

Faculdade de Engenharia da Universidade do Porto



Aplicação de Heurísticas para Otimização e Parâmetros de Modelos de Previsão Eólica

Alexandre Melo Amaral

VERSÃO PARA AVALIAÇÃO DO JÚRI

Dissertação realizada no âmbito do
Mestrado Integrado em Engenharia Eletrotécnica e de Computadores
Major Energia

Orientador: Professor Doutor Cláudio Monteiro
Coorientadora: Professora Doutora Maria Margarida de Amorim Ferreira

Junho de 2015

© Alexandre Melo Amaral, 2015

Resumo

O recurso eólico define-se pela sua natureza incerta e intermitente. Desta forma, o desenvolvimento de ferramentas para a previsão de produção eólica pode ser importante para resolver a maioria dos problemas de gestão e planeamento dos sistemas elétricos resultantes da integração deste tipo energia.

É neste contexto que se insere o tema desta dissertação, cujo objetivo consiste em definir processos de otimização da parametrização de modelos de previsão probabilística de produção eólica baseado em estimadores de densidade condicional não-paramétrica denominados *Nadaraya-Watson Estimator* (NW) e *Kernel Density Estimation* (KDE), que para além de apresentarem os resultados dos valores esperados da previsão eólica, estimam a incerteza associada.

Nesta dissertação será realizado um estudo aprofundado sobre a ferramenta de previsão probabilística e serão criados processos sistemáticos de parametrização com o objetivo de melhorar a *performance* do modelo de previsão escolhido.

Será ainda, criada e programada uma ferramenta NW-KDE, incluindo novas funcionalidades: variáveis cíclicas, previsão recursiva e base de conhecimento com pesos. A modelização matemática do NW-KDE será criada de forma a ser possível correr, em tempo suficientemente curto, problemas de previsão com horizontes longos e com elevado número de variáveis de entrada.

Além de um desenvolvimento otimizado e entendido da técnica NW-KDE, o trabalho da dissertação centra-se em estudar formas de otimização da parametrização para este modelo matemático, nomeadamente, parâmetros da largura de banda, parâmetros de incerteza, parametrização dos pesos e parametrização da previsão recursiva.

Por fim, as metodologias de parametrização serão aplicadas a quatro parques eólicos com diferentes características de capacidade e variabilidade, avaliando-se assim a existência de parametrizações típicas para a aplicação na previsão de produção eólica.

Palavras-chave: Modelos de Previsão de Produção Eólica, *Nadaraya-Watson Estimator*, *Kernel Density Estimator*, Incerteza na Previsão Eólica

Abstract

The wind resource is defined by its uncertain and intermittent nature. In this sense, the development of tools for wind power forecasting may be important to solve the majority of the problems of the integration of wind power into the power system.

In this context, the aim of this dissertation is to define optimizations process of wind forecasting models based on non-parametric conditional estimators named Nadaraya-Watson Estimator (NW) e Kernel Density Estimation (KDE), which presents the expected results and estimates the forecast uncertainty.

Initially, a detailed study of the probabilistic forecasting tools and systematic parameterization process will be created in order to improve the performance of the chosen forecast model.

Additionally, will be also created and programmed a NW-KDE tool, including new features: cyclical variables, recursive prediction and knowledge base with weights. The NW-KDE mathematical modeling will be created in order to be able to run, in shorter time, forecasting problems with long-term horizons and high number of input variables.

In addition to an optimized development of the NW-KDE technique, the work of the dissertation focuses on studying ways of optimizing the parameterization for this mathematical model, namely, bandwidth parameters, uncertainty parameters, weights parameters and recursive prediction parameters.

Lastly, the parameterization methodologies will be applied to four wind farms with different performance and variability characteristics, evaluating the existence of typical parameterization for application in wind forecasting.

Keywords: *Wind Power Forecasting models, Nadaraya-Watson Estimator, Kernel Density Estimation, Forecasting Uncertainty*

Agradecimentos

Em primeiro lugar, gostaria de agradecer ao meu orientador, o professor Dr. Cláudio Monteiro, pela sua disponibilidade, conselhos e por todo o apoio na realização da dissertação.

Gostaria também de agradecer à minha coorientadora, a professora Dr. Maria Ferreira, por todo o apoio prestado.

À empresa *Smartwatt*, pela disponibilização dos dados de produção eólica essenciais para o desenvolvimento desta dissertação.

Aos meus amigos, pela amizade e companheirismo.

Finalmente, à minha família, especialmente aos meus pais, à minha irmã e ao Vítor, por todo o apoio e dedicação.

Índice

Resumo.....	iii
Abstract.....	v
Agradecimentos	vii
Índice.....	ix
Lista de figuras	xi
Lista de tabelas.....	xv
Abreviaturas e Símbolos	xvii
Capítulo 1	1
Introdução	1
1.1 Enquadramento	1
1.2 Objetivos	2
1.3 Estrutura do documento	2
Capítulo 2.....	3
Estado da arte	3
2.1 Modelos Numéricos de Previsão de Meteorologia (NWP)	3
2.2 Metodologias de Previsão Eólica.....	6
2.3 Avaliação dos Modelos de Previsão.....	10
2.4 Incerteza na Previsão Eólica	14
Capítulo 3.....	19
Modelo de Previsão Probabilística	19
3.1 Definição do KDE	19
3.2 Limite mínimo de ativação da função <i>Kernel</i>	24
3.3 Parametriação da entrada recorrente.....	24
3.4 Parametriação dinâmica da largura de banda da função <i>kernel</i>	25
3.5 Parametriação das variáveis cíclicas	30
3.6 Parametriação do peso de cada caso histórico.....	31
3.7 Parametriação para ajustar a incerteza da previsão eólica	32
Capítulo 4.....	33
Modelo de Previsão e Metodologia.....	33
4.1 Modelo de previsão.....	33
4.2 Metodologia dos resultados	37

Capítulo 5	41
Resultados Obtidos	41
5.1 Teste 1: Modelo Inicial	41
5.2 Teste 2: Entrada Recorrente	44
5.3 Teste 3: Parametrização dinâmica da largura de banda	47
5.4 Teste 4: Parametrização das variáveis cíclicas	49
5.5 Teste 5: Ajuste de parâmetros	51
5.6 Teste 6: Parametrização do peso de cada caso histórico	54
5.7 Teste 7: Parametrização para ajuste da incerteza	57
5.8 Teste 8: Modelo Final	59
5.9 Análise dos resultados obtidos	62
 Capítulo 6	 65
Conclusão e Trabalhos Futuros	65
6.1 Conclusão	65
6.2 Trabalhos futuros	66
 Referências	 69
 Anexo A - Histórico dos Parques	 71

Lista de figuras

Figura 2.1 - Modelos de Área Limitada Europeus	5
Figura 2.2 - Etapas dos modelos de previsão físicos	7
Figura 2.3 - Representação do modelo de previsão estatístico	8
Figura 2.4 - Função de distribuição cumulativa (caso ideal)	16
Figura 2.5 - Exemplo de um diagrama de reliability de um parque eólico	17
Figura 2.6 - Distribuição preditiva com diferentes níveis de <i>sharpness</i>	17
Figura 2.7 - Exemplo de um diagrama de <i>sharpness</i> de um parque eólico	18
Figura 3.1 - Histogramas de previsão com distribuições Beta	24
Figura 3.2 - Representação da velocidade do vento em relação à potência real registada	26
Figura 3.3 - Pontos para o cálculo do fator multiplicativo	27
Figura 3.4 - Fluxograma do processo	28
Figura 3.5 - Objetivo do cálculo da função quadrática	29
Figura 3.6 - Representação da densidade marginal da variável hora sem a parametrização das variáveis cíclicas	30
Figura 3.7 - Representação da densidade marginal da variável hora com a parametrização das variáveis cíclicas	31
Figura 4.1 - Representação do histórico de previsão.....	33
Figura 4.2 - Representação do <i>output</i> da previsão	34
Figura 4.3 - Amostra do <i>Target</i> do Parque 1	34
Figura 4.4 - Amostra da velocidade do vento do Parque 1	35
Figura 4.5 - Amostra da direção do vento do Parque 1	35
Figura 4.6 - Amostra da temperatura do Parque 1	35
Figura 4.7 - Representação das variáveis de entrada do teste relativos a uma semana	36

Figura 4.8 - Amostra dos valores obtidos da previsão eólica através redes e os valores registados reais do Parque 1	36
Figura 4.9 - Exemplo da representação da previsão da produção eólica e a incerteza associada.....	38
Figura 4.10 - Exemplo da representação de um diagrama de <i>reliability</i>	39
Figura 4.11 - Exemplo da representação de um diagrama de <i>sharpness</i>	39
Figura 5.1 - Previsão da produção eólica e a incerteza associada do Parque 1	42
Figura 5.2 - Previsão da produção eólica e a incerteza associada do Parque 2	42
Figura 5.3 - Previsão da produção eólica e a incerteza associada do Parque 3	42
Figura 5.4 - Previsão da produção eólica e a incerteza associada do Parque 4	43
Figura 5.5 - Diagrama de <i>reliability</i> do teste 1 para os quatro parques	43
Figura 5.6 - Diagrama de <i>sharpness</i> do teste 1 para os quatro parques.....	44
Figura 5.7 - Previsão da produção eólica e a incerteza associada do Parque 1	45
Figura 5.8 - Previsão da produção eólica e a incerteza associada do Parque 2	45
Figura 5.9 - Previsão da produção eólica e a incerteza associada do Parque 3	46
Figura 5.10 - Previsão da produção eólica e a incerteza associada do Parque 4	46
Figura 5.11 - Previsão da produção eólica e a incerteza associada do Parque 1	47
Figura 5.12 - Previsão da produção eólica e a incerteza associada do Parque 2	48
Figura 5.13 - Previsão da produção eólica e a incerteza associada do Parque 3	48
Figura 5.14 - Previsão da produção eólica e a incerteza associada do Parque 4	48
Figura 5.15 - Previsão da produção eólica e a incerteza associada do Parque 1	50
Figura 5.16 - Previsão da produção eólica e a incerteza associada do Parque 2	50
Figura 5.17 - Previsão da produção eólica e a incerteza associada do Parque 3	50
Figura 5.18 - Previsão da produção eólica e a incerteza associada do Parque 4	51
Figura 5.19 - Previsão da produção eólica e a incerteza associada do Parque 1	53
Figura 5.20 - Previsão da produção eólica e a incerteza associada do Parque 2	53
Figura 5.21 - Previsão da produção eólica e a incerteza associada do Parque 3	53
Figura 5.22 - Previsão da produção eólica e a incerteza associada do Parque 4	54
Figura 5.23 - Previsão da produção eólica e a incerteza associada do Parque 1	55
Figura 5.24 - Previsão da produção eólica e a incerteza associada do Parque 2	55
Figura 5.25 - Previsão da produção eólica e a incerteza associada do Parque 3	56

Figura 5.26 - Previsão da produção eólica e a incerteza associada do Parque 4	56
Figura 5.27 - Diagrama de <i>reliability</i> do teste 7 para os quatro parques	58
Figura 5.28 - Diagrama de <i>sharpness</i> do teste 5 para os quatro parques	58
Figura 5.29 - Diagrama de <i>sharpness</i> do teste 7 para os quatro parques	59
Figura 5.30 - Previsão da produção eólica e a incerteza associada do Parque 1	60
Figura 5.31 - Previsão da produção eólica e a incerteza associada do Parque 2	60
Figura 5.32 - Previsão da produção eólica e a incerteza associada do Parque 3	60
Figura 5.33 - Previsão da produção eólica e a incerteza associada do Parque 4	61
Figura 5.34 - Diagrama de <i>reliability</i> do teste 8 para os quatro parques	61
Figura 5.35 - Diagrama de <i>sharpness</i> do Modelo Final para os quatro parques	62
Figura A.1 - Amostra do Target do Parque 2	71
Figura A.2 - Amostra da velocidade do vento do Parque 2	71
Figura A.3 - Amostra da direção do vento do Parque 2	72
Figura A.4 - Amostra da temperatura do Parque 2	72
Figura A.5 - Amostra do <i>Target</i> do Parque 3	72
Figura A.6 - Amostra da velocidade do vento do Parque 3	73
Figura A.7 - Amostra da direção do vento do Parque 3	73
Figura A.8 - Amostra da temperatura do Parque 3	73
Figura A.9 - Amostra do <i>Target</i> do Parque 4	74
Figura A.10 - Amostra da velocidade do vento do Parque 4	74
Figura A.11 - Amostra da direção do vento do Parque 4	74
Figura A.12 - Amostra da temperatura do Parque 4	75
Figura A.13 - Amostra dos valores obtidos da previsão eólica através redes e os valores registados reais do Parque 2	75
Figura A.14 - Amostra dos valores obtidos da previsão eólica através redes e os valores registados reais do Parque 3	75
Figura A.15 - Amostra dos valores obtidos da previsão eólica através redes e os valores registados reais do Parque 4	76

Lista de tabelas

Tabela 2.1 - Modelos Globais	4
Tabela 2.2 - Sistemas NWP operacionais na Europa	5
Tabela 2.3 - Características dos horizontes temporais	6
Tabela 2.4 - Modelos físicos operacionais	7
Tabela 2.5 - Modelos Estatísticos	8
Tabela 2.6 - Características das abordagens na combinação de modelos	10
Tabela 2.7 - Metodologias baseadas na abordagem paramétrica	14
Tabela 2.8 - Metodologias baseadas na abordagem não-paramétrica	15
Tabela 2.9 - <i>Skill score</i>	18
Tabela 4.1 - Características de cada parque	34
Tabela 5.1 - Parametrizações do teste 1	41
Tabela 5.2 - Erros obtidos dos valores previstos para o Teste 1	43
Tabela 5.3 - Desvio médio quadrático da <i>reliability</i> (Modelo Inicial)	44
Tabela 5.4 - Parametrizações do teste 2	44
Tabela 5.0.5 - Erros obtidos dos valores previstos para o Teste 2	46
Tabela 5.6 - Parametrizações do teste 3	47
Tabela 5.7 - Erros obtidos dos valores previstos para o Teste 3	49
Tabela 5.8 - Parametrizações do teste 4	49
Tabela 5.9 - Erros obtidos dos valores previstos para o Teste 4	51
Tabela 5.10 - Teste ao número mínimo de pontos representativos	51
Tabela 5.11 - Teste ao número de partições para o ajuste da largura de banda da entrada recorrente	52

Tabela 5.12 - Parametrizações do teste 5	52
Tabela 5.13 - Erros obtidos dos valores previstos para o Teste 5	54
Tabela 5.14 - Parametrizações do teste 6	54
Tabela 5.15 - Erros obtidos dos valores previstos para o Teste 6	56
Tabela 5.16 - Resultado dos valores ótimos dos parâmetros da incerteza para o teste 7	57
Tabela 5.17 - Parametrizações do teste 7	57
Tabela 5.18 - Desvio médio quadrático da <i>reliability</i> (teste 7)	58
Tabela 5.19 - Parametrizações do teste 8	59
Tabela 5.20 - Erros obtidos dos valores previstos do Modelo Final	61
Tabela 5.21 - Desvio médio quadrático da <i>reliability</i> (Modelo Final)	62
Tabela 5.22 - Resultados do erro da previsão determinística de todos os testes	62
Tabela 5.23 - Resultados do valor dos parâmetros definidos do Modelo Final	63

Abreviaturas e Símbolos

Lista de abreviaturas (ordenadas por ordem alfabética)

CDF	<i>Cumulative Distribution Function</i>
GPU	Graphics Processing Unit
KDE	<i>Kernel Density Estimation</i>
MAE	<i>Mean Absolute Error</i>
MAPE	<i>Mean Absolute Percentage Error</i>
MSE	Mean Square Error
NWE	<i>Nadaraya-Watson Estimator</i>
NWP	<i>Numerical Weather Prediction</i>
PDF	<i>Probability Density Function</i>
RMSE	<i>Root Mean Square Error</i>
SDE	<i>Standard Deviation of Errors</i>

Capítulo 1

Introdução

Esta dissertação foi realizada no âmbito do Mestrado Integrado em Engenharia Eletrotécnica e de Computadores da Faculdade de Engenharia da Universidade do Porto.

1.1 Enquadramento

O aumento do preço do petróleo bruto e a sensibilização mundial sobre o impacto ambiental das fontes de energia convencionais, aumentaram a necessidade da exploração das energias renováveis. A energia eólica registou um dos maiores crescimentos e é considerada uma alternativa atraente devido à sua alta eficiência e baixa poluição[1, 2].

Apesar das inúmeras vantagens com a integração da energia eólica no *mix* energético, a natureza incerta e intermitente deste tipo de energia representa a sua grande desvantagem. Em sistemas elétricos com baixa penetração de eólica, este efeito pode ser mitigado com o uso de centrais de resposta rápida, no entanto a alta penetração de energia eólica no sistema aumenta consideravelmente a incerteza associada à produção de energia e oferece uma série de desafios nas operações e no planeamento de sistemas elétricos[3].

Uma forma de reduzir a incerteza associada à previsão da produção eólica é o uso de ferramentas de previsão. Este tipo de ferramentas são uma boa solução para resolver a maior parte dos problemas deste tipo de energia, uma vez que permite obter uma estimativa da produção esperada num determinado futuro. Uma previsão de produção eólica competente poderá também fornecer suporte técnico para comercialização de energia eólica no mercado energético[4].

A maior parte das ferramentas de previsão segue uma abordagem determinística da produção eólica, isto é, produz um único valor esperado para o horizonte da previsão. A grande desvantagem deste tipo de previsão, é que nenhuma informação é fornecida sobre qualquer desvio dos valores previstos, o que limita a tomada de decisões dos operadores das redes. A previsão probabilística proporciona informação adicional sobre a incerteza da produção eólica. Com o aumento de produções estocásticas na rede, a previsão desta é importante para a gestão e controlo dos sistemas elétricos[5].

1.2 Objetivos

O principal objetivo desta dissertação consiste em definir processos de otimização da parametrização de modelos de previsão probabilística de produção eólica. É expectável ainda, que esta cumpra os seguintes objetivos:

- Estudar e melhorar o modelo de previsão probabilística, com o objetivo de acrescentar novas funcionalidades, acelerar o tempo de processamento e alargar o número de variáveis de entrada que podem ser suportadas;
- Estudar a aplicação de várias heurísticas para a otimização de parâmetros de treino e ajuste de modelos de previsão probabilística de produção eólica;
- Otimizar e tipificar os parâmetros para modelos de previsão probabilística de produção eólica desenvolvidos.

1.3 Estrutura do documento

Este documento encontra-se dividido em seis capítulos.

No capítulo 1, é descrita a introdução da dissertação com o enquadramento e os objetivos que se esperam alcançar.

No capítulo 2, é abordado o estado da arte do documento, com a revisão bibliográfica dos modelos de previsão de produção eólica e dos modelos numéricos de previsão meteorológica.

No capítulo 3, é feito um estudo aprofundado sobre a ferramenta de previsão probabilística desenvolvendo a formulação matemática de um modelo inovador baseado no *Nadaraya-Watson Estimator*. As otimizações feitas à ferramenta computacional são também descritas.

No capítulo 4, é feita uma descrição da informação da produção de parque eólicos, cedida pela *Smartwatt*, e da metodologia utilizada para a apresentação dos resultados.

No capítulo 5, são apresentados os resultados obtidos dos testes de cada parametrização utilizada, do modelo final obtido e a análise dos resultados obtidos.

Por último, no capítulo 6 são descritas as conclusões da dissertação, bem como os possíveis trabalhos futuros.

Capítulo 2

Estado da arte

Este capítulo baseia-se no estudo bibliográfico sobre a previsão da produção eólica. Como principais referências, utilizou-se os artigos “Current status and future advances for wind speed and power forecasting” [3] e “Review on probabilistic forecasting of wind power generation” [6].

2.1 Modelos Numéricos de Previsão de Meteorologia (NWP)

A metodologia na análise de modelos numéricos de previsão de meteorologia baseia-se no uso de aproximações numéricas de uma equação dinâmica que prevê o estado futuro da circulação atmosférica com o conhecimento do estado atual. O estado atual é definido por um conjunto de variáveis, necessárias para inicializar o modelo NWP que é integrado nas equações de previsão temporais do modelo e que tem como objetivo prever a distribuição futura das variáveis meteorológicas[7].

A precisão dos modelos NWP depende da precisão dos processos físicos e numéricos utilizados, da resolução do modelo e da qualidade dos dados meteorológicos utilizados para definir as condições iniciais do modelo. A inclusão destes modelos NWP em modelos de previsão de produção eólica depende do horizonte para qual é feita a previsão [7, 8].

2.1.1 - Sistemas NWP Operacionais

Os modelos NWP podem ser classificados de acordo com as características espaço-temporais das variáveis meteorológicas. O grau de qualidade das previsões de cada modelo NWP varia com extensão geográfica e com a resolução temporal destas variáveis [9].

2.1.1.1 - Modelos Globais

Os modelos NWP globais são operados pelo NCEP (*National Centers for Environmental Prediction*), ECMWF (*European Center for Medium range Weather Forecasting*) e outros serviços nacionais meteorológicos importantes (Reino Unido, Alemanha, França, Brasil, Japão,

Russia). São operados no tempo universal coordenado (UTC) entre duas e quatro vezes por dia [7].

Tabela 2.1 - Modelos Globais [8]

Símbolo	Instituição	Modelo	Resolução Horizontal
ECMWF	European Center for Medium Range Weather Forecast	IFS	~25 km
METOF	Meteorological Office, UK	UM	~40 km
MSC	Meteorological Service of Canada	GEM	~30 km
NCEP	National Center for Environmental Prediction, USA	GFS	~50 km
METFR	Meteo France	ARPEGE	~15 km sobre França
DWD	Deutscher Wetterdienst, Germany	GME	40 km
AUSBM	Bureau of Meteorology, Australia	GASP	~80 km
JMA	Japan Meteorological Agency	JMA-GSM	~60 km
KMA	Korea Meteorological Agency	GDAPS	~45 km

2.1.1.2 - Modelos de área limitada (LAM - Limited Area models)

Os modelos de área limitada foram desenvolvidos com o objetivo de estudar os processos atmosféricos de mesoescala, que são todos os eventos atmosféricos com uma resolução geográfica entre um e algumas centenas de quilômetros [7].

Na Europa, o SRNWP (*Short Range Numerical Weather Prediction*), desenvolvido em 1993, tornou-se o principal veículo para cooperação entre os consórcios de modelos de área limitada[8]. Estes consórcios de previsão de previsão numérica meteorológicos são:

- *Aire Limitée Adaptation Dynamique Développement International* (ALADIN);
- *Consortium for Small-scale Modeling* (COSMO);
- *High Resolution Limited Area Model* (HIRLAM);
- *Regional Cooperation for Limited Area modeling in Central Europe* (RC LACE);
- *MetOffice United Kingdom* (UKMO).

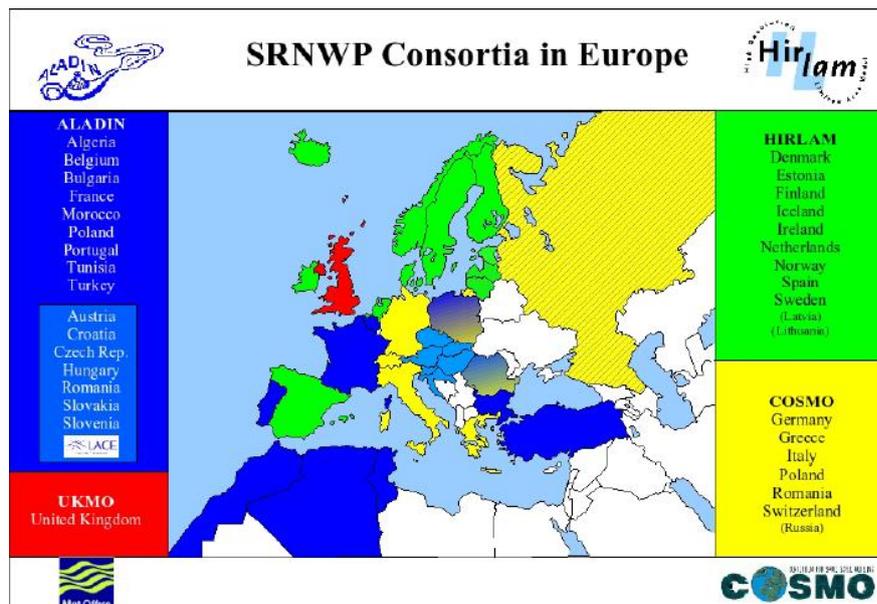


Figura 2.1 - Modelos de Área Limitada Europeus [8]

2.1.2 - Modelos NWP baseados em Ensemble

Os modelos NWP baseados em ensemble assentam-se num conjunto de previsões geradas a partir da perturbação dos dados de entrada e dos parâmetros de configuração do modelo, dentro dos respetivos intervalos de incerteza, de modo a obter um conjunto diferente de previsões. Uma vez que existe um grande número de variáveis e parâmetros para a execução dos modelos NWP, o objetivo consiste em selecionar os responsáveis pela maior parte da incerteza no sistema de previsão [3, 10].

Na tabela seguinte é possível observar alguns exemplos de modelos NWP em ensemble operacionais:

Tabela 2.2 - Sistemas NWP operacionais na Europa [8]

País	Modelo	Resolução Horizontal
Itália	COSMO-LEPS	10 km
Áustria	ALADIN-LAEF	18 km
Noruega	LAMEPS	12 km
Reino Unido	MOGREPS-G	90 km
	MOGREPS-R	24 km
França	PEARL (ARPEGE-EPS)	23 km
Hungria	ALADIN	12 km

2.2 Metodologias de Previsão Eólica

2.2.1 - Definição de Previsão de Produção Eólica

Uma previsão de produção eólica feita no instante t para um futuro próximo $t+k$ é a potência gerada esperada, $P_{t+k|t}$, que o parque eólico deverá produzir, durante um intervalo de tempo definido, se operar com uma velocidade de vento constante. As previsões são tipicamente geradas com recurso a um ou mais modelos de previsão de produção eólica, para um horizonte temporal, T , que indica a extensão total do período de previsão no futuro. Um modelo de previsão da produção eólica é um programa computacional que utiliza várias variáveis de entrada com o objetivo de prever a produção eólica no futuro [4, 10].

A classificação das fronteiras dos horizontes temporais não é unânime na bibliografia sobre a previsão de produção eólica, no entanto, em termos de requisitos de operação do sistema de energia, é comum dividir-se as previsões em quatro horizontes diferentes: previsão imediata (*Nowcasting*), muito curto prazo, curto prazo e médio prazo [4]. Na tabela seguinte é possível observar as características de cada horizonte temporal:

Tabela 2.3 - Características dos horizontes temporais

Horizonte Temporal	Escala de tempo	Aplicações	Informação Utilizada
Previsão imediata (<i>Nowcasting</i>)	De alguns segundos a 15 minutos	Controlo de turbina; Controlo do parque eólico.	Informação medida nos últimos segundos
Muito Curto prazo	De 15 minutos a 6 horas	Despacho; Seguimento de carga.	Informação da produção medida nas últimas horas
Curto prazo	De 6 horas a 72 horas	Preços do dia seguinte; Despacho; Gestão do parque eólico.	Informação meteorológica nas últimas 72 horas
Médio prazo	De 72 horas a 10 dias	Planeamento da manutenção das turbinas eólicas; Preços da semana seguinte.	Informação meteorológica até 10 dias

2.2.2 - Modelos de Previsão Físicos

A abordagem física para a previsão de produção eólica é baseada em métodos numéricos de previsão meteorológica (NWP) e utiliza dados como a temperatura, pressão, a rugosidade da superfície e obstáculos para caracterizar as condições do local do parque eólico. Contrariamente à abordagem estatística, a abordagem física não necessita de entradas baseadas em dados históricos, no entanto a grande desvantagem desta abordagem reside na dificuldade em adquirir os dados físicos necessários [3, 11]. Os modelos de previsão tipicamente seguem as seguintes etapas [3, 4]:

- Redução de escala (*downscaling*): A velocidade e a direção do vento são dimensionadas para a altura do eixo da turbina;
- Conversão para potência: Esta etapa consiste em converter a velocidade do vento em potência utilizando uma curva de potência;

- Modelos estatísticos: De forma a reduzir o erro resultante, que pode ser atribuído aos modelos NWP ou à abordagem utilizada, os modelos de previsão recorrem a modelos estatísticos (MOS - *Model Output Statistics*).

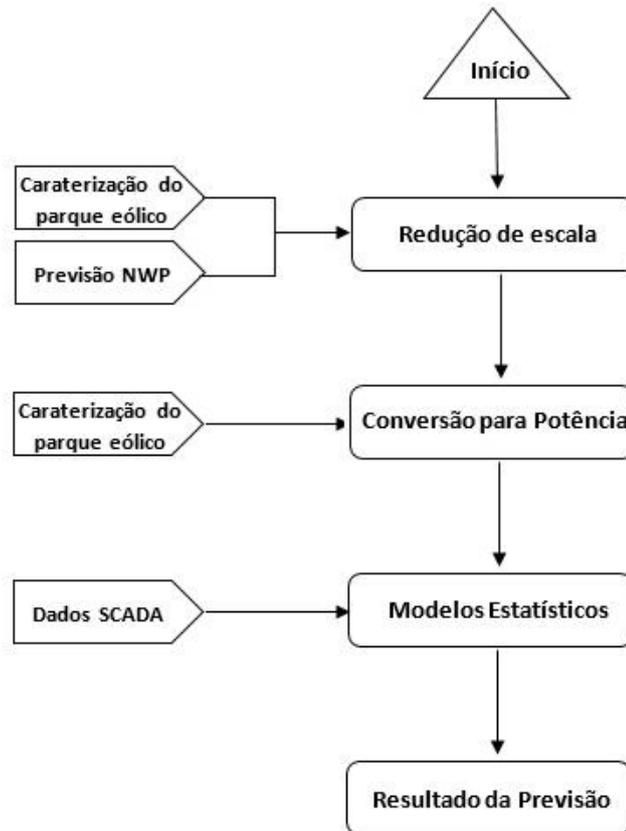


Figura 2.2 - Etapas dos modelos de previsão físicos [3]

Alguns exemplos de modelos físicos operacionais podem ser observados na tabela seguinte:

Tabela 2.4 - Modelos físicos operacionais [8]

Modelo	Instituição
Prediktor	Risoe National Laboratory (Dinamarca)
Previento	Universidade de Oldemburgo (Alemanha)
LocalPred	CENER - National Renewable Energy Centre (Espanha)
eWind	AWS TrueWind Inc. (Estados Unidos da América)

2.2.3 - Modelos de Previsão Estatísticos

A abordagem estatística revela a relação entre a previsão da velocidade do vento ou produção eólica e as variáveis explicadoras que incluem as variáveis atmosféricas dos modelos NWP e as variáveis *online* de medição disponibilizadas pelo SCADA. Os modelos estatísticos utilizam as previsões dos modelos NWP para o instante $t+k$ e as medidas *online* disponibilizadas no instante t para a previsão de produção eólica nas próximas horas, como se observa na Figura 2.3. Esta metodologia utiliza os dados históricos existentes para a formulação dos modelos [3].

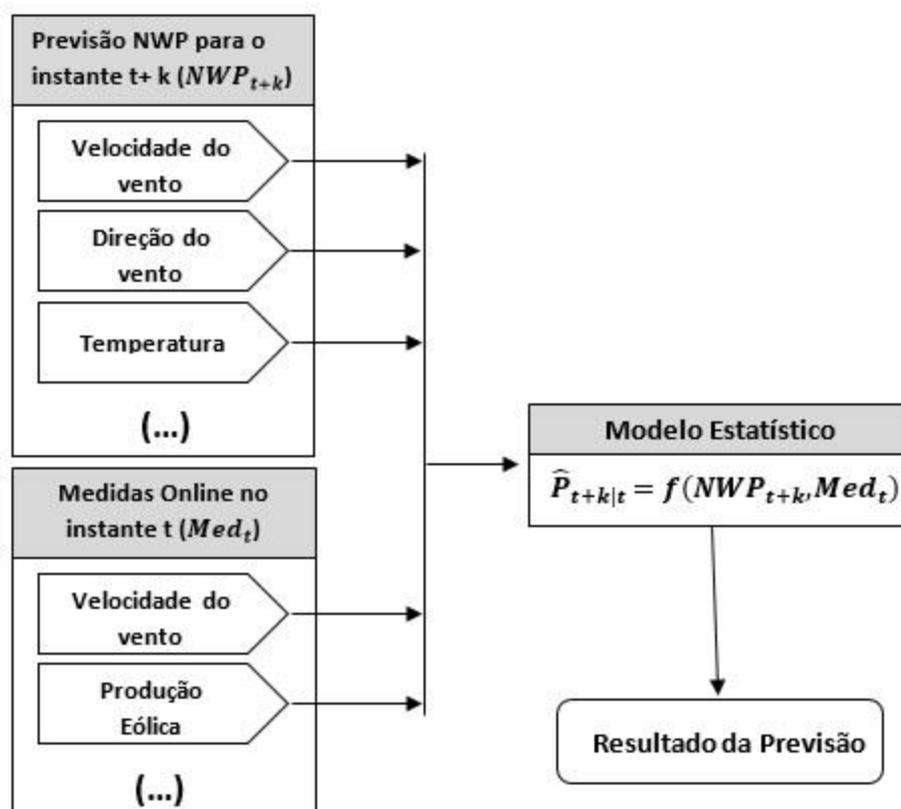


Figura 2.3 - Representação do modelo de previsão estatístico [3]

A facilidade na formulação dos modelos estatísticos e o seu baixo custo de desenvolvimento são as grandes vantagens desta abordagem. Na tabela seguinte estão representados alguns exemplos de modelos associados a esta metodologia [11]:

Tabela 2.5 - Modelos Estatísticos [3]

Abordagem	Modelo
Convencional	Box-Jenkins Autoregressive (AR) Moving average (MA) Autoregressive Moving Average (ARMA) Autoregressive Integrated Moving Average (ARIMA) Autoregressive Conditional Heteroscedastic (ARCH) Improved Time Series (ITSM)

Inteligência Artificial (ANN)	Particle Swarm Optimization (PSO) Modified Hybrid Neural Network (MHNN) Complex-Value Neural Network (CVNN) Real-Valued Neural Network (RVNN) Adaptive Wavelet Neural Network (AWNN) Feed-forward Neural Network (FNN) Kernel Ridge Regression (KRR)
ANN-Fuzzy	Wavelet Transform (WT)-Particle Swarm Optimization (PSO)- Adaptive Network based Fuzzy Inference System (ANFIS) Simultaneous Perturbation Stochastic Approximation (SPSA)- based Multilayer Feed-forward Neural Network (MFNN) with Fuzzy input algorithm
Outros	Modified Taylor Kriging (TK) Grey forecasting (GM) Hybrid Bayesian Clustering by Dynamic (BCD) and Support Vector Regression (SVR) Bayesian Model Averaging (BMA) with Markov Chain Monte Carlo (MCMC)

2.2.4 - Modelos de Previsão Regional

Os modelos de previsão regional têm como objetivo a previsão da potência agregada produzida por um conjunto de turbinas eólicas localizadas numa região. Esta metodologia de “força bruta” é utilizada para evitar os elevados custos, o esforço computacional excessivo e os problemas de controlo que a previsão da produção de cada um dos parques eólicos na região iria provocar [3, 4].

Estes modelos são geralmente baseados numa abordagem de *upscaling*, que calcula a potência eólica total da região através de previsões realizadas para um número representativo de parques eólicos de referência, com recurso a variáveis NWP e/ou medidas *online* acessíveis a estes parques. Além do mais, a agregação dos parques eólicos leva à redução dos erros de previsão devido aos efeitos de *spatial smoothing* [4].

2.2.5 - Combinação de Modelos de Previsão Eólica

As metodologias de combinação de modelos de previsão eólica têm como objetivo melhorar a *performance* final da previsão tirando partido dos pontos fortes de cada modelo. Como os modelos individuais têm um boa *performance* em situações específicas, é necessário estar sempre a testar vários modelos para decidir qual o modelo mais apropriado. O tempo desperdiçado nesta escolha é uma desvantagem dos modelos individuais e pode ser ultrapassado com o uso dos modelos combinados. Apesar disso, estes nem sempre têm uma melhor *performance* que os melhores modelos individuais de previsão. A combinação de modelos reduz ainda os riscos associados a eventos extremos como tempestades, onde os modelos individuais podem ter erros significativos na previsão [3, 12].

Os modelos combinados podem ser constituídos por modelos físicos, estatísticos e estatísticos alternativos. Na tabela seguinte é possível observar diferentes abordagens na combinação de modelos [3, 12]:

Tabela 2.6 - Características das abordagens na combinação de modelos [12]

Abordagens na combinação de modelos	Estratégia	Vantagens	Desvantagens
Combinação de modelos com base em fatores de ponderação	Atribuir fatores de ponderação aos modelos de acordo com a <i>performance</i>	Facilidade de implementação; Adequado para uma variada gama de previsões temporais.	Não garante as melhores previsões; Requer um modelo extra para calcular os fatores de ponderação.
Combinação de modelos incluindo técnicas de pré-processamento de dados	Previsão das subséries obtidas pelos modelos de decomposição	Alta <i>performance</i> ; Facilidade para encontrar exemplos.	Requer conhecimento matemático sobre os modelos de decomposição; Resposta lenta a novos dados.
Combinação de modelos incluindo seleção de parâmetros e técnicas de otimização	Otimização dos parâmetros do modelo de previsão	Estrutura Básica; Facilidade para encontrar exemplos.	Dificuldade na programação; Dependente do conhecimento do projetista; Esforço computacional intenso.
Combinação de modelos incluindo técnicas de processamento do erro	Previsão do erro residual causado pelo modelo de previsão	Alta precisão; Eficaz para reduzir o erro sistemático.	Tempo computacional ineficiente.

2.3 Avaliação dos Modelos de Previsão

Devido à incerteza que caracteriza os modelos de previsão de produção de eólica é essencial que estes sejam propriamente avaliados, não apenas a *performance* do modelo escolhido, mas também para obter um conhecimento mais profundo das características da incerteza [4].

A avaliação da qualidade dos modelos de previsão é feita a partir da comparação entre as previsões feitas por estes, num determinado intervalo de tempo, e os valores reais observados. Para a avaliação da qualidade da *performance* estatística dos modelos é necessário assegurar que os mesmos analisem um considerável número de previsões de forma a garantir a viabilidade dos resultados [4].

2.3.1 - Tipo de Erros de Previsão

Os erros associados à previsão de produção de eólica são geralmente associados aos modelos numéricos de previsão de meteorologia. Existem dois tipos de erros:

- Erros de nível;
- Erros de fase.

Os erros de nível resultam da má interpretação da severidade do evento meteorológico, enquanto os erros de fase resultam da má interpretação quanto ao momento do início e do pico do evento meteorológico [8].

2.3.2 - Modelos de Referência

Para testar a *performance* das novas abordagens para a previsão de produção de eólica existem vários modelos que servem de referência. Alguns desses modelos serão revistos neste subcapítulo.

2.3.2.1 - Modelo de Persistência

O modelo de persistência é o modelo mais utilizado para o teste da *performance* dos novos modelos de previsão. Neste modelo a produção de eólica no instante $t+k$ será a mesma que a produção ocorrida no instante t :

$$\hat{P}_{t+k|t} = P_t, \quad (2.1)$$

Onde:

$\hat{P}_{t+k|t}$ é a potência prevista no instante $t+k$ para uma previsão realizada no instante t ,

P_t é a potência ocorrida no instante t .

É o modelo que todos os outros modelos de previsão têm de superar para um horizonte de curto prazo [3, 8].

2.3.2.2 - Soma ponderada entre a persistência e a produção média de potência

Para horizontes de tempo superiores a algumas horas o modelo proposto é a soma ponderada entre a persistência e produção média de potência [3]:

$$\hat{P}_{t+k|t} = a_k P_t + (1 - a_k) \bar{P}, \quad (2.2)$$

Onde:

P_t é a potência ocorrida no instante t ,

\bar{P} é media estimada da potência,

a_k é o coeficiente de correlação entre P_t e P_{t+k} definido como:

$$a_k = \frac{(1/N) \sum_{t=1}^{N-k} \tilde{P}_t \tilde{P}_{t+k}}{(1/N) \sum_{t=1}^{N-k} \tilde{P}_t^2}, \quad (2.3)$$

Onde:

N é o número total de amostras,

\tilde{P}_t é dado pela diferença entre potência ocorrida e a média estimada da potência:

$$\tilde{P}_t = P_t - \bar{P}, \quad (2.4)$$

2.3.3 - Medidas de Avaliação da *Performance* dos Modelos

O erro de previsão para um instante $t+k$ realizada no instante t ($e_{t+k|t}$) pode ser obtido pela diferença entre o valor real medido no local no instante $t+k$ (p_{t+k}) e o seu valor previsto no instante t ($P_{t+k|t}$) [4]:

$$e_{t+k|t} = P_{t+k} - \hat{P}_{t+k|t}, \quad (2.5)$$

Onde:

$e_{t+k|t}$ é o erro correspondente a um dado instante $t+k$ para uma previsão realizada no instante t ,

P_{t+k} é a potência real medida no instante $t+k$,

$\hat{P}_{t+k|t}$ é a potência prevista no instante $t+k$ para uma previsão realizada no instante t .

O erro poderá ainda ser normalizado pela potência instalada (P_{inst}) de acordo com a equação 2.6 [3].

$$\varepsilon_{t+k|t} = \frac{1}{P_{inst}} e_{t+k|t}, \quad (2.6)$$

Para averiguar a qualidade das previsões realizadas, vários critérios de avaliação são utilizados [3, 4]:

- *Bias*, que traduz a média das diferenças entre os valores das previsões e os valores reais medidos:

$$Bias_k = média(e_{t+k|t}) = \bar{e}_k = \frac{1}{N} \sum_{t=1}^N e_{t+k|t}, \quad (2.7)$$

- Erro médio absoluto (MAE-Mean Absolute Error):

$$MAE_k = \frac{1}{N} \sum_{t=1}^N |e_{t+k|t}|, \quad (2.8)$$

- Erro médio absoluto percentual (MAPE - Mean Absolute Percentage Error):

$$MAPE_k = \frac{1}{N} \sum_{t=1}^N \left| \frac{e_{t+k|t}}{P_{t+k}} \right|, \quad (2.9)$$

- Erro quadrático médio (MSE - Mean Square Error):

$$MSE_k = \frac{1}{N} \sum_{t=1}^N (e_{t+k|t})^2, \quad (2.10)$$

- Raiz quadrada do erro quadrático médio (RMSE - Root Mean Square Error):

$$RMSE_k = \sqrt{MSE_k} = \sqrt{\frac{1}{N} \sum_{t=1}^N (e_{t+k|t})^2}, \quad (2.11)$$

- Desvio padrão dos erros (SDE - Standard Deviation of Errors):

$$SDE_k = \sigma(e_{t+k|t}) = \sqrt{\frac{1}{N} \sum_{t=1}^N (e_{t+k|t} - \bar{e}_k)^2}, \quad (2.12)$$

Onde:

N é o número total de amostras,

\bar{e}_k é a média do erro,

σ é o desvio padrão.

A raiz quadrada do erro quadrático (RMSE) poderá ainda ser expressa pela combinação do bias e o desvio padrão dos erros[3]:

$$RMSE_k^2 = Bias_k^2 + SDE_k^2, \quad (2.13)$$

Para comparar o nível de *performance* de um dado modelo, é importante quantificar a melhoria verificada da abordagem utilizada em relação aos modelos de referência. Esta poderá assim ser definida como[13]:

$$imp_{EC}^{ref}(k) = \frac{EVC_k^{ref} - EVC_k}{EVC_k^{ref}}, \quad (2.14)$$

Onde:

$imp_{EC}^{ref}(k)$ é a melhoria verificada da abordagem utilizada em relação ao modelo de referência (expressa normalmente em percentagem),

EVC_k^{ref} é o critério de avaliação do modelo de referência,

EVC_k é o critério de avaliação do modelo com nova abordagem.

Valores positivos da melhoria, $imp_{EC}^{ref}(k)$, indicam que as abordagens utilizadas são melhores que os modelos de referência, enquanto que valores negativos da melhoria demonstram uma pior *performance* das abordagens utilizadas em relação aos modelos de referência.

2.4 Incerteza na Previsão Eólica

2.4.1 - Modelos de Previsão Probabilística

Atualmente os modelos de previsão probabilística são os mais utilizados para a representação da incerteza na previsão de produção eólica. Tipicamente em forma de função densidade de probabilidade, conhecer a incerteza da previsão permite ao operador do sistema avaliar o risco inerente à mesma. Os modelos de previsão probabilística [3, 6] poderão ser classificados em paramétricos e não-paramétricos.

2.4.1.1 - Abordagem Paramétrica

Esta abordagem assume que a incerteza da previsão irá seguir uma forma de distribuição pré-definida caracterizada por parâmetros. A distribuição gaussiana é uma das escolhas mais comuns para esta abordagem e pode ser caracterizada pelo parâmetro de localização e o parâmetro de escala, como é descrito pela seguinte equação [6]:

$$\hat{f}_{t+k|t} = f(p_{t+k}, \mu, \sigma^2), \quad (2.15)$$

Onde:

$\hat{f}_{t+k|t}$ representa a função de distribuição da previsão

μ é o parâmetro de localização;

σ é o parâmetro de escala.

Uma das grandes vantagens da abordagem paramétrica é que a forma da distribuição depende de poucos parâmetros, o que resulta numa estimativa simplificada e com baixos custos computacionais. Porém, assumir uma forma de distribuição de previsão de produção eólica poderá, muitas vezes, não ser razoável [6].

Alguns exemplos de metodologias que se baseiam nesta abordagem podem ser revistos na tabela seguinte:

Tabela 2.7 - Metodologias baseadas na abordagem paramétrica [6]

Método	Observações
Séries temporais Homocedásticas	A forma da distribuição é assumida
Séries temporais Heterocedástico	Maior competitividade para a previsão a curto prazo
Inteligência Artificial	Menor custos computacionais devido à simplificação da estimativa

2.4.1.2 - Abordagem Não-Paramétrica

Contrariamente à abordagem anterior, a abordagem não-paramétrica não assume qualquer pressuposto em relação à forma da distribuição da incerteza da previsão e é utilizada quando não é possível formular a distribuição dos erros de previsão. Como as previsões da velocidade do vento e da produção eólica são muito diferentes consoante o horizonte de tempo e a localização, uma abordagem não-paramétrica é adequada para a previsão da incerteza correspondente [3, 6].

Nesta abordagem funções de densidade de probabilidade (PDF) ou funções de distribuição cumulativa (CDF) são estimadas num número finito de pontos. Com recurso à interpolação entre estes, é possível obter a descrição completa das funções PDF ou CDF.

Os custos computacionais para correr os modelos baseados na abordagem não-paramétrica são superiores aos custos correspondentes para modelos baseados na abordagem paramétrica [6]. Na tabela seguinte é possível rever alguns dos métodos não-paramétricos mais utilizados:

Tabela 2.8 - Metodologias baseadas na abordagem não-paramétrica [6]

Método	Observações
<i>Quantile Regression</i>	Não assume a forma da distribuição
<i>Kernel Density Estimation</i>	Requer um número elevado de amostras para construção da base de conhecimento
Inteligência Artificial	Elevados custos computacionais em consequência da complexidade da estimação

2.4.1.3 - Previsão baseada em Ensemble

Os modelos de previsão baseados em Ensemble são projetados de forma a poderem aproveitar as previsões disponibilizadas pelos modelos NWP em ensemble. Estes modelos podem fornecer vários cenários equiprováveis de variáveis meteorológicas, em vez de um único cenário como é providenciado por um modelo NWP convencional [3, 6].

Existem quatro abordagens que são utilizadas para produzir modelos de previsão baseados em ensemble [6]:

- Utilizar vários resultados de modelos NWP em funcionamento com pequenas diferenças nas condições iniciais;
- Utilizar vários resultados de modelos NWP em funcionamento com pequenas diferenças na representação numérica da atmosfera;
- Utilizar vários modelos NWP desenvolvidos por diferentes instituições;
- Utilizar o resultado das previsões feitas para diferentes origens temporais passadas.

As previsões baseadas em Ensemble dão uma melhor ideia aos utilizadores sobre os eventos meteorológicos que podem ocorrer no futuro. Ao comparar as diferentes previsões, os modelos identificam a variação esperada e calculam a probabilidade de um determinado evento meteorológico acontecer [3].

Para horizontes temporais de médio prazo, as previsões probabilística baseadas em ensemble têm uma melhor *performance* que as previsões baseadas em modelos NWP convencionais [6].

2.4.2 - Previsão do Índice de Risco da Produção Eólica

O índice de risco é uma representação simplificada da incerteza na previsão da produção eólica orientada para os utilizadores e pode ser definido como o nível esperado do erro de previsão. Quanto mais elevado este for, menos confiável é o resultado da previsão, podendo tal índice ser representado sob a forma de códigos de cores (por exemplo vermelho, amarelo, verde) ou números inteiros (por exemplo de 1 a 5). Através dele é possível obter conhecimento sobre a incerteza da produção eólica, necessário para a coordenação do parque eólico. Se o índice de risco for alto serão acionadas ferramentas preventivas com o objetivo de reduzir o risco potencial resultante do erro de previsão [14].

2.4.3 - Avaliação da Incerteza da Previsão Eólica

A avaliação da previsão da incerteza da produção eólica pode assinalar deficiências nos modelos de previsão, sugerir melhorias e ainda mostrar quais são os melhores modelos nessa área. As principais medidas de avaliação da previsão probabilística são a *reliability*, *sharpness* e *skill score* [6].

2.4.3.1 - Reliability

Para a tomada de decisões por parte dos utilizadores, é importante que as previsões probabilísticas de produção eólica sejam devidamente calibradas. *Reliability* refere-se à consistência estatística da distribuição prevista e observações, e representa a capacidade de um modelo de previsão probabilístico de alcançar as frequências observadas. Na Figura 2.4 é possível observar um exemplo de uma função de distribuição cumulativa em que um intervalo preditivo de 50% contém 50% dos valores reais observados. Neste caso ideal a proporção empírica do quantil preditivo é 50% e igual à proporção nominal. Na realidade, porém, a proporção empírica do quantil preditivo deve ser o mais próximo possível da proporção nominal[6, 15]:

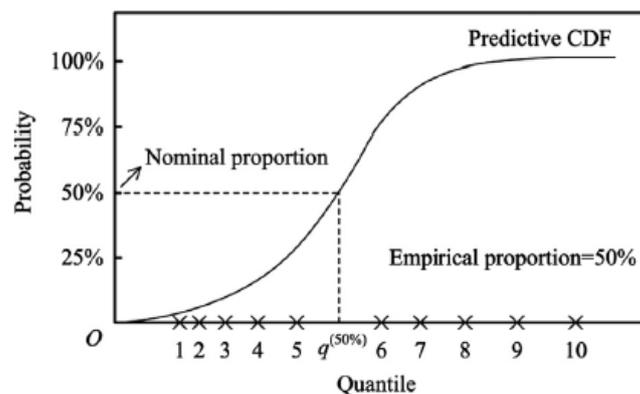


Figura 2.4 - Função de distribuição cumulativa (caso ideal) [6]

Um diagrama de *reliability* traça um gráfico das frequências reais da previsão de acordo com as probabilidades nominais. As previsões são agrupadas num horizonte de dias inteiros e as frequências reais, em que a produção de energia observada está abaixo ou acima de um dado quantil, são calculadas para quantis de 0 a 100% em incrementos de 5%. Se a previsão da incerteza da produção eólica apresenta a maior *reliability*, a cobertura de todas as ocorrências deverá ser constante [16]. Um exemplo deste tipo de diagramas pode ser observado na Figura 2.5.

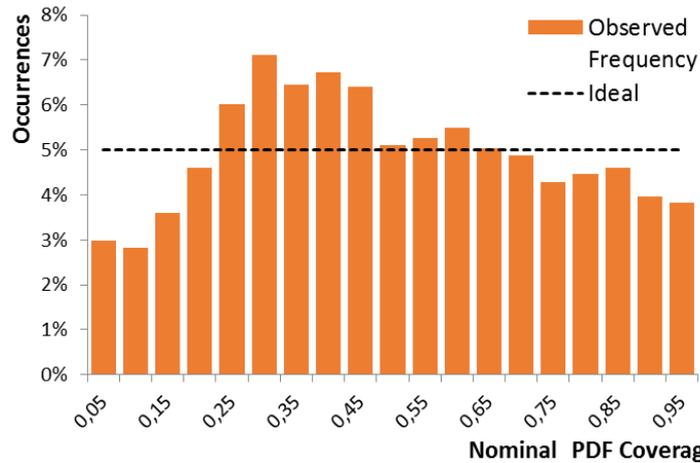


Figura 2.5 - Exemplo de um diagrama de reliability de um parque eólico

2.4.3.2 - Sharpness

A previsão probabilística de produção eólica deverá ser nítida, o que significa que a amplitude das distribuições preditivas deverá ser a menor possível. Uma distribuição preditiva mais concentrada produz mais informação e é mais competitiva que uma distribuição menos concentrada. Assim, a *sharpness* refere-se à concentração da distribuição preditiva e corresponde à capacidade da previsão probabilística para concentrar a informação probabilística sobre o resultado futuro. A amplitude desta distribuição é o foco da avaliação deste parâmetro na previsão da incerteza de produção eólica [17, 18]. Na figura seguinte é possível observar duas distribuições preditivas com diferentes níveis de *sharpness*:

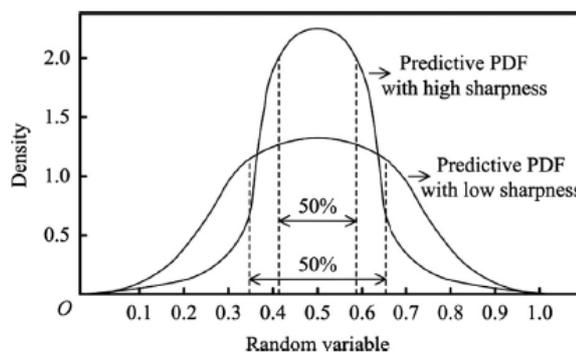


Figura 2.6 - Distribuição preditiva com diferentes níveis de *sharpness* [6]

O diagrama de *sharpness* representa a dispersão da incerteza da previsão em relação à escala de valores a prever. Quanto maior a amplitude do intervalo da incerteza, maior é a dispersão.

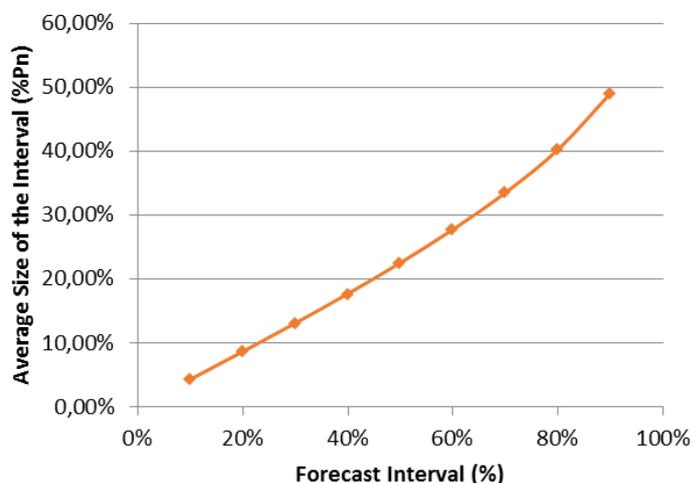


Figura 2.7 - Exemplo de um diagrama de *sharpness* de um parque eólico

2.4.3.3 - Skill score

A qualidade da previsão da incerteza não pode ser avaliada unilateralmente por uma certa propriedade, mas através da *reliability* e *sharpness* em conjunto. O *skill score* pode providenciar a informação completa da *performance* envolvendo estas duas propriedades e assegura que as previsões perfeitas obtenham os melhores resultados. Utilizar este parâmetro de avaliação único permite comparar o *skill* geral das abordagens rivais, porém não permite identificar as contribuições da *reliability* e da *sharpness* para o *skill score* [6, 18].

Diferentes *skill scores* são atribuídos para variáveis aleatórias e representações de incerteza. Para a avaliação da previsão da incerteza de previsão eólica, os mais utilizados são referenciados na seguinte tabela:

Tabela 2.9 - *Skill score* [6]

Representação da Incerteza	<i>Skill score</i>
Função de densidade de probabilidade	<i>Logarithmic score</i>
Função de distribuição cumulativa	<i>CRPS</i>
Quantis	<i>Tick or check loss score</i>
Intervalos	<i>Interval score</i>
Função de densidade de probabilidade com multivariáveis	<i>Energy score</i>

Capítulo 3

Modelo de Previsão Probabilística

Neste capítulo, é feito um estudo aprofundado sobre a ferramenta de previsão probabilística.

Baseada na ferramenta *Kernel Density Estimator* (KDE), criou-se uma variante desta família de modelos a que se designou *Nadaraya-Watson-Kernel Density Estimator* (NW-KDE). Com esta variante, o tempo computacional é acelerado e permite estimar diretamente os parâmetros da distribuição de probabilidade, em vez de estimar uma distribuição discreta não paramétrica como acontece ao utilizar a ferramenta *KDE*. No modelo NW-KDE desenvolvido integram-se algumas funcionalidades adicionais: pesos para os casos na base de conhecimento, modelização de variáveis cíclicas, inclusão de realimentação recursiva e parametrização de incerteza.

3.1 Definição do KDE

O KDE é um estimador de densidade condicional não-paramétrica onde nenhuma função de regressão paramétrica é assumida. A limitação dos erros de estimação, erros devidos à hipótese incorreta da função de regressão assumida, constitui a principal vantagem desta metodologia. No entanto, esta metodologia requer um conjunto de dados e um poder de processamento maior que um método que utiliza uma abordagem paramétrica [19].

Em problemas de previsão, a estimativa do futuro condicional da função de densidade tem um importante papel, ao descrever a relação entre as variáveis explicadoras e as variáveis dependentes. Tal estimativa pode ser vista como uma generalização da regressão, uma vez que tem como objetivo alcançar uma função densidade de probabilidade completa e não apenas um valor esperado. Na previsão probabilística a função densidade de probabilidade é usada para representar a incerteza. As primeiras referências à abordagem KDE na previsão de produção eólica foram apresentadas em 2007 [20].

Esta metodologia de previsão com KDE foi mais aprofundada, e em particular aplicada à previsão de eólica em 2012 na publicação [21]. A mesma metodologia foi desenvolvida e aplicada em *software* operacional pela empresa *Smartwatt*. Este *software* é baseado em técnicas de programação para GPU, com o objetivo de realização de previsões probabilísticas. Este trabalho foi desenvolvido em 2013 no âmbito de uma dissertação de mestrado [22] e mais

tarde, em 2014, a metodologia foi aplicada e publicada num relatório do projeto de investigação europeu FP7, SINGULAR [7]. Uma versão atualizada da metodologia foi também descrita em [23].

O KDE servirá de base para a criação de uma variante, que designaremos por NWE (*Nadaraya-Watson Estimator*). A abordagem é muito semelhante mas com algumas diferenças na formulação matemática e com melhorias muito significativas em tempo computacional.

3.1.1 - Formulação do Modelo com KDE

Consideremos um conjunto de N ocorrências históricas (V_i, P_i) , em que V_i representa, para cada caso i ($i = 1, \dots, N$), um vetor de variáveis explicadoras, com dimensão d , e P_i representa o correspondente vetor de variáveis dependentes, que vamos assumir com uma única dimensão.

Consideremos um caso futuro $(v_{t+k}, p_{t+k|t})$, em que V_{t+k} representa o vetor de variáveis explicadoras no instante futuro $t+k$, e $p_{t+k|t}$ representa o valor da variável dependente previsto no instante t para o instante $t+k$.

A estimativa da função densidade condicional $f_{p|v}$, condicionada pelas variáveis explicadoras futuras v_{t+k} , pode ser obtida utilizando a função de densidade conjunta $f_{p,v}$ e a função de densidade marginal f_v [24]:

$$f_{p|v}(P_i | v_{t+k}) = \frac{f_{p,v}(P_i, V_i)}{f_v(v_{t+k})}, \quad (3.1)$$

Onde:

$f_{p|v}(P_i | v_{t+k})$ é a função de densidade condicionada pelo caso v_{t+k} ;

$f_{p,v}(P_i, V_i)$ é a função de densidade conjunta para todos os casos históricos (P_i, V_i) ;

$f_v(v_{t+k})$ é a função de densidade marginal na vizinhança de v_{t+k} .

As funções de densidade podem ser representadas, na vizinhança de um ponto v_{t+k} , por uma função *kernel* (K), desde que a função *kernel* seja suficientemente apertada de forma que P_i seja razoavelmente constante no intervalo multidimensional H de ativação do *kernel* K . Mais tarde adotaremos a função gaussiana como função *kernel*.

A função densidade marginal, representada por funções de densidade *kernel*, apresenta-se na forma da equação 3.2:

$$\hat{f}(v_{t+k}) = \frac{1}{N \cdot |H|} \sum_{i=1}^N K(H^{-1}(v_{t+k} - v_i)) \quad (3.2)$$

Em que v_i é o vetor de variáveis explicadoras, para o caso histórico i , no qual está centrado o *kernel*, e v_{t+k} é o caso futuro representando o valor central da vizinhança, no intervalo H .

Para a função densidade conjunta, a representação por uma função *kernel*, será:

$$\hat{f}(P, v_{t+k}) = \frac{1}{N \cdot |H| \cdot h_p} \cdot \sum_{i=1}^N \left[K(H^{-1}(v_{t+k} - v_i)) \cdot K\left(\frac{P - P_i}{h_p}\right) \right] \quad (3.3)$$

Em que p_i é a variável dependente para o caso histórico i , em que p é um qualquer caso no espaço de variável dependente P , e em que h_p é a largura de banda do *kernel* aplicado na variável dependente.

A função densidade condicional, na vizinhança de v_{t+k} , resulta da substituição das equações 3.2 e 3.3 em 3.1. Assumindo $v = v_{t+k}$, o valor de p poderá ser qualquer no espaço da variável dependente P :

$$\hat{f}(p | v_{t+k}) = \frac{\sum_{i=1}^N \left(K(H^{-1}(v_{t+k} - v_i)) \cdot K\left(\frac{p - p_i}{h_p}\right) \right)}{h_p \cdot \sum_{i=1}^N K(H^{-1}(v_{t+k} - v_i))} \quad (3.4)$$

O denominador será constante e independente da variável dependente p .

Uma vez que a função densidade condicional é discreta e não normalizada, (sendo normalizada posteriormente), o denominador deixa de ter qualquer importância, simplificando-se a expressão:

$$\hat{f}(p | v_{t+k}) = \sum_{i=1}^N \left(K(H^{-1}(v_{t+k} - v_i)) \cdot K\left(\frac{p - p_i}{h_p}\right) \right) \quad (3.5)$$

Assumindo funções gaussianas para a função *kernel* das variáveis explicadoras, variáveis de dimensão d , tais funções podem ser decompostas no produto de d funções gaussianas:

$$K(u) = \mathcal{N}\left(\frac{v_{t+k} - v_i}{H}, 0, 1\right) = \prod_{j=1}^d \mathcal{N}\left(\frac{v_{t+k,j} - v_{i,j}}{h_j}, 0, 1\right) = \prod_{j=1}^d \mathcal{N}(v_{t+k,j}, v_{i,j}, h_j) \quad (3.6)$$

Onde $u = v_{t+k}$.

Para a variável dependente teremos:

$$K(u) = \mathcal{N}\left(\frac{p - p_i}{h_p}, 0, 1\right) = \mathcal{N}(p, p_i, h_p) \quad (3.7)$$

Onde $u = P$.

Substituindo 3.6 e 3.7 em 3.5, podemos representar a função densidade condicionada como:

$$\hat{f}(p | v_{t+k}) = \sum_{i=1}^N \left(\mathcal{N}(p, p_i, h_p) \cdot \prod_{j=1}^d \mathcal{N}(v_{t+k,j}, v_{i,j}, h_j) \right) \quad (3.8)$$

Note-se que $\mathcal{N}(v_{t+k,j}, v_{i,j}, h_j) = \mathcal{N}(v_{i,j}, v_{t+k,j}, h_j)$ poderá trazer vantagens de cálculo, permitindo mesmo especificar larguras de banda diferenciadas para cada caso $t+k$, da função da densidade de pontos na vizinhança, resolvendo assim problemas de representatividade

estatística da base de conhecimento. Este será um tema a desenvolver nesta dissertação. Assim, a função 3.8 poderá ser escrita como:

$$f(p|v_{t+k}) = \sum_{i=1}^N (\mathcal{N}(p, p_i, h_p) \cdot \prod_{j=1}^d \mathcal{N}(v_{i,j}, v_{t+k,j}, h_{t+k,j})) \quad (3.9)$$

A equação 3.9 tem um grau de complexidade matemática $P \times d \times N$, em que P é a população da discretização na variável dependente, d é a dimensão das variáveis explicadoras e N é a população de casos históricos.

Esta abordagem é denominada por KDE, resultando na função densidade condicional $f(p|v_{t+k})$, pode ser ainda representada e parametrizada por uma função densidade de probabilidade beta, sendo necessário obter a média e variância da distribuição discreta.

3.1.2 - Formulação do Modelo baseado em NW-KDE

Outra abordagem para a definição da função densidade de probabilidade beta é estimar diretamente o valor médio e a variância da função densidade condicionada. Para tal poderemos usar o estimador Nadaraya-Watson [25, 26].

O estimador Nadaraya-Watson permite-nos estimar o valor esperado de uma determinada função densidade condicional na vizinhança de um caso do vetor de variáveis explicadoras:

$$E(p|v) = \int p \cdot f(p|v) dp = \frac{\int p \cdot f(p,v) dp}{\int f(p,v) dp} \quad (3.10)$$

Daqui resulta, usando estimativas, a seguinte equação:

$$E(p|v_{t+k}) = \frac{\sum_{i=1}^N \left(p_i \prod_{j=1}^d K\left(\frac{v_{i,j} - v_{t+k,j}}{h_j}\right) \right)}{\sum_{i=1}^N \left(\prod_{j=1}^d K\left(\frac{v_{i,j} - v_{t+k,j}}{h_j}\right) \right)} \quad (3.11)$$

Segundo a representação explicada em 3.9 podemos escrever ainda:

$$E(p|v_{t+k}) = \frac{\sum_{i=1}^N \left(p_i \prod_{j=1}^d \mathcal{N}(v_{i,j}, v_{t+k,j}, h_{t+k,j}) \right)}{\sum_{i=1}^N \left(\prod_{j=1}^d \mathcal{N}(v_{i,j}, v_{t+k,j}, h_{t+k,j}) \right)} \quad (3.12)$$

A variância $V(p|v_{t+k}) = E(p^2|v_{t+k}) - E(p|v_{t+k})^2$, pode também ser estimada como:

$$V(p|v_{t+k}) = \left(\frac{\sum_{i=1}^N \left(p_i^2 \prod_{j=1}^d \mathcal{N}(v_{i,j}, v_{t+k,j}, h_{t+k,j}) \right)}{\sum_{i=1}^N \left(\prod_{j=1}^d \mathcal{N}(v_{i,j}, v_{t+k,j}, h_{t+k,j}) \right)} \right) - \left(\frac{\sum_{i=1}^N \left(p_i \prod_{j=1}^d \mathcal{N}(v_{i,j}, v_{t+k,j}, h_{t+k,j}) \right)}{\sum_{i=1}^N \left(\prod_{j=1}^d \mathcal{N}(v_{i,j}, v_{t+k,j}, h_{t+k,j}) \right)} \right)^2 \quad (3.13)$$

Desta forma, apenas com complexidade matemática $d \times N$, obtemos diretamente estimativas para a média e variância, necessárias para parametrizar a distribuição densidade de probabilidade Beta. Note-se que com esta abordagem h_p , usado na equação 3.9, não é requerido, simplificando a parametrização do modelo.

3.1.3 - Função densidade de Probabilidade Beta

Após aplicar a metodologia do NW-KDE, a densidade resultante é aplicada a uma distribuição Beta, considerada uma boa aproximação para a modelização de variáveis com limites máximos e mínimos, como é caso da produção eólica [7]:

$$f(x, \alpha, \beta) = \left(\frac{x^{\alpha-1} (1-x)^{\beta-1}}{B[\alpha, \beta]} \right), \quad (3.14)$$

Onde:

$B[\alpha, \beta]$ é uma constante de normalização para assegurar que f é uma função densidade de probabilidade;

x é um valor definido entre $[0, 1]$;

α, β são parâmetros com $\alpha, \beta > 0$.

Os parâmetros alfa e beta são calculados, para cada instante da previsão, através da média e da variância da distribuição Beta [7]:

$$\alpha = \frac{(1-\mu) \cdot \mu^2}{\mathcal{G}} - \mu, \quad (3.15)$$

$$\beta = \frac{(1-\mu)}{\mu} \cdot \alpha, \quad (3.16)$$

Onde:

μ é a média da densidade original;

\mathcal{G} é a variância da densidade original.

Na figura seguinte é possível observar exemplos de distribuições de densidade de probabilidade Beta aplicadas a histogramas de previsão obtidos através do KDE. No caso da metodologia NWE os parâmetros da distribuição são obtidos diretamente.

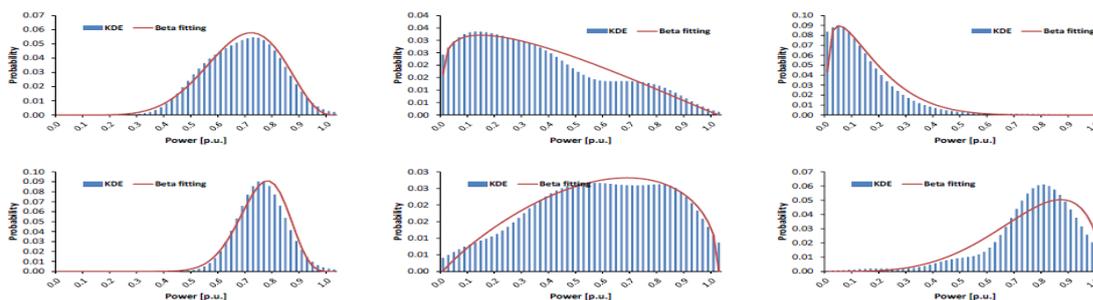


Figura 3.1 - Histogramas de previsão com distribuições Beta [23]

3.2 Limite mínimo de ativação da função *Kernel*

Esta limitação serve para acelerar os processos da ferramenta computacional. A gaussiana da distribuição *kernel* não é limitada pelo que assume valores maiores que zero em todos os pontos do histórico. Este facto obriga o algoritmo computacional a percorrer todos os pontos históricos ($v_{i,j}$), sendo este problema tanto maior quanto o número de dimensões, fazendo com que o KDE não seja computacionalmente aceitável para problemas com muitas variáveis de entrada. Para resolver o problema considera-se um nível mínimo para a ativação da gaussiana.

Com esta limitação os pontos $v_{i,j}$ ficam limitados à gama na proximidade do input $v_{t+k,j}$. Como a ativação do *kernel* se realiza sequencialmente o número de pontos a percorrer diminui progressivamente conforme se vão percorrendo as várias variáveis. Esta limitação acelera tremendamente o processo de cálculo, deixando de ser um problema o crescente número de variáveis. O algoritmo foi testado com 20 variáveis de entrada correndo em tempo aceitável.

Se o valor da função *kernel* da variável explicadora não obedecer à restrição da equação 3.17, então os pontos não serão considerados na iteração seguinte.

$$N(v_{i,j}, v_{t+k,j}, h_n) > \frac{1}{5000^{j+1}}, \quad (3.17)$$

3.3 Parametrização da entrada recorrente

As séries temporais, como a produção eólica, tem normalmente uma característica de segunda ordem, ou seja, a variabilidade entre instantes de tempo é uma informação importante que deve ser também modelizada pelo modelo de previsão. Para esta modelização é necessário usar como *input* no instante t a previsão realizada no instante $t - 1$. Este processo recorrente, usado em redes neuronais recorrentes, fica também adaptado para implementação no NWE. Este aspeto é especialmente importante na modelização das incertezas de variabilidade temporal.

O valor da entrada recorrente, V_{AR} , de cada caso histórico i será igual ao valor normalizado da variável dependente do caso histórico anterior (P_{i-1}). Para o primeiro caso histórico o valor da entrada recorrente será igual ao valor normalizado da variável dependente deste primeiro caso histórico (P_1):

Para o primeiro caso histórico:

$$V_{AR} = \frac{P_1 - \min(P_i)}{\max(P_i) - \min(P_i)}, \quad (3.18)$$

Para os restantes casos históricos:

$$V_{AR} = \frac{P_{i-1} - \min(P_i)}{\max(P_i) - \min(P_i)}, \quad (3.19)$$

Onde:

$\max(P_i)$ é o valor máximo da série de valores da variável dependente;

$\min(P_i)$ é o valor mínimo da série de valores da variável dependente.

A largura de banda da entrada recorrente, $h_{t+k,j}$, será dada pela seguinte equação:

$$h_{t+k,j} = \frac{1}{Part}, \quad (3.20)$$

Onde:

$Part$ é o número de partições do modelo autorregressivo definido pelo utilizador para a variável dependente. Este valor é uma das parametrizações a definir de forma a alcançar o melhor erro para o modelo de previsão.

3.4 Parametrização dinâmica da largura de banda da função *kernel*

A parametrização da largura de banda ótima das variáveis explicadoras é um dos pontos mais importantes para otimizar a previsão eólica utilizando o NW-KDE. O objetivo aqui é determinar a largura de banda ótima da função de densidade de pontos numa vizinhança em cada caso $t+k$.

Consideremos como exemplo uma dimensão do vetor das variáveis explicadoras em relação ao vetor de variáveis dependentes como se pode observar na seguinte figura:

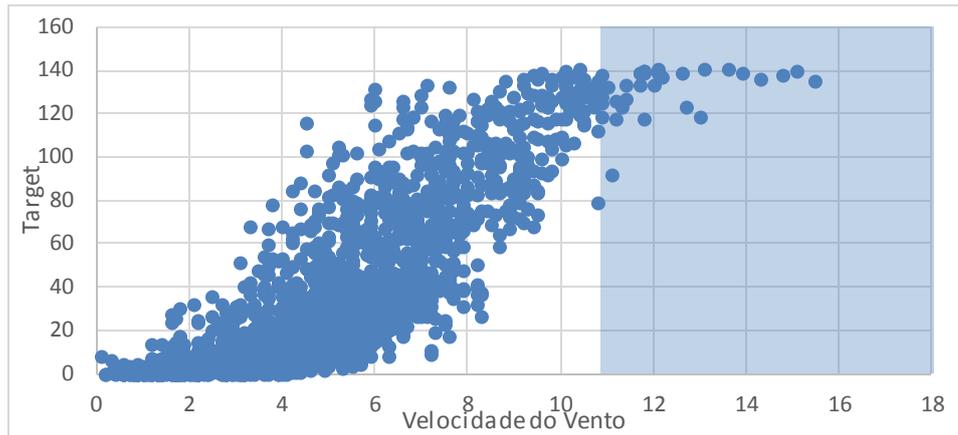


Figura 3.2 - Representação da velocidade do vento em relação à potência real registrada

A largura de banda ótima, centrada no vetor de variáveis explicadoras para o instante futuro $t + k$, deverá ser apertada nas zonas com grande quantidade de pontos (zona branca) e mais ampla nas zonas com menos pontos (zona colorida).

A largura de banda de cada dimensão da variável explicadora é calculada utilizando a seguinte equação:

$$h_{t+k,d} = h_{inicial} * h_{fator}, \quad (3.21)$$

Onde:

$h_{inicial}$ é a largura de banda inicial definida para a variável com base na ocorrência histórica;

h_{fator} é um fator multiplicativo criado para otimizar o valor da largura de banda das variáveis explicadoras.

3.4.1 - Cálculo da largura de banda inicial

A largura de banda inicial, $h_{inicial}$, pode ser determinada pela seguinte equação:

$$h_{inicial} = 0.075 * (\max(V_{i,j}) - \min(V_{i,j})), \quad (3.22)$$

Onde:

$\max(V_{i,j})$ é o valor máximo da série de valores da variável explicadora;

$\min(V_{i,j})$ é o valor mínimo da série de valores da variável explicadora.

3.4.2 - Cálculo do fator multiplicativo

Na determinação do fator multiplicativo, tem-se como objetivo obter a largura de banda (h_n) associada ao número mínimo de pontos representativos (Nhn), definido pelo utilizador e

para o qual o valor registado da função *kernel* da variável explicadora obedece à seguinte restrição:

$$N(v_{i,j}, v_{t+k,j}, h_n) > \frac{1}{5000^{j+1}}, \quad (3.23)$$

Onde:

j é o índice da variável explicadora;

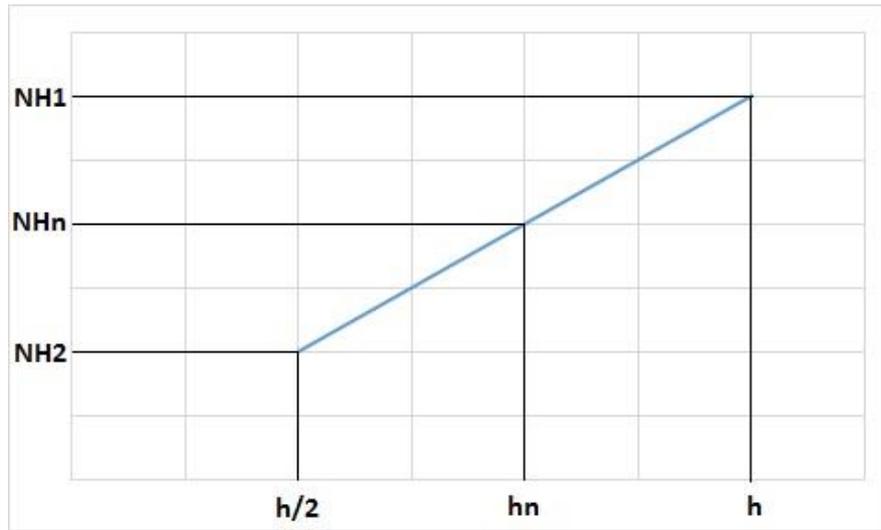


Figura 3.3 - Pontos para o cálculo do fator multiplicativo

Com recurso a dois pontos diferentes, o fator multiplicativo é calculado de acordo com a seguinte equação:

$$h_{fator} = \frac{\partial h}{\partial hn} = h - \frac{h - \frac{h}{2}}{NH1 - NH2} * (NH1 - NHn) = 1 - 0.5 * \frac{NH1 - NHn}{NH1 - NH2}, \quad (3.24)$$

Onde:

NHn é o número mínimo de pontos representativos definido pelo utilizador. Este valor terá de ser testado no final das parametrizações de forma a alcançar o melhor erro para o modelo de previsão;

$NH1$ é o número de pontos em que o valor registado da função *kernel* da variável explicadora obedece à seguinte restrição:

$$N(v_{i,j}, v_{t+k,j}, h_{inicial}) > \frac{1}{5000^{j+1}}, \quad (3.25)$$

NH_2 é o número de pontos em que o valor registado da função *kernel* da variável explicadora obedece à seguinte restrição:

$$N(v_{i,j}, v_{t+k,j}, \frac{h_{inicial}}{2}) > \frac{1}{5000^{j+1}}, \quad (3.26)$$

3.4.3 - Otimização da largura de banda inicial

A otimização da largura de banda inicial é conseguida com recurso às propriedades da função quadrática. O objetivo é determinar o valor da largura de banda para o qual o erro entre o valor real registado e valor esperado, calculado pelo NW-KDE, é menor. Considerando que tal erro é definido por uma função quadrática, para obter a equação da função quadrática é necessário calcular as coordenadas horizontais e verticais de três pontos, na escala onde por experiência se sabe que o modelo encontrará os melhores parâmetros h .

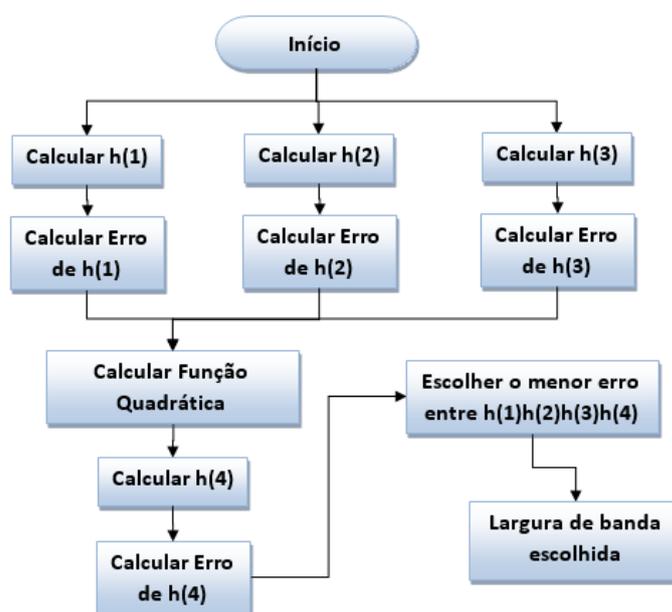


Figura 3.4 - Fluxograma do processo

Para os três pontos iniciais que definem a parábola, considerou-se uma largura de banda inicial, $h_{inicial}$, de 2,5%, 7,5% e 12,5% da escala total em cada uma das variáveis j :

$$h_{inicial,j}(1) = 0.025 * (\max(V_{i,j}) - \min(V_{i,j})), \quad (3.27)$$

$$h_{inicial,j}(2) = 0.075 * (\max(V_{i,j}) - \min(V_{i,j})), \quad (3.28)$$

$$h_{inicial,j}(3) = 0.125 * (\max(V_{i,j}) - \min(V_{i,j})), \quad (3.29)$$

Para cada um dos 3 pontos (terceira solução do parâmetro h_j), é feita a previsão, mantendo fixo o parâmetro h para as restantes variáveis diferentes de j . As ordenadas (y_1, y_2, y_3)

consideradas para a definição da parábola serão o erro entre o valor real registado e o valor esperado obtido:

$$e_j = MAPE_j(h_1, \dots, h_j, \dots, h_n), \quad (3.30)$$

Onde o MAPE é estimado pela equação 2.9.

Os três pontos $(x_1, y_1), (x_2, y_2), (x_3, y_3)$, considerados para a variável j , servirão para a formulação da função quadrática serão os seguintes:

$$Ponto1 = [h_{inicial,j}(1); e(h_{inicial,j}(1))] = (x_1, y_1), \quad (3.31)$$

$$Ponto2 = [h_{inicial,j}(2); e(h_{inicial,j}(2))] = (x_2, y_2), \quad (3.32)$$

$$Ponto3 = [h_{inicial,j}(3); e(h_{inicial,j}(3))] = (x_3, y_3), \quad (3.33)$$

Os coeficientes da função quadrática são determinados pelas seguintes equações:

$$a = \frac{x_3(-y_1 + y_2) + x_2(y_1 - y_3) + x_1(-y_2 + y_3)}{(x_1 - x_2)(x_1 - x_3)(x_2 - x_3)}, \quad (3.34)$$

$$b = \frac{x_3^2(y_1 - y_2) + x_2^2(-y_1 + y_3) + x_1^2(y_2 - y_3)}{(x_1 - x_2)(x_1 - x_3)(x_2 - x_3)}, \quad (3.35)$$

$$c = \frac{x_3(x_2(x_2 - x_3)y_1 + x_1(-x_1 + x_3)y_2) + x_1(x_1 - x_2)x_2y_3}{(x_1 - x_2)(x_1 - x_3)(x_2 - x_3)}, \quad (3.36)$$

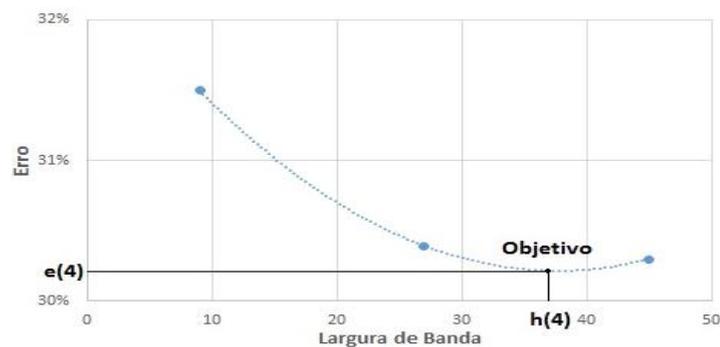


Figura 3.5 - Objetivo do cálculo da função quadrática

A função quadrática, para cada variável j , é então representada pela expressão seguinte:

$$f(x) = a * (h_j)^2 + b * (h_j) + c, \quad (3.37)$$

O valor teórico da largura de banda, para o qual o erro entre o valor real registado e valor esperado é menor, será dado pela coordenada horizontal do vértice da parábola. A coordenada horizontal do vértice, $h_j(4)$, poderá ser calculada com recurso à derivada da função quadrática:

$$h_j(4) = -\frac{b}{2a}, \quad (3.38)$$

Na prática, nem sempre este valor teórico de largura de banda corresponde ao menor erro registado. Assim, o valor da largura de banda inicial será o que registou o menor erro de entre os quatro valores de largura de banda utilizados nesta otimização:

$$h_j = \min(e_j(1); e_j(2); e_j(3); e_j(4)), \quad (3.39)$$

Se o valor teórico de largura de banda for negativo, este não será incluído na equação anterior.

O processo é repetido sequencialmente, percorrendo todas as variáveis j .

3.5 Parametrização das variáveis cíclicas

Para as variáveis cíclicas é necessário que a ferramenta computacional reconheça que o primeiro e último valor da série de valores da variável estão interligados. Por exemplo, para a variável hora, a ferramenta computacional tem que reconhecer que na vizinhança da hora 23, considerando uma janela de amplitude 2, estão os valores registados na hora 1.

Nas figuras seguintes é possível analisar o efeito da introdução da parametrização na ferramenta computacional:



Figura 3.6 - Representação da densidade marginal da variável hora sem a parametrização das variáveis cíclicas



Figura 3.7 - Representação da densidade marginal da variável hora com a parametrização das variáveis cíclicas

Na Figura 3.7 é possível observar que a adição da parametrização permite à ferramenta computacional reconhecer que os valores registados nas horas 1 e 2 estão na vizinhança dos valores registados na hora 23, ao passo que na representação sem a parametrização (Figura 3.6) o mesmo não acontece.

Para efetuar a parametrização é calculado em primeiro lugar a diferença (*dif*) entre o valor mínimo, $\min(V_{i,j})$, e o valor máximo, $\max(V_{i,j})$, da série de valores da variável explicadora cíclica:

$$dif = \max(V_{i,j}) - \min(V_{i,j}), \quad (3.40)$$

Para o cálculo da função *kernel* da variável explicadora cíclica acrescentou-se duas distribuições normais iguais, deslocadas de acordo com a diferença calculada no ponto anterior:

$$K(u) = \prod_{j=1}^d N(v_{i,j}, v_{t+k,j}, h_j) + \prod_{j=1}^d N(v_{i,j}, -dif + v_{t+k,j}, h_j) + \prod_{j=1}^d N(v_{i,j}, v_{t+k,j} + dif, h_j) \quad (3.41)$$

3.6 Parametrização do peso de cada caso histórico

O peso atribuído a cada caso histórico é um valor compreendido entre zero e um valor que reflete a importância deste caso em relação às restantes ocorrências históricas. Um caso histórico com peso zero não terá qualquer influência no resultado final da previsão eólica.

O peso de cada ocorrência histórica será dado pela equação 3.42:

$$Peso = 1 - \frac{|P_{t+k} - \hat{P}_{t+k|t}|}{\text{Max}(P_{t+k}, \hat{P}_{t+k|t}) + 0.1 * \text{Max}(P_{t+k}, \hat{P}_{t+k|t})}, \quad (3.42)$$

Onde:

P_{t+k} é a potência real medida no instante t+k,

$\hat{P}_{t+k|t}$ é a potência prevista no instante t+k para uma previsão realizada no instante t, $Max(P_{t+k}, \hat{P}_{t+k|t})$ é a potência máxima registrada entre P_{t+k} e $\hat{P}_{t+k|t}$.

3.7 Parametrização para ajustar a incerteza da previsão eólica

O objetivo desta parametrização é aumentar a consistência estatística da distribuição e observações da previsão eólica. Os valores de alfa (α), beta (β), mínimo da função Beta ($MinFB$) e máximo da função Beta ($MaxFB$) serão alterados para que estes se adaptem melhor aos valores reais. O valor de cada parâmetro necessário para esta parametrização será obtido através de um conjunto de validação que será um conjunto de ocorrências registradas relativas a uma semana da previsão.

$$Novo_ \alpha_{t+k} = \alpha_{t+k} * a_{\alpha} + b_{\alpha}, \quad (3.43)$$

$$Novo_ \beta_{t+k} = \beta_{t+k} * a_{\beta} + b_{\beta}, \quad (3.44)$$

$$NovoMinFB_{t+k} = (MinFB_{t+k}) - Expand * ((MinFB_{t+k}) - (\min(MinFB_{t+k}))), \quad (3.45)$$

$$NovoMaxFB_{t+k} = (MaxFB_{t+k}) - Expand * ((MaxFB_{t+k}) - (\max(MaxFB_{t+k}))), \quad (3.46)$$

Onde:

$a_{\alpha}, b_{\alpha}, a_{\beta}, b_{\beta}$ são os parâmetros de ajuste de incerteza que serão otimizados a partir de testes efetuados ao conjunto de validação;

$Novo_ \alpha_{t+k}, Novo_ \beta_{t+k}, NovoMinFB_{t+k}, NovoMaxFB_{t+k}$ são os novos parâmetros da função Beta, de cada ocorrência testada, que são atualizados na ferramenta computacional;

$\min(MinFB_{t+k})$ é o valor mínimo da função Beta de todas as ocorrências testadas;

$\max(MaxFB_{t+k})$ é o valor máximo da função Beta de todas as ocorrências testadas;

O parâmetro *Expand* poderá tomar valores entre 0 e 1, em que 0 não afeta a largura dos limites da distribuição beta e 1 alarga os limites da distribuição beta para os limites da variável dependente.

Capítulo 4

Modelo de Previsão e Metodologia

Neste capítulo é feita a análise de séries de produção de quatro parques eólicos cedidas pela *Smartwatt* e que será usada para testar as várias otimizações feitas à ferramenta computacional. Será, ainda, descrita a metodologia em que a apresentação de resultados se irá basear.

4.1 Modelo de previsão

4.1.1 - Tipo de variáveis

Para o teste das várias otimizações criadas para o NW-KDE serão utilizados os dados relativos a quatro diferentes parques eólicos. Para cada um destes serão utilizadas o mesmo tipo de dados e que podem ser divididos em diferentes grupos:

- Inputs (ou entradas): São as variáveis utilizadas para inicializar o NW-KDE (por exemplo velocidade do vento, temperatura, hora);
- Target: São os valores reais registados de potência eólica relativos às variáveis utilizadas;
- Histórico: O histórico para a previsão da potência eólica do parque será constituído por variáveis de entrada de instantes passados com o seu respetivo *target* e formam a base de dados onde o NW-KDE se irá basear.

Casos	Variáveis Explicadoras				Target	
	X1	X2	X3	X _d	Y	
Histórico	i=1	φ_{11}	φ_{12}	φ_{13}	φ_{1d}	Y ₁
	i=2	φ_{21}	φ_{22}	φ_{23}	φ_{2d}	Y ₂
	i=3	φ_{31}	φ_{32}	φ_{33}	φ_{3d}	Y ₃
	(...)	(...)				
	i=N	φ_{N1}	φ_{N2}	φ_{N3}	φ_{Nd}	Y _N

Figura 4.1 - Representação do histórico de previsão

O objetivo da ferramenta é calcular os outputs (ou saídas) da previsão, que são as variáveis de saída do NW-KDE e que transmitem o resultado obtido da previsão para o instante que esta é feita. Os outputs de cada parque serão calculados a partir de variáveis de entradas de um conjunto de variáveis de teste disponibilizadas.

Instante de Teste				
Variáveis Explicadoras				Output
φ_1	φ_2	φ_3	φ_d	$\gamma (?)$

Figura 4.2 - Representação do *output* da previsão

4.1.2 - Histórico do Modelo de Previsão

As variáveis presentes no histórico dos parques eólicos são uma parte essencial do modelo de previsão. Se o histórico não representar de forma correta o comportamento dos parques, a previsão dos *outputs* não apresentará valores aceitáveis. Neste subcapítulo, apenas será representado o histórico de um dos parques que serão testados. Para os restantes, o histórico será reproduzido nos anexos da dissertação.



Figura 4.3 - Amostra do *Target* do Parque 1

Tabela 4.1 - Características de cada parque

Produção (MW)	Potência máxima	Valor médio	Variância	Fator de Capacidade média	Fator de Sharpness médio
Parque 1	141	41	1561	0,295	0,281
Parque 2	233	71	4424	0,304	0,286
Parque 3	216	63	3771	0,293	0,285
Parque 4	142	46	1836	0,323	0,301

As variáveis explicadoras de entrada serão dados meteorológicos recolhidos ou previstos com a hora associada. Para cada um dos quatro parques testados, serão utilizados o mesmo tipo de variáveis:

- Hora;
- Velocidade do vento (m/s);
- Direção do vento (graus);
- Temperatura (°C).



Figura 4.4 - Amostra da velocidade do vento do Parque 1



Figura 4.5 - Amostra da direção do vento do Parque 1

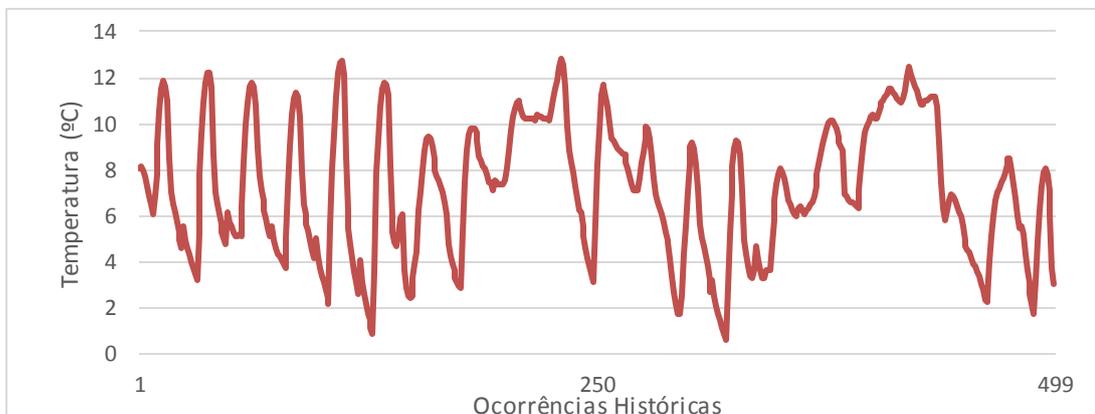


Figura 4.6 - Amostra da temperatura do Parque 1

4.1.3 - Variáveis para teste

O teste da ferramenta será realizado com a informação relativa a 23 semanas. Para cada semana será disponibilizada a informação relativa a 6 dias. Para cada dia será disponibilizada a informação relativa a 12 horas. No total serão testados 1656 ocorrências. Para o cálculo do valor de cada parâmetro necessário para o teste da parametrização do ajuste da incerteza será utilizado ocorrências de uma semana diferente às anteriores.

Variáveis Teste		Hora	Velocidade do vento	Direção do vento	Temperatura
Semana	Dia 1	9	φ_{i2}	φ_{i3}	φ_{i4}
		(...)			
		20	φ_{i2}	φ_{i3}	φ_{i4}
	Dia 6	9	φ_{i2}	φ_{i3}	φ_{i4}
		(...)			
		20	φ_{i2}	φ_{i3}	φ_{i4}

Figura 4.7 - Representação das variáveis de entrada do teste relativos a uma semana

Para efeitos de análise serão, ainda, disponibilizados os valores reais das potências eólicas registados para os instantes em que são feitas as previsões.

4.1.4 - Redes Neurais

Para efeitos de análise dos *outputs* calculados pelo NW-KDE é feita a mesma previsão da potência eólica com recurso a outra ferramenta computacional, redes neurais (4n+1 neurónios, 5 previsões em paralelo). As variáveis de entrada dos testes serão as mesmas que as utilizadas pelo NW-KDE. Na figura seguinte é possível observar uma amostra dos valores obtidos da previsão eólica através das redes neurais para o Parque 1. Nos anexos da dissertação estão representados os valores da previsão dos restantes parques.

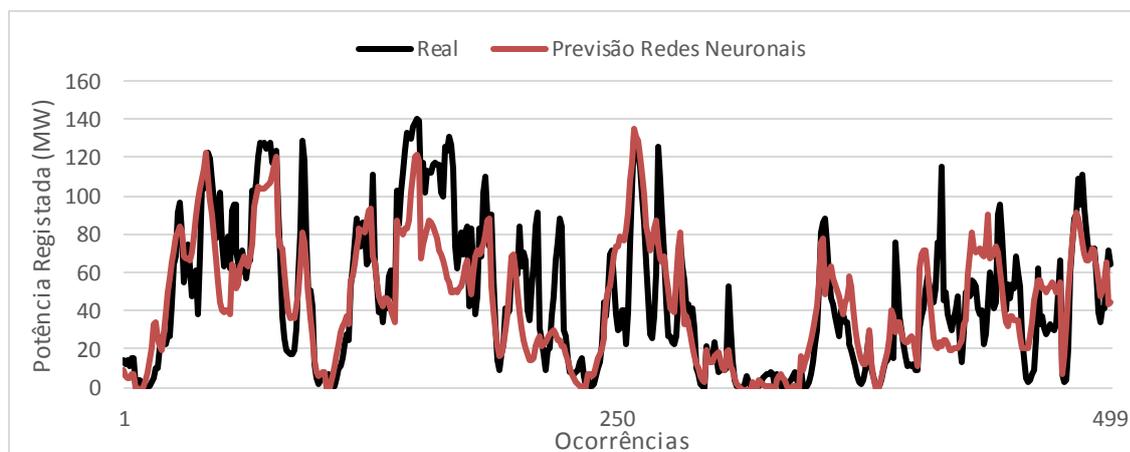


Figura 4.8 - Amostra dos valores obtidos da previsão eólica através redes e os valores registados reais do Parque 1

4.2 Metodologia dos resultados

4.2.1 - Testes realizados

Cada parametrização do NW-KDE será testada de forma sequencial. O primeiro teste realizado apresentará os valores registados do NW-KDE sem parametrizações incluídas, a partir daí serão acrescentadas as parametrizações de acordo com a seguinte ordem:

- **Teste 1:** Modelo inicial;
- **Teste 2:** Teste 1 + Parametrização da entrada recorrente;
- **Teste 3:** Teste 2 + Parametrização dinâmica da largura de banda;
- **Teste 4:** Teste 3 + Parametrização das variáveis cíclicas;
- **Teste 5:** Teste 4 + Ajuste da largura de banda da entrada recorrente e do número mínimo de pontos representativos definido pelo utilizador;
- **Teste 6:** Teste 5 + Parametrização do peso de cada caso histórico;
- **Teste 7:** Teste 6 + Parametrização para ajuste da incerteza;
- **Modelo Final:** Modelo com os melhores resultados.

Se os resultados de cada teste não melhorarem os resultados do teste anterior, a parametrização associada ao teste não será considerada no teste seguinte.

As previsões do teste 2 ao teste 6 servem para melhorar o erro da previsão determinística e portanto não será analisada a incerteza associada destes. Por outro lado, apenas a incerteza do teste 7 será analisada. Para o modelo inicial e o modelo ótimo será analisada a previsão determinística e a incerteza associada.

4.2.2 - Apresentação de resultados

Para uma mais fácil compreensão dos resultados obtidos será seguida a mesma metodologia na apresentação de resultados para todos os testes realizados.

Para a previsão determinística, as potências previstas pelo NW-KDE serão comparadas às potências previstas pelas redes neuronais e, ainda, aos valores reais registados das potências produzidas pelos parques eólicos para o instante em que são feitas as previsões. Esta análise será apresentada graficamente e será feita para os quatro parques eólicos que estão a ser testados. Será ainda apresentada a incerteza associada para cada um dos parques.

Devido ao alto número de ocorrências testadas apenas os dados relativos entre a semana 9 e semana 14 serão apresentados.

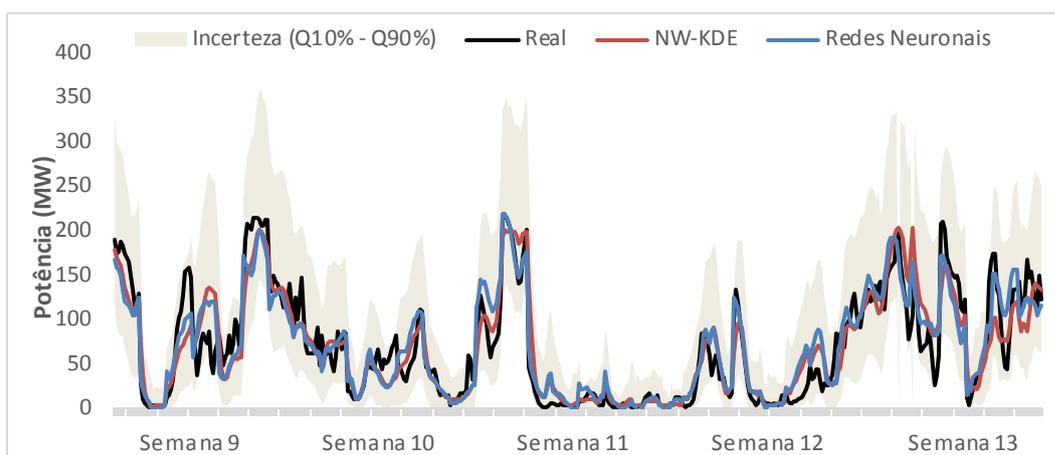


Figura 4.9 - Exemplo da representação da previsão da produção eólica e a incerteza associada

Em seguida será analisado o erro médio absoluto em percentagem da produção média (MAPE) do *NW-KDE* em relação ao MAPE das redes neuronais e, ainda, ao MAPE do teste anterior.

O MAPE das potências previstas, em relação aos valores reais registados, será dado pela equação seguinte:

$$MAPE = \frac{MAE}{AP}, \quad (4.1)$$

O erro médio absoluto (MAE) é dado pela equação 4.2.

$$MAE_k = \frac{1}{N} \sum_{t=1}^N |P_{t+k} - \hat{P}_{t+k|t}|, \quad (4.2)$$

O valor médio real (AP) é caracterizado na equação 4.3.

$$AP = \frac{1}{N} \sum_{t=1}^N (P_{t+k}), \quad (4.3)$$

Onde:

N representa o número de casos do teste;

P_{t+k} é a produção real para o instante $t+k$;

\hat{P}_{t+k} é a produção prevista para o instante $t+k$.

Os erros das potências reais e previstas (*NW-KDE* e redes neuronais) serão para calculados para cada um dos parques testados.

A análise da incerteza da previsão eólica será analisada através de um diagrama de *reliability*, o desvio médio quadrático associado e de um diagrama de *sharpness*. Estes diagramas estão explicados na seção 2.4.3.

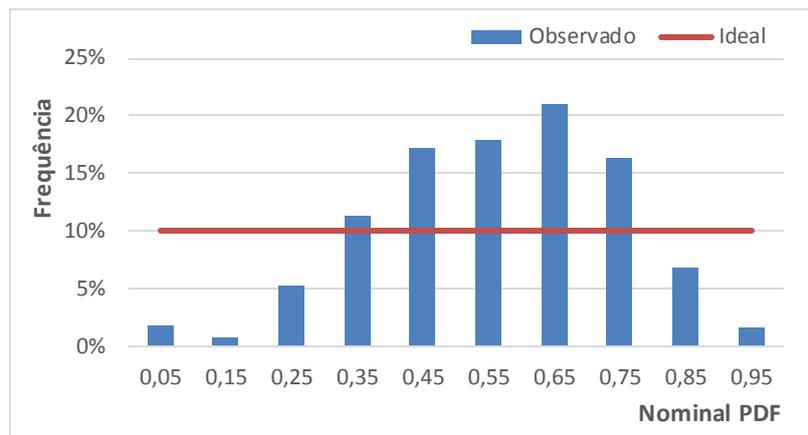


Figura 4.10 - Exemplo da representação de um diagrama de *reliability*

A incerteza ideal da previsão ocorre quando as frequências da função de densidade de probabilidade são constantes em todos os intervalos de probabilidade.

O desvio médio quadrático da *reliability*, D_r , será dado por:

$$D_r = \frac{1}{N} \sum_{i=1}^N (e_{FQ})^2, \tag{4.5}$$

Onde:

N representa o número de intervalos de probabilidade analisados, neste caso serão 10;

e_{FQ} é o erro da frequência registada em relação à frequência ideal de cada intervalo de probabilidade.

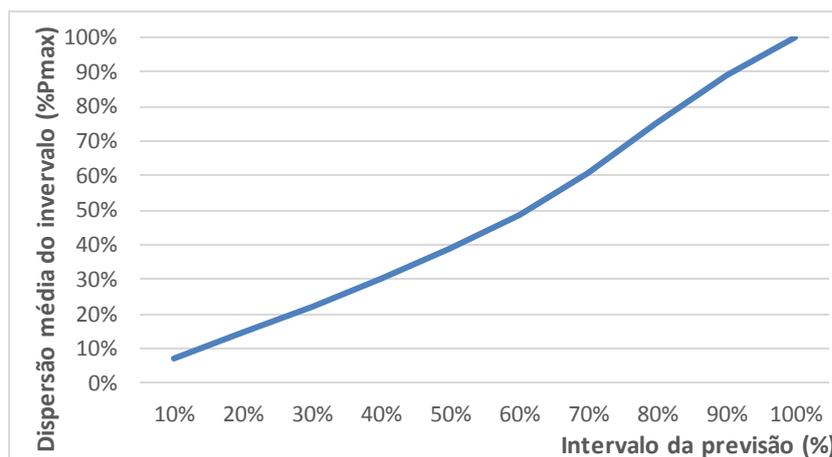


Figura 4.11 - Exemplo da representação de um diagrama de *sharpness*

O diagrama de *sharpness* representa a dispersão da incerteza da previsão em relação à escala de valores a prever. A dispersão registada será apresentada como uma percentagem da potência máxima registada de cada parque.

Capítulo 5

Resultados Obtidos

Ao longo deste capítulo são apresentados os resultados obtidos do teste de cada parametrização utilizada para os quatro parques mencionados anteriormente. No final, será apresentado o modelo com os melhores resultados obtidos.

5.1 Teste 1: Modelo Inicial

Neste primeiro teste, o modelo será testado sem nenhuma das novas parametrizações criadas. Os parâmetros utilizados para o teste deste modelo estão representados na tabela 5.1.

Tabela 5.1 - Parametrizações do teste 1

	Parque 1	Parque 2	Parque 3	Parque 4
Número de casos históricos	12714	12410	12522	12494
Número de casos para teste	1656	1656	1656	1656
Entrada recorrente	Não			
Número de partições	----	----	----	----
Largura de banda dinâmica	Não			
Largura de banda (Variável 1)	1,725	1,725	1,725	1,725
Largura de banda (Variável 2)	1,471	1,775	1,454	1,378
Largura de banda (Variável 3)	26,999	27,000	27,000	26,999
Largura de banda (Variável 4)	2,993	2,882	2,932	2,969
Número mínimo de pontos	----	----	----	----
Variáveis Cíclicas	Não			
Pesos	Não			
Ajuste da incerteza	Não			

42 Resultados Obtidos

Parâmetro a_α	----	----	----	----
Parâmetro a_β	----	----	----	----
Parâmetro b_α	----	----	----	----
Parâmetro b_β	----	----	----	----
Expand	----	----	----	----

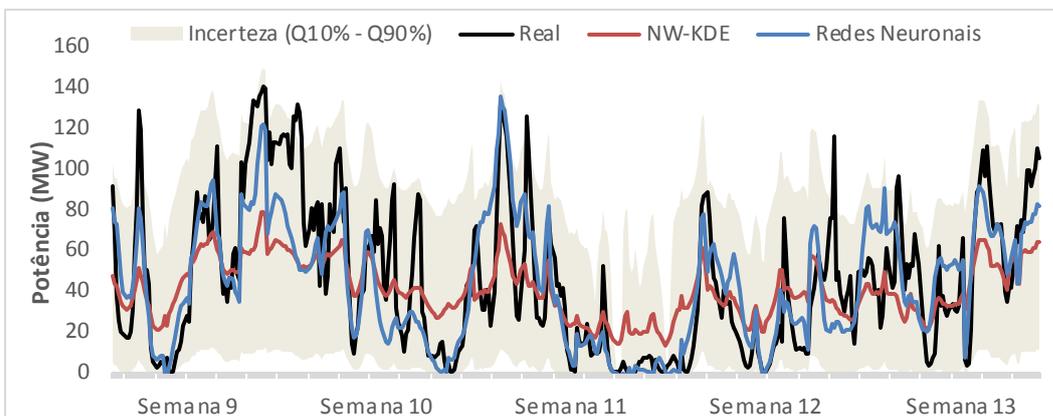


Figura 5.1 - Previsão da produção eólica e a incerteza associada do Parque 1

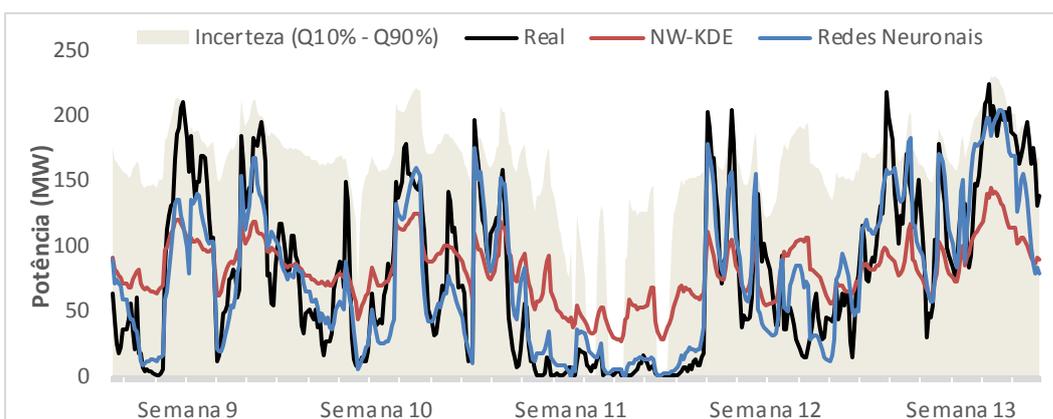


Figura 5.2 - Previsão da produção eólica e a incerteza associada do Parque 2

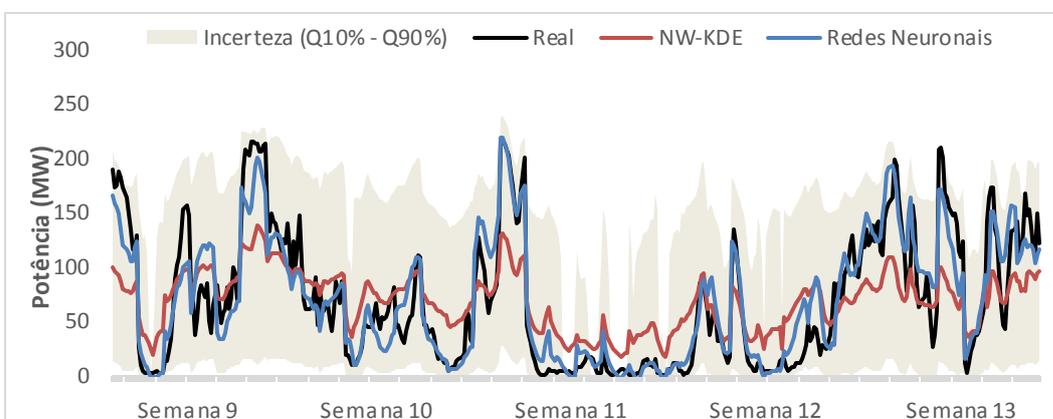


Figura 5.3 - Previsão da produção eólica e a incerteza associada do Parque 3

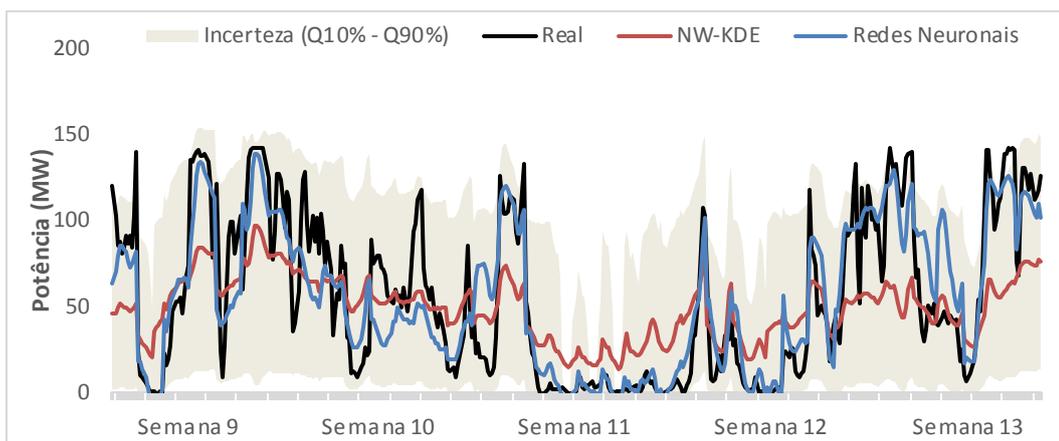


Figura 5.4 - Previsão da produção eólica e a incerteza associada do Parque 4

Tabela 5.2 - Erros obtidos dos valores previstos para o Teste 1

Erro	Redes Neuronais	NW-KDE
Parque 1	31,74%	53,17%
Parque 2	28,34%	56,73%
Parque 3	27,42%	52,74%
Parque 4	35,63%	53,47%

O erro deste modelo inicial do NW-KDE é significativamente elevado se compararmos com o erro registado pelas redes neuronais. É possível observar que o NW-KDE não consegue prever os picos de produção. Verifica-se ainda, que os maiores parques (Parques 2 e 3) têm menor erro que os parques mais pequenos (Parques 1 e 4), porém existem outros fatores como a complexidade do terreno e a localização que também influenciam.

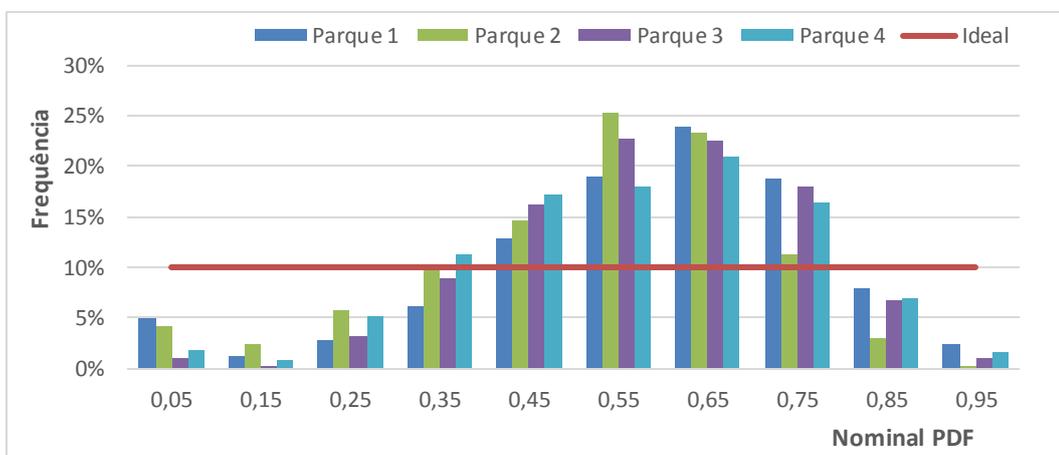
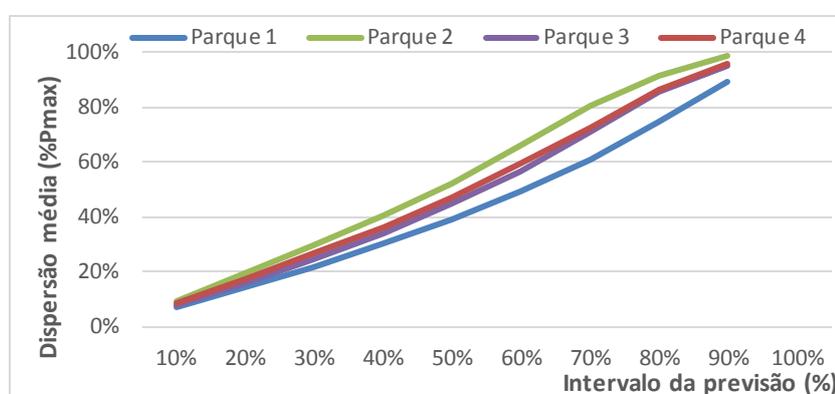


Figura 5.5 - Diagrama de reliability do teste 1 para os quatro parques

Tabela 5.3 - Desvio médio quadrático da *reliability* (Modelo Inicial)

Desvio	
Parque 1	58,89%
Parque 2	69,38%
Parque 3	73,64%
Parque 4	53,00%

Em relação ao diagrama de *reliability*, é possível observar que as frequências dos quantis centrais da função de densidade de probabilidade são muito elevadas, sendo necessária uma amplificação dos valores de alfa e beta através dos parâmetros de ajuste da incerteza.

**Figura 5.6** - Diagrama de *sharpness* do teste 1 para os quatro parques

Para o diagrama de *sharpness*, observa-se que o valor da dispersão média de todos os parques é bastante elevado para maiores intervalos de previsão, o que é típico de previsões eólicas em parques individuais. Para agregados regionais, a dispersão seria menor.

5.2 Teste 2: Entrada Recorrente

Para este teste será considerada a parametrização da entrada recorrente no modelo anterior. O número de partições do modelo autorregressivo definido pelo utilizador para a variável dependente, será definido inicialmente como 6.

Tabela 5.4 - Parametrizações do teste 2

	Parque 1	Parque 2	Parque 3	Parque 4
Número de casos históricos	12714	12410	12522	12494
Número de casos para teste	1656	1656	1656	1656
Entrada recorrente	Sim			
Número de partições	6	6	6	6
Largura de banda dinâmica	Não			

Largura de banda (Variável 1)	1,725	1,725	1,725	1,725
Largura de banda (Variável 2)	1,471	1,775	1,454	1,378
Largura de banda (Variável 3)	26,999	27,000	27,000	26,999
Largura de banda (Variável 4)	2,993	2,882	2,932	2,969
Número mínimo de pontos	----	----	----	----
Variáveis Cíclicas	Não			
Pesos	Não			
Ajuste da incerteza	Não			
Parâmetro a_α	----	---	----	----
Parâmetro a_β	----	---	----	----
Parâmetro b_α	----	---	----	----
Parâmetro b_β	----	---	----	----
Expand	----	---	----	----

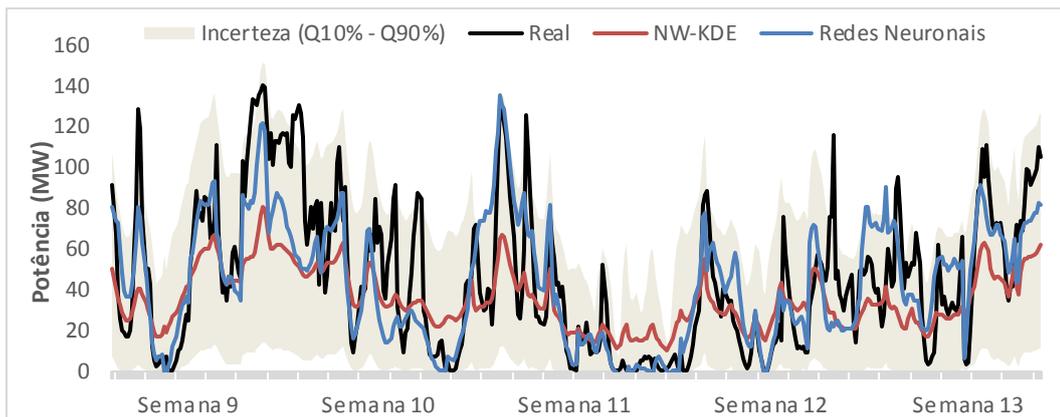


Figura 5.7 - Previsão da produção eólica e a incerteza associada do Parque 1

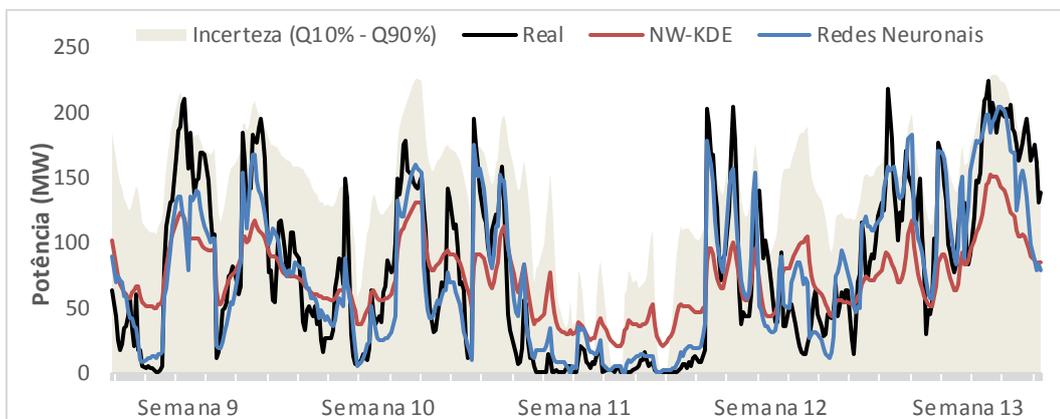


Figura 5.8 - Previsão da produção eólica e a incerteza associada do Parque 2

46 Resultados Obtidos

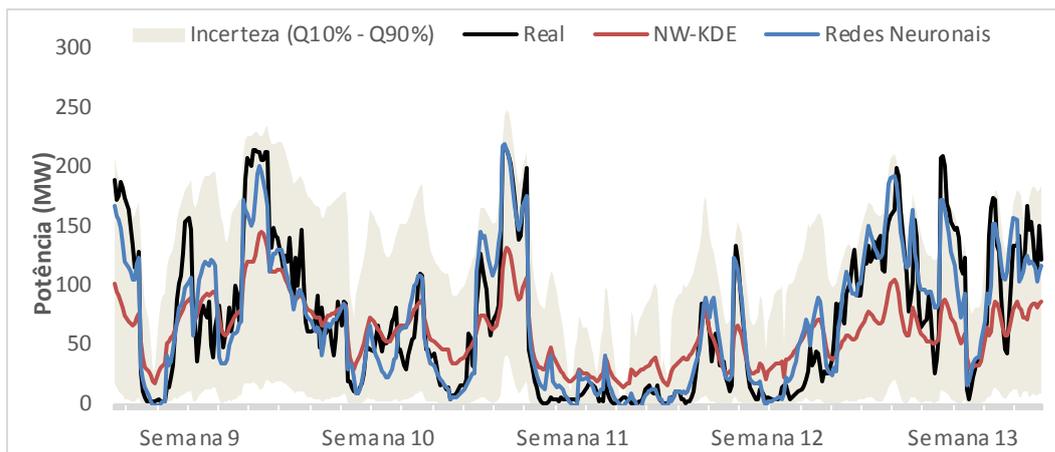


Figura 5.9 - Previsão da produção eólica e a incerteza associada do Parque 3

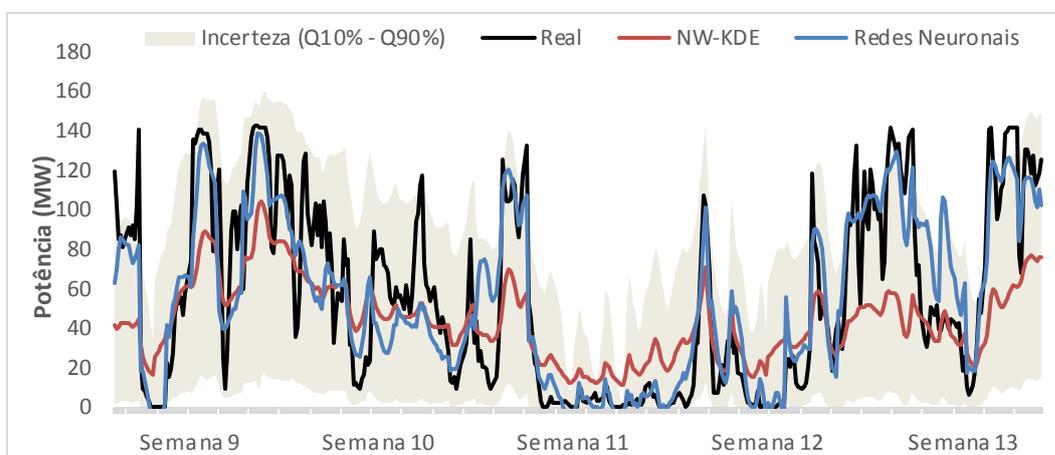


Figura 5.10 - Previsão da produção eólica e a incerteza associada do Parque 4

Tabela 5.0.5 - Erros obtidos dos valores previstos para o Teste 2

Erro	Redes Neurais	NW-KDE	Teste 1
Parque 1	31,74%	49,79%	53,17%
Parque 2	27,42%	46,79%	56,73%
Parque 3	28,34%	48,49%	52,74%
Parque 4	35,63%	48,23%	53,47%

Com esta parametrização o *NW-KDE* já consegue acompanhar melhor os picos da produção eólica, contudo o erro ainda é muito elevado. Relativamente ao teste anterior, esta parametrização consegue melhorar o erro de forma muito satisfatória, no entanto a *performance* das redes neurais é muito superior.

5.3 Teste 3: Parametrização dinâmica da largura de banda

Neste teste será considerada a parametrização dinâmica da largura de banda da função *kernel* das variáveis explicadoras ao modelo anterior. O número mínimo de pontos representativos, definido pelo utilizador, necessário para correr esta parametrização será definido como 50. No Teste 5 este valor será testado e ajustado.

Tabela 5.6 - Parametrizações do teste 3

	Parque 1	Parque 2	Parque 3	Parque 4
Número de casos históricos	12714	12410	12522	12494
Número de casos para teste	1656	1656	1656	1656
Entrada recorrente	Sim			
Número de partições	6	6	6	6
Largura de banda dinâmica	Sim			
Largura de banda (Variável 1)	0,575	0,575	2,875	0,575
Largura de banda (Variável 2)	0,490	0,592	0,485	0,722
Largura de banda (Variável 3)	23,464	18,479	22,170	26,999
Largura de banda (Variável 4)	4,988	4,004	4,886	3,843
Número mínimo de pontos	50	50	50	50
Variáveis Cíclicas	Não			
Pesos	Não			
Ajuste da incerteza	Não			
Parâmetro a_α	----	----	----	----
Parâmetro a_β	----	----	----	----
Parâmetro b_α	----	----	----	----
Parâmetro b_β	----	----	----	----
Expand	----	----	----	----

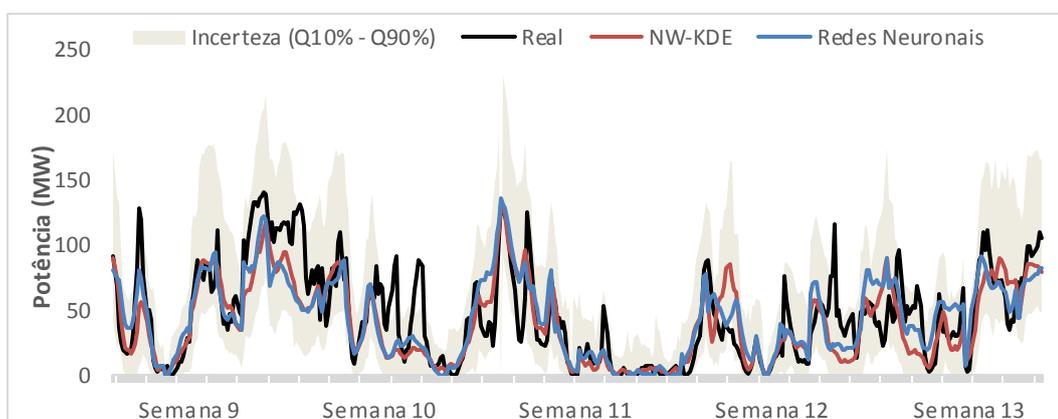


Figura 5.11 - Previsão da produção eólica e a incerteza associada do Parque 1

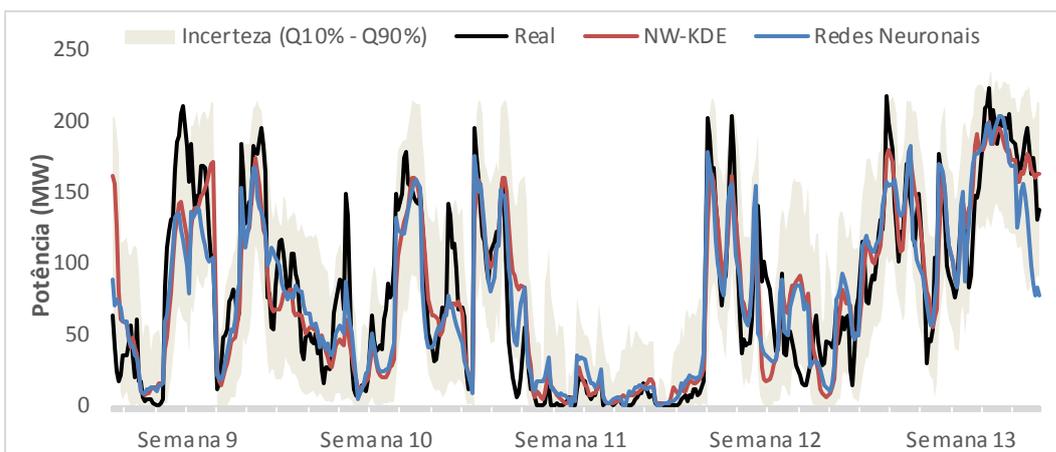


Figura 5.12 - Previsão da produção eólica e a incerteza associada do Parque 2

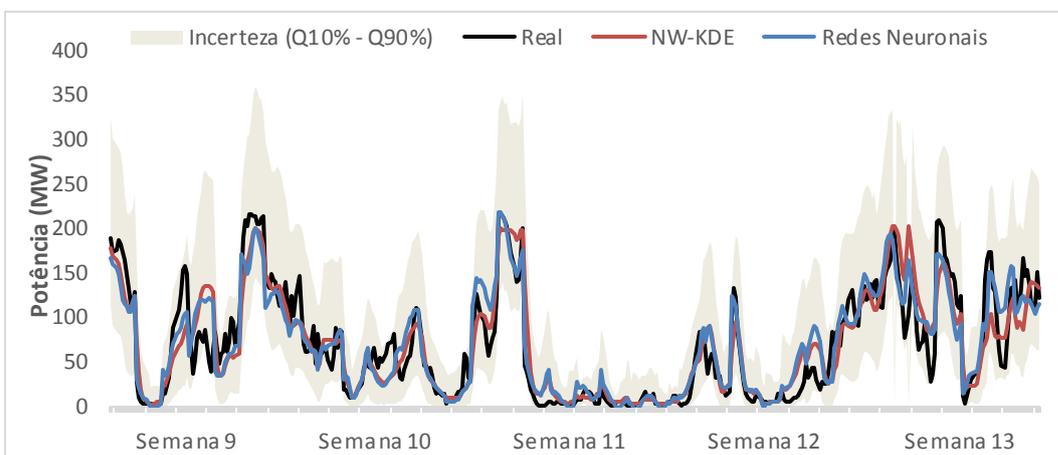


Figura 5.13 - Previsão da produção eólica e a incerteza associada do Parque 3

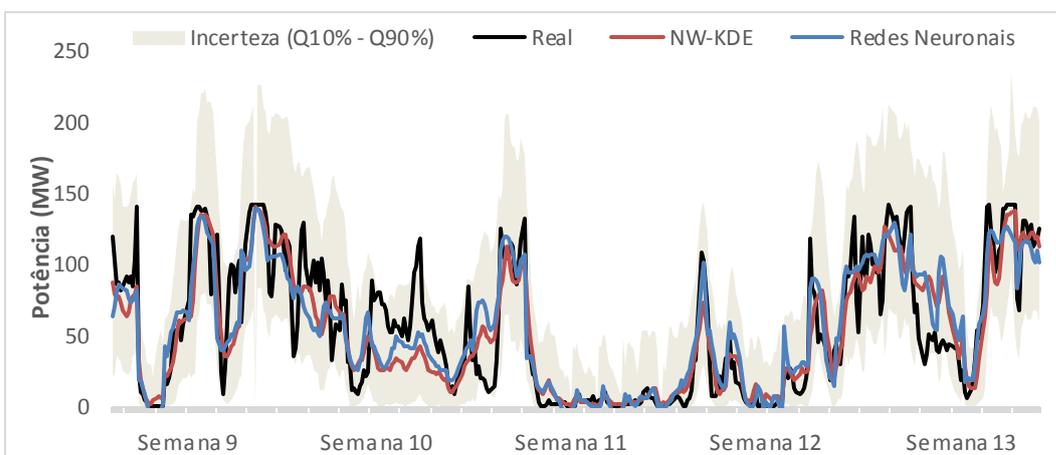


Figura 5.14 - Previsão da produção eólica e a incerteza associada do Parque 4

Tabela 5.7 - Erros obtidos dos valores previstos para o Teste 3

Erro	Redes Neurais	NW-KDE	Teste 2
Parque 1	31,74%	31,59%	49,79%
Parque 2	28,34%	28,26%	46,79%
Parque 3	27,42%	28,03%	48,49%
Parque 4	35,63%	35,60%	48,23%

O erro com a entrada desta parametrização diminui significativamente em relação ao modelo anterior. A *performance* do *NW-KDE* com as parametrizações deste modelo bate a *performance* das redes neurais em três dos quatro parques testados. Este já consegue acompanhar de forma satisfatória grande parte dos picos da produção eólica.

5.4 Teste 4: Parametrização das variáveis cíclicas

Neste novo modelo será considerado a parametrização das variáveis cíclicas. Duas das variáveis de entrada serão afetadas: hora e a direção do vento.

Tabela 5.8 - Parametrizações do teste 4

	Parque 1	Parque 2	Parque 3	Parque 4
Número de casos históricos	12714	12410	12522	12494
Número de casos para teste	1656	1656	1656	1656
Entrada recorrente	Sim			
Número de partições	6	6	6	6
Largura de banda dinâmica	Sim			
Largura de banda (Variável 1)	0,575	0,575	2,875	0,575
Largura de banda (Variável 2)	0,490	0,592	0,485	0,722
Largura de banda (Variável 3)	23,464	18,479	22,170	26,999
Largura de banda (Variável 4)	4,988	4,004	4,886	3,843
Número mínimo de pontos	50	50	50	50
Variáveis Cíclicas	Sim			
Pesos	Não			
Ajuste da incerteza	Não			
Parâmetro a_α	----	----	----	----
Parâmetro a_β	----	----	----	----
Parâmetro b_α	----	----	----	----
Parâmetro b_β	----	----	----	----
Expand	----	----	----	----

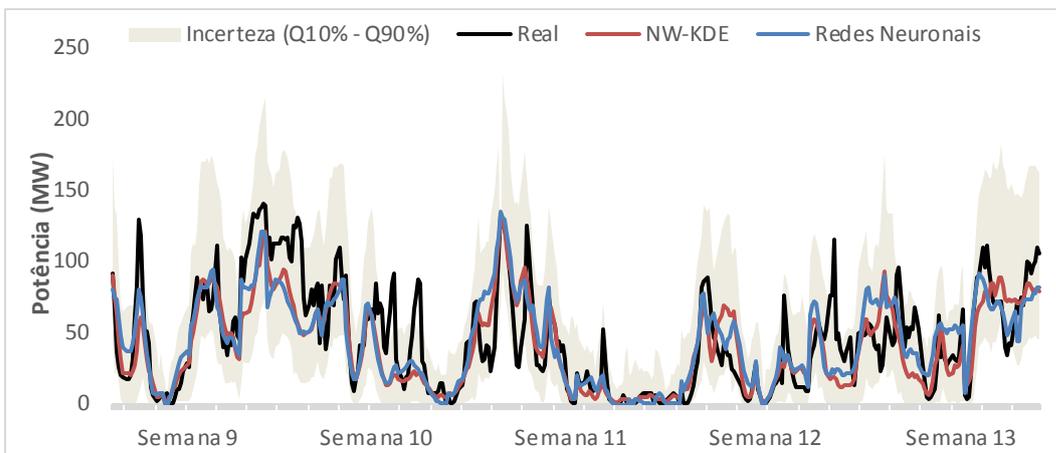


Figura 5.15 - Previsão da produção eólica e a incerteza associada do Parque 1

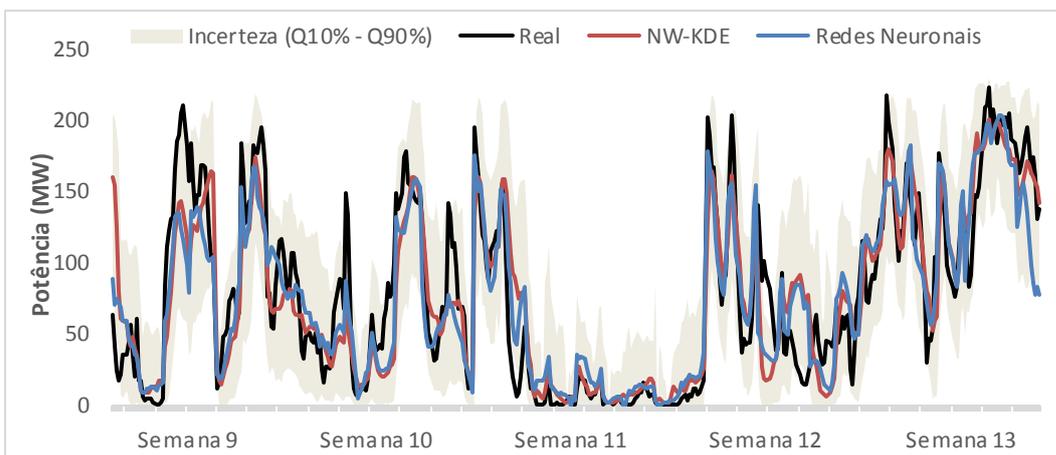


Figura 5.16 - Previsão da produção eólica e a incerteza associada do Parque 2

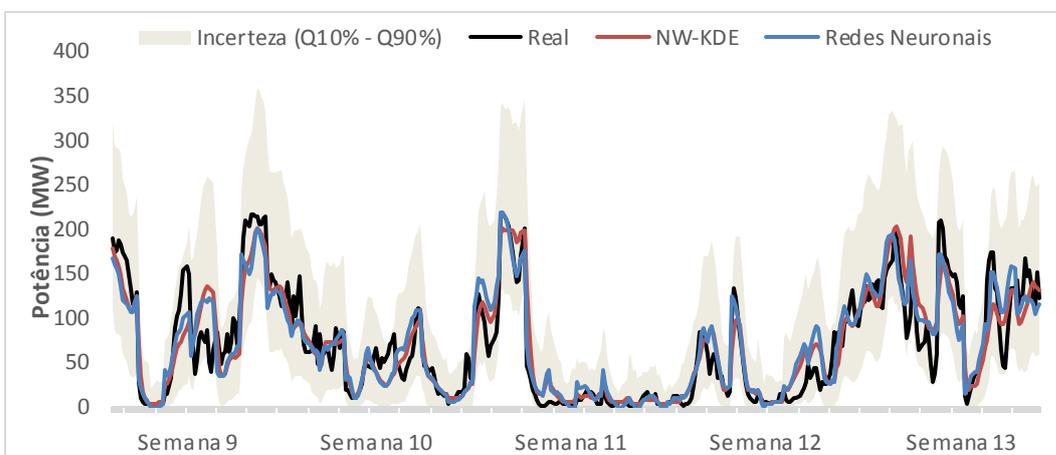


Figura 5.17 - Previsão da produção eólica e a incerteza associada do Parque 3

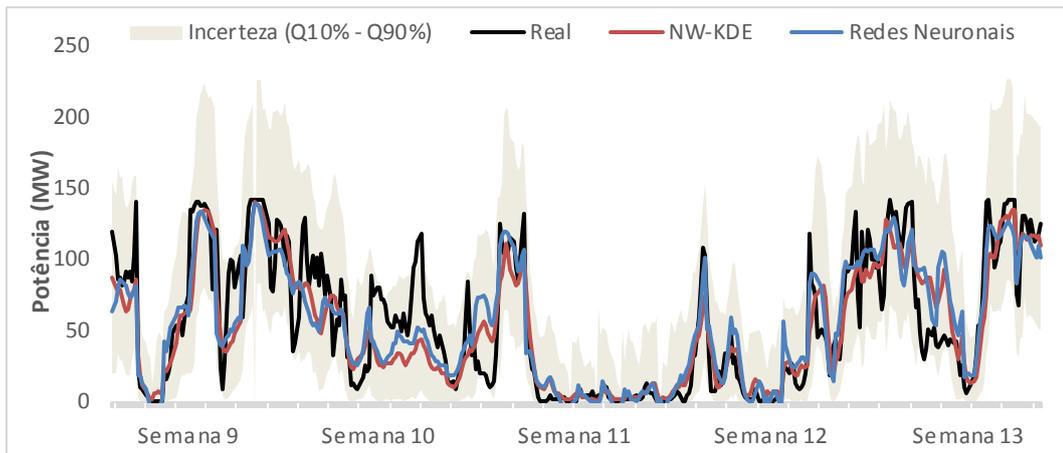


Figura 5.18 - Previsão da produção eólica e a incerteza associada do Parque 4

Tabela 5.9 - Erros obtidos dos valores previstos para o Teste 4

Erro	Redes Neurais	NW-KDE	Teste 3
Parque 1	31,74%	30,94%	31,59%
Parque 2	28,34%	28,17%	28,26%
Parque 3	27,42%	27,57%	28,03%
Parque 4	35,63%	35,60%	35,60%

Neste teste o erro diminui ligeiramente em três dos parques e mantém-se no parque 4. A previsão das redes neurais para o parque 3 continua a ter uma melhor *performance* que o *NW-KDE*.

5.5 Teste 5: Ajuste de parâmetros

Para este teste, serão ajustados a largura de banda da entrada recorrente e o número mínimo de pontos representativos definido pelo utilizador. Estes parâmetros serão testados para um conjunto de valores pré-definidos.

Tabela 5.10 - Teste ao número mínimo de pontos representativos (descrito na secção 3.4.2)

Número mínimo de pontos	50	75	100	125
Parque 1	30,94%	30,89%	30,88%	30,89%
Parque 2	28,17%	28,12%	28,10%	28,11%
Parque 3	27,57%	27,56%	27,54%	27,53%
Parque 4	35,72%	35,64%	35,60%	35,55%

Tabela 5.11 - Teste ao número de partições para o ajuste da largura de banda da entrada recorrente (descrito na seção 3.3)

Número de partições	3	4	5	6
Parque 1	30,17%	30,12%	30,47%	30,88%
Parque 2	27,44%	27,71%	27,90%	28,10%
Parque 3	27,10%	27,18%	27,35%	27,53%
Parque 4	33,96%	34,58%	35,11%	35,55%

Para a previsão determinística de cada parque, serão utilizados os parâmetros que registaram os melhores resultados.

Tabela 5.12 - Parametrizações do teste 5

	Parque 1	Parque 2	Parque 3	Parque 4
Número de casos históricos	12714	12410	12522	12494
Número de casos para teste	1656	1656	1656	1656
Entrada recorrente	Sim			
Número de partições	4	3	3	3
Largura de banda dinâmica	Sim			
Largura de banda (Variável 1)	0,575	0,575	2,875	0,575
Largura de banda (Variável 2)	0,490	0,592	0,485	0,722
Largura de banda (Variável 3)	23,464	18,479	22,170	26,999
Largura de banda (Variável 4)	4,988	4,004	4,886	3,843
Número mínimo de pontos	100	100	125	125
Variáveis Cíclicas	Sim			
Pesos	Não			
Ajuste da incerteza	Não			
Parâmetro a_α	----	----	----	----
Parâmetro a_β	----	----	----	----
Parâmetro b_α	----	----	----	----
Parâmetro b_β	----	----	----	----
Expand	----	----	----	----

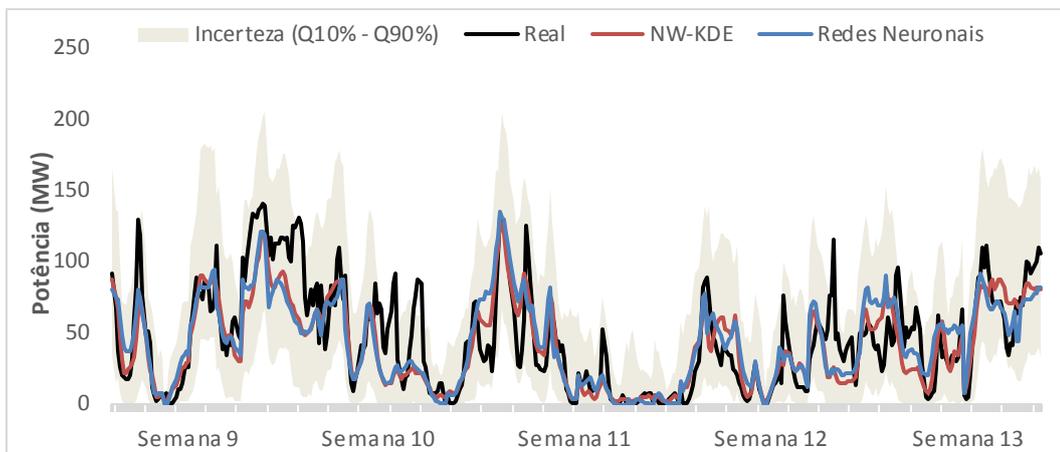


Figura 5.19 - Previsão da produção eólica e a incerteza associada do Parque 1

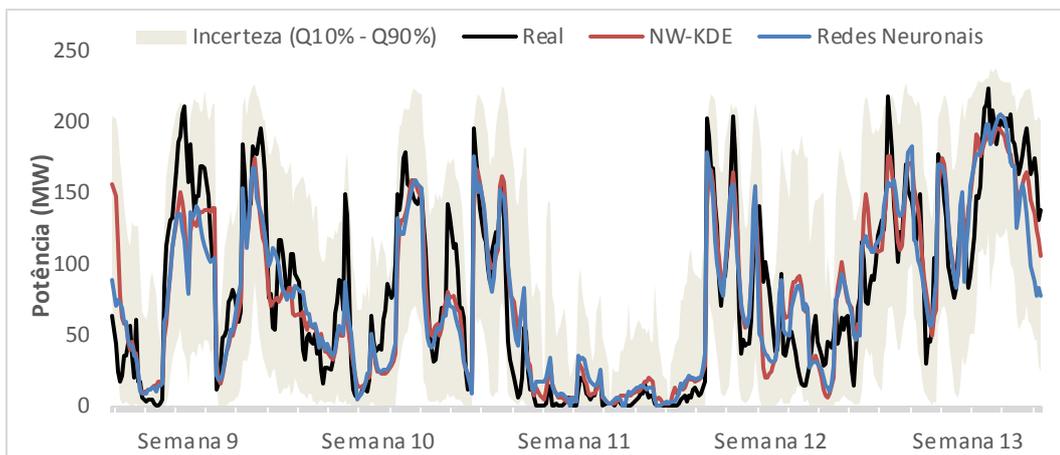


Figura 5.20 - Previsão da produção eólica e a incerteza associada do Parque 2

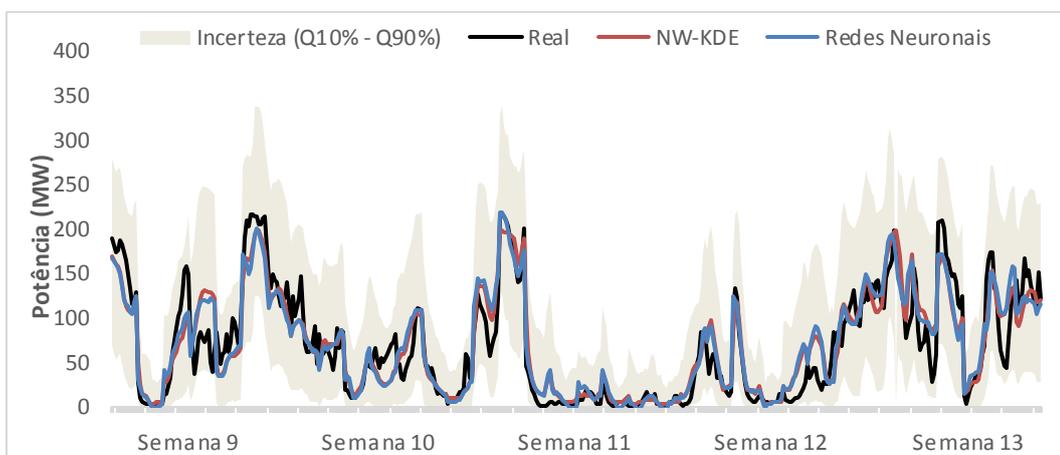


Figura 5.21 - Previsão da produção eólica e a incerteza associada do Parque 3

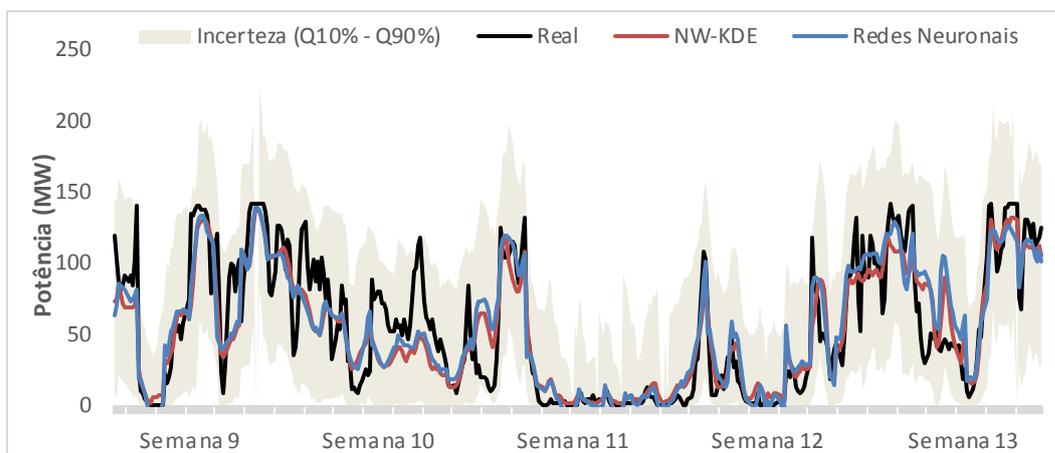


Figura 5.22 - Previsão da produção eólica e a incerteza associada do Parque 4

Tabela 5.13 - Erros obtidos dos valores previstos para o Teste 5

Erro	Redes Neurais	NW-KDE	Teste 4
Parque 1	31,74%	30,12%	30,94%
Parque 2	28,34%	27,44%	28,17%
Parque 3	27,42%	27,10%	27,57%
Parque 4	35,63%	33,96%	35,72%

Com o ajuste do número mínimo de pontos representativos e a largura de banda da entrada recorrente o erro do *NW-KDE* é reduzido e consegue bater em todos os parques a *performance* das redes neurais.

5.6 Teste 6: Parametrização do peso de cada caso histórico

Para este modelo será incluída a parametrização do peso de cada caso histórico ao modelo anterior.

Tabela 5.14 - Parametrizações do teste 6

	Parque 1	Parque 2	Parque 3	Parque 4
Número de casos históricos	12714	12410	12522	12494
Número de casos para teste	1656	1656	1656	1656
Entrada recorrente	Sim			
Número de partições	4	3	3	3
Largura de banda dinâmica	Sim			

Largura de banda (Variável 1)	0,575	0,575	2,875	0,575
Largura de banda (Variável 2)	0,490	0,592	0,485	0,722
Largura de banda (Variável 3)	23,464	18,479	22,170	26,999
Largura de banda (Variável 4)	4,988	4,004	4,886	3,843
Número mínimo de pontos	100	100	125	125
Variáveis Cíclicas	Sim			
Pesos	Sim			
Ajuste da incerteza	Não			
Parâmetro a_α	----	----	----	----
Parâmetro a_β	----	----	----	----
Parâmetro b_α	----	----	----	----
Parâmetro b_β	----	----	----	----
Expand	----	----	----	----

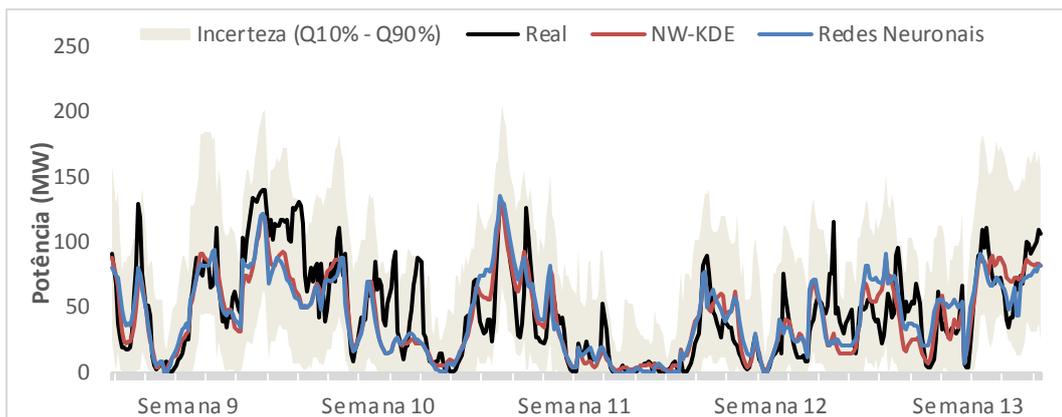


Figura 5.23 - Previsão da produção eólica e a incerteza associada do Parque 1

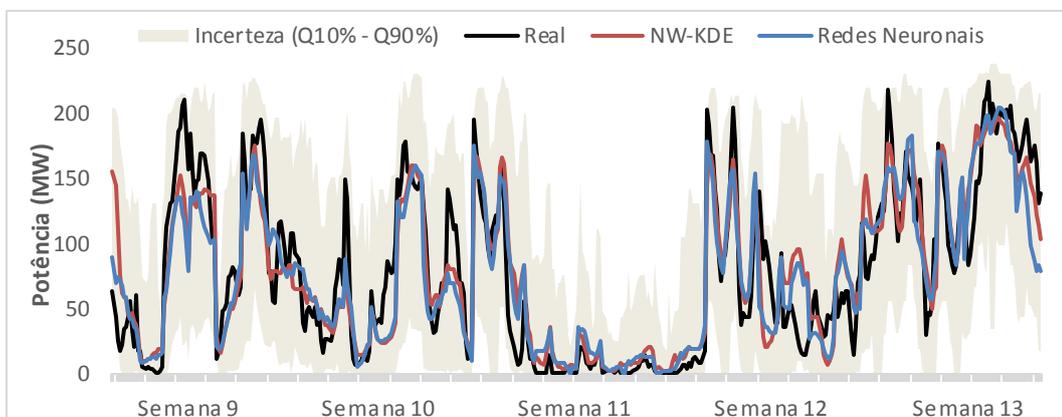


Figura 5.24 - Previsão da produção eólica e a incerteza associada do Parque 2

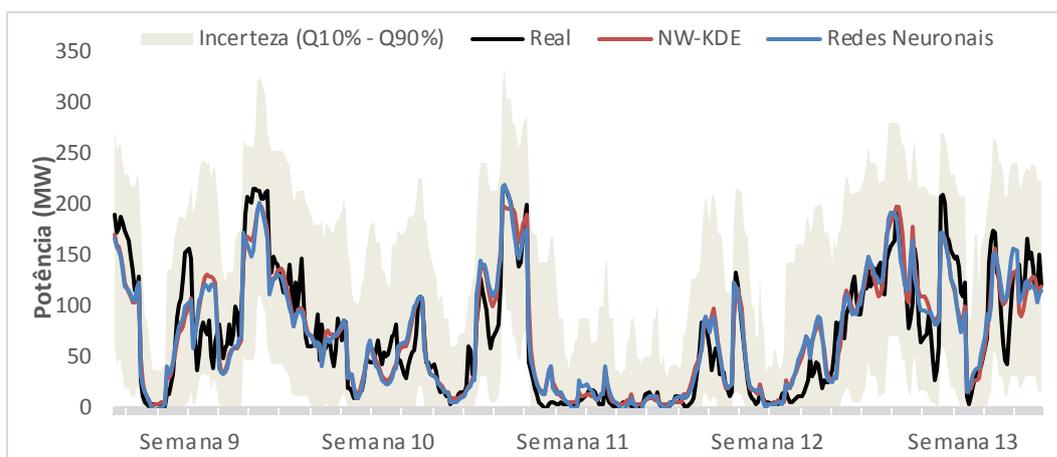


Figura 5.25 - Previsão da produção eólica e a incerteza associada do Parque 3

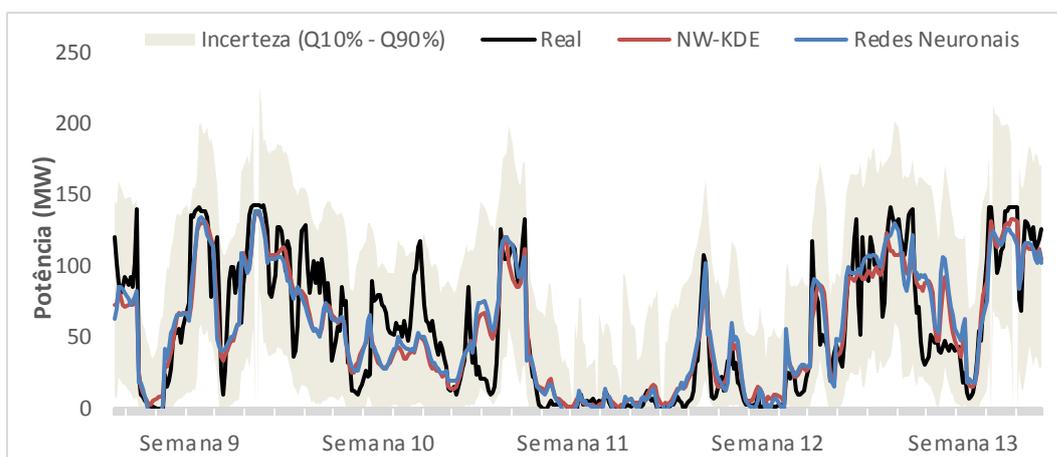


Figura 5.26 - Previsão da produção eólica e a incerteza associada do Parque 4

Tabela 5.15 - Erros obtidos dos valores previstos para o Teste 6

Erro	Redes Neurais	NW-KDE	Teste 5
Parque 1	31,74%	29,95%	30,12%
Parque 2	28,34%	27,76%	27,44%
Parque 3	27,42%	27,26%	27,10%
Parque 4	35,63%	34,19%	33,96%

Apesar da lógica inerente de dar mais peso aos casos históricos com menor erro, na prática esta parametrização agrava o erro em três dos quatro parques. Esta parametrização não será incluída no modelo final.

5.7 Teste 7: Parametrização para ajuste da incerteza

Para este último teste será ajustada a incerteza da previsão. Neste modelo, não será considerada a parametrização do peso de cada caso histórico. A previsão determinística não será analisada dado que esta parametrização não influencia os valores esperados da previsão.

Foram otimizados os parâmetros para o ajuste da incerteza através do teste dos valores de um conjunto de validação.

Tabela 5.16 - Resultado dos valores ótimos dos parâmetros da incerteza para o teste 7

Incógnitas	a_α	b_α	a_β	b_β
Parque 1	1,9	0,1	1,7	0,1
Parque 2	1,7	0,1	1,9	0,1
Parque 3	1,6	0,1	1,7	0,1
Parque 4	1,7	0,1	1,9	0,1

Tabela 5.17 - Parametrizações do teste 7

	Parque 1	Parque 2	Parque 3	Parque 4
Número de casos históricos	12714	12410	12522	12494
Número de casos para teste	1656	1656	1656	1656
Entrada recorrente	Sim			
Número de partições	4	3	3	3
Largura de banda dinâmica	Sim			
Largura de banda (Variável 1)	0,575	0,575	2,875	0,575
Largura de banda (Variável 2)	0,490	0,592	0,485	0,722
Largura de banda (Variável 3)	23,464	18,479	22,170	26,999
Largura de banda (Variável 4)	4,988	4,004	4,886	3,843
Número mínimo de pontos	100	100	125	125
Variáveis Cíclicas	Sim			
Pesos	Não			
Ajuste da incerteza	Sim			
Parâmetro a_α	1,9	1,9	1,6	1,7
Parâmetro a_β	1,7	1,9	1,9	1,9
Parâmetro b_α	0,1	0,1	0,1	0,1
Parâmetro b_β	0,1	0,1	0,1	0,1
Expand	1	1	1	1

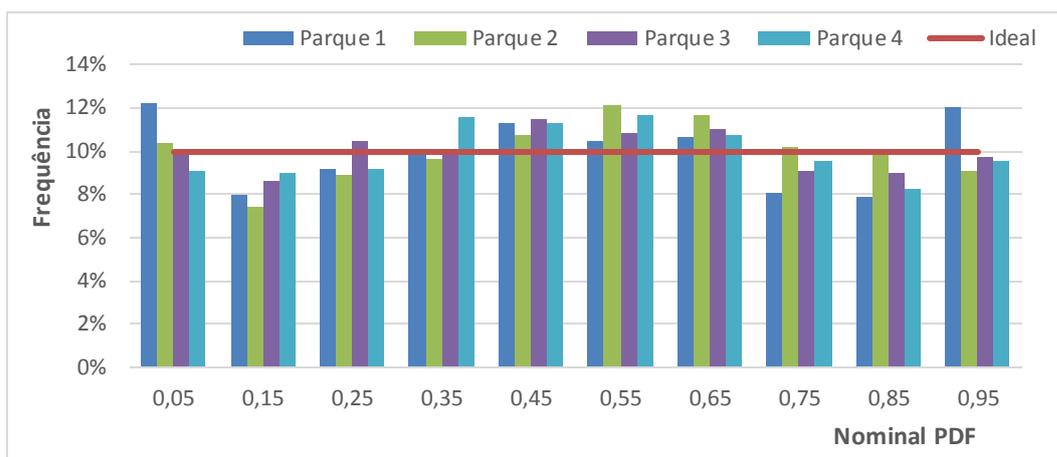


Figura 5.27 - Diagrama de *reliability* do teste 7 para os quatro parques

Tabela 5.18 - Desvio médio quadrático da *reliability* (teste 7)

Erro	Incerteza	Teste 5
Parque 1	2,36%	9,51%
Parque 2	1,66%	11,91%
Parque 3	0,81%	10,70%
Parque 4	1,33%	7,45%

Com esta parametrização, o desvio médio quadrático da *reliability* é reduzido em todos os parques testados. As frequências da função de densidade de probabilidade estão muito próximas das frequências ideais em todos os intervalos de probabilidade.

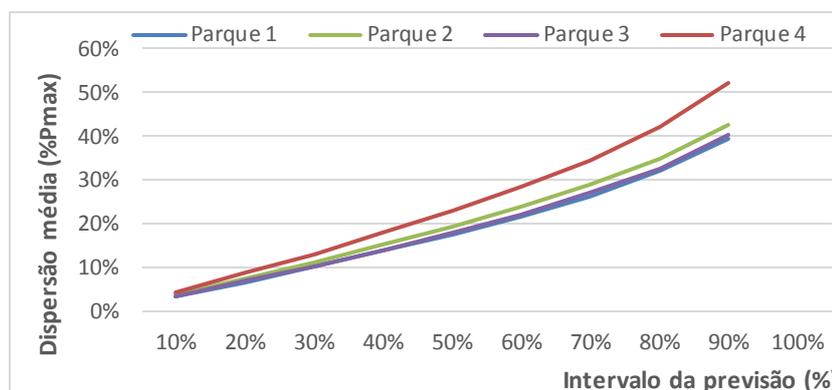


Figura 5.28 - Diagrama de *sharpness* do teste 5 para os quatro parques

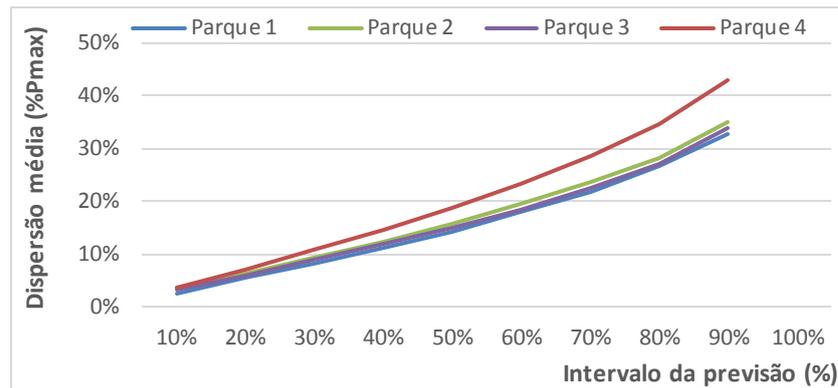


Figura 5.29 - Diagrama de *sharpness* do teste 7 para os quatro parques

Pelo diagrama de *sharpness* do teste 7, é possível observar que a dispersão média dos valores da incerteza é reduzida em relação ao diagrama de *sharpness* do teste 5.

5.8 Teste 8: Modelo Final

Finalmente, neste último modelo farão parte todas as parametrizações que melhoraram os resultados observados, deste modo, apenas a parametrização do peso não será incluída. Este modelo será comparado com o modelo inicial.

Tabela 5.19 - Parametrizações do teste 8

	Parque 1	Parque 2	Parque 3	Parque 4
Número de casos históricos	12714	12410	12522	12494
Número de casos para teste	1656	1656	1656	1656
Entrada recorrente	Sim			
Número de partições	4	3	3	3
Largura de banda dinâmica	Sim			
Largura de banda (Variável 1)	0,575	0,575	2,875	0,575
Largura de banda (Variável 2)	0,490	0,592	0,485	0,722
Largura de banda (Variável 3)	23,464	18,479	22,170	26,999
Largura de banda (Variável 4)	4,988	4,004	4,886	3,843
Número mínimo de pontos	100	100	125	125
Variáveis Cíclicas	Sim			
Pesos	Não			
Ajuste da incerteza	Sim			
Parâmetro a_α	1,9	1,9	1,6	1,7

60 Resultados Obtidos

Parâmetro a_β	1,7	1,9	1,9	1,9
Parâmetro b_α	0,1	0,1	0,1	0,1
Parâmetro b_β	0,1	0,1	0,1	0,1
Expand	1	1	1	1

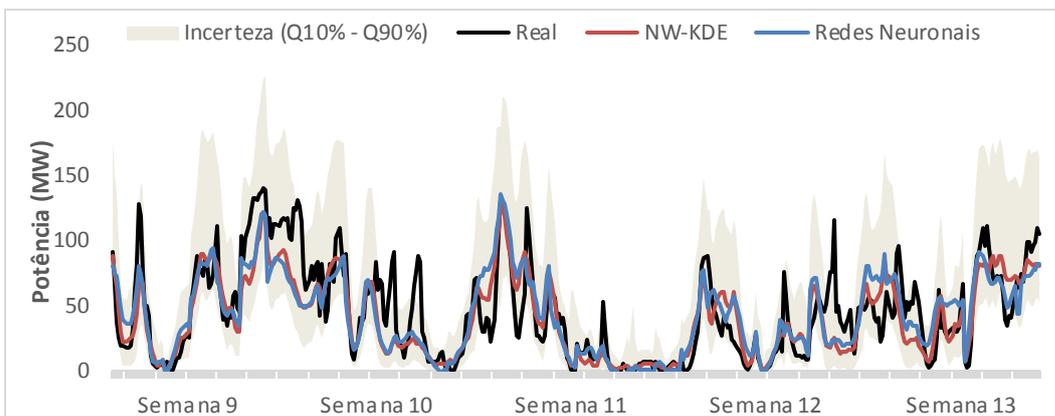


Figura 5.30 - Previsão da produção eólica e a incerteza associada do Parque 1

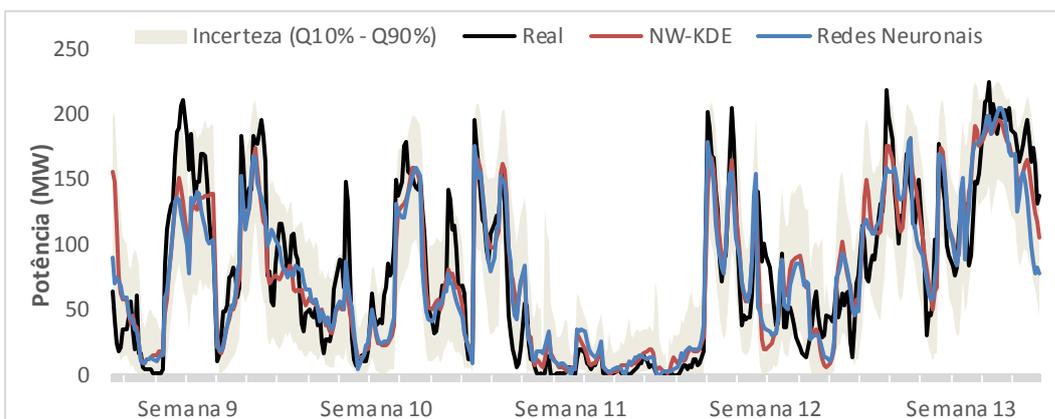


Figura 5.31 - Previsão da produção eólica e a incerteza associada do Parque 2

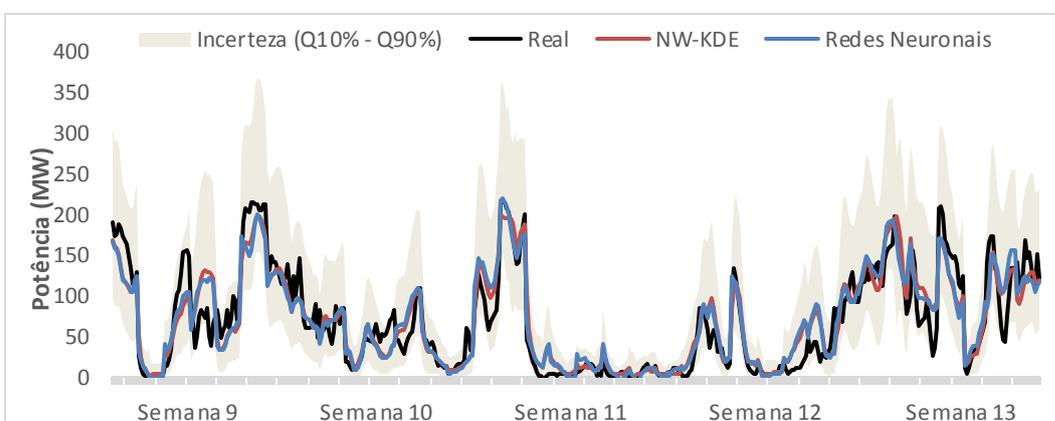


Figura 5.32 - Previsão da produção eólica e a incerteza associada do Parque 3

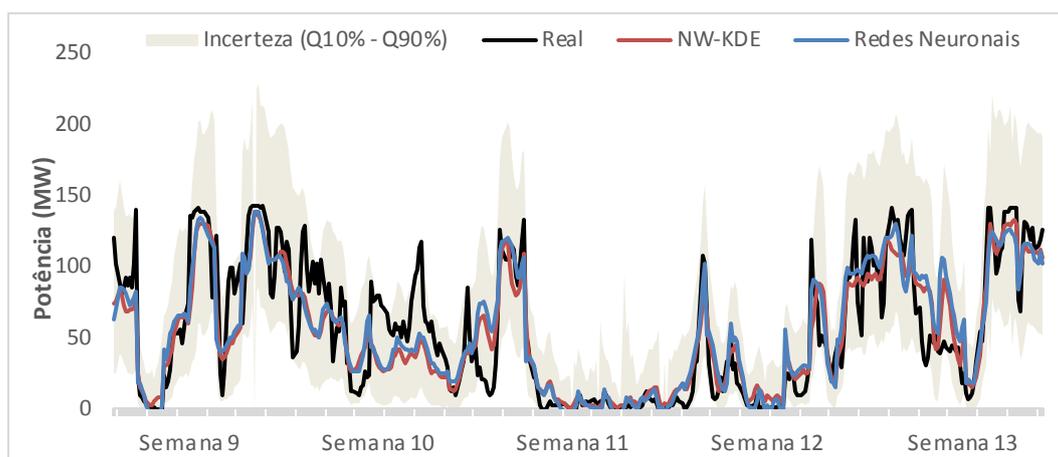


Figura 5.33 - Previsão da produção eólica e a incerteza associada do Parque 4

Tabela 5.20 - Erros obtidos dos valores previstos do Modelo Final

Erro	Redes Neurais	NW-KDE	Modelo Inicial
Parque 1	31,74%	30,12%	53,17%
Parque 2	28,34%	27,44%	56,73%
Parque 3	27,42%	27,10%	52,74%
Parque 4	35,63%	33,96%	53,47%

A *performance* do NW-KDE com as parametrizações criadas é muito superior à *performance* do modelo inicial. É possível observar que em todos os parques o erro das redes neurais é superado, mostrando que o modelo desenvolvido consegue modelizar incertezas e por outro lado, obter desempenhos da previsão do valor esperado superiores às melhores técnicas de previsão existentes atualmente no estado da arte.

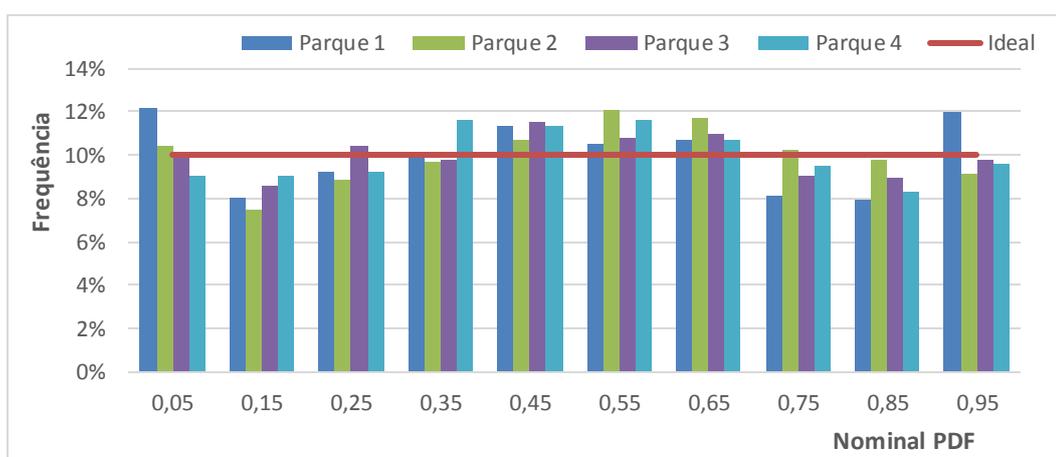
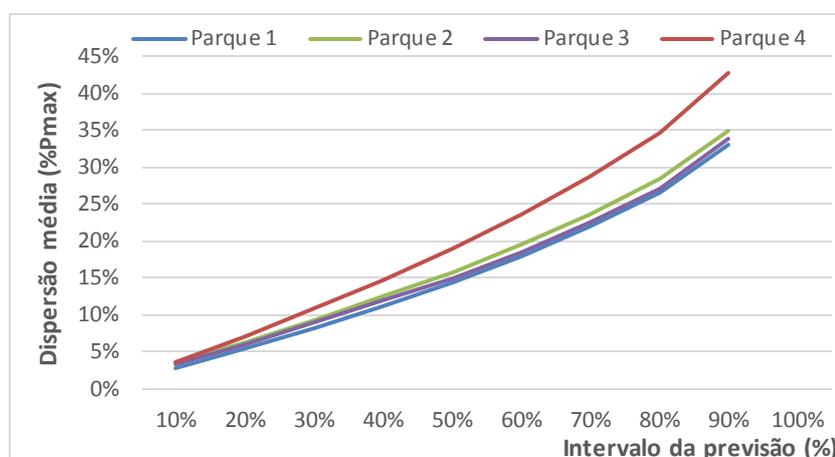


Figura 5.34 - Diagrama de *reliability* do teste 8 para os quatro parques

Tabela 5.21 - Desvio médio quadrático da *reliability* (Modelo Final)

Erro	Incerteza	Modelo Inicial
Parque 1	2,36%	58,89%
Parque 2	1,66%	69,38%
Parque 3	0,81%	73,64%
Parque 4	1,33%	53,00%

O desvio médio quadrático da *reliability* é reduzido de forma substancial em relação ao modelo inicial.

Figura 5.35 - Diagrama de *sharpness* do Modelo Final para os quatro parques

Relativamente ao modelo inicial representado na Figura 5.6, a dispersão média do modelo final é reduzida para mais de metade em todos os intervalos de previsão. Observa-se que o parque 4, pelos indicadores de incerteza e de erro, é mais difícil de prever que os restantes, provavelmente, devido à localização em zonas de maior complexidade orográfica.

5.9 Análise dos resultados obtidos

Tabela 5.22 - Resultados do erro da previsão determinística de todos os testes

	Modelo Inicial	Teste 2	Teste 3	Teste 4	Teste 5	Teste 6	Modelo Final
Parque 1	53,17%	-3,39%	-18,19%	-0,65%	-0,82%	-0,17%	30,12%
Parque 2	56,73%	-9,94%	-18,53%	-0,09%	-0,73%	0,32%	27,44%
Parque 3	52,74%	-4,25%	-20,46%	-0,46%	-0,47%	0,16%	27,10%
Parque 4	53,47%	-5,24%	-12,63%	0,12%	-1,76%	0,23%	33,96%
Média	-----	-5,71%	-17,45%	-0,27%	-0,95%	0,13%	-----

De todas as parametrizações testadas, a utilizada para definir a largura de banda ótima da função *kernel* das variáveis explicadoras, é a mais importante. A *performance* desta é superior a todas as outras parametrizações juntas.

A parametrização da entrada recorrente revela-se importante para o modelo de previsão, evidenciando uma descida assinalável do erro.

Destaca-se ainda, a importância dos ajustes dos parâmetros da largura de banda da entrada recorrente e o número mínimo de pontos representativos definido pelo utilizador, melhorando ligeiramente a *performance* do modelo.

Por outro lado, não se revelam muitos importantes para o modelo de previsão, a parametrização das variáveis cíclicas e a do peso de cada caso histórico. A primeira parametrização melhora ligeiramente o erro de previsão em três dos quatro parques. A última, por outro lado, piora globalmente o erro e nem faz parte do modelo final.

Em relação à parametrização de ajuste da incerteza, através da

Tabela 5.18 é possível observar o efeito positivo desta na consistência estatística da distribuição prevista. Analisando as Figura 5.28 e Figura 5.29, observa-se o efeito positivo na dispersão média dos valores da incerteza para todos os intervalos de probabilidade.

Tabela 5.23 - Resultados do valor dos parâmetros definidos do Modelo Final

	Parque 1	Parque 2	Parque 3	Parque 4	Notas
Número de partições	4	3	3	3	(1)
Número mínimo de pontos	100	100	125	125	(2)
Parâmetro a_α	1,9	1,9	1,6	1,7	(3)
Parâmetro a_β	1,7	1,9	1,9	1,9	
Parâmetro b_α	0,1	0,1	0,1	0,1	(4)
Parâmetro b_β	0,1	0,1	0,1	0,1	
Expand	1	1	1	1	(5)

Cada um dos parâmetros do modelo final foram testados para um conjunto de valores diferentes e registados aqueles que permitiram ao NW-KDE obter a melhor *performance*. Porém, observando a Tabela 5.23, é possível verificar a similaridade dos valores de cada parâmetro para todos os parques. Assim para a previsão de produção eólica através do NW-KDE, podemos retirar conclusões quanto à ordem de grandeza dos parâmetros a utilizar. No entanto, é de salientar que os parâmetros devem ser otimizados para cada um dos casos de previsão. As conclusões sobre o estudo de parametrização são as seguintes:

- (1) O número de partições da entrada recorrente deverá ser 3 ou 4. A escolha de um valor não ótimo para este parâmetro aumenta o erro da previsão de forma significativa.
- (2) O número mínimo de pontos da largura de banda deverá ser entre 100 e 125. Da mesma forma que o parâmetro anterior, uma ligeira diferença nos valores ótimos pode aumentar significativamente o erro da previsão.
- (3) O valor dos parâmetros a_α e a_β deverá ficar compreendido entre 1,6 e 1,9. Valores muito elevados destes parâmetros não produzem resultados aceitáveis.

64 Resultados Obtidos

- (4) O valor dos parâmetros b_α e b_β deverá ser 0,1. Este valor nunca poderá ser muito elevado devido ao efeito negativo que tem no desvio médio quadrático da *reliability*;
- (5) O *expand* não produziu melhorias no ajuste da incerteza e o seu valor foi fixado em 1, não tendo assim, nenhuma influência no processo de otimização.

Capítulo 6

Conclusão e Trabalhos Futuros

O objetivo deste capítulo é dar a perceber a conclusão do trabalho realizado e verificar se os objetivos propostos foram cumpridos. São ainda apresentados possíveis trabalhos futuros que poderão dar continuidade ao trabalho realizado.

6.1 Conclusão

A dissertação efetuada teve como objetivo principal o estudo de modelos de previsão dos valores esperados e da incerteza da produção eólica, bem como a criação de otimizações para um modelo já existente de forma a melhorar a sua *performance* final.

Inicialmente, foi feito o estudo de várias heurísticas para a otimização de modelos de previsão probabilística de produção eólica com base em diversas técnicas de inteligência computacional.

O modelo de previsão escolhido baseou-se num estimador de densidade condicional não-paramétrica baseado em estimadores de densidade condicional não-paramétrica denominados *Nadaraya-Watson Estimator* e *Kernel Density Estimation*. Este modelo utiliza um conjunto extenso de dados históricos como base de dados para o processo de previsão dos valores esperados e da incerteza associada a estes. O histórico da previsão contém variáveis meteorológicas e valores de produção eólica reais.

De acordo com o que foi definido nos objetivos, foram acrescentadas novas otimizações ao modelo de previsão probabilística utilizado, acelerado o tempo de processamento deste e aumentado o número de variáveis de entrada que o modelo pode suportar. Os parâmetros deste modelo foram ainda otimizados e caracterizados. A diminuição do erro de previsão do valor esperado e o ajuste da incerteza foram os principais focos do trabalho realizado.

Procedeu-se à análise do efeito de cada parametrização criada e da sua influência no modelo inicial. De todas as parametrizações criadas, a parametrização da largura de banda da função *kernel* é a que tem mais impacto na melhoria no modelo. Também é possível concluir, que uma descuidada parametrização do modelo NW-KDE poderá resultar em desempenhos desfavoráveis e, portanto, é essencial uma boa parametrização para cada caso de previsão, ou caso não seja possível, usar parametrizações típicas adequadas ao tipo de problema de

previsão. Nesta dissertação foram apresentadas as parametrizações típicas a usar em modelos de previsão de produção eólica de parques individuais.

Para o ajuste da incerteza, foram criadas formas originais de parametrização que ajustam de forma global os parâmetros da distribuição Beta. Esta parametrização também deve ser realizada individualmente para cada parque eólico, mas nesta dissertação conseguiu-se retirar conclusões sobre valores indicativos a usar na parametrização. A metodologia de parametrização de incerteza criada não interfere com o desempenho da previsão de valor esperado do modelo, o que é uma inovadora mais-valia do NW-KDE. A não utilização desta parametrização poderá implicar modelizações pouco adequadas da incerteza, em que o modelo apresenta naturalmente valores demasiado apertados de incerteza, não correspondendo a uma representação fidedigna.

No geral, o novo modelo NW-KDE relativamente à sua variante original KDE, inclui a introdução de um modelo recorrente, variáveis cíclicas, pesos, parametrização dinâmica e parametrização de incerteza. Adicionalmente foi otimizada a algoritmia do modelo acelerando o processo de cálculo e tornando-o aplicável a problemas com elevado número de variáveis. Quanto ao desempenho é possível concluir que se consegue melhores desempenhos na previsão do valor esperado e melhores modelizações de previsão probabilística. Tendo em conta estes resultados, pode afirmar-se que foram largamente ultrapassadas as expectativas tendo-se ultrapassado os desempenhos do estado da arte geral associados a este tema.

6.2 Trabalhos futuros

Após a conclusão da dissertação verifica-se que existem ainda áreas que poderão ser mais desenvolvidas e eventuais imprecisões que poderão ser corrigidas.

A maior parte dos testes realizados produziu bons resultados. Porém as otimizações implementadas na ferramenta de previsão poderão ser sempre aperfeiçoadas.

Seria interessante perceber a possibilidade de desenvolver mais ainda a parametrização da largura de banda da função *kernel* das variáveis explicadoras, já que esta apresentou bons resultados para o problema da previsão eólica apenas com 4 variáveis. No entanto, para problemas de previsão de 20 variáveis, a otimização torna-se complexa e morosa sendo necessário o desenvolvimento de processos de aceleração.

Em relação às variáveis de entrada do modelo de previsão para esta dissertação, as utilizadas foram as variáveis disponibilizadas. No entanto, existe a possibilidade da inclusão de variáveis diferentes e importantes que poderão trazer maior fiabilidade, ao modelo de previsão de produção eólica.

Há ainda, a possibilidade de ser feito um estudo com o objetivo de criar indicadores de avaliação de desempenho da previsão probabilística para uma melhor avaliação da incerteza da previsão eólica.

No que concerne à ferramenta computacional, o processo iterativo de otimização e cálculo dos valores esperados e da incerteza da previsão eólica é um processo moroso, pelo que seria aliciente criar novas heurísticas com o objetivo de aumentar a velocidade dos processos.

Finalmente, relativamente à tipificação das parametrizações, seria necessário o mesmo tipo de trabalho realizado para diferentes aplicações de previsão (consumos, solar, hídrica, preço, etc.) e para diferentes variantes estruturais destes modelos.

Referências

- [1] W.-Y. Chang, "An RBF Neural Network Combined with OLS Algorithm and Genetic Algorithm for Short-Term Wind Power Forecasting," *Journal of Applied Mathematics*, vol. 2013, p. 9, 2013.
- [2] C. Peiyuan, T. Pedersen, B. Bak-Jensen, and C. Zhe, "ARIMA-Based Time Series Model of Stochastic Wind Power Generation," *Power Systems, IEEE Transactions on*, vol. 25, pp. 667-676, 2010.
- [3] J. Jung and R. P. Broadwater, "Current status and future advances for wind speed and power forecasting," *Renewable and Sustainable Energy Reviews*, vol. 31, pp. 762-777, 3// 2014.
- [4] C. Monteiro, R. Bessa, V. Miranda, A. Botterud, J. Wang, G. Conzelmann, *et al.*, "Wind power forecasting : state-of-the-art 2009," ANL/DIS-10-1; TRN: US200924%351 United States10.2172/968212TRN: US200924%351Thu Dec 31 07:14:26 EST 2009ANLEGLISH, 2009.
- [5] J. Mendes, R. J. Bessa, H. Keko, J. Sumaili, V. Miranda, C. Ferreira, *et al.*, "Development and testing of improved statistical wind power forecasting methods; 2011.
- [6] Y. Zhang, J. Wang, and X. Wang, "Review on probabilistic forecasting of wind power generation," *Renewable and Sustainable Energy Reviews*, vol. 32, pp. 255-270, 4// 2014.
- [7] B. Santos, C. Monteiro, C. Soares, J. Catalão, G. Osório, J. Rojas, *et al.*, "Report on the mathematical formulation of the different forecasting models and data analysis," Singular2012.
- [8] G. Giebel, R. Brownsword, G. Kariniotakis, M. Denhard, and C. Draxl, "The State-Of-The-Art in Short-Term Prediction of Wind Power A Literature Overview, 2nd Edition," Project report for the Anemos.plus and SafeWind projects2011.
- [9] C. Monteiro, L. A. Fernandez-Jimenez, I. J. Ramirez-Rosado, A. Muñoz-Jimenez, and P. M. Lara-Santillan, "Short-Term Forecasting Models for Photovoltaic Plants: Analytical versus Soft-Computing Techniques," *Mathematical Problems in Engineering*, vol. 2013, p. 9, 2013.
- [10] W. Glassley, J. Kleissl, C. P. C. v. Dam, H. Shiu, J. Huang, G. Braun, *et al.*, "Appendix B - Wind Energy Forecasting: A Review of State-of-the Art and Recommendations for Better Forecasts," California Renewable Energy Forecasting, Resource Data and Mapping, 2010.
- [11] W. Chang, "A Literature Review of Wind Forecasting Methods," *Journal of Power and Energy Engineering*, vol. 2, pp. 161-168, (2014).
- [12] A. Tascikaraoglu and M. Uzunoglu, "A review of combined approaches for prediction of short-term wind speed and power," *Renewable and Sustainable Energy Reviews*, vol. 34, pp. 243-254, 6// 2014.

[13]H. Madsen, P. Pinson, G. Kariniotakis, H. A. Nielsen, and T. Nielsen, "Standardizing the Performance Evaluation of ShortTerm Wind Power Prediction Models," *Wind Engineering*, vol. 29, pp. 475-489, 12/01/ 2005.

[14]P. Pinson and G. Kariniotakis, "On-line assessment of prediction risk for wind power production forecasts," *Wind Energy*, vol. 7, pp. 119-132, 2004.

[15]J. Juban, L. Fugon, and G. Kariniotakis, "Uncertainty estimation of wind power forecasts: Comparison of Probabilistic Modelling Approaches," presented at the European Wind Energy Conference & Exhibition EWEC 2008, Bruxelles, Bélgica, 2008.

[16]H. A. Nielsen, T. S. Nielsen, H. Madsen, G. Giebel, J. Badger, L. Landbergt, *et al.*, "From wind ensembles to probabilistic information about future wind power production -- results from an actual application," in *Probabilistic Methods Applied to Power Systems, 2006. PMAPS 2006. International Conference on, 2006*, pp. 1-8.

[17]T. Gneiting, F. Balabdaoui, and A. E. Raftery, "Probabilistic forecasts, calibration and sharpness," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 69, pp. 243-268, 2007.

[18]P. Pinson, H. A. Nielsen, J. K. Møller, H. Madsen, and G. N. Kariniotakis, "Non-parametric probabilistic forecasts of wind power: required properties and evaluation," *Wind Energy*, vol. 10, pp. 497-516, 2007.

[19]J. Juban, L. Fugon, and G. Kariniotakis, "Probabilistic short-term wind power forecasting based on *kernel* density estimators," in *European Wind Energy Conference and exhibition, EWEC 2007, 2007*, p. <http://ewec2007proceedings.info/>>.

[20]J. Juban, N. Siebert, and G. N. Kariniotakis, "Probabilistic Short-term Wind Power Forecasting for the Optimal Management of Wind Generation," in *Power Tech, 2007 IEEE Lausanne, 2007*, pp. 683-688.

[21]R. J. Bessa, V. Miranda, A. Botterud, W. Jianhui, and E. M. Constantinescu, "Time Adaptive Conditional *Kernel* Density Estimation for Wind Power Forecasting," *Sustainable Energy, IEEE Transactions on*, vol. 3, pp. 660-669, 2012.

[22]A. d. S. Pires, "Improving electrical production forecasting using GPU," Faculdade de Engenharia da Universidade do Porto, 2013.

[23]P. M. N. d. Fonte, "Advanced forecast and scheduling of power systems with highly variable sources," Faculdade de Engenharia da Universidade do Porto, 2015.

[24]P. Billingsley, *Probability and Measure, Third Edition*: New York, Toronto, London: John Wiley and Sons Inc, 1995.

[25]E. A. Nadaraya, "On Estimating Regression," *Theory of Probability & Its Applications*, vol. 9, pp. 141-142, 1964.

[26]G. S. Watson, "Smooth regression analysis," *Sankhya Ser.*, 26:359-372, 1964.

Anexo A - Histórico dos Parques

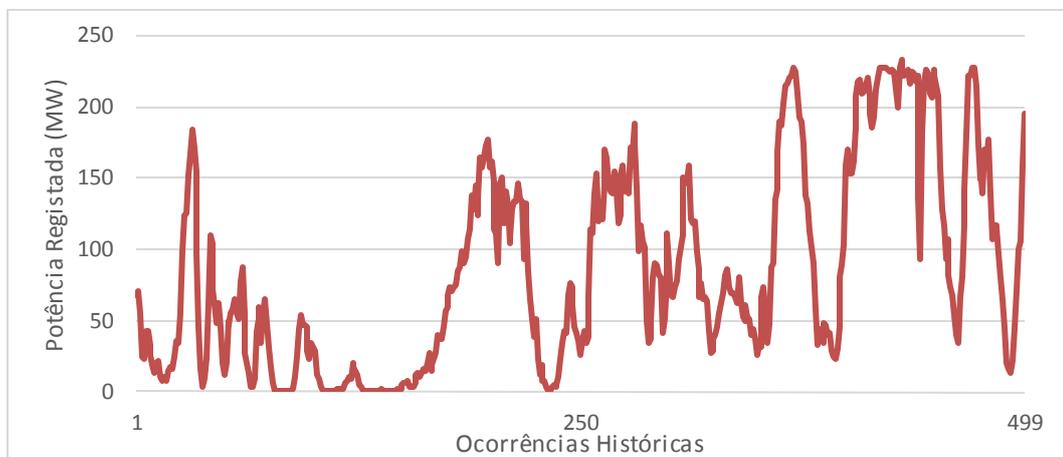


Figura A.1 - Amostra do Target do Parque 2

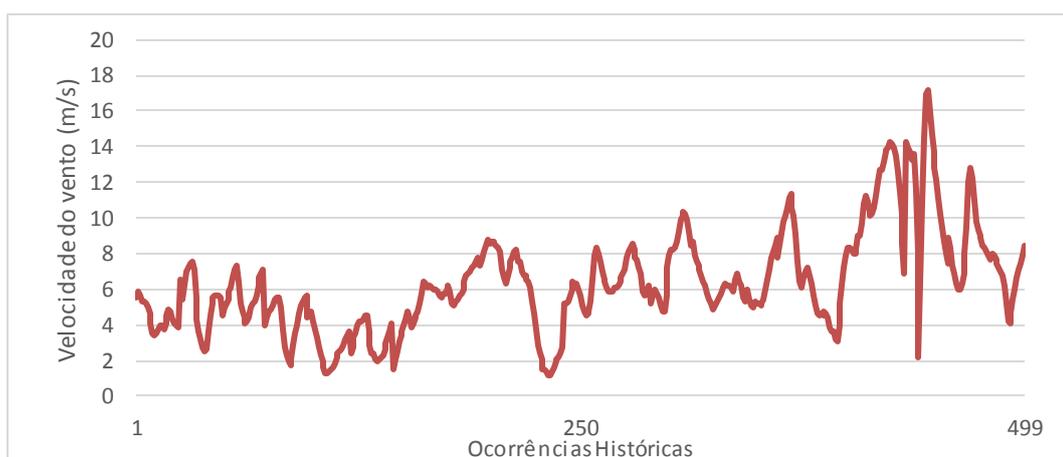


Figura A.2 - Amostra da velocidade do vento do Parque 2



Figura A.3 - Amostra da direção do vento do Parque 2

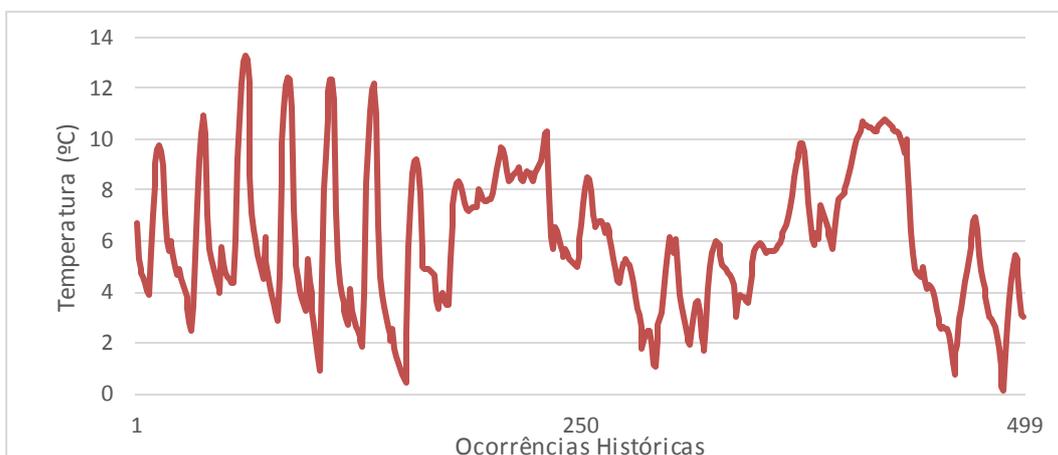


Figura A.4 - Amostra da temperatura do Parque 2

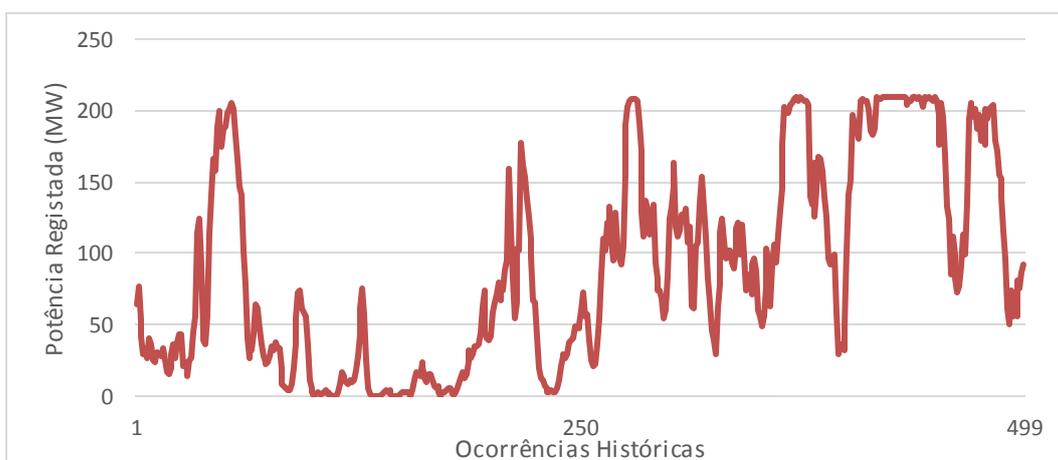


Figura A.5 - Amostra do *Target* do Parque 3

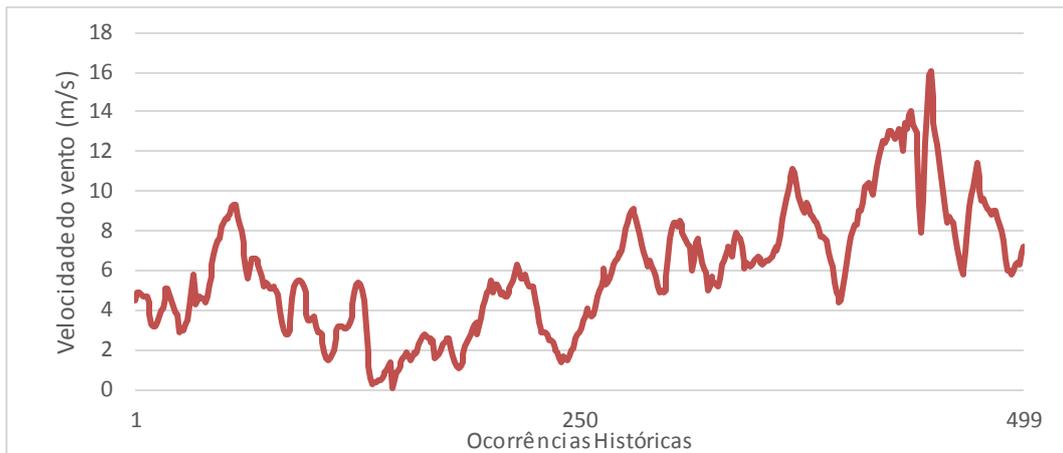


Figura A.6 - Amostra da velocidade do vento do Parque 3



Figura A.7 - Amostra da direção do vento do Parque 3

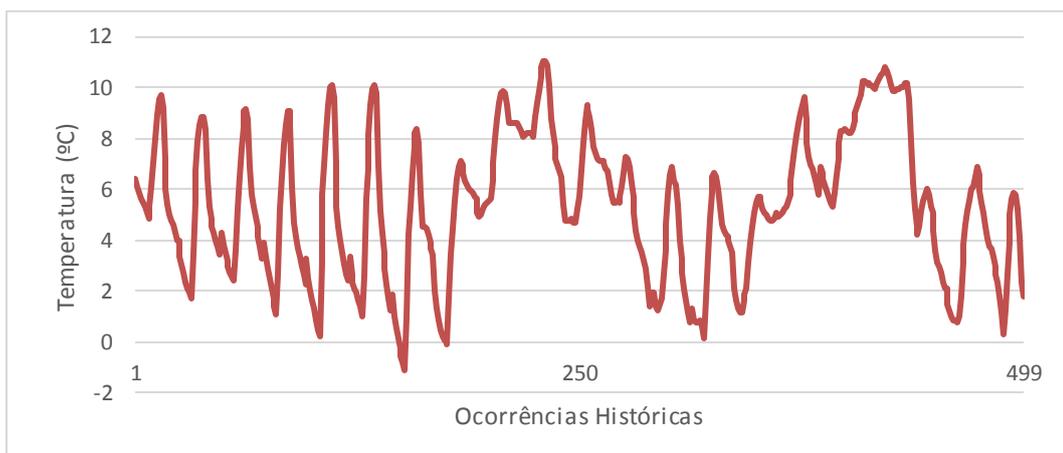


Figura A.8 - Amostra da temperatura do Parque 3



Figura A.9 - Amostra do Target do Parque 4

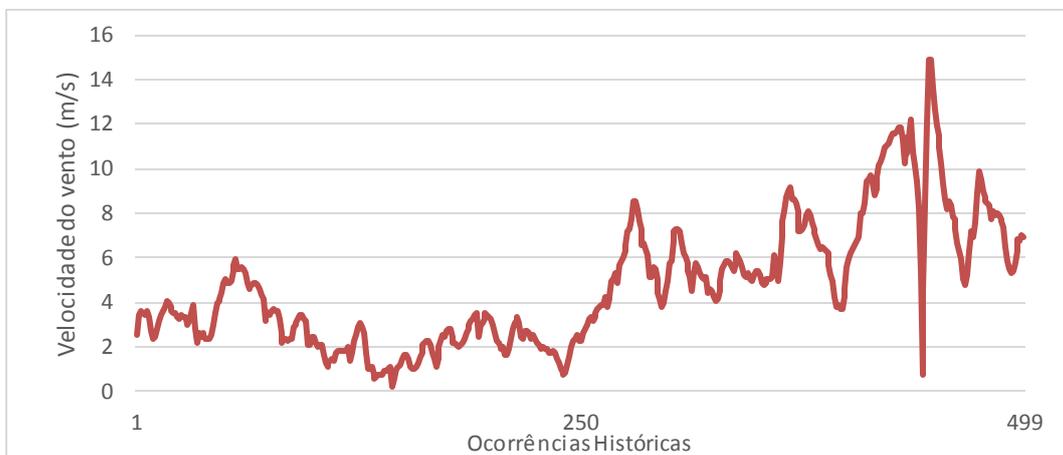


Figura A.10 - Amostra da velocidade do vento do Parque 4

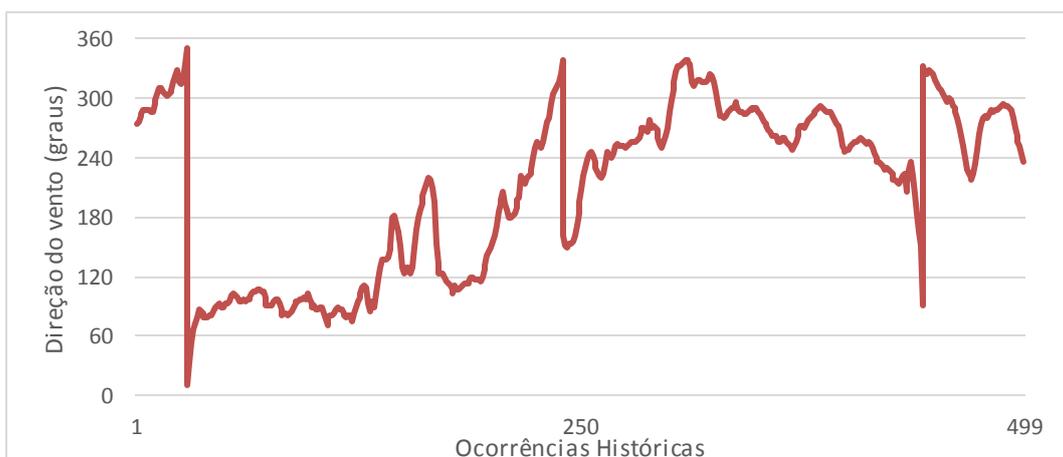


Figura A.11 - Amostra da direção do vento do Parque 4

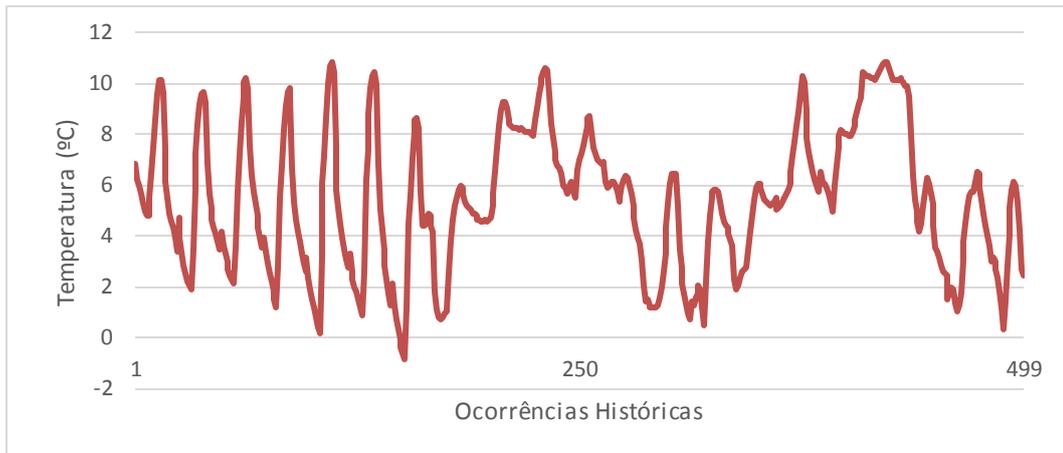


Figura A.12 - Amostra da temperatura do Parque 4

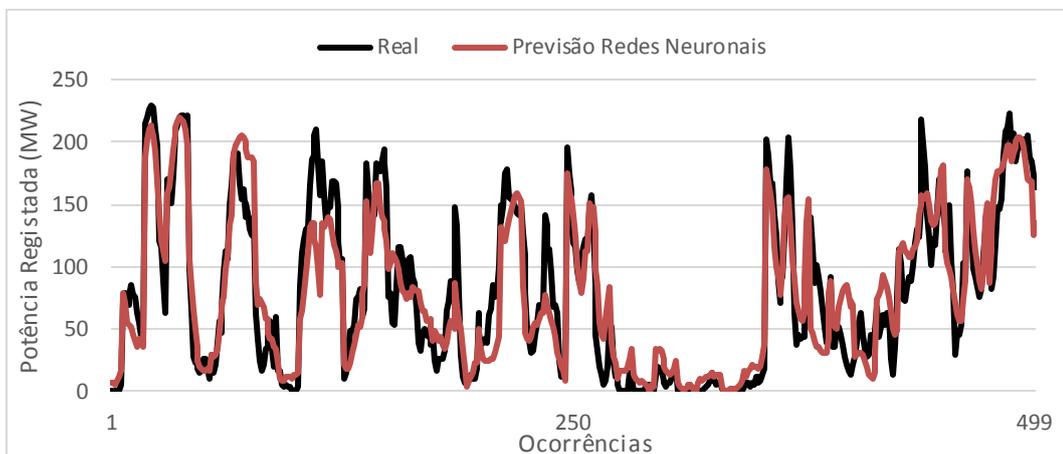


Figura A.13 - Amostra dos valores obtidos da previsão eólica através redes e os valores registrados reais do Parque 2

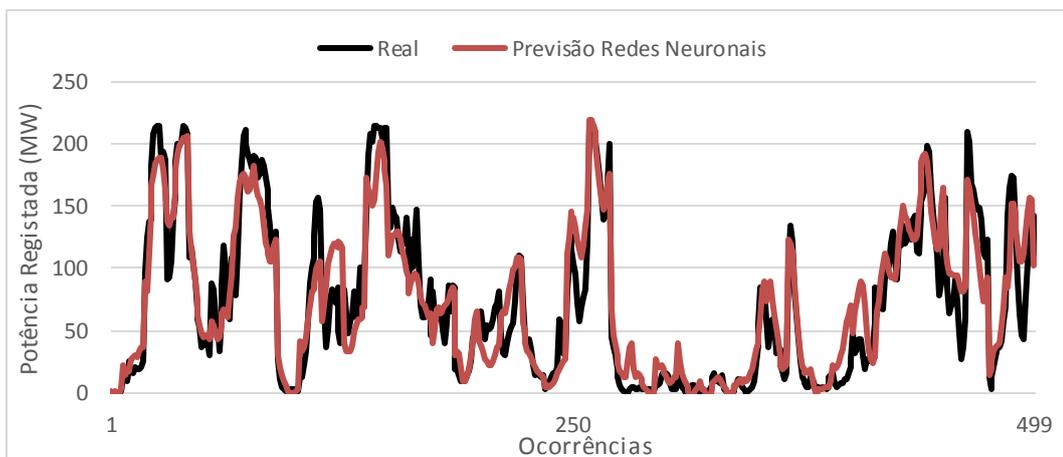


Figura A.14 - Amostra dos valores obtidos da previsão eólica através redes e os valores registrados reais do Parque 3

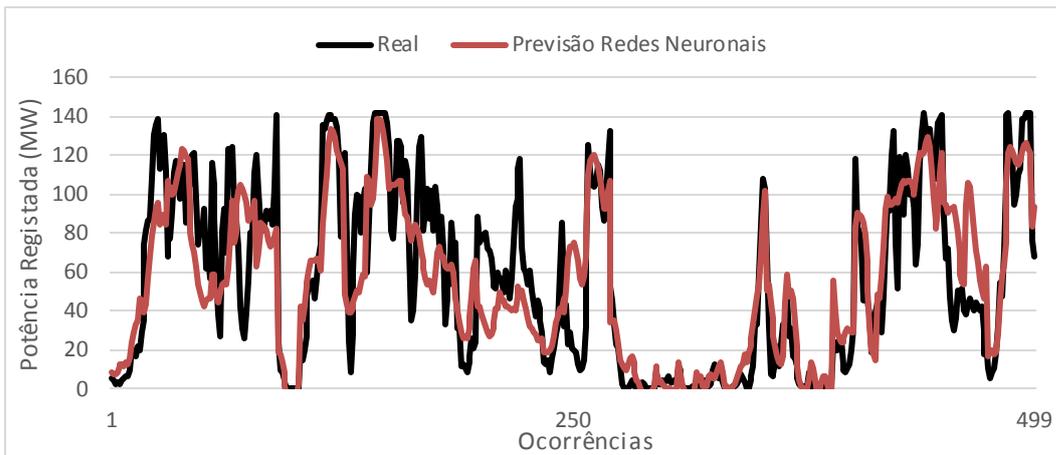


Figura A.15 - Amostra dos valores obtidos da previsão eólica através redes e os valores registados reais do Parque 4