

UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE ENGENHARIA GEOGRÁFICA, GEOFÍSICA E ENERGIA



Previsão de curto tempo da produção eólica nacional

João Cravinho Santos

Mestrado Integrado em Engenharia da Energia e do Ambiente

Dissertação orientada por:
Professora Doutora Ana Isabel Lopes Estanqueiro
Mestre António Manuel Vitoriano Couto

2021

Resumo

De acordo com as metas estabelecidas para 2050 cerca de 32% da energia elétrica produzida e consumida, a nível europeu, terá de provir de fontes de energias renováveis. Nesse sentido é expectável que no futuro a penetração de fontes de energia renovável nos sistemas electroprodutores aumente. Para permitir que essa crescente penetração renovável ocorra sem prejuízos para a manutenção da qualidade de serviço e robustez do sistema electroprodutor é necessário desenvolver ferramentas e instrumentos que permitam essa mesma integração e que auxiliem a gestão de todo o sistema.

Esta dissertação debruça-se sobre uma das fontes renováveis com maior expressividade no *mix* energético português, a energia eólica. Atualmente para auxiliar na gestão do sistema electroprodutor ferramentas de previsão de produção desempenham um papel preponderante. O valor destas ferramentas será ainda mais relevante no futuro onde é esperada uma elevada penetração eólica nos sistemas elétricos tendencialmente 100% baseados em energias renováveis.

Assim sendo a presente dissertação tem como foco a aplicação e estudo da fiabilidade de metodologias de previsão de produção eólica de curto horizonte temporal (6 horas) para Portugal continental. Foram desenvolvidas, implementadas e analisadas duas metodologias de previsão de produção eólica distintas: o algoritmo dos *k*-vizinhos mais próximos e as redes neuronais artificiais. Foi também estudada a otimização de alguns dos parâmetros mais relevantes de cada uma das metodologias e o impacto dessa otimização na eficácia das previsões de produção eólica. Estas metodologias foram comparadas a uma abordagem de persistência que é utilizada neste trabalho como referência (*benchmark*) Para auxiliar neste processo de otimização procedeu-se a análise do erro médio quadrático, viés e correlação e implementou-se uma medida para análise da capacidade de sobre ou sub estimação das variações da produção em seis horas.

Verificou-se que, das metodologias de previsão de produção eólica em estudo, a que apresenta melhor desempenho nas métricas analisadas foi a metodologia baseada em redes neuronais artificiais. Esta destaca-se pelo facto de conseguir identificar relações não lineares nos dados, característica que não está presente nas restantes metodologias de previsão implementadas. Face à metodologia de referência, as previsões obtidas através da metodologia das redes neuronais artificiais permitiram uma diminuição do erro médio quadrático em 15,27%. Já no caso da configuração mais adequada da metodologia dos *k*-vizinhos mais próximos a diminuição foi de 10.8%. A otimização das metodologias de previsão implementadas provou ser complexa, contudo demonstrou ser um passo extremamente relevante para identificar melhorias significativas na eficácia das previsões.

Palavras-chave: Previsão produção eólica, aprendizagem automática, *k*-vizinhos mais próximos, redes neuronais artificiais.

Abstract

According to the established goals to the horizon of 2050, about 32% of the total electric energy produced and consumed, in the European Union, should be provided from renewable energy sources. It is therefore expected that in the future, these technologies will play an increasing role in the energy production sector. Hence it is important to develop tools that allow the increasing renewable energy sources penetration and, at the same time, guarantee the safety, robustness, quality, and management of each power system.

This dissertation delves in wind production, one of the most expressive renewable energy sources in the Portuguese energetic mix. To help the present and future scenario of high wind penetration forecasting systems will play a crucial role in the management of the energy production system.

Thus, the present dissertation focuses on studying the effectiveness of wind power production forecasting methodologies to a short time horizon (6 hours) using the (continental) Portuguese national aggregated time series. Two distinct methodologies were implemented and analyzed: k-nearest neighbors' and artificial neural network. This dissertation focused, as well, in the optimization of key parameters for each of the methodologies and the impact in the forecast effectiveness resulted from the optimization. The focused methodologies were then compared against a benchmark forecasting methodology: the persistence. For the optimization and effectiveness of forecasting different error metrics were studied: bias, error, root mean squared error and an attempt to interpret the ability, of each methodology, to under or overestimate the forecasts.

It was found that, the one who generates more effective forecasts is the artificial neural network methodology. This one stands out mainly due to its non-linear nature and enabling to identify and solve non-linear connections in the data, which is not possible to achieve with the other methodologies implemented in this dissertation. Compared to the benchmark forecasting methodology, the forecasts obtained with the artificial neural networks were able to decrease the root mean squared error in 15,27%. In the case of the most effective configuration of the k-nearest neighbor methodology, a decrease, of the same error metric, of 10.8% was identified. The optimization of the forecasting methodologies has proven to be a difficult process, but nonetheless very relevant and important step to achieve better wind power forecasts.

Key words: Wind power production forecast, machine learning, k-nearest neighbors, artificial neural networks.

Índice

Resumo	iii
Abstract	iv
Índice	v
Lista de Figuras	vii
Lista de Tabelas	x
Simbologia e Notação	xi
1. Introdução	12
1.1. Enquadramento Geral	12
1.2. Objetivos da Dissertação	14
1.3. Organização da Dissertação	14
2. Energia Eólica e os desafios à sua integração no Sistema Electroprodutor	14
2.1. O vento e o aproveitamento eólico	14
2.2. Ciclos Naturais e Flutuação de Potência	15
2.3. Efeitos da penetração eólica no sistema electroprodutor	16
2.3.1. Reservas de Potência	17
3. Previsão de Produção Eólica – Relevância e Revisão de Literatura	18
3.1. Escalas Temporais	18
3.2. Metodologias de Previsão	19
3.2.1. Métodos Físicos	19
3.2.2. Métodos Estatísticos	21
3.2.3. Métodos Híbridos	22
4. Dados e Metodologias de Previsão	23
4.1. Tratamento de Dados	23
4.2. Metodologia dos k vizinhos mais próximos	24
4.2.1. Aplicações	25
4.2.2. Principais passos do algoritmo	27
4.2.3. Definição do número de vizinhos	28
4.2.4. Semelhança entre instâncias	28
4.2.5. Ponderação	29
4.2.6. Implementação do método dos k vizinhos mais próximos	29
4.3. Metodologia das Redes Neurais Artificiais	33
4.3.1. Elementos de uma RNA	34
4.3.2. Topologia da Rede	36
4.3.3. Aprendizagem	38

4.3.4.	Algoritmo de retropropagação	39
4.3.5.	Implementação das Redes Neurais Artificiais	45
4.4.	Métricas de Avaliação.....	47
5.	Apresentação, análise e discussão de resultados.....	50
5.1.	Análises de Sensibilidade.....	50
5.2.	Métricas de erro para as metodologias implementadas.....	51
5.2.1.	Persistência	51
5.2.2.	<i>k</i> Vizinhos Mais Próximos	52
5.2.3.	Redes Neurais Artificiais	60
5.3.	Desempenho das previsões nas variações de produção	63
5.3.1.	Metodologia kNN	63
5.3.2.	Metodologia RNA.....	64
5.4.	Análise do máximo, mínimo e média produção eólica normalizada	65
5.4.1.	Caso produção eólica normalizada máxima.....	65
5.4.2.	Caso produção eólica normalizada mínima	66
5.5.	Análise da sazonalidade do erro de previsão	68
5.6.	Síntese da análise dos resultados	70
6.	Conclusões finais	71
	Trabalhos futuros	72
	Referências Bibliográficas	74
I.	Anexo – Análises de Sensibilidade referentes às metodologias implementadas	79
II.	Análise detalhada dos cenários estudados.....	81

Lista de Figuras

FIGURA 1-1: EVOLUÇÃO DA PRODUÇÃO DE ELETRICIDADE EM PORTUGAL CONTINENTAL. _____	12
FIGURA 2-1: ESCALAS TEMPORAIS RELEVANTES, REFERENTES AOS CICLOS NATURAIS DAS FER. _____	15
FIGURA 2-2: EFEITO DE SUAUIZACÃO POR AGREGACÃO: POTÊNCIA ENTREGUE POR 1 AEROGERADOR (LINHA AZUL); POR 1 PARQUE EÓLICO (LINHA VERDE); POR 4 PARQUES EÓLICOS (LINHA VERMELHA); POR 24 PARQUES EÓLICOS (LINHA PRETA). _____	16
FIGURA 2-3: EFEITOS NO SE CONSEQUENTES DA PENETRAÇÃO DA ENERGIA EÓLICA. _____	16
FIGURA 2-4: EXEMPLO DE VARIABILIDADE E INCERTEZA _____	17
FIGURA 4-1: PERFIL MÉDIO HORÁRIO DA PRODUÇÃO EÓLICA NORMALIZADA PARA OS RESPECTIVOS PERÍODOS TEMPORAIS. _____	24
FIGURA 4-2: EXEMPLO DE CLASSIFICAÇÃO DE UMA INSTÂNCIA SEGUNDO O ALGORITMO DE k VIZINHOS MAIS PRÓXIMOS. A ESTRELA A VERMELHO REPRESENTA O OBJETO EM ESTUDO; OS CÍRCULOS CORRESPONDEM ÀS INSTÂNCIAS MAIS PRÓXIMAS DO OBJETO EM ESTUDO. _____	25
FIGURA 4-3: EXEMPLO DA UTILIZAÇÃO DO ALGORITMO k NN COM VISTA A REGRESSÃO.. _____	26
FIGURA 4-4: EXEMPLO DA VARIAÇÃO DA SEMELHANÇA, OU DISTÂNCIA, ENTRE AS k INSTÂNCIAS MAIS PRÓXIMAS. OS CÍRCULOS AMARELOS REPRESENTAM OS DADOS PERTENCENTES AO CONJUNTO HISTÓRICO DE DADOS; O QUADRADO A VERMELHO REPRESENTA A INSTÂNCIA ALVO, PERTENCENTE AO CONJUNTO DE TESTE. _____	27
FIGURA 4-5: FLUXOGRAMA REFERENTE AO FUNCIONAMENTO DO ALGORITMO DOS k VIZINHOS MAIS PRÓXIMOS. _____	27
FIGURA 4-6: EXEMPLO DA INSTABILIDADE DO ALGORITMO k NN, COM VISTA A CLASSIFICAÇÃO, PROVOCADA POR UM VALOR MÍNIMO PARA O PARÂMETRO DO NÚMERO DE VIZINHOS MAIS PRÓXIMOS (k). A ESTRELA A VERMELHO REPRESENTA O DADO ALVO A CLASSIFICAR; OS CÍRCULOS A VERDE E VERMELHO CORRESPONDEM A DADOS DO CONJUNTO HISTÓRICO DE DADOS DE DUAS CLASSES DISTINTAS: CLASSE A E CLASSE B, RESPECTIVAMENTE. _____	28
FIGURA 4-7: DIAGRAMA COM A ESQUEMATIZAÇÃO DOS PRINCIPAIS PASSOS DA METODOLOGIA DO ALGORITMO DOS k NN. _____	30
FIGURA 4-8: MODELO DE UM NEURÓNIO _____	34
FIGURA 4-9: ESQUEMA REPRESENTATIVO DAS ETAPAS DURANTE O PROCESSAMENTO DE DADOS LEVADO A CABO POR UM NEURÓNIO. _____	34
FIGURA 4-10: REPRESENTAÇÃO DAS FUNÇÕES DE TRANSFERÊNCIA MAIS UTILIZADAS: A) FUNÇÃO DE TRANSFERÊNCIA LIMIAR; B) FUNÇÃO DE TRANSFERÊNCIA LINEAR; C) FUNÇÃO DE TRANSFERÊNCIA SIGMOIDE, COM DIFERENTES DECLIVES. _____	35
FIGURA 4-11: ESQUEMA REPRESENTATIVO DE UMA REDE NEURONAL ARTIFICIAL COM UMA CAMADA (ESQUERDA) E MULTICAMADA, COM n -CAMADAS ESCONDIDAS (DIREITA). CADA NÓDULO CIRCULAR REPRESENTA UM NEURÓNIO ARTIFICIAL E CADA SETA ENTRE CAMADAS REPRESENTA UMA LIGAÇÃO ENTRE A SAÍDA (<i>OUTPUT</i>) DE UM NEURÓNIO E A ENTRADA (<i>INPUT</i>) DE OUTRO NEURÓNIO, PERTENCENTE À CAMADA SEGUINTE. _____	37
FIGURA 4-12: ARQUITETURAS POSSÍVEIS DE UMA RNA SEGUNDO AS LIGAÇÕES ENTRE CAMADAS. _____	37
FIGURA 4-13: EXEMPLO DE REDE NEURONAL ARTIFICIAL RECORRENTE COM OS ELEMENTOS DE RETARDAMENTO DE UNIDADES PRESENTES _____	38
FIGURA 4-14: PROCESSO DE CALIBRAÇÃO DOS PESOS PELO ESTUDO DO GRADIENTE DA FUNÇÃO DE ERRO. _____	40
FIGURA 4-15: EXEMPLIFICAÇÃO DA INFLUÊNCIA DA TAXA DE APRENDIZAGEM NA BUSCA PELO MÍNIMO DA FUNÇÃO. CADA PONTO, REPRESENTADO, EM CADA UM DOS EXEMPLOS, REPRESENTA A SOLUÇÃO, AJUSTADA, PARA OS PESOS, EM CADA INSTANTE. EM A) É POSSÍVEL OBSERVAR O EFEITO DE UMA TAXA DE APRENDIZAGEM MUITO BAIXA; EM B) UMA TAXA DE APRENDIZAGEM INTERMÉDIA; EM C) UMA TAXA DE APRENDIZAGEM ALTA; EM D) UMA TAXA DE APRENDIZAGEM MUITO ALTA. _____	42
FIGURA 4-16: FLUXOGRAMA REFERENTE À CALIBRAÇÃO DE UMA REDE NEURONAL ARTIFICIAL UTILIZANDO O ALGORITMO DE RETRO PROPAGAÇÃO. _____	43
FIGURA 4-17: COMPARAÇÃO ENTRE OS DIFERENTES ALGORITMOS ESTUDADOS PARA O ALGORITMO DE RETROPROPAGAÇÃO. _____	43
FIGURA 4-18: FLUXOGRAMA REFERENTE AO TREINO DE UMA REDE NEURONAL ARTIFICIAL UTILIZANDO O ALGORITMO DE LEVENBERG-MARQUARDT _____	45
FIGURA 4-19: ESTRUTURA GERAL DA REDE NEURONAL ARTIFICIAL UTILIZADA. _____	45

FIGURA 4-20: DIAGRAMA COM A ESQUEMATIZAÇÃO DOS PRINCIPAIS PASSOS DA METODOLOGIA DAS REDES NEURONAIS ARTIFICIAIS.	46
FIGURA 5-1: ERRO MÉDIO HORÁRIO DA PREVISÃO DE PRODUÇÃO EÓLICA NORMALIZADA, PARA A METODOLOGIA DE PREVISÃO PERSISTÊNCIA.	52
FIGURA 5-2: RMSE MÉDIO HORÁRIO DA PREVISÃO DE PRODUÇÃO EÓLICA NORMALIZADA, PARA A METODOLOGIA DA PERSISTÊNCIA.	52
FIGURA 5-3: REPRESENTAÇÃO GRÁFICA DOS RESULTADOS OBTIDOS PARA AS MÉTRICAS DO RMSE (BARRAS) E CORRELAÇÃO (LINHAS), PARA AS PREVISÕES OBTIDAS ATRAVÉS DA METODOLOGIA k NN.	55
FIGURA 5-4: ERRO MÉDIO DA PREVISÃO DE PRODUÇÃO EÓLICA NORMALIZADA, PARA A METODOLOGIA k NN PARA DIFERENTES VALORES DE k .	57
FIGURA 5-5: RMSE MÉDIO HORÁRIO DA PREVISÃO DE PRODUÇÃO EÓLICA NORMALIZADA, PARA A METODOLOGIA k NN PARA DIFERENTES VALORES DE k .	59
FIGURA 5-6: ERRO MÉDIO HORÁRIO DA PREVISÃO DE PRODUÇÃO EÓLICA NORMALIZADA, PARA A METODOLOGIA DE PREVISÃO RNA.	61
FIGURA 5-7: RMSE MÉDIO HORÁRIO DA PREVISÃO DE PRODUÇÃO EÓLICA NORMALIZADA, PARA A METODOLOGIA RNA.	62
FIGURA 5-8: MAPEAMENTO DAS OCORRÊNCIAS OBTIDAS PARA CADA CLASSE, REFERENTE À METODOLOGIA IMPLEMENTADA k NN, ONDE k ASSUME O VALOR DE 1. A BARRA DE COR INDICA A ESCALA DE COR CONSOANTE O NÚMERO DE OCORRÊNCIAS VERIFICADO. NO EIXO DOS X ESTÃO REPRESENTADAS AS TAXAS DE VARIAÇÃO DA PRODUÇÃO EÓLICA NORMALIZADA OBSERVADA, ENQUANTO NOS EIXOS Y ENCONTRAM-SE REPRESENTADAS AS TAXAS DE VARIAÇÃO DA PREVISÃO DE PRODUÇÃO EÓLICA NORMALIZADA.	64
FIGURA 5-9: PRODUÇÃO E PREVISÃO EÓLICA NORMALIZADA PARA O CASO DE PRODUÇÃO MÁXIMA, PARA O PERÍODO DE VALIDAÇÃO.	65
FIGURA 5-9: PRODUÇÃO E PREVISÃO EÓLICA NORMALIZADA PARA O CASO DE PRODUÇÃO MÍNIMA, PARA O PERÍODO DE VALIDAÇÃO.	67
FIGURA 5-11: PERFIL DE RMSE MÉDIO DO REFERENTE À ESTAÇÃO DO INVERNO, PARA A METODOLOGIA DE PREVISÃO RNA E k NN, COM FOCO NAS VARIAÇÕES COM A DISTÂNCIA EUCLIDIANA E DE MINKOWSKI.	68
FIGURA 5-12: PERFIL DE RMSE MÉDIO REFERENTE À ESTAÇÃO DO OUTONO, PARA A METODOLOGIA DE PREVISÃO RNA E k NN, COM FOCO NAS VARIAÇÕES COM A DISTÂNCIA EUCLIDIANA E DE MINKOWSKI.	69
FIGURA 5-13: PERFIL DE RMSE MÉDIO REFERENTE À ESTAÇÃO DA PRIMAVERA, PARA A METODOLOGIA DE PREVISÃO RNA E k NN, COM FOCO NAS VARIAÇÕES COM A DISTÂNCIA EUCLIDIANA E DE MINKOWSKI.	69
FIGURA 5-14: PERFIL DE RMSE MÉDIO REFERENTE À ESTAÇÃO DO VERÃO, PARA A METODOLOGIA DE PREVISÃO RNA E k NN, COM FOCO NAS VARIAÇÕES COM A DISTÂNCIA EUCLIDIANA E DE MINKOWSKI.	69
FIGURA I-1: ANÁLISE DE SENSIBILIDADE REFERENTE AO VALOR DA ORDEM p DA DISTÂNCIA DE MINKOWSKI. CADA LINHA A CORES REPRESENTA A VARIAÇÃO DA DISTÂNCIA DE UM PONTO, PERTENCENTE À SÉRIE DE DADOS HISTÓRICA, À ORIGEM.	79
FIGURA II-1: RMSE MÉDIO HORÁRIO DA PREVISÃO DE PRODUÇÃO EÓLICA NORMALIZADA, PARA A METODOLOGIA k NN UTILIZANDO A MÉTRICA DE DISTÂNCIA EUCLIDIANA E DE MINKOWSKI, PARA TODOS OS VALORES DE k .	82
FIGURA II-2: MAPEAMENTO DAS OCORRÊNCIAS OBTIDAS PARA CADA CLASSE, REFERENTE À METODOLOGIA IMPLEMENTADA k NN, ONDE k ASSUME O VALOR DE 1. A BARRA DE COR INDICA A ESCALA DE COR CONSOANTE O NÚMERO DE OCORRÊNCIAS VERIFICADO. NO EIXO DOS X ESTÃO REPRESENTADAS AS TAXAS DE VARIAÇÃO DA PRODUÇÃO EÓLICA NORMALIZADA OBSERVADA, ENQUANTO NOS EIXOS Y ENCONTRAM-SE REPRESENTADAS AS TAXAS DE VARIAÇÃO DA PREVISÃO DE PRODUÇÃO EÓLICA NORMALIZADA.	82
FIGURA II-3: SIMILAR À FIGURA II-2 MAS ONDE k ASSUME O VALOR DE 5	83
FIGURA II-4: SIMILAR À FIGURA II-2 MAS ONDE k ASSUME O VALOR DE 10	83
FIGURA II-5: SIMILAR À FIGURA II-2 MAS ONDE k ASSUME O VALOR DE 20	84
FIGURA II-6: SIMILAR À FIGURA II-2 MAS ONDE k ASSUME O VALOR DE 162	84
FIGURA II-7: MAPEAMENTO DAS OCORRÊNCIAS OBTIDAS PARA CADA CLASSE, REFERENTE À METODOLOGIA IMPLEMENTADA RNA. A BARRA DE COR INDICA A ESCALA DE COR CONSOANTE O NÚMERO DE OCORRÊNCIAS VERIFICADO. NO EIXO DOS X ESTÃO REPRESENTADAS AS TAXAS DE VARIAÇÃO DA PRODUÇÃO EÓLICA	

NORMALIZADA OBSERVADA, ENQUANTO NOS EIXOS Y ENCONTRAM-SE REPRESENTADAS AS TAXAS DE VARIÇÃO DA PREVISÃO DE PRODUÇÃO EÓLICA NORMALIZADA. _____	85
FIGURA II-8: REPRESENTAÇÃO GRÁFICA DO RMSE MÉDIO PARA O CASO MÍNIMO DE PRODUÇÃO EÓLICA MÉDIA NORMALIZADA, PARA O ANO DE 2018. _____	85
FIGURA II-9: REPRESENTAÇÃO GRÁFICA DO RMSE MÁXIMO PARA O CASO MÉDIO DE PRODUÇÃO EÓLICA MÉDIA NORMALIZADA, PARA O ANO DE 2018. _____	86
FIGURA II-10: PERFIL DIÁRIO DO RMSE REFERENTE À ESTAÇÃO DO INVERNO, PARA TODAS AS METODOLOGIAS DE PREVISÃO DE PRODUÇÃO EÓLICA IMPLEMENTADAS. _____	86
FIGURA II-11: PERFIL DIÁRIO DO RMSE REFERENTE À ESTAÇÃO DO OUTONO, PARA TODAS AS METODOLOGIAS DE PREVISÃO DE PRODUÇÃO EÓLICA IMPLEMENTADAS. _____	87
FIGURA II-12: PERFIL DIÁRIO DO RMSE REFERENTE À ESTAÇÃO DA PRIMAVERA, PARA TODAS AS METODOLOGIAS DE PREVISÃO DE PRODUÇÃO EÓLICA IMPLEMENTADAS. _____	87
FIGURA II-13: PERFIL DIÁRIO DO RMSE REFERENTE À ESTAÇÃO DO VERÃO, PARA TODAS AS METODOLOGIAS DE PREVISÃO DE PRODUÇÃO EÓLICA IMPLEMENTADAS. _____	88

Lista de Tabelas

TABELA 3-1: ESCALAS TEMPORAIS E SUAS APLICAÇÕES PARA AS VÁRIAS CLASSES DE PREVISÃO DE PRODUÇÃO EÓLICA.....	18
TABELA 3-2: REVISÃO BIBLIOGRÁFICA REFERENTE A PREVISÕES DE VELOCIDADE DO VENTO OU POTÊNCIA EÓLICA E CORRESPONDENTE MÉTRICA DE ERRO ENCONTRADA PARA O RESPECTIVO ESTUDO.....	19
TABELA 5-1: VALORES DAS MÉTRICAS DE ERRO, CALCULADAS PARA AS PREVISÕES OBTIDAS COM PERSISTÊNCIA.	51
TABELA 5-2: VALORES DO VIÉS, P.U., CALCULADOS PARA AS PREVISÕES OBTIDAS COM A METODOLOGIA k NN. A NEGRITO É INDICADA A CONFIGURAÇÃO COM O MELHOR DESEMPENHO.....	52
TABELA 5-3: VALORES DO RMSE, P.U., CALCULADO PARA AS PREVISÕES OBTIDAS COM A METODOLOGIA k NN. A NEGRITO É INDICADA A CONFIGURAÇÃO COM O MELHOR DESEMPENHO.....	53
TABELA 5-4: VALORES DA CORRELAÇÃO, CALCULADO PARA AS PREVISÕES OBTIDAS COM A METODOLOGIA k NN.	54
TABELA 5-5: VARIAÇÃO PERCENTUAL DO RMSE MÉDIO DA PREVISÃO DE PRODUÇÃO EÓLICA RESULTANTE DA METODOLOGIA k NN, FACE À PERSISTÊNCIA.	59
TABELA 5—6: VALORES MÉDIOS PARA AS MÉTRICAS DE ERRO, CALCULADAS PARA AS PREVISÕES OBTIDAS COM A METODOLOGIA DAS REDES NEURONAIIS ARTIFICIAIS.	60
TABELA 5-9: VARIAÇÃO PERCENTUAL DO RMSE MÉDIO DA PREVISÃO DE PRODUÇÃO EÓLICA RESULTANTE DA METODOLOGIA RNA, FACE À PERSISTÊNCIA.....	62
TABELA I-1: VALOR DE ERRO, SEGUNDO CONDIÇÃO IMPOSTA, PARA A ORDEM P DA DISTÂNCIA DE MINKOWSKI. 79	
TABELA I-2: RMSE OBTIDO PARA A ANÁLISE DE SENSIBILIDADE DAS RNA, REFERENTE AO ESTUDO ACERCA DO NÚMERO ÓTIMO DE NEURÓNIOS PRESENTES NA CAMADA ESCONDIDA. NA PRIMEIRA LINHA DA TABELA ENCONTRA-SE O TAMANHO, I.E. NÚMERO DE NEURÓNIOS, NA CAMADA DE ENTRADA. DE REPRESENTA O NÚMERO DE NEURÓNIOS DA CAMADA DE ENTRADA DA REDE, SEIS. OS RESULTADOS ENCONTRAM-SE MAPEADOS POR COR: QUANTO MAIS VERMELHO MAIOR SERÁ O RMSE E PIOR SERÁ O DESEMPENHO DA MÉTRICA DE AVALIAÇÃO – O RMSE. PELO CONTRÁRIO, QUANTO MAIS VERDE, MENOR O VALOR ENCONTRADO PARA O RMSE E MELHOR SERÁ O DESEMPENHO DESTA MÉTRICA.	80
TABELA I-3: CORRELAÇÃO OBTIDA PARA A ANÁLISE DE SENSIBILIDADE DAS RNA, REFERENTE AO ESTUDO ACERCA DO NÚMERO ÓTIMO DE NEURÓNIOS PRESENTES NA CAMADA ESCONDIDA. NA PRIMEIRA LINHA DA TABELA ENCONTRA-SE O TAMANHO, I.E. NÚMERO DE NEURÓNIOS, NA CAMADA DE ENTRADA. DE REPRESENTA O NÚMERO DE NEURÓNIOS DA CAMADA DE ENTRADA DA REDE, SEIS. OS RESULTADOS ENCONTRAM-SE MAPEADOS POR COR: QUANTO MAIS VERMELHO MENOR SERÁ A CORRELAÇÃO E PIOR SERÁ O DESEMPENHO DA MÉTRICA DE AVALIAÇÃO – A CORRELAÇÃO. PELO CONTRÁRIO, QUANTO MAIS VERDE, MAIOR O VALOR ENCONTRADO PARA A CORRELAÇÃO, E MELHOR SERÁ O DESEMPENHO DESTA MÉTRICA. ...	80
TABELA II-1: VARIAÇÕES PERCENTUAIS DO RMSE MÉDIO, OBTIDO COM A IMPLEMENTAÇÃO DA METODOLOGIA DE PREVISÃO k NN, ENTRE OS MÉTODOS DE CÁLCULO DE DISTÂNCIAS ENTRE AS INSTÂNCIAS.....	81
TABELA II-2: VARIAÇÕES PERCENTUAIS DA CORRELAÇÃO MÉDIA, OBTIDA COM A IMPLEMENTAÇÃO DA METODOLOGIA DE PREVISÃO RNA FACE AOS RESULTADOS DA METODOLOGIA DE PREVISÃO DA PERSISTÊNCIA E k NN, PARA A REFERIDA MÉTRICA.	81

Simbologia e Notação

FER	Fontes de energia renováveis
EU	União Europeia
GEE	Gases de efeito de estufa
SE	Sistema electroprodutor
NWP	Numerical weather prediction
AR	Modelos autorregressivos
MA	Média móvel
ARMA	Autorregressivos de média movel
ARIMA	Autorregressivos integrados de média móvel
IA	Inteligência artificial
ML	Machine Learning
kNN	k vizinhos mais próximos
$wkNN$	k vizinhos mais próximos ponderado
RNA	Redes neuronais artificiais
LNEG	Laboratório Nacional de Energia e Geologia
DGEG	Direção Geral Energia e Geologia
p_t	Produção eólica normalizada [p.u.]
HT	Horizonte temporal
NH	Número intervalos horários
L_p	Distância de Minkowski, de ordem p
w_{k_i}	Ponderação
LM	Levenberg-Marquardt
e	Desvio [p.u.]
$\overline{e(t)}$	Viés [p.u.]
RMSE	Erro quadrático médio [p.u.]
R^2	Correlação
TVO	Taxa de variação da produção eólica normalizada
TVP	Taxa de variação da previsão de produção eólica normalizada

1. Introdução

1.1. Enquadramento Geral

Onde antes a alternativa renovável se prendia única e exclusivamente com a finitude dos recursos fósseis, hoje, a aposta em fontes de energia renováveis (FER), deve-se principalmente à criação de modelos energéticos sustentáveis aliada à crescente preocupação ambiental. Sucessivos desenvolvimentos tecnológicos têm permitido um maior e mais eficiente aproveitamento dos recursos energéticos renováveis com base na tecnologia solar, eólica, geotérmica e a biomassa.

Foram postos em ação grandes esforços para que seja possível originar uma mudança significativa nesses dois campos – ambiental e energético. No que diz respeito à União Europeia (UE), por exemplo, foi imposto a cada estado-membro rigorosos objetivos em dois horizontes temporais distintos –2030 e 2050 – que têm em vista a obrigatoriedade de desenvolvimento das duas áreas. Os objetivos para o primeiro prazo, parte do programa *European Green Deal* [1], são:

- Redução de, pelo menos, 40% das emissões de gases de efeito de estufa (GEE) face aos níveis de 1990;
- Aumentar para 32% a fração do consumo e produção de energia da UE proveniente de FER;
- Redução de 32.5% do consumo de energia primária devido à aposta na eficiência energética.

No plano nacional é de ressaltar dois importantes planos para a reafirmação do compromisso europeu:

- PNEC – Plano Nacional de Energia e Clima [2];
- RNC2050 – Roteiro para a Neutralidade Carbónica [3].

Devido ao cenário político e socioeconómico nacional quer os planos quer os seus objetivos estão sob constante avaliação e atualização.

Apesar de todos os esforços direcionados para alterar o paradigma energético e ambiental é necessário que estes sejam contínuos e permanentes. O sector onde as FER têm ganho mais relevância tem sido no sector elétrico, abastecendo em 2017 cerca de 40,9% da produção de energia a nível nacional [4]. Na Figura 1-1 é possível observar a evolução do *mix* energético português ao longo dos últimos anos e o contínuo aumento de produção renovável no mesmo.

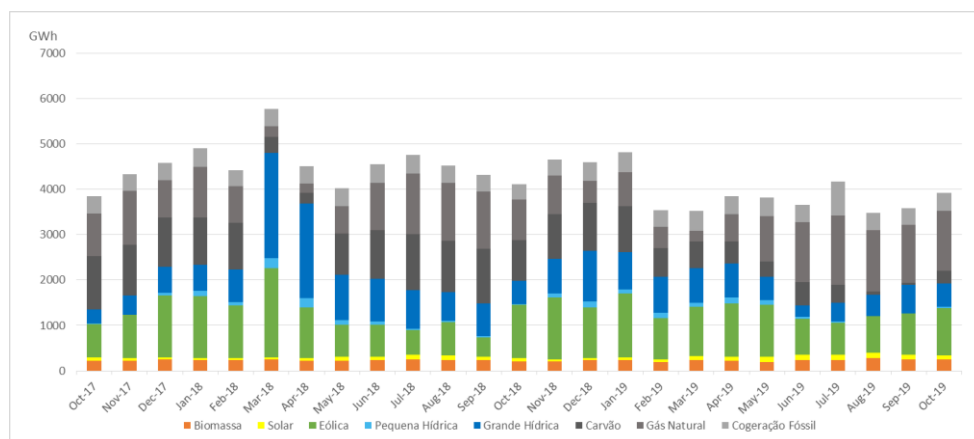


Figura 1-1: Evolução da produção de eletricidade em Portugal Continental (retirado de [5]).

A estocacidade natural associada aos recursos renováveis é uma das características fundamentais que distinguem as FER das fontes convencionais. Tanto o vento como o sol, por exemplo, são recursos variáveis no tempo e no espaço e, conseqüentemente, a produção também o vai ser. Assim sendo, a

produção de fontes renováveis não permite, regra geral¹, ao gestor de rede a possibilidade de controlar essa mesma produção conforme as necessidades do sistema electroprodutor (SE). O principal desafio consequente dessa característica prende-se com a necessidade de assegurar uma operação segura e estável do SE a custos reduzidos, isto é, garantir a qualquer instante que exista um equilíbrio entre a produção e o consumo sem um impacto socioeconómico significativo.

De forma a compensar esta estocacidade têm sido desenvolvidas e implementadas estratégias que possibilitem atingir esse equilíbrio:

- Aumento do nível de reservas operacionais do SE;
- Soluções de armazenamento de energia;
- Interligações com países vizinhos;
- Ferramentas de previsão;
- Gestão do consumo, ou *demand side management*.

Numa análise mais tradicional é possível afirmar que um SE com centrais de produção controláveis e de resposta rápida, como é o caso de grandes hídricas ou de centrais a gás de ciclo aberto, possui maior aptidão para dar resposta ao problema de flexibilidade do SE² [6]. Assim, essa flexibilidade, por vezes bastante onerosa, pode ser conseguida atendendo a alguns fatores em diferentes sectores: no sector da produção o equilíbrio pode ser assegurado pela existência de centrais despacháveis com baixo tempo de resposta e possibilidade de operar em diferentes regimes da sua curva de potência característica. No sector da rede elétrica e da operação do sistema electroprodutor passa pela melhoria das redes de transporte e distribuição de forma a permitir uma maior integração de FER, especialmente dispersas, e solucionar problemas de congestionamento na rede.

Nos últimos anos a energia produzida em Portugal com origem em centrais eólicas tem abastecido, entre 24-26% do consumo de energia elétrica. Tem-se então que esta fração é significativa no que toca à gestão do SE. Devido à, já mencionada, estocacidade do vento é necessário implementar e desenvolver ferramentas que aplicadas às fontes de energia eólica visem a minimização do impacto da integração desta fonte no SE, de modo a contribuir para uma eficiente gestão do SE e acima de tudo, o cumprimento das metas nacionais e internacionais.

Em SEs com elevada penetração de energia eólica no consumo, como se verifica em Portugal, é crucial a adoção de ferramentas de previsão [7]. Estas ferramentas permitem conhecer *a priori*, e com alguma precisão, as características do vento ou até mesmo saber qual a produção de energia que no futuro vai ocorrer. Entre outras possibilidades, estas ferramentas permitem:

- Apoiar nas decisões operacionais do operador e gestor de rede, proporcionando estabilidade e segurança de abastecimento na rede elétrica;
- Garantir a robustez e segurança de abastecimento do SE, visto que permite antecipadamente ter conhecimento dos valores futuros relativamente à produção eólica, para cada instante e para cada ponto de ligação à rede;
- Auxiliar, juntamente com dados de previsão do consumo, a gestão de reservas de potência necessárias reduzindo os custos do sistema;
- Auxiliar a otimização de sistemas de armazenamento de energia (como por exemplo as centrais de bombagem em albufeiras);

¹ Atualmente, a semelhança do que se verifica em vários países europeus, o SE já contempla o recurso a medidas de limitação/redução de energia (*curtailment*) de parques.

² Capacidade deste adaptar os padrões de produção e consumo de energia elétrica de modo a assegurar um equilíbrio constante entre a oferta e a procura.

- Retroativamente aumentar a possibilidade de aumentar a penetração eólica, consequente da maior facilidade de gestão da rede;
- Auxiliar os produtores de energia eólica nas suas ofertas nos mercados de eletricidade.

1.2. Objetivos da Dissertação

O objetivo da presente dissertação prende-se com a implementação, otimização e avaliação do desempenho de metodologias de previsão da produção eólica para curto prazo, i.e., previsões com um horizonte temporal de previsão de seis horas.

O presente estudo será elaborado com recurso a séries históricas horárias de produção eólica de Portugal Continental de modo a permitir avaliar a exatidão das metodologias implementadas para casos reais.

1.3. Organização da Dissertação

No primeiro capítulo da presente dissertação apresenta-se o panorama internacional e nacional geral relativo às FER, com maior foque no sector eólico. É também apresentada a motivação e organização da dissertação.

No segundo capítulo é abordado o papel e relevância da fonte eólica num sistema electroprodutor. De igual modo são apresentadas as dificuldades e problemáticas associadas à penetração de FER no mesmo.

No terceiro capítulo abordam-se os métodos de previsão eólica e as suas principais características.

No quarto capítulo apresenta-se a metodologia implementada na presente dissertação, nomeadamente, as técnicas estatísticas de previsão, o tratamento de dados e métricas de avaliação utilizadas.

No quinto capítulo desenvolve-se a análise e discussão dos resultados obtidos.

Por fim, no sexto capítulo são apresentadas as conclusões finais da dissertação seguida de recomendações para futuros trabalhos.

2. Energia Eólica e os desafios à sua integração no Sistema Electroprodutor

2.1. O vento e o aproveitamento eólico

O recurso eólico é, como descrito anteriormente, um recurso crucial para o futuro desenvolvimento de um sistema electroprodutor sustentável. A caracterização e a correta quantificação do recurso eólico é, portanto, um passo crucial para um eficiente aproveitamento deste.

A energia presente no vento é aproveitada pelo Homem desde cedo e de diferentes modos: desde a navegação com barcos movidos a velas, moinhos de vento para tarefas de natureza agrícola ou para bombagem de água, até à geração de eletricidade através de moinhos movidos pelo vento. Mais recentemente, a geração de energia diretamente a partir do vento foi possível graças ao desenvolvimento tecnológico de turbinas eólicas.

A energia associada ao movimento do vento é de natureza cinética e pode ser dada pela Equação 2-1:

$$E_c = \frac{1}{2}mv^2 \qquad \text{Equação 2-1}$$

onde m e v são, respetivamente, a massa e a velocidade do ar em movimento. Tendo em conta que é possível decompor a massa em função da densidade e do volume (que por sua vez depende da área) é possível escrever, sobre a potência passível de ser extraída pelo vento numa determinada área A , conforme a Equação 2-2:

$$P_{vento} = \frac{1}{2} \rho A v^3 \quad \text{Equação 2-2}$$

A potência é então proporcional à área de varredura A , do cubo da velocidade e ainda da densidade do ar. Idealmente, toda a energia deveria ser convertida em formas úteis e passíveis de serem aproveitadas. Contudo, apenas uma parte dessa energia é passível, com a tecnologia eólica presente, de ser convertida em energia útil: a conversão da energia cinética em energia mecânica é limitada pelo limite de Betz. Assim, o valor da eficiência teórica máxima para uma turbina eólica é de aproximadamente 59% [8].

2.2. Ciclos Naturais e Flutuação de Potência

Como já foi abordado uma das vertentes no que toca aos desafios inerentes à penetração em larga escala das FER num SE prende-se com os ciclos naturais dos recursos, *i.e.*, a variabilidade da fonte primária de cada recurso (como o vento, o sol, a pluviosidade, entre outros). A escala temporal em que os recursos variam pode ir desde uma escala de segundos à escala anual e cada uma irá impactar a gestão do SE de modos distintos. Na Figura 2-1 é possível observar as escalas temporais relevantes.

Em patamares de escalas muito elevadas (geracionais e inter-anuais) o impacto na gestão do SE não é muito elevado visto que as centrais que dependem dos recursos primários com este tipo de flutuação são capazes de garantir o abastecimento, desde que cuidadosamente planeadas. Em escalas temporais intermédias (anual, sazonal e horária) os desafios inerentes prendem-se com a operação e escalonamento das centrais despacháveis existentes no SE, impactando a gestão e flexibilidade do mesmo. No patamar de escalas mais rápidas (ao nível dos segundos e minutos) é onde se observam: *i)* fenómenos de flutuação de tensão nos pontos de interligação da rede e *ii)* variações na qualidade da energia entregue nas redes [6]. De facto, os ciclos associados a estas escalas temporais representam um entrave a uma maior penetração de FER e diversificação de fontes de energia, principalmente quando, como foi abordado na secção 1.1, o aproveitamento do recurso eólico desempenha um papel crucial e proeminente na gestão e evolução do sistema electroprodutor. Adicionalmente esta escala temporal é de extrema relevância para novos paradigmas como por exemplo, Centrais Renováveis Virtuais [6]. Assim, no âmbito desta dissertação de mestrado, o foco será tanto nas escalas rápidas (dos minutos) como nas intermédias (horas).

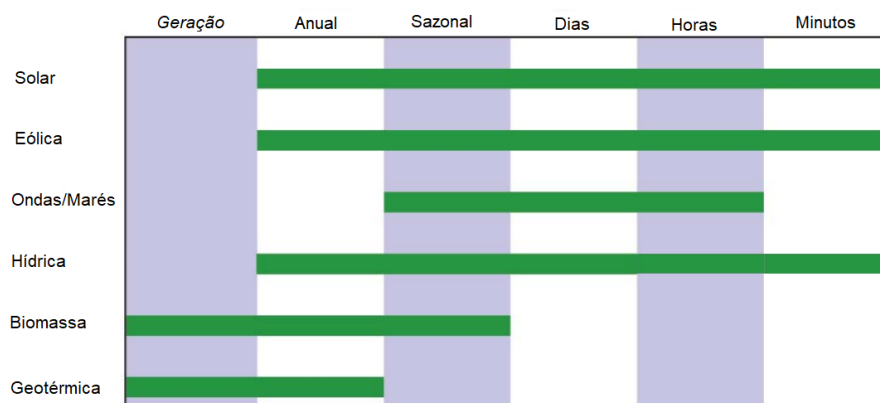


Figura 2-1: Escalas temporais relevantes, referentes aos ciclos naturais das FER (adaptado de [9]).

Como é possível observar pela Figura 2-1 produção eólica apresenta uma elevada variabilidade. Podem ocorrer variações à escala sazonal (onde os picos de produção podem ocorrer quer no Verão quer no Inverno, dependendo da região) mas também variações diárias e horárias, com perfis de produção diurnos ou noturnos que dependem, do mesmo modo, do local da instalação dos parques eólicos. No que toca à escala temporal de minutos o fenómeno mais relevante para a fonte eólica refere-se às flutuações de potência. Estas podem ser reduzidas/mitigadas através de métodos de agregação de turbinas e centrais eólica espacialmente dispersas – o efeito estatístico de suavização de produção (*power smoothing effect*), Figura 2-2 [9]-[12].

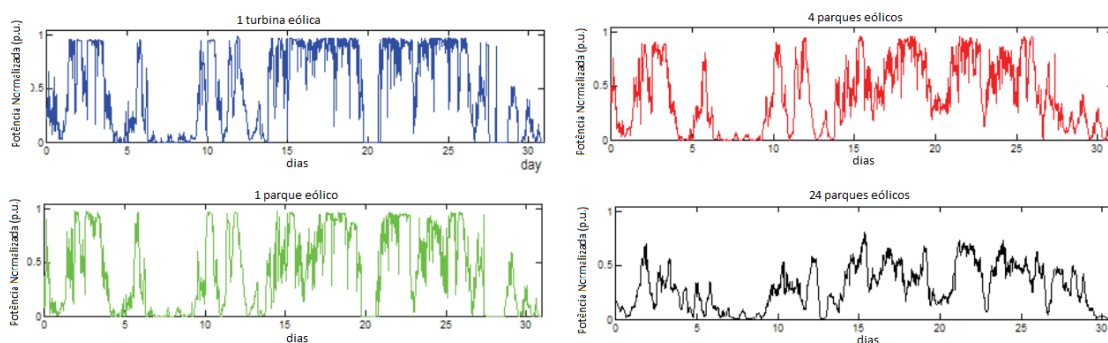


Figura 2-2: Efeito de suavização por agregação: potência entregue por 1 aerogerador (linha azul); por 1 parque eólico (linha verde); por 4 parques eólicos (linha vermelha); por 24 parques eólicos (linha preta). (Adaptado de: [11])

2.3. Efeitos da penetração eólica no sistema electroprodutor

A robustez e segurança de um sistema electroprodutor são garantidas, pelos operadores do sistema, através de um processo usualmente designado por “despacho” [13]. Este processo consiste na simultânea conjugação entre a monitorização das centrais de produção, com o agendamento prévio para o funcionamento dos centros produtores despacháveis, e a previsão de consumos. No entanto, até recentemente, a dificuldade deste processo prendia-se essencialmente com as incertezas associadas à previsão dos consumos e a eventos inesperados nos centros de produção ou na rede. Com o desenvolvimento e crescente integração na rede de FER acarretou um incremento no grau de complexidade e dificuldade no que toca à gestão do SE.

Na Figura 2-3 é possível observar um esquema onde são reunidos os vários impactos da integração de geração eólica num sistema electroprodutor de acordo com diferentes escalas espaciais e temporais.

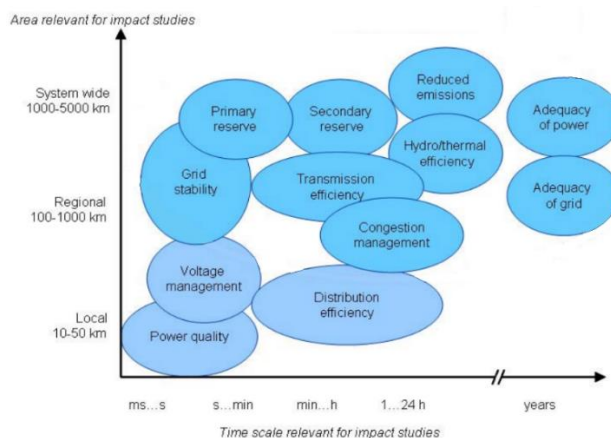


Figura 2-3: Efeitos no SE consequentes da penetração da energia eólica (adaptado de [10]).

2.3.1. Reservas de Potência

Derivado do facto de as FER estarem totalmente dependentes dos respetivos recursos primários, estas centrais nem sempre permitem uma adaptação e resposta às necessidades da rede, i.e., não oferecem garantia de potência ao SE.

Sendo essencial o equilíbrio do binómio entre o consumo e produção [6], é necessário que no sistema electroprodutor estejam integradas centrais que possuam a capacidade de ajustar a sua produção – despachabilidade – consoante as necessidades da rede e as flutuações, tanto dos recursos renováveis como do consumo, garantindo assim, em qualquer instante, um equilíbrio entre a produção e o consumo de energia. As centrais que oferecem esta garantia, ou reserva, de potência são, geralmente, centrais convencionais despacháveis que permitem ou aumentar a potência entregue à rede (*upward reserve*) ou diminuir a mesma (*downward reserve*).

No contexto de gestão de rede estas centrais onerosas que funcionam como reservas de potências vêm ganhando relevância dado que à medida que a penetração de FE num SE, requer a necessidades extra de reservas de potência (tanto para subir - *upward* - ou descer - *downward*) [14].

Para lá da incerteza associada ao consumo existem duas características associadas às centrais FER que são responsáveis pela necessidade extra de reservas de potência: a variabilidade e incerteza dessa produção, conforme representado na Figura 2-4.

A variabilidade representa variações esperadas que possam ocorrer no sistema (como, por exemplo, aumentos ou decréscimos na produção eólica) que podem, ou não, afetar a necessidade de reservas (caso essas variações ocorram dentro de escalas temporais inferiores ao agendamento prévio das centrais verifica-se uma maior necessidade de reservas de energia).

A incerteza refere-se às variações inesperadas que podem ocorrer no sistema devido, por exemplo, a erros de previsão. Estas variações inesperadas implicam que o perfil de produção/consumo seja diferente do que estaria escalonado pelo que será necessário recorrer a reservas de energia.

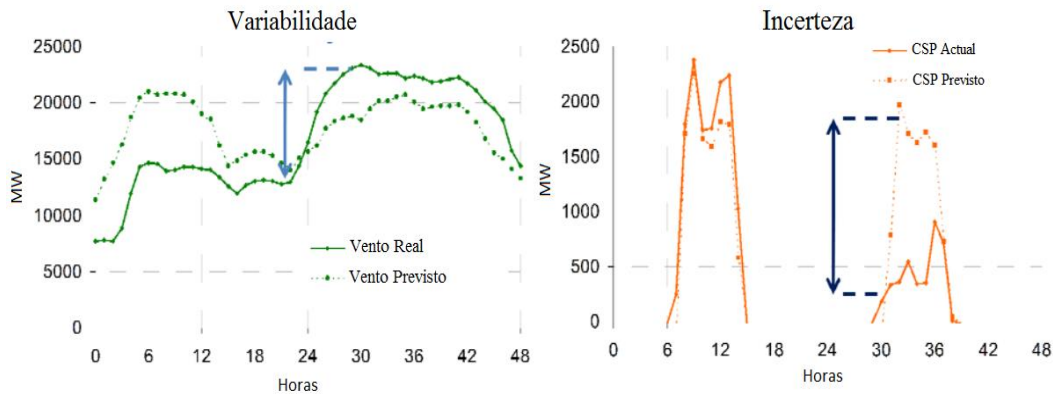


Figura 2-4: Exemplo de variabilidade e incerteza (adaptado de [14]).

É importante distinguir os diferentes tipos de reserva de potência e as suas funções [14]. Estas são um serviço auxiliar da gestão de um SE, às quais pode ser recorrer através de [15]:

- 1) Obrigatoriedade do fornecimento deste serviço da parte dos grandes produtores e outros utilizadores da rede;
- 2) Contractos bilaterais, onde o volume da reserva e a remuneração são negociados diretamente com os produtores.
- 3) Mercado, ou leilões, de Reservas.

3. Previsão de Produção Eólica – Relevância e Revisão de Literatura

Como já foi visto nas secções anteriores a energia eólica pode desempenhar um papel importante e proeminente no SE atual e até num futuro com elevadas penetrações de FER. Nesse sentido são necessárias ferramentas que possibilitem uma fácil integração destas fontes e que permitam uma gestão da rede sem escalar *i)* a complexidade da gestão da mesma e *ii)* os custos do SE. Assim, as ferramentas de previsão de produção eólica podem desempenhar um papel preponderante para a crescente integração de FER, mais concretamente de energia eólica, no sistema electroprodutor e ser uma ferramenta muito importante no auxílio de tomada de decisão do gestor da rede [16].

A previsão determinística de produção eólica, abordada nesta dissertação, é realizada num instante, t , para um horizonte temporal futuro, $t+k$. As ferramentas/modelos de previsão fornecem informação da potência média, P_{t+k} , que a turbina eólica ou o parque eólico em estudo irá fornecer à rede no instante $t+k$. Na literatura existem várias classificações para as técnicas de previsão eólica [17]. Estas podem ser classificadas quanto: ao horizonte temporal de previsão e tipo de metodologias.

3.1. Escalas Temporais

Em primeira instância é necessário ter sensibilidade para o horizonte temporal para o qual as previsões vão ser realizadas. A classificação consoante os horizontes temporais e o seu alcance variam conforme a aplicação a qual se destina a previsão. Na Tabela 3-1 encontram-se esquematizadas as quatro diferentes classes de previsões mais usuais no setor, segundo o horizonte temporal, e para que tipo de aplicações se destinam [18]-[20].

Tabela 3-1: Escalas temporais e suas aplicações para as várias classes de previsão de produção eólica.

Horizonte Temporal	Escala Temporal	Aplicações
Muito curto	Varia entre os segundos aos 30 minutos	<ul style="list-style-type: none"> ▪ Operações de despacho e regulação, em tempo real, na rede.
Curto	Varia entre 30 minutos a 6 horas	<ul style="list-style-type: none"> ▪ Impactos na determinação do preço da energia nos mercados intra-diários; ▪ Apoio à decisão sobre o estado das cargas da rede; ▪ Apoio à decisão de ligar ou desligar grupo de geradores com resposta rápida; ▪ Operações de segurança para o mercado energético.
Médio	Varia entre 6 horas e 1 dia	<ul style="list-style-type: none"> ▪ Apoio à decisão de ligar ou desligar geradores; ▪ Horizonte temporal de segurança para o mercado de energia do dia seguinte (<i>day-ahead market</i>). ▪ Impactos na determinação do preço da energia;
Longo	Mais que um dia	<ul style="list-style-type: none"> ▪ Alocação de reservas de potência; ▪ Planeamento de operações de manutenção.

No contexto deste trabalho, será analisado o horizonte temporal de previsão “curto” (*i.e.*, até 6h de previsão). A escolha prende-se com o facto de ser neste horizonte temporal que se dão alguns fenómenos relevantes como o planeamento de despachabilidade económica e operações de segurança do mercado energético, bem como a sua relevância para novos paradigmas de mercados de eletricidade que se

preveem ter horizontes temporais mais próximos do tempo real de operação. Na Tabela 3-2 encontram-se identificados alguns trabalhos que exploram o horizonte temporal em análise e os respectivos erros obtidos pelos autores.

Tabela 3-2: Revisão bibliográfica referente a previsões de velocidade do vento ou potência eólica e correspondente métrica de erro encontrada para o respetivo estudo.

[29]	Velocidade	1 hora	0.051 RMSE
		3 horas	0.074 RMSE
		5 horas	0.1 RMSE
[30]	Potência	15mins	0.0576 NRMSE
[31]	Velocidade	24h	1.39m/s RMSE
[32]	velocidade	1h	0.2583 RMSE
[33]	Velocidade	1-24h	0.3552 RMSE
			0.4192 RMSE
[34]	Potência	4h	0.211 RMSE
			0.188 RMSE
			0.24 RMSE
			0.178 RMSE
[16]	Potência	1h	0.0269 RMSE
			0.0261 RMSE
			0.0192 RMSE
			0.0218 RMSE
[36]	Potência	1h	0.05-0.075 RMSE
[37]	Potência	1h	0.538 RMSE
		3h	0.417 RMSE
		6h	0.519 RMSE

3.2. Metodologias de Previsão

Quanto às metodologias de previsão podemos distinguir na bibliografia, três grandes classes de modelos de previsão: físicos, estatísticos e os híbridos, que se encontram descritas de forma mais detalhada nas subsecções seguintes.

3.2.1. Métodos Físicos

As metodologias de previsão englobadas na classe dos modelos físicos assentam na utilização de modelos meteorológicos numéricos – *Numerical Weather Prediction* (NWP), que parametrizam e

modelam detalhadamente a atmosfera e os mecanismos de circulação na mesma. Neste tipo de modelos variáveis como as componentes do vento, temperatura, pressão, rugosidade da superfície terrestre e obstáculos são utilizadas como dados de entrada para gerar as previsões.

Este tipo de modelos existe desde 1950 quando os modelos NWP eram utilizados para fazer previsões meteorológicas com horizontes temporais na escala dos dias. Eram, no entanto, modelos muito primitivos baseados em teorias *quasi*-geostróficas onde era impossível, quer por falta de conhecimento quer por falta de recursos computacionais, incluir processos físicos relevantes (processos de radiação e transição de fase) para realizar as previsões com fiabilidade. Com o decorrer dos anos e com a substancial melhoria tecnológica estes modelos e as parametrizações que os regem foram aperfeiçoadas e os processos físicos relevantes em falta foram adicionados progressivamente. Atualmente estes modelos ainda constituem o núcleo da previsão meteorológica e evoluíram substancialmente acompanhando o crescente conhecimento acerca dos processos físicos que regem a atmosfera e a sua circulação. Os modelos NWP são, hoje em dia, menos simplistas e com parametrizações físicas mais detalhadas e precisas. Adicionalmente estes modelos beneficiaram de sistemas de aquisições de dados mais eficazes e representativos de todo o globo. Simultaneamente o horizonte temporal de previsão destes modelos foi sendo continuamente alargado e o problema da incerteza deste tipo de metodologias tem sido abordado através da incorporação de métodos estatísticos de previsão nos próprios modelos NWP [21]-[22].

A escolha de quais modelos NWP a utilizar é crucial para uma previsão fiável com base em modelos NWP. Existem dois grandes grupos de modelos NWP: os modelos globais e os modelos regionais ou de meso escala [21]. Os critérios de escolha entre ambos são a resolução espacial e temporal do modelo, a área geográfica a estudar, o horizonte temporal a estudar, a precisão da previsão, o tempo computacional e o número de ciclos a correr.

Devido à natureza física deste tipo de metodologias, a necessidade de dados e variáveis estatísticas a recolher é enorme. A precisão dos resultados destes modelos aumenta proporcionalmente ao número de dados inserido nestes modelos, bem como a qualidade destes mesmos dados [7]. Isto acarreta consequências: a parametrização destas metodologias e o processo de cálculo é complexo e difícil e tanto o tempo de cálculo como os custos de computação são altos [7]. Nesse sentido estas metodologias são utilizadas com vista a horizontes temporais de previsão longos.

Geralmente previsões que se baseiam em modelos físicos seguem três etapas distintas para a sua formulação [21]:

I. Downscaling

O conceito de *downscaling* provém da necessidade de adicionar informação mais locais/regionais aos dados extraídos dos NWP. Este processo pode ser realizado através de vários métodos estatísticos que estabelecem relações entre as variáveis locais (como velocidade do vento) e as variáveis com características de larga escala (como campos de pressão) [23]-[24]. Outra possibilidade passa pela utilização de modelos de microescala usados no setor eólico (e.g., WASP) ou modelos de dinâmica dos fluídos (*computational fluid dynamic – CFD*). Este tipo de modelos permitem trabalhar com escalas espaciais mais reduzidas e por isso são, regra geral, passíveis de melhorar as previsões NWP [21].

II. Conversão para potência

Com base nos resultados dos NWP, nomeadamente, a velocidade e direção do vento para a altura do eixo do rotor da turbina eólica – o próximo passo passa pela utilização de curvas de potência dos fabricantes das turbinas eólicas/ou estimadas com base no histórico para se obter a potência da turbina ou do parque eólico.

III. Upscaling

Caso as previsões necessárias sejam referentes a um único parque eólico esta etapa não é necessária. No entanto para os casos que visam uma área geográfica considerável ou que existe mais que um parque eólico a considerar é recomendado por diferentes autores a utilização de algoritmos estatísticos que combinam a produção prevista para um conjunto de centrais eólicas existentes numa área geográfica com vista a estimar a previsão total da área. Visto o *upscaling* tender para uma nivelção da produção regional, o erro associado à previsão tende, também, a diminuir [22].

3.2.2. Métodos Estatísticos

Para colmatar as ineficiências dos métodos físicos descritos anteriormente e, ao mesmo tempo, obter previsões operacionais com precisão adequada para gerir a variabilidade da energia eólica e as consequências que esta variabilidade acarreta na estimativa de produção várias metodologias estatísticas têm sido desenvolvidas. Estas metodologias de previsão baseiam-se na análise estatística de séries temporais de dados históricos, sejam eles de velocidade do vento ou de produção eólica. Mais precisamente este tipo de abordagens procuram estabelecer relações entre séries históricas de dados com o que se observa atualmente, no instante para o qual se quer realizar a previsão.

Em comparação com os métodos físicos de previsão, os métodos estatísticos apresentam-se como métodos com complexidade reduzida e menos dispendiosos (quer seja de tempo quer seja de recursos). Isto deve-se ao facto de nestas abordagens os processos meteorológicos não se encontrarem explicitamente expressos.

Nas subsecções seguintes encontram-se apresentados alguns métodos estatísticos comumente implementados na previsão eólica.

3.2.2.1. Persistência

A metodologia de previsão baseada na persistência é a metodologia estatística de previsão mais básica e simples de ser implementada e utilizada em previsão de produção eólica. Apesar de pertencer ao grupo de métodos estatísticos de previsão é, muitas vezes, abordado em separado. Isto deve-se ao facto de esta metodologia ser considerada como a referência, ou *benchmark*, face a todas as outras metodologias utilizadas, i.e., para estudar a viabilidade da implementação de novas metodologias de previsão os resultados obtidos através dessas têm de ser equiparados aos resultados atingidos pelo método de persistência. Apenas as metodologias que apresentem resultados mais favoráveis que os gerados pela persistência são passíveis de serem implementados.

O método de persistência tem base na suposição que a velocidade do vento, ou a potência, mantém-se igual, num instante futuro, ao valor observado no instante para o qual se realiza uma previsão. Se a velocidade e potência do vento, no instante t , forem dados por, respetivamente, v_t e p_t tem-se que a velocidade e potência no instante futuro, $t + \Delta t$, são dados por:

$$v_{t+\Delta t} = v_t \quad \text{Equação 3-1}$$

$$p_{t+\Delta t} = p_t \quad \text{Equação 3-2}$$

onde Δt corresponde ao intervalo de tempo, para o qual se quer realizar a previsão.

Para horizontes temporais de previsão muito curtos (Tabela 3-1) este modelo fornece resultados, em média, com alguma precisão. Para horizontes temporais de previsão de até seis horas esta metodologia, apesar da sua simplicidade, consegue gerar resultados melhores que muitas outras metodologias de previsão [25]. No entanto, e como seria de esperar, devido à variabilidade do vento à medida que o

horizonte temporal de previsão aumenta a precisão desta metodologia decresce. Para séries temporais com elevada flutuação, quer de velocidade quer de produção (*e.g.*, parques eólicos ou em zonas com orografia complexa), a precisão deste método é reduzida.

3.2.2.2. Métodos baseados na modelação de séries temporais

Para horizontes temporais de previsão muito curtos e curtos (Tabela 3-1) existe a possibilidade de, com certo grau de fiabilidade, utilizar-se métodos baseados exclusivamente na análise estatística de séries temporais de dados reais. Em específico, este tipo de metodologias tenta descobrir qual a relação entre uma série histórica de dados de produção, ou velocidade, e o valor da mesma no instante para o qual se quer fazer a previsão, de modo a obter previsões para os instantes seguintes. Ao contrário dos modelos físicos, neste tipo de metodologias de previsão, apenas é necessário um passo para converter os dados de entrada (*input*) em dados de saída (*output*).

Entre os vários métodos estatísticos utilizados na previsão eólica destacam-se os modelos autorregressivos (AR), média móvel (MA), autorregressivos de média móvel (ARMA), autorregressivos integrados de média móvel (ARIMA), a metodologia de Box-Jenkins e a utilização de filtros de Kalman [19] [21].

3.2.2.3. Métodos Aprendizagem Automática

Um grupo de metodologias aplicadas a problemas desta natureza são as metodologias de aprendizagem automática, também denominados por, em inglês, *machine learning* (ML). São modelos computacionais multidisciplinares que se caracterizam essencialmente pela capacidade computacional de autoaprendizagem através da experiência e treino sem que para tal exista uma programação explícita. Como tal os algoritmos *machine learning* apresentam a possibilidade de aprender e realizar previsões sob um conjunto de dados de maneira não-explicita sem seguir um conjunto de instruções estáticas de aprendizagem estatística.

O processo de um modelo de autoaprendizagem inclui alguns passos. Em primeiro lugar é necessário obter dados referentes ao passado para uma fase posterior de treino. Em segundo lugar é definido uma relação entre os dados de entrada e saída – o *input* e o *output* - desejado através de uma função alvo. O terceiro passo passa por escolher o modelo de autoaprendizagem. De seguida esse modelo sofre um processo de treino, utilizando o conjunto de dados e exemplos previamente determinados [26].

Este tipo de metodologias tendem a apresentar uma elevada eficácia na previsão [27]. No entanto, a implementação destas metodologias acarreta a desvantagem de não ser possível a descrição da relação entre os elementos do modelo, i.e., não é possível descrever ou compreender as relações encontradas, por estes modelos, entre as variáveis de entrada e as variáveis de saída [27].

Entre os métodos deste grupo destacam-se, pela sua simplicidade e eficiência, as árvores de decisão (ou *decision trees*) ou as redes neuronais artificiais (RNA) [27]-[28].

3.2.3. Métodos de previsão híbridos

Devido à natureza estocástica do vento, os métodos estatísticos de previsão necessitam de séries históricas extensas de dados para servir como base a modelos que consigam fornecer previsões precisas e fiáveis [38].

Os modelos híbridos de previsão eólica (sejam elas previsões de velocidade do vento, sejam previsões de produção) são modelos que combinam dois, ou mais, modelos de naturezas semelhantes ou distintas. Isto é: a sua génese resulta da combinação de modelos tanto de natureza estatística como de natureza física. As escalas temporais aplicáveis aos métodos de previsão também podem ser diferentes: é possível juntar métodos cujo horizonte de previsão seja diferente. De modo geral, os modelos híbridos

são compostos por um modelo linear e um modelo não linear para poder analisar as respetivas componentes lineares e não lineares da série temporal de dados.

Os métodos híbridos podem ser categorizados em quatro classes diferentes [38]:

I. Metodologias baseadas em pesos

Estas metodologias têm por base a atribuição de pesos/ponderações aos vários modelos de previsão utilizados de acordo com o seu desempenho. É uma metodologia simples, de fácil implementação e que possui a vantagem de se adaptar a novos conjuntos de dados. É uma metodologia adequada para uma larga gama de horizontes temporais de previsão. No entanto esta não garante a melhor eficácia de previsão para a totalidade do horizonte temporal de previsão e tem a necessidade de um modelo adicional para a atribuição das ponderações.

II. Metodologias baseadas em técnicas de pré-processamento de dados

Estas metodologias têm base num pré-processamento dos dados: o conjunto de dados é decomposto através de um modelo matemático e a previsão é realizada via as várias subséries de dados. Este tipo de metodologias apresenta um robusto comportamento face à natureza variável, e por vezes súbita, da velocidade do vento. Assim sendo é possível obter elevadas eficácias na previsão. No entanto este tipo de metodologias possui a desvantagem de necessitar que o utilizador compreenda o complexo modelo matemático que realiza a decomposição dos dados e a ausência de um comportamento dinâmico no sentido em que a utilização de novas séries de dados, ou a atualização da mesma, pode resultar numa lenta resposta da metodologia.

III. Metodologias baseadas em técnicas de otimização e seleção de parâmetros

Estas metodologias têm por base a otimização dos parâmetros do modelo de previsão – parâmetros meteorológicos como a temperatura, velocidade e direção do vento, precipitação, entre outros. Apesar de proporcionar ao utilizador uma maior compreensão do impacto dos diversos parâmetros na eficácia da previsão esta abordagem é de difícil implementação.

IV. Metodologias baseadas em técnicas de pós processamento de dados

Estas metodologias têm por base um pós processamento de dados de previsão. Mais concretamente esta abordagem estuda o impacto dos erros residuais das previsões obtidas, através dos modelos de previsão utilizados, na eficácia global. A eficácia das previsões, por consequência, sofre um aumento considerável face às restantes abordagens. No entanto, devido à necessidade de cálculo dos erros residuais, são necessários recursos computacionais e temporais bastante superiores aos necessários para as restantes abordagens anteriormente apresentadas.

4. Dados e Metodologias de Previsão

4.1. Tratamento de Dados

Para iniciar qualquer tipo de previsão de produção eólica é essencial obter um conjunto de dados por um período suficientemente longo de modo a poder criar um histórico de produção eólica sobre o qual é possível aplicar os diferentes métodos de previsão. Para tal recorreu-se a uma série de dados determinística, fornecida pelo Laboratório Nacional de Energia e Geologia (LNEG), de produção eólica agregada a nível nacional, no período compreendido entre o início do ano de 2015 até ao fim do ano de 2018. Estes dados foram obtidos na página de *internet* do operador do sistema em Portugal [39], com resolução temporal de quinze minutos tendo-se procedido ao cálculo da média horária.

Para mitigar o impacto da integração de novos parques eólicos no SE, os dados obtidos para a produção eólica de Portugal Continental foram normalizados. Para tal recorreu-se às estatísticas tornadas públicas, pela Direção Geral de Energia e Geologia (DGEG), relativamente aos valores mensais de potência eólica instalada em cada mês no período temporal em estudo – 2015 a 2018. Tendo em consideração os métodos de previsão utilizados no âmbito da dissertação, decidiu-se seccionar o conjunto de dados em dois componentes: os anos 2015 a 2017 com o intuito de servir como histórico de dados de produção para calibrar e testar os modelos – conjunto de treino. Os restantes dados, referentes aos dados de produção normalizada para o ano de 2018, foram utilizados com o objetivo de avaliar o desempenho das previsões – conjunto de validação. Assim, os modelos de previsão serão calibrados e testados com base nos dados de produção existentes entre 2015 e 2017, e a avaliação da performance e análise de resultados incide sobre 2018. A validação dos resultados será feita com base nas previsões obtidas face aos dados de produção verificados para o respetivo período.

Na Figura 4-1 é possível observar os perfis médios horários da produção eólica normalizada para os conjuntos de calibração e validação das metodologias de previsão a implementar na presente dissertação.

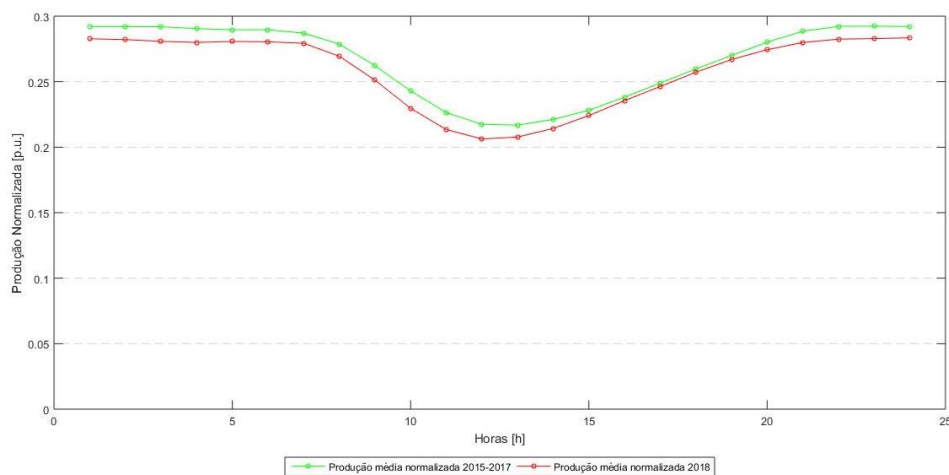


Figura 4-1: Perfil médio horário da produção eólica normalizada para os respetivos períodos temporais.

Tendo em conta com as características gerais, descritas na secção 3.2, dos vários métodos de previsão e tendo em conta também a natureza desta dissertação, optou-se por utilizar métodos de previsão de natureza estatística. Os métodos a utilizar para o presente trabalho foram selecionados, de todos os analisados/encontrados na revisão bibliográfica efetuada, segundo os seguintes critérios:

- Tempo e complexidade de computação do método de previsão;
- Complexidade de implementação do método de previsão;
- Potencial de melhoria no que toca às métricas de avaliação referentes a cada método de previsão;
- Potencial do método relativamente a atingir patamares de erros de previsão cada vez mais baixos.

Com base nestes critérios selecionaram-se os seguintes métodos: metodologia dos k vizinhos mais próximos e redes neuronais artificiais. Estes são apresentados de forma detalhada nas subsecções que seguem.

4.2. Metodologia dos k vizinhos mais próximos

O algoritmo k vizinho(s) mais próximo(s) (usualmente designado por k Nearest Neighbor - kNN) é uma variante de uma família de algoritmos que visa a resolução de problemas de otimização, dados pelo

nome de Procura pelo Vizinho Mais Próximo (em inglês *Nearest Neighbor Search* – NNS) que pertencem, por sua vez, ao conjunto de algoritmos de aprendizagem automática.

O algoritmo assume que as ocorrências históricas semelhantes entre si desencadeiam o mesmo efeito. Portanto o objetivo do algoritmo kNN consiste em encontrar, num conjunto de dados históricos as k instâncias mais semelhantes de uma instância-alvo, ou objeto, que se quer categorizar ou prever. Esta semelhança assume a forma de “distância” entre os pontos preditores e previstos, i.e., a semelhança entre os pontos preditores e previstos é calculada através de uma função de distância.

De acordo com vários autores, das principais vantagens deste algoritmo destacam-se as seguintes: i) é relativamente simples de compreender e implementar, sem deixar de conseguir resultados robustos; ii) não necessita de fase de treino realizando a sua previsão com base em valores históricos observados e iii) é uma abordagem não paramétrica, que não implica quaisquer pressupostos acerca da distribuição das variáveis a prever [40] – vertente essa importante para tratar dados de natureza não-linear. Por outro lado, algumas desvantagens da utilização deste algoritmo passam por o grande esforço computacional [41]: i) necessidade de elevada memória física do computador, visto ser necessário aceder a um conjunto de dados históricos para fazer a previsão; e ii) tempo de computação do algoritmo pode tornar-se cada vez mais longo à medida que o tamanho dos dados também incrementa.

4.2.1. Aplicações

A versatilidade do algoritmo kNN é uma das principais valências desta metodologia. É possível utilizar o algoritmo visando duas diferentes finalidades:

4.2.1.1. Classificação

A utilização do algoritmo kNN, com vista a classificação os dados de saída desejados, consiste na determinação de classes para o objeto em estudo. Conforme exemplificado na Figura 4-2 dependendo da classe a que pertencem os “k” vizinhos mais próximos (onde k representa um número definido pelo utilizador ou pelas circunstâncias do problema) é atribuída ao objeto em estudo a classe predominante nos “k” vizinhos mais próximos ou, alternadamente, uma classe com base em pesos atribuídos a cada vizinho, conforme a distância de cada instância ao objeto em estudo.

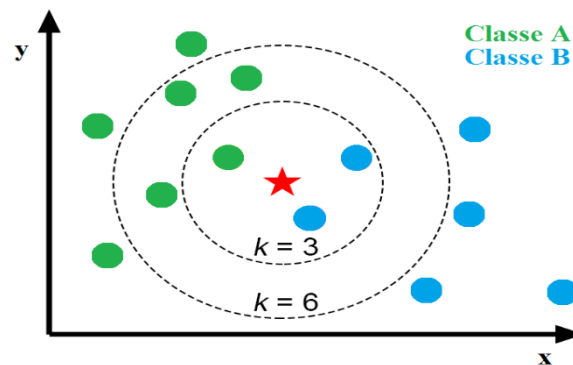


Figura 4-2: Exemplo de classificação de uma instância segundo o algoritmo de k vizinhos mais próximos. A estrela a vermelho representa o objeto em estudo; os círculos correspondem às instâncias mais próximas do objeto em estudo.

No exemplo presente na Figura 4-2 a classe atribuída ao objeto em estudo, para k igual a 3 é Classe B, enquanto, quando k assume o valor 6, a classe atribuída é Classe A.

O processo de atribuição de classe ao alvo baseia-se na representação da classe predominante, ou mais frequente, observada na vizinhança do alvo [40]. Para cada classe, j, é possível calcular a probabilidade de o alvo pertencer a essa j-ésima classe, por meio da Equação 4-1:

$$P(Y = j|X = x^*) = \frac{1}{k} \sum_{x_i \in N_k(x)} I(y_i = j) \quad \text{Equação 4-1}$$

onde x^* representa a instância alvo, pertencente ao conjunto de dados de teste; k o número de vizinhos a considerar; $N_k(x)$ representa a vizinhança definida pelo número de vizinhos, k ; x_i os respectivos vizinhos, presentes no conjunto histórico de dados.

4.2.1.2. Regressão

A utilização do algoritmo k NN com vista a previsão consiste na atribuição de um valor, resultante de uma regressão, ao objeto alvo de estudo. Conforme exemplificado na Figura 4-3 o valor atribuído ao objeto em estudo (caixa de bigodes azul) é determinado com base nos dados observados da velocidade do vento – quadrados azuis – de acordo com as k instâncias mais próximas – circunferências azuis – do valor observado na série de referência – círculo azul. Este procedimento permite estimar a velocidade do vento com base em dados da velocidade do vento medidos num determinado local e na série de referência (longo termo) [42].

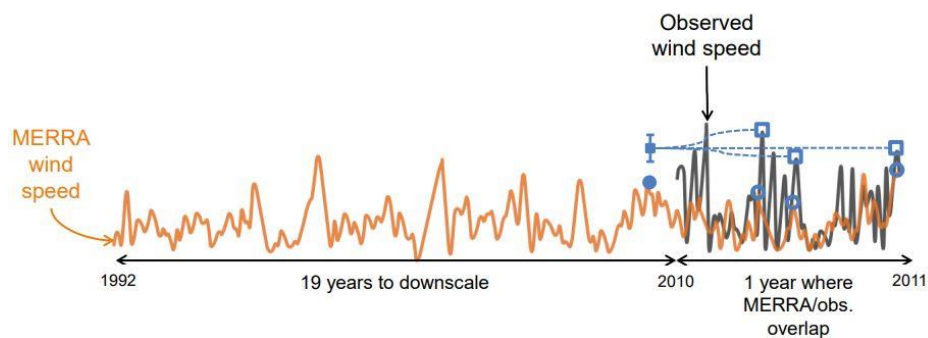


Figura 4-3: Exemplo da utilização do algoritmo k NN com vista a regressão. Retirado de [42].

O processo de atribuição de um valor ao alvo baseia-se no cálculo de uma média entre as k instâncias mais semelhantes encontradas, segundo a Equação 4-2:

$$f(x) = \frac{1}{k} \sum_{x_i \in N_k(x)} y_i \quad \text{Equação 4-2}$$

onde k representa o número de vizinhos a considerar; $N_k(x)$ representa a vizinhança, definida pelo número de vizinhos, k ; x_i os vizinhos, presentes no conjunto histórico de dados; y_i o atual valor dos vizinhos a considerar;

De notar que deste modo, implicitamente, tem-se que todas as k instâncias mais próximas encontradas pelo algoritmo k NN possuem a mesma relevância no que toca à previsão, independentemente da semelhança ou distância individual efetiva ao alvo, que pode apresentar grandes variações.

Na Figura 4-4 pode-se observar uma exemplificação da problemática acima descrita.

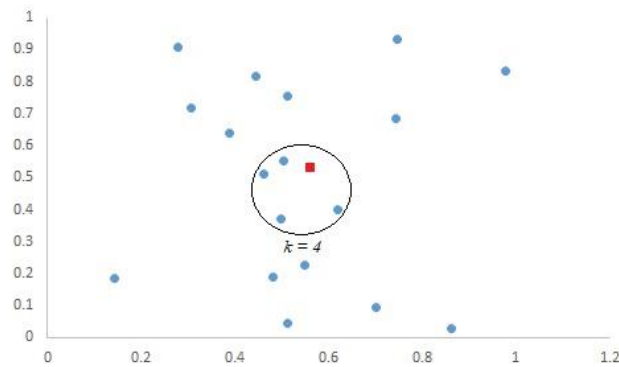


Figura 4-4: Exemplo da variação da semelhança, ou distância, entre as k instâncias mais próximas. Os círculos amarelos representam os dados pertencentes ao conjunto histórico de dados; o quadrado a vermelho representa a instância alvo, pertencente ao conjunto de teste.

Neste caso, assumindo k igual a 4, os pontos mais próximos do dado alvo encontram-se selecionados e serão estas quatro instâncias as que vão ser utilizadas para obter uma previsão para o alvo. Existe, no entanto, dentro das quatro instâncias mais semelhantes alguma variação entre elas: não se encontram, as quatro, à mesma distância ao alvo.

4.2.2. Principais passos do algoritmo

Na Figura 4-5 é possível observar um diagrama que engloba os principais passos do algoritmo.

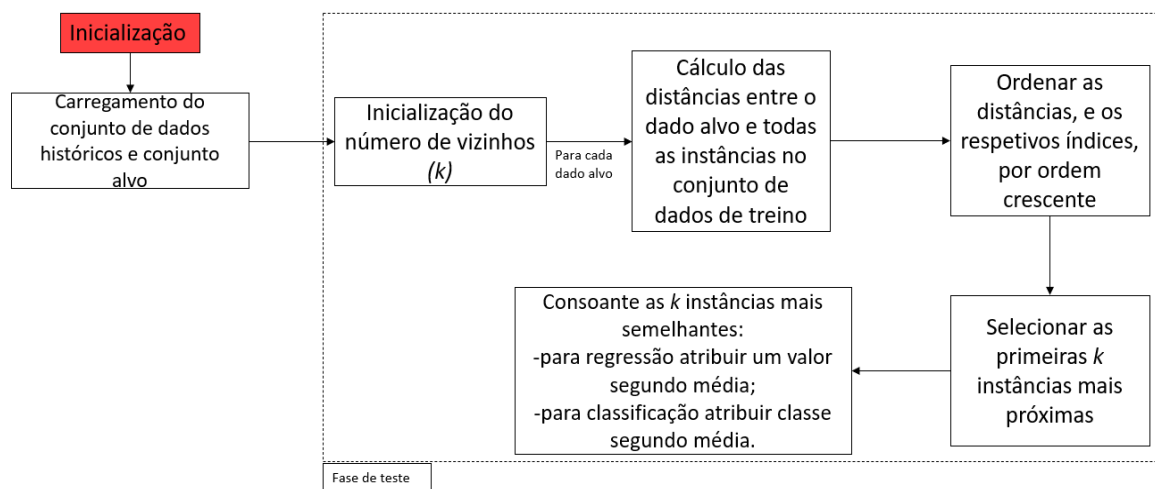


Figura 4-5: Fluxograma referente ao funcionamento do algoritmo dos k vizinhos mais próximos.

É possível dividir a aplicação deste algoritmo em duas fases distintas [43]:

- **Fase de construção de base de dados históricos:** esta fase consiste no carregamento e armazenamento do conjunto dos dados históricos e as suas características – no caso da utilização com vista a classificação as características guardadas são as classes enquanto no caso de regressão as características são os seus valores.
- **Fase de previsão/classificação:** esta consiste em, para uma nova instância, encontrar as k instâncias mais próximas, no conjunto histórico de dados, acima referido. Essa instância é então classificada ou é-lhe atribuída um valor, com base nessas instâncias mais próximas, dependendo se a utilização do algoritmo visa a classificação ou regressão respetivamente.

4.2.3. Definição do número de vizinhos

O número de vizinhos, k , é um dos parâmetros a ter em conta para a calibração do algoritmo k NN, sendo determinante para a eficácia e precisão dos resultados obtidos. Contudo, a definição *a priori* do valor ótimo para k não é possível uma vez que este varia de caso para caso e, mais especificamente, a variação tem uma grande dependência do tamanho do conjunto de dados utilizados [44]-[47]. Este parâmetro, k , determina a ajustabilidade do modelo para a geração de dados futuros, i.e., previsões, e tem uma relação direta tanto com o viés como a variância dos resultados obtidos [40].

O número de vizinhos, k , tem sempre de assumir um valor inteiro pelo que o valor mínimo possível para este parâmetro é 1. O número máximo para k é igual a N , que representa o total de dados presentes no conjunto histórico de dados [40]. Geralmente, no caso de k ser reduzido a variância dos resultados da previsão assume um valor elevado enquanto o viés assume um valor baixo. No caso contrário, em que k tende para N , a variância apresenta um valor reduzido e o viés assume um valor mais elevado do que a o obtido comparativamente a k igual a 1. Assim, à medida que o valor de k tende para o seu mínimo, o algoritmo torna-se cada vez mais instável e vulnerável ao ruído ou discrepâncias presentes no conjunto histórico de dados. Na Figura 4-6 é possível observar a exemplificação dessa instabilidade.

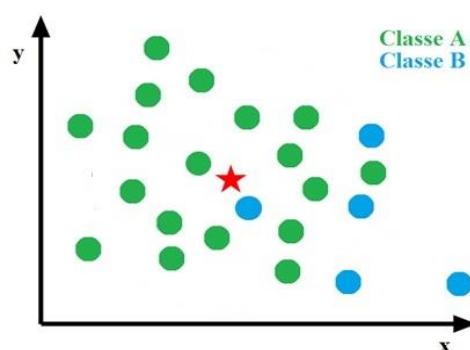


Figura 4-6:Exemplo da instabilidade do algoritmo k NN, com vista a classificação, provocada por um valor mínimo para o parâmetro do número de vizinhos mais próximos (k). A estrela a vermelho representa o dado alvo a classificar; os círculos a verde e vermelho correspondem a dados do conjunto histórico de dados de duas classes distintas: classe A e classe B, respetivamente.

No exemplo da Figura 4-6 é possível observar que a maior parte dos pontos que rodeia o objeto, ou dado alvo, pertence à classe A. No entanto o ponto, pertencente ao conjunto histórico, mais próximo do dado alvo pertence à classe B. Quando o parâmetro k assume o seu valor mínimo, um, o algoritmo tende a atribuir a classe B pois é a classe a que pertence o vizinho mais próximo. No entanto, seria de esperar que, pelo maior número de dados pertencentes à classe A envolventes do dado alvo, a classe a atribuir seria a classe A. Contrariamente tem-se que à medida que o número de vizinhos a encontrar, k , aumenta o problema de instabilidade é solucionado. Para a utilização do algoritmo k NN com vista a classificação existe uma regra comumente utilizada, em inglês (*rule of thumb*), que determina que o valor ótimo para o número de vizinhos a encontrar, k , é dado pela raiz quadrada do tamanho do conjunto histórico de dados [44]. Contudo, na utilização do algoritmo k NN, é sempre desejável a realização de uma análise de sensibilidade para otimização deste parâmetro – quer seja em problemas de natureza de regressão, como o presente trabalho, ou de classificação.

4.2.4. Semelhança entre instâncias

O principal objetivo do cálculo da distância entre pontos, como já mencionado, é o de encontrar, para as instâncias presentes no conjunto de dados de teste, as instâncias, pertencentes ao conjunto histórico de dados, que apresentam o maior grau de semelhança. A semelhança, neste caso, é dada pela distância entre os dois pontos – a instância alvo pertencente ao conjunto histórico de dados face aos dados

presentes no conjunto de teste. A previsão efetuada pelo algoritmo vai depender da distância do alvo às k -instâncias mais próximas, conforme a Equação 4-3:

$$Previsão_{t,t+h} = \underset{h=1}{\operatorname{argmin}} \sum^H (\|d(X_h, Y_{t,h})\|) \quad \text{Equação 4-3}$$

onde X_h representa os dados presentes no conjunto histórico de dados; $Y_{t,h}$ representa os dados alvo a considerar para realizar as previsões presentes no conjunto de dados de teste; $d(X_h, Y_{t,h})$ representa a distância efetiva entre X_h e $Y_{t,h}$; a previsão corresponde ao argumento cujo valor da distância entre o dado alvo e todas as instâncias presentes no conjunto histórico de dados apresenta o valor mínimo. O resultado de $d(X_h, Y_{t,h})$, pode ser obtido com diferentes formulações para o cálculo da distância entre dois pontos.

Em [45]-[46] os autores demonstram que a escolha da metodologia utilizada para o cálculo da distância pode ser crucial para a performance do algoritmo e eficácia dos resultados obtidos. Na secção 4.2.6.3 são discutidas as metodologias utilizadas no cálculo das distâncias adotadas na presente dissertação.

4.2.5. Ponderação

Esta variante, também denominada, em inglês, *weighted k-nearest neighbors* (*wkNN*), oferece uma alternativa à problemática da variação da semelhança entre as k instâncias pré-selecionadas para realizar a previsão e tenta ajustar a solução gerada com a atribuição de pesos para cada uma das k instâncias mais semelhantes [48]. Sugerida e desenvolvida em 2007 para utilizações de previsão por *Troncoso et al* [49], a premissa desta extensão do algoritmo kNN baseia-se na atribuição de pesos a cada k instância semelhante, presente no conjunto histórico de dados, ao alvo.

A atribuição de pesos pode ser realizada de várias maneiras. Em [50] é possível encontrar uma descrição pormenorizada das várias possibilidades no que toca a funções de ponderação. Na secção 4.2.6.5 é discutida a abordagem tida na presente dissertação, no que toca à utilização desta metodologia (*wkNN*).

4.2.6. Implementação do método dos k vizinhos mais próximos

No âmbito da dissertação foi implementado um algoritmo para realizar as previsões da metodologia kNN com recurso ao *software* Matlab®. Na Figura 4-7 encontram-se esquematizados os principais passos da metodologia de implementação e previsão utilizando o algoritmo kNN . Na presente dissertação, após cada previsão este período era adicionado ao conjunto de dados históricos.

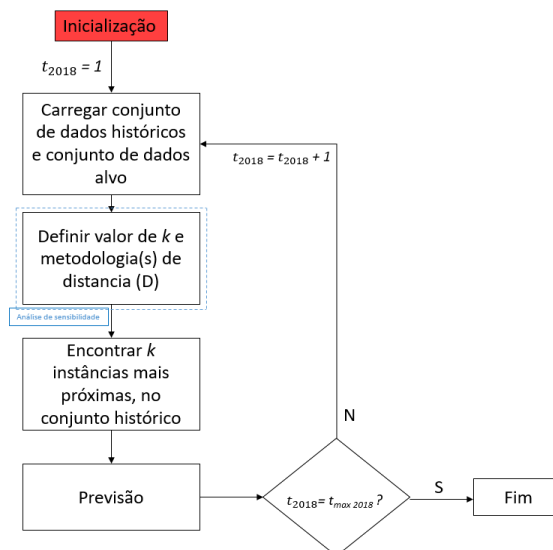


Figura 4-7: Diagrama com a esquematização dos principais passos da metodologia do algoritmo dos *k*NN.

4.2.6.1. Matriz Histórico

A Matriz Histórico, nome dado ao conjunto histórico de dados de treino para o algoritmo *k*NN, foi definida segundo a Equação 4-1:

$$Matriz\ Histórico = \begin{bmatrix} p_{t-HT} & p_{t-(HT-1)} & \cdots & p_{t-(HT-5)} \\ \vdots & \vdots & \vdots & \vdots \\ p_{t_{max}-HT} & p_{t_{max}-(HT-1)} & \cdots & p_{t_{max}-(HT-5)} \end{bmatrix} \quad \text{Equação 4-4}$$

onde p_t simboliza a produção eólica normalizada para o instante t , instante para o qual se quer realizar a previsão; HT como sendo o horizonte temporal de previsão – 6 horas. O valor de t vai sendo iterativamente atualizado: inicia-se no momento para o qual se quer realizar a primeira previsão do ano, i.e. após as primeiras seis horas do ano para o qual se quer realizar a previsão – 2018 –, tomando o valor de 7. Na segunda previsão o valor de t irá ser incrementado em um e assim sucessivamente até, conforme esquematizado na Figura 4-7, o t_{max} para o respetivo período de tempo ser atingido. O valor de t_{max} é dado por:

$$t_{max} = NH - HT - 1 \quad \text{Equação 4-5}$$

onde NH representa o número de registos horários disponíveis. Do mesmo modo, NH vai sendo atualizado iterativamente: inicialmente engloba apenas os anos de 2015 a 2017, no entanto a cada iteração, i.e., previsão, as seis horas seguintes – pertencentes ao ano de 2018 – vão sendo adicionadas ao conjunto histórico de dados, utilizando o índice t_{2018} conforme esquematizado na Figura 4-7. Deste modo este conjunto vai sendo aumentado e atualizado com vista a incrementar o alcance da procura dos vizinhos e possibilitando aumentar a eficácia da previsão.

4.2.6.2. Matriz Alvo

A Matriz Alvo, ou *target matrix*, corresponde ao conjunto de dados de teste para o algoritmo *k*NN, na Figura 4-7 identificado como conjunto de dados alvo. Esta matriz possui a mesma estrutura que a Matriz Histórico com a diferença de possuir os dados de produção eólica normalizada, para o período de previsão: o ano de 2018. O tamanho da matriz alvo tem de estar em concordância com o horizonte temporal para o qual se quer realizar a previsão – seis horas. A Matriz Alvo é definida pela Equação 4-3, segundo as variáveis já discriminadas para a Equação 4-1:

$$Matriz\ Alvo = \begin{bmatrix} p_t & p_{t+(HT-5)} & \cdots & p_{t+(HT-1)} \\ \vdots & \vdots & \vdots & \vdots \\ p_{t_{max}} & p_{t_{max}+(HT-5)} & \cdots & p_{t_{max}+(HT-1)} \end{bmatrix} \quad \text{Equação 4-6}$$

onde NH neste caso, como já foi dito, compreende o período de tempo para o qual se quer realizar as previsões utilizando o algoritmo *k*NN: o ano de 2018. Ao contrário do que acontece com a Matriz Histórico esta não sofre qualquer alteração de tamanho a cada iteração.

4.2.6.3. Métricas de Distância

Conforme referido na secção 4.2 a escolha das métricas para o cálculo das distâncias a implementar no algoritmo *k*NN é um dos fatores mais importantes a ter em conta para a obtenção de previsões. Na presente dissertação optou-se por adotar três diferentes tipos de metodologias de cálculo da distância entre pontos e estudar o conseqüente efeito nas previsões obtidas através do algoritmo *k*NN: distância de Minkowski, distância euclidiana e distância por correlação. Todas as métricas de utilizadas na presente dissertação encontram-se previamente implementadas na ferramenta computacional em utilização.

Para dois vetores, $p = (p_1, p_2, \dots, p_n)$ e $q = (q_1, q_2, \dots, q_n)$ de características de dimensão n , têm-se:

4.2.6.3.1. Distância de Minkowski

A distância de Minkowski (L_p), também denominada por *p-norm distance* [51], enquadra-se num tipo de família de métricas de distâncias entre dois pontos, onde três casos especiais existem, cada um com a sua ordem p (Equação 4-7). Quando p assume o valor de:

- $1 \rightarrow$ a distância de Manhattan, também conhecida por norma L_1 , distância de quarteirão (em inglês *city block*), norma *Taxicab* [45].
- quando p tende para infinito tem-se em mãos a distância de Chebychev, também conhecida distância de tabuleiro de xadrez (em inglês *chessboard distance*) [45], Lagrange [50] ou por distância de máximo valor [45]. Esta representa a máxima distância, entre dois vetores, em qualquer eixo de coordenadas.
- quando p assume o valor de 2 tem-se a distância Euclidiana, ver secção 4.2.6.3.2.

A distância Minkowski é definida pela Equação 4-7:

$$d(p, q)_{MINKOWSKI} = \sqrt[p]{\sum_{i=1}^n |p_i - q_i|^p} \quad \text{Equação 4-7}$$

onde p é um valor positivo e representa a ordem da distância.

No âmbito da presente dissertação optou-se por realizar uma análise de sensibilidade para o cálculo da ordem p da distância da Minkowski mais indicada para o conjunto de dados e objetivo da dissertação.

O teste consistiu em verificar qual a distância obtida entre os dados de produção normalizada presentes no conjunto histórico de dados e a origem, i.e. ponto de origem de um referencial, enquanto se variava a ordem p da distância de Minkowski entre 1 e 100. Os resultados obtidos para a distância entre os dados históricos e a origem foram analisados no sentido de verificar qual o ponto de equilíbrio entre o valor p e a convergência da distância. Nesse sentido, foi identificada a menor ordem de p que satisfaça a seguinte condição de erro imposta (Equação 4-8):

$$Erro(p) < 1 - \frac{\text{distância Minkowski}_{p=100}}{\text{distância Minkowski}_p} \quad \text{Equação 4-8}$$

O valor de Erro (p) foi considerado igual a 5%. O fator p , resultante da análise de sensibilidade foi implementado no algoritmo e aplicado em todas as previsões referentes à metodologia do algoritmo k NN para a distância de Minkowski.

4.2.6.3.2. Distância Euclidiana

A distância Euclidiana é o terceiro caso especial da família de distâncias de Minkowski de norma (L_p). Também chamada por distância de régua (em inglês *ruler distance*) ou por norma L_2 [45]. Esta é uma extensão do Teorema de Pitágoras e é dada pela Equação 4-9:

$$d(p, q)_{EUCLIDIANA} = \sqrt{\sum_{i=1}^n |p_i - q_i|^2} \quad \text{Equação 4-9}$$

4.2.6.3.3. Distância por Correlação

A distância por correlação é uma medida de dependência estatística entre dois vetores. Também chamada por distância de Pearson [45] esta é derivada do coeficiente de correlação de Pearson, que mede a relação linear entre dois vetores. Esta métrica de distância é dada pela subtração valor unitário ao coeficiente de Pearson, conforme a Equação 4-10:

$$d(p, q)_{CORRELAÇÃO} = 1 - \frac{\sum_{i=1}^n (p_i - \bar{p})(q_i - \bar{q})}{\sqrt{\sum_{i=1}^n (p_i - \bar{p})^2 * \sum_{i=1}^n (q_i - \bar{q})^2}} \quad \text{Equação 4-10}$$

onde $\bar{p} = \frac{1}{n} \sum_{i=1}^n p_i$ e, do mesmo modo, $\bar{q} = \frac{1}{n} \sum_{i=1}^n q_i$.

4.2.6.4. Número de Vizinhos

Como já foi referido na secção 4.2.3 o parâmetro k , que determina o número de vizinhos a considerar para a realização de previsões, é o único que necessita de otimização. Este é, também, *case dependent* pelo que não existe uma regra geral para a escolha de um valor rigoroso para qualquer problema.

Nesse sentido optou-se por estudar a variação do número de vizinhos a encontrar e qual a configuração do algoritmo k NN que permite obter previsões com maior grau de fiabilidade e eficácia: fez-se variar o parâmetro k entre os valores 1, 5, 10, 20 e 162. O último valor testado para k tem origem na *rule-of-thumb* geral descrita acima (ver secção 4.2.3) onde o valor de k pode ser igual à raiz quadrada do tamanho dos dados.

4.2.6.5. Previsão

Conforme discutido na secção 4.2.5 a implementação do wk NN poderá trazer vantagens para as soluções obtidas. No âmbito da dissertação optou-se então por verificar qual o impacto da atribuição de pesos aos k vizinhos mais próximos encontrados para as previsões.

Nesse sentido estabeleceu-se que:

- Para quando k assume o valor de 1, a previsão fosse dada pelo padrão de produção identificado para o respetivo vizinho encontrado no conjunto de dados de histórico;
- Para valores de k superiores a 1 optou-se por comparar dois casos distintos:
 - No primeiro caso foi atribuído um igual peso a cada um dos k vizinhos mais próximos, i.e. contribuindo todos eles com igual peso para a previsão obtida. A previsão obtida, neste caso, é calculada através de uma média aritmética entre os respetivos k vizinhos;
 - No segundo caso estudado foram atribuídas, a cada um dos k vizinhos encontrados, ponderações distintas. Das várias distâncias existentes na literatura, na presente dissertação optou-se por implementar uma função com base no inverso da distância euclidiana, d , onde a cada vizinho, k_i , é atribuído um peso, w_{k_i} , para a previsão obtida, inversamente proporcional à sua respetiva distância euclidiana ao alvo. Deste modo quando a distância é mínima o peso atribuído vai ser maior. À medida que o valor da distância aumenta o peso atribuído à instância para o cálculo da previsão a solução gerada é cada vez menor. Isto significa que à medida que a semelhança entre a instância alvo e as instâncias pertencentes ao conjunto de teste aumenta, o peso que efetivamente essa instância tem na previsão/solução gerada é cada vez mais significativo. Na Equação 4-11 tem-se a respetiva função implementada para a ponderação:

$$w_{k_i} = \frac{1}{\sum \frac{1}{|d_{k_i}|}}$$

Equação
4-11

4.3. Metodologia das Redes Neurais Artificiais

A primeira menção bibliográfica de uma rede neuronal artificial (RNA) foi em 1943 por Warren Culloch e Walter Pitts que desenvolveram o primeiro modelo conceptual [52]. No seu trabalho descrevem-se os conceitos de um neurónio interligados entre si, formando uma rede de neurónios inteiramente interligada.

As redes neuronais artificiais (RNA) são uma abordagem transversal a múltiplos campos de conhecimento que se fundamenta na estrutura neuronal de um cérebro humano, replicando a sua capacidade de processar informação. O seu funcionamento baseia-se em dois princípios bastante semelhantes ao cérebro humano: i) o conhecimento adquire-se através de um processo de aprendizagem ou treino; ii) o funcionamento do sistema baseia-se na interligação dos seus elementos. Simulando o comportamento do cérebro humano as RNA são métodos estatísticos não-lineares complexos composto por várias unidades de processamento, ou neurónios, interligadas entre si. Essas ligações aliadas ao algoritmo de aprendizagem – função que permite ajustar a conexão entre neurónios – permitem adquirir informação através da experiência e treino. Esta capacidade de aprendizagem distingue esta técnica de métodos estatísticos convencionais, permitindo a sua aplicação em diversas áreas, nomeadamente em ferramentas de previsão [53], [54].

A utilização de RNA oferece ao(s) seu(s) utilizador(es) as seguintes propriedades [55]:

- i. Não linearidade: esta é uma propriedade importante, principalmente se o mecanismo físico por detrás dos dados de entrada tiver uma natureza não-linear, como é o caso da previsão da energia eólica. Um neurónio pode ser de natureza linear ou não linear e uma rede formada pela interligação de neurónios não lineares é, em si, não linear. A natureza do neurónio, i.e. linear ou não linear, é determinada pela sua função de transferência (ver secção 4.3.1.3). Para obter relações não lineares entre os dados de entrada e os de saída é necessário existir, na constituição da rede, pelo menos um neurónio de natureza não linear.
- ii. Mapeamento de Entrada-Saída: ilustrativo da possibilidade de treino/aprendizagem supervisionada (ver secção 4.3.3). Aqui a rede sofre um processo contínuo e iterativo de preparação onde um ou mais que um conjunto de dados (tanto de entrada como de saída) são apresentados à rede, de modo que esta crie um mapeamento/relação das conexões entre os dados de entrada e saída para o problema. Este processo repete-se até que não haja ajustes significativos nas ligações. Esta propriedade confere uma dimensão não-paramétrica à rede: não existem quaisquer premissas à distribuição dos dados de entrada antes do processamento dos mesmos.
- iii. Adaptabilidade: as RNA possuem a capacidade interna de ajustarem as ligações da própria rede em função do ambiente e dos dados apresentados. Independentemente do treino prévio que sofreram estas redes podem, em resposta a determinado ambiente, sofrer novamente um treino que ajuste os pesos entre os seus neurónios adequadamente ao novo ambiente.
- iv. Tolerância a falhas: a robusta capacidade de computação das RNA não é significativamente afetada em casos de falhas de um ou mais neurónios, existe sempre outros caminhos de ligação entre neurónios para o fluxo de informação/dados.

4.3.1. Elementos de uma RNA

Os principais elementos constituintes de numa RNA são:

4.3.1.1. Neurónios

Os neurónios correspondem aos elementos mais básicos, mas também mais importantes no funcionamento de uma RNA. Estes são os elementos responsáveis pelo processamento de informação na rede. A Figura 4-8 ilustra a estrutura geral de um neurónio.

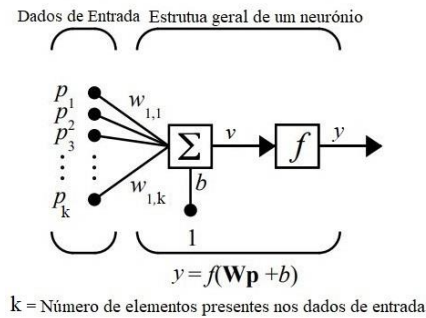


Figura 4-8: Modelo de um neurónio (adaptado de [52]).

O esquema de funcionamento dos neurónios segue as seguintes etapas, ilustradas na Figura 4-9 [56]:

- Em primeiro lugar cada dado de entrada (*input*), p_R , é multiplicado pelo respetivo peso, $w_{1,k}$;
- É adicionado ao *input* ponderado (i.e. com o devido peso em fator) o viés, b_k , para formar os dados de entrada da rede, v_k .
- O neurónio verifica se o valor resultante, n , atinge ou não um valor pré-determinado pelo utilizador – o limiar. Se o valor do limiar não for atingido o sinal deste neurónio não é enviado; se o valor do limiar for atingido passa para a camada de saída. Este processo de verificação é denominado por função transferência, $f(n_k)$.
- Dependendo da função de transferência utilizada a amplitude dos sinais, v_k , é limitado entre $[0,1]$ ou $[-1,1]$, e permite obter o sinal de saída, y_k . Este processo é similar ao “disparo” de um neurónio do cérebro.



Figura 4-9: Esquema representativo das etapas durante o processamento de dados levado a cabo por um neurónio.

As equações matemáticas, Equação 4-12 e Equação 4-13, que descrevem o neurónio k são [55]:

$$v_k = \sum_{j=0}^m (p_j w_{kj}) + b_k \quad \text{Equação 4-12}$$

$$y_k = f(v_k)$$

Equação
4-13

4.3.1.2. Ligações e pesos

As ligações entre neurónios é a característica que permite aos neurónios comunicarem informação entre si. Os pesos servem como ponderação dos sinais de entrada para o respetivo neurónio. Durante a fase de calibração da rede os neurónios vão transmitindo informação entre eles e os pesos das ligações entre os respetivos neurónios vão sendo adequadamente ajustados. O processo de calibração da rede – i.e. o ajuste dos pesos entre neurónios – pode ser de dois tipos: supervisionado ou não supervisionado (ver secção 4.3.3)

4.3.1.3. Função Transferência

Também denominada, alternadamente, por função de ativação. A função transferência toma forma, como já se viu, entre os dados de entrada e os dados de saída de cada neurónio. Existem dois tipos: as funções lineares e as não-lineares. Na Figura 4-10 encontram-se representadas as funções transferência mais utilizadas:

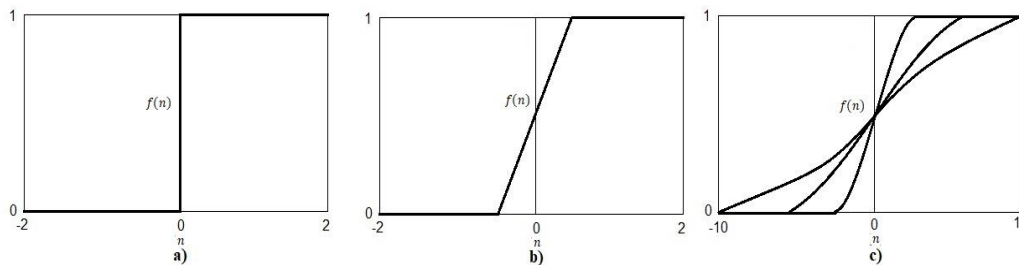


Figura 4-10: Representação das funções de transferência mais utilizadas: a) Função de Transferência Limiar; b) Função de Transferência Linear; c) Função de Transferência Sigmoide, com diferentes declives. Adaptado de [56]

4.3.1.3.1. Função transferência limiar

Também denominada função *Hard Limit* [52] ou Heaviside [57]. Esta função, representada na Figura 4-10 a), é uma das funções de ativação mais simples. Consiste na utilização de um limiar, ou limite, – θ – para avaliar os dados de entrada. Casos estes sejam superiores ou inferiores ao limite imposto o neurónio é, ou não, respetivamente, ativado. Caso o seja, o sinal é enviado para a camada seguinte. Esta função é também chamada de *squashing function* uma vez que esta limita a amplitude do sinal de saída do neurónio, a_k , num intervalo definido e limitado: $[0,1]$ ou alternativamente $[-1,1]$. A Equação 4-14 define matematicamente a respetiva função limiar:

$$f(v) = \begin{cases} 1, & \text{se } v \geq \theta \\ 0, & \text{se } v < \theta \end{cases} \quad \text{Equação 4-14}$$

4.3.1.3.2. Função transferência linear

Esta função encontra-se representada na Figura 4-10 b). O esquema de funcionamento deste tipo de funções de transferência baseia-se na multiplicação dos dados de entrada pelos pesos de cada neurónio que por sua vez é gerado um sinal de saída que é proporcional ao sinal de entrada. Esta é uma opção mais viável que a opção mais simplista – a função transferência limiar – pelo facto que com esta função transferência é possível obter um sinal de saída com múltiplos valores em vez de apenas um. A Equação 4-15 define matematicamente a respetiva função linear:

$$f(v) = \begin{cases} 1, & \text{se } v \geq +\frac{1}{2} \\ v, & \text{se } +\frac{1}{2} > v > -\frac{1}{2} \\ 0, & \text{se } v \leq -\frac{1}{2} \end{cases} \quad \begin{array}{l} \text{Equação} \\ 4-15 \end{array}$$

4.3.1.3.3. Função transferência sigmoide

A função de transferência sigmoide, caracterizada por um gráfico em forma da letra “s”, é a função mais utilizada na construção de uma RNA, devido ao facto de ser uma função estritamente crescente que exhibe um balanço entre a sua linearidade e não linearidade [55]. Um exemplo de uma função transferência sigmoide é a função logística, definida pela Equação 4-16:

$$f(v) = \frac{1}{1 + \exp(-av)} \quad \begin{array}{l} \text{Equação} \\ 4-16 \end{array}$$

onde a representa o parâmetro de declive da função sigmoide. Alterando o valor de a obtém-se os vários declives, como exemplificado na Figura 4-10 c). À medida que a tende para infinito a função passa a assemelhar-se a uma função limiar. No entanto as diferenças entre estas duas funções prendem-se com o facto de enquanto a função limiar assumir dois valores apenas, 0 ou 1, a função sigmoide pode assumir qualquer valor situado entre 0 e 1; a segunda diferença prende-se com a diferenciabilidade da função: enquanto a função limiar não é diferenciável, a função sigmoide é. Esta última sendo uma característica importante para a teoria das RNA.

4.3.2. Topologia da Rede

Nas secções anteriores foram apresentados os benefícios da utilização de RNA bem como os seus constituintes e características. No entanto, interessa também saber como se organizam esses mesmos elementos. O agrupamento de neurónios ocorre por camadas, existindo sempre uma de entrada e de saída e, entre elas, uma ou mais camadas escondidas. As camadas podem, então, ser classificadas de três modos:

- Camada da entrada: onde os dados são apresentados à rede;
- Camadas escondidas: ou camadas intermédias - onde ocorre a maior parte do processamento de dados;
- Camada de saída: onde o processamento de dados é finalizado e o resultado final é apresentado.

No caso de uma RNA ser composta por mais do que uma camada escondida, entre a camada de entrada e a de saída, a rede tem a designação de rede multicamada, ou em inglês *multilayer network*. Na Figura 4-11 é possível observar esquematicamente a diferença entre uma rede com apenas uma camada e uma rede multicamada.

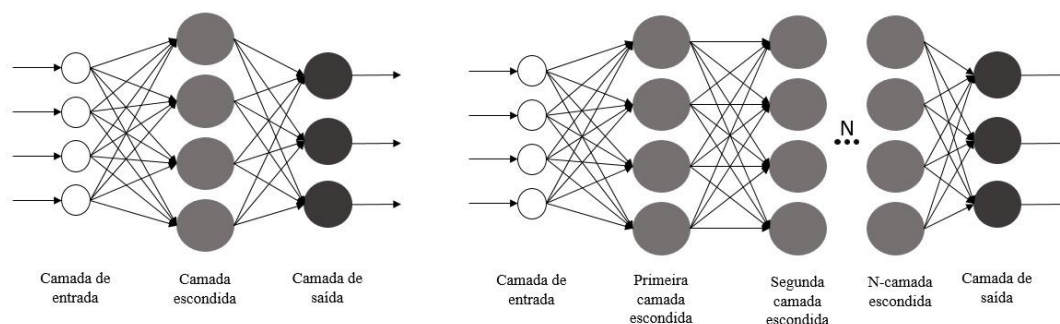


Figura 4-11: Esquema representativo de uma rede neuronal artificial com uma camada (esquerda) e multicamada, com n -camadas escondidas (direita). Cada nódulo circular representa um neurónio artificial e cada seta entre camadas representa uma ligação entre a saída (*output*) de um neurónio e a entrada (*input*) de outro neurónio, pertencente à camada seguinte.

De acordo com [55] e [58] e é possível classificar as RNA segundo a aglomeração dos neurónios em camadas e as consequentes ligações existentes entre essas camadas. Na Figura 4-12 encontra-se esquematizada essa classificação.

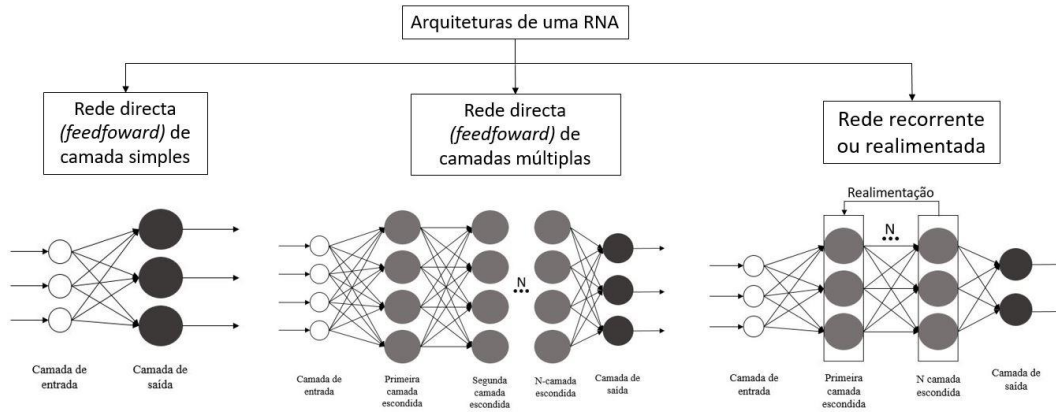


Figura 4-12: Arquiteturas possíveis de uma RNA segundo as ligações entre camadas.

4.3.2.1. Rede direta

Também denominada, em inglês, rede *feedforward*. Neste tipo de arquitetura de rede os dados circulam numa única direção, entre a camada de entrada e a camada de saída. No caso da rede direta de camada simples existe apenas uma única camada onde ocorre o processamento de dados – a camada de saída. Na rede direta de múltiplas camadas existem N camadas escondidas, onde a maior parte do processamento dos dados ocorre nos neurónios escondidos pertencentes a essas camadas. À medida que o número de camadas escondidas aumenta, a RNA é capaz de extrair dos dados relações incrementalmente mais complexas [55]. Esta característica deve-se ao maior número de ligações estabelecidas entres os neurónios. Esta característica tem uma importância acrescida à medida que o tamanho da camada de entrada aumenta [55]. Nos casos presentes na Figura 4-12, referentes a arquiteturas diretas, as redes dizem-se totalmente conectadas, em inglês *fully connected*, pois todos os neurónios de uma camada estão ligados a todos os neurónios da camada seguinte. Caso uma ligação entre algum nódulo da rede falhe, essa rede diz-se parcialmente conectada, em inglês *partially connected*.

4.3.2.2. Rede recorrente ou realimentada

Esta arquitetura de rede difere das redes diretas pela existência de, pelo menos um, ciclo de *feedback* – um ciclo de retroalimentação dos dados de saída de uma camada de neurónios para a(s) camada(s) anterior(es). A presença destes ciclos recorrentes tem um profundo impacto na capacidade de aprendizagem da RNA e, conseqüentemente, na sua performance.

A existência de ciclos de realimentação envolve a utilização de um particular tipo de elementos: elementos de retardamento de unidades, representados por z^{-1} , conforme exemplificado Figura 4-13. Na utilização destes elementos resulta num comportamento dinâmico não linear da RNA – assumindo que existem elementos não lineares nessa mesma rede.

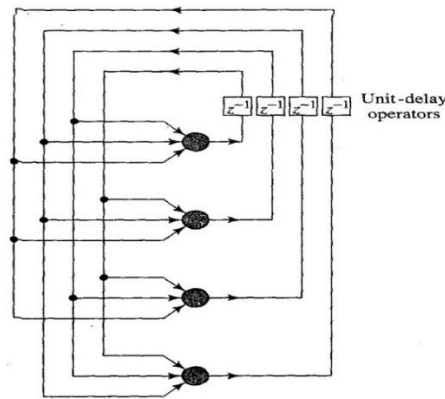


Figura 4-13: Exemplo de rede neuronal artificial recorrente com os elementos de retardamento de unidades presentes (retirado de [55])

As possibilidades da utilização dos ciclos de realimentação são inúmeras. Conforme ilustrado em Figura 4-13, a RNA é constituída apenas por uma única camada onde o resultado (*output*) de cada neurónio alimenta, por sua vez, todos os neurónios dessa camada exceto o neurónio que originou esse *output*. É possível também, utilizar estes ciclos de modo a alimentar cada neurónio com o *output* que esse neurónio originou. A esse tipo de ciclos denomina-se *self-feedback loop*, em português ciclo de realimentação própria.

4.3.3. Aprendizagem

Como referido na secção 4.3.1 o comportamento eficiente das redes neuronais artificiais resulta das interligações entre os elementos de processamento de dados existentes aliada a capacidade da rede em procurar ligações e padrões complexos nos dados para obter resultados precisos. Para tal, a rede necessita de realizar um conjunto de procedimentos de modo a ajustar adequadamente os parâmetros relevantes e poder assim “apreender” dos dados uma determinada função ou uma solução generalizada para uma classe de problemas. A esse conjunto de procedimentos dá-se o nome de algoritmo de aprendizagem. Geralmente esta aprendizagem é realizada recorrendo a um subconjunto dos dados disponíveis – fase de treino permitindo proceder a calibração da rede. Assim, é necessário dividir o conjunto dos dados de entrada em três diferentes subconjuntos, de modo à rede poder realizar o treino necessário:

- Subconjunto de treino: este subconjunto é utilizado na fase de calibração e tem a função de servir como base para o ajustamento dos pesos e do viés de cada neurónio, de cada camada, a partir do algoritmo de treino previamente escolhido.
- Subconjunto de validação: Este subconjunto é utilizado com o propósito de gravar os pesos e erros da rede. Durante a fase de calibração a função de erro, referida em cima, é monitorizada. Durante a fase inicial de calibração/treino da rede essa função de erro para o conjunto de treino e de validação tende a decrescer. No entanto, com o sobre ajuste dos dados, aquando a calibração da rede, tem-se que o erro tende a aumentar. É, portanto, no ponto onde o erro é mínimo que os valores dos pesos e o viés são guardados.
- Subconjunto de teste: Este subconjunto não é utilizado durante a fase de treino ou calibração. No entanto é utilizado para comparar diferentes modelos. É útil também mapear o erro de teste durante a fase de calibração. Se o erro de teste atingir um mínimo numa iteração significativamente diferente do mínimo obtido no subconjunto de validação pode indicar que se a divisão dos dados de entrada, com vista para a calibração da rede, foi mal realizada [55].

Como já foi referido na secção 4.3.1.2 as RNA podem sofrer dois tipos distintos de treino: treino supervisionado ou treino não supervisionado:

4.3.3.1. Treino não supervisionado

Neste tipo de treino apenas são fornecidos à RNA uma série de dados de entrada. Este método não exige treino prévio pelo que cabe à própria rede decidir quais as características e pesos que tem de atribuir para agrupar e processar os dados de entrada. É necessário que haja redundâncias nos dados para que seja possível que a rede identifique padrões. A este processo é também dado o nome de auto-organização ou adaptativo [53].

4.3.3.2. Treino supervisionado

Este é o tipo de treino geralmente utilizado na calibração de uma rede neuronal. Neste caso é apresentado sinais de entrada e saída à rede. A rede, i.e., os pesos entre os neurónios, são ajustados continuamente e iterativamente de modo a minimizar uma função de erro, através de um algoritmo de retropropagação (em inglês *backpropagation*), secção 4.3.4. Uma das funções de erro mais comumente utilizada é a minimização da função do erro médio quadrático, em inglês *root mean squared error (RMSE)*, entre os dados de entrada e os dados na matriz alvo – nome dado ao conjunto de dados de saída [53].

4.3.4. Algoritmo de retropropagação

O algoritmo de retropropagação, ou *backpropagation*, é o algoritmo de aprendizagem mais utilizado no treino de redes neuronais artificiais. O algoritmo foi introduzido na literatura nos anos 70, mas apenas em 1982 foi reconhecida a sua importância e benefícios, através do trabalho de David Rumelhart, Geoffrey Hinton, e Ronald Williams. No seu artigo os autores demonstram experimentalmente a utilidade da utilização deste algoritmo de treino para as redes neuronais artificiais, mais especificamente a utilidade da utilização dos algoritmos de retropropagação para calibrar a rede com foco nas camadas escondidas e a sua importância para obter resultados mais precisos [59].

À medida que o sinal de entrada é propagado através das várias camadas da RNA, durante a fase de calibração da rede, a função de erro, E , é monitorizada. A função de erro está dependente do sinal de saída gerado pela rede e do sinal de saída desejado (subconjunto de treino e validação – ver secção 4.3.3). A Equação 4-17 define matematicamente a função erro.

$$E_k = \frac{1}{2} \sum_k (y_k - t_k)^2 \quad \text{Equação 4-17}$$

onde y_k representa, conforme as secções anteriores, o sinal de saída do neurónio, i.e., representa os dados de saída (*outputs*) – as previsões obtidas; t_k representa o vetor da matriz alvo, ou *target matrix*, onde se encontram os valores desejados para a mesma previsão.

A calibração dos parâmetros da rede dá-se, então, de modo a poder minimizar-se essa função de erro. Durante a fase de treino com recurso ao algoritmo de retropropagação é possível distinguir duas fases:

1. É apresentado à RNA um conjunto de dados, divididos em três subconjuntos, como visto em 4.3.3. Estes dados alimentam a RNA, propagando-se sendo processados em cada camada até dar origem a um sinal de saída, na camada de saída;
2. O sinal de saída produzido pela rede é comparado com o sinal de saída desejado. O algoritmo consiste na retropropagação do erro entre esses dois sinais, modificando os parâmetros da rede (pesos e viés), de modo a minimizar o erro na próxima iteração, i.e., utilizando os novos pesos e viés atribuídos, sendo que o erro se propaga, na rede, em sentido contrário ao do sinal de entrada.

Para esta calibração gerar resultados fiáveis o algoritmo estuda, como ferramenta, a variação do gradiente da função de erro de modo a tomar a decisão de aumentar ou decrescer os valores dos pesos. Este estudo do gradiente tem como base o cálculo diferencial que mede a sensibilidade do resultado final (*output*) em relação à alteração de um argumento na camada de entrada e/ou escondida. Por outras palavras, através do estudo do gradiente da função de erro o algoritmo consegue determinar a relação entre a alteração do valor dos pesos e viés de cada neurónio da RNA e o efeito que essa alteração provoca no *output* da rede e perceber a orientação que necessitam, tanto os pesos como o viés, para minimizar os erros obtidos. Matematicamente, o gradiente não é mais que as derivadas parciais da função de erro em função do peso de cada neurónio e o viés, conforme Equação 4-18 e Equação 4-19, respetivamente:

$$\frac{\partial E}{\partial w} \quad \text{Equação 4-18}$$

$$\frac{\partial E}{\partial b} \quad \text{Equação 4-19}$$

Este processo de ajuste dos pesos através do estudo do seu gradiente, que tem o nome de *gradient descent*, encontra-se ilustrado na Figura 4-14. Este processo é feito iterativamente no sentido da minimização da função de erro até que a solução obtida, usando os pesos obtidos a cada iteração, convirja para uma solução estável, isto é, que o gradiente da função erro seja praticamente nulo.

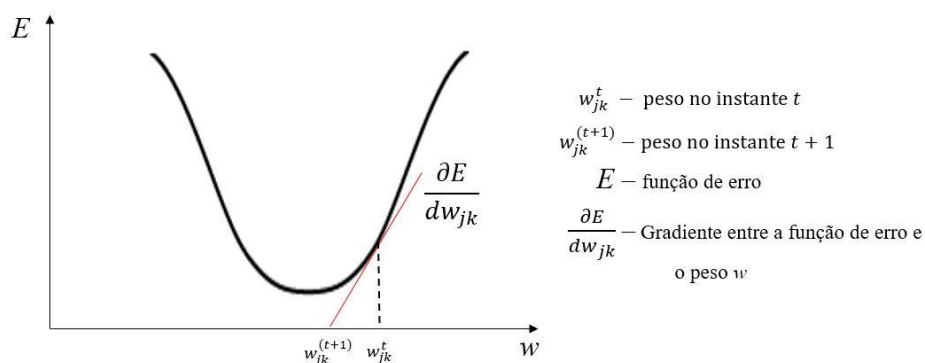


Figura 4-14: Processo de calibração dos pesos pelo estudo do gradiente da função de erro.

Na Figura 4-14 existe apenas um mínimo – o mínimo global da função de erro. No entanto, na generalidade dos casos práticos, é possível distinguir mínimos locais e mínimos absolutos ou globais da função de erro. Idealmente procura-se uma convergência para a minimização absoluta da função de erro. Se o peso, w , sofrer uma atualização por um pequeno passo, ou Δw , a alteração na função de erro pode ser dada pela Equação 4-20:

$$\Delta E = \Delta w_{jk} \nabla E(w_{jk}) = \Delta w_{jk} \frac{\partial E}{\partial w_{jk}} \quad \text{Equação 4-20}$$

onde ΔE representa a variação na função de erro; Δw_{jk} representa a variação dos pesos e $\nabla E(w_{jk})$ representa o vetor do gradiente da função de erro em ordem a w . O gradiente é um vetor que indica qual a direção onde é possível observar uma maior taxa de aumento da função de erro [60]. Sabendo que a solução se encontra na direção da minimização da função de erro, o algoritmo vai iterativamente alterando os pesos na direção de $-\nabla E(w_{jk})$ de modo a poder minimizar a função de erro. Este ciclo termina quando o gradiente for nulo, isto é, quando a função de erro for minimizada. Esse ponto

coincide com o momento em que $\nabla E(w_{jk}) \approx 0$. A calibração do viés é também realizada utilizando este mesmo processo. A Equação 4-20 pode ser reescrita em função do viés:

$$\Delta E = \Delta b_{jk} \nabla E(b_{jk}) = \Delta b_{jk} \frac{\partial E}{\partial b_{jk}} \quad \text{Equação 4-21}$$

De ressaltar que quando se tem uma função de erro que apresenta tanto mínimos locais como mínimos absolutos existe mais que um ponto em que o gradiente se anula. Torna-se, portanto, difícil de perceber se o mínimo atingido, aquando a minimização da função erro, é um mínimo local ou mínimo absoluto [60]. Para solucionar este problema é costume repetir a metodologia de previsão, com os mesmos subconjuntos de dados, e determinar qual a melhor configuração e solução obtida, do conjunto de testes realizado. Deste modo é possível verificar se existe variância significativa, entre os testes, para a solução obtida para os parâmetros da rede, para os erros calculados para essas mesmas soluções e ainda determinar se os mínimos atingidos pela convergência da rede são globais ou locais e se a solução encontrada aquando a minimização da função de erro é robusta.

A Equação 4-22 e Equação 4-23 são necessárias para a descrição matemática do processo de estudo do gradiente. Enquanto a primeira descreve a variação nos pesos de acordo com o gradiente, a segunda descreve o cálculo do peso no instante $t + 1$:

$$\Delta w_{jk} = -\eta \nabla E(w_{jk}) = -\eta \frac{\partial E}{\partial w_{jk}} \quad \text{Equação 4-22}$$

$$w_{jk}^{(t+1)} = w_{jk}^t + \Delta w_{jk} \quad \text{Equação 4-23}$$

onde η representa a taxa de aprendizagem do algoritmo. Também denominada por tamanho do passo (*step size*). Este parâmetro configurável representa a variação possível, a cada iteração, dos pesos ou do viés e toma um valor positivo entre 0 e 1. Este parâmetro limita, portanto, a velocidade a que a rede é calibrada. À medida que a taxa de aprendizagem se aproxima de 1 o modelo aprende com maior velocidade com o custo de atingir uma solução sub-ótima, tanto para os pesos como para o viés. Em casos limite, um valor demasiado alto para a taxa de aprendizagem pode levar ao aumento do erro, à falta de convergência ou até mesmo à divergência da solução. À medida que η se aproxima de 0 o modelo pode convergir com maior precisão. Se essa taxa for demasiado baixa o mínimo para a converge poderá ser um mínimo local e não um mínimo absoluto. Para taxas de aprendizagem cada vez mais baixas, o tempo de computação para a calibração da rede vai aumentar significativamente. A Figura 4-15 exemplifica como a taxa de aprendizagem pode influenciar a minimização e a procura do mínimo absoluto da função erro.

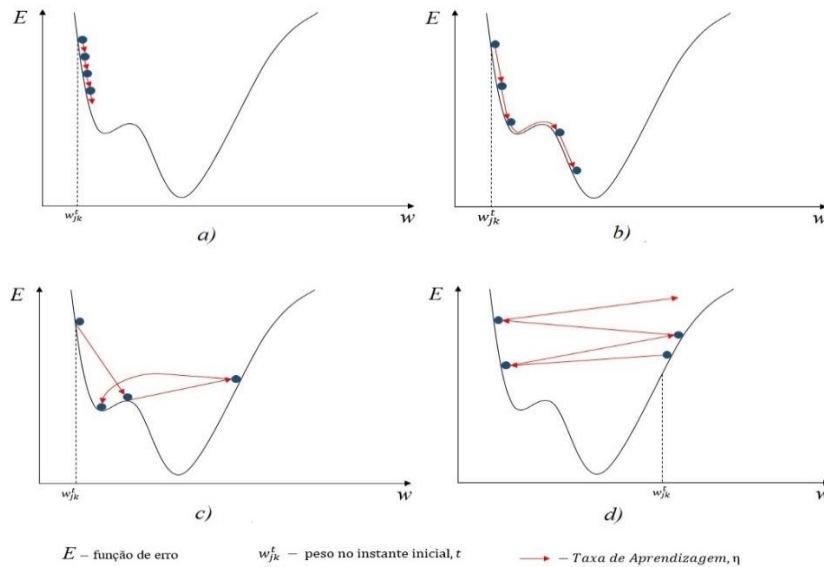


Figura 4-15: Exemplificação da influência da taxa de aprendizagem na busca pelo mínimo da função. Cada ponto, representado, em cada um dos exemplos, representa a solução, ajustada, para os pesos, em cada instante. Em a) é possível observar o efeito de uma taxa de aprendizagem muito baixa; em b) uma taxa de aprendizagem intermédia; em c) uma taxa de aprendizagem alta; em d) uma taxa de aprendizagem muito alta.

Sabendo o modo como o algoritmo minimiza a função de erro é possível reescrever a Equação 4-23 utilizando a Equação 4-22:

$$w_{jk}^{(t+1)} = w_{jk}^t - \eta \frac{\partial E}{\partial w_{jk}} \quad \text{Equação 4-24}$$

Do mesmo modo é possível reescrever a Equação 4-22 e Equação 4-23 em função do viés, b :

$$\Delta b_{jk} = -\eta \nabla E(b_{jk}) = -\eta \frac{\partial E}{\partial b_{jk}} \quad \text{Equação 4-25}$$

$$b_{jk}^{(t+1)} = b_{jk}^t + \Delta b_{jk} = b_{jk}^t - \eta \frac{\partial E}{\partial b_{jk}} \quad \text{Equação 4-26}$$

Introduzida agora a notação adequada tanto a Equação 4-18 como a Equação 4-19 podem ser reescritas:

$$\nabla E(w_{jk}) = \frac{\partial E}{\partial w} = -\frac{w_{jk}^{t+1} - w_{jk}^t}{\eta} \quad \text{Equação 4-27}$$

$$\nabla E(b_{jk}) = \frac{\partial E}{\partial b} = -\frac{b_{jk}^{t+1} - b_{jk}^t}{\eta} \quad \text{Equação 4-28}$$

Na Figura 4-16 é possível observar um fluxograma representativo do processo de calibração da rede utilizando o algoritmo de retro propagação.

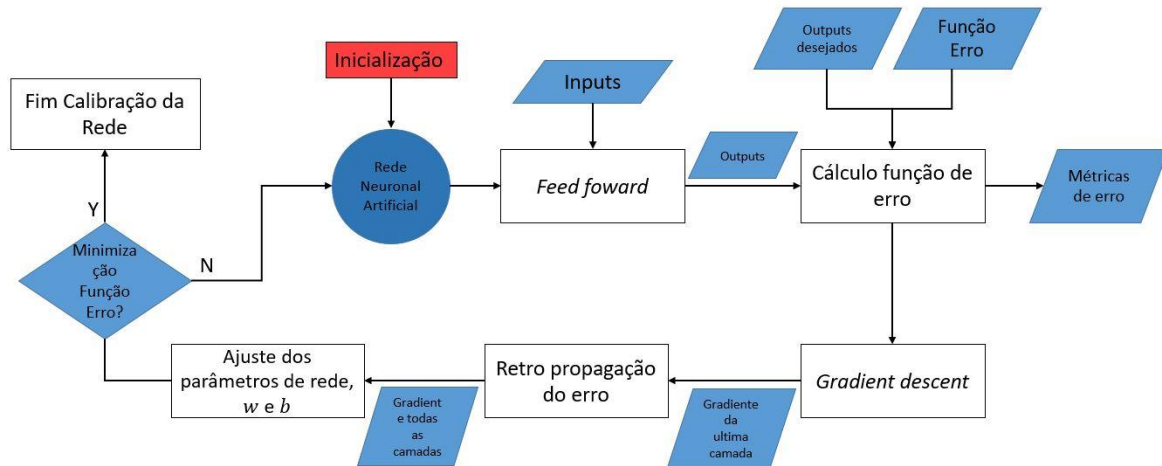


Figura 4-16: Fluxograma referente à calibração de uma rede neuronal artificial utilizando o algoritmo de retro propagação.

4.3.4.1. Algoritmo de retro propagação de Levenberg-Marquardt

Inicialmente publicado em 1944, por Kenneth Levenberg [61], e, posteriormente, redescoberto por Marquardt em 1963 [62], é hoje considerado a técnica *standard* utilizada para solucionar problemas não-lineares dos mínimos dos quadrados (*least-squares*) [63]. O algoritmo de retro propagação de Levenberg-Marquardt (LM) é um algoritmo iterativo que fornece uma solução numérica para o problema da minimização de uma função não linear de valores reais [61] [64].

Amplamente utilizado, transversalmente, num amplo conjunto de disciplinas este algoritmo é um dos algoritmos de treino mais eficientes para as metodologias baseadas em redes neuronais artificiais [65]. Segundo [64], este algoritmo é uma escolha apropriada no que toca à minimização de funções não-lineares onde é requerido uma convergência rápida mas eficaz e estável. Este algoritmo é também apropriado em escalas médias de dados apresentando-se como uma opção mais robusta e eficaz face a dois outros algoritmos utilizados na bibliografia (Gauss-Newton e Levenberg-Marquardt) para calibração de redes neuronais artificiais. Na Figura 4-17 é possível observar os benefícios da utilização deste algoritmo face às opções apresentadas em [64].

Algoritmos	Taxa de Convergência (%)	Iteração Média	Tempo Médio (ms)
Algoritmo EBP ($\alpha=1$)	100	1646.52	320.6
Algoritmo EBP ($\alpha=100$)	79	171.48	36.5
Algoritmo Gauss-Newton	3	4.33	1.2
Algoritmo Levenberg-Marquardt	100	6.18	1.6

Figura 4-17: Comparação entre os diferentes algoritmos estudados para o algoritmo de retropropagação, em [64]. O algoritmo EBP representa o algoritmo *error back propagation*, também chamado *steepest descent* ou em português descida mais íngreme; o parâmetro α é o inverso do parâmetro μ – coeficiente de *damping*; o algoritmo Gauss-Newton é a restante opção para o algoritmo de treino da RNA, presente no trabalho; a taxa de convergência representa a convergência da solução utilizando o respetivo algoritmo; a iteração média representa o número médio de iterações do algoritmo até à convergência; o tempo médio representa o tempo médio do algoritmo, em ms, até à convergência. Adaptado de [64].

O algoritmo de Levenberg-Marquardt (LM) consiste numa aproximação do método de Gauss-Newton com a particularidade de não ser necessário calcular a matriz Hessiana, i.e., sem necessitar de determinar as derivadas de segunda ordem [64], [66].

Na base deste algoritmo faz-se uma aproximação da matriz Hessiana a partir da matriz Jacobiana, derivada do algoritmo de Gauss-Newton, segundo a Equação 4-29. Esta aproximação é válida no caso da calibração de redes neuronais artificiais uma vez que a função de performance do algoritmo apresenta a forma da soma de quadrados [66].

$$H \approx J^T J \quad \text{Equação 4-29}$$

A solução do algoritmo LM consiste na solução da Equação 4-30:

$$J^T e = [\mu I + J^T J] \delta \quad \text{Equação 4-30}$$

onde J representa a matriz Jacobiana que contém as derivadas de primeira ordem. Esta possui um tamanho de N por W , representativo do número de entradas no subconjunto de treino e o número total de parâmetros da RNA (pesos e viés), respetivamente; e representa o vetor que contém o erro de cada *output* associado a cada vetor de *input* utilizado na fase de calibração da RNA. Este assume a forma da soma de quadrados (função performance); I representa a matriz identidade; δ representa o vector constituído pelos pesos ajustados a encontrar e μ representa o coeficiente de *damping* do algoritmo LM – este coeficiente assume um valor positivo e é ajustado a cada iteração, auxiliando o processo de otimização dos parâmetros da RNA a calibrar.

A regra de ajuste dos pesos e viés deste algoritmo de treino é dada pela Equação 4-31 e Equação 4-32, respetivamente:

$$w_{k+1} = w_k - (\mu I + J_k^T J_k)^{-1} J_k e_k \quad \text{Equação 4-31}$$

$$b_{k+1} = b_k - (\mu I + J_k^T J_k)^{-1} J_k e_k \quad \text{Equação 4-32}$$

Este algoritmo de treino é caracterizado pela utilização de duas metodologias diferentes aquando da calibração dos parâmetros, dependendo do valor de μ . No caso de μ assumir um valor pequeno (até próximo de 0), ocorre uma rápida redução da função de erro e o algoritmo LM aproxima-se do processo de calibração verificado no algoritmo Gauss-Newton para a calibração da rede. Por outro lado, se μ assumir um valor maior, o algoritmo aproxima-se do processo de *gradient descent*. É esta característica única que permite ao algoritmo LM conjugar ambos os algoritmos de aprendizagem anteriormente referidos e obter uma mais rápida e estável convergência para a solução de calibração [52] [64].

De modo a encontrar a solução para a Equação 4-30, a cada iteração, os seguintes passos são necessários:

1. Calcular a matriz Jacobiana;
2. Calcular o gradiente, dado por:

$$g = J^T \cdot e \quad \text{Equação 4-33}$$

3. Determinar a aproximação da matriz Hessiana, através da Equação 4-29;

- Calcular a solução de Equação 4-33 em ordem a δ :

$$g = [\mu I + H]\delta$$

Equação
4-34

- Ajustar, de acordo com δ , os parâmetros da rede – tanto os pesos como o viés;
- Calcular o erro (na forma da soma do quadrado dos erros) utilizando os parâmetros ajustados, i.e. para $k+1$;
- Se a soma do quadrado dos erros não decrescer descartar os novos parâmetros da rede, aumentando, ou diminuindo de acordo com o comportamento da rede, o valor do coeficiente μ de modo a refazer o processo a partir do 4º passo.

Na Figura 4-18 é possível observar um fluxograma onde estes passos acima descritos encontram-se esquematizados.

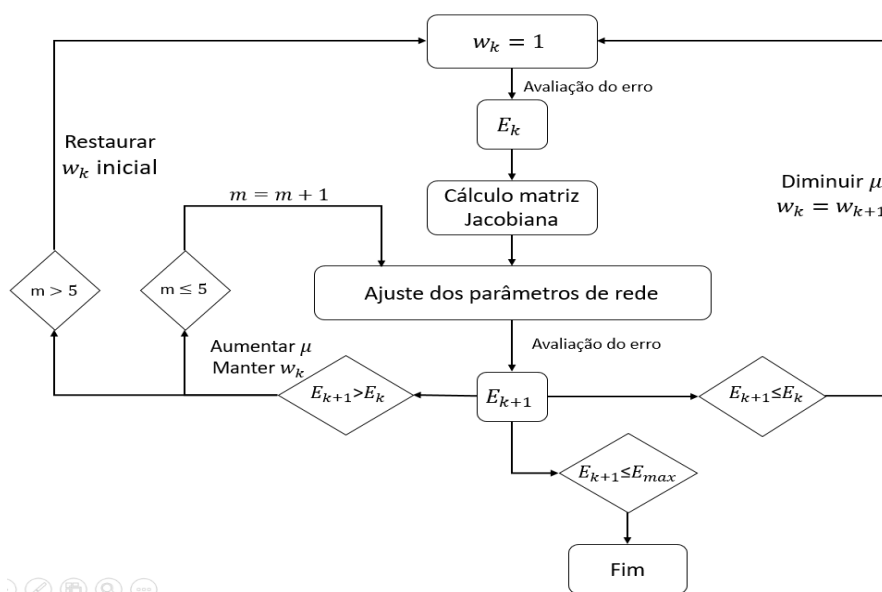


Figura 4-18: Fluxograma referente ao treino de uma rede neuronal artificial utilizando o algoritmo de Levenberg-Marquardt. w_k representa o peso no instante inicial; w_{k+1} representa o peso após ajuste; do mesmo modo E_k representa o erro total com os pesos iniciais enquanto E_{k+1} representa o erro total com os pesos ajustados; m representa um contador de iterações do ciclo e μ representa o coeficiente de *damping*.

4.3.5. Implementação das Redes Neurais Artificiais

No âmbito da dissertação as previsões utilizando esta metodologia foram realizadas com recurso à *Neural Network Toolbox* [67] presente na ferramenta Matlab®. A estrutura geral das RNA com realimentação utilizadas encontram-se esquematizadas na Figura 4-19.

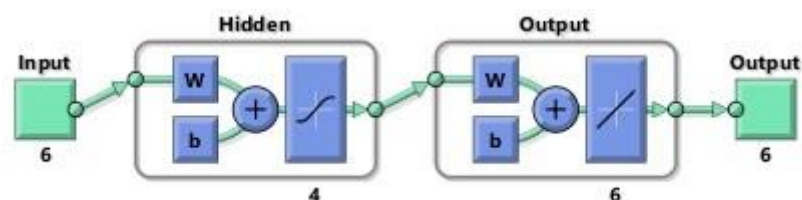


Figura 4-19: Estrutura geral da Rede Neuronal Artificial utilizada. Adaptado de [67].

Na Figura 4-20 encontram-se esquematizados os principais passos na metodologia.

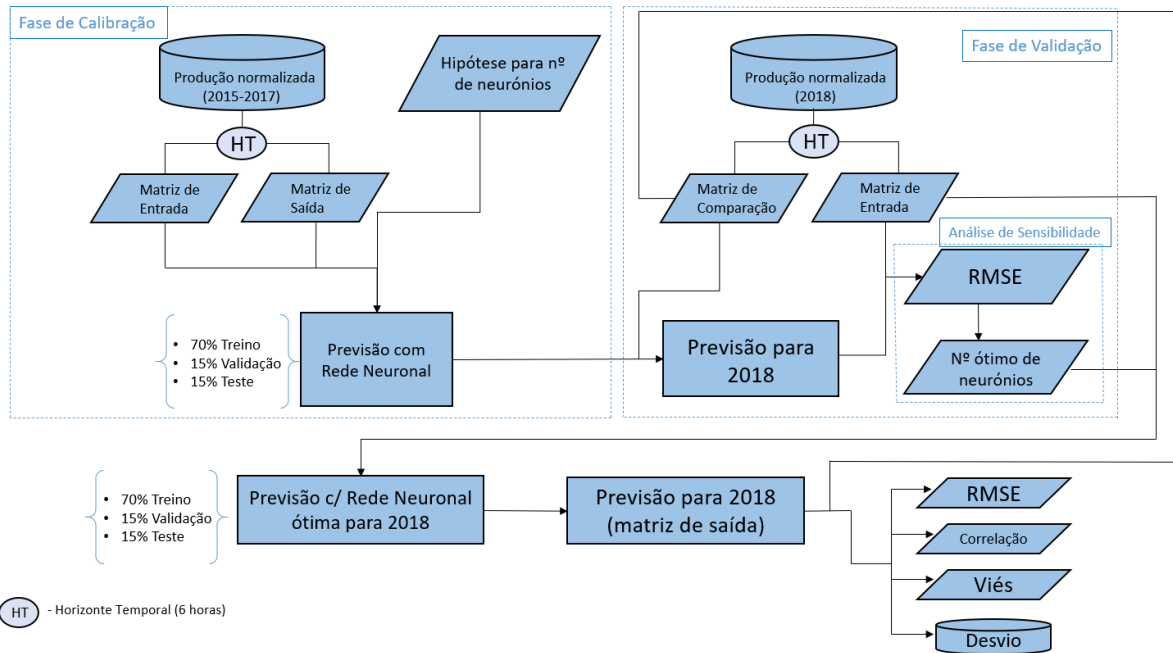


Figura 4-20: Diagrama com a esquematização dos principais passos da metodologia das redes neuronais artificiais.

4.3.5.1. Divisão do conjunto de dados

Na calibração das RNA é necessário dividir o conjunto de dados em três diferentes subconjuntos, conforme mencionado na secção 4.3.3. Assim sendo a divisão dos dados históricos, *i.e.*, a produção eólica normalizada para os anos de 2015 a 2017, foi feita segundo a escolha padrão presente na ferramenta utilizada: 70% dos dados para a fase de treino, 15% para a fase de validação e os restantes 15% para a fase de teste. A divisão dos dados é feita de maneira aleatória pela ferramenta.

4.3.5.2. Número de neurónios na camada escondida

Na atual literatura não existe uma definição conclusiva para qual o número de neurónios presentes nas camadas escondidas ótimo, existindo apenas diferentes recomendações (também chamadas *rule-of-thumb*). Segundo [52] um número reduzido de neurónios pode levar a um subajuste dos dados enquanto um número elevado de neurónios pode levar a sobreajustes. Foi então realizada uma análise de sensibilidade para determinar o número ótimo de neurónios na camada escondida para o problema em estudo. Determinou-se que o número ótimo de neurónios é aquele que apresentar o menor valor de RMSE e o máximo valor de correlação. Optou-se por configurar a rede neuronal com: 2/3, 70% e 90% do tamanho da camada de entrada [68]. Como o número de neurónios tem de ser obrigatoriamente um valor inteiro positivo, sempre que justifique é realizado um arredondamento do valor obtido. Devido à elevada variabilidade da produção eólica foi utilizado também o dobro dos neurónios face ao tamanho da camada de entrada.

A Matriz de Entrada foi definida considerando a trajetória dos valores observados no horizonte temporal *HT* anterior, segundo a Equação 4-35:

$$\text{Matriz de Entrada} = \begin{bmatrix} p_{t-HT} & p_{t-(HT-1)} & \cdots & p_{t-(HT-5)} \\ \vdots & \vdots & \vdots & \vdots \\ p_{t_{max}-HT} & p_{t_{max}-(HT-1)} & \cdots & p_{t_{max}-(HT-5)} \end{bmatrix} \quad \text{Equação 4-35}$$

onde p_t simboliza a produção eólica normalizada para o instante t ; HT como sendo o horizonte temporal de previsão – 6 horas – e NH o número de intervalos horários compreendidos entre o respetivo período de tempo – no caso da fase de calibração a matriz de entrada compreende os dados de produção eólica

normalizada entre os anos de 2015 e 2017. No caso da fase de validação, os dados compreendem o ano de 2018. Esta matriz representa o conjunto histórico de dados com um intervalo de 6 em 6 horas, para do respetivo período de tempo.

A Matriz de Saída tem como principal função replicar os dados de saída correspondente aos dados históricos de entrada, i.e., os resultados obtidos pela previsão através da RNA. É constituída pela matriz complementar da Matriz de Entrada, durante a fase de calibração e pela atual previsão obtida pela RNA para o período de 2018. Utilizando as variáveis apresentas em Equação 4-35, tem-se Equação 4-36:

$$Matriz\ de\ Saída = \begin{bmatrix} p_t & p_{t+(HT-5)} & \cdots & p_{t+(HT-1)} \\ \vdots & \vdots & \vdots & \vdots \\ p_{t_{max}} & p_{t_{max}+(HT-5)} & \cdots & p_{t_{max}+(HT-1)} \end{bmatrix} \quad \begin{array}{l} \text{Equação} \\ 4-36 \end{array}$$

A Matriz Comparação, também denominada, *target matrix*, é o equivalente à matriz complementar mas para o ano de 2018. Possui exatamente a mesma estrutura que a Matriz de Saída com a diferença que esta é composta pelos atuais valores de produção eólica normalizada para o respetivo período de tempo. Através desta matriz é possível avaliar a precisão das previsões obtidas através da metodologia de previsão, determinando as métricas de erro apropriadas.

4.3.5.3.Previsão com RNA

Como visto anteriormente o algoritmo de retro propagação implementado na metodologia para a presente dissertação pode encontrar dificuldades em localizar o mínimo global da função e encontrar, em vez, um mínimo local da mesma. Como já foi visto a aproximação pelos mínimos locais é suficiente, desde que a solução encontrada seja estável [60]. Para verificar a estabilidade e robustez da convergência da rede foram realizados, então, cinco testes diferentes para cada configuração da RNA a estudar, i.e. para cada configuração da RNA – com os diferentes números de neurónios na camada escondida – foram realizados cinco testes diferentes para verificar se havia variação significativa na convergência das mesmas.

Realizados então os cinco testes a previsão para o ano de 2018 foi efetuada de acordo com a configuração, e teste, da RNA que apresentasse melhor performance. A performance foi avaliada segundo duas métricas de erro utilizadas na presente dissertação, e discutidas posteriormente: o erro quadrático médio e a Correlação.

Após a escolha da melhor configuração, estudada na presente dissertação, para a metodologia RNA procedeu-se, antes das previsões, ao treino da mesma. Para tal realizou-se a mesma divisão dos dados para calibração que a mencionada em 4.3.5.1. No que toca ao algoritmo de aprendizagem utilizado recorreu-se ao algoritmo de treino de Levenberg-Marquardt, descrito em 4.3.4.1. A rede é, então, treinada, seguindo o fluxograma presente na Figura 4-18, até um dos três critérios impostos, segundo o padrão da ferramenta, serem cumpridos: até o treino da rede atingir no máximo mil iterações; quer seja o mínimo do gradiente ser encontrado – 1.00×10^{-7} –; ou até que o índice de performance da rede, referente à métrica RMSE, atinja o mínimo – 0.00. Quando qualquer um destes três critérios é cumprido o treino da rede chega ao fim.

É, no fim, esta a configuração da RNA que é utilizada para realizar as previsões de produção eólica.

4.4. Métricas de Avaliação

Realizadas as previsões de produção eólica normalizada para o ano de 2018, cada uma das metodologias de previsão implementadas durante a dissertação têm de ser avaliadas segundo parâmetros que permitam avaliar a precisão de cada uma. Não existe nenhuma métrica singular que consiga descrever ou avaliar a performance de uma metodologia de previsão. Nesse sentido, optou-se por utilizar uma

combinação de diferentes métricas de avaliação para poder avaliar a(s) metodologia(s) implementada(s) [69] [19]. As métricas utilizadas na presente dissertação para a avaliação das metodologias de previsão foram:

- Desvio

O desvio, ou erro, e , é uma métrica que avalia a diferença entre o valor previsto e o valor observado, para cada horizonte temporal. É dado pela Equação 4-37:

$$e(t) = \text{Potência Eólica normalizada}_{\text{prevista}}(t) - \text{Potência Eólica normalizada}_{\text{observada}}(t) \quad \text{Equação 4-37}$$

- Viés

O viés, em inglês *bias*, representado por $\overline{e(t)}$, corresponde ao erro sistemático presente em cada previsão. Este representa o valor médio do erro, para o horizonte temporal de previsão e é dado pela Equação 4-38.

$$\overline{e(t)} = \frac{1}{NH} \sum_{t=1}^N e(t) \quad \text{Equação 4-38}$$

onde NH corresponde ao número de intervalos horários para o respetivo período de tempo; $e(t)$ corresponde ao erro, ou desvio, para o mesmo período de tempo.

O viés é a métrica de avaliação mais simples de implementar, no entanto é uma das métricas que permitem avaliar se a metodologia de previsão apresenta tendência para subestimar ou sobrestimar produção os valores observados. Idealmente procura-se um viés, para o horizonte temporal, o mais próximo possível de zero: significa isso que para o horizonte temporal, a média das flutuações da previsão encontram-se todas próximas dos valores reais observados [70].

- Erro Quadrático Médio

Do inglês *Root Mean Squared Error* (RMSE), esta é uma métrica de avaliação comumente utilizada na avaliação de metodologias de previsão permitindo identificar a variação dos erros de amplitude, devido à natureza do quadrado das diferenças [69]. Este tipo de erros é de maior interesse e valor para os gestores da rede elétrica [70]. No entanto esta vantagem acarreta a perda de direccionalidade do erro, i.e., a informação relacionada com a sub/sobrestimação da previsão obtida pela metodologia, informação essa que pode ser essencial quando a penetração eólica no sistema elétrico é elevada e, particularmente, em eventos de rampa onde os erros de previsão podem fazer a diferença para a gestão do sistema elétrico [69]. Esta métrica é calculada através da Equação 4-39, e o seu valor ideal é zero.

$$RMSE = \sqrt{\frac{1}{NH} \sum_{t=1}^N e(t)^2} \quad \text{Equação 4-39}$$

- Correlação

O coeficiente de correlação, representado por R^2 , é uma medida estatística que se baseia nas similaridades presentes entre as previsões obtidas, para um horizonte temporal de previsão, e o valor observado para o mesmo horizonte temporal [25]. É dado pela Equação 4-40.

$$R^2_{x,y} = \frac{\sum((x_i - \bar{x})(y_i - \bar{y}))}{\sqrt{S_{xx}S_{yy}}} \quad \text{Equação 4-40}$$

onde a variável x representa as previsões feitas; a variável y representa os valores observados; x_i e y_i representam os valores, para o instante i , previstos e observados; \bar{x} e \bar{y} representam as médias das respectivas variáveis; $S_{xx} = \sum_i(x_i - \bar{x})^2$ e $S_{yy} = \sum_i(y_i - \bar{y})^2$. Este coeficiente varia entre [-1 1]. O valor próximo de zero significa previsões sem qualidade, e o valor unitário representa previsões perfeitas. Um valor negativo elevado significa que a previsão se encontra em oposição de fase.

- Taxa de Variação da produção

De modo a compreender a capacidade das metodologias de previsão para lidar com períodos extremos de variação de produção, usualmente designados por períodos de rampa [71], determinou-se a taxa de variação entre o primeiro e último instante de previsão Equação 4-41. Estes valores foram comparados com a taxa de variação obtida para a produção real observada Equação 4-42.

$$\text{Taxa de Variação}_{Previsões,t} = Previsão_t(HT) - Previsão_t(HT - 5) \quad \text{Equação 4-41}$$

$$\text{Taxa de Variação}_{Observado,t} = Observado_t(HT) - Observado_t(HT - 5) \quad \text{Equação 4-42}$$

Após determinar as taxas de variação, estas são agrupadas, de forma individual, por classe de variação considerando intervalos de 10%, que variam entre [-100 100]. Este procedimento permite identificar a frequência de ocorrência de cada intervalo permitindo identificar a capacidade de previsão das metodologias implementadas, nomeadamente, se a metodologia sub/sobrestima as variações do intervalo de previsão e em que condições o faz.

Na Equação 4-43 encontra-se a matriz de classes, para as taxas de variação da previsão eólica normalizada e da produção eólica normalizada observada. Nesta matriz é possível observar o número de ocorrências para cada classe das taxas de variação. Quando a taxa de variação, para t , da produção eólica normalizada observada e a taxa de variação da previsão obtida pertencem à mesma classe de variação é, então, contabilizada uma ocorrência, para essa mesma classe. Foi contabilizado o número total de ocorrências, para cada classe, de modo a obter um mapa que nos permite avaliar a capacidade da previsão de cada metodologia de previsão implementada face ao observado.

$$\left[\begin{array}{cccc} TVP_{0,9}^1 & \& TVO_{-0,9}^{-1} & TVP_{0,9}^1 & \& TVO_{-0,8}^{-0,9} & \dots & TVP_{0,9}^1 & \& TVO_{0,8}^{0,9} & TVP_{0,9}^1 & \& TVO_{0,9}^1 \\ TVP_{0,8}^{0,9} & \& TVO_{-0,9}^{-1} & TVP_{0,8}^{0,9} & \& TVO_{-0,8}^{-0,9} & \ddots & TVP_{0,8}^{0,9} & \& TVO_{0,8}^{0,9} & TVP_{0,8}^{0,9} & \& TVO_{0,9}^1 \\ \vdots & & \vdots & & \vdots & & & \vdots & & \vdots & & & \\ TVP_{-0,9}^{-0,8} & \& TVO_{-0,9}^{-1} & TVP_{-0,9}^{-0,8} & \& TVO_{-0,8}^{-0,9} & \dots & TVP_{-0,9}^{-0,8} & \& TVO_{0,8}^{0,9} & TVP_{-0,9}^{-0,8} & \& TVO_{0,9}^1 \\ TVP_{-1}^{-0,9} & \& TVO_{-0,9}^{-1} & TVP_{-1}^{-0,9} & \& TVO_{-0,8}^{-0,9} & \dots & TVP_{-1}^{-0,9} & \& TVO_{0,8}^{0,9} & TVP_{-1}^{-0,9} & \& TVO_{0,9}^1 \end{array} \right] \quad \text{Equação 4-43}$$

onde TVO representa a taxa de variação referente à produção eólica normalizada observada; TVP representa a taxa de variação referente à previsão de produção eólica normalizada; $TVP_{x_2}^{x_1}$ e $TVO_{x_2}^{x_1}$ representam as classes de taxa de variação definidas pelo intervalo de classe x_1 e x_2 , sendo x_1 o maior limite da classe e x_2 o menor limite da mesma classe. Para efeitos de contabilização de ocorrências impôs-se a seguinte regra, representada na Equação 4-44, para a delimitação dos limites das classes:

$$x_1 \leq \text{taxa de variação} < x_2$$

Equação
4-44

- **Varição Percentual**

De modo a poder comparar dois valores entre si e determinar qual a variação, positiva ou negativa, entre eles determina-se a variação percentual. Esta métrica é útil para ilustrar a variação de uma métrica de avaliação de uma metodologia de previsão. Para tal, entre um valor novo, a , e um valor antigo, b , a variação percentual da métrica é dada pela Equação 4-45:

$$\text{Variação percentual} = \frac{(a - b)}{b} \times 100$$

Equação
4-45

Nesta métrica um valor positivo representa uma pioria de a face a b , enquanto um valor negativo representa uma melhoria.

5. Apresentação, análise e discussão de resultados

Nesta secção serão apresentados, discutidos e analisados os resultados obtidos pela implementação das metodologias de previsão adotadas na presente dissertação.

5.1. Análise de sensibilidade das metodologias

Como mencionado nos capítulos anteriores alguns parâmetros das metodologias de previsão implementadas na presente dissertação necessitaram de uma análise de sensibilidade para se obter a configuração mais adequada para o problema em análise

- **Metodologia kNN**

A análise de sensibilidade necessária para esta metodologia refere-se, como explicado na secção 4.2.6.3.1: a otimização do parâmetro p , referente à ordem da distância de Minkowski. Este parâmetro é dependente da aplicação em análise e, portanto, para um melhor ajuste à série de dados utilizada nesta dissertação optou-se por fazer uma análise de sensibilidade e otimização deste parâmetro. O procedimento de cálculo, para a referida análise e otimização, encontra-se descrito na mesma secção 4.2.6.3.1.

O menor valor encontrado, para a ordem p , que satisfaz a condição limite estipulada foi $p = 14$. Uma representação gráfica dos resultados obtidos para esta análise de sensibilidade consta na Figura I-1, presente nos anexos – Anexo I. Também presente no Anexo I encontra-se a Tabela I-1, onde se encontra discriminado o valor do erro, segundo a regra de erro imposta e descrita em 4.2.6.3.1, para o valor de p encontrado. É, portanto, esse o valor adotado para a ordem p da distância de Minkowski na metodologia de previsão com o algoritmo kNN .

- **Metodologia RNA**

A análise de sensibilidade necessária para a metodologia das RNA, ao contrário do que se verificou para o kNN , foi feita durante a fase de calibração da metodologia – ver Figura 4-20.

Na Tabela I-2 e Tabela I-3, presentes nos anexos – Anexo I – encontram-se resumidos os resultados obtidos para a análise de sensibilidade referente ao número de neurónios presentes na camada escondida. Determinou-se, então, que o número ótimo de neurónios presentes na camada escondida, no âmbito da presente dissertação, que minimiza o RMSE e maximiza a correlação foi alcançado no teste

realizado com 5 neurónios (correspondente a 90% do tamanho da camada de entrada) presentes na camada escondida. É, portanto, esta a configuração adotada para a metodologia de previsão com RNA.

5.2. Métricas de erro para as metodologias implementadas

Realizadas as análises de sensibilidade e incorporando os resultados obtidos nas respetivas metodologias foi possível implementar as metodologias e gerar previsões. Após as previsões geradas, já na fase de validação, foi possível avaliar os resultados obtidos, através das métricas de avaliação descritas na secção 4.4.

5.2.1. Persistência

A primeira metodologia de previsão que temos de avaliar, é a metodologia que serve como referência ou *benchmark* – a persistência. Na Tabela 5-1 encontram-se discriminados os valores médios das métricas de erro, descritas em 4.4, das previsões obtidas pela persistência face ao observado, para o período de validação.

Tabela 5-1: Valores das métricas de erro, calculadas para as previsões obtidas com a metodologia da persistência.

Viés [p.u.]	RMSE [p.u.]	Correlação
4.66E-04	0.0714	0.8932

É de salientar que o horizonte temporal em utilização na presente dissertação é curto, como já visto na secção 3.1. Pela natureza da metodologia de persistência, como discutido na secção 3.2.2.1, os horizontes temporais em estudo são os prediletos para este tipo de metodologia. Adicionalmente, devido à série em análise representar a produção nacional agregada, de acordo com vários autores, potencia o que na literatura é designado por efeito estatístico de alisamento (*power smothing effect*) onde as variações de produção são naturalmente atenuadas entre si quando se considera um conjunto alargado de parques eólicos geograficamente dispersos [6]. Nesse sentido os valores médios encontrados para as métricas encontram-se de acordo com o esperado: os valores do erro e viés, apesar de positivos, são muito baixos. A natureza positiva de ambos pode indicar uma tendência para sobrestimação na previsão, utilizando esta metodologia, face ao observado.

- **Perfil diário dos erros da metodologia de persistência**

Na Figura 5-1 encontra-se representado o erro médio por hora, relativamente à metodologia da persistência. Esta foi construída fazendo a média do erro médio para todas as horas do dia, relativamente à totalidade do período de validação – o ano de 2018. Deste modo obtemos a variação média, para um dia – ou 24 horas – do erro, ou desvio, para a metodologia.

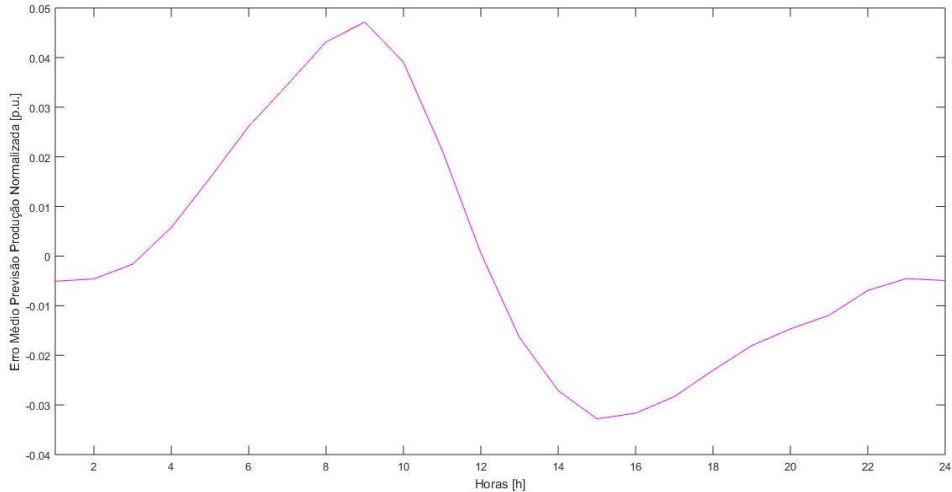


Figura 5-1: Erro médio horário da previsão de produção eólica normalizada, para a metodologia de previsão de persistência.

Usando a mesma abordagem, na Figura 5-2 encontra-se representado o RMSE médio horário para as previsões obtidas pela metodologia de previsão de referência, a persistência.

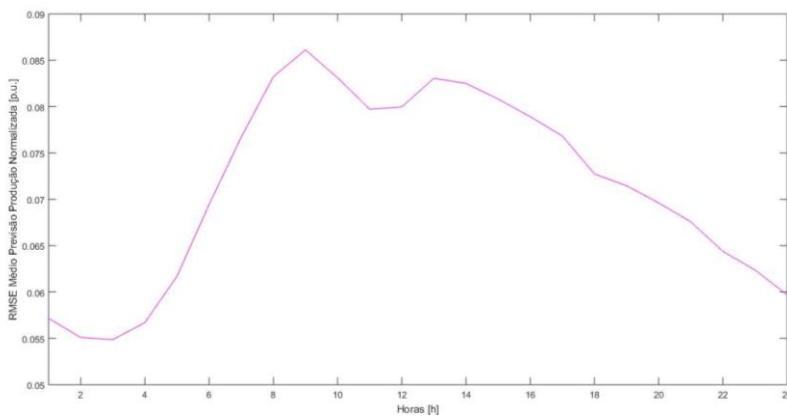


Figura 5-2: RMSE médio horário da previsão de produção eólica normalizada, para a metodologia da persistência.

Os resultados observados na Figura 5-1 e Figura 5-2 encontram-se de acordo com o esperado e em concordância com o comportamento demonstrado no perfil diário (Figura 4-1). Fisicamente este perfil encontra-se associado aos efeitos térmicos de estratificação da atmosfera, resultado dos efeitos de estratificação térmica da atmosfera devido ao ciclo do sol.

5.2.2. *k* Vizinhos Mais Próximos

A segunda metodologia de previsão de produção eólica implementada no âmbito da presente dissertação foi a metodologia de previsão *k*NN.

Na Tabela 5-2 encontram-se discriminados os valores médios do viés, respetivamente, relativamente às previsões obtidas por esta metodologia, face ao observado, para o período de validação.

Tabela 5-2: Valores do viés, p.u., calculados para as previsões obtidas com a metodologia *k*NN. A negrito é indicada a configuração com o melhor desempenho.

Distância	Euclidiana	Minkowski	Correlação

k=1	Não-Ponderada	-6.94E-03	-3.86E-03	3.99E-04
	Ponderada	-6.94E-03	-3.86E-03	3.99E-04
k=5	Não-Ponderada	-6.60E-03	-8.37E-03	2.10E-03
	Ponderada	-6.08E-03	-7.77E-03	1.96E-03
k=10	Não-Ponderada	-7.93E-03	-8.50E-03	6.90E-04
	Ponderada	-7.93E-03	-8.50E-03	6.90E-04
k=20	Não-Ponderada	-1.09E-02	-1.31E-02	-6.74E-03
	Ponderada	-1.01E-02	-1.21E-02	-4.43E-03
k=162	Não-Ponderada	-1.91E-02	-1.90E-02	-2.33E-02
	Ponderada	-1.80E-02	-1.81E-02	-1.94E-02

Uma comparação entre as tabelas acima mencionadas com a Tabela 5-1 confirma as expectativas iniciais: sendo a persistência uma metodologia demasiado simplista é normal que o desempenho obtido por esta seja inferior, face ao obtido pela metodologia dos k vizinhos mais próximos nas suas diferentes configurações.

No caso da métrica de erro do viés, comparando os resultados médios obtidos para a persistência e para a metodologia k NN, verifica-se que: a persistência apresenta um desempenho inferior ao da metodologia k NN, para qualquer uma das variações do algoritmo, i.e., para qualquer um dos valores que k assume ou distância analisada. Do mesmo modo, para o viés: no caso da persistência o resultado obtido para o viés é superior a zero enquanto no caso do k NN o viés é negativo. Suportando a possibilidade da persistência tender a sobrestimar as previsões enquanto o algoritmo k NN tende a subestimar as previsões. Verifica-se, no entanto, que para esta metodologia o viés tende a aumentar à medida que o valor de k também aumenta. Isto poderá significar que a subestimação da metodologia de previsão perde expressão à medida que k aumenta.

Na Tabela 5-3 encontram-se discriminados os valores médios da métrica de erro RMSE, relativamente às previsões obtidas pela metodologia de previsão k NN, face ao observado, para o período de validação. Do mesmo modo, na

Tabela 5-4, encontram-se discriminados os valores médios para a correlação, relativamente às previsões obtidas por esta metodologia, face ao observado, para o mesmo período de validação.

Tabela 5-3: Valores do RMSE, p.u., calculado para as previsões obtidas com a metodologia k NN. A negrito é indicada a configuração com o melhor desempenho.

Distância		Euclidiana	Minkowski	Correlação
k=1	Não-Ponderada	0.0858	0.0855	0.2047
	Ponderada	0.0858	0.0855	0.2047
k=5	Não-Ponderada	0.0669	0.0671	0.1634
	Ponderada	0.0671	0.0672	0.1659
k=10	Não-Ponderada	0.0648	0.0646	0.1579
	Ponderada	0.0648	0.0646	0.1599
k=20	Não-Ponderada	0.0639	0.0638	0.1546
	Ponderada	0.0638	0.0637	0.1561
k=162	Não-Ponderada	0.0658	0.0652	0.1524
	Ponderada	0.0653	0.0648	0.1526

Tabela 5-4: Valores da correlação, calculado para as previsões obtidas com a metodologia *k*NN.

Distância		Euclidiana	Minkowski	Correlação
k=1	Não-Ponderada	0.1528	0.1564	0.0946
	Ponderada	0.1528	0.1564	0.0946
k=5	Não-Ponderada	0.2376	0.2374	0.1563
	Ponderada	0.2345	0.2356	0.1441
k=10	Não-Ponderada	0.2588	0.2647	0.1934
	Ponderada	0.2576	0.2627	0.1763
k=20	Não-Ponderada	0.2761	0.2758	0.2149
	Ponderada	0.2745	0.2769	0.2014
k=162	Não-Ponderada	0.2912	0.2905	0.2484
	Ponderada	0.2915	0.2912	0.2463

Na Figura 5-3 encontra-se uma representação gráfica dos resultados presentes nas Tabela 5-3 e Tabela 5-4.

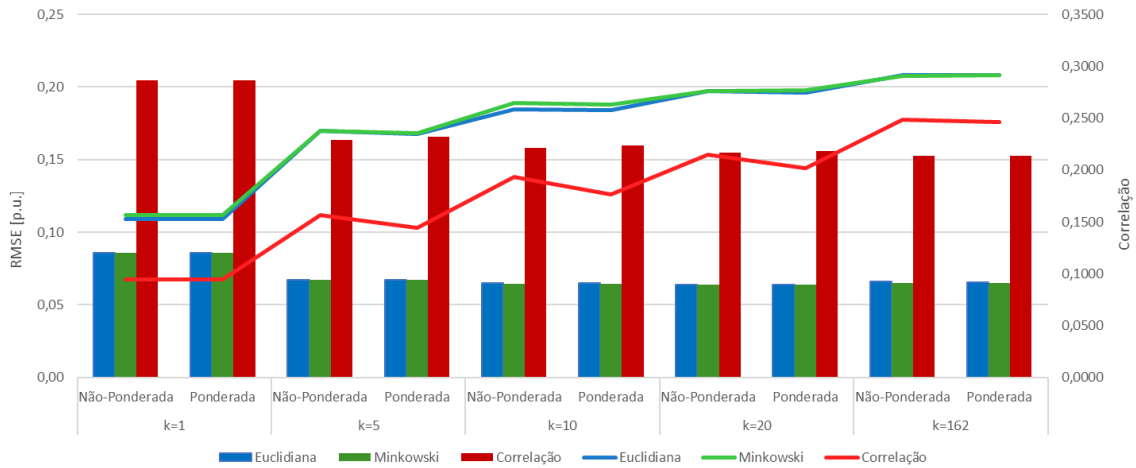


Figura 5-3: Representação gráfica dos resultados obtidos para as métricas do RMSE (barras) e correlação (linhas), para as previsões obtidas através da metodologia kNN.

No RMSE é possível observar que à medida que o valor de vizinhos (k) aumenta tem-se que o RMSE tende a decrescer, para qualquer das configurações da metodologia kNN estudadas. É de ressaltar, no entanto, na diferença, grande, que da configuração da metodologia kNN com a distância por correlação (a barra vermelha, na Figura 5-3) resultam valores bastante díspares face aos observados nas duas outras configurações. A eficácia de previsão é, portanto, melhor quando se utilizam configurações que têm por base o cálculo da distância euclidiana e de Minkowski que apresentam comportamentos e tendências bastante similares entre si. Outro aspeto importante de ressaltar prende-se com o impacto que a ponderação tem na eficácia das previsões obtidas: para as três configurações da metodologia kNN estudadas verifica-se que a ponderação não provoca variação significativa desta métrica de erro.

Nos valores do coeficiente de correlação é possível observar que os resultados vão, no geral, ao encontro com o que foi observado para a métrica anteriormente discutida. No que toca ao impacto da ponderação é possível observar que a tendência geral é, novamente, de não haver variação significativa na eficácia das previsões com a implementação de um fator de ponderação.

- **Perfil diário dos erros da metodologia kNN**

Do mesmo modo que descrito para a metodologia da persistência, na Figura 5-4 encontra-se a representação do erro médio por hora, para o mesmo período de validação, em função do número de vizinhos (k): 1, 5, 10 ou 162, respetivamente. De notar que os resultados observados para quando k assume o valor de 20 são similares aos obtidos para quando k assume o valor de 10. Nesse sentido os resultados para 20 vizinhos não se encontram representados na figura abaixo. Deste modo é possível avaliar e analisar o impacto que o estudo de k ótimo tem, na métrica do erro.

Comparando o perfil diário do erro para a metodologia kNN com o perfil da técnica de persistência, verifica-se, pela interpretação das mesmas, que a variação do erro médio para cada hora segue um comportamento típico para estas metodologias: tanto para a persistência como para o algoritmo kNN os resultados indicam que o erro médio horário sofre flutuações, aparentemente, típicas ao longo das 24 horas. No caso da persistência verifica-se que, apesar de próximo de zero, o erro é tendencialmente negativo durante a maior parte do dia, exceto entre as 4 e as 12 horas aproximadamente, quando se verifica um pico acentuado no erro médio. Isto significa que a metodologia tende a, durante o referido pico, a sobrestimar a previsão de produção eólica. No que diz respeito à metodologia kNN e à respetiva variação de erro médio horário verifica-se o mesmo comportamento: valores médios de erro maioritariamente negativos para a maior parte do dia, ocorrendo, no mesmo período horário que a

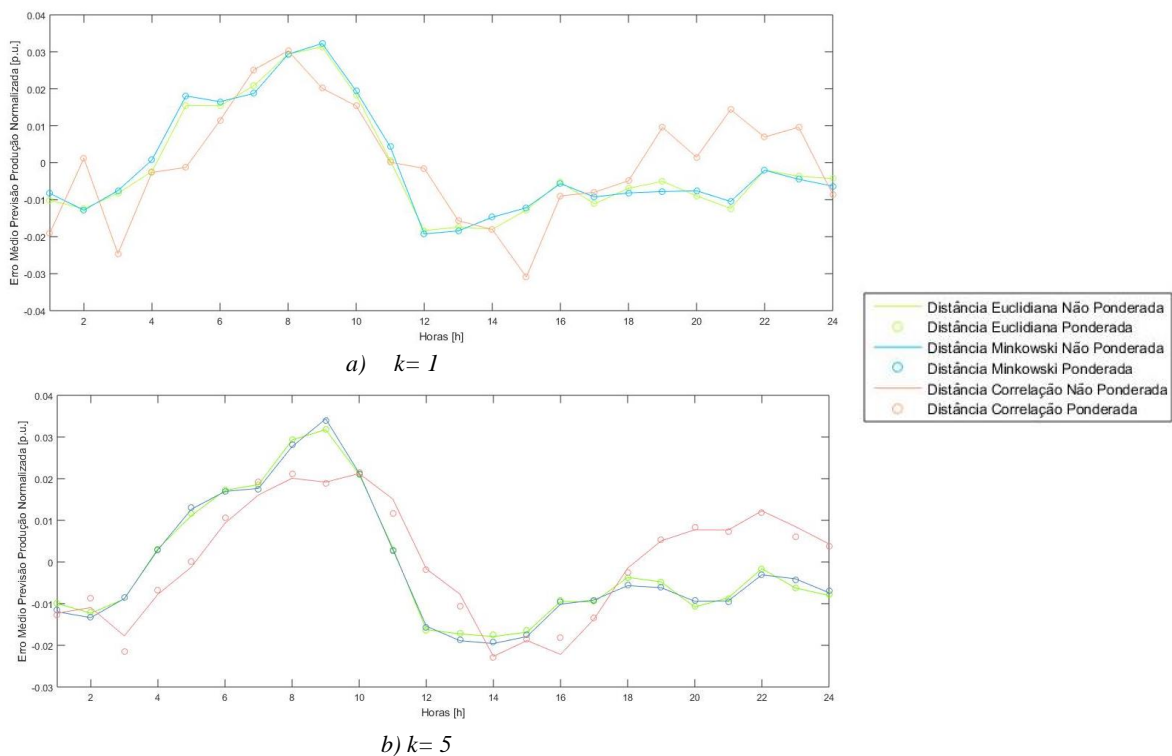
persistência, um pico acentuado. Este, no entanto, não tão significativo e acentuado. Justificando assim a diferença entre os valores negativos observados para o erro médio no caso da metodologia kNN face aos valores positivos observados para o erro médio no caso da metodologia da persistência.

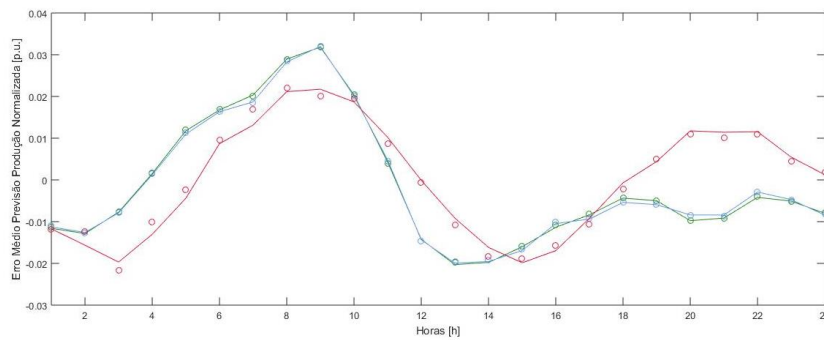
No que diz respeito à variação do erro médio horário relativamente ao estudo da metodologia kNN , i.e., ao estudo da variação da eficácia da metodologia em função do valor de k e do método de cálculo de distâncias, é possível verificar, pela análise da Tabela 5-3 e Tabela 5-4, e da Figura 5-3, que os erros médios horários das previsões realizadas com o cálculo de distâncias Euclidiana e de Minkowski apresentam variações bastante semelhantes.

À medida que o valor de k aumenta verifica-se que os erros médios horários, para a metodologia de cálculo de distâncias Euclidiana e de Minkowski, tornam-se cada vez mais semelhantes.

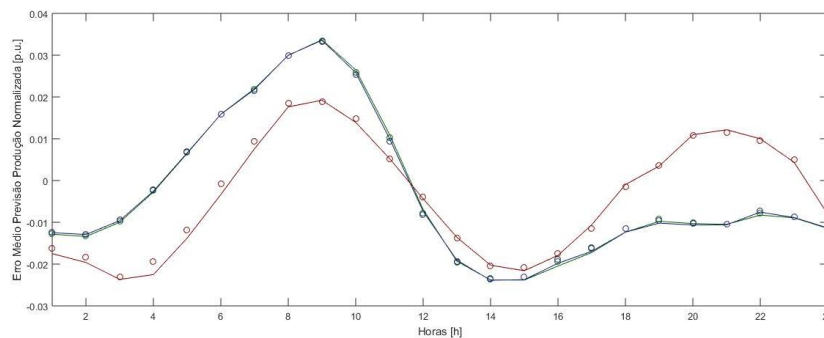
Pelo contrário, o cálculo da distância entre duas instâncias baseada na distância por correlação apresenta flutuações horárias e comportamento distintos do verificado nos três casos anteriores – persistência, metodologia kNN com recurso à distância euclidiana e de Minkowski. Verifica-se que quanto maior for a variação do valor de k , verifica-se que o erro médio horário também varia significativamente, principalmente quando comparando essa variação com a homóloga para as outras métricas de distância.

À medida que o número de vizinhos (k) aumenta verifica-se uma suavização na flutuação no erro médio horário: quando k assume os valores mais baixos os resultados para o erro médio horário seguem uma tendência, mas encontram-se mais dispersos; quando k assume valores mais altos (10 e 162) verifica-se nas figuras que a tendência geral de flutuação ao longo do dia encontra-se presente e os resultados são menos dispersos. Isto pode dever-se ao facto de haver um maior número de instâncias que contribuam para a previsão – valores muito díspares tendem a não ter tanto impacto quanto teriam em valores de k menores.





c) $k=10$



d) $k=162$

Figura 5-4: Erro médio da previsão de produção eólica normalizada, para a metodologia k NN para diferentes valores de k .

O impacto da ponderação para as previsões obtidas com a metodologia k NN não é expressivo: não se verifica nenhuma tendência geral de diminuição ou aumento do erro médio horário da ponderação face à não ponderação.

Pela análise da Figura 5-4, referentes ao erro médio horário, é possível observar que este não sofre qualquer impacto quando existe ponderação baseada no cálculo da distância no caso da distância Euclidiana e de Minkowski, independentemente do valor de k . Pelo contrário é possível observar que a ponderação nas previsões tem impacto no erro médio horário quando é aplicada a distância de correlação entre as instâncias.

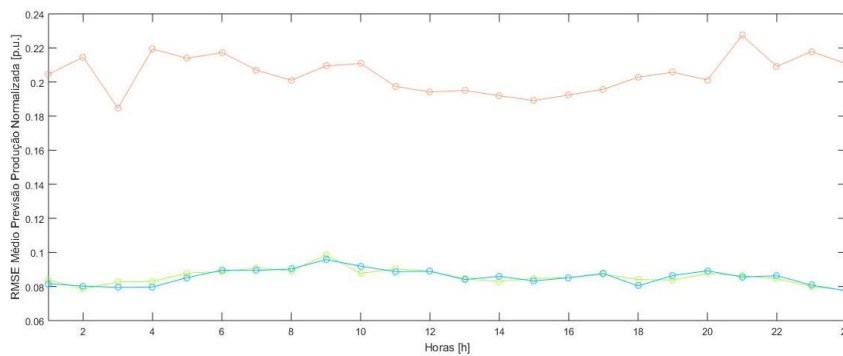
Para os resultados obtidos no que diz respeito ao viés tem-se que, ao contrário do que se verifica com o erro médio horário, para a distância Euclidiana e de Minkowski, quanto menor o valor de k menor é o viés obtido. À medida que k aumenta verifica-se um aumento do valor do viés. Contudo os resultados obtidos indiquem um elevado desempenho desta metodologia para esta métrica de avaliação. Para o caso do cálculo da distância por correlação verifica-se o contrário: este mostra uma tendência decrescente com o aumento do valor de k . Os valores do viés e do erro médio horário são, no entanto, pequenos (ou próximos de zero) para dizer com fiabilidade se existe sub ou sobrestimação das metodologias.

Na Figura 5-5 é possível observar uma representação do RMSE médio horário calculado para a metodologia de previsão k NN implementada. Com vista a estudar a variação do número de vizinhos a considerar, cada figura corresponde a um diferente valor de k e estão representadas as três variantes do cálculo da distância entre instâncias.

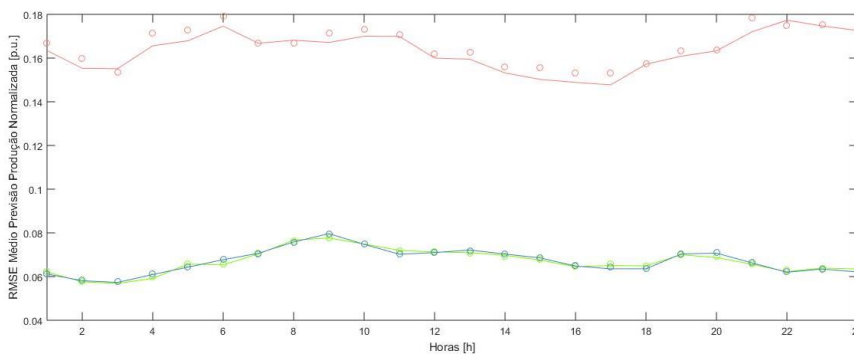
Analisando as figuras relativas ao RMSE médio horário acima representadas, é possível confirmar algumas tendências previamente descritas: é possível confirmar que a natureza da métrica de cálculo da distância entre semelhanças tem, de facto, expressão significativa no desempenho das previsões obtidas na metodologia em análise. A distância Euclidiana e de Minkowski, que são da mesma família de métricas de distância, dão origem a resultados bastante diferentes – e melhores – que os apresentados

pela distância entre instâncias por correlação. A título de exemplo destaca-se que o RMSE é menor para as previsões obtidas com recurso ao cálculo da distância através da distância Euclidiana e de Minkowski face às obtidas com recurso ao cálculo da distância por Correlação. No caso do RMSE médio as diferenças encontram-se apresentadas na Tabela II-1 do Anexo II. De notar que, mais uma vez, se observa que as previsões obtidas através do uso das métricas de distância Euclidiana e de Minkowski originam previsões e resultados bastante semelhantes entre si, tanto no comportamento do RMSE médio horário como na magnitude do mesmo. A análise das métricas de erro analisadas anteriormente confirma este comportamento.

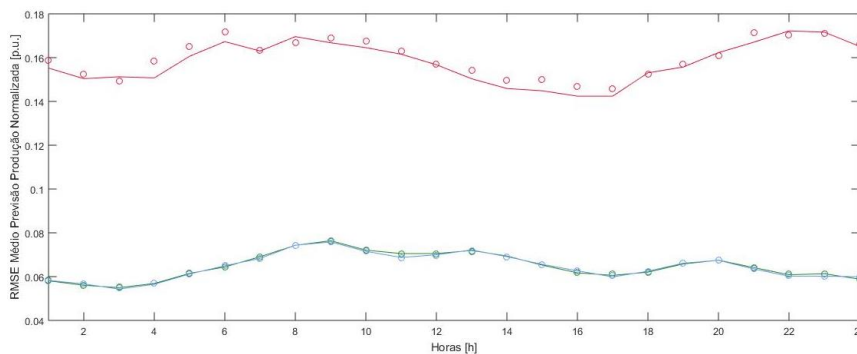
No que diz respeito à variação do RMSE face ao número de vizinhos a considerar verifica-se, para as três métricas de cálculo da distância, o mesmo: à medida que o número de vizinhos, k , aumenta, o RMSE médio obtido para as previsões decresce. De notar, no entanto, que o RMSE não decresce indefinidamente à medida que o valor de k aumenta. Uma análise entre os resultados obtidos para o RMSE médio horário, entre quando k assume o valor de 20 e quando k assume o valor de 162, identifica uma estagnação dos mesmos valores onde o RMSE médio bem como o RMSE médio horário não sofrem variações significativas. O número ótimo de vizinhos a considerar, no estudo realizado no âmbito da presente dissertação, seria, por consequência, o de 20 vizinhos.



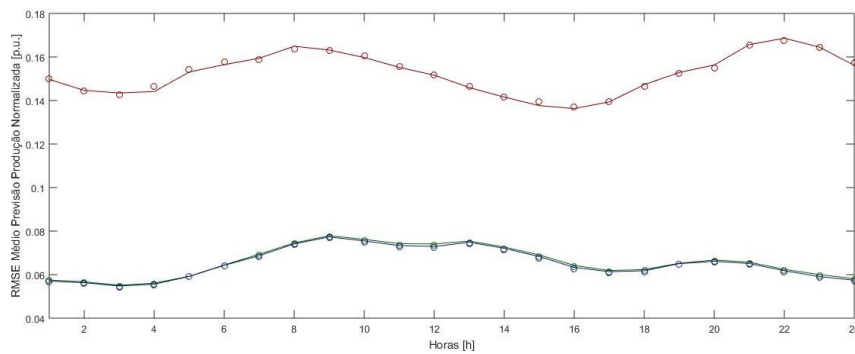
a) $k=1$



b) $k=5$



c) $k=10$



d) $k=162$

Figura 5-5: RMSE médio horário da previsão de produção eólica normalizada, para a metodologia k NN para diferentes valores de k .

Na Figura II-1, presente nos Anexos II, é possível observar o comportamento do RMSE médio horário para a metodologia k NN com foco nas variações do algoritmo que utilizam a métrica de distância Euclidiana e de Minkowski. Verificando-se que estas possuem um comportamento semelhante ao observado para a persistência. Mais uma vez observa-se que o aumento do valor de k reflete-se num decréscimo do RMSE médio horário.

No que toca a importância da ponderação implementada na metodologia k NN, é possível observar o comportamento indicado pela análise de resultados referente às métricas de erro já mencionadas anteriormente: quando k assume o valor de 1 – e não existe possibilidade de realizar ponderação – os resultados, obviamente, não iriam variar. No entanto, para os restantes valores atribuídos a k no âmbito da dissertação, tanto as previsões obtidas como os resultados produzidos pela ponderação, face à não ponderação, são praticamente iguais, *i.e.*, não se verifica que a ponderação, implementada no âmbito da dissertação, tenha expressão significativa no incremento da precisão geral da metodologia de previsão do algoritmo k NN implementada.

- **Desempenho da metodologia k NN versus persistência**

Como referido em 3.2.2.1, a persistência é a metodologia de referência no que toca à viabilidade de implementação de uma metodologia de previsão eólica. Nesse sentido é importante comparar as métricas de erro entre metodologias e avaliar se as previsões obtidas conseguem bater os *standards* da persistência. Com recurso à Tabela 5-1 e Tabela 5-3 é possível fazer essa comparação para a metodologia k NN.

Tabela 5-5: Variação percentual do RMSE médio da previsão de produção eólica resultante da metodologia k NN, face à persistência.

Metodologias k NN vs Persistência				
Distância		Euclidiana	Minkowski	Correlação
k=1	Não-Ponderada	20,14%	19,79%	186,67%
	Ponderada	20,14%	19,79%	186,67%
k=5	Não-Ponderada	-6,29%	-6,10%	128,80%
	Ponderada	-6,08%	-5,93%	132,30%
k=10	Não-Ponderada	-9,32%	-9,56%	121,18%
	Ponderada	-9,30%	-9,53%	123,94%
k=20	Não-Ponderada	-10,55%	-10,63%	116,53%
	Ponderada	-10,69%	-10,80%	118,65%
k=162	Não-Ponderada	-7,85%	-8,67%	113,43%
	Ponderada	-8,60%	-9,30%	113,63%

Para a totalidade de variações do algoritmo *k*NN existem configurações da metodologia que apresentam resultados médios de RMSE que conseguem um melhor desempenho em comparação com a persistência (*benchmark*), enquanto outras não conseguem atingir a performance referência. O caso mais flagrante refere-se à variação do algoritmo *k*NN onde a distância entre instâncias é calculada através da distância de correlação: independentemente do valor que *k* assume, em nenhum dos estudos realizados esta configuração do algoritmo consegue obter um desempenho médio do RMSE melhor que o obtido com recurso à persistência, conforme é observável na Tabela 5-5.

Para as restantes variações de método de cálculo de distância, apenas quando *k* assume o valor de um as previsões obtidas não apresentam resultados satisfatórios, face aos obtidos pela persistência. Nas restantes configurações do algoritmo *k*NN em estudo – referentes às restantes variações do valor de *k* – as previsões obtidas conseguem suplantar as previsões referência, i.e., o RMSE médio obtido pelas previsões de produção eólica quando *k* assume os valores de 5,10, 20 e 162, tanto para a utilização da distância Euclidiana como para a utilização da distância de Minkowski, é mais baixo que o RMSE médio obtido para as previsões com a persistência. Na Tabela 5-5 é possível encontrar discriminadas as variações encontradas para o RMSE médio referente às variações da metodologia *k*NN que superaram a previsão de referência: as variantes da metodologia representadas na tabela conseguiram alcançar melhorias no RMSE médio na ordem de 6% a, aproximadamente, 11%.

Para os resultados referentes ao RMSE médio obtido para as previsões de produção eólica normalizada resultantes da implementação da metodologia *k*NN, o impacto da ponderação nos mesmos é, pela análise da Tabela 5-3 e Figura 5-3, pouco expressivo/significativo. O valor do RMSE médio final obtido para as previsões ponderadas e não ponderadas é praticamente igual, verificando-se variações pouco expressivas: melhorias de 0,75% para o melhor caso e decréscimos de 0,21% para o pior dos casos.

No que toca à métrica de avaliação correlação as correlações encontradas para a metodologia de persistência são mais elevadas do que as verificadas no caso de previsões resultantes da metodologia *k*NN. Mais, nenhuma das configurações de estudo do algoritmo *k*NN atingiu as mesmas grandezas que a persistência.

5.2.3. Redes Neurais Artificiais

A terceira metodologia de previsão analisada, para a dissertação, refere-se à metodologia das redes neuronais artificiais (RNA) com a configuração implementada resultante da análise de sensibilidade, descrita em 5.1.

Na Tabela 5—6 encontram-se discriminados os valores médios das métricas de erro em estudo na presente dissertação, relativas às previsões realizadas com a metodologia RNA, face ao observado para o período de validação.

Tabela 5—6: Valores médios para as métricas de erro, calculadas para as previsões obtidas com a metodologia das redes neuronais artificiais.

Viés [p.u.]	RMSE [p.u.]	Correlação
5.13E-03	0.0605	0.9212

A presença de um viés positivo ou negativo pode indicar uma tendência geral desta metodologia de previsão para uma capacidade de sobre ou subestimação das previsões, face ao observado. No entanto

os valores que se observam no presente estudo são demasiado próximos de zero (0) para ser possível extrair conclusões.

Na Figura 5-6 encontra-se representado o erro médio horário para a metodologia de previsão de produção eólica normalizada em estudo.

O comportamento geral do erro médio horário é semelhante ao observado para as outras metodologias de previsão analisadas no âmbito desta dissertação, apresentando as mesmas flutuações ao longo do dia: um período de pico na sua magnitude, no início do dia, seguido de um decréscimo abrupto e uma ligeira subida ao longo do restante dia. A principal diferença entre a Figura 5-6 e as homólogas para as restantes metodologias de previsão implementadas e analisadas prende-se com a magnitude e a localização temporal em que o pico ocorre. Para a metodologia RNA observa-se que o pico de erro médio horário, para além de ser de menor magnitude, ocorre mais cedo. Verifica-se, no entanto, que o comportamento geral é semelhante às outras métricas. Este comportamento estará relacionado com a variabilidade diária da produção eólica: as alturas em que se verificam que ocorrem os maiores picos no erro de previsão das metodologias em estudo, ocorrem nas mesmas janelas temporais onde se verifica, em média, uma diminuição acentuada da produção na transição da noite e início de dia. Assim, a aplicação das metodologias aplicadas nesta dissertação, para este período do dia devem ser analisados com mais cautela. Durante o resto do dia existem flutuações e, dependendo da metodologia de previsão as flutuações do erro podem ser maiores ou menores, mas nunca tão significativas como o pico verificado no início do dia.

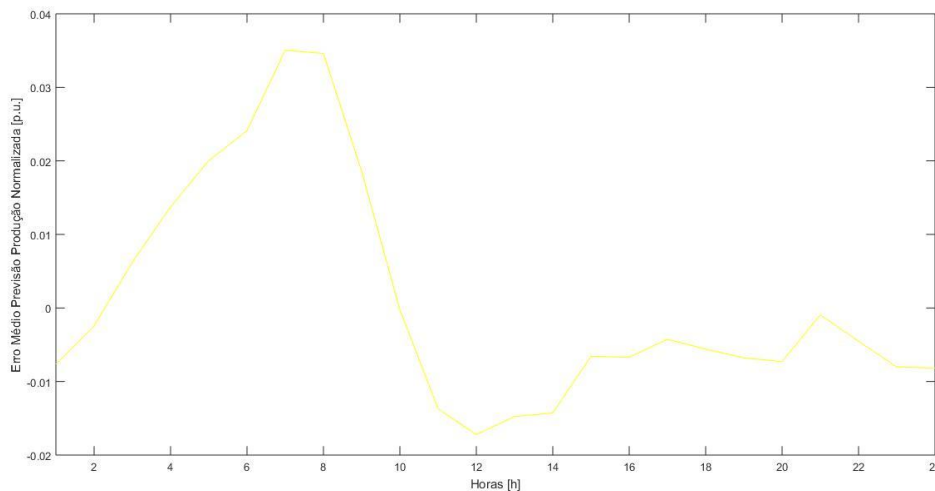


Figura 5-6: Erro médio horário da previsão de produção eólica normalizada, para a metodologia de previsão RNA.

No que toca ao RMSE médio observa-se uma melhoria, da metodologia RNA implementada na presente dissertação, de cerca de 15% face ao resultado referência, obtido para a metodologia de referência – a persistência. O valor de RMSE médio para a metodologia RNA consegue, inclusive, ultrapassar os resultados obtidos através de todas as diferentes configurações da metodologia *kNN*. É, portanto, a metodologia que, por esta métrica, apresenta previsões de produção eólica normalizada com maior fiabilidade.

Na Figura 5-7 encontra-se representado o RMSE médio horário computado para a metodologia de previsão de produção eólica RNA.

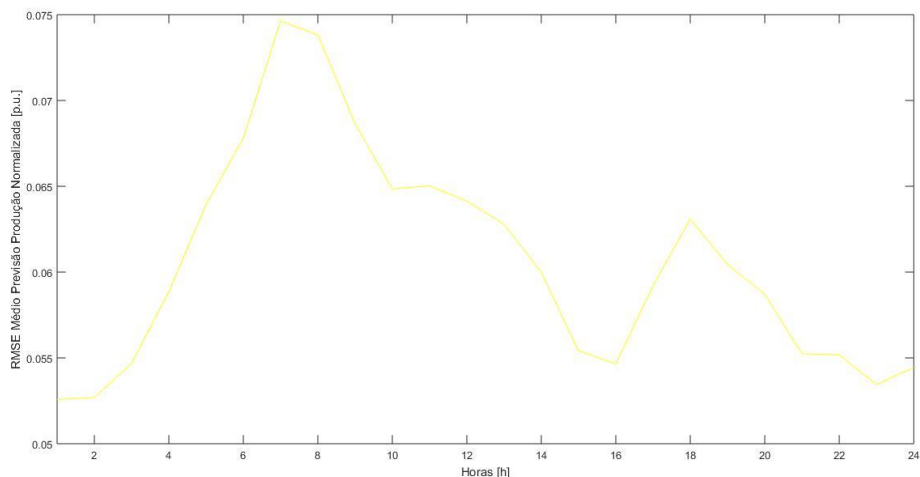


Figura 5-7: RMSE médio horário da previsão de produção eólica normalizada, para a metodologia RNA.

Uma análise da Figura 5-7 face aos resultados homólogos para a persistência e metodologia *k*NN verifica que neste caso o comportamento do RMSE médio horário produz um comportamento ligeiramente diferente das outras metodologias. Para além de apresentar valores do RMSE médio horário mais reduzidos verifica-se também que as flutuações de RMSE médio horário são maiores. Apesar de temporalmente desfasado em relação aos resultados obtidos para a mesma métrica de erro referente às duas metodologias anteriormente apresentadas, observa-se, de novo, um pico significativo de RMSE médio horário no início do dia, seguido de um decréscimo quase tão significativo como a subida e seguido de uma pequena flutuação. O mesmo comportamento é observado tanto para a persistência como para a metodologia *k*NN implementada, com a principal diferença da magnitude do erro, i.e., os valores observados.

No que toca aos valores da correlação observa-se que a implementação da metodologia de previsão das RNA consegue gerar resultados médios que superam as restantes metodologias de previsão implementadas e em estudo no âmbito da presente dissertação. A metodologia de previsão das RNA consegue superar outras restantes metodologias analisadas: face à metodologia *k*NN observam-se melhorias que variam entre, aproximadamente, 104% e 163%; face à correlação média obtida para a metodologia de referência, a implementação da metodologia RNA consegue uma melhoria de apenas 3%. As referidas variações percentuais de correlação média entre a metodologia RNA com as restantes, implementadas na dissertação, podem ser consultadas na Tabela II-2. De referir, no entanto, que apesar de a metodologia RNA apresentar uma estrutura não linear, útil para encontrar complexas relações não lineares presentes no conjunto de dados históricos de um problema, esta é também possível de modelar e processar relações lineares [72]. Justificando, deste modo, os valores de correlação obtidos da implementação da metodologia RNA.

- **Desempenho da metodologia RNA versus persistência**

Na Tabela 5-7 encontram-se discriminados os valores médios da variação percentual do RMSE médio da previsão de produção eólica resultante da metodologia RNA, face à persistência.

Tabela 5-7: Variação percentual do RMSE médio da previsão de produção eólica resultante da metodologia RNA, face à persistência.

Metodologias RNA vs Persistência
-15,27%

Para a metodologia das RNA implementada verifica-se que, face à metodologia da persistência, esta consegue atingir uma melhoria, no que toca à métrica de avaliação do RMSE, na ordem dos 15,27%.

No que diz respeito ao coeficiente de correlação, observável na Tabela 5—6 para a metodologia da RNA e na Tabela 5-1 para a persistência, é possível observar que, também neste caso, a metodologia da RNA conseguiu superar a metodologia da persistência. Este resultado seria expectável, especialmente porque, como descrito em 4.3.1.3, a função de transferência da RNA tenta encontrar um balanço entre a linearidade e a não linearidade da solução.

5.3. Desempenho das previsões nas variações de produção

Para analisar as potencialidades das metodologias de previsão implementadas no âmbito da presente dissertação procurou-se também analisar a capacidade das metodologias de previsão implementadas na deteção das variações de produção num horizonte temporal de 6 horas.

Para tal, utilizando a métrica de avaliação das taxas de variação, apresentadas em 4.4, mapeou-se a diferença entre a taxa de variação das previsões obtidas, pelas metodologias de previsão implementadas, face à taxa de variação verificada para a produção observada, para o período temporal de validação – o ano de 2018.

5.3.1. Metodologia kNN

Para a metodologia dos k -vizinhos mais próximos é possível observar que, quanto ao número ótimo de vizinhos, observa-se que, independentemente da métrica do cálculo da distância ou da implementação de um fator de ponderação, quando k assume valores mais baixos o número de ocorrências encontra-se mais disperso por um maior número de classes. Classes essas onde, quando o valor atribuído a k aumenta, não se observam ocorrências. De notar que apesar de se verificar uma dispersão de ocorrências por diversas classes, em todos os três casos estudados, independentemente do valor atribuído a k , as taxas de variação estão próximas do centro do mapa, o zero; no que toca ao comportamento que toca à métrica de distância entre instâncias para determinar os vizinhos mais próximos verifica-se que os resultados obtidos para este estudo são coerentes com o que se verificou nas análises anteriormente feitas: a implementação das métricas de distância Euclidiana e de Minkowski originam resultados semelhantes entre si – o número de ocorrências tende para concentrar-se nas mesmas classes para as duas métricas de distância, independentemente do valor assumido pelo parâmetro k ; ao invés da implementação da métrica de distância por correlação que fornece resultados díspares dos observados para as duas outras métricas de distância: para todos os valores assumidos por k os resultados obtidos para a distância por correlação apresentam maior número de ocorrências num maior número de classes. Apesar de tudo, a diferença entre as três métricas de cálculo de distância não é significativa: em nenhuma das métricas estudadas, as classes predominantes nunca se traduzem em variações que ultrapassem $\pm 0,1$, ou 10% da TVP, traduzindo-se, conforme observável nas figuras, numa polarização de pequena sub ou sobre estimação; no que toca à expressão da ponderação nas previsões obtidas é possível verificar que a implementação de uma ponderação, baseada no inverso da distância, não provoca alteração nos resultados obtidos para esta análise. Uma comparação entre os mapas de ocorrências, relativos às previsões realizadas sem ponderação e as previsões realizadas com a função de ponderação implementada na presente dissertação, não diferem significativamente, independentemente do valor assumido por k , e seja qual for a métrica de cálculo de distância utilizada.

Para ilustrar o que foi discutido, é possível observar a Figura 5-8, referente ao mapeamento de classes de taxa de variação, figurando as várias configurações da metodologia kNN estudadas e onde k assume o valor de 1.

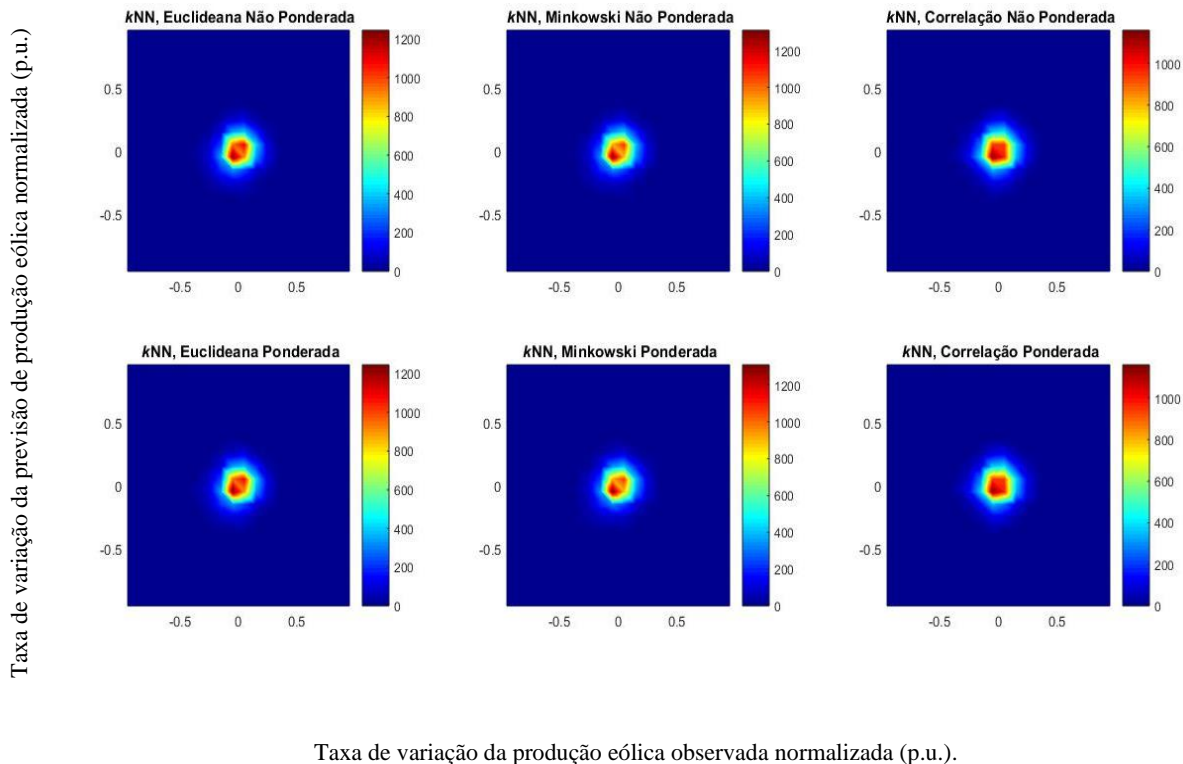


Figura 5-8: Mapeamento das ocorrências obtidas para cada classe, referente à metodologia implementada k NN, onde k assume o valor de 1. A barra de cor indica a escala de cor consoante o número de ocorrências verificado. No eixo dos x estão representadas as taxas de variação da produção eólica normalizada observada, enquanto nos eixos y encontram-se representadas as taxas de variação da previsão de produção eólica normalizada.

É possível, então, observar a concentração de ocorrências que traduzem uma pequena sub ou sobre polarização em classes de variação de $\pm 0,1$, ou 10%. Com a análise das restantes figuras, presentes no Anexo II: Figura II-2 – Figura II-6. é possível observar que quando k aumenta verifica-se um reforço das mesmas classes predominantes, independentemente da configuração da metodologia k NN. Assim sendo não é possível detetar tendências estatísticas relevantes e significativas de modo a entender a capacidade de sub ou sobre estimação desta metodologia.

5.3.2. Metodologia RNA

Para a metodologia RNA implementada verifica-se que a dispersão do número de ocorrências é bastante semelhante à observada na metodologia k NN com distância Euclidiana e de Minkowski, para valores mais elevados de k . Esta semelhança da capacidade de sub ou sobre estimar as previsões obtidas é coerente com a análise das métricas de erro gerais para as acima mencionadas metodologias: à medida que k aumenta verifica-se uma aproximação da metodologia k NN, com a distância Euclidiana e de Minkowski, das métricas de avaliação referentes à metodologia RNA: as classes de ocorrências predominantes são as referentes a $\pm 0,1$, ou 10% da TVP.

Estes resultados, transversais a ambas as metodologias, são, no entanto, inconclusivos: verifica-se que existe uma predominância muito grande, para ambas as metodologias, de um pequeno conjunto de classes. Essa concentração significativa de ocorrências num conjunto polarizado de classes, entre a sub e sobre estimação, não permite detetar resultados estatisticamente relevantes para poder inferir, com certeza, sobre a capacidade de sub ou sobre estimação destas metodologias implementadas. Contudo nesta análise é possível suportar, conforme já foi referido, as anteriores análises acerca do comportamento da otimização das metodologias implementadas na presente dissertação.

A imagem referente a esta análise, para a metodologia RNA, encontra-se presente no Anexo II: Figura II-7.

5.4. Análise do máximo, mínimo e média produção eólica normalizada

A análise que se segue diz respeito a dois casos específicos de interesse: os momentos onde se verificou que, para o período de validação – o ano de 2018 –, se observam os eventos extremos de produção eólica, i.e., quando se observam a produção média eólica máxima e mínima. Para tal a produção eólica normalizada, para cada conjunto de observações de seis horas, foi somada e calculou-se a média para o conjunto desse período. Verificou-se, então, quando se observa o máximo e mínimo desta. Com base nesta abordagem identificaram-se os dois seguintes casos de estudo:

- o máximo da produção eólica média normalizada: 24 de Março entre as 08h e as 13h;
- o mínimo da produção eólica média normalizada: 11h e as 16h do dia 25 de Outubro;

5.4.1. Caso produção eólica normalizada máxima

Na Figura 5-9 encontra-se representado o caso de máxima produção eólica normalizada, onde figuram a matriz alvo – seis horas que antecederam a previsão e que são a base para as previsões –, as previsões realizadas com todas as metodologias de previsão e suas variantes, implementadas na presente dissertação, e ainda a produção eólica normalizada observada, para as seis horas para as quais as previsões foram realizadas.

Na Figura II-8 presente no Anexo II, encontra-se representado graficamente o RMSE médio calculado para cada uma das metodologias de previsão implementadas e em estudo.

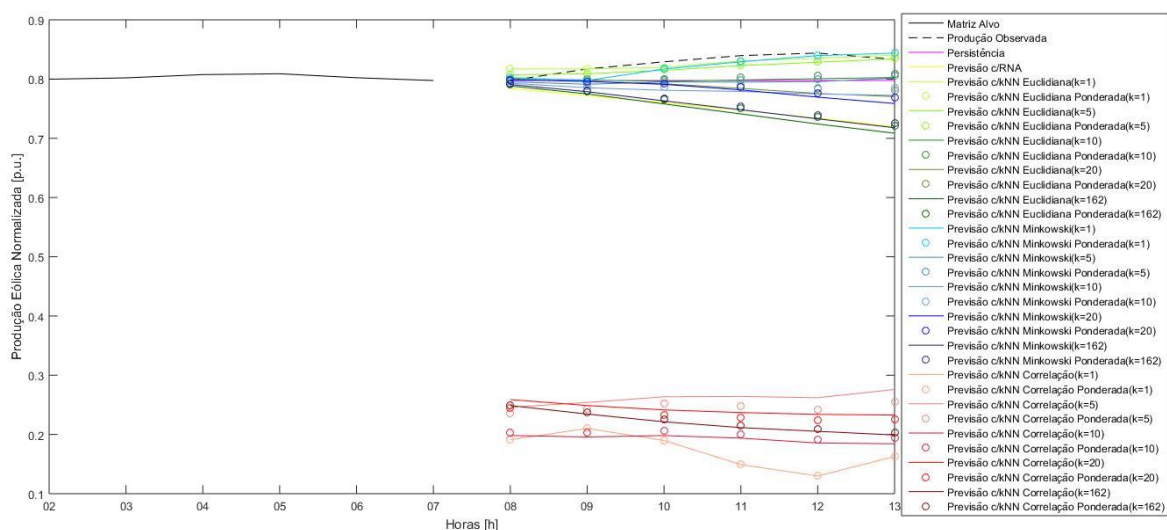


Figura 5-9: Produção eólica normalizada observada e prevista para o caso de produção mínima, para o período de validação.

Uma análise imediata à Figura 5-9 sobressai uma característica que é coerente com as análises anteriormente feitas: a inadequação das previsões obtidas com a variação da metodologia k NN utilizando a Correlação como métrica para o cálculo das distâncias. Neste caso, tanto as previsões obtidas como o RMSE observado, representado na Figura II-8, são díspares e com reduzida fiabilidade, comparando com os obtidos através das restantes metodologias de previsão implementadas: as previsões obtidas apresentam erros e incerteza elevados quando comparado com a produção observada. O RMSE, para estas previsões, apresenta-se significativamente mais alto que o observado para as

restantes metodologias de previsão, surgindo como confirmação da inadequação das previsões. Este comportamento é transversal a qualquer valor assumido por k referente a esta metodologia de previsão.

No que toca às restantes metodologias de previsão implementadas verifica-se que apresentam um comportamento bastante semelhante entre si. É possível, no entanto, identificar as metodologias de previsão que apresentam melhor desempenho, nomeadamente as que conseguem superar as previsões obtidas face à metodologia de referência – a persistência. Estas são a metodologia k NN com a métrica de distância Euclidiana, para valores de k de 1 e 5 – tanto as previsões não ponderadas como as previsões ponderadas –, para quando k assume o valor de 10 – neste caso só para as previsões ponderadas; e a metodologia k NN com a métrica de distância de Minkowski, para valores de k igual a 1. As restantes metodologias de previsão implementadas na presente dissertação não conseguem superar, para este caso específico, a persistência.

De notar que, contrariamente ao que foi observado para as análises anteriores, para este caso específico à medida que o valor de k aumenta o respetivo RMSE das variações da metodologia k NN implementada também tende a aumentar. Do mesmo modo, a implementação de uma regra de ponderação apresenta algum efeito no valor do RMSE obtido para as previsões. Para valores de k mais baixos a implementação de uma ponderação não apresenta expressão na avaliação das previsões. Verifica-se, no entanto que à medida que k aumenta a regra de ponderação tem alguma expressão, como é o caso, por exemplo, da metodologia k NN com a distância Euclidiana onde, quando k assume o valor de 10, a previsão não ponderada não consegue bater a persistência e a previsão ponderada consegue.

De referir que todas as metodologias de previsão de produção eólica implementadas no âmbito da presente dissertação, neste caso específico, tendem a subestimar as previsões, face ao observado.

Estes resultados servem para ilustrar que apesar de, como anteriormente referido, a metodologia de previsão k NN apresentar pior desempenho geral que a metodologia de previsão RNA, e que de todas as variações da metodologia k NN implementadas e estudadas o grau de fiabilidade e eficácia das previsões obtidas pode variar consoante os vários parâmetros em estudo (seja o valor de k seja a métrica de cálculo da distância ou até mesmo a implementação de uma regra de ponderação), esta metodologia pode apresentar o melhor desempenho para alguns casos particulares.

5.4.2. Caso produção eólica normalizada mínima

Na Figura 5-10 encontra-se representado o caso de máxima produção eólica normalizada, onde figuram a matriz alvo – seis horas que antecederam a previsão –, as previsões realizadas com todas as metodologias de previsão e suas variantes, implementadas na presente dissertação, e ainda a produção eólica normalizada observada, para as seis horas para as quais as previsões foram realizadas.

Na Figura II-9, presente no Anexo II, encontra-se representado graficamente o RMSE médio calculado para cada uma das metodologias de previsão implementadas em estudo.

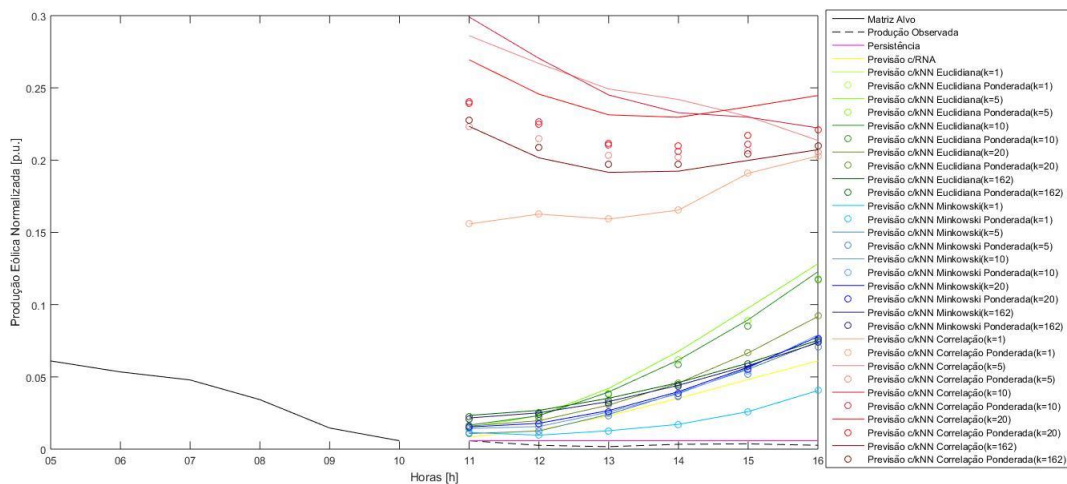


Figura 5-10: Produção e previsão eólica normalizada para o caso de produção mínima, para o período de validação.

Como em 5.4.1 é possível identificar a característica mais evidente: a inadequação das previsões obtidas através da metodologia kNN que utiliza a correlação como cálculo de distância. Novamente verifica-se que as previsões obtidas através desta metodologia de previsão fornecem resultados bastante díspares tanto das previsões obtidas através das restantes metodologias de previsão como da produção eólica normalizada observada. O RMSE calculado para esta metodologia de previsão apresenta-se, mais uma vez, significativamente mais alto do que o apresentado pelas restantes metodologias de previsão em estudo. Esta característica é, mais uma vez, transversal a qualquer valor de k aplicado para a metodologia em causa.

Verifica-se que o comportamento das restantes metodologias de previsão implementadas e em estudo é, no entanto, no presente caso de mínima produção média eólica normalizada bastante diferente do observado no caso anterior: aqui nenhuma das metodologias de previsão implementadas consegue superar o desempenho obtido para a metodologia de referência, i.e., a metodologia de previsão que fornece melhores resultados é a metodologia de referência, a persistência.

As metodologias de previsão que mais se aproximam da persistência, tanto em comportamento como o em métrica de erro são, por ordem crescente, a metodologia kNN com a distância Euclidiana e de Minkowski – com k a assumir o valor de um – e a metodologia RNA. Mais uma vez se verifica que apesar de se atingir um melhor desempenho geral das previsões através da implementação da metodologia RNA, a metodologia kNN pode, em alguns casos, render um melhor desempenho.

Devido à abordagem simplista da persistência, como anteriormente observado, é possível contextualizar o melhor desempenho de previsão desta, face às metodologias de previsão implementadas no caso em que a produção eólica normalizada é mínima: os valores relativamente à produção eólica normalizada são baixos e tendem a manter-se baixos. Pelo contrário, as metodologias de previsão implementadas tendem, neste caso, a sobre estimar a produção eólica normalizada. É possível verificar que a metodologia RNA e kNN apresentam um crescimento nos valores previstos. Isto significa que historicamente o comportamento nas seis horas de previsão é atípico e por isso nenhuma metodologia teve capacidade de estimar com rigor os valores observados.

Para a metodologia kNN também os resultados são coerentes: com a métrica de distância Euclidiana e de Minkowski, quanto mais alto for o valor de k a metodologia tende a sobre estimar as previsões.

De notar que neste caso, contrariamente ao observado em 5.4.2, a ponderação não apresenta expressão nas previsões obtidas: os resultados obtidos pelas previsões ponderadas não superam,

significativamente, os resultados obtidos pelas previsões não ponderadas, mesmo para valores mais elevados de k .

5.5. Análise da sazonalidade do erro de previsão

A análise que se segue diz respeito ao estudo da sazonalidade nos resultados das previsões e a fiabilidade das previsões obtidas através da implementação das metodologias de previsão estudadas na presente dissertação. Nesse sentido foi determinado o perfil diário do erro por estação do ano. Este perfil representa a média horária do erro para todo o período de tempo que engloba a estação. Para o Inverno considerou-se os meses de Dezembro, Janeiro e Fevereiro. Do mesmo modo, a Primavera engloba todos os meses de Março, Abril e Maio; o Verão engloba os meses de Junho, Julho e Agosto, e o Outono engloba os meses de Setembro, Outubro e Novembro. Deste modo foi possível determinar o RMSE médio com vista a avaliar as metodologias de previsão implementadas na presente dissertação.

Na Figura II-10 à Figura II-13, em anexo, encontram-se representados os perfis diários do RMSE para estação e para todas as metodologias de previsão de produção eólica implementadas na presente dissertação, para o inverno, primavera, verão e outono, respetivamente.

Uma análise rápida às figuras acima mencionadas corrobora uma conclusão já anteriormente descrita: a disparidade dos resultados na metodologia k NN, com o cálculo da distância da Correlação, face as restantes distâncias. A configuração da metodologia k NN com a distância de correlação origina previsões menos fiáveis que as restantes metodologias implementadas – tanto duas outras variantes da metodologia k NN estudada como para a metodologia RNA. Esta metodologia não constitui, portanto, interesse comparativo com as restantes metodologias.

Nesse sentido nas Figura 5-11 — Figura 5-14 figuram isolados os perfis da métrica RMSE médio, referente aos dias típicos de cada estação, da metodologia RNA e k NN, com a distância Euclidiana e de Minkowski.

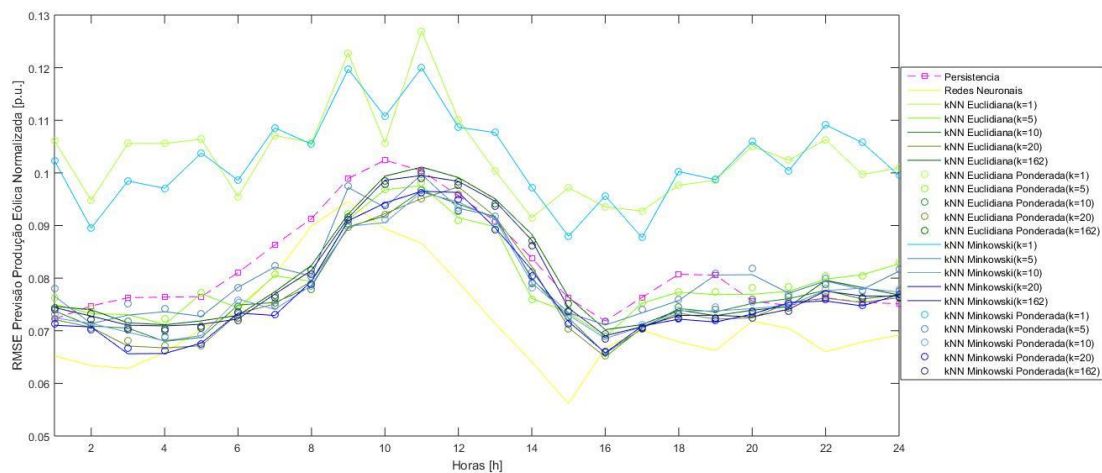


Figura 5-11: Perfil de RMSE médio do referente à estação do Inverno, para a metodologia de previsão RNA e k NN, com foco nas variações com a distância Euclidiana e de Minkowski.

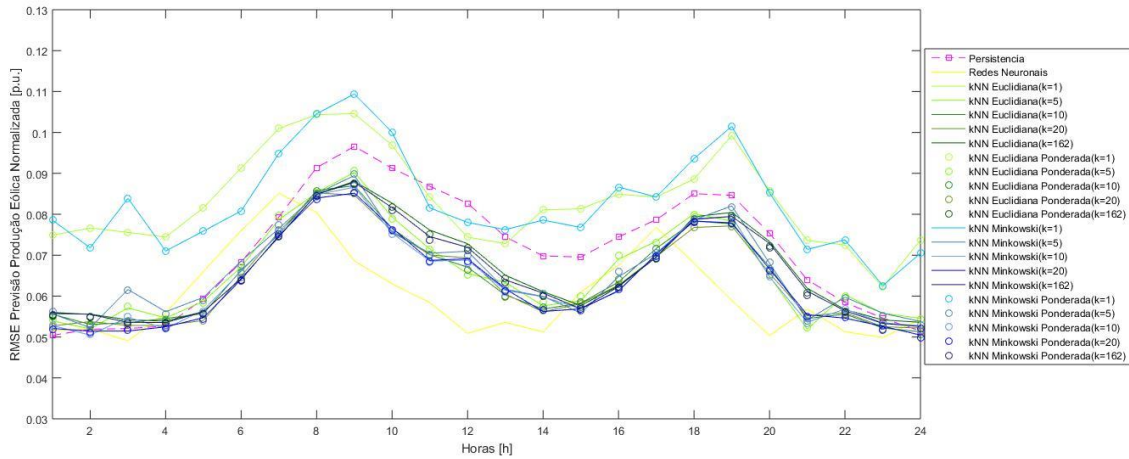


Figura 5-12: Perfil de RMSE médio referente à estação do Outono, para a metodologia de previsão RNA e k NN, com foco nas variações com a distância Euclidiana e de Minkowski.

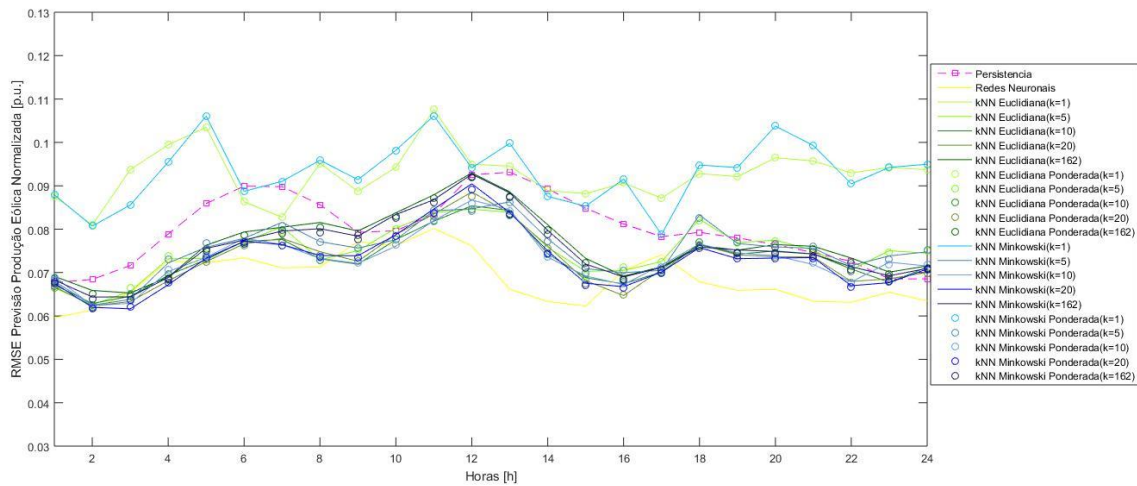


Figura 5-13: Perfil de RMSE médio referente à estação da Primavera, para a metodologia de previsão RNA e k NN, com foco nas variações com a distância Euclidiana e de Minkowski.

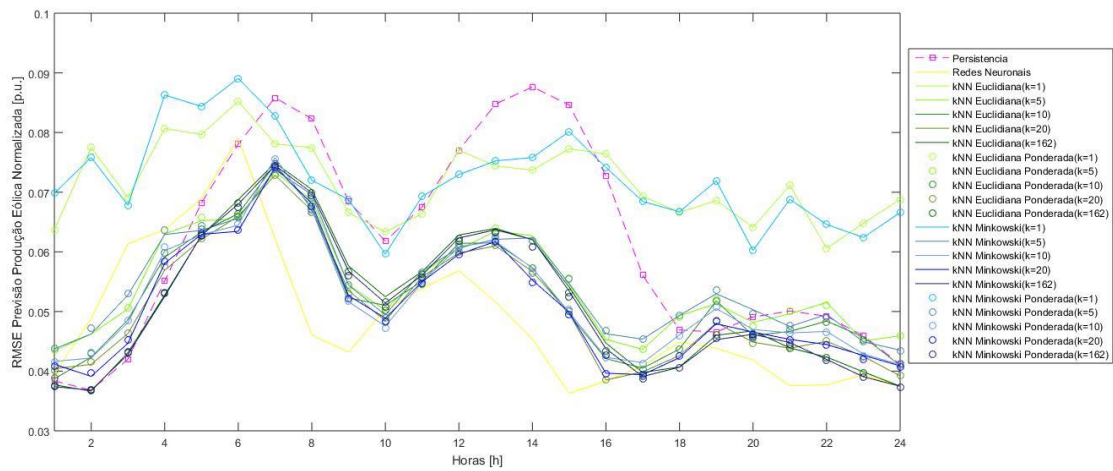


Figura 5-14: Perfil de RMSE médio referente à estação do Verão, para a metodologia de previsão RNA e k NN, com foco nas variações com a distância Euclidiana e de Minkowski.

A análise das figuras torna possível retirar algumas conclusões acerca das previsões obtidas através da implementação das metodologias de previsão em estudo na presente dissertação, para o comportamento sazonal:

- Os erros de previsão, medidos pelo RMSE horário, são, geralmente, menores de noite que durante o dia. De acordo com o perfil diário de produção eólica, a variabilidade da produção é menor neste período;
- O Verão apresenta, no geral, o menor erro de previsão, para todas as metodologias de previsão de produção eólica implementada. Esta estação corresponde, também, à estação onde se verifica uma menor produção eólica normalizada para o ano de 2018. Em contrapartida, a primavera – época do ano onde se verifica uma maior produção eólica normalizada para o ano de 2018 – é a que apresenta maior erro de previsão, também para todas as metodologias de previsão de produção eólica implementadas na presente dissertação;
- Destaca-se ainda que na Primavera o pico de erro de previsão de produção eólica dá-se por volta das 12h, enquanto para o Outono e Inverno ocorrem durante a manhã – 9h e 11h, respetivamente. Para o verão este ocorre durante a noite – aproximadamente pelas 05h;
- Com a exceção do dia Verão, a variação entre as metodologias de previsão implementadas na presente dissertação, não é significativa sendo possível verifica que o comportamento do RMSE é semelhante entre para as metodologias de previsão implementadas. No entanto verifica-se que, no Verão, as previsões obtidas através da metodologia RNA diferem – entre as 00h e as 6h da manhã – face às restantes metodologias em estudo: o RMSE médio horário é alto, face ao que se observa nas restantes figuras e até nas restantes metodologias. É nessa altura que se observa o pico de RMSE médio horário para o referido dia típico. Para esta estação, a metodologia RNA fornece durante largos períodos previsões menos fiáveis que a maioria das variações da metodologia kNN ;
- Nem todas as variações da metodologia kNN conseguem superar a persistência: para valores baixos de k a metodologia kNN , tanto com a distância Euclidiana como a distância de Minkowski, raramente consegue superar os resultados face à persistência – apenas das 7h às 16h no Verão. No entanto, à medida que o parâmetro k aumenta o seu valor os perfis de RMSE horário para cada dia típico, referente à metodologia kNN considerando a distância Euclidiana e de Minkowski, conseguem superar a persistência, verificando-se uma melhoria na performance da referida metodologia de previsão. Quando k assume o valor de 20 ou 162 verifica-se que, em ambas as variações da metodologia kNN – distância Euclidiana ou de Minkowski, o perfil do RMSE médio horário consegue melhorar face à metodologia de referência;
- A metodologia RNA consegue, de modo geral, superar a persistência, em todas as estações. Os únicos casos em que tal não se verifica ocorrem durante a madrugada do Verão e do Outono;

5.6. Síntese da análise dos resultados

De seguida apresenta-se uma síntese da análise dos resultados obtidos com base na informação dos subcapítulos anteriores:

- A semelhança de comportamento e valores obtidos para as métricas de avaliação estudadas entre a metodologia kNN , com as distâncias Euclidiana e de Minkowski;
- A abordagem de ponderação dos k vizinhos mais próximos, implementada na presente dissertação, não produz melhorias significativas nas previsões obtidas pela metodologia kNN ;
- À medida que o parâmetro k aumenta, referente à metodologia kNN verifica-se que ocorre uma melhoria dos perfis do RMSE médio horário – este diminui. No entanto verifica-se, mais uma vez, que a melhoria entre os perfis para quando k assume o valor de 20 e de 162 apresenta

variações muito pouco expressivas, corroborando a análise anteriormente feita que se prende com a melhoria do RMSE face ao aumento do parâmetro k ;

- A metodologia RNA apresenta previsões de produção eólica com melhor desempenho no geral. No entanto é possível que as previsões obtidas através da metodologia k NN consigam superar as resultantes da metodologia RNA em alguns casos.
- Os erros de previsão, para as metodologias de previsão implementadas na presente dissertação com melhor desempenho, variam até ao máximo de 10%, face à produção observada, da capacidade instalada em Portugal.
- Relativamente à metodologia de *benchmark*, a persistência, foi possível superar o desempenho desta em ambos os casos estudados na presente dissertação: tanto com a implementação da metodologia RNA como com parte das configurações estudadas da metodologia k NN, incluindo a melhor configuração desta.

6. Conclusões finais

Atendendo aos objetivos da dissertação, foram implementadas com sucesso, duas metodologias de previsão de produção eólica para Portugal continental considerando um horizonte temporal de seis horas. Essas metodologias – k Vizinhos mais próximos e redes neuronais artificiais - foram selecionadas devido às suas características intrínsecas. A metodologia k Vizinhos mais próximos (k -nearest neighbors ou k NN) é caracterizada na literatura por ser uma abordagem de simples implementação, não paramétrica, que origina resultados robustos. A metodologia baseada em redes neuronais artificiais (RNA), de natureza não linear, não paramétrica, com uma forte capacidade de adaptabilidade que permite treinar e iterativamente mapear as relações entre os dados de entrada e saída. São, ambas, metodologias de previsão amplamente utilizadas, transversalmente, em múltiplos campos. Desse modo foi possível não só avaliar a eficácia e fiabilidade de cada uma das metodologias para a previsão de produção eólica, mas também avaliar a sua otimização.

De acordo com os resultados obtidos, analisados e discutidos nas secções anteriores, ambas as metodologias implementadas conseguiram superar, no geral, a metodologia de previsão com base na persistência (usada como *benchmark*) nas métricas utilizadas na presente dissertação, nomeadamente, o coeficiente de correlação, o viés e o erro quadrático médio.

No que diz respeito à tendência geral do erro médio horário das previsões, geradas por ambas as metodologias implementadas na presente dissertação, foi possível observar que o comportamento e variação, ao longo do dia, acompanham os períodos de transição de produção eólica: durante a noite e início do dia, onde as flutuações de produção eólica são mais significativas, o erro associado às previsões obtidas por ambas as metodologias implementadas é mais elevado. Durante o resto do dia existem flutuações, tanto de produção eólica como do erro associado às previsões obtidas, mas nunca tão significativas como nessas janelas temporais de transição.

Na metodologia dos k vizinhos mais próximos, verificou-se que na maioria das diferentes configurações esta consegue superar a metodologia de *benchmark*. Foi estudada também a otimização de dois parâmetros desta metodologia e analisado o impacto dessa otimização na eficácia geral das previsões realizadas. Relativamente ao primeiro parâmetro estudado, a métrica de distância entre instâncias, verificou-se que as previsões obtidas com recurso às duas métricas de distância da mesma família, a distância Euclidiana e de Minkowski, apresentavam não só o mesmo comportamento, mas também resultados semelhantes entre si. A terceira métrica de distância estudada, a distância por correlação, deu origem a resultados completamente díspares e inadequados, chegando mesmo a não conseguir superar a metodologia de *benchmark*. Este resultado sugere que é mais importante o nível de potência nos

intervalos anteriores (como identificado pelas distâncias Euclidiana e de Minkowski) do que o comportamento/variabilidade nos períodos antecedentes à previsão. O segundo parâmetro estudado para efeitos de otimização da metodologia foi o número de vizinhos - k . Nesse sentido verificou-se que, numa tendência geral presente em todas as configurações da metodologia estudadas, à medida que o número de vizinhos aumenta também o desempenho das previsões resultantes aumenta. O terceiro e último parâmetro estudado para esta metodologia prendeu-se com a implementação de um fator de ponderação dos k vizinhos, de acordo com grau de similaridade nas observações históricas, para o cálculo das previsões. Nas configurações estudadas da metodologia a implementação de um fator de ponderação não origina qualquer variação significativa no desempenho das previsões de produção eólica.

Na metodologia das redes neuronais artificiais – de natureza não linear –, foi analisada igualmente diferentes configurações no que diz respeito ao número de neurónios presentes na camada escondida. A melhor configuração da rede neuronal artificial identificada superou a metodologia de benchmark nas métricas de avaliação em uso na presente dissertação..

A análise dos resultados centrou-se igualmente na análise da capacidade das metodologias para sub ou sobre estimação (taxas de variação) e na dependência dos erros de acordo com a sazonalidade. Estas análises ajudam a confirmar as tendências gerais observadas e descritas anteriormente, verificando-se que a metodologia que apresentou um maior grau geral de fiabilidade e eficácia foi, na generalidade dos casos estudados, a metodologia das RNA. Estes resultados podem ser maioritariamente explicados com a: i) capacidade para lidar com fenómenos de natureza não linear da metodologia em causa e, ii) capacidade de treino que permite minimizar os erros.

Para finalizar é também importante também fazer uma comparação entre a eficácia e fiabilidade das metodologias de previsão implementadas no presente trabalho face ao que foi identificado, aquando da revisão bibliográfica. Nesse sentido verificou-se que, face aos valores de referência identificados na revisão bibliográfica, para o mesmo horizonte temporal que se foca a presente dissertação, as metodologias de previsão implementadas apresentam um desempenho superior.

Trabalhos futuros

Inevitavelmente aos sistemas de previsão de produção eólica a incerteza e os erros de previsão encontram-se transversalmente presentes. No entanto o objetivo central da investigação, desenvolvimento e implementação destas metodologias visa a diminuição da diferença entre a potência eólica prevista e a potência eólica observada. Nesse sentido será do interesse explorar e comparar um maior número de metodologias de previsão que, apesar de maior complexidade – tanto de modelação como de implementação – possam ajudar a reduzir esta incerteza. Existe a possibilidade de desenvolver modelos de previsão de produção eólica de segunda ordem, i.e., cujas variáveis de entrada podem ser as mesmas previsões de produção eólica num processo de retroalimentação. Esta possibilidade acarreta, no entanto, uma complexidade acrescida devido à natureza estocástica do recurso eólico. Simultaneamente, os benefícios da utilização de previsões obtidas com base em modelos numéricos de mesoescala para o horizonte temporal em análise devem ser estudados.

Uma análise de metodologias de previsão integrada com uma análise de rampas de produção poderia também apresentar algum interesse para a gestão da rede, uma vez que estes fenómenos podem funcionar como disrupções no equilíbrio do SE nacional.

Seria também do interesse que fosse ainda incluída nesta análise o efeito que estas metodologias de previsão podem produzir nos níveis de necessidade de reservas de potência, juntamente com uma análise económica de custos e poupanças do SE.

Referências Bibliográficas

- [1] European Commission. Green Paper: A 2030 framework for climate and energy policies. Brussels. Disponível em <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52013DC0169&from=EN> (acedido a 11 de Janeiro de 2021).
- [2] Gabinete do Ministro do Ambiente e Acção Climática (2020). Plano Nacional Energia e Clima 2030 aprovado em Conselho de Ministros. Lisboa. Disponível em <https://www.portugal.gov.pt/download-ficheiros/ficheiro.aspx?v=%3D%3DBAAAAB%2BLCAAAAAAABACztDQyAgDZs3qiBAAA%3D%3D> (acedido em 11 de janeiro de 2021).
- [3] Agência Portuguesa do Ambiente & Comité Executivo da Comissão para as Alterações Climáticas. (2012). Roteiro Nacional de Baixo Carbono: Análise técnica das opções de Transição para uma economia de baixo carbono competitiva em 2050. Lisboa. Disponível em https://www.apambiente.pt/_zdata/DESTAQUES/2012/RNBC_COMPLETO_2050_V04.pdf (acedido em 11 de janeiro de 2021).
- [4] DGEG/MAAC e PORDATA, 2020, Produção de energia eléctrica a partir de fontes renováveis (%). Qual a percentagem da electricidade que é obtida de fontes não fósseis e não minerais? Disponível em [https://www.pordata.pt/Portugal/Produ%C3%A7%C3%A3o+de+energia+el%C3%A9ctrica+a+p+artir+de+fontes+renov%C3%A1veis+\(percentagem\)-1232](https://www.pordata.pt/Portugal/Produ%C3%A7%C3%A3o+de+energia+el%C3%A9ctrica+a+p+artir+de+fontes+renov%C3%A1veis+(percentagem)-1232) (acedido em 10 de junho de 2020).
- [5] APREN, 2021, Balanço da Produção de Eletricidade de Portugal Continental. Abril 2021. Disponível em <https://www.apren.pt/pt/energias-renovaveis/producao> (acedido a 10 de maio de 2021).
- [6] A. R. Silva. “Modelação e otimização do funcionamento de uma central renovável virtual”. (Dissertação para obtenção do grau de Mestre em Engenharia da Energia e do Ambiente). Faculdade de Ciências da Universidade de Lisboa, Lisboa, 2016.
- [7] N. Son, S. Yang, e J. Na. Hybrid forecasting model for short-term wind power prediction using modified long short-term memory. *Energies*, vol. 12, no. 20, 2019.
- [8] F. Kullmann. “Efficiency and Functionality of a Contra-Rotation wind turbine”. (Dissertação para obtenção graduação em engenharia ambiental). Satakunta University of Applied Sciences, Satakunta, Finlândia, 2015.
- [9] IEA. “Variability of Wind Power and Other Renewables – Management options and strategies”. Relatório técnico, p. 57, 2005.
- [10] H. Holttinen, P. Meibom, A. Orth, F. Van Hulle, B. Lange, M. O ’malley e E. Ela. “Design and operation of power systems with large amounts of wind power”, Final report, Phase one 2006-2008, 2008.
- [11] A. Estanqueiro. “The future energy mix paradigm: How to embed large amounts of wind generation while preserving the robustness and quality of the power systems?”. Chapter 22, in *Wind Power*, no. June 2009, p. 22, 2009.
- [12] J. Kiviluoma *et al.* “Variability in large-scale wind power generation”. *Wind Energy*, vol. 19, n. 9, pp. 1649–1665, Set. 2016.

- [13] A. Estanqueiro. “Assessment of wind power fluctuations in power systems: study on the smoothing effects”. In Proc. of European Wind Energy Conference and Exhibition 2008, Brussels, March 2008.
- [14] I. Okumus e A. Dinler. “Current status of wind energy forecasting and a hybrid method for hourly predictions”. *Energy Convers. Manag.*, vol. 123, pp. 362–371, Sep. 2016.
- [15] E. Ela, M. Milligan, e B. Kirby. “Operating Reserves and Variable Generation”. Golden, CO (United States), Ago. 2011.
- [16] M. Godinho e R. Castro. “Comparative performance of AI methods for wind power forecast in Portugal,” *Wind Energy*, vol. 24, no. 1, pp. 39–53, 2021.
- [17] W.-Y. Chang. “A Literature Review of Wind Forecasting Methods,” *J. Power Energy Eng.*, vol. 02, no. 04, pp. 161–168, Jan. 2014.
- [18] S. S. Soman, H. Zareipour, O. Malik, and P. Mandal. “A review of wind power and wind speed forecasting methods with different time horizons,” in *North American Power Symposium 2010*, no. October, pp. 1–8, 2010.
- [19] S. Hanifi, X. Liu, Z. Lin, e S. Lotfian. “A Critical Review of Wind Power Forecasting Methods—Past, Present and Future,” *Energies*, vol. 13, no. 15, p. 3764, Jul. 2020.
- [20] S.M. Lawan, W. A. W. Z. Abidin, W. Y. Chai, A. Baharun e T. Masri. “Different Models of Wind Speed Prediction; A Comprehensive Review”. *International Journal of Scientific & Engineering Research*, 5(1), 1760–1768, 2014.
- [21] G. Giebel, R. Brownsword, G. Kariniotakis, M. Denhard, e C. Draxl. “The State-Of-The-Art in Short Term Prediction of Wind Power: A Literature Overview”, 2nd edition. ANEMOS.plus, 2011.
- [22] X. Wang, P. Guo, E X. Huang. “A review of wind power forecasting models”. *Energy Procedia*, vol. 12, pp. 770–778, 2011.
- [23] California Landscape Conservation Partnership. No Title. 2016. Disponível em <http://climate.calcommons.org/article/downscaling> (acedido em 30 de novembro de 2018).
- [24] GIS Program at National Center for Atmospheric Research (NCAR). (n.d.). Retirado a novembro 30, 2018, disponível em <https://gisclimatechange.ucar.edu/question/63>.
- [25] M. P. Emhart. “Exploring Machine Learning Models for Wind Speed Prediction”. (Dissertação para obtenção do grau de Mestre de Ciências em Inteligência Artificial). Universidade Politécnica da Catalunha, Barcelona, Espanha, 2018.
- [26] Y. Liu. “Machine Learning for Wind Power Prediction”.(Dissertação para obtenção do grau de Mestre em Ciência Informática). Universidade de Nova Brunswick, Fredericton, Canadá, 2016.
- [27] S. Vidya e E. S. V. Janani. “A review on the hybrid approaches for wind speed forecasting,” *Int. J. Sci. Technol. Res.*, vol. 8, no. 9, pp. 1584–1590, 2019.
- [28] M. Yesilbudak, S. Sagiroglu, e I. Colak. “A new approach to very short term wind speed prediction using k-nearest neighbor classification,” *Energy Convers. Manag.*, vol. 69, pp. 77–86, 2013.

- [29] Y. Ren, P. N. Suganthan, e N. Srikanth. “Ensemble methods for wind and solar power forecasting - A state-of-the-art review,” *Renew. Sustain. Energy Rev.*, vol. 50, pp. 82–91, 2015.
- [30] Z. Li, L. Ye, Y. Zhao, X. Song, J. Teng, and J. Jin. “Short-term wind power prediction based on extreme learning machine with error correction,” *Prot. Control Mod. Power Syst.*, vol. 1, no. 1, p. 1, Dec. 2016.
- [31] M. Ferreira, A. Santos, e P. Lucio. “Short-term forecast of wind speed through mathematical models,” *Energy Reports*, vol. 5, pp. 1172–1184, Nov. 2019.
- [32] T. H. M. El-Fouly, E. F. El-Saadany, M. M. A. Salama, T. H. M. El-Fouly, E. F. El-Saadany, e M. M. A. Salama. “Grey Predictor for Wind Energy Conversion Systems Output Power Prediction,” *IEEE Trans. Power Syst.*, vol. 21, no. 3, pp. 1450–1452, Aug. 2006.
- [33] K. Pan, Z. Qian, e N. Chen. “Probabilistic Short-Term Wind Power Forecasting Using Sparse Bayesian Learning and NWP,” *Math. Probl. Eng.*, vol. 2015, no. May 2016, 2015.
- [34] S. X. Wang, M. Li e L. Zhao e C. Jin. “Short-term wind power prediction based on improved small-world neural network,” *Neural Comput. Appl.*, vol. 31, no. 7, pp. 3173–3185, 2019.
- [35] R. Jursa e K. Rohrig. “Short-term wind power forecasting using evolutionary algorithms for the automated specification of artificial intelligence models,” *Int. J. Forecast.*, vol. 24, no. 4, pp. 694–709, 2008.
- [37] N. Korprasertsak e T. Leephakpreeda. “Robust short-term prediction of wind power generation under uncertainty via statistical interpretation of multiple forecasting models,” *Energy*, vol. 180, pp. 387–397, Aug. 2019.
- [38] D. Yuan, Z. Qian, B. Jing, e Y. Pei. “Short-term wind speed forecasting using STLSSVM hybrid model,” in *2018 International Conference on Power System Technology (POWERCON)*, no. November, pp. 1661–1667, 2018.
- [39] REN, “REN - Centro de Informação,” 2020. Disponível em : <http://www.centrodeinformacao.ren.pt/PT/Paginas/CIHomePage.aspx>. (acedido em 08 de Janeiro de 2020).
- [40] H. dos Santos. “Comparação da performance de algoritmos de machine learning para a análise preditiva em saúde pública e medicina”. (Dissertação para obtenção do grau de Doutor em Ciências). Faculdade de Saúde Pública da Universidade de São Paulo, São Paulo, Brasil, 2018.
- [41] H. A. Abu Alfeilat *et al.* “Effects of Distance Measure Choice on K-Nearest Neighbor Classifier Performance: A Review,” *Big Data*, vol. 7, no. 4, pp. 221–248, Dec. 2019.
- [42] C. Junk, L. D. Monache e S. Alessandrini. “Analog-based ensemble model output statistics,” *Mon. Weather Rev.*, vol. 143, no. 7, pp. 2909–2917, 2015.
- [43] A. Teixeira. “Machine Learning methods for quantitative structure-property relationship modeling”. (Dissertação para obtenção do grau de Doutor em Informática). Faculdade de Ciências da Universidade de Lisboa, Lisboa, Portugal, 2014.

- [44] A. B. Hassanat, M. A. Abbadi, G. A. Altarawneh, and A. A. Alhasanat. “Solving the Problem of the K Parameter in the KNN Classifier Using an Ensemble Learning Approach,”. *International Journal of Computer Science and Information Security*. 12. 33-39, Sep. 2014.
- [45] H. A. Abu Alfeilat *et al.* “Effects of Distance Measure Choice on K-Nearest Neighbor Classifier Performance: A Review,” *Big Data*, vol. 7, no. 4, pp. 221–248, Dec. 2019.
- [46] M. Alkasassbeh e G. Altarawneh e A. Hassanat. “On Enhancing The Performance Of Nearest Neighbour Classifiers Using Hassanat Distance Metric”. *Indian Journal of Pure and Applied Mathematics*. 9. 3291-3298, Jan. 2015.
- [47] L. Kozma. “k Nearest Neighbors algorithm (kNN)” [Apresentação PowerPoint]. Retirado de László Kozma: <http://www.lkozma.net/knn2.pdf>, 2008.
- [48] G. F. Fan, Y. H. Guo, J. M. Zheng, and W. C. Hong. “Application of the weighted k-nearest neighbor algorithm for short-term load forecasting,” *Energies*, vol. 12, no. 5, 2019.
- [49] A. T. Lora, J. M. R. Santos, A. G. Exposito, J. L. M. Ramos, and J. C. R. Santos. “Electricity Market Price Forecasting Based on Weighted Nearest Neighbors Techniques,” *IEEE Trans. Power Syst.*, vol. 22, no. 3, pp. 1294–1301, Aug. 2007.
- [50] K. Hechenbichler & K. Schliep. “Weighted k-Nearest-Neighbor Techniques and Ordinal Classification”. Discussion paper. Sonderforschungsbereich 399, 2004.
- [51] R. Shahid, S. Bertazzon, M. L. Knudtson, e W. A. Ghali. “Comparison of distance measures in spatial analytical modeling for health service planning”, *BMC Health Serv. Res.*, vol. 9, n. May 2014, 2009.
- [52] R. Matsumura, K. Harada, Y. Domae, e W. Wan. “Learning Based Industrial Bin-Picking Trained with Approximate Physics Simulator”, em *Advances in Intelligent Systems and Computing*, vol. 867, pp. 786–798, 2019.
- [53] A. P. Marugán, F. P. G. Márquez, J. M. P. Perez, e D. Ruiz-Hernández. “A survey of artificial neural network in wind energy systems”. *Appl. Energy*, vol. 228, n. October, pp. 1822–1836, 2018.
- [54] G. Zhang, B. Eddy Patuwo, e M. Y. Hu. “Forecasting with artificial neural networks”. *Int. J. Forecast.*, vol. 14, n. 1, pp. 35–62, Mar. 1998.
- [55] H. Simon. “Neural Networks – A comprehensive Foundation” (9th edition). Pearson Education. McMaster University, Hamilton, Ontario, 2005
- [56] R. Alves. “Aplicação de Modelos de Redes Neurais para Previsão de Consumos de Energia”. (Dissertação para obtenção do grau de Mestre). Instituto Superior Técnico, Lisboa, Portugal, 2016
- [57] S. Jung e O. J. Kwon. “Improving Neural Networks by Adopting Amplifying and Attenuating Neurons”, n. May, pp. 0–10, Mai, 2019.
- [58] B. Kröse, P. van der Smagt, e P. Smagt. “An introduction to neural networks”. *J Comput Sci*. 48. 1993.

- [59] D. E. Rumelhart e G. E. Hintont. “Learning Representations by Back-Propagating Errors”, em *Cognitive Modeling*, n. 2, The MIT Press, pp. 3–6, 2002.
- [60] A. Le My Phung e K. James. “Comparison of Support Vector Regression and Neural Networks”, (Trabalho para obtenção parcial do grau de Mestre em Ciência). Universidade do Minnesota Duluth, Duluth, Estados Unidos da América, 2016.
- [61] K. Levenberg. “A Method for the Solution of Certain Non-Linear Problems in Least Squares”. *Quarterly of Applied Mathematics*, 2, 164-168, Jul. 1944.
- [62] D. W. Marquardt. “An algorithm for least-squares estimation of nonlinear parameters”. *Journal of the Society for Industrial and Applied Mathematics*, 11(2), 431—441, Jun 1963.
- [63] M. I. Lourakis. “A Brief Description of the Levenberg-Marquardt Algorithm Implemented by levmar”, *Matrix*, vol. 3, p. 2, Jan. 2005.
- [64] H. Yu e B. M. Wilamowski. “Levenberg Marquardt Training”. *Industrial Electronics Handbook*, vol. 5 Intelligent Systems, 2nd Edition, chapter 12, pp. 12-1 to 12-15, CRC Press, 2011.
- [65] A. Sharif Ahmadian. “Numerical Methods and Procedures”, em *Numerical Models for Submerged Breakwaters*, Elsevier, pp. 93–108, 2016.
- [66] G. Nazaré. “Wind power forecast using Neural Networks tuned with advanced optimization techniques” (Dissertação para obtenção do grau de Mestre). Instituto Superior Técnico, Lisboa, Portugal, 2016.
- [67] MATLAB and Neural Network Toolbox Release 2012b, The MathWorks, Inc., Natick, Massachusetts, Estados Unidos da América.
- [68] S. Karsoliya. “Approximating Number of Hidden layer neurons in Multiple Hidden Layer BPNN Architecture”. *International Journal of Engineering Trends and Technology (IJETT)*. V3(6):714-717, Nov-Dez. 2012.
- [69] M. Bielecki. “Statistical Characterization of Errors in Wind Power Prediction”. (Dissertação para obtenção do grau de Mestre de Ciências em Engenharia). Faculdade do Norte do Arizona, Arizona, Estados Unidos da América, 2010.
- [70] H. Holttinen, J. Miettinen e S. Sillanpää. “Wind power forecasting accuracy and uncertainty in Finland”. VTT Technical Research Centre of Finland. 76 p. (VTT Technology; No. 95), 2013.
- [71] M. Lacerda, A. Couto, e A. Estanqueiro. “Wind Power Ramps Driven by Windstorms and Cyclones,” *Energies*, vol. 10, no. 10, p. 1475, Sep. 2017.
- [72] R. G. Ahangar, M. Yahyazadehfar, e H. Pournaghshband. “The Comparison of Methods Artificial Neural Network with Linear Regression Using Specific Variables for Prediction Stock Price in Tehran Stock Exchange”, *Int. J. Comput. Sci. Inf. Secur.*, vol. 7, n. 2, pp. 38–46, Mar. 2010.

I. Anexo – Análises de Sensibilidade referentes às metodologias implementadas

- Metodologia kNN

As figuras e tabelas que figuram nesta subsecção servem o propósito de suportar a análise de sensibilidade efetuada para a metodologia kNN.

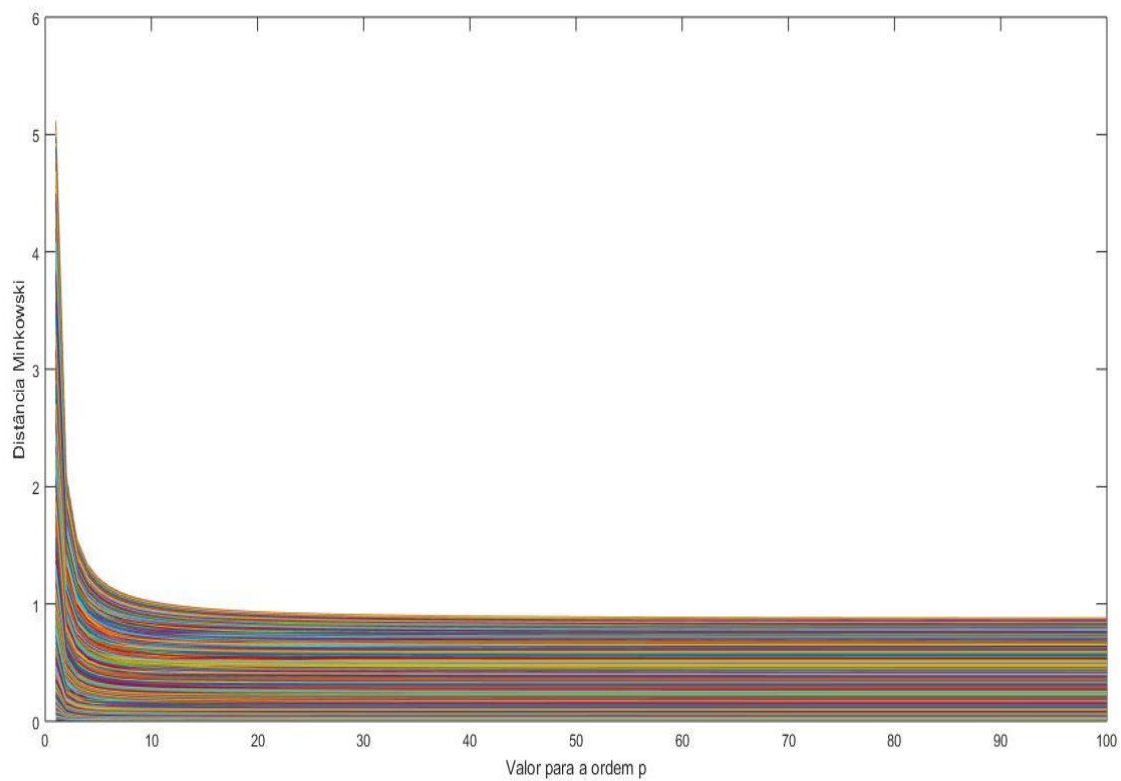


Figura I-1: Análise de sensibilidade referente ao valor da ordem p da distância de Minkowski. Cada linha a cores representa a variação da distância de um ponto, pertencente à série de dados histórica, à origem.

Tabela I-1: Valor de erro (p.u.), segundo condição imposta, para a ordem p da distância de Minkowski.

Erro ($p=14$)
0.5023

- **Metodologia RNA**

As figuras e tabelas que figuram nesta subsecção servem o propósito de suportar a análise de sensibilidade efetuada para a metodologia RNA.

Tabela I-2: RMSE (p.u.) obtido para a análise de sensibilidade das RNA, referente ao estudo acerca do número ótimo de neurónios presentes na camada escondida. Na primeira linha da tabela encontra-se o tamanho, i.e. número de neurónios, na camada de entrada. DE representa o número de neurónios da camada de entrada da rede, seis. Os resultados encontram-se mapeados por cor: quanto mais vermelho maior será o RMSE e pior será o desempenho da métrica de avaliação – o RMSE. Pelo contrário, quanto mais verde, menor o valor encontrado para o RMSE e melhor será o desempenho desta métrica.

Regra	2/3 DE	0.9DE	DE	2DE-1	2DE
Nº de neurónios	4	5	6	11	12
RMSE [p.u.]	0.0766	0.0765	0.0764	0.0764	0.0764
	0.0764	0.0763	0.0764	0.0764	0.0763
	0.0764	0.0762	0.0763	0.0763	0.0764
	0.0763	0.0762	0.0764	0.0763	0.0764
	0.0763	0.0764	0.0763	0.0765	0.0763

Tabela I-3: Resultados como na Tabela I-2 mas para os valores do coeficiente de correlação.

Regra	2/3 DE	0.9DE	DE	2DE-1	2DE
Nº de neurónios	4	5	6	11	12
Coeficiente de correlação	0.9208	0.9209	0.9211	0.9212	0.9211
	0.9210	0.9213	0.9212	0.9212	0.9213
	0.9211	0.9215	0.9213	0.9213	0.9211
	0.9215	0.9215	0.9212	0.9213	0.9211
	0.9214	0.9211	0.9213	0.9209	0.9213

II. Análise detalhada dos cenários estudados

- **Métricas de erro**

Um ponto central e evidente resultante da análise de resultados, no que diz respeito às várias configurações da metodologia kNN implementadas na dissertação, foi a inadequação das previsões obtidas pela configuração da metodologia kNN com a distância de Correlação. A Tabela II-1 representa a variação percentual entre o RMSE obtido pelas previsões com as métricas de distância física (euclidiana e Minkowski) vs o RMSE obtido pelas previsões com as métricas de distância por correlação.

Tabela II-1: Variações percentuais do RMSE médio, obtido com a implementação da metodologia de previsão kNN, entre os métodos de cálculo de distâncias entre as instâncias.

Euclidiana vs Correlação	Minkowski vs Correlação
224.40%	246.04%
224.40%	246.04%
311.20%	279.87%
314.76%	281.32%
230.05%	227.86%
248.43%	243.86%
128.06%	136.01%
144.43%	151.09%
91.92%	91.67%
98.40%	98.52%

Na Tabela II-2 é possível observar a variação percentual da métrica de erro Correlação referente à metodologia RNA, face às duas outras metodologias em estudo na presente dissertação: a persistência, como *benchmark*, e as várias configurações estudadas da metodologia kNN.

Tabela II-2: Variações percentuais da correlação média, obtida com a implementação da metodologia de previsão RNA face aos resultados da metodologia de previsão da persistência e kNN, para a referida métrica.

Persistência		3.08%			
	Distância		Euclidiana	Minkowski	Correlação
	kNN	k=1	Não-Ponderada	143.09%	141.93%
Ponderada			143.09%	141.93%	162.75%
k=5		Não-Ponderada	117.97%	118.03%	141.99%
		Ponderada	118.84%	118.52%	145.90%
k=10		Não-Ponderada	112.28%	110.71%	130.58%
		Ponderada	112.60%	111.24%	135.76%
k=20		Não-Ponderada	107.76%	107.84%	124.35%
		Ponderada	108.17%	107.55%	128.25%
k=162		Não-Ponderada	103.93%	104.09%	115.05%
		Ponderada	103.84%	103.92%	115.62%

Comportamento do RMSE médio horário para a metodologia k NN, com foco nas variações do algoritmo que utilizam a métrica de distância Euclidiana e de Minkowski.

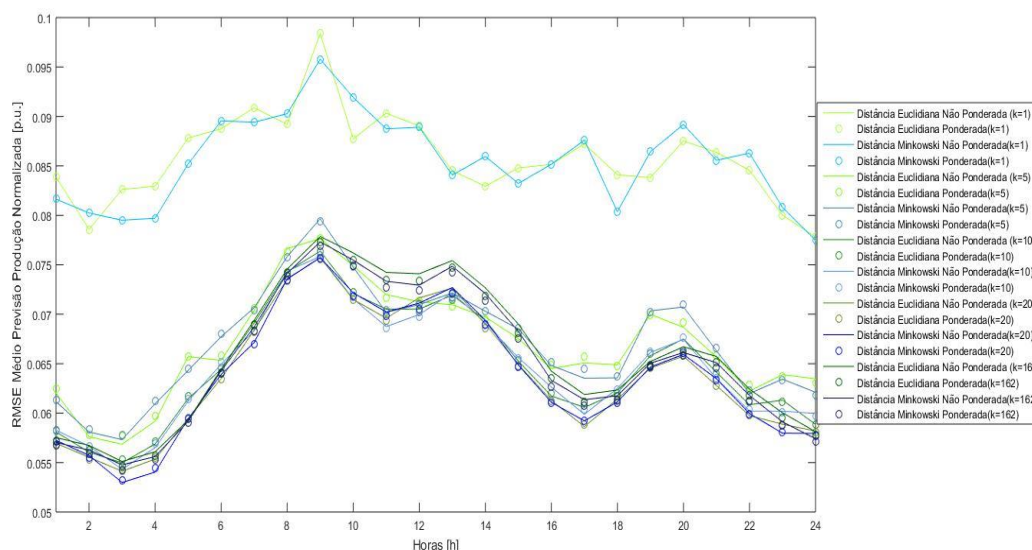
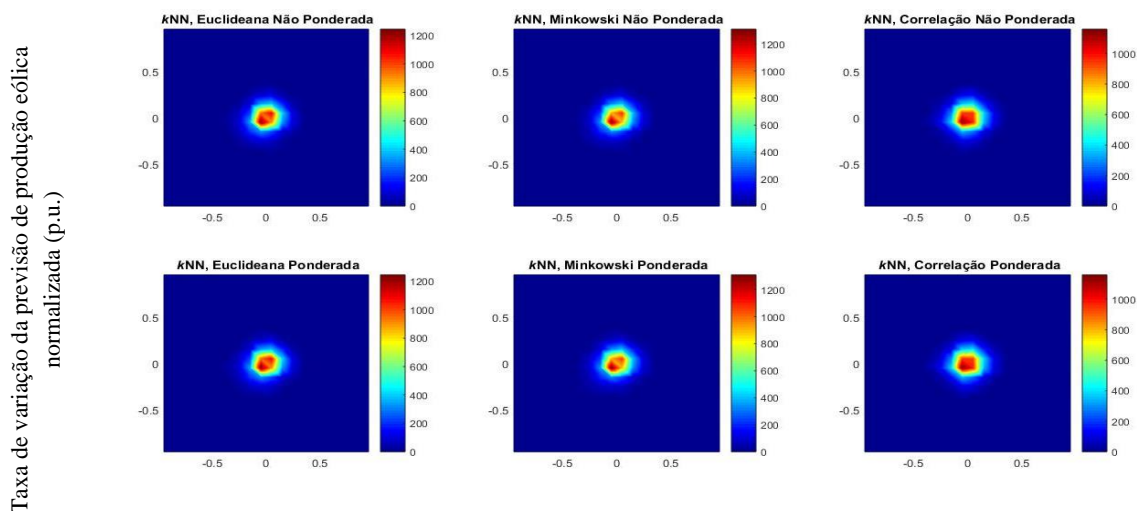


Figura II-1: RMSE médio horário da previsão de produção eólica normalizada, para a metodologia k NN utilizando a métrica de distância euclidiana e de Minkowski, para todos os valores de k .

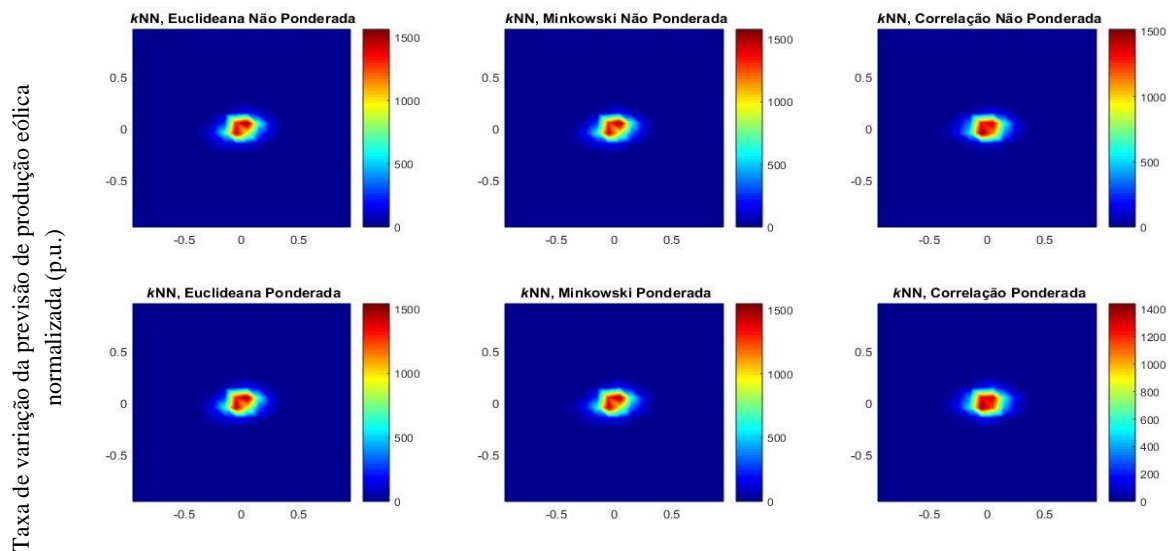
- **Desempenho das metodologias relativo às variações de produção**

Nas Figura II-2 –Figura II-6 encontram-se mapeadas, para a metodologia k NN, as ocorrências de cada classe de taxa de variação da produção eólica normalizada (TVO) e da taxa de variação da previsão de produção eólica normalizada (TVP).



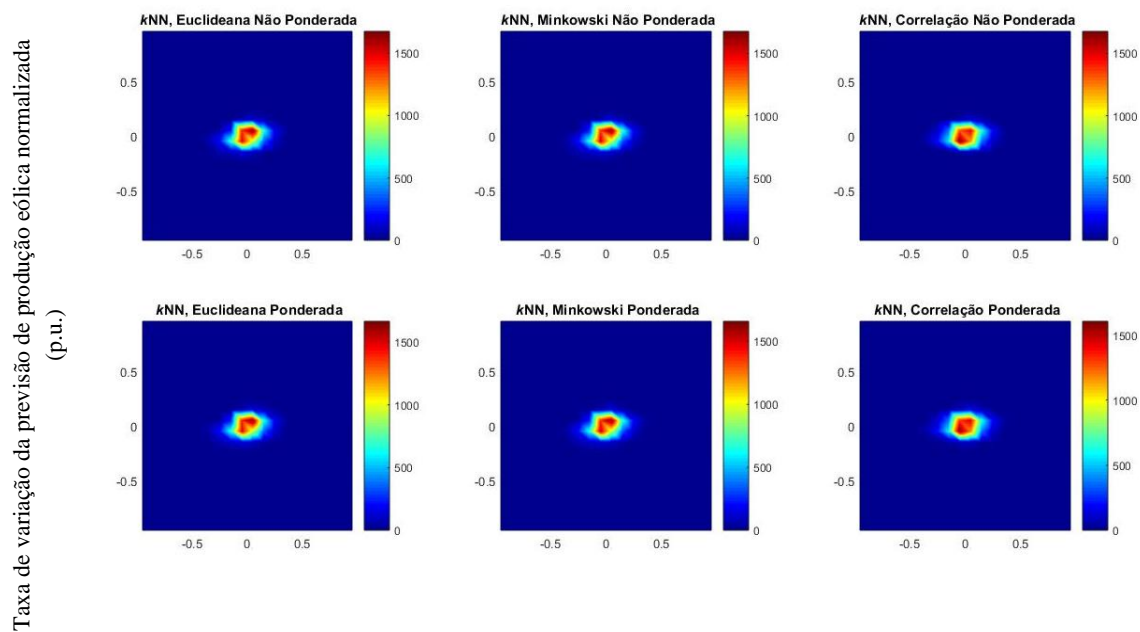
Taxa de variação da produção eólica normalizada observada normalizada (p.u.).

Figura II-2: Mapeamento das ocorrências obtidas para cada classe, referente à metodologia implementada k NN, onde k assume o valor de 1. A barra de cor indica a escala de cor consoante o número de ocorrências verificado. No eixo dos x estão representadas as taxas de variação da produção eólica normalizada observada, enquanto nos eixos y encontram-se representadas as taxas de variação da previsão de produção eólica normalizada.



Taxa de variação da produção eólica normalizada observada normalizada (p.u.).

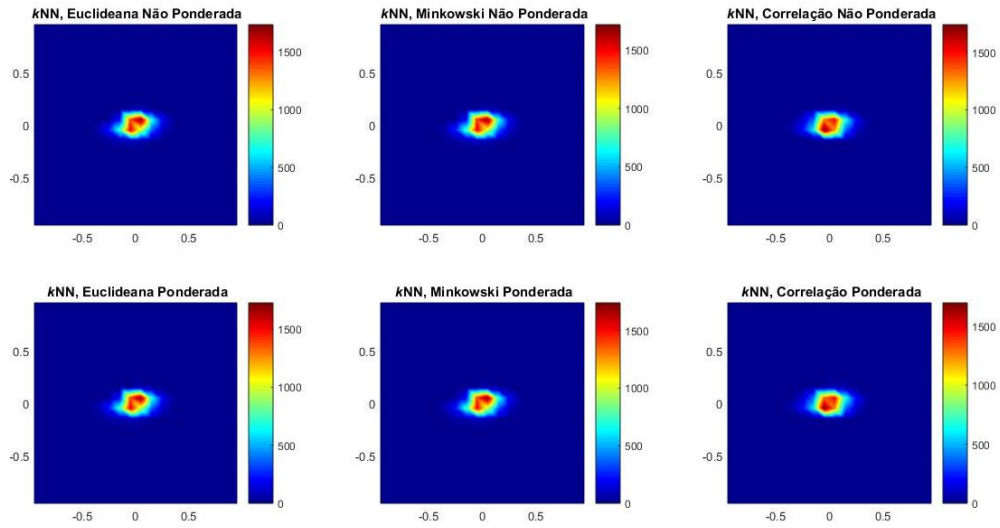
Figura II-3: Similar à Figura II-2 mas onde k assume o valor de 5



Taxa de variação da produção eólica normalizada observada normalizada (p.u.).

Figura II-4: Similar à Figura II-2 mas onde k assume o valor de 10

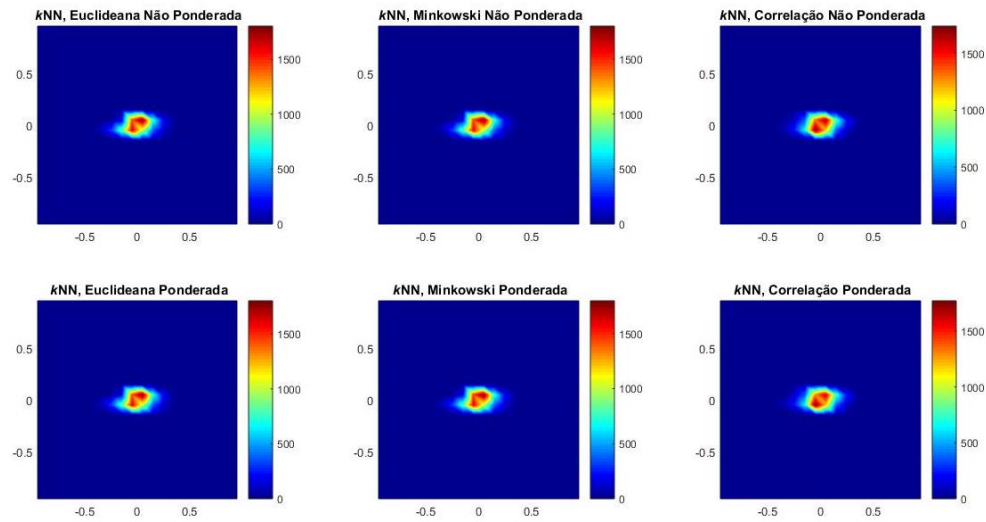
Taxa de variação da previsão de produção eólica normalizada
(p.u.)



Taxa de variação da produção eólica normalizada observada normalizada (p.u.).

Figura II-5: Similar à Figura II-2 mas onde k assume o valor de 20

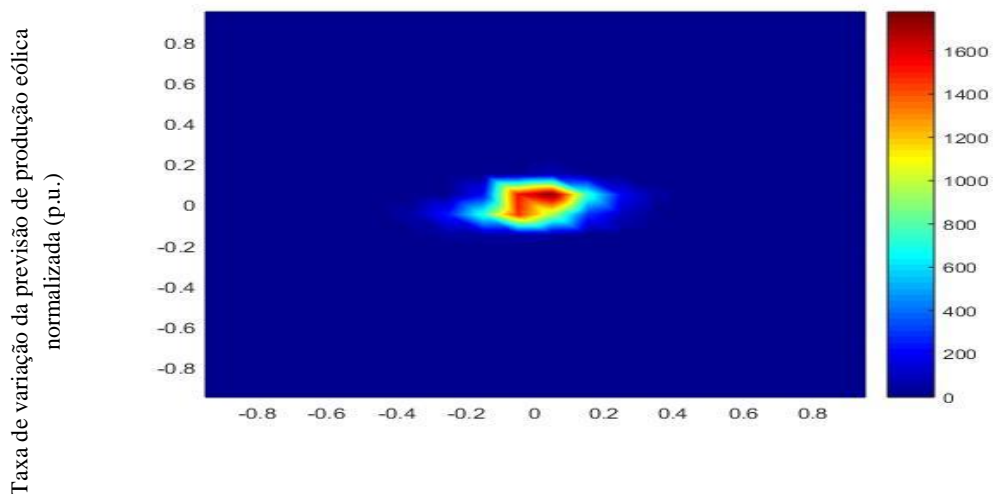
Taxa de variação da previsão de produção eólica normalizada
(p.u.)



Taxa de variação da produção eólica normalizada observada normalizada (p.u.).

Figura II-6: Similar à Figura II-2 mas onde k assume o valor de 162

Nas Figura II-7 encontra-se mapeadas, para a metodologia RNA, as ocorrências de cada classe as ocorrências de cada classe de taxa de variação da produção eólica normalizada (TVO) e da taxa de variação da previsão de produção eólica normalizada (TVP).



Taxa de variação da produção eólica normalizada observada normalizada (p.u.).

Figura II-7: Mapeamento das ocorrências obtidas para cada classe, referente à metodologia implementada RNA. A barra de cor indica a escala de cor consoante o número de ocorrências verificado. No eixo dos x estão representadas as taxas de variação da produção eólica normalizada observada, enquanto nos eixos y encontram-se representadas as taxas de variação da previsão de produção eólica normalizada.

- **Extremos de produção: RMSE médio**

Na Figura II-8 – Figura II-9 observa-se a representação gráfica do RMSE médio para os casos de extremos de produção: máxima e mínima produção eólica normalizada, respetivamente.

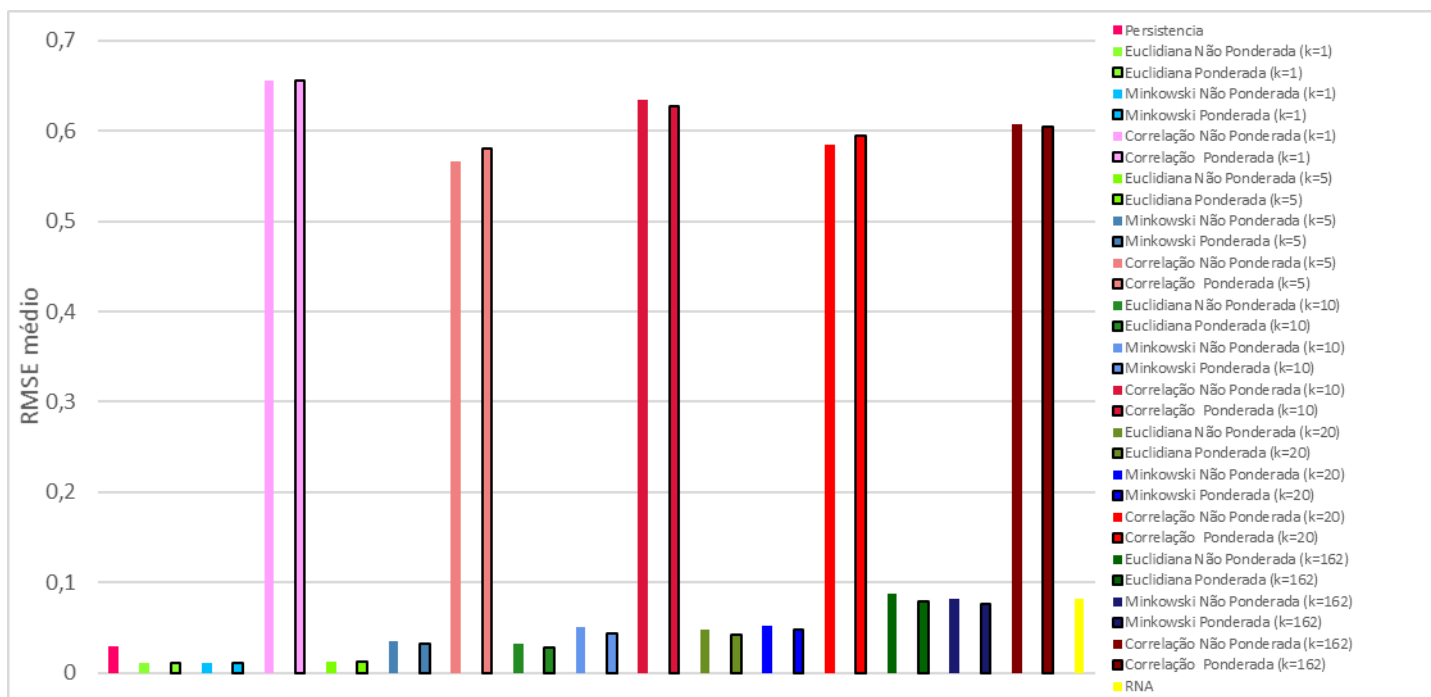


Figura II-8: Representação gráfica do RMSE médio para o caso máximo de produção eólica média normalizada, para o ano de 2018.

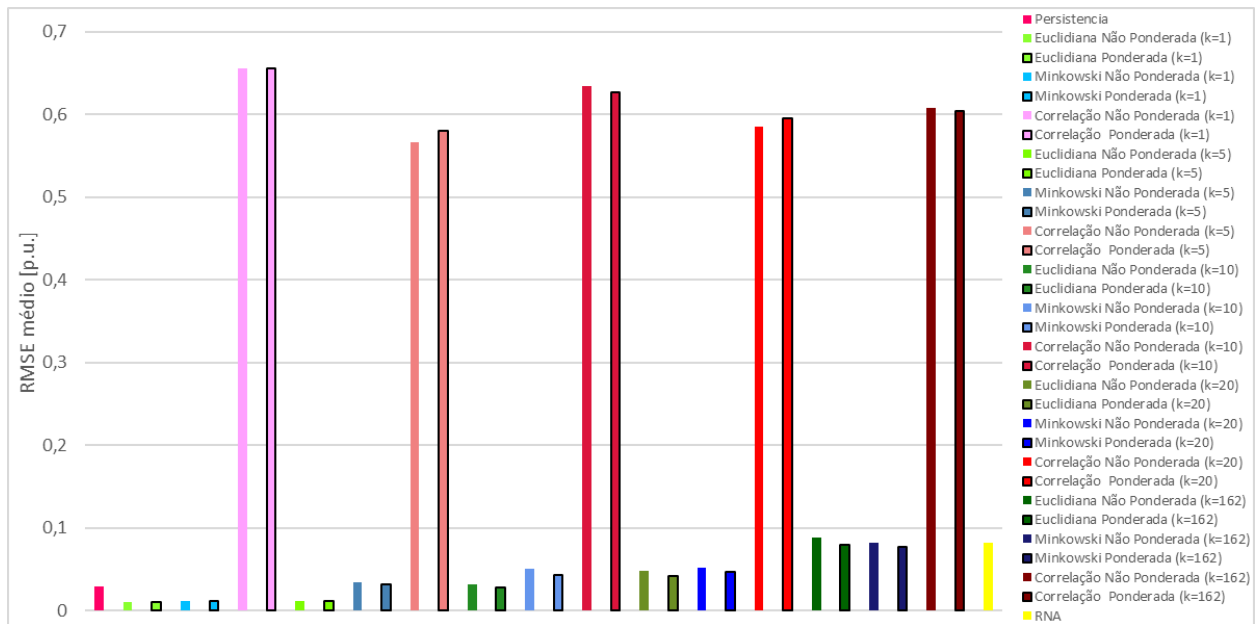


Figura II-9: Representação gráfica do RMSE máximo para o caso mínimo de produção eólica média normalizada, para o ano de 2018.

- Métricas de erro conforme sazonalidade

Na Figura II-10 à Figura II-13 encontram-se representados os perfis diários do RMSE, para todas as metodologias de previsão de produção eólica implementadas na presente dissertação, conforme a sazonalidade: inverno, primavera, verão e outono, respetivamente.

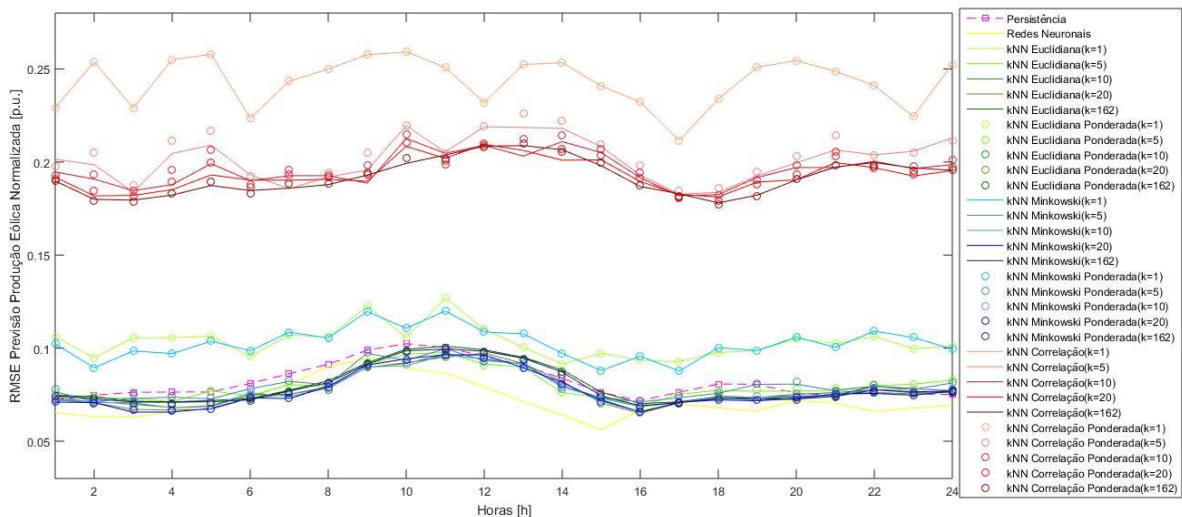


Figura II-10: Perfil diário do RMSE referente à estação do Inverno, para todas as metodologias de previsão de produção eólica implementadas.

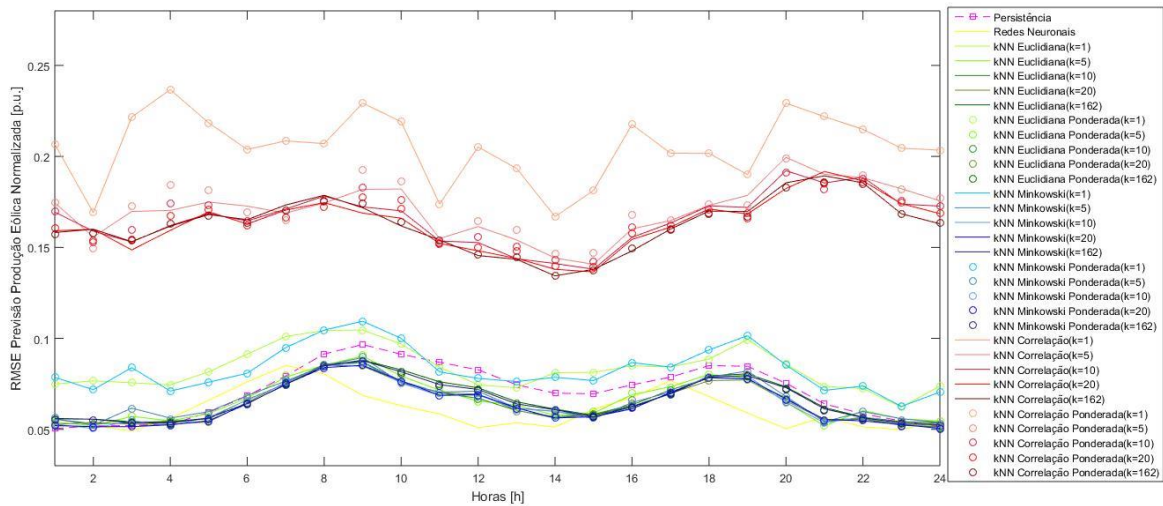


Figura II-11: Perfil diário do RMSE referente à estação do Outono, para todas as metodologias de previsão de produção eólica implementadas.

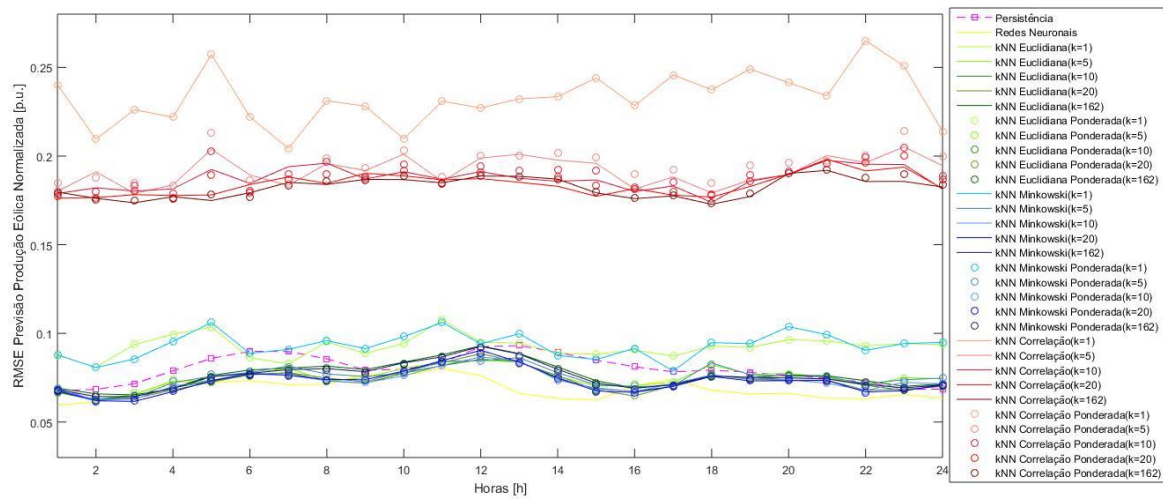


Figura II-12: Perfil diário do RMSE referente à estação da Primavera, para todas as metodologias de previsão de produção eólica implementadas.

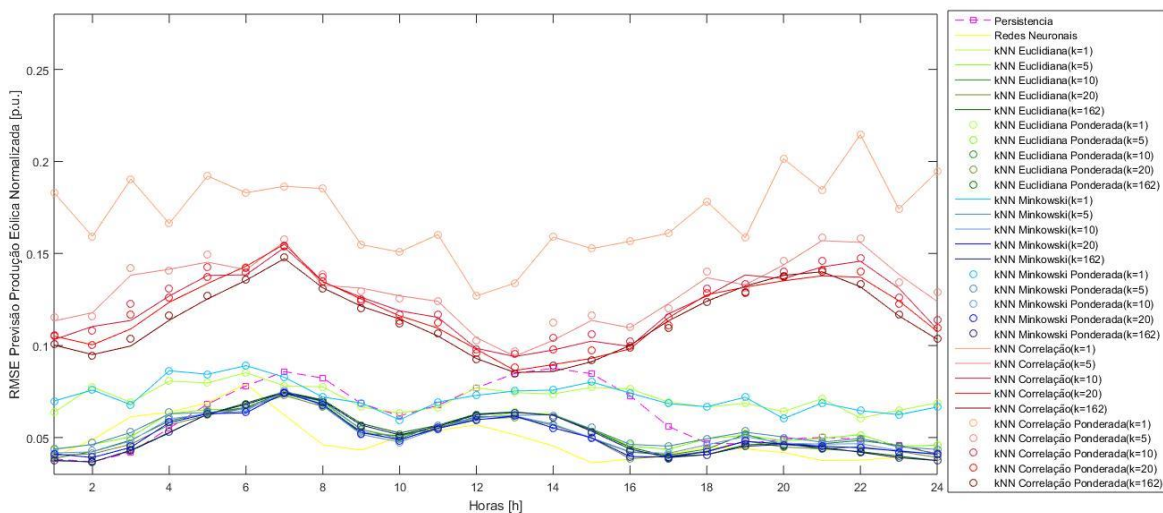


Figura II-13: Perfil diário do RMSE referente à estação do Verão, para todas as metodologias de previsão de produção eólica implementadas.