de potência coletadas a cada 30 minutos durante todo o ano de 2018. Os resultados revelaram que no tocante a diferença entre estratégias, a recursiva é uma escolha a ser evitada devido a rápida propagação do erro. Já a estratégia de múltiplas saídas é bem-vinda em casos onde a capacidade de processamento é reduzida, embora seja necessário abrir mão da precisão nos primeiros horizontes. Por fim, a estratégia direta apresentou a melhor média geral do RMSE principalmente quando aliada a redes neurais *Multilayer Perceptron*.

Palavras-chave: *Deep Learning*, Energia Eólica, Séries Temporais, Previsão de múltiplos horizontes.

SUMÁRIO

1. Introdução	12
2. Conceitos Básicos	13
2.1 Séries Temporais	13
2.2 Redes Neurais Artificiais	14
2.2.1 <i>Multilayer Perceptron</i>	14
2.2.2 <i>Long Short-Term Memory</i>	15
2.3 Estratégias para Previsão de Múltiplos Horizontes	16
2.3.1 Estratégia Direta	17
2.3.2 Estratégia Recursiva	18
2.3.3 Estratégia de Múltiplas Saídas	19
3. Estado da Arte	20
4. Experimento	21
4.1 Dados	21
4.2 Pré-processamento	23
4.3 Modelos	23
5. Resultados e Discussão	25
5.1 Central Star Hill	25
5.2 Central Boco Rock	27
5.3 Central Macarthur	28
6. Conclusão	30
6.1 Trabalhos Futuros	30
7. Referências	31
Apêndice	33

LISTA DE FIGURAS

- Figura 1. Estrutura de uma rede neural *Multilayer Perceptron*
- Figura 2. Esquema de um *Perceptron*
- Figura 3. Estrutura de uma camada de unidades *Long Short-Term Memory*
- Figura 4. Esquema interno de uma unidade LSTM
- Figura 5. Camadas LSTM empilhadas
- Figura 6. Previsões por meio da estratégia direta
- Figura 7. Previsões por meio da estratégia recursiva
- Figura 8. Previsões por meio da estratégia de múltiplas saídas
- Figura 9. Metodologia do experimento
- Figura 10. Localização geográfica dos parques eólicos no sul da Austrália
- Figura 11. Séries temporais de potência referentes a cada parque eólico
- Figura 12. *Splits* da validação *walk-forward*
- Figura 13. Modelo da persistência para o horizonte k
- Figura 14. Média geral do RMSE (STARHLWF)
- Figura 15. Média do RMSE para os horizontes 1 e 12 (STARHLWF)
- Figura 16. Média geral do RMSE (BOCORWF1)
- Figura 17. Média do RMSE para os horizontes 1 e 12 (BOCORWF1)
- Figura 18. Média geral do RMSE (MACARTH1)
- Figura 19. Média do RMSE para os horizontes 1 e 12 (MACARTH1)
- Figura 20. Média do RMSE para cada horizonte (STARHLWF)
- Figura 21. Média do RMSE para cada horizonte (BOCORWF1)
- Figura 22. Média do RMSE para cada horizonte (MACARTH1)

LISTA DE TABELAS

Tabela 1 - Estatísticas das séries de potência em megawatts

Tabela 2 - Melhores parâmetros para cada par de rede e estratégia (STARHLWF)

Tabela 3 - Melhores parâmetros para cada par de rede e estratégia (BOCORWF1)

Tabela 4 - Melhores parâmetros para cada par de rede e estratégia (MACARTH1)

Tabela 5 - Média geral do RMSE (STARHLWF)

Tabela 6 - Média geral do RMSE (BOCORWF1)

Tabela 7 - Média geral do RMSE (MACARTH1)

TABELA DE SIGLAS

Sigla	Significado	Página
ARMA	Autorregressivo de Média Móvel	13
RNA	Rede Neural Artificial	14
MLP	<i>Multilayer Perceptron</i>	14
LSTM	<i>Long Short-Term Memory</i>	15
ARIMA	Autorregressivo Integrado de Média Móvel	20
SVM	<i>Support Vector Machine</i>	20
kNN	<i>k-Nearest Neighbors</i>	20
RBF	<i>Radial Basis Function</i>	20
RMSE	Raiz do Erro Quadrático Médio	20
MAE	Erro Absoluto Médio	20
MAPE	Erro Percentual Absoluto Médio	20
MW	Megawatt	20
PCA	Análise de Componentes Principais	20
AEMO	Operador do Mercado de Eletricidade Australiano	21
STARHLWF	Parque eólico Star Hill	21
BOCORWF1	Parque eólico Boco Rock	21
MACARTH1	Parque eólico Macarthur	21
NaN	<i>Not-a-Number</i>	23

1. Introdução

O aumento significativo na instalação de usinas eólicas deve-se principalmente ao mínimo impacto ambiental causado quando comparadas a fontes não renováveis. Por outro lado, a natureza intermitente dos ventos faz com que a potência gerada possa flutuar, levando a instabilidades na rede elétrica e prejuízos de mercado [1]. É visando tornar essa fonte mais confiável que muito esforço vem sendo aplicado no desenvolvimento de sistemas de previsão.

Atualmente, as previsões de vento são realizadas através de duas abordagens: a abordagem numérica, modelando o comportamento atmosférico por meio das equações físicas que regem as massas de ar; ou a abordagem estatística, reconhecendo padrões contidos em séries temporais. A primeira é utilizada para resolver previsões a longo prazo, enquanto que a segunda é apropriada para previsões a curto prazo [2] devido à característica persistente dos ventos [3].

É aproveitando o crescente desenvolvimento da área de aprendizagem de máquina que a abordagem baseada em séries temporais vem andando a passos largos. Muito se deve à capacidade destes modelos em extrair dependências não lineares das amostras históricas de potência. Como exemplo pode-se citar as redes neurais artificiais que aprendem através da apresentação sucessiva de sequências de entrada e saída, ajustando suas unidades internas de processamento na direção da minimização do erro de previsão.

Embora estejam aptas a reconhecer dependências não lineares, redes neurais como *Multilayer Perceptron* e *Long Short-Term Memory* também carregam consigo a flexibilidade de construção e definição de hiperparâmetros. Não é trivial determinar qual a combinação de hiperparâmetros melhor generalizam o comportamento dos dados, sendo necessário escolhê-los de forma empírica, treinando, testando e avaliando as combinações desejadas. Uma das principais decisões em aplicações de séries temporais é definir a quantidade de variáveis de entrada e de saída do modelo. Geralmente deseja-se gerar previsões para múltiplos instantes no futuro, conhecidos como horizontes, havendo algumas estratégias para tal. Dentre elas podem ser citadas: a estratégia direta, que cria um modelo para cada horizonte; a estratégia recursiva, que utiliza apenas um modelo para prever todos os horizontes recursivamente; e a estratégia de múltiplas saídas, que tem como base uma única rede neural capaz de se ajustar a todos os horizontes simultaneamente [4].

O objetivo geral deste trabalho é prover uma comparação entre as diferentes estratégias para a previsão de múltiplos horizontes utilizando redes neurais *Multilayer Perceptron* e *Long Short-Term Memory* no tocante a séries temporais de potência eólica.

Para atingir o objetivo principal será necessário atender aos seguintes pontos:

- Selecionar parques eólicos com características suficientemente distintas;
- Tratar os dados de potência para que possam ser usados como entrada para os modelos;
- Variar os hiperparâmetros das redes em busca da configuração de melhor desempenho.

2. Conceitos Básicos

Para que possamos prosseguir com o pleno entendimento deste estudo se faz necessário o conhecimento de alguns conceitos básicos. Estes conceitos serão citados com certa frequência a partir deste segundo capítulo e estão dispostos de forma que: na Seção 2.1 é introduzida a definição de série temporal e como funcionam os modelos autorregressivos clássicos; na Seção 2.2 são apresentadas as redes neurais artificiais *Multilayer Perceptron* e *Long Short-Term Memory*; e na Seção 2.3 são descritas as diferentes estratégias para previsão de múltiplos horizontes, bem como suas vantagens e desvantagens.

2.1 Séries Temporais

Uma série temporal é uma coleção de observações tomadas sucessivamente a cada intervalo de tempo pré-definido. É um tipo de dados obtido através de qualquer variável que tenha seu valor modificado ao longo do tempo. Por isso está presente em diversas áreas, como por exemplo estatística, econometria, processamento de sinais e meteorologia. As observações que compõem a série temporal podem ser feitas sob qualquer intervalo de tempo, desde que fixo. Seja usando uma escala mensal, diária, horária, a cada minuto, segundo, ou fração de segundo, o intervalo entre coletas depende muitas vezes das características inerentes à aplicação em questão.

Uma notação comum para especificar uma série temporal é tomar uma variável X e indexá-la através de números naturais, formando uma sequência da forma:

$$X = \{X_0, X_1, X_2, \dots, X_{t-1}, X_t\} \quad (1)$$

onde X_0 é o valor da variável X quando no instante inicial, X_1 é o seu valor após 1 intervalo de tempo, e assim sucessivamente, até X_t que se refere ao valor no instante atual.

Devido a esta característica sequencial no tempo, as observações de uma série temporal tendem a estar de alguma forma relacionadas. Diversos modelos são construídos com o intuito de identificar estas dependências e então gerar previsões para instantes futuros. Entretanto nem sempre é possível modelar estas relações. Elas podem ser simples como uma combinação linear, um pouco mais complexas como uma combinação não linear, ou sequer apresentar relação como no caso de dados completamente aleatórios.

Por exemplo, modelos autorregressivos de médias móveis (ARMA), assumem que o valor futuro de uma variável possa ser uma função linear de observações e erros defasados no tempo [5]. Tipicamente, modelos ARMA(p, q) tentam extrair o processo gerador de uma série temporal linear, de forma a revelar que esta segue a estrutura:

$$X_t = c + \epsilon_t + \sum_{i=1}^p \varphi_i X_{t-i} + \sum_{j=1}^q \theta_j \epsilon_{t-j} \quad (2)$$

onde X_t e ϵ_t são o valor da série e o erro no instante atual t . Os termos φ_i e θ_j são coeficientes que caracterizam a importância de cada valor passado X_{t-i} e erro passado ϵ_{t-j} . Os inteiros p e q são referidos como ordem do modelo e c é uma constante de ajuste.

Todavia, séries de vento tendem a apresentar dependências não lineares e um modelo ARMA se torna insuficiente para descrever a relação entre os dados históricos. Para lidar com o reconhecimento de relações não lineares podem ser utilizadas redes neurais artificiais, tidos como modelos de aprendizagem de máquina.

2.2 Redes Neurais Artificiais

As redes neurais artificiais (RNA) são modelos computacionais capazes de reconhecer padrões complexos adquirindo conhecimento através da exposição a exemplos de treinamento. As RNAs têm inspiração no sistema nervoso animal, simulando o funcionamento de neurônios e suas conexões. Nesse sentido, na área de aprendizagem profunda, as redes são tipicamente estruturadas por camadas interconectadas de unidades de processamento, sendo uma camada de entrada, uma de saída, além de uma ou mais camadas ocultas. Quanto à orientação de propagação dos sinais, estes modelos podem ser classificados como *feedforward* ou recorrente.

2.2.1 Multilayer Perceptron

Multilayer Perceptron (MLP) são redes do tipo *feedforward*, ou em tradução livre alimentada em avanço, pois permitem que os sinais sigam em apenas uma direção: da entrada para a saída. Não havendo, inclusive, compartilhamento de informações entre unidades de uma mesma camada.

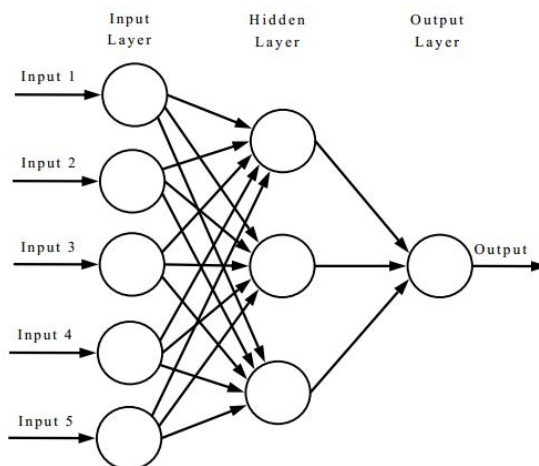


Figura 1. Estrutura de uma rede neural *Multilayer Perceptron*

As camadas de uma MLP são compostas por unidades de processamento chamadas *Perceptrons* [6], cujo esquema pode ser visualizado na Figura 2. O funcionamento destas unidades é caracterizado pela Equação 3 e acontece da seguinte maneira:

1. Os sinais x_i são apresentados à entrada;
2. Cada sinal é ponderado por um coeficiente (peso sináptico) w_i ;
3. A combinação linear resultante gera um nível de atividade net ;
4. Esta atividade é avaliada por uma função de ativação f , levando a emissão de um sinal de saída y .

$$y = f\left(\sum_{i=1}^n w_i x_i - b\right) \quad (3)$$

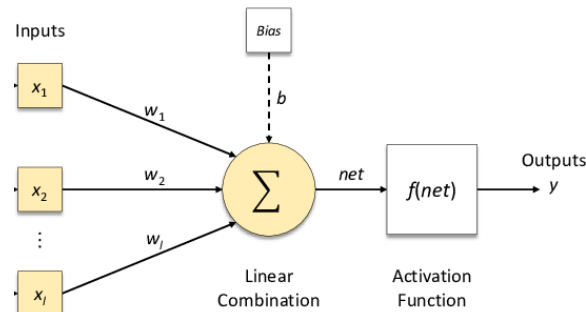


Figura 2. Esquema de um *Perceptron*

Durante a fase de treinamento padrões de entrada associados a padrões de saída são apresentados iterativamente ao modelo. Aqui os padrões de entrada e os pesos sinápticos atuais produzem sinais de saída cujos valores diferem daqueles desejados. Visando minimizar o erro a ser cometido ao longo das próximas iterações, o erro atual é retropropagado da camada de saída até a entrada, corrigindo os pesos sinápticos e levando a rede ao estado que melhor se ajusta aos dados.

2.2.2 Long Short-Term Memory

Modelos *Long Short-Term Memory* (LSTM) são redes neurais do tipo recorrente capazes de capturar dependências a longo prazo. Elas proporcionam a manutenção de um estado de memória que pode ser parcialmente esquecido, incrementado e propagado [7] entre unidades de uma mesma camada.

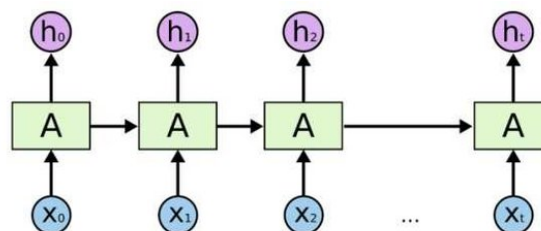


Figura 3. Estrutura de uma camada de unidades *Long Short-Term Memory*

Uma unidade LSTM é composta por 3 válvulas que realizam o controle do estado de memória por meio de funções sigmóide. As funções sigmóide σ geram saídas entre 0 e 1 indicando o quanto de informação deve-se permitir passar, onde 0 indica que toda a informação deve ser bloqueada e 1 indica que toda informação deve seguir adiante [8].

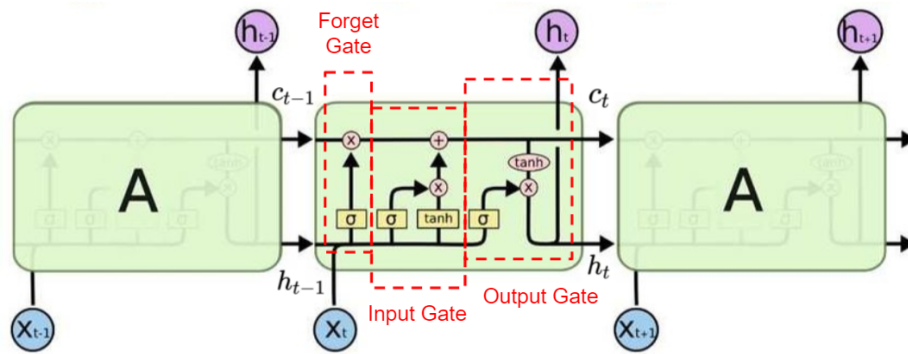


Figura 4. Esquema interno de uma unidade LSTM

A Figura 4 mostra o esquema interno de uma unidade LSTM. Cada válvula é designada a cumprir as seguintes especificações:

- A válvula do esquecimento (*forget gate*) controla quanto da informação contida no estado de memória prévio c_{t-1} deve ser descartada;
- A válvula de entrada (*input gate*) controla quanto da informação de entrada x_t deve ser considerada para integrar o estado de memória atual c_t ;
- A válvula de saída (*output gate*) controla quanto da informação contida na memória atual deve fazer parte do sinal de saída h_t .

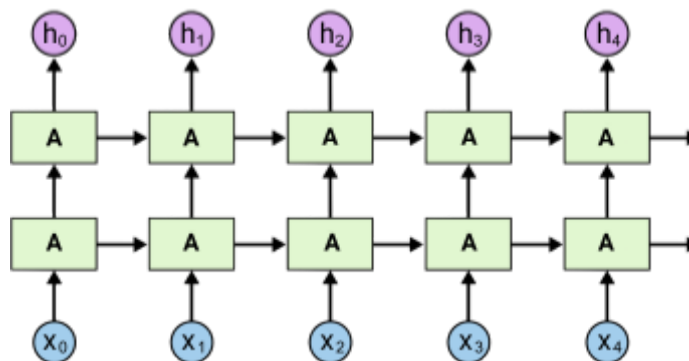


Figura 5. Camadas LSTM empilhadas

2.3 Estratégias para Previsão de Múltiplos Horizontes

Em previsões de séries temporais de vento, modelar vários instantes no futuro compreende papel fundamental na tomada de decisões. Por exemplo, considerando dados

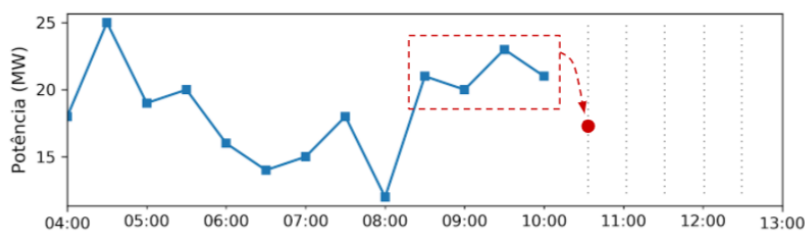
tomados a cada 30 minutos desde o momento inicial até o instante atual t , deseja-se prever a potência gerada 30 minutos à frente, 1 hora à frente, 1 hora e meia à frente, e assim por diante. Estes instantes futuros são denominados horizontes ou passos e são denotados respectivamente por $t + 1$, $t + 2$, $t + 3$,...

Modelos de redes neurais como *Multilayer Perceptron* e *Long Short-Term Memory* são flexíveis o suficiente para se adaptar a diferentes estratégias de previsão de múltiplos horizontes. As subseções a seguir descrevem a estratégia direta, a recursiva e a de múltiplas saídas.

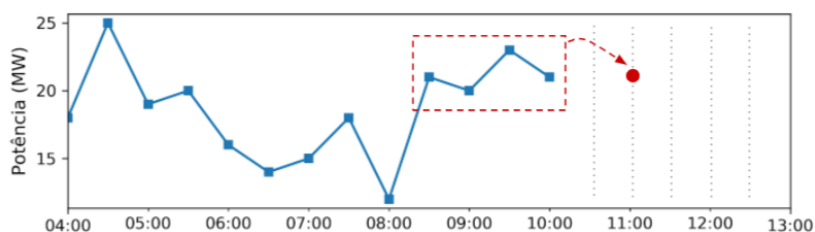
2.3.1 Estratégia Direta

A estratégia direta consiste na elaboração de um modelo diferente para prever cada um dos horizontes. Todos os modelos desenvolvidos são redes neurais de saída única, que recebem como entrada as n observações mais recentes da série e se especializam em prever um dos horizontes de interesse. A Figura 6 ilustra um exemplo com as entradas e saídas das RNAs da estratégia direta quando utilizadas 4 observações da série p para prever o primeiro e o segundo horizonte.

Esta abordagem exige maior custo computacional uma vez que é necessário treinar modelos distintos para cada propósito. Além disso, não há compartilhamento de qualquer resultado entre as previsões. À medida que o horizonte aumenta, mais distantes no tempo se tornam as referências aos dados de entrada. Por outro lado, a estratégia direta proporciona que cada RNA se dedique exclusivamente a minimizar o erro de previsão de seu horizonte alvo. Esta estratégia também possibilita a escolha de hiperparâmetros específicos para cada rede neural construída [9].



(a) Representação da aplicação para o horizonte 1



(b) Representação da aplicação para o horizonte 2

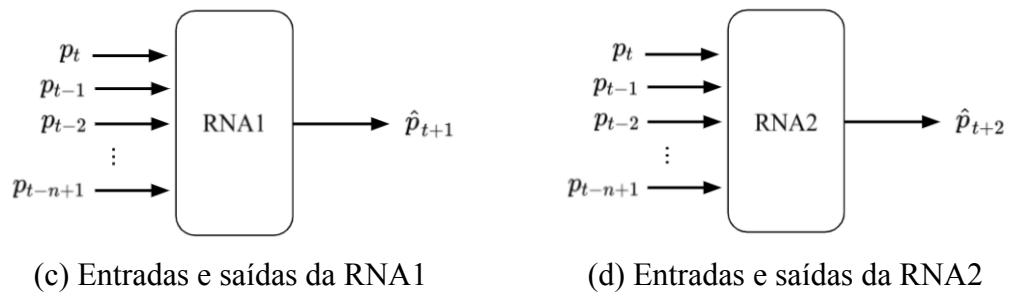
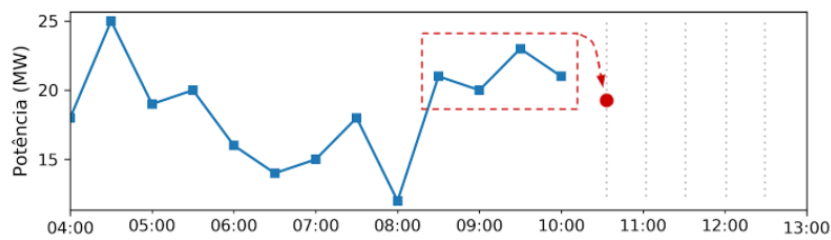


Figura 6. Previsões por meio da estratégia direta

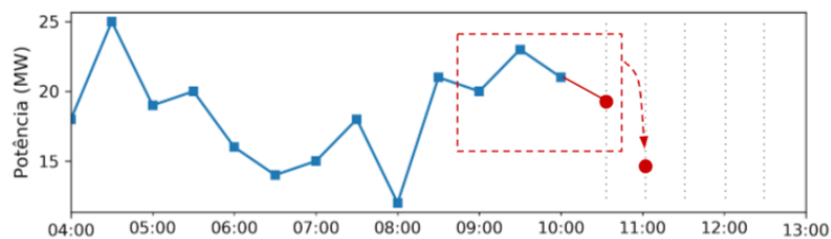
2.3.2 Estratégia Recursiva

A estratégia recursiva parte da ideia de que existe uma rede neural capaz de extrair o processo gerador da série temporal. Ela se baseia na construção de apenas um modelo de saída única especializado na previsão do primeiro horizonte. Durante o treinamento da RNA são apresentados padrões de entrada contendo n observações sucessivas, e padrões de saída contendo o valor da série para 1 horizonte à frente. Já durante a fase de teste, este modelo devidamente treinado se encarrega de prever todos os horizontes de maneira recursiva, onde a previsão do primeiro horizonte é utilizada como entrada para a previsão do segundo, e assim por diante, sempre tomando as n informações mais recentemente obtidas. A Figura 7a e 7b apresentam como se dá o deslizamento da janela de entrada, enquanto 7c ilustra as entradas e a saída do modelo.

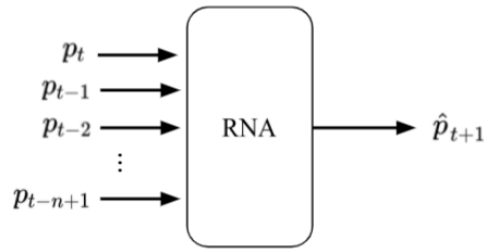
Embora pareça tentador treinar apenas um modelo e gerar uma quantidade indefinida de previsões através dele, a estratégia recursiva tende a propagar o erro cometido ao longo dos horizontes. Por conta disso deve-se dedicar um certo cuidado à escolha dos hiperparâmetros e garantir que o treinamento da RNA aconteça da melhor forma possível. Em contraponto à abordagem direta, a estratégia recursiva sempre faz uso das informações mais recentemente obtidas para prever horizontes mais distantes [9].



(a) Representação da aplicação para o horizonte 1



(b) Representação da aplicação para o horizonte 2



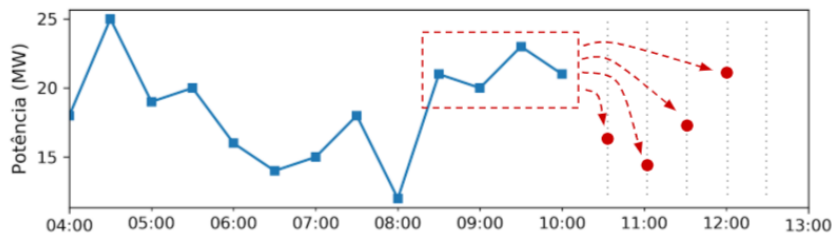
(c) Entradas e saída do modelo

Figura 7. Previsões por meio da estratégia recursiva

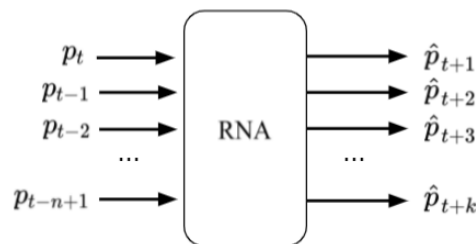
2.3.3 Estratégia de Múltiplas Saídas

A estratégia de múltiplas saídas consiste na criação de um único modelo capaz de se ajustar a todos os horizontes simultaneamente [9]. Redes neurais artificiais MLP e LSTM dão suporte a esta estratégia por meio de uma camada de saída com k unidades de processamento, onde k representa o número de horizontes. A Figura 8 ilustra como são feitas as previsões dos k horizontes quando passadas como entrada as n observações mais recentes da série.

A abordagem de múltiplas saídas oferece a vantagem do treinamento de apenas um modelo assim como na estratégia recursiva, porém sem o viés de propagação do erro. Também vale salientar que a abordagem em questão tenta minimizar o erro total de todas as saídas ao mesmo tempo, e por isso é preciso prestar atenção à possibilidade de que o modelo negligencie a acurácia de alguns horizontes em nome do desempenho global.



(a) Representação da aplicação da estratégia



(b) Entradas e saídas do modelo

Figura 8. Previsões por meio da estratégia de múltiplas saídas

3. Estado da Arte

Já discutimos alguns dos principais conceitos a respeito da previsão de séries temporais. Então agora revisaremos brevemente o estado da arte de sua aplicação em dados de potência eólica.

Métodos estatísticos lineares como os modelos ARMA foram utilizados por Lujano-Rojas et al. em [10] sobre dados em escala horária adquiridos de 5 localizações na Espanha. O modelo ARMA superou a persistência em previsões de até 10 horas à frente.

Palomares-Salas et al. em [11] desenvolveu modelos autorregressivos integrados de médias móveis (ARIMA), comparou-os com redes neurais artificiais e utilizou dados amostrados a cada 10 minutos provenientes de uma estação do tempo em Sevilha, Espanha. Seu estudo reportou que o ARIMA e as RNAs obtiveram resultados similares.

Diversos modelos de aprendizado de máquina (MLP, SVM, kNN, RBF) foram experimentados por Liu et al. em [12] utilizando 7 bases de dados com informações de umidade, temperatura, velocidade e potência dos ventos. Diferentes MLPs com até 4 camadas ocultas foram testadas, porém não houve melhora significativa associada à profundidade das redes. Adaptações de *Support Vector Machines* (SVM) para regressão alcançaram melhor RMSE geral, mas MLPs obtiveram melhor acurácia em horizontes distantes.

Abordagens híbridas ARIMA-MLP e ARIMA-SVM foram estudadas por Shi et al. em [13]. A ideia proposta foi prever as componentes linear e não linear dos dados separadamente. Sua conclusão foi a de que a abordagem híbrida, mesmo sendo mais complexa, ofereceu pouca melhora em relação aos modelos isolados.

Li et al. comparou várias arquiteturas MLP utilizando as métricas MAE, RMSE e MAPE. Através do estudo, chegou-se à conclusão de que não havia superioridade entre arquiteturas, e que o desempenho dependia dos dados. Foram alcançadas melhorias de até 20% quando escolhidos os parâmetros que melhor se ajustavam aos dados [14].

Em [15], Xiaoyun et al. utilizaram dados de um único parque eólico com capacidade instalada de 120 megawatts (MW). O conjunto contou com medidas de potência, densidade do ar, pressão, temperatura, velocidade e direção do vento coletadas em intervalos de 5 minutos. Foi feito uso da técnica de análise de componentes principais (PCA) para extrair informações relevantes antes do treinamento dos modelos e apenas o horizonte de 24 horas à frente foi estabelecido como alvo. A abordagem PCA-LSTM mostrou maior acurácia quando comparada a LSTM e SVM para regressão.

J. Manero em sua tese de doutorado [16] comparou o uso de redes neurais convolucionais, recorrentes e Multilayer Perceptron aplicadas ao maior repositório público de dados de vento com informações vindas de 126692 localizações nos EUA. Por meio de dados espaçados em 1h foram realizadas previsões para horizontes de até 12 horas à frente usando a estratégia de múltiplas saídas. Redes do tipo MLP apresentaram melhor acurácia para a grande maioria das localizações, levantando questionamentos acerca do uso das redes recorrentes LSTM em aplicações de séries temporais.

Como vemos, não existe um modelo generalista. O desempenho das redes neurais artificiais parecem depender das características geográficas, da frequência de amostragem da série temporal, dos horizontes de previsão e das técnicas de pré-processamento adotadas.

4. Experimento

Os dados deste estudo seguiram o fluxo ilustrado pela Figura 9. Inicialmente, as séries brutas de potência eólica passaram por uma etapa de pré-processamento onde: foram integradas para uma nova escala temporal; tiveram suas lacunas não valoradas preenchidas através de interpolação; foram normalizadas para uma amplitude conveniente para os modelos; e por fim, foram transformadas em padrões supervisionados de entrada e saída. Depois, os padrões supervisionados serviram para treinar e testar as várias combinações de redes neurais artificiais, estratégias e hiperparâmetros. Os resultados finais passaram então por etapas de pós-processamento e avaliação a fim de estabelecer comparações de desempenho.

Dentre as ferramentas necessárias para a implementação deste trabalho podem ser mencionadas as bibliotecas científicas disponíveis na linguagem *Python*, como: *numpy*, para manipulação numérica e de *arrays*; *pandas*, para manipulação de dados estruturados; e *tensorflow*, *keras* e *scikit-learn*, para a criação das redes neurais artificiais.

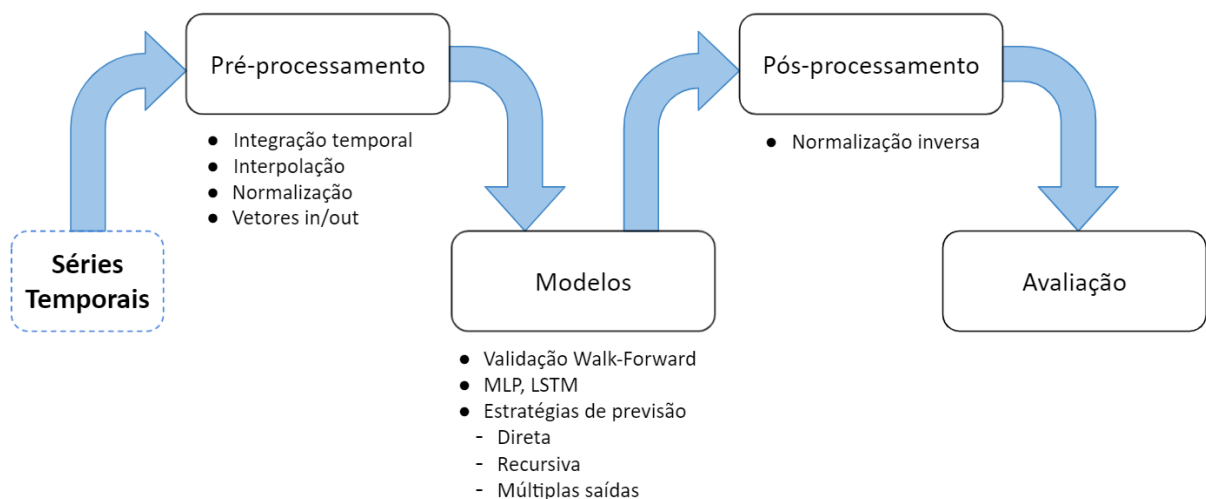


Figura 9. Metodologia do experimento

4.1 Dados

Os dados originais de potência foram derivados de uma base pública disponibilizada pelo Operador do Mercado de Eletricidade Australiano (AEMO) [17]. O conjunto de dados como um todo é formado por 51 séries temporais de potência eólica, uma para cada parque eólico, com amostras coletadas a cada 5 minutos durante o ano de 2018.

Com o intuito de fortalecer os resultados deste estudo, foram escolhidas séries de 3 parques suficientemente distintos. Os critérios de seleção foram a quantidade de amostras não valoradas, a localização geográfica e a capacidade instalada em megawatts. Como apresentado nas Figuras 10 e 11, as séries escolhidas se referem às centrais: Star Hill (STARHLWF) com capacidade de 34 MW, Boco Rock (BOCORWF1) com até 113 MW; e Macarthur (MACARTH1) com 420 MW.

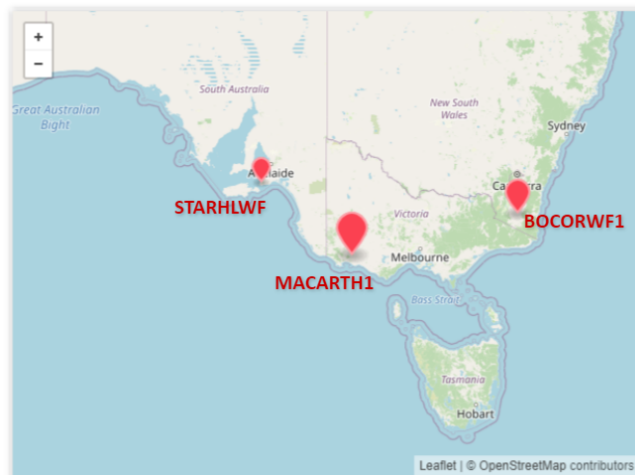


Figura 10. Localização geográfica dos parques eólicos no sul da Austrália

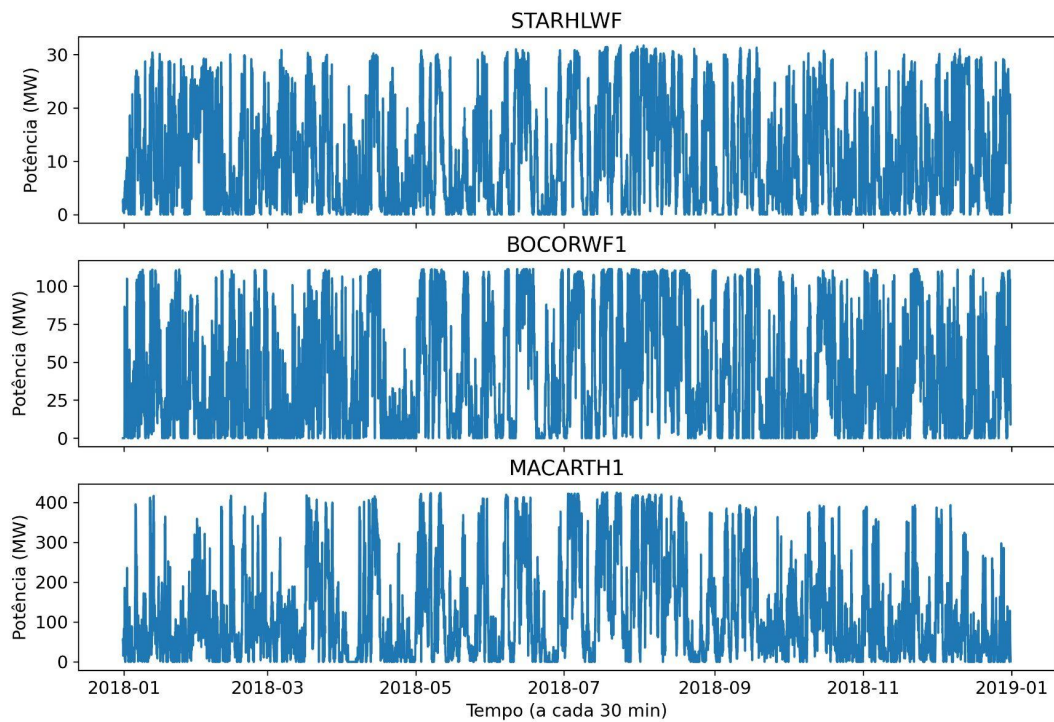


Figura 11. Séries temporais de potência referentes a cada parque eólico

Tabela 1 - Estatísticas das séries de potência em megawatts

	STARHLWF	BOCORWF1	MACARTH1
count	17519.00	17519.00	17519.00
mean	11.04	41.93	122.28
std	9.74	39.00	123.23
min	0.00	0.00	0.00
25%	2.10	5.63	23.55
50%	8.31	28.57	75.90
75%	19.39	78.73	190.54
max	31.76	111.35	424.78

Exceto por *count*, todas as estatísticas dispostas na Tabela 1 são dadas em megawatts (MW) e significam:

- ***count***: a contagem total de amostras que compõem as séries temporais;
- ***mean***: a média aritmética;
- ***std (standard deviation)***: o desvio padrão;
- ***min (minimum)***: a amostra de menor amplitude;
- **25%**: o 1º quartil, ou seja, após ordenar os dados da série em ordem crescente, esta é a amostra que separa 1 quarto dos dados para a esquerda e 3 quartos para a direita;
- **50%**: a mediana, ou seja, após ordenar os dados da série em ordem crescente, esta é a amostra que separa metade dos dados para a esquerda e metade para a direita;
- **75%**: o 3º quartil, ou seja, após ordenar os dados da série em ordem crescente, esta é a amostra que separa 3 quartos dos dados para a esquerda e 1 quarto para a direita;
- ***max (maximum)***: a amostra de maior amplitude.

4.2 Pré-processamento

A primeira transformação realizada na etapa de pré-processamento foi a integração dos dados da escala de 5 minutos para a escala 30 minutos. Muito porque a escala de 5 minutos não exibiu mudanças significativas no comportamento dos ventos quando analisados ao menos 12 horizontes (até 1 hora à frente).

A segunda modificação foi a interpolação de dados ausentes. Das 3 séries escolhidas, apenas a STARHLWF precisou deste tratamento, visto que apresentava 34 ausências pontuais ao longo dos 12 meses. Este procedimento foi necessário pois as implementações das redes neurais em *Python* não dão suporte a dados do tipo *not-a-number* (NaN) durante o treinamento.

A terceira intervenção se deu com a aplicação de uma normalização de mínimo e máximo caracterizada pela Equação 4 abaixo:

$$X'_i = \frac{X_i - \min(X)}{\max(X) - \min(X)} \quad (4)$$

onde X'_i é o valor normalizado da amostra original X_i , $\min(X)$ é a amplitude mínima atingida pela série X , e $\max(X)$ é a amplitude máxima. A execução desta normalização levou os dados de sua amplitude original para uma amplitude limitada entre 0 e 1.

Finalmente, em um último passo, a série foi transformada em padrões supervisionados de treinamento e teste de acordo com o tamanho das entradas e saídas próprias de cada modelo. O que será discutido na Seção 4.3 a seguir.

4.3 Modelos

Considerando que o regime dos ventos pode sofrer alterações ao longo do ano, é razoável assumir que a implantação de algum tipo de validação cruzada trará robustez aos

resultados. A validação *walk-forward* promove a criação de *splits* (treino-teste) que se deslocam rumo ao futuro até o fim da série temporal. Neste trabalho, cada *split* compreendeu um período de 2 meses consecutivos, sendo o mais antigo usado para treinar os modelos, e o mais recente para testá-los. Devido a limitações de processamento, o conjunto de treinamento não foi expandido mesmo havendo dados disponíveis conforme o avanço dos *splits*.

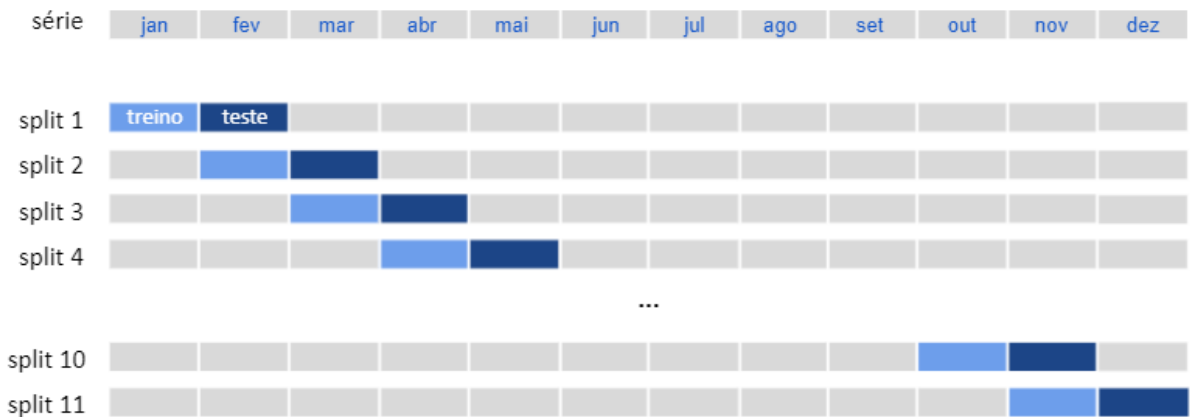


Figura 12. *Splits* da validação *walk-forward*

Para cada instante dentro de um mês de teste foram previstos os 12 primeiros horizontes, o que equivale a até 6 horas à frente devido a escala de 30 minutos adotada. As 24 configurações de modelagem construídas são fruto da combinação de:

- 2 redes neurais artificiais: *Multilayer Perceptron* ou *Long Short-Term Memory*;
- 3 estratégias de previsão múltipla: direta, recursiva ou de múltiplas saídas;
- 2 tamanhos do vetor de entrada (*input_size*): 4 ou 12 variáveis de entrada;
- 2 estruturas de camadas ocultas (*hidden_layers*): 1 camada com 12 neurônios ou 2 camadas com 24 e 12 neurônios.

O parâmetro que especifica o tamanho do vetor de entrada foi escolhido como variável de estudo pois não se sabe quantas observações defasadas no tempo carregam relevância para a previsão. O mesmo vale para a definição das camadas ocultas, visto que não há uma regra geral que quantifique, a priori, o impacto que estas escolhas podem causar no desempenho das redes neurais.

A persistência foi escolhida como modelo de referência para fins de comparação. Seu funcionamento parte do pressuposto de que a potência observada no instante atual p_t persistirá até o horizonte k de interesse. Gerando uma previsão da forma $\hat{p}_{t+k} = p_t$.



Figura 13. Modelo da persistência para o horizonte k

5. Resultados e Discussão

Os resultados tiveram como base a métrica da raiz do erro quadrático médio (RMSE) dada pela Equação 5 abaixo:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (\hat{X}_i - X_i)^2} \quad (5)$$

onde N é a quantidade total de amostras da série temporal, X_i é o valor verdadeiro observado para a i -ésima amostra e \hat{X}_i é a respectiva previsão obtida. O RMSE é capaz de ponderar erros de acordo com a magnitude em que ocorrem, além de prover a mesma unidade de medida dos dados originais, ou seja, megawatts (MW). Foi a partir desta métrica de erro que desenvolveram-se duas análises:

1. A análise I teve foco no desempenho geral dos modelos. Assim, para cada combinação de {rede, estratégia, *input_size*, *hidden_layers*} foi calculada a média geral do RMSE ao longo dos horizontes e splits. O intuito desta primeira análise foi saber se há algum indício de superioridade entre as redes neurais artificiais MLP e LSTM, ou entre as estratégias direta, recursiva e de múltiplas saídas. Além disso, buscou-se verificar qual parametrização de *input_size* e *hidden_layers* proporcionaram melhores resultados.
2. A análise II teve o propósito de avaliar o desempenho das redes e estratégias quanto à previsão de cada horizonte. Para isso, para cada combinação de {rede, estratégia, *input_size*, *hidden_layers*, horizonte} foi calculada a média do RMSE dos splits.

Nos gráficos de ambas as análises, conjuntos colunares com 4 pontos de mesma cor significam as 4 possíveis combinações de *input_size* e *hidden_layers*. Não há significado numérico no eixo das abscissas, apenas separação entre categorias. As seções a seguir apresentam os resultados para a central eólica Star Hill, Boco Rock e Macarthur, nesta ordem.

5.1 Central Star Hill

Para o parque eólico Star Hill, como mostra a Figura 14, a média geral do RMSE se mostrou menor para a MLP do que para LSTM quando comparadas somente as estratégias direta e de múltiplas saídas. A estratégia recursiva se mostrou pouco adequada junto a rede MLP, tendo desempenho pior que a persistência para 3 das 4 combinações de *input_size* e *hidden_layers*. Por outro lado, todas as configurações que utilizaram a rede neural LSTM superaram a persistência.

A Tabela 2 apresenta o RMSE médio geral da melhor combinação de *input_size* e *hidden_layers* para cada par de {rede, estratégia}. Todas conseguiram bater a persistência. Também é possível observar que 4 variáveis de entrada aliadas a duas camadas ocultas com 24 e 12 neurônios são o destaque da rede MLP, conferindo a esta os 2 primeiros lugares.

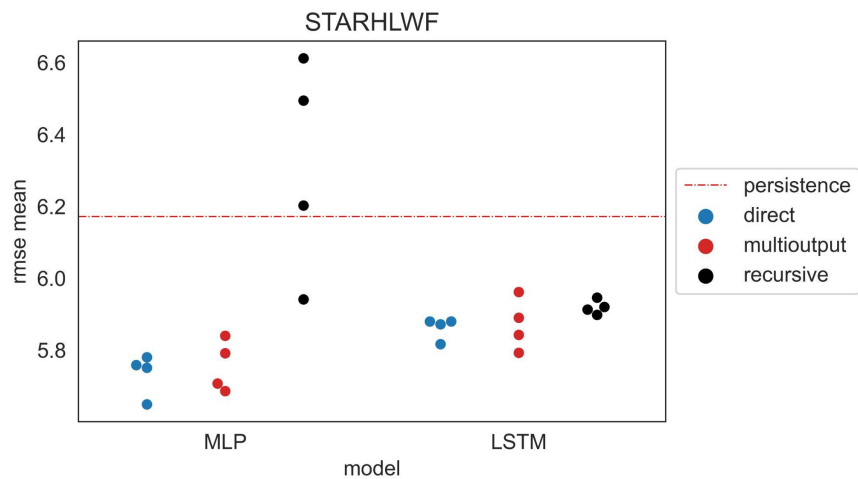
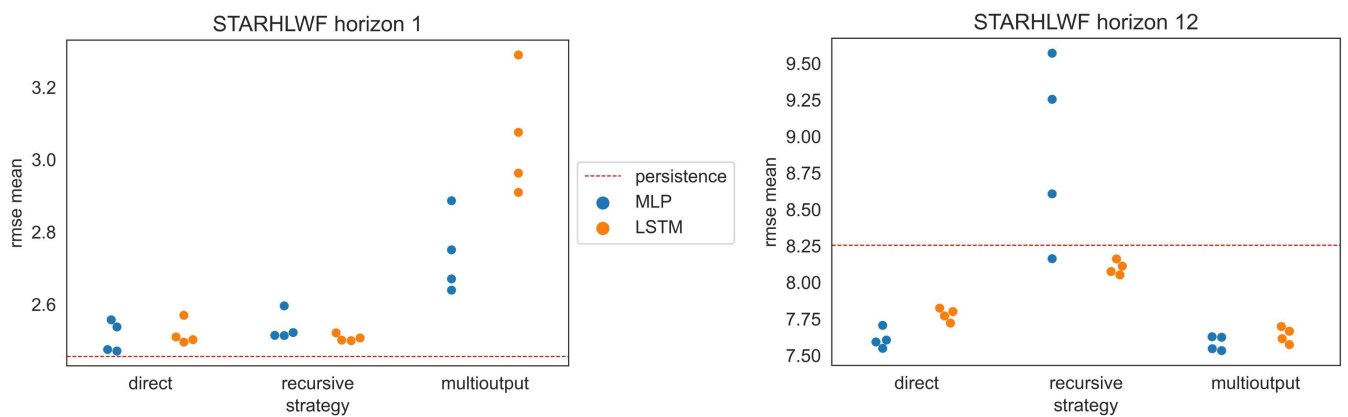


Figura 14. Média geral do RMSE (STARHLWF)

Tabela 2 - Melhores parâmetros para cada par de rede e estratégia (STARHLWF)

	model	strategy	input_size	hidden_layers	rmse mean	rmse std
0	MLP	direct	4	[24, 12]	5.649331	0.514808
1	MLP	multioutput	4	[24, 12]	5.686135	0.501048
2	LSTM	multioutput	12	[12]	5.792648	0.522791
3	LSTM	direct	12	[12]	5.816735	0.511214
4	LSTM	recursive	4	[24, 12]	5.898261	0.612975
5	MLP	recursive	4	[12]	5.941343	0.616552
6	persistence	NaN	NaN	NaN	6.172146	0.776148



(a) Horizonte 1

(b) Horizonte 12

Figura 15. Média do RMSE para os horizontes 1 e 12 (STARHLWF)

Ao longo dos horizontes foi observado que a estratégia de múltiplas saídas não ofereceu bom desempenho em previsões de curtíssimo prazo. Em contrapartida, a estratégia recursiva teve seu RMSE médio degradado à medida em que se aumentou o horizonte de previsão. Os dados de potência de STARHLWF mostraram ser bem persistentes, de forma que as configurações testadas dificilmente superaram o modelo de referência nos horizontes 1 e 2. Ver Figura 15 e Figura 20 no Apêndice para mais detalhes.

5.2 Central Boco Rock

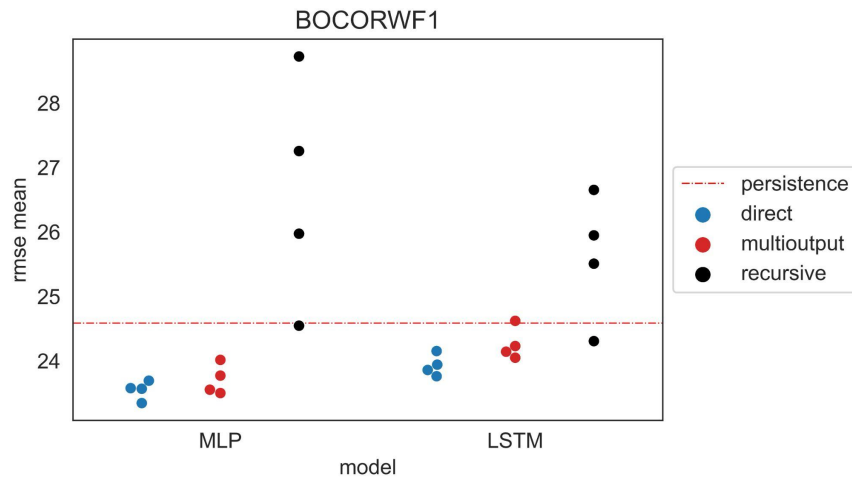


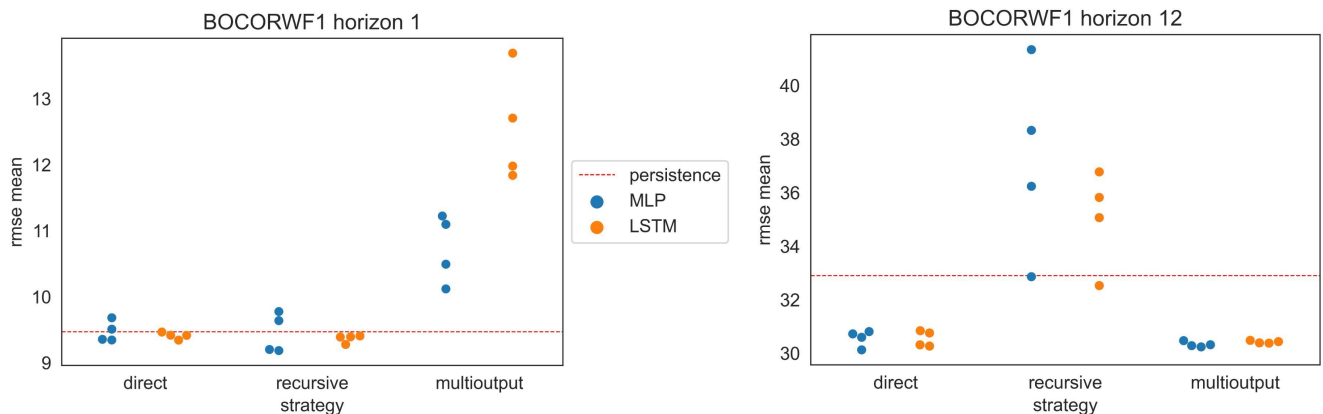
Figura 16. Média geral do RMSE (BOCORWF1)

Para o parque Boco Rock, a média geral do RMSE das estratégias direta e de múltiplas saídas superaram o modelo de referência. A estratégia recursiva mostrou pior desempenho independente da RNA utilizada, tendo apenas a parametrização de $input_size = 4$ e $hidden_layers = [12]$ atuando melhor que a persistência. Aqui, assim como para Star Hill, todos os pares de {rede, estratégia} conseguiram superar a persistência ao menos com seus melhores parâmetros. Inclusive, as duas primeiras colocações também foram ocupadas por redes neurais MLP com 4 variáveis de entrada e 2 camadas ocultas com 24 e 12 neurônios. Ver Figura 16 e Tabela 3.

Tabela 3 - Melhores parâmetros para cada par de rede e estratégia (BOCORWF1)

	model	strategy	input_size	hidden_layers	rmse mean	rmse std
0	MLP	direct	4	[24, 12]	23.343230	2.391541
1	MLP	multioutput	4	[24, 12]	23.498747	2.342645
2	LSTM	direct	12	[12]	23.760306	2.467463
3	LSTM	multioutput	12	[12]	24.047824	2.635030
4	LSTM	recursive	4	[12]	24.305342	2.470719
5	MLP	recursive	4	[12]	24.545043	2.820402
6	persistence	NaN	NaN	NaN	24.586304	3.148453

Novamente, o desempenho da estratégia de múltiplas saídas mostrou grande dispersão e alto RMSE médio nos primeiros horizontes. As previsões de múltiplas saídas melhoraram à medida que se avançou o horizonte de previsão, quase sempre com a rede LSTM se mostrando pior que a MLP. A superioridade da rede MLP também pode ser observada através da estratégia direta. Já a estratégia recursiva, apresentou RMSE médio pior que a persistência em todos os horizontes, exceto o primeiro, variando muito com os parâmetros *input_size* e *hidden_layers*. Ver Figura 17 e Figura 21 no Apêndice para mais detalhes.



(a) Horizonte 1

(b) Horizonte 12

Figura 17. Média do RMSE para os horizontes 1 e 12 (BOCORWF1)

5.3 Central Macarthur

Para o parque eólico Macarthur, como mostra a Figura 18, a estratégia direta aliada a rede MLP se mostrou a abordagem mais recomendada, pois superou a persistência independente da parametrização. Sendo o completo oposto da estratégia recursiva. As redes LSTM, no geral, tiveram bastante dificuldade em superar o modelo de referência, mesmo quando aliada às estratégias direta ou de múltiplas saídas. Entretanto, foi neste parque que pela primeira vez a estratégia recursiva alcançou a melhor média geral do RMSE para redes LSTM.

A Tabela 4 revela mais uma vez que as parametrizações com melhor RMSE geral são de MLPs com 4 variáveis de entrada e 2 camadas ocultas com 24 e 12 neurônios. Sendo uma conclusão unânime entre os 3 parques eólicos analisados neste estudo. Ver também a Tabela 7 no Apêndice.

Através da Figura 19, mais uma vez foi constatado o péssimo desempenho da estratégia de múltiplas saídas em realizar previsões a curto prazo. Enquanto que a estratégia recursiva tem seu desempenho degradado à medida que se realiza previsões mais distantes no futuro. Ver também a Figura 22 no Apêndice.

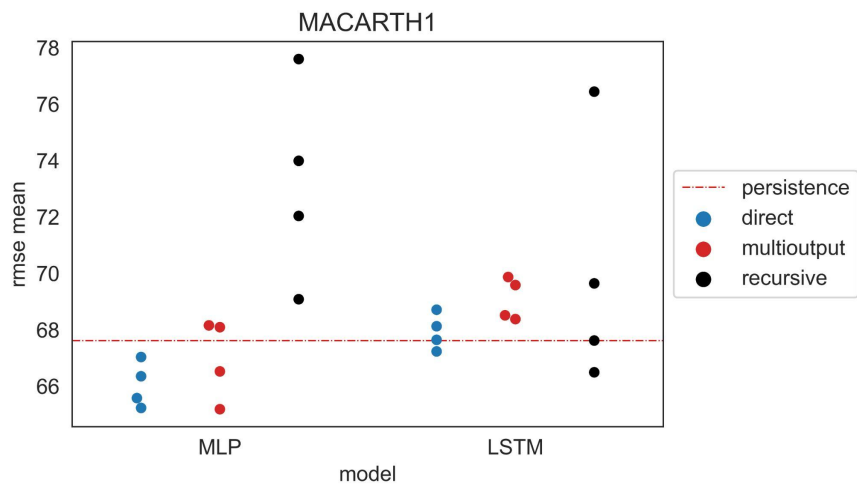
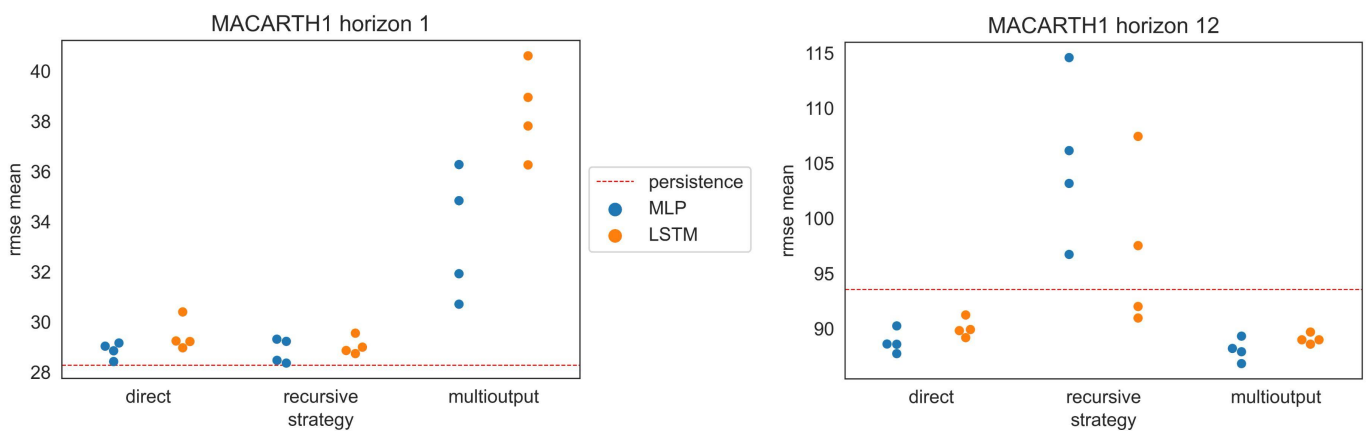


Figura 18. Média geral do RMSE (MACARTH1)

Tabela 4 - Melhores parâmetros para cada par de rede e estratégia (MACARTH1)

	model	strategy	input_size	hidden_layers	rmse mean	rmse std
0	MLP	multioutput	4	[24, 12]	65.184855	5.029576
1	MLP	direct	4	[24, 12]	65.230224	4.811660
2	LSTM	recursive	4	[12]	66.493975	4.395350
3	LSTM	direct	4	[12]	67.233842	5.487319
4	persistence	NaN	NaN	NaN	67.622216	6.111518
5	LSTM	multioutput	12	[12]	68.380338	6.796884
6	MLP	recursive	4	[12]	69.085383	7.336785



(a) Horizonte 1

(b) Horizonte 12

Figura 19. Média do RMSE para os horizontes 1 e 12 (MACARTH1)

6. Conclusão

Ao longo deste estudo foram introduzidos conceitos chave para a previsão de séries temporais, bem como a importância desta atividade para a integração da energia eólica à matriz energética.

Dados históricos de potência de três parques eólicos foram selecionados como entrada para uma diversidade de modelos de redes neurais artificiais aliados a estratégias de previsão de múltiplos horizontes. Algumas particularidades já documentadas das estratégias direta, recursiva e de múltiplas saídas ficaram evidentes também nesta aplicação em séries temporais de potência eólica.

Independentemente do parque analisado, a estratégia de múltiplas saídas se mostrou pior que a persistência nos primeiros horizontes, melhorando gradativamente até superá-la por volta do horizonte 6. A estratégia recursiva somente pareceu adequada para a previsão do primeiro horizonte, tendo suas previsões rapidamente deterioradas já no horizonte 2 devido a propagação dos erros. Por fim, a estratégia direta apresentou a melhor média geral do RMSE principalmente quando aliada a redes neurais *Multilayer Perceptron*.

Conforme o exposto, conclui-se que em casos onde há poder computacional muito limitado, deve-se preferir o uso da estratégia de múltiplas saídas, por ter como base um único modelo. Em casos de maior capacidade de processamento, deve-se preferir o uso da estratégia direta. Adicionalmente, em ambos os casos, deve-se preferir aliar tais estratégias a redes neurais MLP, pois embora as *Long Short-Term Memory* tenham suporte a dados sequenciais, estas não obtiveram melhor performance.

6.1 Trabalhos Futuros

Frente aos resultados apresentados e às conclusões constatadas, existem alguns pontos que ainda podem ser melhor explorados. Durante o desenvolvimento da estratégia recursiva não houveram cuidados extras quanto ao refino dos hiperparâmetros. Adicionalmente, sabe-se que esta estratégia é extremamente dependente do desempenho de uma única RNA. Neste trabalho foram testadas somente 2 variações de *input_size* e 2 variações de *hidden_layers*, sem justificativas para estas escolhas. Uma ampla avaliação empírica de hiperparâmetros poderia ser conduzida para validar os resultados aqui reportados.

O péssimo desempenho da estratégia de múltiplas saídas em previsões a curtíssimo prazo pode estar associado à quantidade de saídas as quais a rede tenta se ajustar simultaneamente. Uma abordagem que poderia ser investigada trata da construção de dois modelos distintos de múltiplas saídas, sendo: um primeiro para prever do horizonte 1 ao 6; e um segundo para prever do horizonte 7 ao 12. Algo como uma estratégia intermediária entre a direta e a de múltiplas saídas.

Por fim, seria interessante replicar este estudo utilizando outras séries de potência eólica, ou mesmo expandindo o conjunto de treinamento ao longo da validação *walk-forward*.

7. Referências

- [1] A. M. Foley, P. G. Leahy, A. Marvuglia, E. J. McKeogh. “Current Methods and Advances in Forecasting of Wind Power Generation”, *Renewable Energy*, vol 37, pp. 1-8, 2012.
- [2] N. A. Treiber, S. Spath, J. Heinermann, L. von Bremen, O. Kramer. “Comparison of Numerical Models and Statistical Learning for Wind Speed Prediction”, *ESANN*, 2015.
- [3] L. Landberg, S. J. Watson. “Short-Term Prediction of Local Wind Conditions”, *Boundary-Layer Meteorology*, vol 70, pp. 171–95, 1994.
- [4] S. B. Taieb, A. Sorjanna, G. Bontempi. “Multiple-output Modeling for Multi-step-ahead Time Series Forecasting”, *Neurocomputing*, vol 73, pp. 1950-1957, 2010.
- [5] E. Erdem, J. Shi. “ARMA Based Approaches for Forecasting the Tuple of Wind Speed and Direction. *Applied Energy*”, *Applied Energy*, vol 88, pp. 1405-1414, 2011.
- [6] M. Gardner, S. Dorling. “Artificial Neural Networks (the Multilayer Perceptron) - a Review of Applications in the Atmospheric Sciences”, *Atmospheric Environment*, vol 32, pp. 2627–2636, 1998.
- [7] S. Hochreiter, J. Schmidhuber. “Long Short-Term Memory”, *Neural Computation*, vol 9, pp. 1735-1780, 1997.
- [8] Understanding LSTM Networks. Disponível em: [Understanding LSTM Networks -- colah's blog](#)
- [9] S. B. Taieb, G. Bontempi, A. F. Atiya, A. Sorjanna. “A Review and Comparison of Strategies for Multi-step Ahead Time Series Forecasting Based on the NN5 Forecast Competition”, *Expert Systems with Applications*, vol 39, pp. 7067-7083, 2012.
- [10] J. M. Lujano-Rojas, J. L. Bernal-Agustín, R. Dufo-López, J. A. Domínguez-Navarro. “Forecast of Hourly Average Wind Speed Using ARMA Model with Discrete Probability Transformation”, *Electrical Engineering and Control*, vol 2, pp. 1003-1010, 2011.
- [11] J. C. Palomares-Salas, J. J. G. de la Rosa, J. G. Ramiro, J. Melgar, A. Aguera, A. Moreno. “Arima vs. Neural Networks for Wind Speed Forecasting”, *IEEE International Conference on Computational Intelligence for Measurement Systems and Applications*, pp. 129-133, 2009.
- [12] Y. Liu, H. Zhang. “An Empirical Study on Machine Learning Models for Wind Power Predictions”, *IEEE Conf. on Machine Learning and Applications*, pp. 758-763, 2016.
- [13] J. Shi, J. Guo, S. Zheng. “Evaluation of Hybrid Forecasting Approaches for Wind Speed and Power Generation Time Series”, *Renewable and Sustainable Energy Reviews*, vol 16, pp. 3471-3480, 2012.
- [14] G. Li, J. Shi. “On Comparing Three Artificial Neural Networks for Wind Speed Forecasting”, *Applied Energy*, vol 87, pp. 2313-2320, 2010.

- [15] Q. Xiaoyun, K. Xiaoning, Z. Chao, J. Shuai, M. Xiuda. “Short-Term Prediction of Wind Power Based on Deep Long Short-Term Memory”, IEEE PES Asia-Pacific Power and Energy Engineering Conference, pp. 1148-1152, 2016.
- [16] J. Manero. “Deep Learning Architectures Applied to Wind Time Series Multi-Step Forecasting”, Universitat Politècnica de Catalunya, 2020.
- [17] AEMO Energy Generation Data. Disponível em: <https://anero.id/energy/data>

Apêndice

A - Especificações Técnicas Adicionais

- ***input_size***: 4 ou 12
Quantidade de instantes passados que compõem o vetor de entrada. São também chamadas de variáveis regressoras, ou tamanho da janela.
- ***hidden_layers***: 1 com 12 neurônios, ou 2 com 24 e 12 neurônios
Quantidade de camadas ocultas e número de unidades de processamento por camada oculta.
- ***activation_function***: tanh (*hidden layers*), linear (*output layer*)
Funções de ativação para as unidades da camada oculta e da camada de saída.
- ***optimizer***: adam
Otimizador utilizado para encontrar um mínimo local (ou global) da superfície multidimensional do erro durante o treinamento.
- ***loss_function***: mse
Função de perda que quantifica o erro cometido pela rede durante o treinamento. O algoritmo otimizador deve encontrar um ponto de mínimo para esta função.
- ***validation_split***: 0.1
Porcentagem do conjunto de treinamento que será utilizada como validação. O conjunto de validação serve para que se monitore o desempenho da rede durante o treinamento, sendo apresentado sempre após a conclusão de uma época. Se o erro no conjunto de treinamento diminuir, mas o erro no conjunto de validação aumentar, então a rede está se especializando no conjunto de treinamento e perdendo sua capacidade de generalização (*overfitting*).
- ***callbacks***: EarlyStopping(val_loss, patience=5, restore_best_weights=True)
Permite a parada precoce do treinamento em caso de detecção de *overfitting*. O *EarlyStopping* irá monitorar o erro no conjunto de validação (*val_loss*). Se este erro aumentar por 5 épocas seguidas, o treinamento é interrompido e os pesos são restaurados para o momento anterior ao *overfitting*.
- ***max_epochs***: 200
Número de vezes que o conjunto de treinamento será apresentado à rede. Um número insuficiente de épocas pode acarretar na incapacidade de extração de informação dos dados (*underfitting*).

B - Central Star Hill

A Tabela 5 abaixo apresenta a média geral do RMSE ao longo dos horizontes e splits. A partir desta tabela foram ilustrados todos os pontos da Figura 14. Também foi a partir dela que se obtiveram os dados de melhor parametrização exibidos na Tabela 2.

Tabela 5 - Média geral do RMSE (STARHLWF)

	model	strategy	input_size	hidden_layers	rmse mean	rmse std
0	MLP	direct	4	[24, 12]	5.649331	0.514808
1	MLP	multioutput	4	[24, 12]	5.686135	0.501048
2	MLP	multioutput	12	[24, 12]	5.707191	0.552157
3	MLP	direct	12	[24, 12]	5.751079	0.539615
4	MLP	direct	4	[12]	5.758651	0.539389
5	MLP	direct	12	[12]	5.780318	0.533635
6	MLP	multioutput	12	[12]	5.791463	0.569976
7	LSTM	multioutput	12	[12]	5.792648	0.522791
8	LSTM	direct	12	[12]	5.816735	0.511214
9	MLP	multioutput	4	[12]	5.839950	0.537181
10	LSTM	multioutput	4	[12]	5.842558	0.613606
11	LSTM	direct	12	[24, 12]	5.872039	0.514548
12	LSTM	direct	4	[24, 12]	5.879847	0.539240
13	LSTM	direct	4	[12]	5.879927	0.560085
14	LSTM	multioutput	12	[24, 12]	5.890146	0.579387
15	LSTM	recursive	4	[24, 12]	5.898261	0.612975
16	LSTM	recursive	12	[24, 12]	5.913033	0.590222
17	LSTM	recursive	4	[12]	5.920260	0.658334
18	MLP	recursive	4	[12]	5.941343	0.616552
19	LSTM	recursive	12	[12]	5.946171	0.452948
20	LSTM	multioutput	4	[24, 12]	5.961731	0.601403
21	persistence	NaN	NaN	NaN	6.172146	0.776148
22	MLP	recursive	12	[24, 12]	6.202475	0.833986
23	MLP	recursive	12	[12]	6.494411	1.064871
24	MLP	recursive	4	[24, 12]	6.612019	0.719206

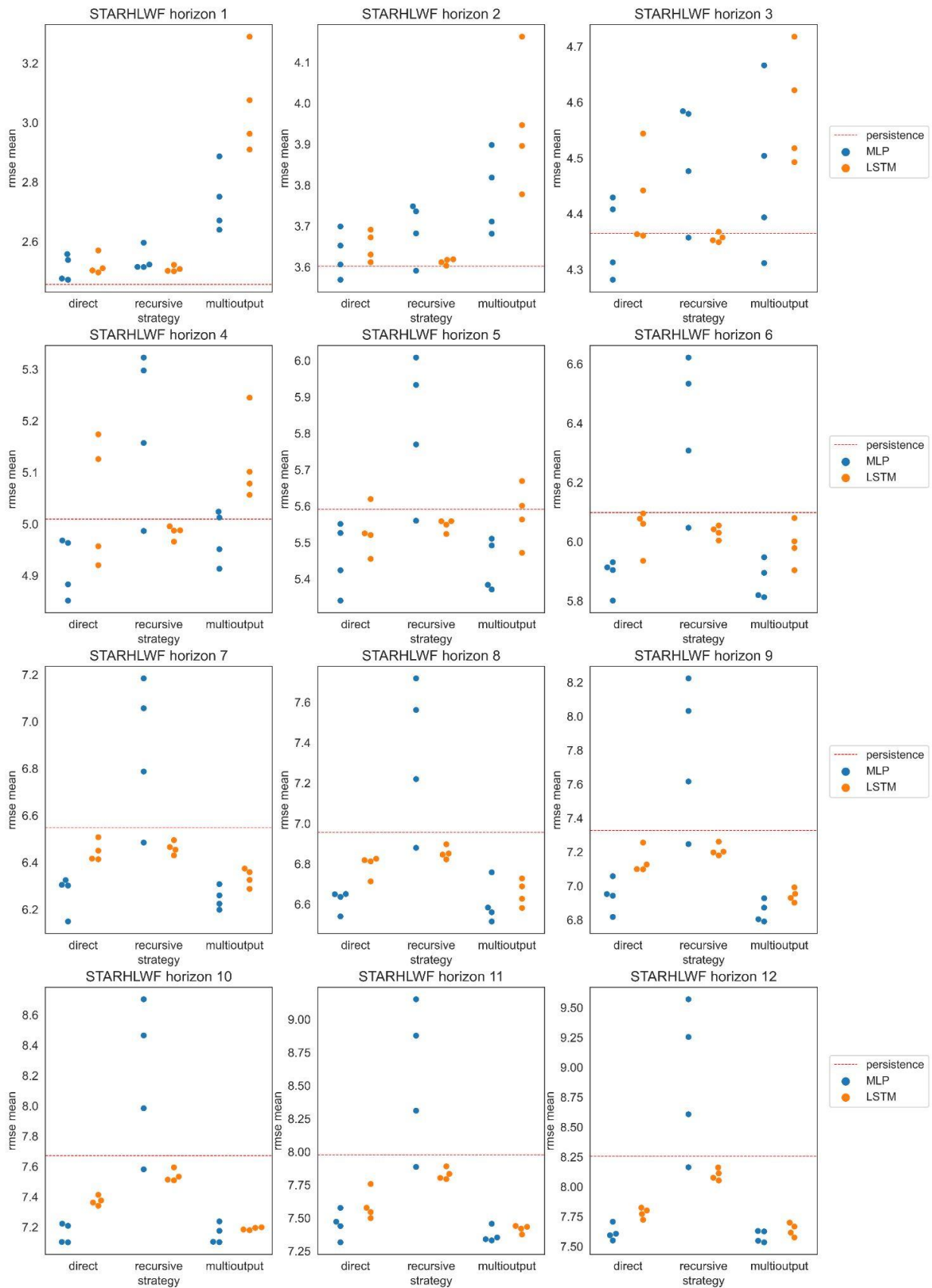


Figura 20. Média do RMSE para cada horizonte (STARHLWF)

C - Central Boco Rock

A Tabela 6 abaixo apresenta a média geral do RMSE ao longo dos horizontes e splits. A partir desta tabela foram ilustrados todos os pontos da Figura 16. Também foi a partir dela que se obtiveram os dados de melhor parametrização exibidos na Tabela 3.

Tabela 6 - Média geral do RMSE (BOCORWF1)

	model	strategy	input_size	hidden_layers	rmse mean	rmse std
0	MLP	direct	4	[24, 12]	23.343230	2.391541
1	MLP	multioutput	4	[24, 12]	23.498747	2.342645
2	MLP	multioutput	12	[24, 12]	23.551217	2.415426
3	MLP	direct	12	[24, 12]	23.565881	2.403832
4	MLP	direct	4	[12]	23.575382	2.324721
5	MLP	direct	12	[12]	23.691422	2.439270
6	LSTM	direct	12	[12]	23.760306	2.467463
7	MLP	multioutput	4	[12]	23.769939	2.496087
8	LSTM	direct	4	[12]	23.857367	2.436881
9	LSTM	direct	12	[24, 12]	23.939870	2.397582
10	MLP	multioutput	12	[12]	24.013741	2.440266
11	LSTM	multioutput	12	[12]	24.047824	2.635030
12	LSTM	multioutput	4	[12]	24.141041	2.508065
13	LSTM	direct	4	[24, 12]	24.151080	2.534778
14	LSTM	multioutput	12	[24, 12]	24.228767	2.800223
15	LSTM	recursive	4	[12]	24.305342	2.470719
16	MLP	recursive	4	[12]	24.545043	2.820402
17	persistence	NaN	NaN	NaN	24.586304	3.148453
18	LSTM	multioutput	4	[24, 12]	24.619710	2.897083
19	LSTM	recursive	12	[12]	25.507416	3.015197
20	LSTM	recursive	4	[24, 12]	25.947335	2.946402
21	MLP	recursive	4	[24, 12]	25.972746	3.954795
22	LSTM	recursive	12	[24, 12]	26.652728	5.841482
23	MLP	recursive	12	[12]	27.256181	3.150825
24	MLP	recursive	12	[24, 12]	28.724035	4.615786

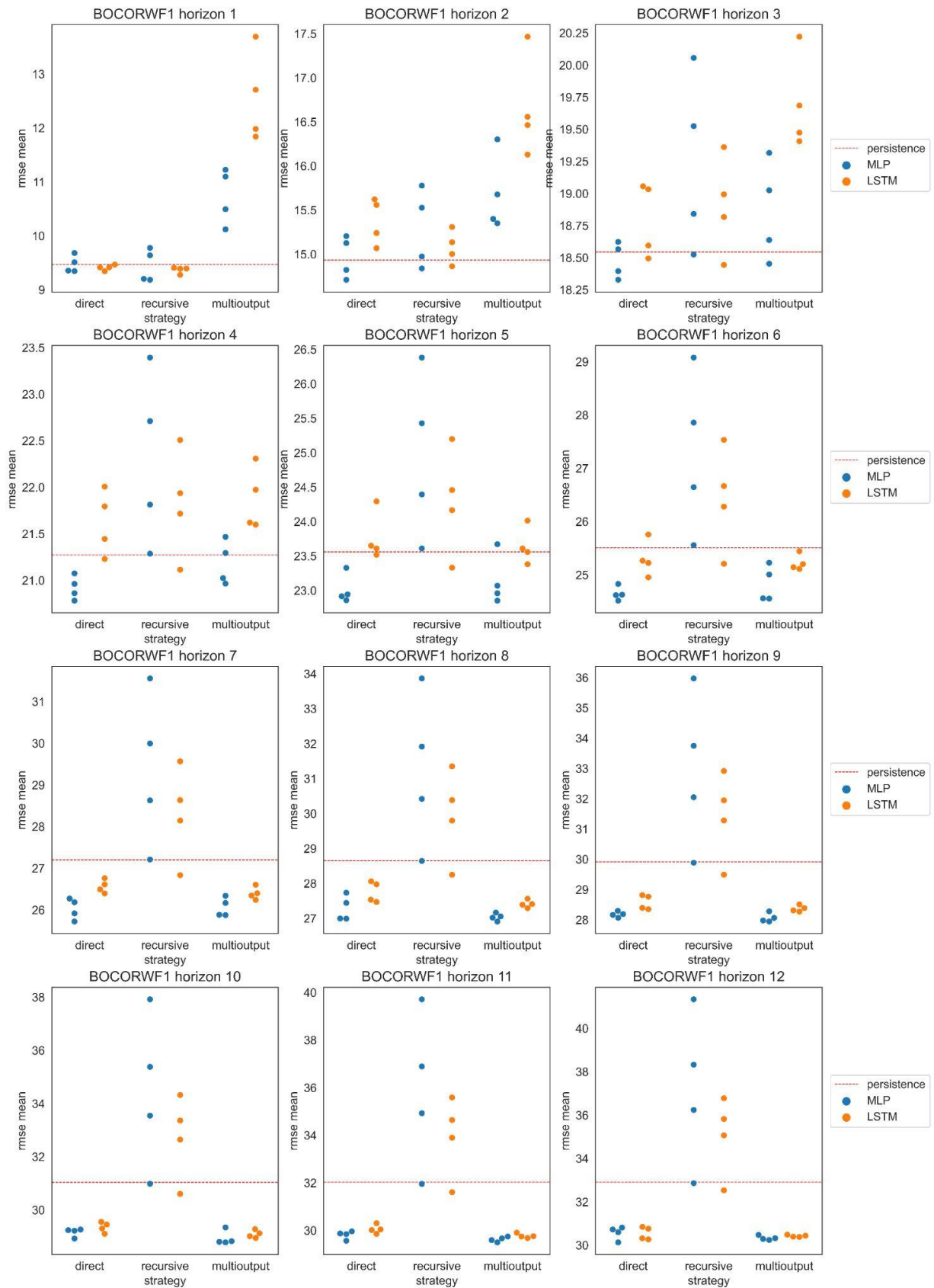


Figura 21. Média do RMSE para cada horizonte (BOCORWF1)

D - Central Macarthur

A Tabela 7 abaixo apresenta a média geral do RMSE ao longo dos horizontes e splits. A partir desta tabela foram ilustrados todos os pontos da Figura 18. Também foi a partir dela que se obtiveram os dados de melhor parametrização exibidos na Tabela 4.

Tabela 7 - Média geral do RMSE (MACARTH1)

	model	strategy	input_size	hidden_layers	rmse mean	rmse std
0	MLP	multioutput	4	[24, 12]	65.184855	5.029576
1	MLP	direct	4	[24, 12]	65.230224	4.811660
2	MLP	direct	4	[12]	65.579656	4.866136
3	MLP	direct	12	[24, 12]	66.354108	5.181735
4	LSTM	recursive	4	[12]	66.493975	4.395350
5	MLP	multioutput	12	[24, 12]	66.525598	5.695851
6	MLP	direct	12	[12]	67.035718	4.894645
7	LSTM	direct	4	[12]	67.233842	5.487319
8	LSTM	recursive	4	[24, 12]	67.620272	5.747654
9	persistence	NaN	NaN	NaN	67.622216	6.111518
10	LSTM	direct	12	[12]	67.644843	5.534046
11	MLP	multioutput	12	[12]	68.093465	6.414304
12	LSTM	direct	4	[24, 12]	68.126941	5.360563
13	MLP	multioutput	4	[12]	68.157415	7.838066
14	LSTM	multioutput	12	[12]	68.380338	6.796884
15	LSTM	multioutput	4	[12]	68.515034	5.090440
16	LSTM	direct	12	[24, 12]	68.713108	5.506059
17	MLP	recursive	4	[12]	69.085383	7.336785
18	LSTM	multioutput	4	[24, 12]	69.586682	6.624234
19	LSTM	recursive	12	[12]	69.647426	8.746248
20	LSTM	multioutput	12	[24, 12]	69.872292	7.320446
21	MLP	recursive	4	[24, 12]	72.038537	8.187290
22	MLP	recursive	12	[12]	73.992572	7.796994
23	LSTM	recursive	12	[24, 12]	76.446233	19.009492
24	MLP	recursive	12	[24, 12]	77.602436	8.400541

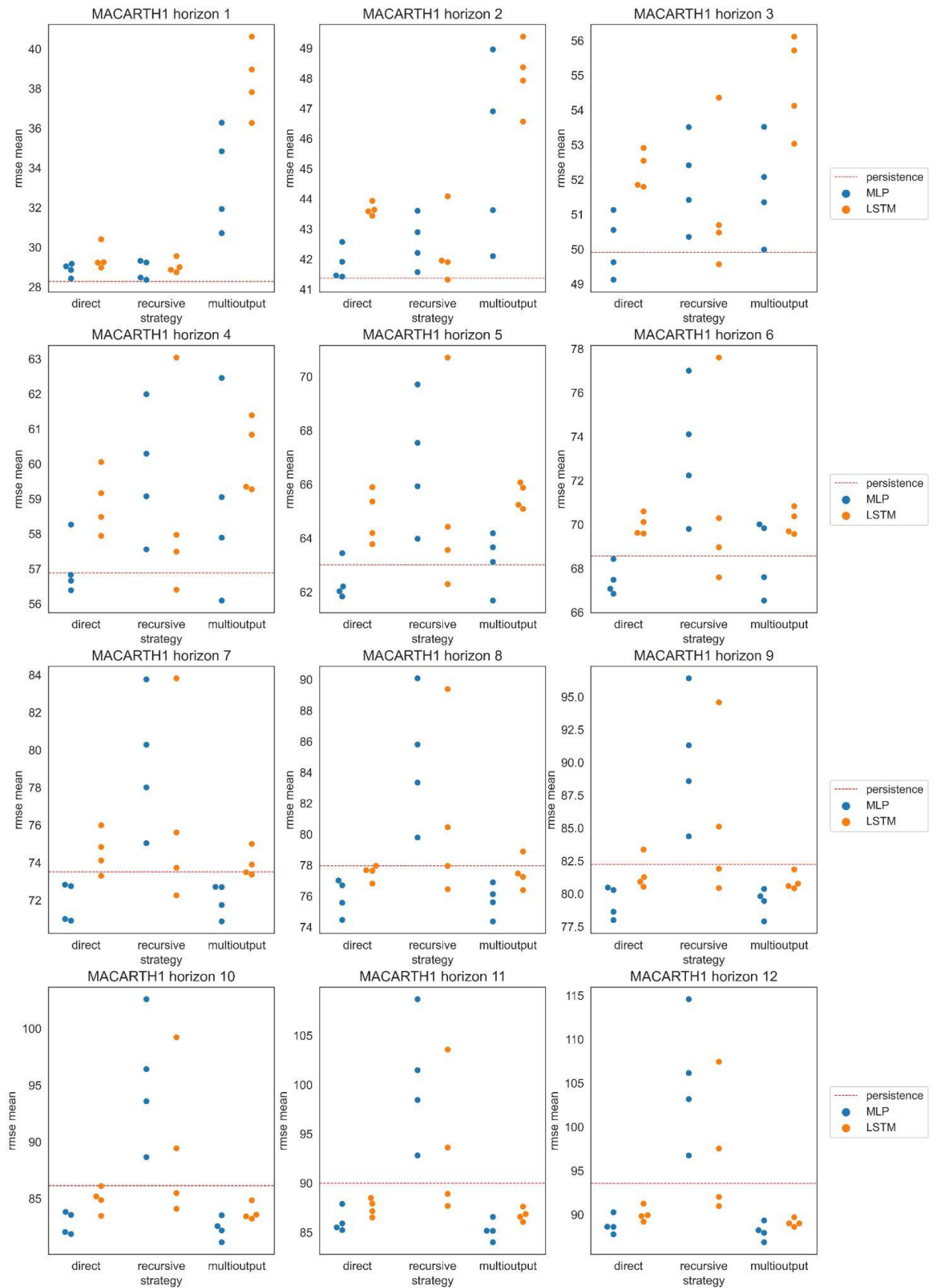


Figura 22. Média do RMSE para cada horizonte (MACARTH1)