

FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO



Previsão de Produção Eólica com Modelização de Incertezas

Marta Neves Madureira Rolo

Mestrado Integrado em Engenharia Eletrotécnica e de Computadores

Orientador: Professor Doutor Cláudio Monteiro

23 de Julho de 2014

A Dissertação intitulada

“Previsão de Produção Eólica com Modelização de Incertezas”

foi aprovada em provas realizadas em 17-07-2014

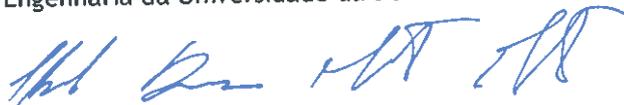
o júri



Presidente **Professor Doutor José Nuno Moura Marques Fidalgo**
Professor Associado do Departamento de Engenharia Eletrotécnica e de
Computadores da Faculdade de Engenharia da Universidade do Porto



Professor Doutor João Paulo da Silva Catalão
Professor Auxiliar do Departamento de Engenharia Electrotécnica da Faculdade de
Ciências da Engenharia da Universidade da Beira Interior



Professor Doutor Cláudio Domingos Martins Monteiro
Professor Auxiliar do Departamento de Engenharia Eletrotécnica e de Computadores
da Faculdade de Engenharia da Universidade do Porto

O autor declara que a presente dissertação (ou relatório de projeto) é da sua exclusiva autoria e foi escrita sem qualquer apoio externo não explicitamente autorizado. Os resultados, ideias, parágrafos, ou outros extratos tomados de ou inspirados em trabalhos de outros autores, e demais referências bibliográficas usadas, são corretamente citados.



Autor - Marta Neves Madureira Rolo

Faculdade de Engenharia da Universidade do Porto

Resumo

A gestão e manutenção dos sistemas elétricos de energia tem vindo a revelar-se cada vez mais complexa, à medida que a integração de geração eólica aumenta. Efetivamente, a variabilidade do vento introduz na rede um elevado nível de incerteza de produção que, por sua vez, provoca um aumento dos custos de desvio e dificulta operações de despacho, *unit commitment* e determinação de preços do mercado de eletricidade.

Neste sentido, e de forma a combater os desequilíbrios de produção originados pelo recurso eólico, têm vindo a ser feitos inúmeros esforços e investimentos no desenvolvimento de técnicas de previsão cada vez mais precisas, tornando a energia eólica mais competitiva.

Tradicionalmente, os sistemas de previsão fornecem valores pontuais da produção esperada. No entanto, atualmente, essa informação revela-se insuficiente. Na presente dissertação, são desenvolvidos novos modelos, baseados em KDE (*Kernel Density Estimation*), que se distinguem dos restantes por permitirem obter uma modelização da incerteza associada à previsão. Neste sentido, é estudado o processo de construção e otimização de modelos do tipo estatístico, com horizonte temporal de curto prazo, aplicados aos parques eólicos da ilha de Creta, Grécia.

Uma das principais fontes de imprecisões neste tipo de sistemas está relacionada com a qualidade do conjunto de dados usado, pelo que se devem utilizar boas previsões meteorológicas. Quanto à informação histórica, desenvolveram-se e compararam-se duas metodologias distintas de eliminação de erros, essenciais para a obtenção de previsões com alta qualidade.

A avaliação do modelo final desenvolvido é feita através da análise dos seus resultados e comparação com outras técnicas de previsão, como é o caso das redes neuronais e do modelo de persistência. No global, pretende-se perceber de que forma é possível otimizar tanto as previsões determinísticas, como as probabilísticas.

Palavras-chave: Previsão de Produção Eólica, Modelização de Incerteza, Previsões Probabilísticas

Abstract

The management and maintenance of electric power systems has become more and more complex, as the integration of wind generation increases. Actually, the wind variability introduces a high level of production uncertainty in the electric grid, which, in its turn, causes higher deviation costs and complicates dispatch operations, unit commitment and determination of the electricity market's prices.

In order to address the imbalances of production caused by the wind resource, countless efforts and investments have been made in the development of more accurate forecasting techniques, making wind energy more competitive.

Traditionally, forecasting systems provide point values of the expected output. However, currently, this information proves to be insufficient. In this dissertation, new models are developed, based on KDE (Kernel Density Estimation), distinguishing themselves from others by obtaining a modelization of the forecast uncertainty. Therefore, the study focuses on the process of building and optimizing models of statistical type, with short term time horizons, applied to wind farms in Crete, Greece.

One of the major sources of inaccuracy in such systems is related to the quality of the data set used, therefore, one should use good numerical weather predictions. As for the historical information, two distinct methods of debugging were developed and compared, in order to obtain high quality forecasts.

The evaluation of the final model is made through the analysis of the results and comparison with other forecasting techniques, such as neuronal networks and the persistence model. Overall, the aim is to understand how both the deterministic and the probabilistic forecasts can be optimized.

Keywords: Wind Power Forecasting Systems, Uncertainty Modelling, Probabilistic Forecasts

Agradecimentos

A presente dissertação representa o culminar de uma etapa de vida, não apenas acadêmica, mas também pessoal. Serve este espaço para agradecer e homenagear as pessoas e instituições que contribuíram, direta ou indiretamente, para o seu sucesso.

Ao meu orientador, professor Cláudio Monteiro, pelo acompanhamento, sabedoria, críticas, sugestões e, não menos importante, boa disposição essenciais ao desenvolvimento deste trabalho.

À empresa *Smartwatt* e seus colaboradores, pelos recursos e cooperação, em especial ao Eng. Bruno Santos, pela paciência e disponibilidade no enriquecimento do trabalho desenvolvido.

À minha mãe, que, apesar dos sacrifícios e preocupação, garantiu todas as condições necessárias à minha concretização acadêmica e pessoal, não só este semestre, mas durante toda a minha vida.

Ao meu irmão, pela prontidão nos conselhos, apoio incondicional e palermices gerais.

Aos meus amigos, que fizeram dos últimos cinco anos uma aventura inesquecível, especialmente, aos que me ajudaram, direta ou indiretamente, durante este último semestre. Aos amigos pacientes que leram e ajudaram a corrigir o presente documento. Ao grupo do "*bullying na tese*". Aos fiéis da *i105*. Aos que me acompanharão o resto da vida.

Obrigada.

*“I consider that a man’s brain originally is like a little empty attic,
and you have to stock it with such furniture as you choose”*

Arthur Conan Doyle, A Study in Scarlet

Conteúdo

1	Introdução	1
1.1	Enquadramento	1
1.2	Objetivos	3
1.3	Informação e Ferramentas Utilizadas	4
1.4	Estrutura da Dissertação	4
2	Estado da Arte	7
2.1	Caracterização de Modelos de Previsão Eólica	7
2.1.1	Descrição Genérica de Previsão Eólica	7
2.1.2	Diferentes Abordagens de Previsão Eólica	8
2.1.3	Utilização da Previsão Eólica	10
2.2	Modelos Numéricos de Previsão Meteorológica (NWP)	11
2.2.1	Sistemas NWP Operacionais	11
2.2.2	Sistemas NWP Baseados em Ensemble	13
2.2.3	Variáveis NWP Usadas em Previsão Eólica	14
2.3	Modelos de Previsão Eólica a Curto Prazo	15
2.3.1	Técnicas Matemáticas de Previsão	17
2.3.2	Modelização de Incertezas	17
2.3.2.1	Abordagem Estatística	18
2.3.2.2	Incertezas Baseadas em <i>Ensemble</i>	18
2.3.3	Ferramentas de Previsão Probabilística	19
2.3.3.1	<i>Kernel Density Estimation</i>	19
2.3.3.2	Distribuição Beta na Modelização de Resultados KDE	20
2.3.4	Modelização de Rampas	21
2.4	Avaliação dos Modelos de Previsão	22
2.4.1	Fatores de Influência nos Erros de Previsão	23
2.4.2	Indicadores de Erros de Previsão	24
2.4.3	Interpretação de Desempenho de Previsão	25
3	Processamento de Dados	27
3.1	Informação Utilizada	27
3.2	Metodologias Desenvolvidas	30
3.2.1	Filtragem de Erros	31
3.2.2	Modelização da Curva de Potência Média	33
3.3	Comparação de Metodologias	33
3.4	Conclusões	35

4	Desenvolvimento de Modelos de Previsão	37
4.1	Estrutura e Funcionamento	37
4.2	Escolha de Variáveis	38
4.3	Modelização de Incerteza	43
4.4	Ajuste de Parâmetros	46
4.5	Ajuste Dinâmico	49
4.6	Conclusões	51
5	Avaliação do Modelo Final	53
5.1	Análise de Resultados	53
5.2	Comparação entre Modelos	58
5.3	Comparação com Outras Técnicas	63
5.4	Aplicações	65
5.5	Conclusões	67
6	Conclusões	69
6.1	Satisfação de Objetivos	69
6.2	Conclusões Gerais	69
6.3	Trabalhos Futuros	70
A	Filtragem de Erros	73
B	Erros da Previsão Probabilística	79
	Referências	83

Lista de Figuras

1.1	Repartição da produção dos diferentes tipos de energia, no ano de 2012	1
1.2	Custos de desvio entre o valor real e o valor previsto de potência produzida por diferentes tipos de fontes de energia	2
2.1	Extrapolação da velocidade do vento a uma determinada altura do eixo, através de dados NWP - abordagem física	9
2.2	<i>Short Range Numerical Weather Prediction</i> (SRNWP) na Europa	12
2.3	Função densidade de probabilidade de produção eólica adaptada a uma distribuição beta	21
2.4	Curva de potência de um aerogerador	23
2.5	Variação do RMSE para diferentes horizontes temporais e modelos de previsão	26
3.1	Curva de potência do parque eólico <i>Rokas</i>	29
3.2	Curva de potência do parque eólico <i>Kouloukonas</i>	29
3.3	Curva de potência do parque eólico <i>Anemos Alkionis</i> , onde as zonas a vermelho assinalam pontos correspondentes a erros	30
3.4	Representação gráfica de uma curva genérica criada segundo a equação 3.1	31
3.5	Definição dos limites da curva de potência (3.5a) e resultado, após aplicação dos filtros (3.5b) - parque eólico <i>Enteka</i> , subestação <i>Siteias</i>	32
3.6	Modelização da curva de potência média do parque eólico <i>Enteka</i> , subestação <i>Siteias</i>	33
3.7	Curva de potência do agregado obtido através da 1ª metodologia - filtragem de erros	34
3.8	Curva de potência do agregado obtido através da 2ª metodologia - modelização da curva de potência média	34
4.1	Representação do funcionamento genérico de um modelo de previsão estatístico	37
4.2	Representação dos <i>targets</i> a utilizar nos modelos de previsão, para o ano de 2013	38
4.3	Esquema representativo dos horizontes temporais das previsões meteorológicas	39
4.4	Representação das variáveis a utilizar nos modelos de previsão, para o ano de 2013	40
4.5	Velocidade de vento durante parte dos meses Março e Abril de 2012, com falhas (<i>blanks</i>) destacadas	41
4.6	Comparação entre as previsões de produção eólica obtidas através do Modelo II e do Modelo VI, para dois meses de 2013	42
4.7	Previsão da produção eólica com representação de incerteza, com intervalo de confiança 80%, para uma semana de Maio de 2013	44
4.8	Previsão da produção eólica com representação de incerteza, com intervalo de confiança 50%, para uma semana de Maio de 2013	44
4.9	Probabilidade obtida em função da probabilidade esperada, para diferentes intervalos de confiança e através de diferentes modelos	45

4.10	Valores de MAPE para diferentes valores de desvio padrão da velocidade de vento, no caso do Modelo III	46
4.11	Valores de MAPE para diferentes valores de desvio padrão da velocidade de vento, no caso do Modelo VI	47
4.12	Valores de MAPE para diferentes valores de desvio padrão da direção de vento (à esquerda) e da temperatura (à direita), no caso do Modelo VI	48
4.13	Representação da qualidade da incerteza, comparando o Modelo VI melhorado com o anterior	49
4.14	Esquema representativo do método de ajuste dinâmico	50
4.15	Linha de tendência do gráfico de dispersão entre a previsão e os valores reais de produção eólica, com função de ajuste dinâmico	50
5.1	Previsão da produção eólica com representação de incerteza para o mês de janeiro de 2013 (MAPE=32,532%)	54
5.2	Previsão da produção eólica com representação de incerteza para o mês de julho de 2013 (MAPE=16,138%)	54
5.3	Exemplo de erro de amplitude (janeiro)	55
5.4	Exemplo de erro de fase (julho)	55
5.5	Comparação entre a produção registada em 2012 e em 2013	55
5.6	Evolução média do erro e da produção de potência eólica, ao longo do dia	56
5.7	Representação dos valores de incerteza obtidos, para cada intervalo de confiança	56
5.8	Comparação entre as previsões de produção eólica, obtidas através do modelo final, antes e após a realização do seu ajuste dinâmico, para o mês de Janeiro de 2013	57
5.9	Previsões do Modelo VI antes e após o ajuste das suas parametrizações	59
5.10	Comparação entre as previsões de produção eólica obtidas através de diferentes modelos, para o mês de Setembro de 2013	59
5.11	Valores de MAPE mensais da previsão determinística, para cada um dos modelos criados	60
5.12	Valores do <i>bias</i> mensal da previsão determinística, para cada um dos modelos criados	60
5.13	Previsões obtidas através do Modelo II e do Modelo IV	61
5.14	Histograma do erro médio de cada um dos modelos desenvolvidos	62
5.15	Histograma do erro médio relativo ao modelo final com e sem ajuste dinâmico	62
5.16	Evolução média do erro ao longo do dia, segundo diferentes modelos de previsão	63
5.17	Evolução média do erro ao longo do dia, segundo diferentes técnicas de previsão	64
5.18	Previsões da produção eólica para um dia, segundo diferentes técnicas de previsão	65
5.19	Exemplo de situação de despacho, com representação de previsão probabilística e necessidade de produção eólica	65
5.20	Aplicação de previsões de produção eólica em decisões de despacho	66
5.21	Modelo final em funcionamento, implementado à ilha de Creta pela <i>Smartwatt</i>	67
B.1	Resultados detalhados da previsão probabilística obtida para todo o ano de 2013, no caso do Modelo I	79
B.2	Resultados detalhados da previsão probabilística obtida para todo o ano de 2013, no caso do Modelo II	79
B.3	Resultados detalhados da previsão probabilística obtida para todo o ano de 2013, no caso do Modelo III	80
B.4	Resultados detalhados da previsão probabilística obtida para todo o ano de 2013, no caso do Modelo IV	80

B.5	Resultados detalhados da previsão probabilística obtida para todo o ano de 2013, no caso do Modelo V	80
B.6	Resultados detalhados da previsão probabilística obtida para todo o ano de 2013, no caso do Modelo VI	81
B.7	Resultados detalhados da previsão probabilística obtida para todo o ano de 2013, no caso do modelo final	81

Lista de Tabelas

2.1	Modelos de previsão numérica globais, operacionais atualmente	11
2.2	Resumo dos sistemas NWP mesoescala/LAM operacionais na Europa	13
2.3	Resumo dos sistemas NWP baseados em ensemble (EPS) operacionais na Europa	14
2.4	Modelos de previsão de produção eólica	16
3.1	Subestações e parques eólicos em estudo	28
3.2	Percentagens de erro médio absoluto (MAPE) em função dos dados utilizados . .	35
4.1	Valores de MAPE para diferentes modelos, de acordo com as variáveis utilizadas para os criar	41
4.2	Valores de MAPE da previsão probabilística (incerteza), para os diferentes modelos criados	45
4.3	Valores de MAPE para diferentes combinações de valores de desvio padrão das variáveis de entrada do Modelo VI	48
4.4	Valores de MAPE da previsão determinística e da previsão probabilística, antes e após o ajuste dos parâmetros do modelo selecionado	49
4.5	Valores de MAPE da previsão determinística e da previsão probabilística para a segunda metade de 2013, antes e após o ajuste dinâmico dos resultados	51
5.1	Características do modelo de previsão selecionado para implementação	53
5.2	Valores de MAPE de ambos os tipos de previsão, para um mês de Inverno (janeiro) e um de Verão (julho)	57
5.3	Valores de MAPE da previsão determinística (pontual) e da previsão probabilística (incerteza) para os diferentes modelos criados	58
5.4	Valores de MAPE da previsão determinística utilizando diferentes métodos	64

Abreviaturas e Símbolos

AR	Auto Regressivo
ARIMA	<i>Autoregressive Integrated Moving Average</i>
ARMA	<i>Autoregressive Moving Average</i>
DMOS	<i>Dynamic Model Output Statistics</i>
ECMWF	<i>European Center for Medium-Range Weather Forecasts</i>
EEM	Empresa de Eletricidade da Madeira
EPS	<i>Ensemble Prediction System</i>
GFS	<i>Global Forecasting System</i>
KDE	<i>Kernel Density Estimation</i>
LAM	<i>Limited Area Model</i>
MAE	<i>Mean Absolute Error</i> (erro médio absoluto)
MAPE	<i>Mean Absolute Percentage Error</i>
MOS	<i>Model Output Statistics</i>
NCAR	<i>National Center of Atmospheric Research</i>
NCEP	<i>National Center for Environmental Prediction</i>
NWP	<i>Numerical Weather Predictions</i>
OFE	Outras Fontes de Energia
RMSE	<i>Root Mean Squared Error</i> (Erro Médio Quadrático)
RN	Redes Neurais
SCADA	<i>Supervisory Control and Data Acquisition</i>
SEE	Sistema Elétrico de Energia
SRNWP	<i>Short Range Numerical Weather Prediction</i>
UE	União Europeia
WPPT	<i>Wind Power Prediction Tool</i>
WPF	<i>Wind Power Forecasting</i>

Capítulo 1

Introdução

1.1 Enquadramento

O contínuo aumento do preço dos combustíveis fósseis, juntamente com as crescentes preocupações ambientais por parte dos países, faz com que a energia eólica, sendo uma fonte renovável e limpa, se torne cada vez mais relevante em todo o mundo.

Neste sentido, a União Europeia possui objetivos ambiciosos, uma vez que pretende que em 2020, 20% da energia total produzida seja proveniente de fontes renováveis e que haja uma redução em cerca de 20% (em relação a 1990) não só da emissão de gases de efeito de estufa, mas também do consumo de energia em geral, através de medidas de eficiência energética. Em Portugal, segundo os dados da REN [1], a produção de origem renovável no ano de 2012 abasteceu 37% do consumo, com a eólica a atingir a quota mais elevada de sempre, 20% (figura 1.1). Desde então, este valor tem vindo a aumentar, com a produção eólica a suprir 24% do consumo em 2013 e com a percentagem a chegar aos 25%, no presente ano de 2014. Este rápido aumento da quantidade de geração eólica introduzida no sistema elétrico de energia tem associado, também, um aumento da incerteza de produção, devido à grande variabilidade do vento.

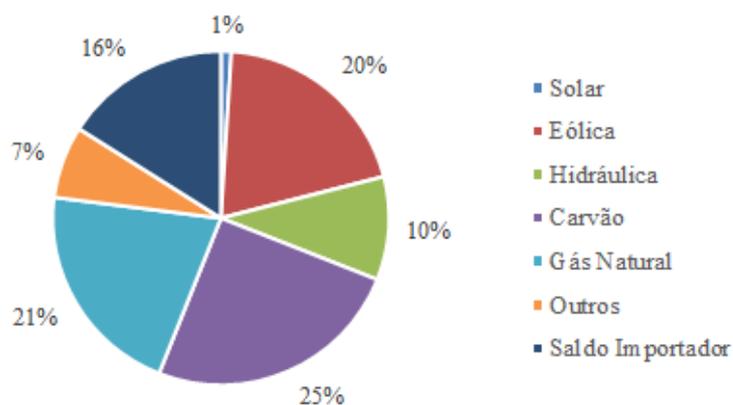


Figura 1.1: Repartição da produção dos diferentes tipos de energia, no ano de 2012 [1]

Desta forma, surgem inúmeros desafios para todos os intervenientes no mercado de eletricidade, sendo particularmente relevante a dificuldade dos operadores do sistema em balancear de forma eficiente a procura e a oferta de energia.

A previsão de produção eólica (WPF - *Wind Power Forecasting*) surge, então, como uma ferramenta essencial na gestão dos sistemas de produção que integram grandes quantidades deste tipo de energia. Para além de contribuir para melhores decisões de despacho, é importante na determinação de preços diários e horários, uma vez que quanto maior for a geração eólica, menores serão os custos marginais de produção de energia por meios convencionais. Se o preço de mercado diminui com a integração de tecnologias com produção intermitente, então o investimento em tecnologias tradicionais também tende a diminuir, por falta de viabilidade económica, o que pode comprometer a continuidade de serviço [2]. Para além disso, existem custos de desvio, isto é, penalizações a ser pagas, pelos produtores, por cada MWh de diferença entre o valor real e o valor previsto de potência produzida, tal como é apresentado no gráfico da figura 1.2, referente a dados de 2011 e 2012 [3].

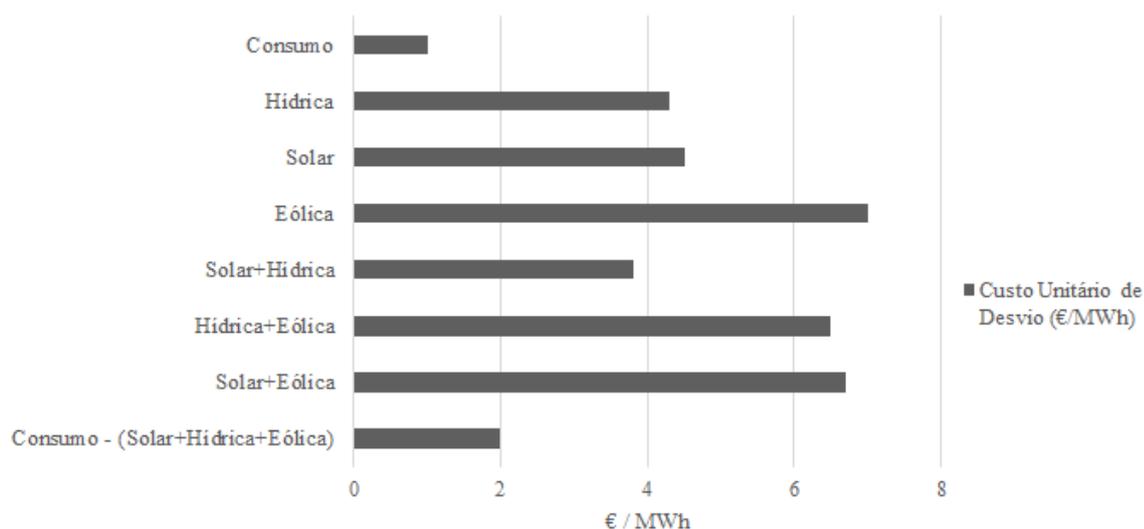


Figura 1.2: Custos de desvio entre o valor real e o valor previsto de potência produzida por diferentes tipos de fontes de energia [3]

Uma vez que todos os participantes na indústria elétrica beneficiam das previsões de produção eólica, têm sido cada vez maiores os apoios e esforços no sentido de aperfeiçoar e desenvolver sistemas mais fiáveis e precisos, de modo a integrar mais eficientemente este tipo de energia no sistema.

Um sistema de previsão caracteriza-se, essencialmente, segundo o horizonte de tempo para o qual realiza previsões e segundo a abordagem que utiliza (física, estatística ou híbrida), podendo, ainda, recorrer a modelos numéricos de previsão meteorológica. A escolha do modelo mais adequado está dependente da aplicação à qual este se destina, variando em cada caso. Por exemplo, previsões realizadas para o dia seguinte ("*day-ahead*") são necessárias para o mercado elétrico,

enquanto que previsões com horizontes temporais de dias ou semanas, são úteis no planeamento de ações de manutenção [4,5].

No entanto, a grande maioria segue uma abordagem determinística (ou "*point forecasting*"), isto é, calculam o valor esperado da potência produzida em cada instante do horizonte temporal. Atualmente, para fazer uma gestão adequada da produção de energia eólica na rede elétrica, este tipo de informação não se tem revelado suficiente [6,7]. É vantajoso para os operadores do sistema conhecer uma estimativa da incerteza associada às previsões pontuais da produção, conhecendo o grau de dispersão dos valores previstos e as suas probabilidades. Assim, tem vindo a revelar-se cada vez mais importante, paralelamente às previsões do valor esperado de produção, realizar previsões probabilísticas. Estas, possibilitam tomar decisões com base em análise de risco (risco de não alimentação da carga ou de deslastre de produção, por exemplo).

A diversidade de abordagens e metodologias que os modelos podem seguir, faz com que a sua análise em função de inúmeros fatores, como características locais ou previsões meteorológicas, constitua um trabalho extenso.

1.2 Objetivos

O tema da presente dissertação está enquadrado no âmbito de um projeto europeu da SiNGULAR (*Smart and Sustainable Insular Electricity Grids Under Large-Scale Renewable Integration*) e tem como principal objetivo o desenvolvimento de novos modelos de previsão da produção de potência para um horizonte temporal de cerca de 7 dias (curto prazo), aplicados a parques eólicos da ilha de Creta, Grécia.

Ao longo deste trabalho, pretende-se fazer um estudo sobre como otimizar tanto a previsão pontual dos valores de produção esperados, como a incerteza a estes associada. É neste valor acrescentado que os modelos em estudo se distinguem dos atuais.

De uma forma simplificada, pretende-se que este trabalho cumpra os seguintes objetivos:

- Desenvolvimento de metodologias de filtragem de erros em séries históricas usadas para o treino dos modelos;
- Criação de modelos de previsão de produção eólica para o agregado de parques de Creta;
- Obtenção de previsões da produção eólica, juntamente com a incerteza a estas associada (previsão probabilística);
- Aplicação de novas técnicas de previsão baseadas em KDE, comparando com redes neurais e com o modelo de persistência;
- Criação de modelos de previsão probabilísticos;
- Desenvolvimento de metodologias de parametrização de modelos probabilísticos.

1.3 Informação e Ferramentas Utilizadas

Para a construção dos modelos de previsão, utilizaram-se dados cedidos pela empresa *Smartwatt*, provenientes do projeto europeu da SiNGULAR, aplicados a casos de estudo na ilha de Creta, Grécia.

As principais ferramentas computacionais utilizadas foram:

- KDE, desenvolvido na *Smartwatt*;
- *Microsoft Office Excel*;
- Matlab;

O *software* KDE foi utilizado para construir os modelos de previsão. Este, caracteriza-se por ser capaz de fornecer previsões de produção e, ainda, a incerteza a estas associada. Tal como o nome indica, baseia-se na técnica não-paramétrica "*Kernel Density Estimation*", cujo objetivo é encontrar uma relação entre determinadas variáveis. O modo como o programa funciona é descrito em pormenor no capítulo 2, em ferramentas de previsão probabilística.

Toda a análise dos resultados obtidos, qualidade das previsões e da incerteza, comparação de modelos, etc., foi realizada em *Excel*. Numa primeira fase, este *software* foi também a principal ferramenta utilizada no tratamento dos dados, eliminação de erros e análise da qualidade da informação.

O *software* Matlab foi utilizado, essencialmente, no desenvolvimento de um modelo baseado em redes neurais, com intuito comparativo em relação ao modelo desenvolvido em KDE.

1.4 Estrutura da Dissertação

A presente dissertação encontra-se organizada em 6 capítulos.

No capítulo 1 é feita uma introdução ao trabalho, descrevendo o tema, a sua relevância e os objetivos que se pretendem atingir. São, ainda, referidas as ferramentas computacionais e a informação utilizadas ao longo do documento.

No capítulo 2 é descrito o estado da arte e é feita uma revisão da literatura sobre modelos de previsão de produção eólica, com especial ênfase nos de curto-prazo. Aqui, é descrito o conceito de modelização de incertezas e algumas das formas mais usuais de avaliação dos modelos. São, ainda, descritos os modelos numéricos de previsão meteorológica mais relevantes.

No capítulo 3 é feita uma análise da qualidade dos dados históricos de produção registados e são apresentadas duas metodologias para o seu processamento, posteriormente comparadas, de forma a selecionar a melhor.

No capítulo 4 são desenvolvidos modelos de previsão eólica para o caso de estudo em questão, de forma a encontrar a melhor solução. Assim, é descrita toda a formulação e etapas de construção do modelo, analisando os efeitos de diferentes especificações na qualidade das previsões determinísticas e probabilísticas. É feita uma otimização do modelo final selecionado e é apresentada, ainda, uma técnica de ajuste final das previsões obtidas.

No capítulo 5 é feita a análise dos resultados obtidos segundo do modelo desenvolvido e otimizado no capítulo 4. Este é, também, comparado a outras técnicas de previsão, neste caso, a um modelo baseado em redes neurais e ao modelo de persistência. Por fim, são apresentadas algumas das aplicações mais relevantes.

No capítulo 6 são, então, apresentadas todas as conclusões retiradas ao longo da dissertação, fazendo uma análise crítica à satisfação de objetivos propostos. É, ainda, feita uma referência a trabalhos com possível interesse de desenvolvimento futuro.

Capítulo 2

Estado da Arte

O presente capítulo baseia-se num estudo da literatura existente sobre previsão de produção eólica. Como principais fontes, utilizou-se o relatório acerca de previsão de energia renovável emitido pela Universidade da Califórnia em 2010 – “*Wind Energy Forecasting: A Review of State-of-the-Art and Recommendations for Better Forecasts*” [8]-, o relatório técnico do SiNGULAR emitido em 2014 – “*Report on the mathematical formulation of the different forecasting models and data analysis*” [9] – e, ainda, o estado da arte emitido pela ANEMOS em 2011 – “*The state-of-the-art in short-term prediction of wind power: A literature overview*” [6]. A restante bibliografia é referenciada ao longo do capítulo e encontra-se enumerada no final do documento.

2.1 Caracterização de Modelos de Previsão Eólica

2.1.1 Descrição Genérica de Previsão Eólica

A previsão eólica trata-se de uma estimativa da produção esperada de uma ou mais turbinas eólicas (ou parques eólicos) num futuro próximo (desde alguns minutos até vários dias). Para tal, utilizam-se modelos computacionais que, através de determinados dados de entrada (*inputs*), originam estimativas da produção eólica num horizonte futuro (*outputs*) – modelos de previsão eólica [8].

Tal como dito anteriormente, o principal desafio que se coloca na produção de energia eólica está relacionado com a variabilidade do vento, que causa instabilidade no mercado elétrico e nos despachos que os operadores dos sistemas de produção têm de realizar. Assim, existe uma necessidade cada vez maior de prever e controlar a produção eólica, de forma a integrar mais eficientemente grandes quantidades desta nos sistemas. Existem vários fatores que influenciam a precisão dos modelos de previsão eólica, entre eles, a topologia do terreno do parque eólico, a irregularidade da superfície, o clima, os padrões de vento, etc. Isto implica que, para cada projeto, seja feita uma comparação e análise de diferentes modelos e/ou combinações de modelos, por forma a encontrar o melhor para a situação específica. É, ainda, importante o investimento contínuo no desenvolvimento de tecnologias de medição, por exemplo, relativas a dados meteorológicos, para aumentar a qualidade e viabilidade destas previsões. Em redes que integram grande quantidade de

unidades de produção eólica, a imprecisão das previsões pode levar a grandes custos associados a despachos económicos incorretos, atividades de manutenção ou incumprimento de prazos em projetos [8, 10].

De uma forma geral, os modelos de previsão eólica diferenciam-se de acordo com o horizonte temporal para o qual as previsões são realizadas, podendo ir desde uns minutos, até alguns dias. Podem, ainda, integrar modelos numéricos de previsão meteorológica (NWP – Numerical Weather Predictions) ou não. Estas diferentes abordagens são descritas em maior pormenor no próximo ponto do capítulo.

2.1.2 Diferentes Abordagens de Previsão Eólica

Uma das principais características de um sistema de previsão está relacionado com o horizonte temporal que lhe está associado (por exemplo, o dia seguinte). Considerando que o vento se mantém constante, é possível, num determinado instante t , fazer uma previsão da produção média, por exemplo, de um parque eólico, para um instante futuro, $t+k$. Por sua vez, este valor k (passo) pode variar, dependendo da resolução de tempo que se pretende obter. Por exemplo, para horizontes entre 24 a 72 horas, é comum que o passo seja de hora a hora [11].

Apesar de não existir unanimidade na definição das fronteiras que separam os diferentes horizontes de tempo, é comum dividir as previsões de produção eólica em três categorias [11]:

- **Muito curto prazo:** o intervalo de tempo é de apenas algumas horas, não havendo unanimidade na bibliografia em relação a número certo. No entanto, é comum considerar-se cerca de 6 horas. Este tipo de previsões pode ser útil, por exemplo, nos mercados intra-diários.
- **Curto prazo:** o horizonte de tempo varia entre o limite estabelecido para o muito curto prazo e cerca de uma semana, necessitando de previsões meteorológicas (NWP) de centros de previsão. É utilizado, maioritariamente, quando se pretende obter previsões para os dias seguintes de forma a realizar o *unit commitment* das unidades de geração convencional, tal como acontece no MIBEL (Mercado de Eletricidade Ibérico), podendo também ser útil no agendamento de ações de manutenção.
- **Médio prazo:** o horizonte de tempo varia desde o limite estabelecido para o curto prazo até um mês. Têm a desvantagem de quanto maior o horizonte temporal, maiores serem os erros associados. Estas previsões são utilizadas para obter informação sobre tendências e padrões de comportamento.

Existe, também, a denominação de “*Nowcasting*”, que corresponde a previsões para os próximos segundos ou minutos, usadas para o controlo e gestão dos parques eólicos. Atualmente, a maioria dos sistemas de previsão utilizados e estudados possuem horizontes de tempo compreendidos entre as 36 e as 72 horas, pelo que é conveniente um estudo mais aprofundado sobre modelos de curto prazo [11].

As previsões de produção eólica a curto-prazo incluem, quase sempre, modelos numéricos de previsão meteorológica (descritos em maior detalhe, posteriormente), podendo, também, ser classificadas entre dois tipos de abordagem [6]:

- Abordagem física
- Abordagem estatística

Na abordagem física, os modelos de previsão baseiam-se em características físicas do terreno e do tempo, utilizando-as ao máximo para alcançar a melhor estimativa possível da velocidade do vento antes de recorrer a modelos estatísticos (MOS – *Model Output Statistics*) para reduzir o erro resultante. Este tipo de sistemas caracteriza-se por utilizar resultados de modelos numéricos de previsão meteorológica, com origem em centros de previsão globais, ajustando-os de acordo com as características locais, de modo a obter melhores previsões (ver Figura 2.1) [12, 13]. Possuem melhor resolução horizontal e vertical do que os modelos dos centros globais, obtendo melhores resultados em situações de baixos níveis de vento [8].

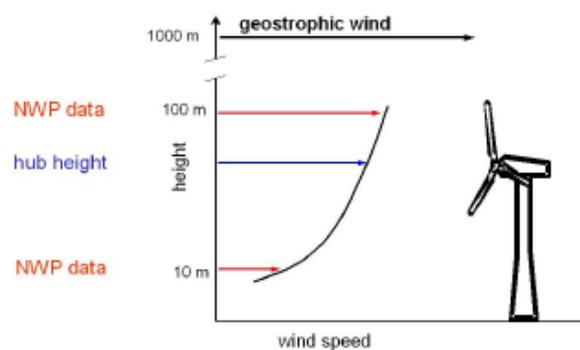


Figura 2.1: Extrapolação da velocidade do vento a uma determinada altura do eixo, através de dados NWP - abordagem física [12]

Na abordagem estatística, são utilizados dados históricos relativos a medições de potência eólica e informação meteorológica, aplicando um ou vários algoritmos estatísticos de forma a desenvolver um modelo de previsão. Estes algoritmos podem incluir técnicas como regressões lineares múltiplas, redes neuronais ou lógica fuzzy, por exemplo. Os modelos estatísticos podem ser utilizados em qualquer fase do encadeamento de modelos, podendo combinar vários passos num só [8, 13].

A maioria dos modelos operacionais comercializados combinam ambas as abordagens, utilizando, em particular, previsões do tempo e séries temporais, uma vez que ambas podem ser necessárias para obter sucesso nas previsões de produção eólica - abordagem híbrida [6].

Na prática, utilizam-se mais frequentemente *ensembles* de previsões do que previsões individuais, uma vez que, em algumas aplicações, produzem previsões e estimativas de incerteza de qualidade superior. O conceito básico consiste em gerar um conjunto de previsões a partir de perturbações nos dados de entrada e nos parâmetros do modelo, dentro dos respetivos intervalos de

incerteza. No entanto, dado que existe um elevado número de variáveis de entrada e parâmetros, não é prático gerar previsões para todas as perturbações possíveis, sendo necessário selecionar quais os responsáveis pela maioria das incertezas do sistema em questão. Esta seleção pode-se tornar difícil, uma que os valores responsáveis pela incerteza variam muito de um ciclo de previsão para outro [8].

2.1.3 Utilização da Previsão Eólica

Quando existem grandes quantidades de unidades de produção eólica inseridas na rede, a incerteza associada aos níveis de energia gerada torna-se problemática, existindo custos associados a unidades ligadas em excesso, bem como a eficiências reduzidas e aumentos na necessidade de manutenção. Previsões melhoradas podem reduzir estes custos [8].

São inúmeros os utilizadores que podem beneficiar de previsões eólicas. No caso dos operadores de sistema, estas tornam-se importantes para uma boa gestão da variabilidade do vento, garantindo a qualidade de serviço do sistema elétrico de energia, de acordo com a procura, e minimizando possíveis congestionamentos. Saber antecipadamente em que alturas o vento vai atingir velocidades mais elevadas, permite, ainda, aos operadores do sistema reduzir custos, tirando de serviço centrais de produção mais dispendiosas [8]. Efetivamente, a otimização do planeamento do sistema elétrico de energia é proporcionada pela gestão da produção convencional (despacho económico, etc.) em função das previsões de produção eólica. Isto acontece mais visivelmente em sistemas de menores dimensões ou isolados da rede [6].

Existem penalizações impostas aos fornecedores de energia, quando ocorrem desvios na produção agendada, pelo que estes também beneficiam das previsões de produção eólica, que permitem conhecer o vento esperado e, portanto, minimizar os custos [8, 10].

A investigação no sentido do desenvolvimento de projetos eólicos tira, também, partido das previsões, uma vez que o vento funciona como combustível neste tipo de sistemas. Apesar de ser gratuito, trata-se de um combustível que não é possível comprar. Assim, as características do local onde será construído determinado projeto é o fator controlável mais importante, determinando se o projeto é economicamente viável ou não [8].

Para além disso, horizontes temporais ainda maiores seriam interessantes no planeamento de operações de manutenção dos grandes componentes das centrais de produção, das turbinas eólicas e das linhas de transmissão. Dependendo das condições meteorológicas, esta manutenção pode durar desde algumas horas até semanas, uma vez que é necessário evitar precipitação, ventos fortes e temperaturas extremas [8]. O problema reside no facto de, quando se consideram desde 5 a 7 dias em avanço, a precisão das previsões meteorológicas diminuir significativamente, tratando-se de um tipo de sistema que só começa agora a surgir [6].

Em suma, as previsões são maioritariamente utilizadas para o planeamento da produção convencional e das trocas de eletricidade no mercado spot, bem como previsão de preços em relação ao dia seguinte (dayahead). Têm, também, um papel importante na alocação das unidades de produção para as horas seguintes. Esta aplicação intradiária requer atualizações fiáveis das previsões de modo a originar previsões melhoradas para horizontes temporais cada vez menores [12, 14].

Como resultado da crescente inclusão de produção eólica no sistema elétrico de energia, verifica-se que a média do preço de mercado tende a diminuir, bem como o número de horas de operação das tecnologias térmicas tradicionais. Assim, novos investimentos em tecnologias tradicionais têm tendência, também, em diminuir, o que torna o sistema elétrico cada vez mais dependente de tecnologias voláteis. Torna-se, então, um problema que deve ser analisado, uma vez que a segurança do fornecimento de energia pode ser posto em causa [2].

2.2 Modelos Numéricos de Previsão Meteorológica (NWP)

A precisão das previsões meteorológicas é fundamental na qualidade das previsões de produção eólica. De facto, é nas variáveis meteorológicas que reside o maior potencial de melhoria das previsões eólicas, em horizontes temporais superiores a 6 horas. Torna-se relevante, então, tentar melhorar as entradas NWP, de forma a aumentar significativamente o desempenho dos sistemas de previsão.

2.2.1 Sistemas NWP Operacionais

Neste ponto, pretende-se caracterizar os modelos numéricos de previsão meteorológica operacionais com maior relevância na previsão de produção eólica na Europa e, em particular, em Portugal. Portugal é membro do *European Centre for Medium-range Weather Forecasts* (ECMWF), cujo modelo de previsão é global, ou seja, efetua previsões para todo o globo. De acordo com o Instituto Português do Mar e da Atmosfera (IPMA), a versão atual operacional deste modelo, em Portugal, tem uma resolução horizontal aproximada de 40 km. A tabela seguinte faz um resumo dos modelos NWP globais utilizados atualmente.

Tabela 2.1: Modelos de previsão numérica globais, operacionais atualmente [6]

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

O *Global Forecast System* (GFS), por exemplo, está em constante atualização e é capaz de fornecer previsões de um conjunto vasto de variáveis, deste vento e precipitação, a humidade do solo e concentrações atmosféricas [15]. Com o progressivo aumento dos recursos computacionais,

pretende-se que os modelos globais ultrapassem as funções atuais dos modelos de área limitada (LAM – *Limited Area Models*). Os LAM's foram desenvolvidos, inicialmente, para investigação de processos atmosféricos de mesoescala, ou seja, processos com escalas horizontais entre 1 até algumas centenas de km [11]. Atualmente, operam com resoluções horizontais de 7 a 12 km e possuem limites estabelecidos pelos modelos globais. Pretende-se que, no futuro, os modelos de área limitada sejam substituídos por altas resoluções, existindo já projetos para os integrar nos modelos globais, com o objetivo de atingir resoluções horizontais até 5 km em determinadas áreas alvo, como é o caso da Europa (por exemplo, o modelo ARPEGE da *Meteo France*) [6].

O projeto do grupo SRNWP (*Short Range Numerical Weather Prediction*), tornou-se o principal veículo de cooperação entre os consórcios Europeus de modelos de área limitada. Estes consórcios de previsão numérica meteorológica são os projetos ALADIN, COSMO, HIRLAM, LACE e o UK MetOffice (figura 2.5) [6]. Existe, ainda, o projeto C-SRNWP (o “C” vem de coordenação) que tem vindo a ser desenvolvido desde 2000 e pretende melhorar a troca de informação entre membros, representar os interesses da SRNWP perante outras organizações e maximizar os benefícios das previsões meteorológicas numéricas, no geral [16].

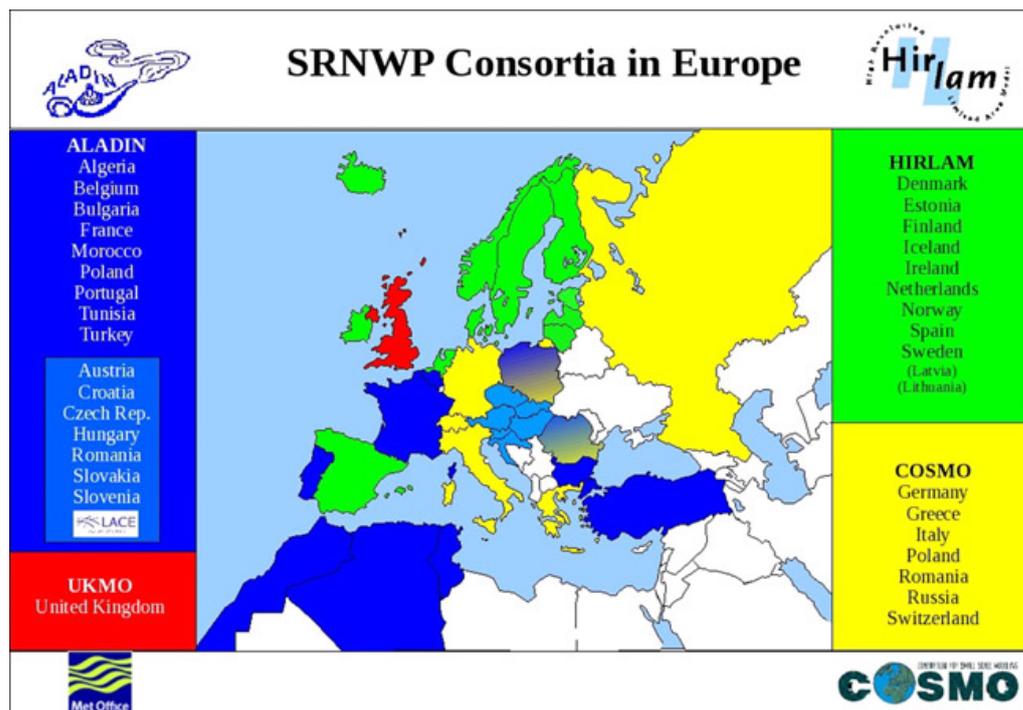


Figura 2.2: *Short Range Numerical Weather Prediction* (SRNWP) na Europa [6]

A figura anterior traduz a lista dos serviços meteorológicos existentes na Europa e a sua cooperação nos diferentes consórcios. Portugal é parceiro no projeto internacional ALADIN, cuja versão que está atualmente operacional no país possui uma resolução horizontal de 9 km. Um outro produto resultante da atividade do IPMA é o modelo de mesoescala de alta resolução AROME, que utiliza as previsões do ALADIN operacional como condições iniciais/fronteira e é executado

localmente com uma resolução horizontal de 2,5 km, para um alcance de 48 horas [16].

Tabela 2.2: Resumo dos sistemas NWP mesoescala/LAM operacionais na Europa [6]

Modelo	Países	Horizonte de Tempo
ALADIN	Áustria, Bélgica, Bulgária, Croácia, Eslováquia, Eslovénia, França, Hungria, Polónia, Portugal, República Checa, Roménia	72 h
AROME	França	30 h
COSMO	Alemanha, Grécia, Itália, Polónia, Roménia, Suíça	78 h
ETA	Sérvia	120 h
HARMONIE	Finlândia	24 h
HIRLAM	Dinamarca, Espanha, Estónia, Finlândia, Holanda, Irlanda, Noruega, Suécia	72 h
HRM	Bósnia e Herzegovina, Bulgária, Itália, Roménia	72 h
MM5	Turquia	48 h
UM	Noruega, Reino Unido	60 h
WRF	Sérvia	48 h

Alguns dos modelos de mesoescala constituem versões limitadas de modelos globais, enquanto que outros são criados especificamente para estudo de alta resolução. O modelo MM5 era um dos mais utilizados na América, até o modelo de mesoescala WRF (*Weather Research and Forecasting*) ser lançado pelo NCAR (*National Center of Atmospheric Research*), juntamente com o NCEP (*National Center for Environmental Prediction*). Trata-se de um modelo de nova geração mais atual e flexível, podendo utilizar diferentes parametrizações e esquemas numéricos [11].

2.2.2 Sistemas NWP Baseados em Ensemble

Tal como referido anteriormente, os sistemas baseados em *ensemble* baseiam-se no conceito de gerar um conjunto de previsões a partir de perturbações nos dados de entrada e nos parâmetros do modelo, dentro dos respetivos intervalos de incerteza. Para tal, é necessário selecionar quais as variáveis e parâmetros responsáveis pela maioria das incertezas do sistema em questão [8]. O THORPEX (*The Observing System Research and Predictability Experiment*) é um programa internacional de investigação com o objetivo de acelerar os avanços na precisão das previsões meteorológicas desde um dia até duas semanas. Uma componente chave deste programa é o TIGGE (*THORPEX Interactive Grand Global Ensemble*), cuja base de dados é constituída por previsões de vários sistemas globais baseados em *ensemble* (conjunto) [6]. O sistema NWP baseado em *ensemble* (EPS – *Ensemble Prediction System*) existente em Portugal, é operado pelo ECMWF desde 1992, tendo sofrido diversas atualizações e melhorias ao longo dos anos. Atualmente, as suas previsões são feitas para 15 dias [17].

Na Europa, existem sete *ensembles* de área limitada operacionais em grandes centros meteorológicos (ver Tabela 2.3), sendo que alguns deles são armazenados na base de dados central do projeto TIGGE-LAM (componente de modelos de área limitada do TIGGE). Para além destes

Tabela 2.3: Resumo dos sistemas NWP baseados em ensemble (EPS) operacionais na Europa [6]

Modelo	País	Horizonte de Tempo
ALADIN	Hungria	60 h
ALADIN-LEAF	Áustria	60 h
COSMO-LEPS	Itália	120 h
LAM-EPS	Noruega	60 h
MOGREPS-G	Reino Unido	72 h
MOGREPS-R	Reino Unido	54 h
PEARP (ARPEGE-EPS)	França	60 h

sistemas (LAM-EPS), está a ser desenvolvido o projeto SRNWP-PEPS, na Alemanha, tratando-se de um sistema ainda experimental, que gera previsões probabilísticas de um multi-modelo de previsão baseado em *ensemble*. Faz uma combinação entre as previsões de área limitada operacionais e os serviços meteorológicos europeus [6].

2.2.3 Variáveis NWP Usadas em Previsão Eólica

A maioria dos modelos de previsão eólica a curto prazo utilizados atualmente (para horizontes de 6 a 72 horas), necessitam de previsões meteorológicas (NWP). A principal característica que distingue as diferentes abordagens tem a ver com o modo como estas utilizam as variáveis NWP [11]. A abordagem física foca-se nas características do fluxo de vento nas imediações e no interior do parque eólico, juntamente com a curva de potência dos aerogeradores, de forma a propor uma estimativa da produção eólica. Ao longo do processo de previsão são utilizadas as variáveis velocidade do vento e a sua direção em relação à turbina eólica, bem como a modelização do perfil de vento local, considerando a estabilidade atmosférica. A ideia é refinar estas variáveis NWP, usando considerações físicas do terreno: inclinação, rugosidade, orografia, obstáculos. Para um modelo NWP, são usadas relações estatísticas entre as variáveis da previsão e as variáveis meteorológicas observadas, de forma a corrigir as primeiras. Assim, é possível que as previsões de variáveis meteorológicas produzidas através desta abordagem contribuam, de certa forma, para os erros dos modelos NWP [11].

Quanto aos modelos de previsão eólica que se baseiam na abordagem estatística, existe apenas um passo: transformação direta de variáveis de entrada em produção eólica. Para tal, utilizam também variáveis NWP, tais como a velocidade do vento, direção, temperatura, etc., de vários níveis de modelos, em conjunto com medições *online* das mesmas. Hoje em dia, os principais desenvolvimentos das abordagens estatísticas em previsão de produção eólica pretendem evoluir para a utilização de previsões meteorológicas múltiplas (*ensembles*) como uma combinação de entradas e saídas [11].

No entanto, é necessário ter em atenção que existem obstáculos ao acesso a dados para previsões [8]:

- Informação restrita: a disponibilidade de dados por estar restrita por motivos de confidencialidade ou devido a dificuldade de recuperação de informação em sistemas de bases de dados complexos.
- Qualidade dos dados: existe um vasto conjunto de problemas associados à qualidade da informação armazenada, podendo haver erros associados.
- Frequência insuficiente das amostras: os dados eólicos são, muitas vezes, guardados com intervalos de 10 minutos ou mesmo de hora em hora, o que não é suficiente em algumas previsões, particularmente quando consideram rampas.

2.3 Modelos de Previsão Eólica a Curto Prazo

Um modelo de previsão eólica é um sistema de *software*, estruturado e encadeado através de diversos sistemas de informação, que utiliza diversas entradas (*inputs*) para determinar valores de produção (*outputs*) futuros. O seu grau de complexidade é muito variável, podendo utilizar-se apenas um modelo ou uma combinação de vários. O modelo de persistência, por exemplo, é um dos mais simples, onde todas as previsões para os instantes seguintes são colocadas ao valor atual. Uma vez que funciona bastante bem a muito curto prazo, tornou-se o modelo de referência que os restantes tentam ultrapassar. Os sistemas de previsão eólica mais modernos apresentam complexidades consideravelmente superiores [8].

O modelo de previsão a curto prazo *Prediktor*, baseia-se numa lógica física, onde a ideia é usar a velocidade e direção do vento obtidas a partir de um sistema NWP, transformar este vento de acordo com as características do local, usar a curva de potência e, finalmente, modifica-la segundo a eficiência do parque eólico. A transformação do vento pode ser feita utilizando o vento de níveis mais altos da camada atmosférica ou utilizando o valor fornecido pelo sistema NWP a 10m acima do nível do solo, opção que, atualmente, produz melhores resultados e permite que o logaritmo seja usado diretamente. É importante referir que, para que haja convergência do MOS, são necessários cerca de 4 meses de dados históricos. O *Prediktor* é, também, utilizado em sistemas SCADA para planeamento de manutenções [6].

O WPPT (*Wind Power Prediction Tool*) é um sistema de modelização para a previsão de produção eólica, desenvolvido na Universidade Técnica da Dinamarca, que utiliza previsões do HIRLAM e pode ser configurado para tirar partido dos seguintes dados [6]:

- Medições *online/offline* da produção em parques eólicos individuais;
- Medições *online/offline* agregadas da produção no caso de áreas maiores;
- Dados NWP relativos a parques eólicos individuais;

- Dados NWP relativos a áreas maiores.

Efetivamente, a complexidade da estrutura do modelo implementado pelo WPPT vai depender da informação disponível, uma vez que este possui uma hierarquia de modelização altamente flexível. As previsões podem ser calculadas na forma de um único valor esperado ou na forma de previsões probabilísticas, onde é dada a distribuição dos resultados [6].

É muito provável que este sistema sofra alterações ao longo do tempo, uma vez que usa características do vento no local e variáveis NWP, pelo que o facto da previsão ser adaptativa temporalmente (*time-adaptive*) constitui uma característica bastante relevante. O horizonte temporal das previsões de produção eólica corresponde ao horizonte das previsões meteorológicas numéricas, ou seja, geralmente até 48 horas [6].

Várias abordagens têm vindo a ser testadas no âmbito da previsão de produção eólica, baseando-se em modelos ARMA (*Autoregressive Moving Average*), diferentes tipos de redes neuronais, lógica *fuzzy*, etc. O projeto MORE-CARE da ARMINES permitiu desenvolver modelos de previsão para as 48/72 horas seguintes, através de sistemas SCADA e NWP, integrando: modelos de muito curto prazo (baseados em abordagens estatísticas de séries temporais), modelos de prazos até 72 horas (baseados em redes neuronais *fuzzy*) e combinação de modelos (atribuindo diferentes pesos às previsões, para obter resultados ótimos). Este sistema encontra-se operacional nas ilhas de Creta e da Madeira. Nestas, uma vez que se trata de uma rede elétrica isolada, as previsões são especialmente importantes para uma integração económica e segura da produção eólica. Em Creta, as previsões são calculadas para um horizonte de 48 horas, a partir dos dados obtidos pelo sistema SCADA da ilha e pelo sistema SKIRON (NWP). Na Madeira, este sistema é operado pela EEM (Empresa de Eletricidade da Madeira), o horizonte temporal é de 8 horas e possui a particularidade de fazer, ainda, previsões para as centrais hidroelétricas [6].

O modelo *eWind*, desenvolvido pela agora denominada *AWS TruePower*, utiliza os modelos de mesoescala ARPS, MASS e WRF, juntamente com os modelos globais ECMWF, GEM e GFS, de modo a produzir um *ensemble* de 9 modelos diferentes [6]. Muitos outros modelos têm vindo a ser desenvolvidos, como por exemplo: *Hugin*, *Previento* (semelhante ao *Prediktor*), *WPFS* (*Wind Power Forecasting Models*), *WPMS* (*Wind Power Management System*), *GH Forecaster* ou a ferramenta EPREV [6]. A tabela seguinte faz um resumo dos modelos mais relevantes [18].

Tabela 2.4: Modelos de previsão de produção eólica

Modelo	Horizonte de Tempo	Abordagem	Developer
Prediktor	48 h	Física	Risø
WPPT	48 h	Estatística	IMM & DTU
MORE-CARE	72 h	Híbrida	ARMINES
eWind	48 h	Híbrida	AWS TruePower
Previento	96 h	Híbrida	Universidade de Oldenburg
WPMS	8 h	Estatística	ISET
GH Forecaster	Alguns dias	Híbrida	Garrad Hassan

2.3.1 Técnicas Matemáticas de Previsão

São várias as técnicas matemáticas que têm vindo a ser utilizadas no âmbito da previsão de produção eólica, pelo que a bibliografia é muito extensa. Algumas destas destinam-se, unicamente, à previsão da velocidade do vento que é, posteriormente, convertida para produção através da curva de potência dos aerogeradores. Outras técnicas têm como objetivo fazer a conversão das previsões NWP para produção eólica – modelos W2P (*wind-to-power*) [11].

Caso o horizonte de previsão não seja demasiado longo, a velocidade do vento e a produção eólica podem ser previstos utilizando apenas métodos de séries temporais, sem recorrer a previsões meteorológicas [6]. Os modelos de séries temporais constituem uma metodologia clássica na previsão a curto prazo, designando-se também por *Box-Jenkins* ou modelos autoregressivos (AR). Dentro destes, é comum a utilização dos modelos ARIMA (*Autoregressive Integrated Moving Average*) e ARMA (*Autoregressive Moving Average*) [11].

Ao longo dos anos foram surgindo inúmeras abordagens com o objetivo de obter cada vez melhores resultados. Estas englobam filtros *Kalman*, estimação do vento para redução de *flicker*, modelos de simulação estocástica, entre outros. Um estudo realizado para os casos de Creta e Shetland, utilizou um modelo linear AR e um modelo baseado em lógica *fuzzy* adaptativa, tendo-se descoberto melhorias pouco significativas em relação ao modelo de persistência para horizontes de 2 horas, mas com redução do erro médio quadrático (RMSE) em cerca de 20% em horizontes de 8 horas. No entanto, para horizontes mais longos, as abordagens com base em dados meteorológicos mostraram-se mais promissoras. Segundo a bibliografia existente, concluiu-se, ainda, que o desempenho dos modelos é altamente dependente dos seus períodos de treino [6]. Outras técnicas de previsão baseiam-se em inteligência artificial, como é o caso das redes neuronais (RN), dos modelos SVM (*Support Vector Machines*) e dos sistemas de inferência difusa (FIS – *Fuzzy Inference Systems*). Os modelos baseados em RN e em SVM revelaram-se, em diversos estudos, globalmente superiores aos modelos de regressão linear e ao modelo de persistência. Existe, ainda, a abordagem híbrida ANFIS (*Adaptive Neural Fuzzy System*) que, combinando redes neuronais com lógica *fuzzy*, tem vindo a ser aplicada em previsões a muito curto prazo, apesar de poder atingir horizontes temporais até 72 horas [11].

Atualmente, as melhores previsões são obtidas através de redes neuronais recorrentes locais, que se baseiam em informação meteorológica. Esta abordagem aplica um algoritmo do tipo RPE (*Recursive Prediction Error*), que treina a rede neuronal *online*, isto é, vai fazendo atualizações *online* dos pesos e tendências. Quanto melhores estes algoritmos de treino, maior é a capacidade de lidar com as mudanças de comportamento e operação que ocorrem nos parques eólicos [11].

2.3.2 Modelização de Incertezas

Para a gestão ideal da produção de energia em parques eólicos, não é suficiente conhecer os valores da previsão, sendo essencial, também, o uso de ferramentas que permitam determinar o risco associado aos mesmos – incerteza. Neste sentido, um tipo de previsão que tem vindo a ganhar relevo é a probabilística, uma vez que constitui uma ferramenta de apoio à decisão muito

útil, por exemplo, em decisões de despacho com base na modelização de incerteza, em oposição à consideração de apenas um valor esperado. Uma das aplicações deste tipo de previsão reside, por exemplo, na criação e análise de diversos cenários de produção, de modo a avaliar alternativas e obter uma maximização do lucro dos produtores [19].

Efetivamente, existem duas principais metodologias utilizadas na indústria para a previsão de incertezas, que se tornam mais relevantes de referir [6]:

- Abordagem estatística, com previsões NWP simples;
- Incertezas resultantes de *ensembles* de previsões;

Em geral, os diferentes modelos matemáticos possuem já estimações de intervalos de confiança associados. Por exemplo, as metodologias estatísticas possuem uma estimativa da incerteza integrada, enquanto os modelos físicos necessitam de processamento adicional para obter essa estimativa [6].

2.3.2.1 Abordagem Estatística

Os métodos de intervalos de confiança mais utilizados assumem que a previsão de erros segue uma distribuição Gaussiana. No entanto, em alguns casos os intervalos de confiança não possuem simetria em relação ao ponto previsto, devido à forma da curva de potência do parque eólico [6]. Um dos métodos aplicáveis tanto a modelos de previsão eólica estatísticos como físicos, baseia-se numa abordagem de reamostragem, juntamente com a consideração do risco associado à estabilidade meteorológica. Este método foi verificado através do uso de previsões HIRLAM e de informação proveniente de 5 parques eólicos irlandeses. Um outro tipo de modelos são os LQR (*Local Quantile Regression*), que utilizam uma técnica de previsão probabilística de forma a estimar diretamente os diferentes quantis da distribuição [6].

Com o objetivo de estimar os intervalos de previsão dinâmica dos sistemas de previsão eólica existentes, foram analisados comparativamente cinco modelos matemáticos diferentes: redes neurais, LQR, regressão multi-linear, classificação simples e modelos adaptativos. Em termos de exatidão foram considerados semelhantes, no entanto apresentaram resultados significativamente distintos no que diz respeito a fiabilidade. Uma combinação dos resultados obtidos através dos cinco modelos permitiu o desenvolvimento de um modelo *ensemble* com qualidade superior de estimativa de incertezas [6].

No subcapítulo seguinte (2.3.3), é descrita a ferramenta de previsão probabilística utilizada ao longo da dissertação, para o desenvolvimento dos modelos. Esta segue, também, uma abordagem estatística.

2.3.2.2 Incertezas Baseadas em *Ensemble*

Com os avanços tecnológicos sofridos ao longo do tempo, surgiram novas formas de estimativa de incertezas. A tentativa de aumentar cada vez mais a resolução das previsões foi substituída pelo

interesse na redução dos restantes erros. Para tal, podem-se utilizar *ensembles* de duas formas distintas [6]:

- *Ensembles* multi-modelo, usando vários modelos NWP diferentes ou parametrizações diferentes do mesmo modelo;
- Fazendo variar os dados de entrada e calculando o *ensemble* baseado em diferentes inicializações de previsões.

Os benefícios deste tipo de previsão de incertezas é ainda, atualmente, estudado por diversos grupos da área.

Vários estudos foram conduzidos no sentido de comparar a utilização de dois tipos modelos de previsão separadamente por oposição a uma combinação de ambos. Os resultados foram sempre os mesmos: a combinação de modelos é melhor ($< RMSE$) do que o melhor dos dois modelos utilizados. Por exemplo, utilizaram-se como entradas 16 serviços de meteorologia europeus diferentes no sistema *Previento*, de forma a correr uma combinação de previsões. Assim, combinaram-se até 5 previsões NWP, reduzindo os erros RMSE em cerca de 4,2 a 5% para toda a Alemanha [6]. O sistema MSEPS (*Multi-Scheme Ensemble Prediction System*) da WEPROG (*Weather and Energy PROGnosis*), companhia dinamarquesa, possui a capacidade de desenvolver incertezas físicas diferentes para cada *ensemble*. Esta característica torna-se especialmente interessante em previsão de produção eólica, uma vez que, ao contrário do que acontece na maioria das abordagens *ensemble*, a incerteza deixa de estar dependente do horizonte de tempo previsto [6].

Fez-se uma avaliação do valor das previsões do ECMWF para os mercados energéticos, tendo surgido como melhor *ensemble* a denominada climatologia condicionada no EPS (*ECMWF Ensemble Prediction System*). Trata-se de um algoritmo que seleciona 10 dias de um conjunto de previsões históricas cujos valores melhor se aproximam da previsão atual. Estes serão, então, usados para modelizar a distribuição de probabilidade da previsão [6].

2.3.3 Ferramentas de Previsão Probabilística

2.3.3.1 *Kernel Density Estimation*

Uma das técnicas não-paramétricas mais utilizada em previsões probabilísticas de produção eólica é a densidade de Kernel estimada (KDE - *Kernel Density Estimation*), que permite obter uma função densidade da probabilidade para cada instante da previsão. Ainda assim, os modelos que usam esta abordagem podem ser muito diferentes entre si. Por exemplo, caso estes sejam treinados *offline*, não vão possuir a capacidade de se adaptar a mudanças temporais da relação entre as variáveis explicatórias e os *targets*. No entanto, é possível criar modelos KDE que, por exemplo, dão maior peso aos dados mais recentes [9].

O método KDE caracteriza-se por fazer uma estimativa suavizada da densidade, através da soma de resultados de uma função K , conhecida como função de *Kernel*, que, por sua vez, calcula a contribuição para a densidade de cada variável histórica X_i (ver equação 2.1).

$$\hat{f}(x) = \frac{1}{N \cdot h} \sum_{i=1}^N K\left(\frac{x - X_i}{h}\right) \quad (2.1)$$

Onde, N é o número de dados históricos e h é a largura de banda, que controla a suavização da estimativa [9, 20]. Neste processo, é importante que estejam bem definidos os parâmetros K e h . Quanto à função de Kernel (K), esta pode ser diferente para cada variável. No entanto, de forma a simplificar e uma vez que o impacto é mínimo, é usual considerar-se uma Gaussiana para todas as variáveis, definida na equação seguinte [21].

$$K(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} \quad (2.2)$$

No que diz respeito ao parâmetro h , por outro lado, é necessário testar e determinar qual o valor ótimo para cada variável e em cada situação. Uma vez que a sua influência nos resultados dos modelos é elevada, este parâmetro é estudado em maior pormenor ao longo do desenvolvimento dos modelos, no capítulo 4.

Caso se utilizem múltiplas variáveis, como é o caso dos modelos desenvolvidos na presente dissertação, multiplica-se o Kernel de cada uma delas por diferentes parâmetros suavizadores h_j , de acordo com a equação 2.3 [21].

$$\hat{f}(x_1, \dots, x_d) = \frac{1}{N} \sum_{i=1}^N \left[\prod_{j=1}^d \frac{1}{h_j} K\left(\frac{x_j - X_{ij}}{h_j}\right) \right] \quad (2.3)$$

Onde d representa o número de variáveis.

A previsão determinística permite obter um valor médio de produção. No entanto, interessa dar mais peso aos pontos em que a distância entre o valor registado e a entrada x é menor. Para tal, usa-se o estimador de Nadaraya-Watson, que permite obter a densidade condicionada (ver equação 2.4) e a média pesada de cada ponto previsto (ver equação 2.6) [21].

$$\hat{f}_{p_{t+k}|x_t}(p, x_j) = \sum_{i=1}^N K\left(\frac{p - P_i}{h_p}\right) \frac{\prod_{j=1}^d K\left(\frac{x_j - X_{ij}}{h_j}\right)}{\sum_{i=1}^N \left[\prod_{j=1}^d K\left(\frac{x_j - X_{ij}}{h_j}\right) \right]} \quad (2.4)$$

$$\hat{p}_{t+k}(x_j) = \sum_{i=1}^N \frac{\prod_{j=1}^d K\left(\frac{x_j - X_{ij}}{h_j}\right)}{\sum_{i=1}^N \left[\prod_{j=1}^d K\left(\frac{x_j - X_{ij}}{h_j}\right) \right]} \cdot P_i \quad (2.5)$$

2.3.3.2 Distribuição Beta na Modelização de Resultados KDE

Após a obtenção da distribuição discreta, através da metodologia anterior, é necessário fazer uma aproximação a uma função densidade de probabilidade paramétrica contínua. O *software* usado ao longo da dissertação faz esta aproximação através de uma distribuição beta, que se

mostrou eficaz na modelização de variáveis com limites máximos e mínimos, como é o caso da potência eólica [9].

A figura 2.3 mostra dois exemplos de distribuições beta adaptadas a histogramas de previsão de produção eólica, por sua vez obtidos através da técnica KDE.

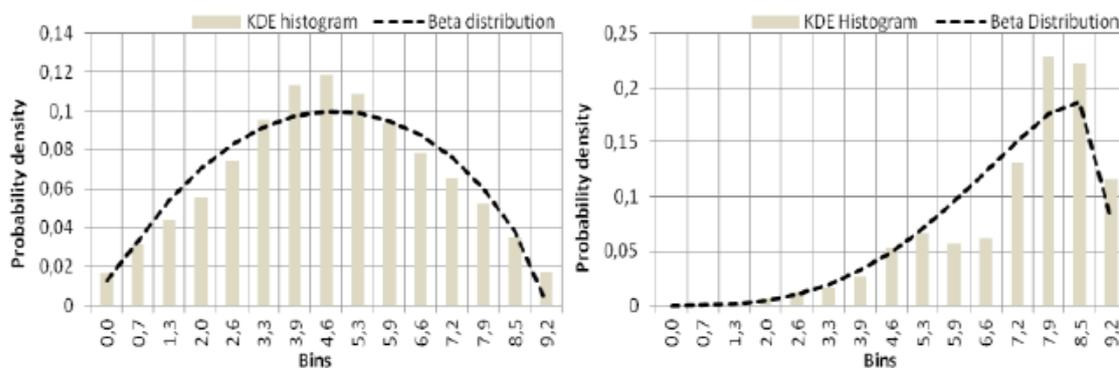


Figura 2.3: Função densidade de probabilidade de produção eólica adaptada a uma distribuição beta [9]

Efetivamente, trata-se de uma boa distribuição paramétrica na modelização deste tipo de variáveis, pelo que se torna relevante defini-la [21]:

$$f(x, \alpha, \beta) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B[\alpha, \beta]} \quad (2.6)$$

Onde $B[\alpha, \beta]$ é uma constante de normalização que garante que a probabilidade total **tende** para 1, x é um valor definido entre $[0, 1]$ e α e β são parâmetros de forma maiores do que 0 [9]. Para além disso, a distribuição beta pode ser simulada usando apenas estes dois parâmetros de forma, que são calculados para cada instante de previsão através da média e variância da distribuição, segundo as equações seguintes [9].

$$\alpha = \frac{(1-\mu) \cdot \mu^2}{\vartheta} - \mu \quad (2.7)$$

$$\beta = \frac{(1-\mu)}{\mu} \cdot \alpha \quad (2.8)$$

2.3.4 Modelização de Rampas

Um evento de rampa caracteriza-se por uma determinada variação da geração de energia num dado intervalo de tempo. O desafio de integrar a produção eólica na rede, ou de maximizar o valor de mercado da energia, torna-se especialmente difícil em períodos onde esta variação ocorre de forma repentina. A definição de um evento de rampa não é uniforme ao longo da bibliografia, havendo autores que definem uma rampa como uma mudança de 20% ou mais da capacidade do parque eólico em intervalos de 1 hora. Outros definem-na como uma mudança de, no mínimo,

50% da capacidade, em intervalos de 4 horas ou menos [8, 14]. Para o trabalho a realizar, não serão estabelecidos este tipo de parâmetros para consideração de rampas, assumindo apenas que as situações de maior severidade caracterizam-se por grandes variações da produção em intervalos de tempo relativamente curtos. A necessidade de perceber quais os processos físicos que dão origem a estes eventos, levou ao desenvolvimento de sistemas de alerta para grandes rampas, como é o caso do ELRAS (*Electricity Reliability Council Of Texas Large Ramp Alert System*). Este utiliza um sistema de assimilação de dados 3DVAR, que usa informação meteorológica como ponto de partida para um modelo NWP. Os resultados deste são usados em mecanismos de deteção que, por sua vez, são usados num modelo estatístico de mudança de regime [6].

Efetivamente, o uso de modelos estatísticos de mudança de regime (*state-transition* ou *regime-switching models*) constitui uma abordagem promissora à previsão de rampas, uma vez que as suas previsões de velocidade do vento se mostraram mais próximas dos valores medidos do que utilizando outras possibilidades (redes neuronais ou filtros de *Kalman*) [6]. Outros projetos de investigação utilizaram o momento e intensidade das rampas previstas como variáveis adicionais, produzindo intervalos de confiança melhorados para os quantis previstos [6]. A combinação e otimização de previsões resultantes tanto de modelos estatísticos, como de físicos, pode, ainda, revelar-se eficaz em termos de precisão dos resultados [22].

A WEPROG, desenvolveu uma ferramenta especializada em previsão de rampas, tendo chegado a algumas conclusões [23]:

- A diversidade dos modelos baseados em *ensembles* é uma medida importante da incerteza associada à amplitude e ao momento da rampa;
- *Ensembles* com bons erros RMSE são demasiado regulares/smooth para serem úteis na previsão de eventos de rampa extremos;
- Quanto maior o erro de fase (erro associado ao momento em que ocorre o evento) permitido na identificação de uma rampa acentuada, melhor é o sistema de previsão.

Atualmente, existe um projeto de previsão de rampas (“*Ramp Cast*”) desenvolvido numa parceria entre a *Xcel Energy*, a *NCAR* e a *Vaisala*. Para horizontes de 0 a 3 horas, são usados dados do vento medidos regionalmente, através de sistemas como o VDRAS (*Variational Doppler Radar Analysis System*). Quando se pretende prever para horizontes temporais de 3 a 60 horas, a utilização de um DMOS (*Dynamic MOS*) permite encontrar as melhores entradas a utilizar e reduzir o erro RMSE em cerca de 5 a 10% [6].

2.4 Avaliação dos Modelos de Previsão

As previsões de produção eólica possuem sempre uma incerteza elevada, devido à variabilidade do vento, pelo que se torna essencial avaliar os diferentes modelos de previsão. Assim, é possível determinar quais as abordagens mais adequadas para cada situação e perceber, ainda, o que caracteriza a incerteza que lhes está associada.

2.4.1 Fatores de Influência nos Erros de Previsão

Os modelos de previsão meteorológica numérica (NWP) são responsáveis pela maioria dos erros associados à previsão de produção eólica, que podem ser divididos em dois tipos [6]:

- Erros de amplitude, associados à severidade dos eventos meteorológicos;
- Erros de fase, associados ao momento em que ocorrem esses eventos no tempo.

Os erros de fase são mais difíceis de quantificar, enquanto que os erros de amplitude são relativamente fáceis de controlar através do uso de séries temporais. Neste sentido, têm vindo a ser desenvolvidos vários estudos com o objetivo de determinar quais os contributos de cada um dos tipos de erro, tendo-se concluído que os erros de fase são os que mais influenciam as previsões [6]. A conversão da potência de vento disponível em potência real varia de forma não-linear, devido à forma da função de transferência dos aerogeradores (ver Figura 2.4). Tipicamente, não existe produção abaixo de um certo valor mínimo de velocidade do vento, verificando-se um rápido crescimento desta até a máquina atingir o seu valor nominal de potência (por volta dos 15 m/s). A partir deste valor de vento, a produção mantém-se constante até ser atingida a velocidade de corte (por volta dos 25 m/s) [11].

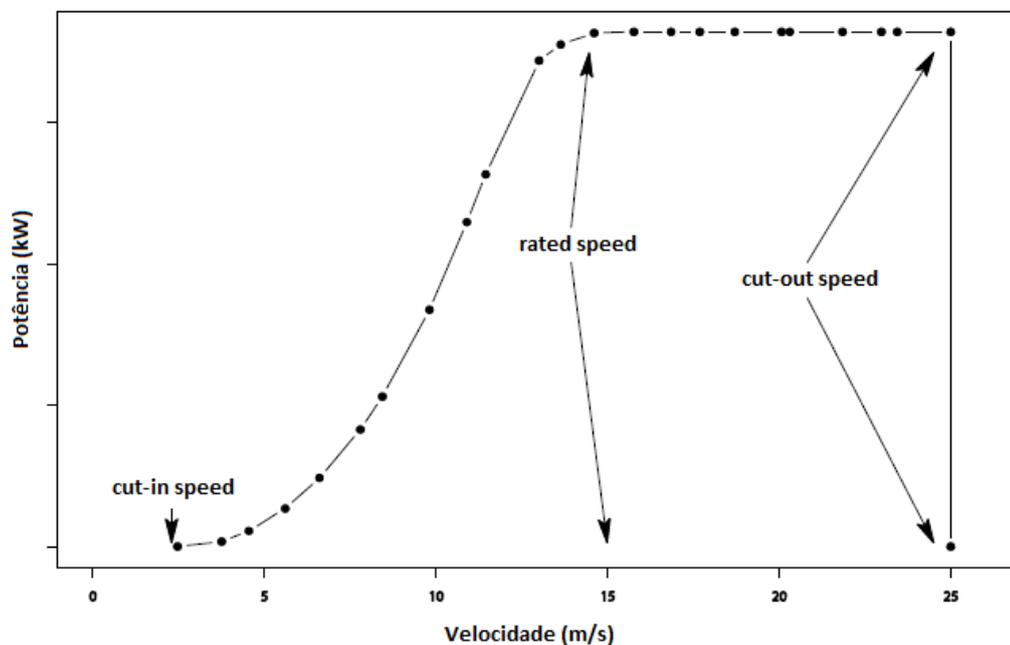


Figura 2.4: Curva de potência de um aerogerador [11]

Num dado parque eólico, a interação entre aerogeradores e as características locais fazem com que a função de transferência varie significativamente em função da direção do vento. A sua característica não-linear faz com que os erros de previsão sejam penalizados de forma heterogênea, ou seja, os erros tornam-se irrelevantes para velocidades muito reduzidas e para velocidades

compreendidas na zona constante da curva de potência. No caso de velocidades moderadas (aproximadamente de 3 a 12 m/s), os erros são altamente penalizados, uma vez que um pequeno erro na velocidade provoca um erro considerável na potência. Perto da velocidade de corte, dão-se, também, erros significativos, devido à mudança repentina do estado do sistema [11].

Um outro fator diretamente relacionado com os erros de previsão é a hora do dia para a qual esta é feita. Verificam-se erros mais elevados nas alturas em que o consumo é superior, geralmente ao início e ao fim do dia (horas de cheia/ponta) [8]. De forma menos significativa, também o horizonte de tempo, as características do terreno (como a rugosidade, p.e.), a localização geográfica, as condições climáticas e o tipo de abordagem utilizada (física, estatística, etc.) podem contribuir para o aumento dos erros.

2.4.2 Indicadores de Erros de Previsão

De forma a avaliar a qualidade dos diversos modelos, comparam-se as suas previsões diretamente com os valores reais medidos no local, para os mesmos intervalos de tempo. A análise do desvio verificado entre estes valores irá determinar a qualidade do modelo. Quando se pretende fazer esta avaliação em relação à componente estatística dos modelos, é necessária a análise de grandes quantidades de previsões, para que os resultados sejam viáveis [11].

O erro verificado num dado instante $t+k$ para uma previsão efetuada no instante t , denomina-se $e_{t+k|t}$ e pode ser obtido calculando a diferença entre o valor real, p_{t+k} , medido em $t+k$, e o valor previsto no instante t para o horizonte k , $P_{t+k|t}$, tal como na equação seguinte:

$$e_{t+k|t} = p_{t+k} - P_{t+k|t} \quad (2.9)$$

A partir deste valor de erro, existem diversas formas de quantificar a precisão de um sistema de previsão, sendo duas das mais comuns: o erro médio absoluto (MAE – *Mean Absolute Error*) e o erro médio quadrático (RMSE – *Root-Mean Square Error*).

O MAE é usado para medir a magnitude média dos erros, sendo definido de acordo com a equação 2.10.

$$MAE = \frac{1}{N} \sum_{i=1}^n |e_{t+k|t}| \quad (2.10)$$

Dividindo o erro de previsão pelo valor real medido no instante $t+k$ e multiplicando todo o somatório por 100, é possível obter a percentagem de erro médio absoluto (MAPE – *Mean Absolute Percentage Error*), tal como na equação 2.11.

$$MAPE = \frac{1}{N} \sum_{i=1}^n \left| \frac{e_{t+k|t}}{p_{t+k}} \right| \cdot 100 \quad (2.11)$$

A variação dos erros de previsão constitui, também, informação importante, sendo analisada através do RMSE, calculado da seguinte forma:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^n (e_{t+k|t})^2} \quad (2.12)$$

Enquanto que o MAE é expresso em percentagem da capacidade nominal da instalação, o RMSE é expresso como o desvio padrão dos erros da previsão [8]. De notar que os erros médios podem ser mal interpretados, caso existam valores positivos e negativos que originem uma média inferior à real [6].

Tipicamente, a superioridade de um modelo em relação a outro é mais evidenciada no valor RMSE do que no MAE, sendo que o RMSE atribui maior importância a erros maiores. Em determinadas situações, o cálculo do RMSE e do MAE podem originar informações opostas, caso possuam valores contrários. Assim, existem autores que distinguem 3 critérios de erro: “*As previsões de produção eólica devem ser próximas dos valores reais. A soma dos desvios entre as previsões e os valores reais deve ser pequena. As previsões devem resultar em baixos custos originados pelos erros de previsão.*” Enquanto os dois primeiros critérios se referem ao equilíbrio da rede elétrica, o último é importante no sentido de obter baixos custos na integração de geração eólica no mercado. No caso de parques eólicos com séries temporais variáveis diferentes, esta comparação torna-se mais complexa [6].

Devido à grande variedade de critérios de previsão ao longo da bibliografia, o projeto ANEMOS propôs-se a estabelecer um conjunto de medidas para avaliar o desempenho dos modelos, sendo estas: o MAE e o RMSE, tal como já referido, juntamente com o *bias*, o coeficiente de determinação R^2 , o indicador de desempenho da regressão para comparação com outros modelos e o histograma da distribuição do erro. É, também, aconselhada a divisão dos dados em conjunto de treino e conjunto de validação, bem como uma normalização dos erros médios em relação à capacidade instalada, para comparar entre diferentes parques eólicos [6].

2.4.3 Interpretação de Desempenho de Previsão

O gráfico seguinte (figura 2.5) traduz o valor do RMSE para diferentes horizontes de tempo e para diferentes modelos de previsão. Neste, a legenda *NewRef* é relativa ao modelo de referência usado e HWP é o equivalente ao modelo agora denominado *Prediktor*. O comportamento verificado para este é geral aos modelos de previsão a curto prazo, com apenas pequenas variações do erro RMSE e do seu comportamento ao longo do tempo. Atualmente, devido a aperfeiçoamentos dos modelos NWP, as previsões possuem erros RMSE na ordem dos 10% da capacidade instalada [6].

Tal como mencionado anteriormente, o modelo de persistência é utilizado como referência para comparar com o desempenho dos outros modelos de previsão. Para previsões a muito curto prazo (até algumas horas), este é o modelo que os outros têm de ultrapassar. No entanto, para

horizontes maiores, o modelo de persistência é consideravelmente pior do que o mais simples dos outros modelos.

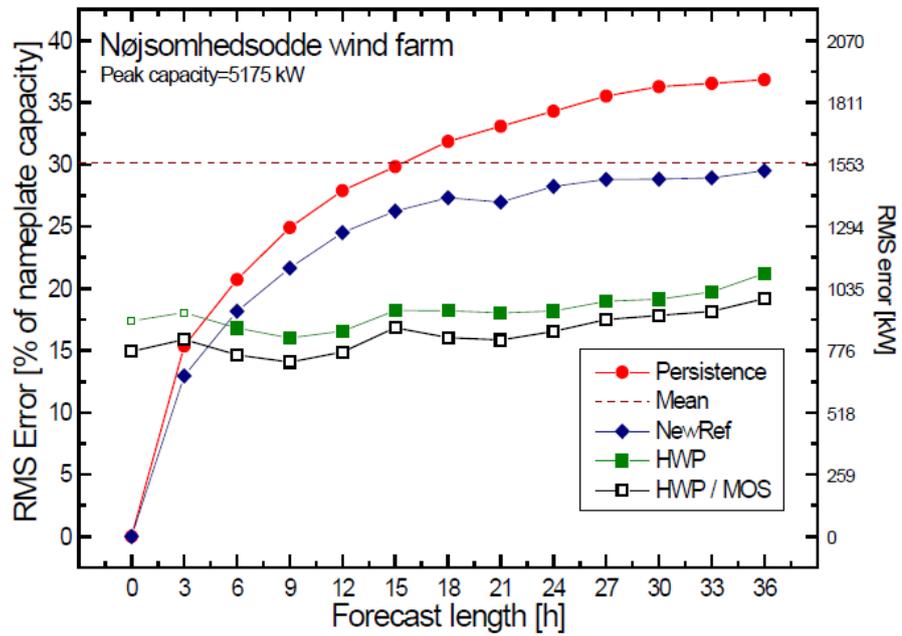


Figura 2.5: Variação do RMSE para diferentes horizontes temporais e modelos de previsão [6]

Capítulo 3

Processamento de Dados

A ilha de Creta possui o maior sistema de energia isolado da Grécia, constituído por três centrais térmicas com uma capacidade instalada total de 850 MW, mais de uma centena de parques fotovoltaicos com cerca de 65 MW e 26 parques eólicos com uma capacidade instalada total de 176,5 MW. No Verão, devido ao turismo, a procura chega a atingir picos de 650MW e, dependendo das condições meteorológicas, as fontes de energia renovável chegam a suprir até cerca de 50% desta necessidade [24].

A bibliografia estudada dá ênfase ao facto do sucesso de uma previsão de produção eólica estar fortemente dependente da qualidade dos dados utilizados na mesma. Assim, neste capítulo, é feita a análise e o processamento de toda a informação fornecida pela SiNGULAR e cedida pela Smartwatt, de modo a compreender qual a melhor forma de a utilizar nos modelos de previsão que se pretende desenvolver.

3.1 Informação Utilizada

O sistema de transmissão de Creta compreende 18 subestações no total, sendo que as descritas na tabela 3.1 são as analisadas em relação à produção de energia eólica, ao longo do trabalho. Os parques eólicos com o símbolo "@" são considerados parques controláveis, sujeitos a regulação, uma vez que, quando a produção total da ilha é excessiva relativamente ao consumo, a produção pode ser reduzida, proporcionalmente, em todos os parques deste tipo.

No global, são utilizados três tipos de dados, todos eles relativos aos anos de 2012 e 2013:

- Histórico da produção de energia de cada parque eólico;
- Velocidade de vento, medida em cada parque eólico;
- Previsões meteorológicas para diferentes pontos da ilha obtidas através de um sistema de previsão global (GFS), descrito no subcapítulo sobre modelos NWP do capítulo 2.

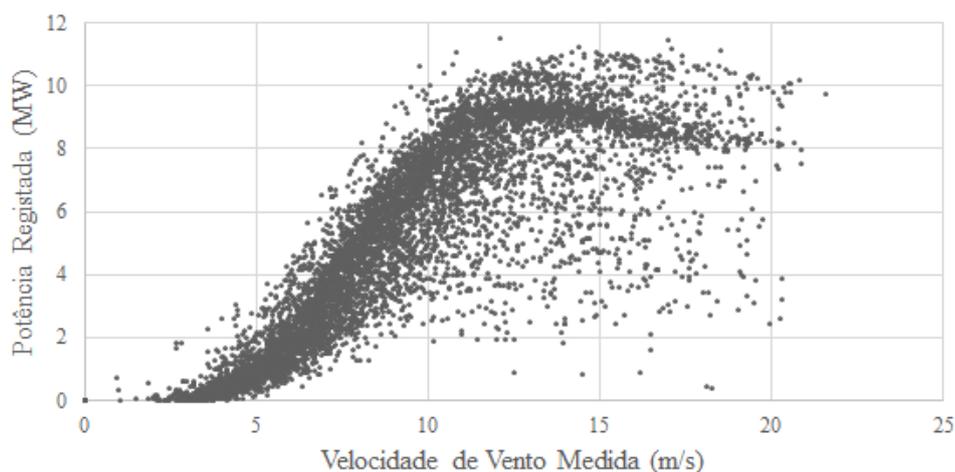
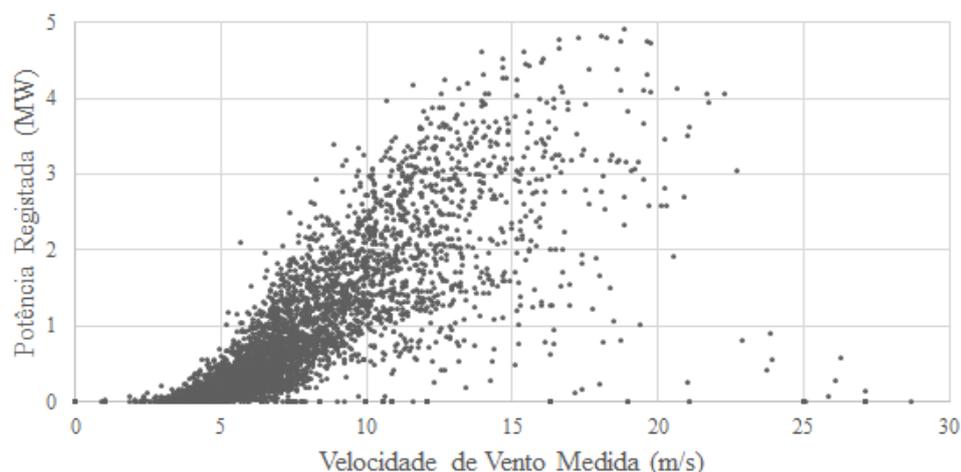
Para cada um destes, é necessário fazer uma análise da qualidade da informação, de modo a eliminar possíveis incorreções. No caso das previsões meteorológicas, estas serão estudadas no capítulo 4, quando é feita a escolha das variáveis de entrada a utilizar nos modelos.

Tabela 3.1: Subestações e parques eólicos em estudo

SUBESTAÇÃO	PARQUE EÓLICO	SIGLA	POTÊNCIA INSTALADA (MW)	UTILIZAÇÃO
SITEIAS	Rokas	RO	12,90	✓
	Aiolos	AI	10,00	✓
	Xirolimni	XH	10,20	-
	Enteka	ET	2,70	✓
	Honos Iweco	XO	4,50	✓
	Rokas (Modi2) @	R2	4,80	✓
	Xirolimni-DEH Ananeosimes	X2	3,00	✓
	OAS Siteias	AO	1,20	✓
MARONIAS	Axladia	AX	10,00	✓
	Krya	KR	10,00	✓
	Anemoessa	AM	5,00	✓
	Platuvola	WR	3,00	✓
	Enercon	EN	2,50	✓
RETHIMNOU	Kouloukonas @	KO	4,80	-
	Akoumia	AK	7,20	-
MOIRON	Iweco (Megali Vrysi)	IW	4,95	✓
	Moiron	MR	5,25	✓
	Diethnis Aioliki @	DI	7,20	✓
KASTELLIYOU	Ydroaioliki	YD	9,35	✓
	Anemos Alkionis @	AL	6,30	✓
	Mousouron @	MS	2,55	-
AGYIAS	Envitec Vardia	EV	5,40	✓
	Envitec Vatali	BT	5,40	✓
AG. VARVARAS	Tema	TE	14,45	-
TOTAL			113,45	

Relativamente aos restantes dados, a análise torna-se mais complexa. Em primeiro lugar, criaram-se curvas de potência para cada parque eólico (ver exemplos das figuras 3.1 e 3.2), colocando no eixo das ordenadas a soma das potências geradas pelos aerogeradores do parque e no eixo das abcissas a velocidade de vento média medida nesse mesmo local. Cada ponto representa um par (v, p) medido numa determinada hora do histórico considerado.

Desta forma, foi possível analisar a qualidade dos dados do histórico, fazendo a comparação com a curva de potência ideal (ver figura 2.4 do capítulo 2), e selecionar a informação relevante, excluindo alguns dos parques eólicos com demasiada percentagem de medidas incoerentes.

Figura 3.1: Curva de potência do parque eólico *Rokas*Figura 3.2: Curva de potência do parque eólico *Kouloukonas*

A figura 3.1 diz respeito ao parque eólico *Rokas*, da subestação *Siteias*, e constitui um caso com uma elevada taxa de pontos sem evidência de erros e, portanto, aceitável para treino de modelos de previsão corretos.

Por outro lado, na figura 3.2 estão os dados do parque eólico *Kouloukonas*, da subestação *Rethimnou*, onde só foi possível obter informação de vento para os últimos três meses de 2012, pelo que não se consegue definir a curva em toda a sua amplitude. Os parques em que ocorreram situações semelhantes a esta acabaram por não ser utilizados, uma vez que poderiam introduzir erros consideráveis nas previsões e dar origem a modelos menos precisos. É mais vantajoso não utilizar estes parques do que filtrar os seus erros, uma vez que, dessa forma, a perda de informação seria mais significativa, no total.

Efetivamente, não foram consideradas as subestações *Rethimnou* e *Ag. Varvaras*, nem os

parques eólicos *Mousouron* e *Xirolimni*, pertencentes às subestações de *Kastelliou* e *Siteias*, respectivamente. Os erros verificados nestes casos devem-se, maioritariamente, a incorreções nas medições da velocidade do vento efetuadas no local, tendo-se verificado valores demasiado elevados da velocidade do vento, ou valores constantes durante um longo período de tempo.

3.2 Metodologias Desenvolvidas

Ao analisar as curvas de potência de cada parque no subcapítulo anterior, verificou-se a existência de valores que não traduzem o comportamento real dos aerogeradores - *outliers*. Na figura 3.3, as zonas a vermelho assinalam esses pontos, que não devem integrar o conjunto de variáveis de entrada no desenvolvimento dos modelos de previsão. A zona A diz respeito a valores considerados errados, enquanto que a zona B compreende um conjunto de pontos improváveis de acontecer na realidade. Muitos dos pontos da zona B correspondem a estados de transição, onde a potência média calculada nesse instante resulta em valores intermédios incoerentes. Embora, estatisticamente, não estejam errados, estes pontos não podem ser utilizados pelos modelos de previsão.

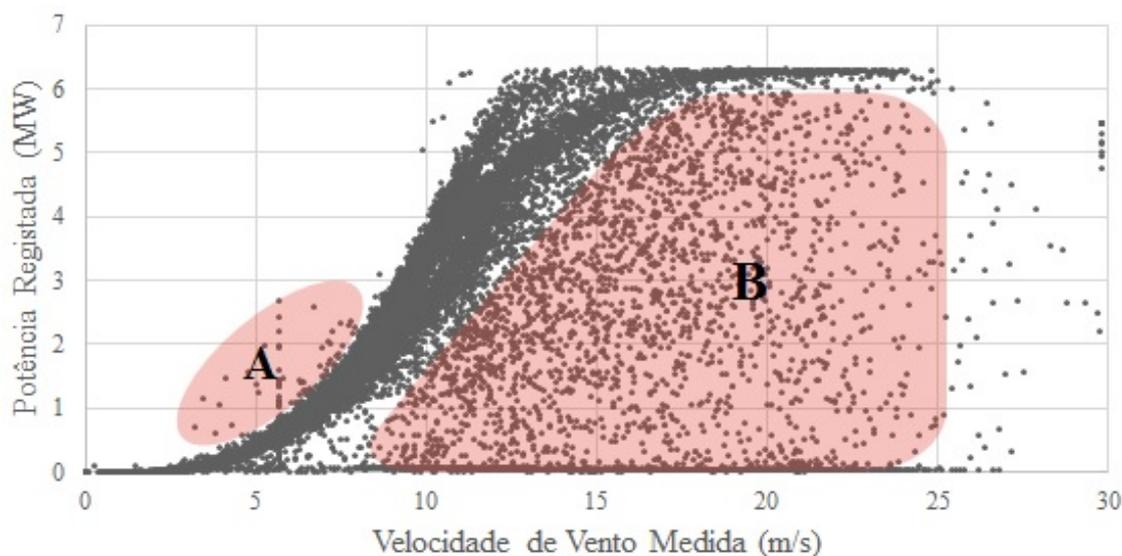


Figura 3.3: Curva de potência do parque eólico *Anemos Alkionis*, onde as zonas a vermelho assinalam pontos correspondentes a erros

Assim, feita a escolha dos parques eólicos a utilizar, foi necessário fazer um processamento dos dados, eliminando todos aqueles que não se pretende utilizar. Para tal, analisaram-se duas metodologias diferentes na tentativa de encontrar os melhores resultados possíveis. A primeira tira proveito da ferramenta de filtros do *software* de folha de cálculo, enquanto que a segunda faz uma aproximação das curvas de potência, diretamente.

3.2.1 Filtragem de Erros

A primeira metodologia testada baseia-se no princípio de eliminar as situações de erro, mantendo o maior número possível de pontos corretos, ou seja, mantendo todos aqueles que se encontram próximos do que seria de esperar.

Para tal, recorrendo à equação 3.1, criam-se duas funções sigmóide: uma superior e uma inferior à curva de potência. Estas são, então, modelizadas de acordo com as características da curva de cada parque eólico, fazendo o ajuste dos parâmetros Y_{min} , Y_{max} , X_{01} , X_{max1} , X_{02} , X_{max2} (ver figura 3.4) de maneira a obter a forma adequada para cada caso.

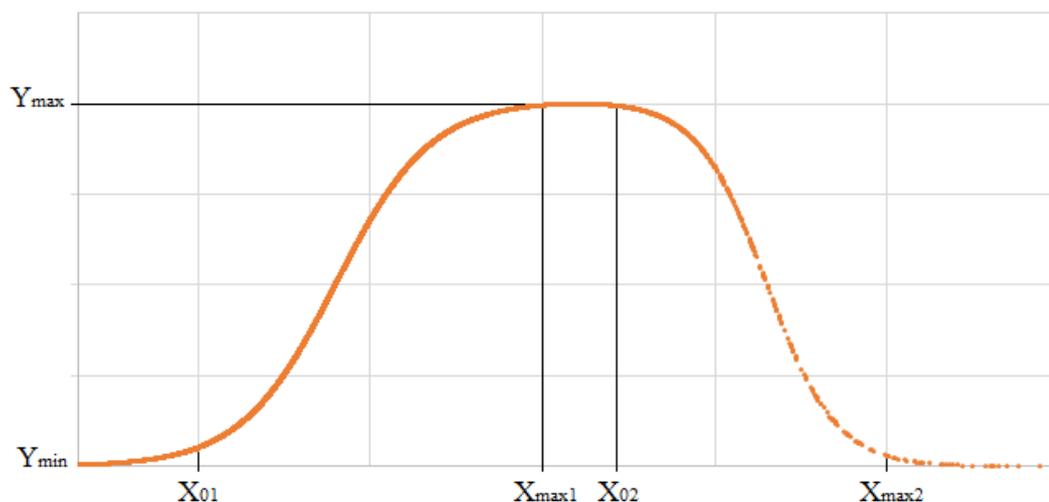


Figura 3.4: Representação gráfica de uma curva genérica criada segundo a equação 3.1

$$f(x) = Y_{min} + \frac{Y_{max} - Y_{min}}{1 + e^{-\frac{8}{t_{S1}} \cdot (v_{vento} - t_{C1})}} - \frac{Y_{max}}{1 + e^{-\frac{8}{t_{S2}} \cdot (v_{vento} - t_{C2})}} \quad (3.1)$$

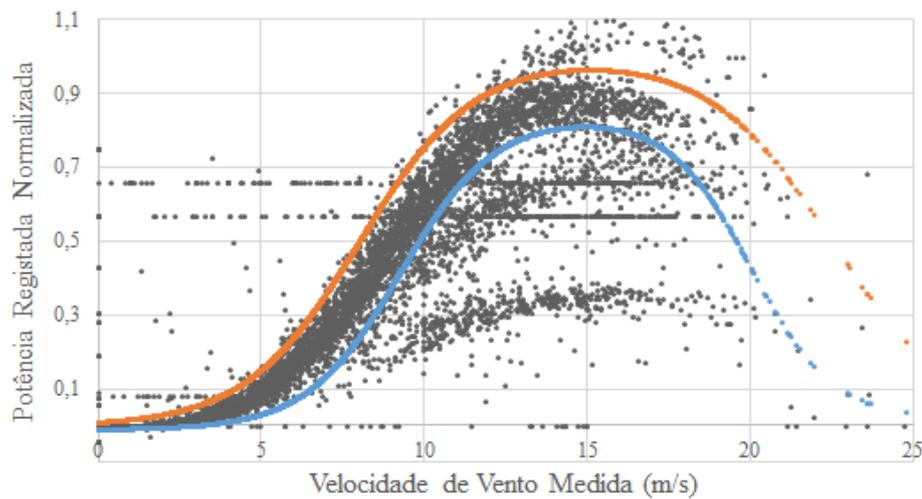
Onde,

$$t_{s1} = X_{max1} - X_{01} \quad t_{s2} = X_{max2} - X_{02} \quad (3.2)$$

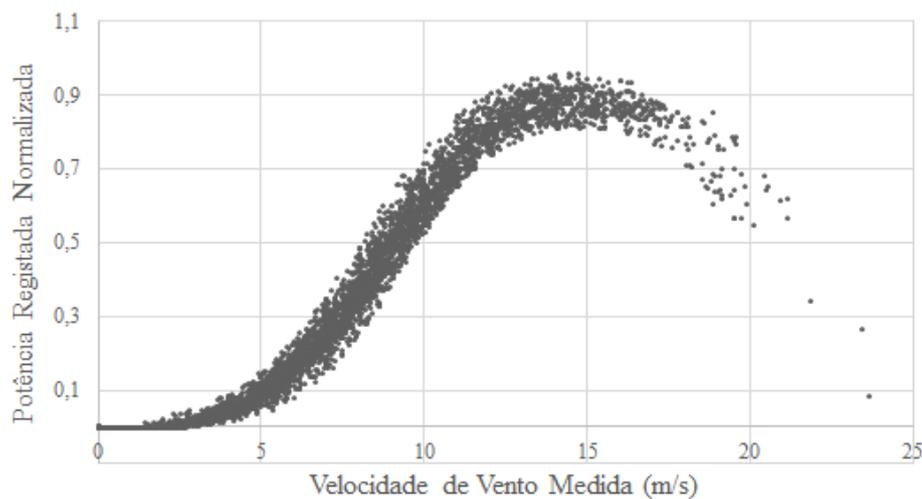
$$t_{C1} = \frac{X_{01} + X_{max1}}{2} \quad t_{C2} = \frac{X_{02} + X_{max2}}{2} \quad (3.3)$$

Assim, é possível filtrar todos os pontos que estejam fora dos limites estabelecidos pelas sigmóides, tal como se pode ver no exemplo da figura 3.5. É, ainda, perceptível o facto de, por vezes, apenas parte do parque eólico estar em funcionamento. Esses pontos são eliminados, uma vez que, para o treino dos modelos, apenas interessam os pontos em que o parque funciona na sua plenitude.

Uma vez que as potências produzidas em cada local são diferentes e de forma a uniformizar os valores, todo o processamento foi feito com dados normalizados. Por fim, estes são utilizados



(a)



(b)

Figura 3.5: Definição dos limites da curva de potência (3.5a) e resultado, após aplicação dos filtros (3.5b) - parque eólico *Enteka*, subestação *Siteias*

para fazer um agregado de todos os parques eólicos em estudo, descrito em maior pormenor no subcapítulo 3.3. É importante que os dados do agregado tenham a melhor qualidade possível, pelo que foi testada outra metodologia de forma a comparar resultados.

Numa fase inicial de análise dos dados recebidos, criou-se um documento, a enviar para a SiNGULAR, com o objetivo de discriminar as falhas de informação e anomalias encontradas, tendo-se descrito a filtragem de erros proposta (ver anexo A). Independentemente da metodologia selecionada para utilizar, posteriormente, no desenvolvimento dos modelos, não é necessário repetir este estudo.

3.2.2 Modelização da Curva de Potência Média

Outra abordagem utilizada consistiu em criar uma única sigmóide, que é modelizada de forma a aproximar-se o melhor possível da curva de potência média de cada aerogerador (ver exemplo da figura 3.6). Trata-se de uma transformação dos dados onde, para cada valor de velocidade do vento, passa a existir apenas um valor de potência eólica a ser produzida (equação 3.4).

$$(v, p) \rightarrow (v, p') \quad (3.4)$$

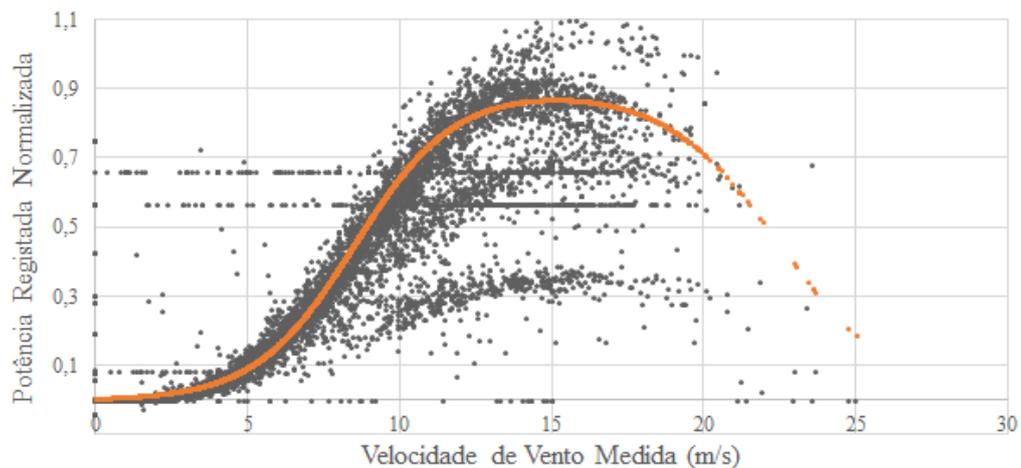


Figura 3.6: Modelização da curva de potência média do parque eólico *Enteka*, subestação *Siteias*

Uma possível desvantagem deste método é o facto de se considerar que os parques estão sempre a produzir de acordo com a curva, o que nem sempre é verdade. Isto pode vir a introduzir outro tipo de erros nos modelos finais.

De forma análoga à anterior, os dados tratados foram utilizados para obter um agregado de todos os parques e subestações em estudo, para posterior comparação.

3.3 Comparação de Metodologias

Com o objetivo de comparar ambas as metodologias de tratamento de dados, as potências foram, numa primeira fase, agregadas por subestações, utilizando a equação 3.5 para cada instante.

$$P_{sub} = \sum_{i=1}^n p_i \quad (3.5)$$

Onde n representa o número de parques da subestação e p_i a potência produzida pelo mesmo, no instante em questão. De forma semelhante, obteve-se, ainda, o agregado total dos dados de

produção:

$$P_{total} = \sum_{j=1}^N P_{sub.j} \quad (3.6)$$

Dividindo cada um destes valores pelo máximo verificado de P_{total} , obtêm-se as potências normalizadas.

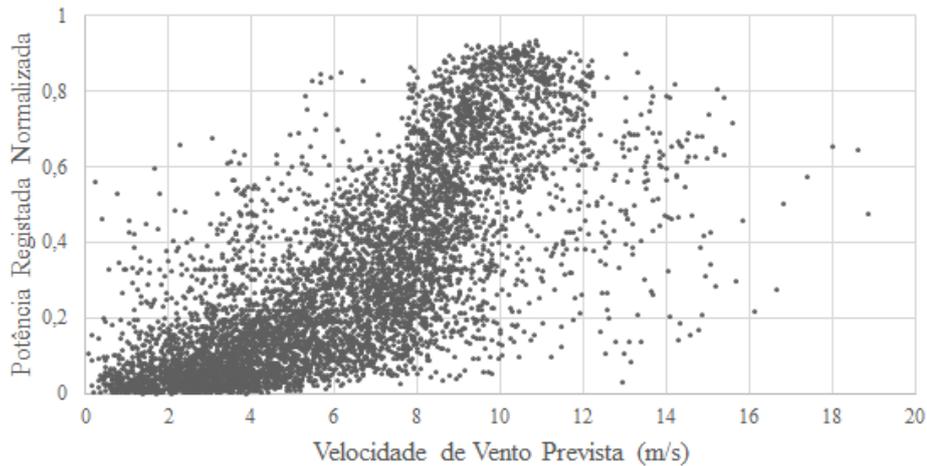


Figura 3.7: Curva de potência do agregado obtido através da 1ª metodologia - filtragem de erros

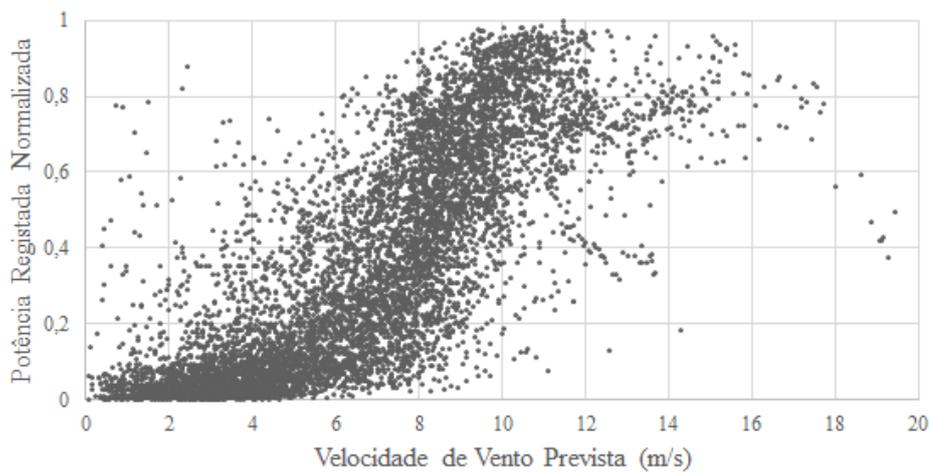


Figura 3.8: Curva de potência do agregado obtido através da 2ª metodologia - modelização da curva de potência média

É, então, necessária a verificação e comparação da qualidade dos dados obtidos através de cada um dos métodos, pelo que se criaram novas curvas de potência, desta vez colocando os valores agregados de potência em função da velocidade do vento prevista nas datas em questão (figuras 3.7 e 3.8).

Por análise visual, a metodologia que faz a modelização das curvas de potência média parece ser a que produz uma melhor curva de agregado, ou seja, mais próxima do esperado (ver figura 2.4 do capítulo 2) e com menor número de erros. No entanto, torna-se necessário comprovar este facto através da realização de previsões e cálculo das respetivas percentagens de erro médio absoluto (MAPE).

Efetivamente, efetuaram-se três previsões distintas para o mês de Julho de 2012, utilizando como variáveis de entrada diferentes dados:

- Previsão I: Utiliza todos os dados de produção, sem eliminação de erros;
- Previsão II: Utiliza o agregado obtido através da 1ª metodologia (filtragem de erros);
- Previsão III: Utiliza o agregado obtido através da 2ª metodologia (modelização da curva de potência média).

O valor de MAPE para cada uma das previsões encontra-se na tabela 3.2. Tal como se esperava, o agregado obtido segundo a 2ª metodologia produziu os melhores resultados.

Tabela 3.2: Percentagens de erro médio absoluto (MAPE) em função dos dados utilizados

	MAPE (%)
PREVISÃO I	24,04
PREVISÃO II	23,35
PREVISÃO III	22,42

Importa, ainda, referir que estas previsões foram realizadas, de forma alternativa, utilizando os valores de potência em unidades do Sistema Internacional (MW), tendo-se obtido valores de erro superiores. Assim, ao longo de todo o trabalho, optou-se pela utilização de potências normalizadas.

3.4 Conclusões

O caso de estudo considerado ao longo da dissertação, envolve uma grande quantidade de subestações e parques eólicos, pelo que se torna fundamental fazer um processamento da informação histórica a utilizar como *target*, no treino dos modelos. Uma vez que uma das maiores fontes de imprecisões na previsão de produção eólica está relacionada com a qualidade dos dados usados, foram desenvolvidas duas metodologias de eliminação de *outliers*, no que diz respeito aos dados de produção registados: filtragem de erros e modelização da curva de potência média. A segunda técnica revelou-se mais eficaz, dando origem à melhor curva de potência do agregado total e reduzindo o valor de MAPE das previsões em cerca de 2%, relativamente ao uso dos dados originais.

Assim, o agregado obtido através desta metodologia será o utilizado ao longo do trabalho, nomeadamente, no desenvolvimento de modelos de previsão. Os testes realizados neste capítulo, permitiram, ainda, concluir que se torna vantajoso utilizar valores de potência normalizados.

Capítulo 4

Desenvolvimento de Modelos de Previsão

Ao longo deste capítulo, é feito um estudo aprofundado de como alcançar um bom modelo de previsão, utilizando o *software* KDE, descrito anteriormente. O objetivo é o de obter o melhor modelo possível para o caso específico da ilha de Creta. Para tal, é utilizada uma abordagem essencialmente estatística, uma vez que as variáveis de entrada dizem respeito a dados históricos de produção e de condições meteorológicas do local.

4.1 Estrutura e Funcionamento

Tal como referido anteriormente, um modelo de previsão caracteriza-se por utilizar diversas variáveis para determinar valores de produção futuros. Pode apresentar graus de complexidade muito distintos, dependendo de como a informação for utilizada. De uma forma muito simples, esta pode ser dividida em três categorias:

- *Inputs*: entradas do sistema, geralmente, variáveis meteorológicas;
- *Targets*: valores registados da variável que se pretende prever;
- *Outputs*: saídas do sistema, neste caso, previsões da produção eólica.

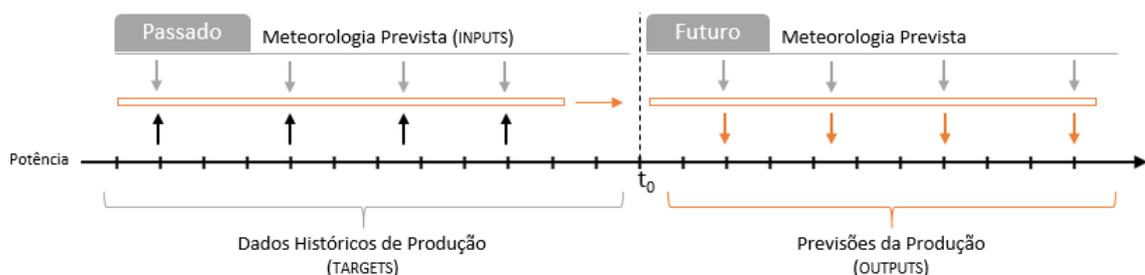


Figura 4.1: Representação do funcionamento genérico de um modelo de previsão estatístico

O esquema da figura 4.1 pretende representar, de uma forma geral, o funcionamento de um modelo de previsão estatístico e a forma como este utiliza os *inputs* e os *targets*. Neste, t_0 representa o instante atual em que as previsões estão a ser feitas.

Efetivamente, para obter previsões de produção eólica é necessário ter em atenção os dados de dois momentos temporais distintos: passado e futuro. Os dados referentes ao passado são utilizados como *input* para o desenvolvimento do modelo propriamente dito, dividindo-se, ainda, em conjunto de treino e conjunto de teste. Durante o treino, estabelecem-se relações entre dados históricos da produção eólica (*targets*) e diferentes variáveis de previsão meteorológica para os mesmos instantes, constituindo, assim, a base de conhecimento. O conjunto de teste, ou validação, é um conjunto de dados que não constam da base de conhecimento e para os quais é feita a verificação de desempenho do modelo.

Numa fase seguinte, o modelo combina essas mesmas variáveis meteorológicas, agora previstas para instantes futuros, com as relações já determinadas, de modo a prever os valores de produção para os instantes seguintes. Trata-se de uma abordagem que funciona bem para previsões a curto-prazo.

Os dados processados no capítulo anterior correspondem aos *targets* a ser utilizados ao longo do desenvolvimento dos modelos de previsão. É necessário, agora, escolher quais as variáveis a usar como *input* e, a partir daí, fazer os melhoramentos possíveis no modelo de previsão.

4.2 Escolha de Variáveis

A obtenção de um bom modelo de previsão é influenciada, principalmente, pelas variáveis de entrada escolhidas para o desenvolver. É fundamental que estas se relacionem com a informação que se pretende prever (*targets*), o que, neste caso, constitui uma dificuldade acrescida, uma vez que a produção de energia eólica apresenta grande variabilidade (ver figura 4.2).

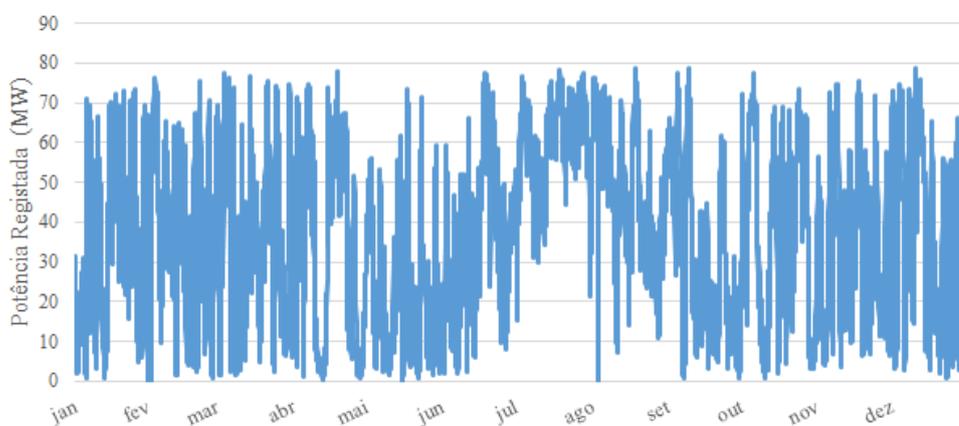


Figura 4.2: Representação dos *targets* a utilizar nos modelos de previsão, para o ano de 2013

Nesta dissertação, as variáveis explicatórias tratam-se de dados meteorológicos GFS referentes a Creta, obtidos através do centro de previsão americano NCEP (ver capítulo 2) para vários pontos da ilha, sendo eles:

- Velocidade do vento a 10m de altura;
- Direção do vento a 10m de altura;
- Velocidade do vento a 100m de altura;
- Direção do vento a 100m de altura;
- Temperatura.

Estes dados podem ser previstos para diferentes horizontes de tempo. Neste trabalho, utilizam-se previsões referente aos anos de 2012 e 2013, feitas diariamente para cada hora do dia seguinte - D (ver figura 4.3). Uma alternativa seria usar, também, previsões feitas para horizontes temporais mais longos, uma vez que se pretende obter um modelo que, operacional, faça previsões até D+7.

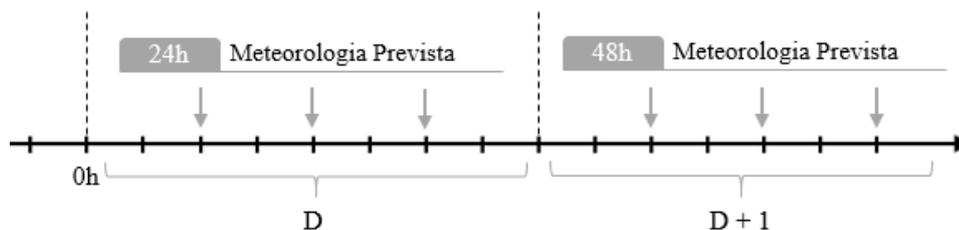


Figura 4.3: Esquema representativo dos horizontes temporais das previsões meteorológicas

Devido a erros ou problemas nos equipamentos, é frequente que grandes conjuntos de dados incluam algumas falhas de informação (*blanks*). Estas falhas obrigam a filtrar (eliminar) determinados dias, o que leva à perda de informação histórica a ser usada como entrada nos modelos. No caso das variáveis possuírem demasiados *blanks*, torna-se mais vantajoso não as incluir, em alternativa a filtrar uma grande quantidade de dados. Por essa razão, optou-se por não incluir as variáveis velocidade e direção do vento a 100 m de altura, uma vez que estas iriam eliminar cerca de 3 meses de dados.

Para além disso, é importante garantir a alta qualidade das variáveis que de facto se pretende utilizar: velocidade do vento (a 10m), direção do vento (a 10m) e temperatura. Assim, é necessário analisar individualmente cada uma destas, de modo a identificar possíveis anomalias e verificar a sua relevância para o modelo em questão.

As variáveis velocidade e direção do vento mostraram-se consistentes, tal como se pode ver nos gráficos da figura 4.4, onde estas estão representadas para um ano. Durante os meses de Verão, verifica-se que os dados são mais constantes, uma vez que as variações climatéricas são menos abruptas. Quanto à temperatura, por outro lado, torna-se menos evidente a sua possível

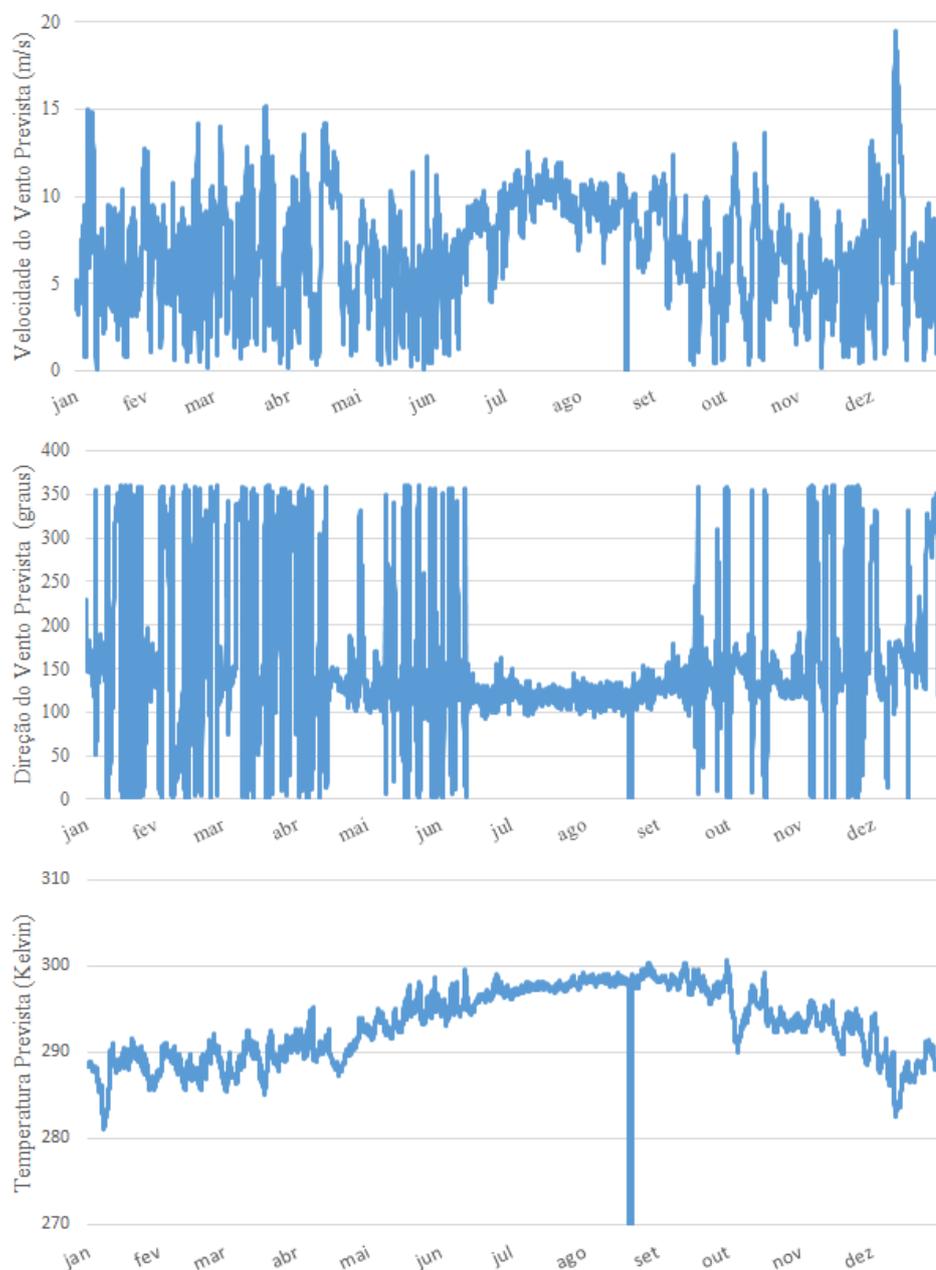


Figura 4.4: Representação das variáveis a utilizar nos modelos de previsão, para o ano de 2013

relação com a potência eólica produzida, o que pode vir a piorar a precisão dos modelos. Ainda assim, pode ser útil em estabelecer padrões sazonais e, até, diários, uma vez que durante o dia a temperatura tende a ser mais alta relativamente ao período noturno. É necessário, então, realizar previsões, considerando combinações destas três variáveis como entrada, de modo a concluir quais utilizar. Para além destas, também a variável "hora do dia" é incluída no estudo, uma vez que pode ser importante na modelização do padrão diário da produção eólica.

É, ainda, possível visualizar algumas das falhas já referidas, tal como acontece, por exemplo, no dia 20 de Agosto de 2013. Estas serão eliminadas do conjunto de entrada antes da realização

de previsões, sem que hajam consequências significativas em termos de perda de informação, uma vez que apenas ocorrem em curtos períodos de tempo (ver figura 4.5).



Figura 4.5: Velocidade de vento durante parte dos meses Março e Abril de 2012, com falhas (*blanks*) destacadas

Com todos os dados prontos, testaram-se vários modelos, utilizando diferentes variáveis de entrada. Tal como já referido, utilizaram-se como *targets* as potências normalizadas do agregado de todos os parques eólicos, descrito no capítulo anterior. Os dados referentes a 2012 foram usados para o treino dos modelos de previsão, enquanto que os de 2013 serviram como conjunto de teste dos mesmos. A tabela seguinte (tabela 4.1) expressa os resultados de cada modelo desenvolvido, em termos de valores de MAPE.

Tabela 4.1: Valores de MAPE para diferentes modelos, de acordo com as variáveis utilizadas para os criar

	VARIÁVEIS DE ENTRADA	MAPE (%)
MODELO I	Hora do dia Direção do vento	52,843
MODELO II	Hora do dia Velocidade do vento	33,269
MODELO III	Hora do dia Velocidade do vento Direção do vento	31,166
MODELO IV	Velocidade do vento Direção do vento	30,986
MODELO V	Hora do dia Velocidade do vento Direção do vento Temperatura	30,806
MODELO VI	Velocidade do vento Direção do vento Temperatura	30,591

A velocidade e direção do vento mostraram-se as variáveis mais relevantes, isto é, com maior poder explicativo, sendo que a inclusão da temperatura revelou, ainda, um pequeno melhoramento dos resultados. Efetivamente, utilizar estas três como entrada deu origem ao modelo mais preciso (Modelo VI). Quando se acrescenta a hora do dia, no entanto, o erro piora ligeiramente. Isto pode dever-se ao facto da temperatura já estabelecer, por si, um padrão relativamente à variação da produção diária, tornando-se repetitivo incluir uma variável como a hora do dia.

É, ainda, de notar o facto de algumas variáveis não possuírem relação suficiente entre si para serem utilizadas sem mais nenhuma outra, como é o caso da hora do dia e da direção do vento, onde os valores de MAPE das previsões se mostraram demasiado elevados (Modelo I).

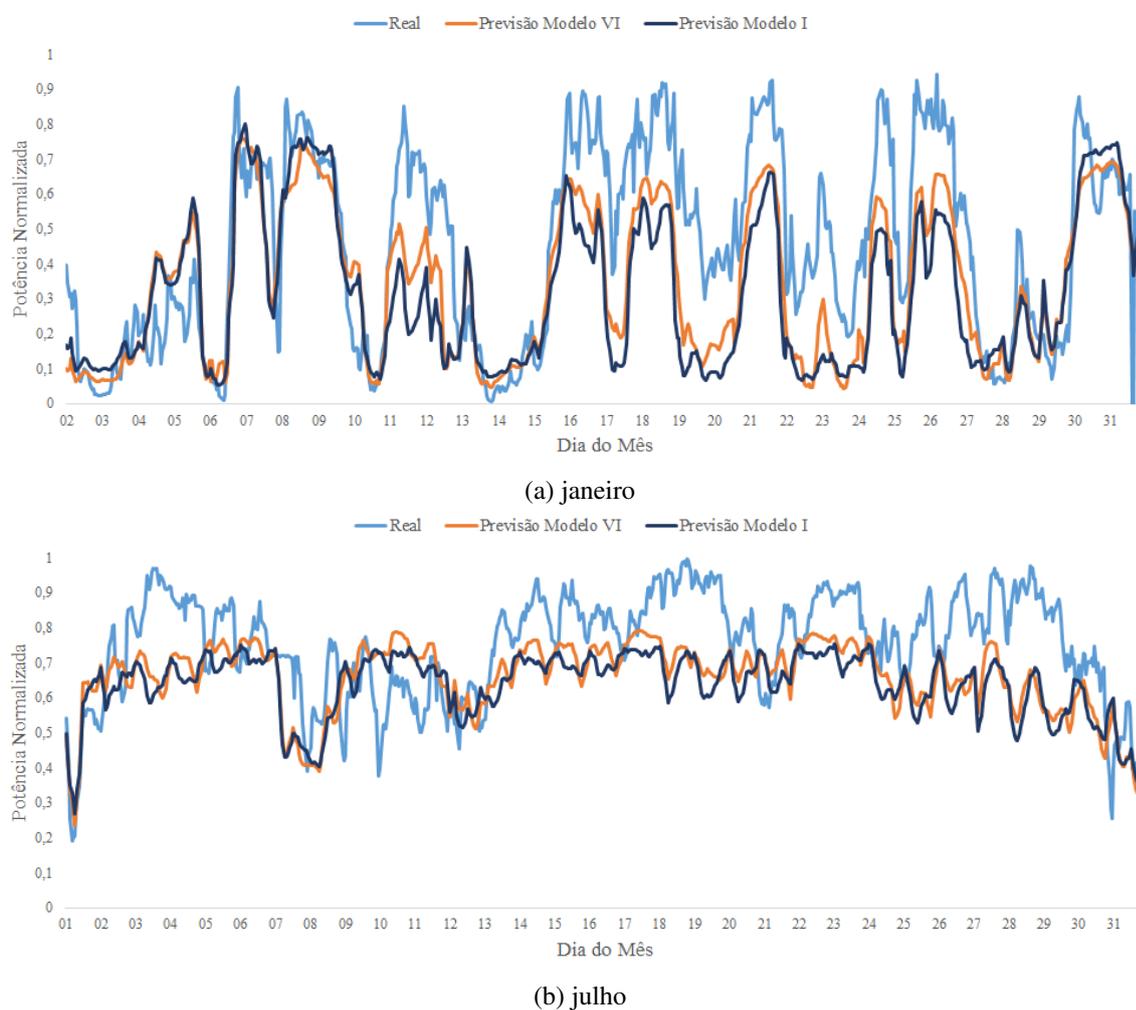


Figura 4.6: Comparação entre as previsões de produção eólica obtidas através do Modelo II e do Modelo VI, para dois meses de 2013

Os resultados apresentados correspondem a modelos preliminares que ainda não sofreram otimizações, daí os seus valores elevados. A imprecisão da previsão pode, também, ser agravada pelo facto de se estar a usar meteorologia proveniente de modelos globais, com pouca resolução, ou

pelas diferenças existentes entre período de treino e de teste. Efetivamente, tratam-se de resultados relevantes em comparações entre si.

É possível, também, verificar visualmente a diferença entre a qualidade das previsões dos diferentes modelos, comparando-as com os valores reais registados. Optou-se por comparar o Modelo II com o Modelo VI, uma vez que não possuem valores de MAPE demasiado próximos, nem exageradamente afastados. As figuras 4.6a e 4.6b apresentam os resultados para Janeiro e Julho de 2013. Tal como seria de esperar, as variações são mais bruscas no mês de inverno, o que torna mais fácil a análise visual. Por exemplo, é de salientar o dia 23 da figura 4.6a, onde o Modelo VI se revela substancialmente superior em termos de prever o pico de produção. Efetivamente, este encontra-se, na maioria das vezes, mais próximo dos valores reais registados.

Em suma, o melhor modelo obtido usou como variáveis explicatórias a velocidade de vento, direção de vento e temperatura (Modelo VI da tabela 4.1). Este será, então, estudado e desenvolvido ao longo dos subcapítulos seguintes.

4.3 Modelização de Incerteza

Um dos objetivos que se pretende alcançar com os modelos desenvolvidos, para além da previsão do valor central da produção (previsão determinística), é a incerteza a esta associada (previsão probabilística).

Para tal, continua-se a utilizar o *software* KDE, desenvolvido pela *Smartwatt*, que expressa a produção de potência eólica em termos de função densidade de probabilidade, através do princípio da densidade de Kernel estimada (ver capítulo 2). Após desenvolver um modelo de previsão neste *software*, obtêm-se as seguintes saídas (*outputs*) para cada instante:

- Previsão determinística: valor previsto da produção eólica;
- Parâmetro α da distribuição beta;
- Parâmetro β da distribuição beta;
- Mínimo da distribuição beta;
- Máximo da distribuição beta.

Efetivamente, obtêm-se uma distribuição de probabilidade beta na forma de quatro parâmetros (últimos quatro valores referidos). Assim, a partir da função inversa da distribuição beta, é possível calcula-la para cada instante de previsão. Desta forma, consegue-se uma representação da incerteza associada à previsão dos valores centrais de produção, tal como se pode ver nas figuras seguintes.

A margem de incerteza pode ser maior ou menor, dependendo do intervalo de confiança definido $[x; 1-x]$. No caso da figura 4.7, este foi de $[0,1; 0,9]$, o que significa que o valor de produção esperado estará dentro do intervalo (área cinzenta do gráfico) com uma probabilidade de 80%. Dependendo daquilo que o utilizador final pretender, este intervalo pode ser ajustado a qualquer

momento. A figura 4.8 apresenta exatamente as mesmas previsões, mas utilizando um intervalo de confiança mais estreito, onde os valores pontuais estarão dentro da margem de incerteza com uma probabilidade de 50%.

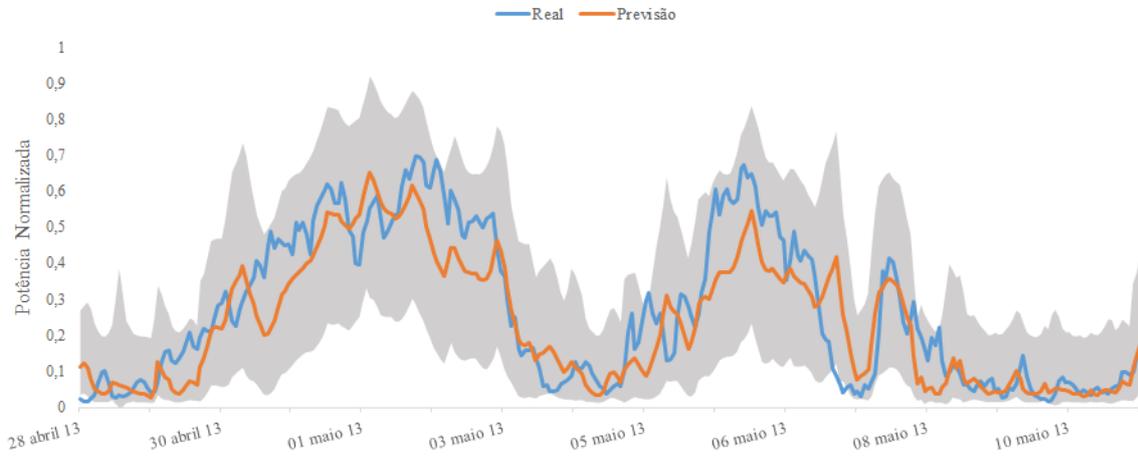


Figura 4.7: Previsão da produção eólica com representação de incerteza, com intervalo de confiança 80%, para uma semana de Maio de 2013

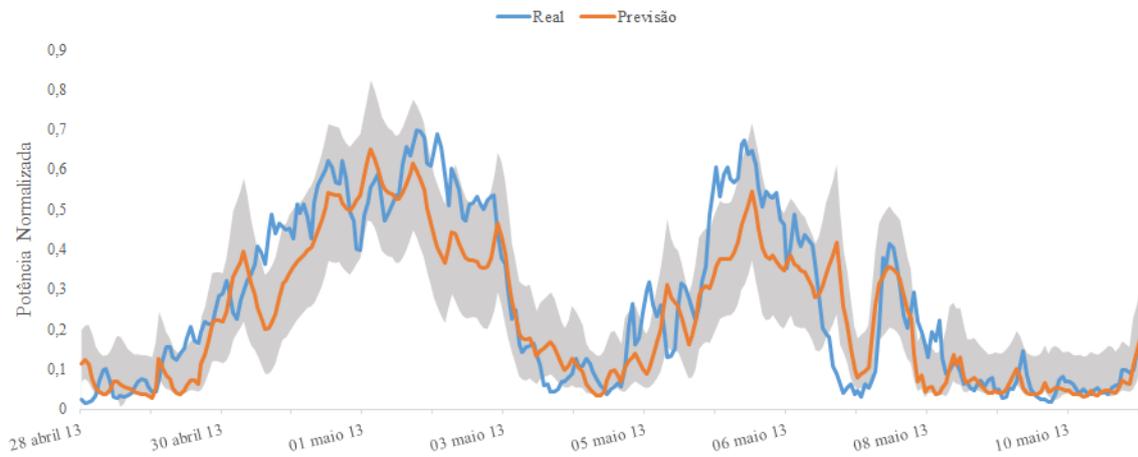


Figura 4.8: Previsão da produção eólica com representação de incerteza, com intervalo de confiança 50%, para uma semana de Maio de 2013

De um modo geral, uma representação da incerteza não é boa ou má, mas sim mais ou menos adequada [9]. Tem especial interesse no auxílio de tomadas de decisão com base em análise de risco, por exemplo, para determinar o despacho das centrais térmicas para o dia seguinte, tendo em conta diferentes cenários de previsão da produção eólica. Estas decisões, dependendo da situação, podem ser tomadas adotando um pensamento pessimista ou otimista em relação à produção eólica futura, com uma dada probabilidade conhecida.

No entanto, é possível efetuar cálculos que permitam perceber se a incerteza está a ser modelizada corretamente. Por exemplo, através da contagem de pontos dentro e fora do intervalo de

confiança estabelecido, foi possível desenhar gráficos como os apresentados na figura 4.9, usando os modelos descritos na tabela 4.1 do subcapítulo anterior. Nesta, são apresentadas as probabilidades, obtidas em teste, em função das probabilidades esperadas, para diferentes intervalos de confiança. Quanto mais próximos os modelos estiverem da linha de referência, mais correta é a sua modelização de incerteza. Isto é, se se pretende obter uma estimativa da incerteza com um intervalo de confiança de 80%, o modelo não deve incluir nessa estimativa nem pontos a mais, nem a menos, tal como acontece.

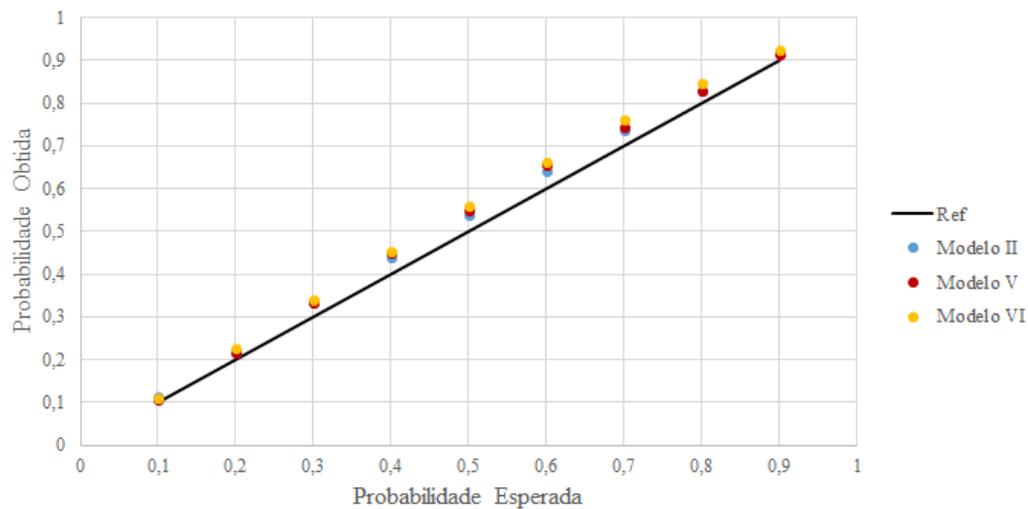


Figura 4.9: Probabilidade obtida em função da probabilidade esperada, para diferentes intervalos de confiança e através de diferentes modelos

De modo a tornar esta informação mais clara, é possível, também, obter os valores de MAPE da incerteza (tabela 4.2), quantificando a previsão probabilística. No anexo B encontram-se os cálculos intermédios efetuados.

Tabela 4.2: Valores de MAPE da previsão probabilística (incerteza), para os diferentes modelos criados

	MAPE INCERTEZA (%)
MODELO I	4,70
MODELO II	5,88
MODELO III	6,82
MODELO IV	8,51
MODELO V	6,57
MODELO VI	8,67

A análise da figura e tabela anteriores, permite concluir que os modelos considerados melhores na previsão determinística do subcapítulo 4.2, apresentam, agora, os piores resultados. Ou seja, a qualidade da previsão de incerteza varia de forma contrária à qualidade da previsão dos valores pontuais de produção de potência.

Em previsões probabilísticas, é essencial avaliar os resultados de acordo com a perspectiva do utilizador, ou seja, de acordo com o valor adicional (económico, técnico, etc) introduzido pela informação da incerteza, no problema em questão [7]. Assim, independentemente das ferramentas quantitativas utilizadas, a interpretação do valor das previsões probabilísticas vai ser sempre distinta de caso para caso.

4.4 Ajuste de Parâmetros

O *software* KDE, utilizado no desenvolvimento dos presentes modelos, necessita de receber valores de determinados parâmetros, de modo a realizar as previsões desejadas. A cada uma das variáveis de entrada utilizadas está associada uma distribuição Gaussiana, cujo valor do desvio padrão pode ser ajustado. Este, indica a dispersão dos dados em relação à sua média, definindo o nível da previsão probabilística que se pretende obter [25]. Ou seja, para uma determinada variável, caso se use um baixo desvio padrão, o modelo vai ser treinado tendo em conta apenas os dados mais próximos da média, naquele instante. Aumentando o desvio padrão, será considerada uma gama de valores mais alargada.

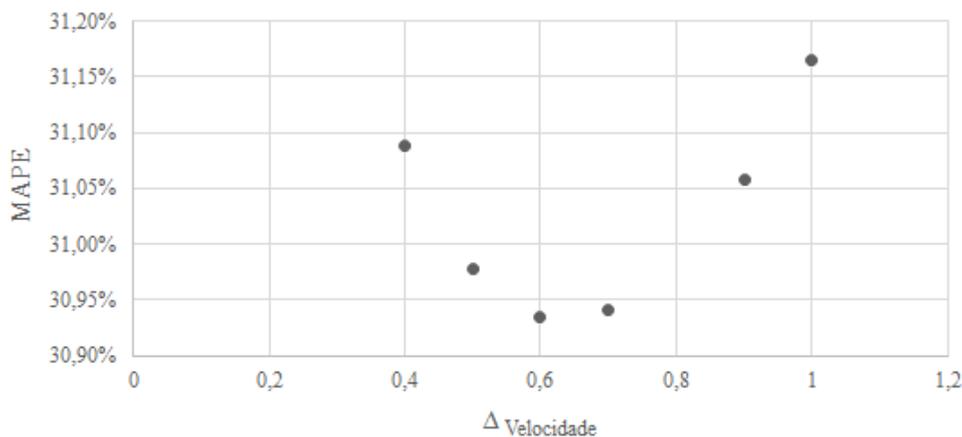


Figura 4.10: Valores de MAPE para diferentes valores de desvio padrão da velocidade de vento, no caso do Modelo III

Devido às características descritas, tratam-se de parâmetros que, por um lado, influenciam substancialmente os resultados das previsões e, por outro, são muito difíceis de determinar, variando de caso para caso. A figura 4.10 apresenta um exemplo da variação do valor de MAPE da previsão de produção, em função de diferentes valores de desvio padrão da variável velocidade de

vento. Assim, fazendo o ajuste do desvio padrão de 1 para 0,6, é possível diminuir o MAPE do Modelo III (ver tabela 4.1) de 31,166% para 30,935%.

Uma vez que o Modelo VI apresentou os melhores resultados de previsão determinística, procurou-se, seguidamente, fazer o ajuste dos seus desvios padrão, ou seja, para cada uma das variáveis que este utiliza (temperatura, velocidade e direção de vento). Este ajuste é realizado, tendo como objetivo a diminuição do valor de MAPE da previsão da produção eólica e estudando os seus efeitos na previsão probabilística.

Em primeiro lugar, alterou-se o desvio padrão da velocidade de vento, de modo a comparar com o ajuste feito para o Modelo III (figura 4.11).

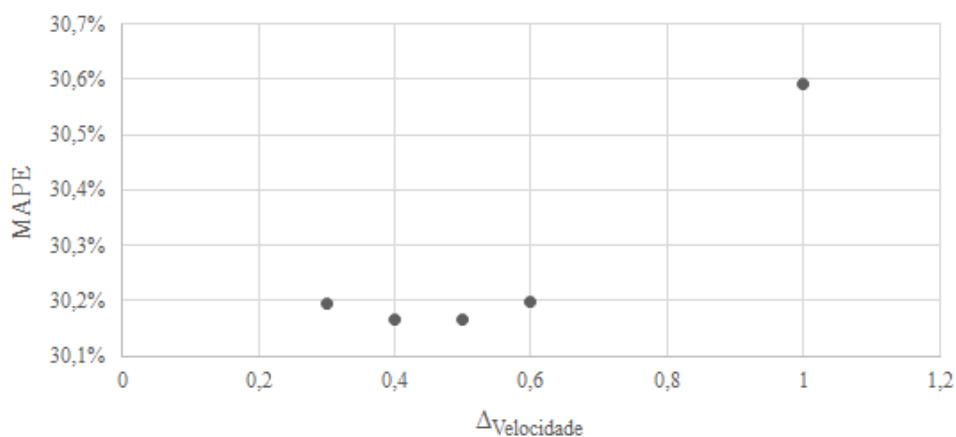


Figura 4.11: Valores de MAPE para diferentes valores de desvio padrão da velocidade de vento, no caso do Modelo VI

Anteriormente, no caso do Modelo III, 0,6 seria o valor que dava origem ao menor erro. No entanto, o mesmo não acontece com o Modelo VI, onde um desvio padrão de 0,4 produz previsões mais precisas. Na realização destas previsões, mantiveram-se os desvios padrão da direção de vento e da temperatura fixos em 23 e 5, respetivamente.

Estes resultados permitem concluir que, mudando as variáveis usadas pelo modelo, os seus desvios padrão ótimos são, também, alterados. Ou seja, para cada modelo é necessário fazer um novo ajuste de parâmetros de modo a atingir os melhores resultados possíveis.

Tendo o parâmetro da velocidade de vento definido, procedeu-se ao ajuste do desvio padrão da direção de vento e, de seguida, da temperatura. O processo é semelhante, estando os resultados para cada variável representados na figura 4.12.

A tabela 4.3 complementa esta informação, apresentando os diferentes valores de MAPE obtidos para cada combinação de desvios testada. Este estudo é feito em três fases distintas, onde se varia o valor do desvio padrão de uma das variáveis e se mantêm os restantes fixos. No caso da temperatura, concluiu-se que o valor usado inicialmente seria já o melhor, uma vez que, ao alterar este parâmetro, os resultados pioram.

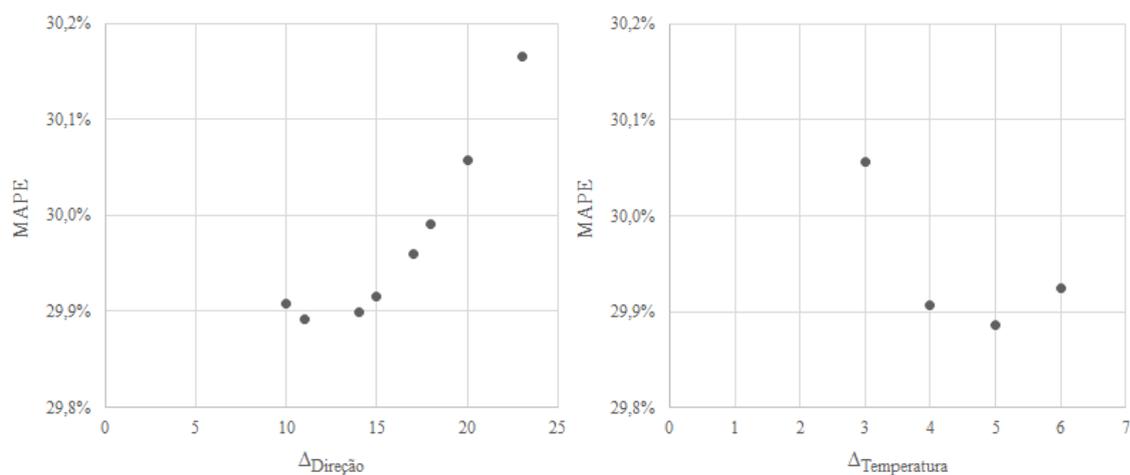


Figura 4.12: Valores de MAPE para diferentes valores de desvio padrão da direção de vento (à esquerda) e da temperatura (à direita), no caso do Modelo VI

Tabela 4.3: Valores de MAPE para diferentes combinações de valores de desvio padrão das variáveis de entrada do Modelo VI

	Δ VELOCIDADE	Δ DIREÇÃO	Δ TEMPERATURA	MAPE (%)
AJUSTE Δ VELOCIDADE	1	23	5	30,591
	0,6	23	5	30,198
	0,5	23	5	30,167
	0,4	23	5	30,165
	0,3	23	5	30,196
AJUSTE Δ DIREÇÃO	0,4	20	5	30,057
	0,4	18	5	29,991
	0,4	17	5	29,959
	0,4	15	5	29,915
	0,4	14	5	29,900
	0,4	12	5	29,887
	0,4	11	5	29,892
	0,4	10	5	29,908
AJUSTE Δ TEMPERATURA	0,4	12	6	29,924
	0,4	12	4	29,907
	0,4	12	3	30,056

A combinação que deu origem a previsões com menor erro utilizou, então, os desvios padrão 0,4, 12 e 5 das variáveis de entrada velocidade, direção e temperatura, respectivamente. No total, foi possível diminuir o MAPE da previsão determinística do Modelo VI de 30,591% para 29,887%.

Resta, fazer uma análise relativa à modelização da incerteza, de modo a verificar se a previsão probabilística sofreu, também, melhorias (figura 4.13).

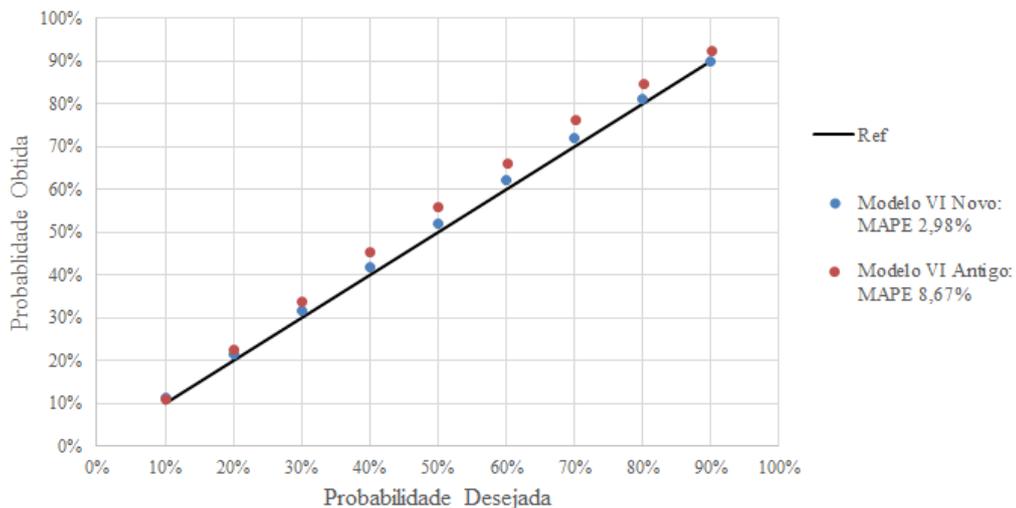


Figura 4.13: Representação da qualidade da incerteza, comparando o Modelo VI melhorado com o anterior

Efetivamente, a melhoria do valor de MAPE relativo à incerteza é, ainda, superior à verificada para a previsão determinística, passando de 8,67% para 2,98%. A tabela 4.4 resume, de forma quantitativa, o aperfeiçoamento conseguido.

Tabela 4.4: Valores de MAPE da previsão determinística e da previsão probabilística, antes e após o ajuste dos parâmetros do modelo selecionado

	PRÉ-AJUSTE	PÓS-AJUSTE
MAPE PONTUAL	30,591 %	29,887 %
MAPE INCERTEZA	8,67 %	2,98 %

Com as variáveis e os parâmetros de entrada otimizados, o modelo encontra-se construído e pronto a ser utilizado. No entanto, após a realização das previsões, é possível, ainda, melhorar a precisão dos valores obtidos, através do método descrito no subcapítulo seguinte.

4.5 Ajuste Dinâmico

Realizadas e analisadas as previsões de produção eólica, procuraram-se formas de melhorar, ainda, os seus valores. De facto, um dos problemas mais frequentemente verificado nas previsões realizadas, está relacionado com a amplitude de produção, uma vez que os valores previstos raramente atingem os máximos registados na realidade, momento em que os aerogeradores estão

a produzir à potência nominal. Assim, tornou-se conveniente testar um último ajuste de valores, com o intuito de melhorar o erro final da previsão determinística.

Através da construção de um gráfico de dispersão entre as previsões obtidas e os valores reais, é possível desenhar uma linha de tendência e verificar em que aspeto esta poderia modelizar mais corretamente a produção. De seguida, é criada uma função que pode ser adaptada de forma a corresponder ao pretendido, dando origem a um novo conjunto de previsões - ajuste dinâmico (ver esquema da figura 4.14).

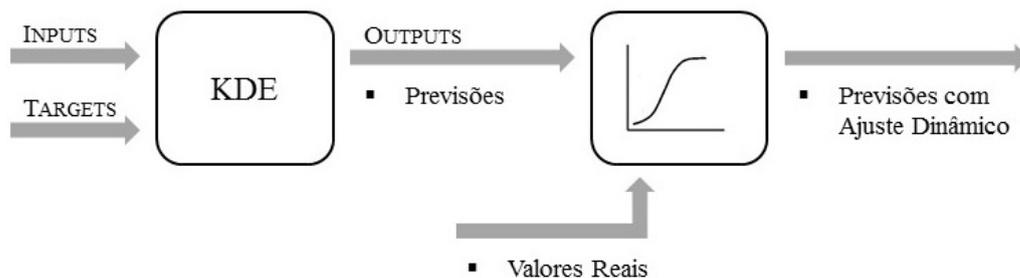


Figura 4.14: Esquema representativo do método de ajuste dinâmico

Na figura 4.15, este método é aplicado para a primeira metade de 2013. Efetivamente, considerando as previsões ajustadas de acordo com a função gerada, o valor de MAPE passa de 31,851% para 30,471%.

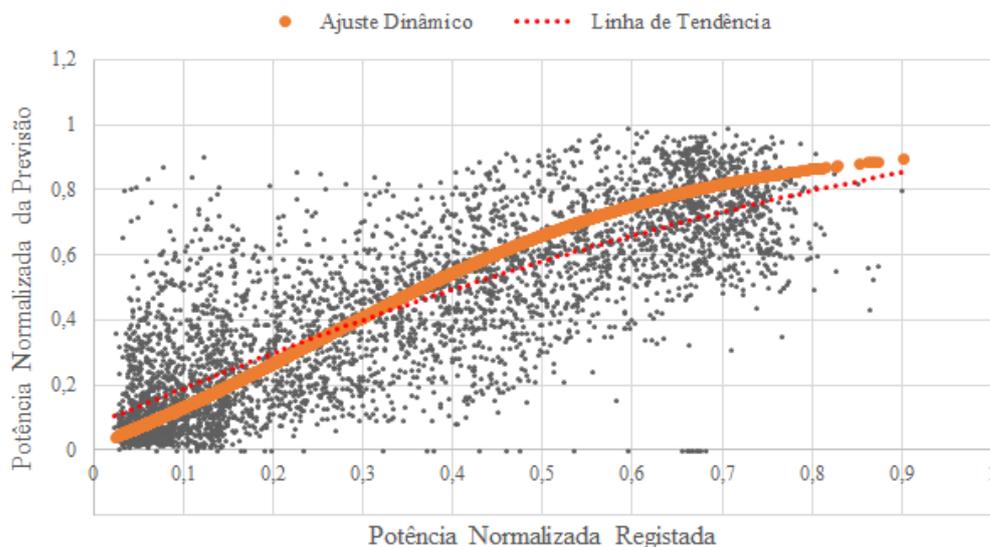


Figura 4.15: Linha de tendência do gráfico de dispersão entre a previsão e os valores reais de produção eólica, com função de ajuste dinâmico

Resta testar se esta função pode, então, ser utilizada nos *outputs* do modelo em previsões futuras, numa tentativa de as melhorar. Para tal, aplicou-se a função modelizada segundo os dados da primeira metade de 2013, à segunda metade do mesmo ano. Efetivamente, o valor de MAPE

pontual destes últimos 6 meses apresenta valores bastante mais satisfatórios. No entanto, a modelização da incerteza piora consideravelmente (tabela 4.5).

Tabela 4.5: Valores de MAPE da previsão determinística e da previsão probabilística para a segunda metade de 2013, antes e após o ajuste dinâmico dos resultados

	PREVISÃO KDE	AJUSTE DINÂMICO
MAPE PONTUAL	28,290 %	24,626 %
MAPE INCERTEZA	1,12 %	25,34 %

Este ajuste pode ser aplicado apenas à previsão determinística, deixando os valores probabilísticos inalterados. Desta forma, os resultados permitem concluir que a função de ajuste dinâmico constitui uma técnica com valor significativo no aperfeiçoamento de valores pontuais esperados de produção. Pode ser aplicada em cada previsão, utilizando, por exemplo, os dados previstos e registados no mês anterior.

4.6 Conclusões

Ao longo deste capítulo foram desenvolvidos e testados modelos com diferentes características, o que permitiu tirar conclusões sobre o processo de construção dos mesmos.

Para atingir a melhor solução possível, numa primeira fase, criaram-se modelos utilizando diferentes variáveis explicativas como entrada (*inputs*), de modo a concluir quais dariam origem a previsões de produção eólica (*outputs*) com erro inferior. Paralelamente à avaliação da previsão determinística, estudou-se, também, a modelização de incerteza (previsão probabilística) por parte de cada um dos seis modelos criados.

Uma das conclusões a extrair dos resultados obtidos desta forma, é o facto de, em geral, à medida que a qualidade das previsões pontuais do valor esperado vai aumentando, a modelização da incerteza vai piorando. Desta forma, deve-se optar pelo modelo mais indicado, dependendo de cada situação em específico e dependendo daquilo que o utilizador pretender. Na verdade, a importância das previsões probabilísticas prende-se, essencialmente, com a forma como estas serão aplicadas em problemas de tomadas de decisão.

Neste trabalho, optou-se por dar prioridade à qualidade das previsões pontuais de produção eólica relativamente à incerteza, pelo que se procurou melhorar o Modelo VI. Para tal, testaram-se diversos valores de desvio padrão de cada uma das variáveis utilizadas pelo mesmo, até encontrar as parametrizações ótimas. Neste processo, concluiu-se que os parâmetros necessitam de ser ajustados para cada modelo, individualmente, uma vez que variam conforme as variáveis e o conjunto de dados utilizado.

Com o modelo terminado, é, ainda, possível fazer um último melhoramento dos resultados obtidos. Comparando as previsões realizadas para um passado recente com os valores reais registrados, é possível concluir em que aspectos a previsão falha e, desta forma, criar uma função que permite fazer o ajuste dinâmico dos valores de potência previstos, diminuindo o erro. A função pode, então, ser utilizada em previsões futuras, de modo a melhorá-las, tendo em atenção o horizonte temporal pretendido, que não deve ser demasiado longo. Este processo foi testado, tendo-se concluído que, apesar de se ter conseguido melhorar o MAPE pontual do valor esperado de produção, a previsão de incerteza apresenta erros bastante superiores. Desta forma, o ajuste dinâmico pode ser relevante, apenas, no caso das previsões determinísticas.

Capítulo 5

Avaliação do Modelo Final

Neste capítulo, é feita uma análise dos resultados obtidos através do modelo desenvolvido em KDE, anteriormente. De modo a verificar o bom funcionamento deste, é feito um estudo das previsões de produção eólica e de incerteza adquiridas para todo o ano de 2013. É, também, feita uma análise das diferenças que este apresenta, relativamente a alguns dos outros modelos desenvolvidos. Posteriormente, torna-se relevante, ainda, a comparação com os resultados adquiridos através de outras técnicas de previsão, como é o caso das redes neuronais e da persistência. Por fim, são expostas possíveis aplicações de interesse, relativamente ao modelo criado.

5.1 Análise de Resultados

O modelo final, selecionado como o melhor no capítulo anterior, compreende as especificações e resultados seguintes (tabela 5.1).

Tabela 5.1: Características do modelo de previsão selecionado para implementação

MODELO FINAL	
VARIÁVEIS DE ENTRADA	Velocidade de Vento Direção de Vento Temperatura
PARÂMETRIZAÇÕES	Δ Velocidade = 0,4 Δ Direção = 12 Δ Temperatura = 5
MAPE PONTUAL	29,887 %
MAPE INCERTEZA	2,98 %

Estes valores dizem respeito a previsões feitas para todo o ano de 2013, de modo a obter estimações de erro mais corretas. No entanto, com o intuito de facilitar a análise de resultados, torna-se conveniente observar espaços de tempo mais curtos. Desta forma, escolheu-se um mês

de inverno - janeiro - e um de verão - julho -, para estudar as previsões obtidas, com um intervalo de confiança de 80%, no que toca à incerteza. Estes valores não englobam a técnica de ajuste dinâmico descrita, uma vez que se pretendem examinar os *outputs* diretos do modelo.

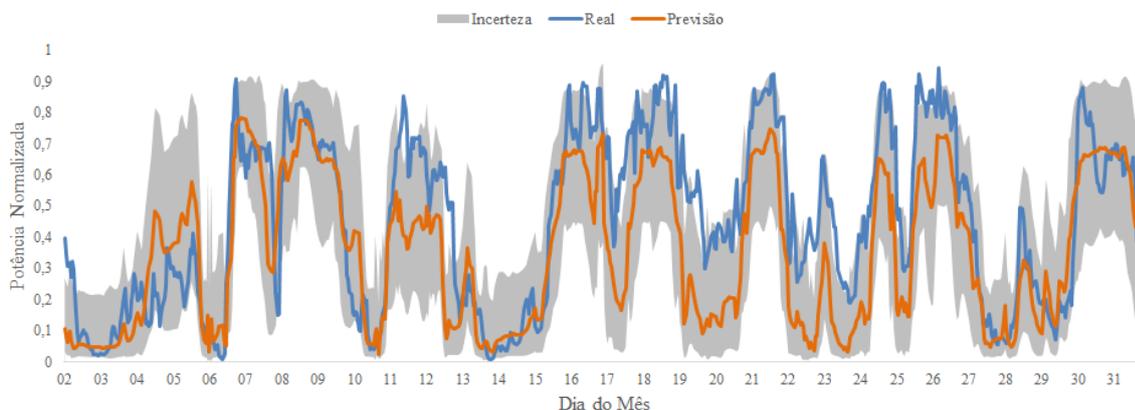


Figura 5.1: Previsão da produção eólica com representação de incerteza para o mês de janeiro de 2013 (MAPE=32,532%)

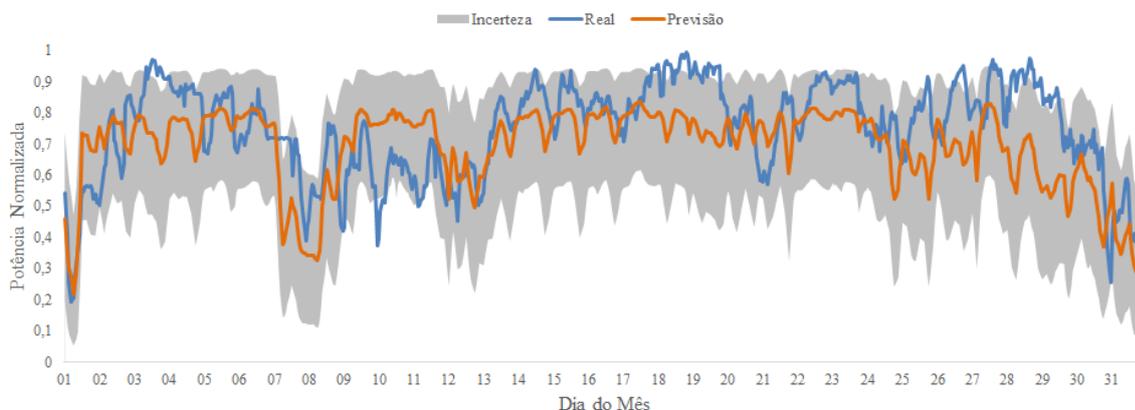


Figura 5.2: Previsão da produção eólica com representação de incerteza para o mês de julho de 2013 (MAPE=16,138%)

Por observação das figuras 5.1 e 5.2, comprova-se a sazonalidade da produção de potência eólica. A variabilidade do vento é substancialmente superior em janeiro, pelo que, na estação do ano que este representa, o erro é, também, significativamente mais elevado.

Fazendo a comparação entre os valores reais registados e os valores previstos, é possível, ainda, tirar conclusões sobre os tipos de erro existentes. Os erros de amplitude são os mais evidentes, estando mais acentuados no mês de inverno. Efetivamente, existem instantes em que a previsão não alcança minimamente os picos de produção, tal como acontece, por exemplo, entre o dia 30 e 31 de janeiro (figura 5.3).

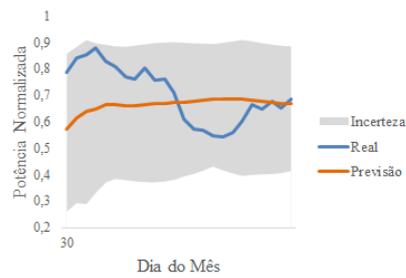


Figura 5.3: Exemplo de erro de amplitude (janeiro)

Quanto aos erros de fase, estes são mais visíveis no mês de verão, como é o caso da queda de produção que ocorre no dia 5 de julho, prevista pelo modelo para acontecer horas antes (figura 5.4). Verifica-se este mesmo efeito várias vezes ao longo de todo o mês (dias 22, 25 e 31, por exemplo).

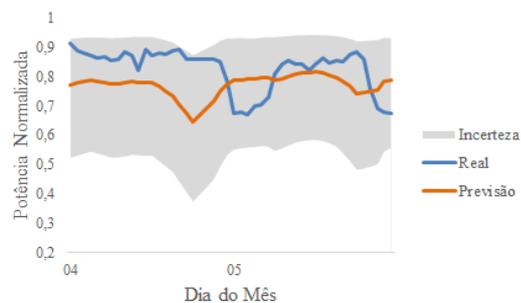


Figura 5.4: Exemplo de erro de fase (julho)

O mais provável é que os erros sejam provenientes, maioritariamente, das variáveis meteorológicas previstas, dado que se utilizaram modelos NWP globais, com baixa resolução. No entanto, outras falhas possíveis podem estar associadas, por exemplo, a erros de medição relacionados com o sistema SCADA. Caso os dados do conjunto de treino e do conjunto de teste sejam muito distintos, a previsão perderá qualidade, também (figura 5.5).

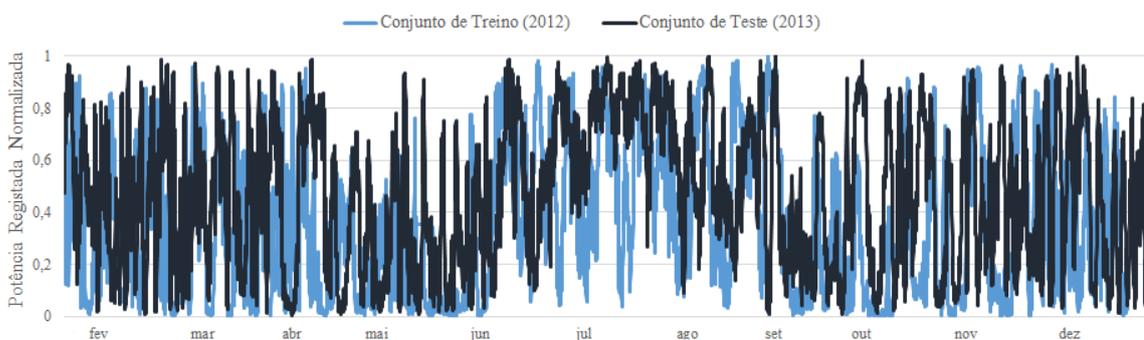


Figura 5.5: Comparação entre a produção registrada em 2012 e em 2013

Efetivamente, verificam-se algumas diferenças entre a produção eólica registada no ano de 2012 (usada como treino) e no ano de 2013 (usada como teste), o que leva à introdução de uma certa quantidade de imprecisões no sistema de previsão.

De modo a analisar a evolução diária do módulo do erro, calculou-se a média para cada hora do dia, utilizando os dados normalizados de todo o ano. Por observação dos resultados da figura 5.6, é possível verificar que durante a noite existem menos erros, relacionando-se com o facto de, nesse intervalo, a produção ser, também, inferior. Tal como seria de esperar, quanto maiores os valores de potência, maiores os erros a que as previsões estão sujeitas.

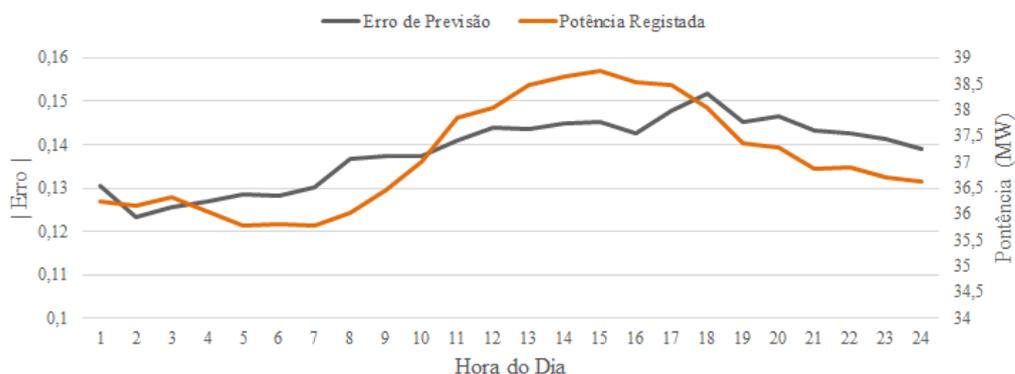


Figura 5.6: Evolução média do erro e da produção de potência eólica, ao longo do dia

A figura 5.7 apresenta a previsão das probabilidades obtidas através do modelo final desenvolvido, em cada intervalo de confiança, comparando-as com os valores de probabilidade esperados.

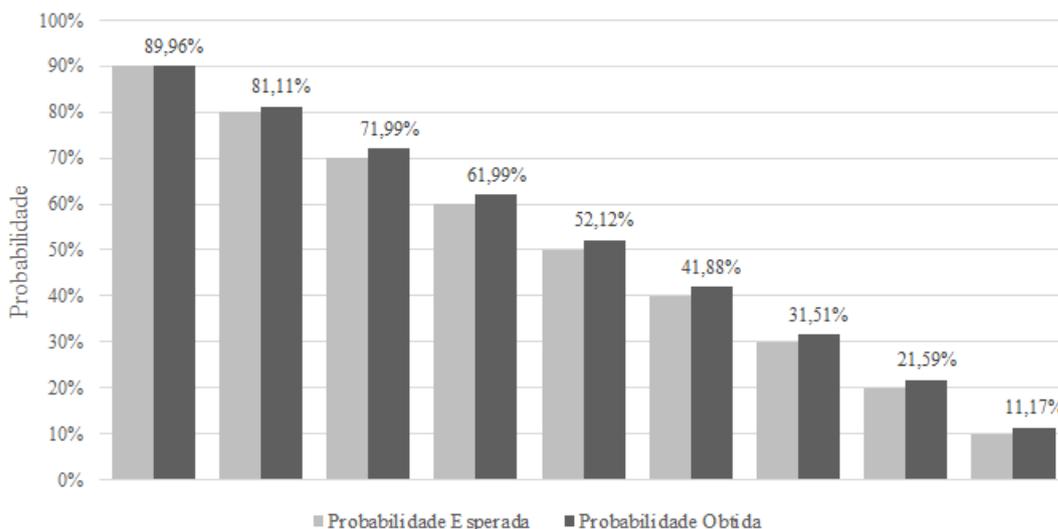


Figura 5.7: Representação dos valores de incerteza obtidos, para cada intervalo de confiança

A qualidade da incerteza, tal como referido no capítulo anterior, torna-se mais difícil de quantificar relativamente às previsões pontuais de produção. No entanto, é possível, para um deter-

minado intervalo de confiança pretendido, quantificar a percentagem de pontos que se encontram dentro da gama de incerteza prevista. O ideal é que não hajam nem mais, nem menos pontos dentro desse intervalo, para além do valor pretendido. A figura 5.7 apresenta, precisamente, esse cálculo. A diferença entre a probabilidade esperada e aquela que foi obtida pelo modelo é pouco significativa, pelo que se pode considerar que as decisões baseadas em análise de risco não ficam comprometidas.

Através desta informação é possível, ainda, calcular os valores de MAPE das previsões probabilísticas, voltando a considerar o mês de janeiro e o de julho. Mais uma vez se verifica que a qualidade da incerteza varia de forma contrária à qualidade da previsão pontual da produção (tabela 5.2).

Tabela 5.2: Valores de MAPE de ambos os tipos de previsão, para um mês de Inverno (janeiro) e um de Verão (julho)

	JANEIRO	JULHO
MAPE PONTUAL	32,532 %	16,138 %
MAPE INCERTEZA	2,47 %	5,01 %

Aplicando o processo de ajuste dinâmico, descrito no capítulo 4, aos valores pontuais previstos para todo o ano de 2013, reduziu-se o MAPE de 29,887% para 27,857%. A comparação entre estes resultados e as potências reais registadas, permite comprovar a diferença considerável na precisão dos valores (ver exemplo da figura 5.8). Sem o ajuste dinâmico, o modelo de previsão dificilmente atinge os picos de produção em toda a sua amplitude.

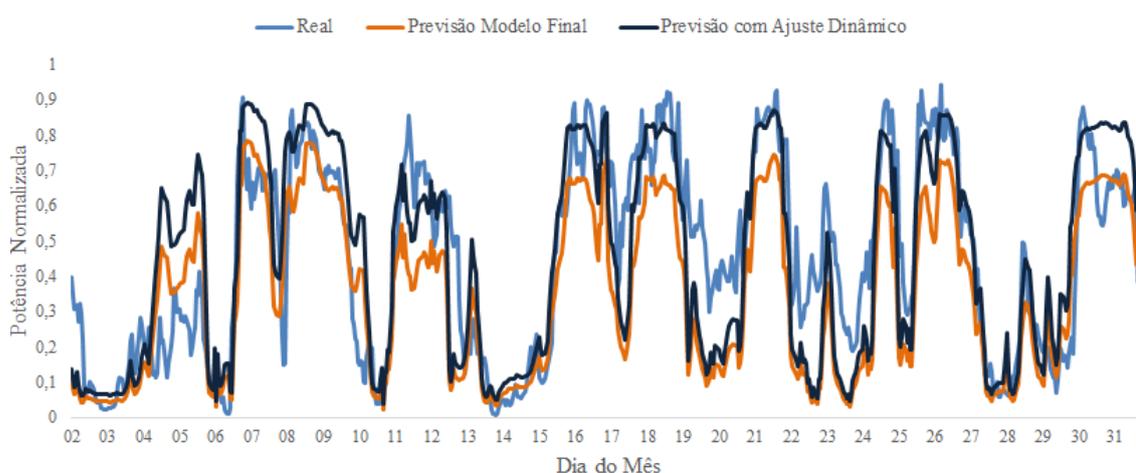


Figura 5.8: Comparação entre as previsões de produção eólica, obtidas através do modelo final, antes e após a realização do seu ajuste dinâmico, para o mês de Janeiro de 2013

5.2 Comparação entre Modelos

A análise realizada para o modelo final, mostra-se relevante, também, em comparação com alguns dos outros modelos desenvolvidos. Tal como referido durante a sua construção, à medida que se consegue aumentar a qualidade das previsões determinísticas, o oposto acontece com a qualidade da previsão de incerteza (tabela 5.3). Apenas o Modelo V se distancia ligeiramente deste padrão, uma vez que apresenta valores razoáveis em ambos os MAPE. Estes resultados dizem respeito a 2012 como período de treino e 2013 como período de teste, constituindo uma previsão da produção do agregado total de parques eólicos.

Tabela 5.3: Valores de MAPE da previsão determinística (pontual) e da previsão probabilística (incerteza) para os diferentes modelos criados

	VARIÁVEIS DE ENTRADA	MAPE PONTUAL (%)	MAPE INCERTEZA (%)
MODELO I	Hora do dia Direção do vento	52,843	4,70
MODELO II	Hora do dia Velocidade do vento	33,269	5,88
MODELO III	Hora do dia Velocidade do vento Direção do vento	31,166	6,82
MODELO IV	Velocidade do vento Direção do vento	30,986	8,51
MODELO V	Hora do dia Velocidade do vento Direção do vento Temperatura	30,806	6,57
MODELO VI	Velocidade do vento Direção do vento Temperatura	30,591	8,67
MODELO FINAL	Velocidade do vento Direção do vento Temperatura	29,887	2,98

Existem, ainda, disparidades significativas que se verificam por observação dos resultados. Por exemplo, no caso do Modelo VI, são bastante evidentes as diferenças que resultam do ajuste das suas parametrizações, dando origem ao modelo final, principalmente no que diz respeito à previsão probabilística (ver figura 5.9). Estas previsões, realizadas para o mês de janeiro, são, especialmente, distintas em termos de limite inferior da incerteza, onde é bastante visível a melhoria da precisão alcançada nas previsões probabilísticas.

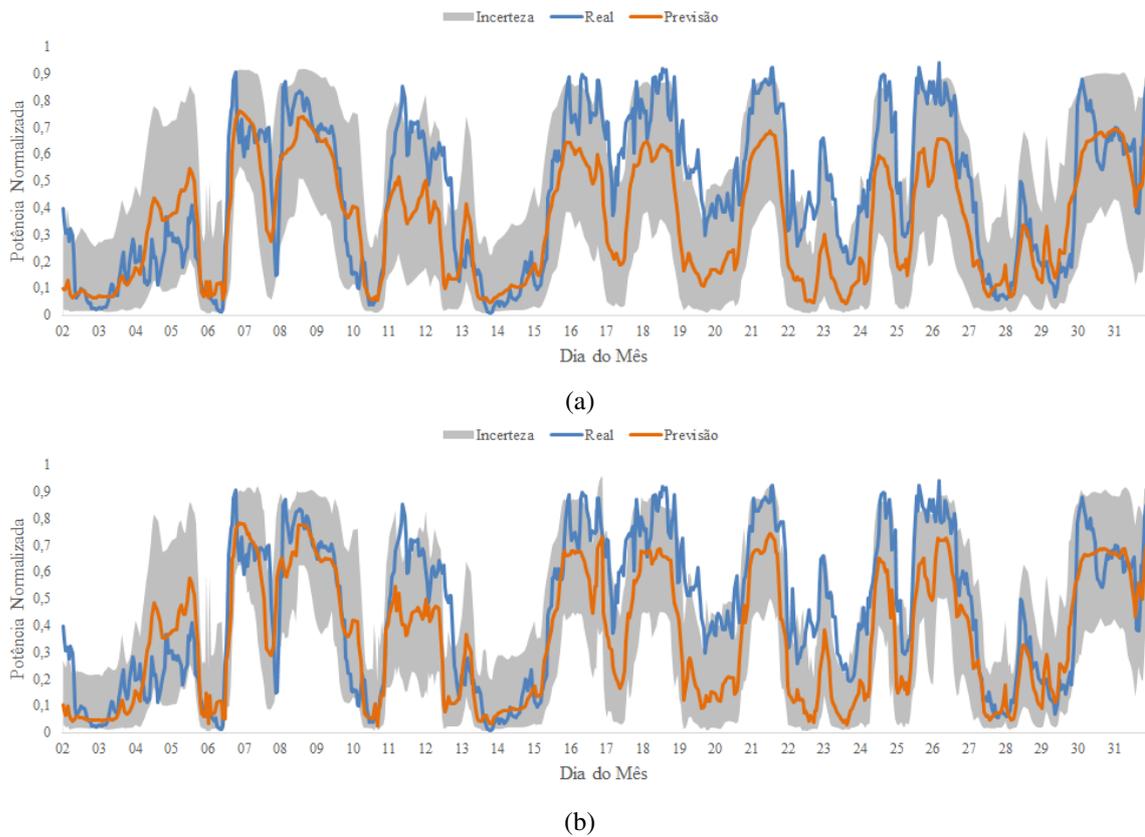


Figura 5.9: Previsões do Modelo VI antes (5.9a) e após (5.9b) o ajuste das suas parametrizações

A grande diferença entre modelos, no entanto, continua a residir nas previsões pontuais do valor esperado de produção, onde as disparidades são maiores e mais simples de verificar, tal como acontece por observação da figura seguinte.

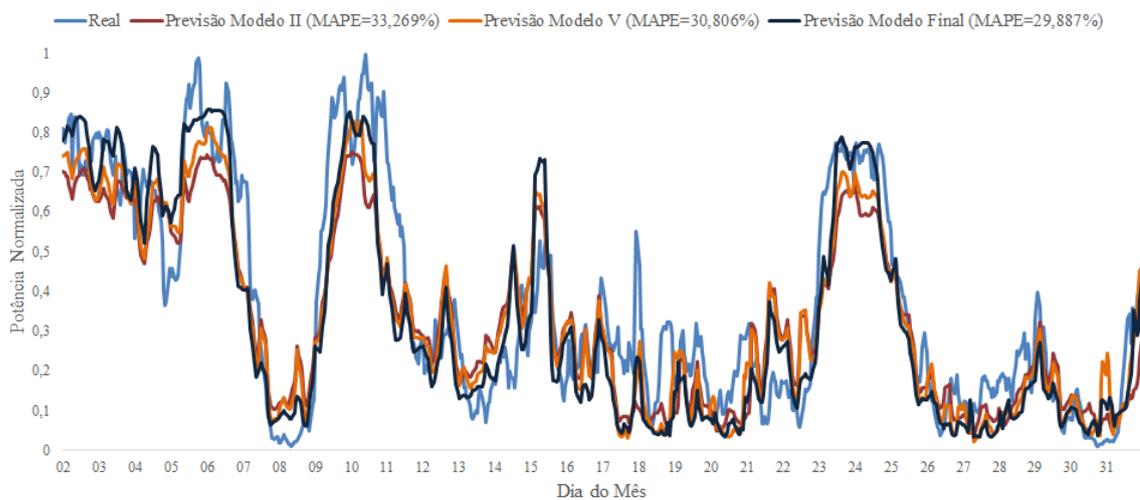


Figura 5.10: Comparação entre as previsões de produção eólica obtidas através de diferentes modelos, para o mês de Setembro de 2013

No entanto, a análise de resultados por observação de apenas determinado mês, como é o caso da figura 5.10, revela-se insuficiente, uma vez que este pode não ser representativo de todo o ano previsto. Neste sentido, torna-se relevante perceber de que modo é que o MAPE varia ao longo dos meses, tendo-se realizado esse cálculo para cada um dos modelos (ver figura 5.11). Apenas o Modelo I não foi considerado nas análises seguintes, uma vez que apresentou um desempenho significativamente baixo, não possuindo, portanto, relevância na comparação de resultados entre modelos.

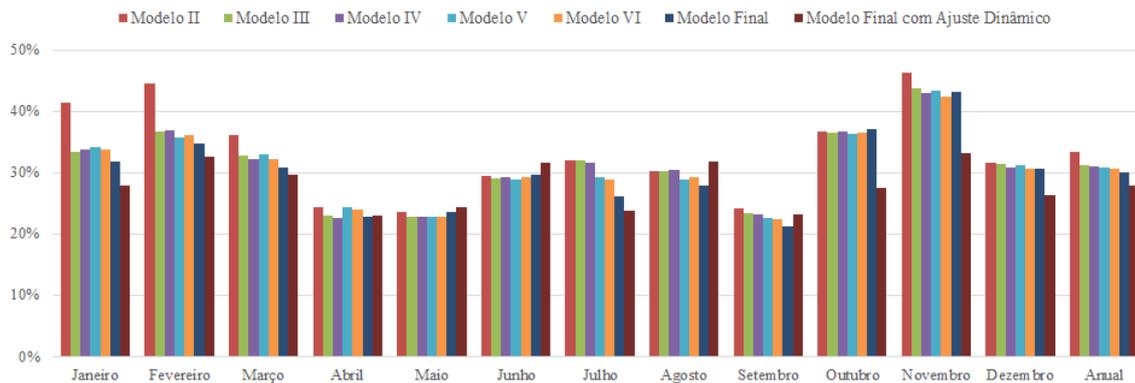


Figura 5.11: Valores de MAPE mensais da previsão determinística, para cada um dos modelos criados

Para além da medida de erro utilizada até agora (MAPE), optou-se, ainda, pela realização do cálculo do *bias*, isto é, do desvio do erro. Este, permite obter informação relativa ao erro sistemático da previsão, através da equação 5.1 [5].

$$BIAS(k) = \bar{\epsilon}_k = \frac{1}{N} \sum_{t=1}^N \epsilon_{t+k|t} \quad (5.1)$$

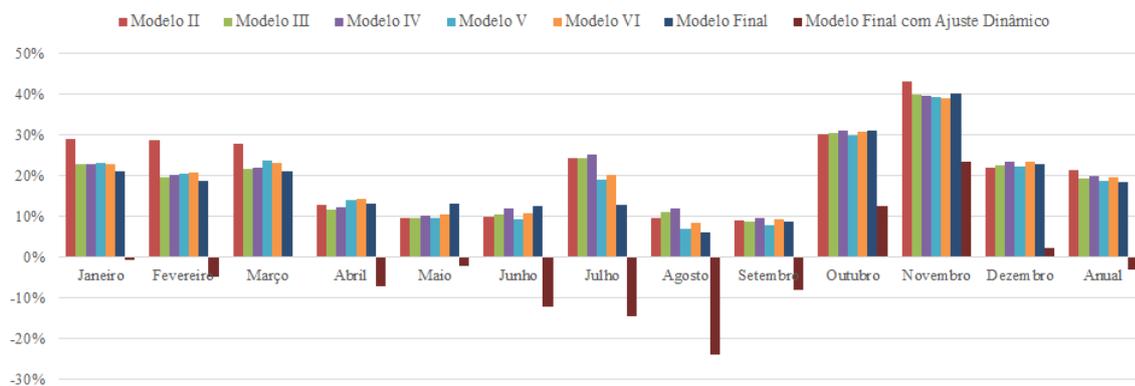


Figura 5.12: Valores do *bias* mensal da previsão determinística, para cada um dos modelos criados

O cálculo foi realizado, mais uma vez, para cada mês de previsão, de modo a facilitar a compreensão da sua evolução ao longo do ano (ver figura 5.12). Os valores são apresentados em

percentagem, relativa à potência média.

Estas figuras permitem conhecer os meses para os quais as previsões se mostraram mais eficazes e, ainda, comprovar a superioridade de uns modelos em relação aos outros. O modelo final, após realização do ajuste dinâmico de resultados, apresenta um desempenho consideravelmente superior a qualquer outro. No entanto, no que toca à previsão probabilística, tal como referido anteriormente, o mesmo não se pode concluir. Utilizando como exemplo os Modelos II e IV, compararam-se as suas previsões para o mês de janeiro. Segundo os cálculos anteriores, é visível a quase constante superioridade do Modelo IV. No entanto, fazendo uma análise da figura 5.13, surgem dúvidas relativas à previsão probabilística. À primeira vista pode parecer que a margem de incerteza, no caso do Modelo IV, engloba melhor os valores de produção reais registados e que, portanto, é mais correta. No entanto, na verdade, trata-se de uma margem demasiado larga, incluindo mais pontos do que os desejados. Isto provoca consequências negativas a nível de tomadas de decisão, uma vez que se vão basear em probabilidades erradas.

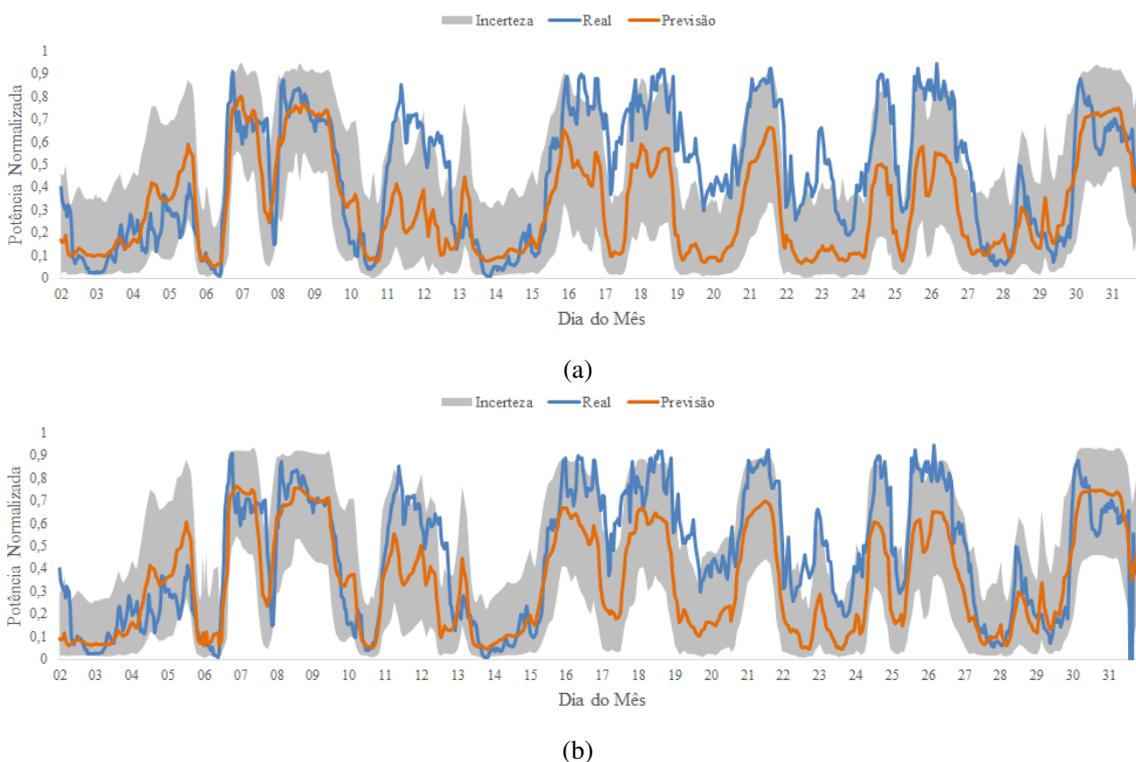


Figura 5.13: Previsões obtidas através do Modelo II (5.13a) e do Modelo IV (5.13b)

Uma outra ferramenta útil na interpretação de erros, consiste na construção de histogramas. Para tal, dividiu-se a percentagem de erro em intervalos (*bins*) e calculou-se a frequência de eventos que ocorre em cada um deles. Desta forma, obtiveram-se histogramas do erro para cada um dos modelos desenvolvidos (figura 5.14). Tal como verificado anteriormente, o Modelo I apresenta valores de erro muito elevados, não constituindo interesse comparativo com outros modelos.

No caso do modelo final, considerou-se relevante analisar as diferenças com e sem ajuste dinâmico, pelo que os histogramas do erro foram colocados no mesmo gráfico (figura 5.15).

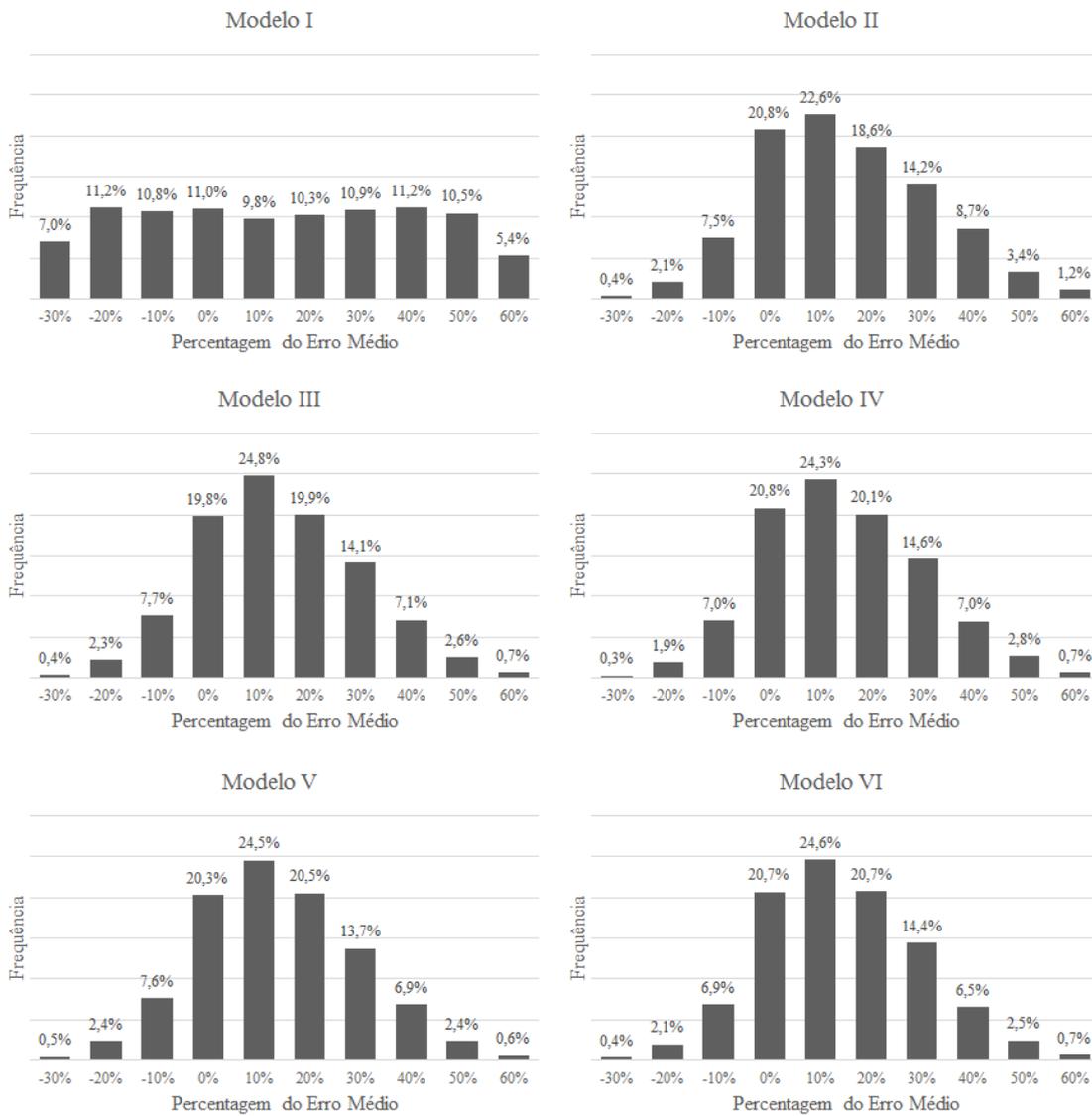


Figura 5.14: Histograma do erro médio de cada um dos modelos desenvolvidos

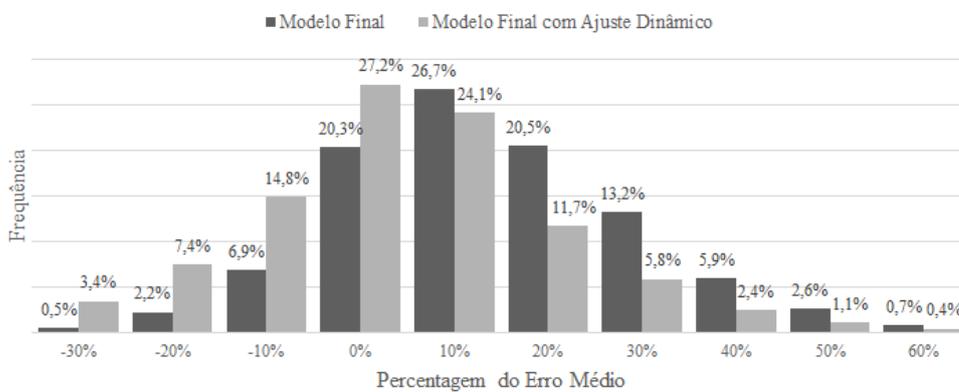


Figura 5.15: Histograma do erro médio relativo ao modelo final com e sem ajuste dinâmico

A evolução média do erro ao longo do dia, realizada para todo o ano de 2013, permite perceber as variações provocadas pela complexidade dos modelos (figura 5.16).

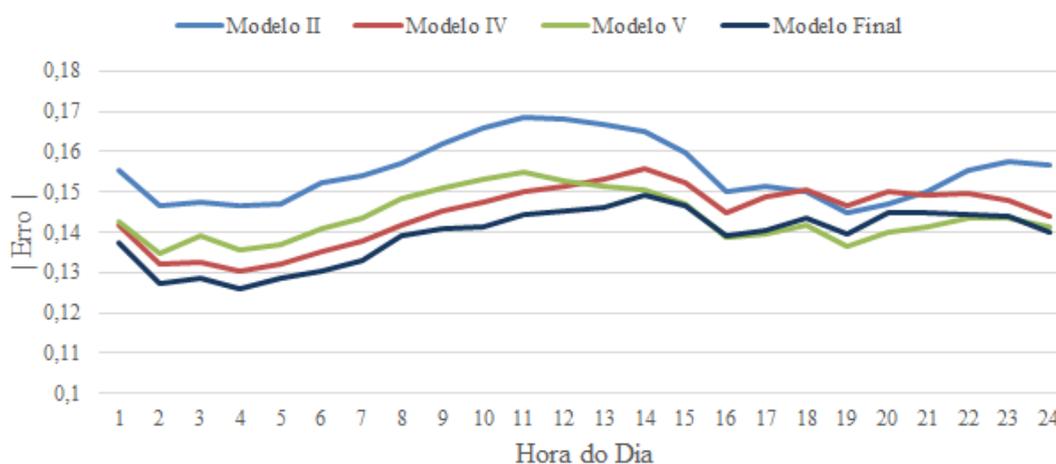


Figura 5.16: Evolução média do erro ao longo do dia, segundo diferentes modelos de previsão

Efetivamente, o modelo V é o que utiliza mais variáveis de entrada, apresentando variações mais bruscas a nível do módulo do erro. O modelo final, apesar de não ser muito menos complexo, revela-se mais consistente, em termos de variação do erro ao longo do tempo.

5.3 Comparação com Outras Técnicas

Na avaliação de desempenho de modelos usando diferentes técnicas de previsão, é essencial utilizar o mesmo conjunto de dados, de forma a que a comparação seja justa e não existam grandes discrepâncias na precisão dos valores utilizados como *input* [5]. Assim, de modo a comparar o modelo construído com outros sistemas, efetuaram-se previsões para o ano de 2013, recorrendo, também, a uma técnica baseada em redes neuronais e à técnica associada ao modelo de Persistência.

Uma vez que a técnica do modelo de Persistência constitui uma referência na avaliação de desempenho dos outros métodos, tal como referido no capítulo 2, considerou-se relevante compará-lo com a técnica KDE desenvolvida. A persistência caracteriza-se por assumir que a produção a um determinado instante futuro é igual à produção no instante atual, podendo ser formulada de acordo com a equação 5.2 [11]. Recorrendo a esta abordagem, fizeram-se previsões determinísticas a cada 24 horas, para as 24 horas seguintes, durante o ano de 2013.

$$\hat{P}_{t+k|t} = P_t \quad (5.2)$$

Com o mesmo intuito, utilizou-se uma técnica constituída por uma rede neuronal que, de forma semelhante ao modelo KDE final, foi treinada com os valores registados em 2012 das variáveis

temperatura, velocidade e direção do vento. Utilizaram-se 15 neurónios numa única camada intermédia, estimando 5 por cada variável de entrada. Os valores de *target* foram, mais uma vez, os dados registados da produção de potência eólica no ano de 2012.

A tabela 5.4 refere-se aos valores de MAPE para cada uma das três técnicas, sendo que na figura 5.17, é possível verificar a evolução do erro ao longo do dia, para cada uma delas. Esta evolução é feita através da média dos erros de todo o ano de 2013.

Tabela 5.4: Valores de MAPE da previsão determinística utilizando diferentes métodos

	MAPE (%)
KDE	29,887
REDE NEURONAL	31,063
PERSISTÊNCIA	35,618

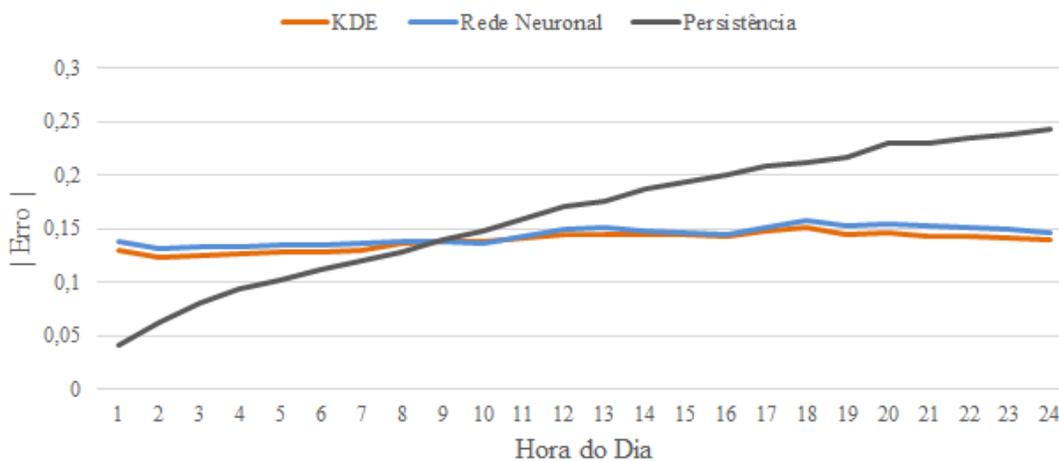


Figura 5.17: Evolução média do erro ao longo do dia, segundo diferentes técnicas de previsão

Tal como seria de esperar, o modelo de Persistência apresenta bons resultados apenas para curtos horizontes temporais (até cerca de 9 horas), uma vez que vai piorando significativamente ao longo do tempo. Na verdade, as previsões obtidas segundo esta técnica, apresentam o comportamento verificado na figura 5.18. No global, acaba por apresentar o pior desempenho dos três modelos.

Quanto ao KDE e à rede neuronal, pode-se concluir que ambos os modelos possuem um erro quase constante ao longo do tempo, cujos valores são bastante próximos. Ainda assim, o modelo KDE mostra-se superior, com menor valor de MAPE e, portanto, com um melhor desempenho geral.

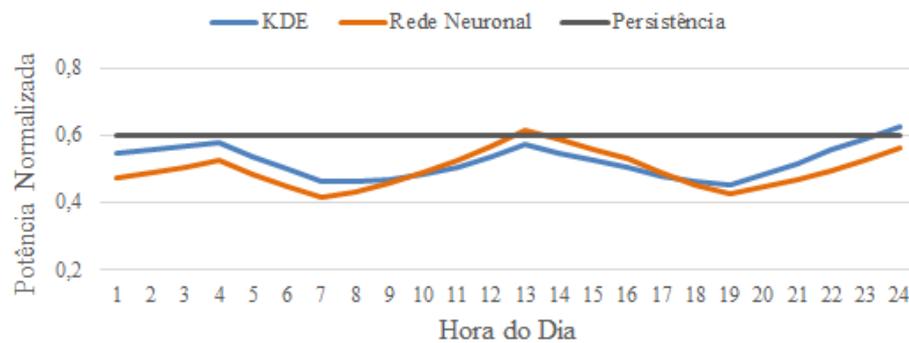


Figura 5.18: Previsões da produção eólica para um dia, segundo diferentes técnicas de previsão

Esta comparação apenas pode ser realizada relativamente à previsão pontual da produção eólica, uma vez que os restantes modelos não fornecem previsões probabilísticas.

5.4 Aplicações

A técnica KDE utilizada e desenvolvida ao longo da presente dissertação, para além de possuir resultados bastante satisfatórios em termos de previsão do valor esperado, possui a grande vantagem de, ainda, fornecer previsões probabilísticas. Efetivamente, é neste aspeto que se distingue, essencialmente, da maioria dos modelos utilizados atualmente.

A incerteza provocada pelas fontes de energia renovável no sistema elétrico leva à necessidade de fazer o despacho tendo em conta uma capacidade de produção extra (reserva) [26, 27]. A importância da previsão probabilística está, maioritariamente, relacionada com o cálculo desta reserva, uma vez que, utilizando abordagens determinísticas, é frequente que a reserva seja fixada numa determinada fração do valor de pico ou no mesmo valor da maior unidade de produção utilizada [9]. Assim, não considerando a incerteza, as soluções determinadas podem vir a ser pouco económicas.

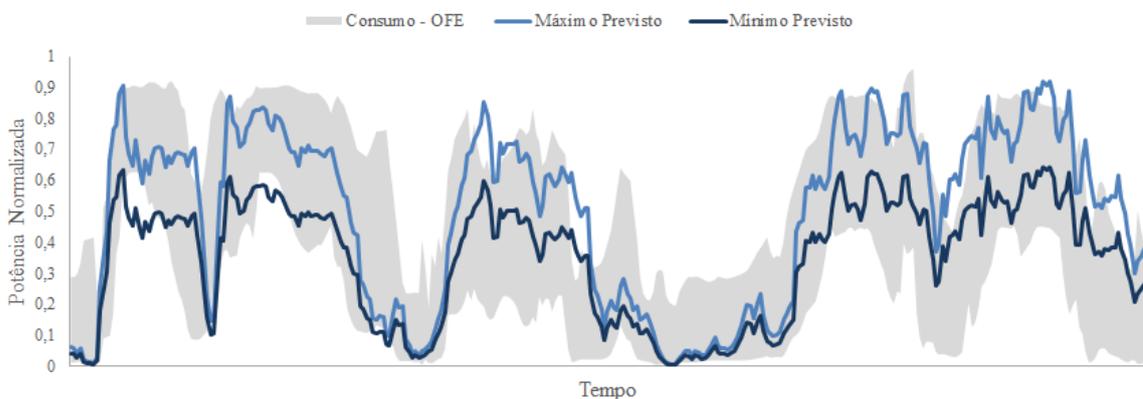


Figura 5.19: Exemplo de situação de despacho, com representação de previsão probabilística e necessidade de produção eólica

Fazendo a diferença entre a previsão do consumo e a previsão da produção por parte das outras fontes de energia (OFE), que não a eólica, obtém-se a produção necessária prevista por parte desta. O facto de se conhecer a previsão probabilística, permite cruzar as informações, tal como na figura 5.19, de modo a tomar decisões com base em probabilidades e cenários. Por exemplo, quando a necessidade de potência é superior ao limite máximo da previsão probabilística de produção eólica, surge o problema de se conseguir manter a continuidade de serviço, no entanto, com esta informação, é possível conhecer com que probabilidade as fontes eólicas não vão ser capazes de suprir o consumo e, portanto, fazer um despacho adequado à situação.

A figura 5.20 representa uma aplicação de previsões determinísticas e probabilísticas, obtidas através de modelos KDE, em operações de despacho (pela empresa *Smartwatt*).

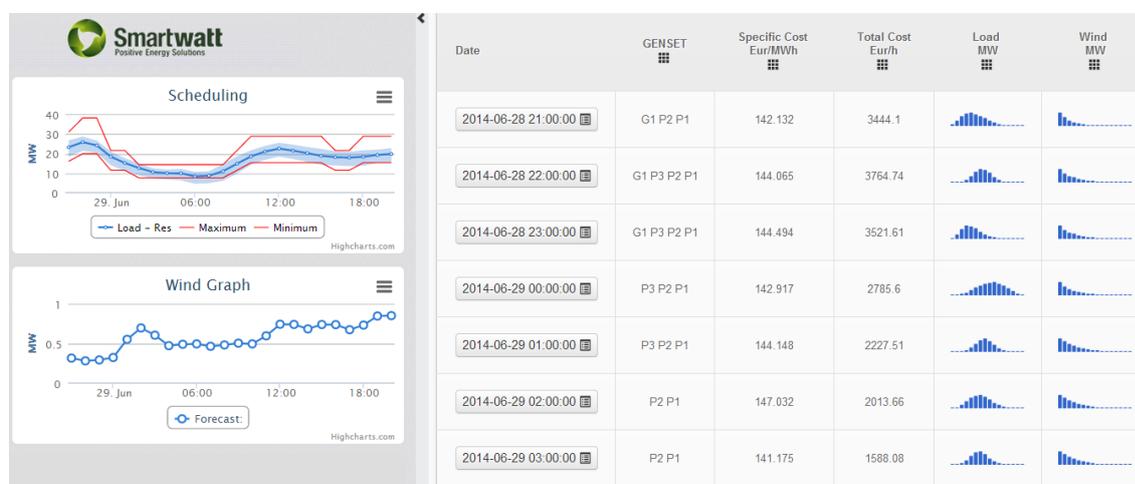


Figura 5.20: Aplicação de previsões de produção eólica em decisões de despacho

Na verdade, através deste tipo de estudo, a tomada de decisões com base em análise de risco pode, ainda, variar significativamente, conforme os operadores e utilizadores do sistema elétrico. Estes podem estar mais dispostos a correr riscos ou não, podendo surgir então uma das duas situações:

- Com uma dada probabilidade conhecida, é possível que a produção de energia eólica não seja suficiente, pelo que é essencial colocar em funcionamento unidades de produção convencional que cubram os valores de potência necessários, de modo a garantir a continuidade de serviço;
- A probabilidade das fontes de energia eólica produzirem potência suficiente, de acordo com o que se estima ser necessário, é satisfatória. Logo, utiliza-se o mínimo possível de unidades convencionais.

Esta análise deve ser feita tendo em conta as várias fontes de energia renovável integrantes no sistema e a qualidade das previsões de cada uma delas. Em ambos dos casos, a utilização de um método de decisão probabilístico, no geral, permite reduzir custos e riscos, relativamente

à adoção de critérios "fixos" de dimensionamento da reserva, cujos pressupostos traduzem pouca credibilidade [28].

Uma outra aplicação importante de previsões probabilísticas tem a ver com os preços do mercado elétrico. Estas dão origem a boas estratégias de trocas de energia, especialmente em termos de prevenir penalizações devido a desvios entre o despacho "day-ahead" e a produção em tempo real [28].

A figura 5.21 apresenta a interface gráfica, desenvolvida pela empresa *Smartwatt*, onde o modelo criado foi implementado e é utilizado, atualmente, na previsão a curto-prazo. Através desta plataforma, é possível conhecer os valores de produção esperados para os 7 dias seguintes, para o instante atual e para instantes passados. É, ainda, representada a incerteza a estes associada, com o intervalo de confiança definido pelos valores de "upper bound" e "lower bound" (alteráveis a qualquer momento).



Figura 5.21: Modelo final em funcionamento, implementado à ilha de Creta pela *Smartwatt*

5.5 Conclusões

Neste capítulo, fez-se uma análise de desempenho do modelo final criado no capítulo 4 e um levantamento das suas principais aplicações.

Começou-se por seleccionar o mês de janeiro e o mês de julho, para comparação de previsões obtidas segundo o modelo desenvolvido. Tal como seria de esperar, verificou-se que os erros de previsão são substancialmente superiores no inverno (MAPE=32,532%), relativamente

ao verão (MAPE=16,138%), devido à maior intensidade e variabilidade de vento. Efetivamente, comprovou-se a relação entre a quantidade de potência produzida e a variação do erro de previsão, tendo-se concluído que, para valores mais elevados de produção, o erro tende a ser, também, superior.

Os valores de MAPE elevados, verificados em alguns dos meses, podem ter diversas justificações. O facto de se terem utilizado variáveis meteorológicas provenientes de um modelo NWP global, ou mesmo as diferenças verificadas entre os dados do conjunto de treino e do conjunto de teste, constituem fatores que agravam a precisão do sistema desenvolvido.

A análise dos tipos de erro existentes, permitiu concluir que os erros de fase são mais visíveis no mês referente ao verão, onde a produção é mais constante, o que pode estar relacionado com o facto de neste, o erro da previsão probabilística ser superior.

Comparando o modelo final com os outros modelos desenvolvidos, é possível concluir, no seguimento do capítulo anterior, que são distintas as características das previsões obtidas através de cada um deles. O cálculo do MAPE e do *bias* para cada mês de previsão, permitiu comprovar a sazonalidade da produção e da qualidade dos resultados. Deste modo, de acordo com cada situação em específico, determinado modelo pode ter mais ou menos valor, dependendo se a importância é mais atribuída à previsão determinística ou à probabilística.

Um aspeto importante na avaliação de modelos de previsão, está na sua comparação com resultados obtidos através de técnicas diferentes. Assim, compararam-se as previsões realizadas através do modelo desenvolvido, baseado em KDE, com as previsões obtidas, para os mesmos instantes, através do modelo de persistência (referência) e de uma técnica baseada em redes neuronais. Uma vez que estes não realizam previsões probabilísticas, apenas foi possível comparar os valores referentes à previsão pontual de produção esperada. O modelo final, desenvolvido ao longo da dissertação, revelou-se superior a nível de desempenho geral, apresentando valores de MAPE inferiores e uma evolução do erro relativamente constante.

O facto da técnica utilizada fornecer informação sobre a incerteza associada à previsão dos valores pontuais de produção, constitui uma mais valia relativamente aos modelos de previsão tradicionais. Assim, a gama de aplicações torna-se muito mais vasta. A comparação entre os máximos e mínimos de incerteza prevista com a produção eólica que se prevê ser necessária, melhora os processos de despacho económico, cálculo da reserva pretendida e trocas de energia. Assim, a utilização de um método de decisão probabilístico constitui uma ferramenta com maior credibilidade, permitindo reduzir custos e riscos.

Capítulo 6

Conclusões

6.1 Satisfação de Objetivos

O rápido aumento dos investimentos em fontes de energia eólica, tem vindo a tornar a previsão da produção cada vez mais importante na complexa gestão e manutenção dos sistemas elétricos. Neste sentido, a dissertação desenvolvida compreendeu uma revisão do estado da arte sobre previsão de produção eólica e posterior construção e otimização de modelos do tipo estatístico, com um horizonte temporal de curto prazo.

O objetivo central do trabalho desenvolvido consistiu na criação de modelos capazes de, para o caso da ilha de Creta, efetuar boas previsões, não só do valor esperado de produção eólica, mas também da incerteza a estes associada (previsões probabilísticas). Assim, através de novas técnicas baseadas em KDE, estudaram-se diversos modelos, fazendo variar as suas características e parametrizações, de modo a alcançar uma solução otimizada.

Tal como tinha sido definido nos objetivos, desenvolveram-se e compararam-se, também, duas metodologias distintas de eliminação de erros em conjuntos de dados históricos, essenciais para a obtenção de previsões com alta qualidade.

Por fim, foi possível fazer uma análise dos resultados pontuais do valor esperado, obtidos segundo métodos de previsão distintos: técnica baseada em KDE, técnica baseada em redes neurais e técnica de persistência.

6.2 Conclusões Gerais

O sistema de previsão desenvolvido para a ilha de Creta, tira partido de valores de produção históricos e de variáveis meteorológicas previstas (NWP). Na verdade, a metodologia proposta, baseada em KDE, distingue-se da maioria das técnicas atualmente utilizadas, por ser capaz de fornecer não só previsões do valor esperado de produção, mas também uma modelização da incerteza associada, com diferentes intervalos de confiança.

De modo a alcançar a solução final, foram criados modelos com diferentes características, analisando as vantagens e o comportamento de cada um, relativamente a ambos os tipos de previsão

(determinística e probabilística). Durante este processo, concluiu-se que, na escolha de variáveis de entrada, apenas o desempenho da previsão pontual ou o desempenho da previsão de incerteza pode ser otimizado, uma vez que o outro varia de forma contrária. Neste caso, considerou-se pertinente selecionar o modelo com melhores valores esperados de produção, ajustando, posteriormente, as suas parametrizações de modo a otimizar a representação da incerteza. No entanto, esta opção deve ter em conta aquilo que os utilizadores finais pretendem e o modo como as previsões serão utilizadas. Dependendo da situação, pode ser mais vantajoso direcionar a construção do modelo, desde o início, para a previsão probabilística.

O estudo sobre os parâmetros a fornecer ao *software*, ou seja, sobre os desvios padrão referentes a cada variável explicativa, permitiu concluir que, dependendo do sistema, existem valores ideais diferentes. Assim, para cada conjunto de dados e de variáveis de entrada aplicados ao treino de um modelo, será necessária uma nova otimização das parametrizações, que pode dar origem a melhorias significativas do desempenho global do mesmo.

Uma pequena diminuição do erro de previsão constitui uma redução significativa das penalizações impostas aos produtores, pelos desvios de produção. Assim, a garantia da qualidade do conjunto de dados utilizado torna-se fundamental, tendo-se testado duas metodologias e selecionado a que deu origem aos melhores resultados, ou seja, a modelização da curva de potência média dos parques eólicos. Também a utilização de técnicas de aperfeiçoamento das previsões, após a sua realização, constitui uma ferramenta importante na diminuição do erro, tal como se verificou com o ajuste dinâmico dos resultados conseguidos.

No global, o modelo final apresentou melhor desempenho relativamente aos restantes modelos desenvolvidos e, ainda, relativamente a outras técnicas de previsão, como é o caso das redes neurais e do modelo de persistência. Os resultados obtidos foram bastante satisfatórios para todo o ano previsto. No entanto, tal como seria de esperar, durante o inverno verificam-se erros mais elevados, devido à considerável variabilidade do recurso eólico. Uma forma de melhorar estes resultados e aumentar a precisão do modelo, consiste em ir atualizando, ao longo do tempo, a sua base de conhecimento.

Por fim, resta salientar a importância de conhecer um conjunto de informação, em comparação com apenas um valor previsto de potência, tal como acontece com os modelos tradicionais. A possibilidade de obter a incerteza associada a cada ponto de previsão, permite minimizar os impactos causados pela volatilidade do vento e tornar a utilização de energia eólica mais segura e fiável para todos os intervenientes.

6.3 Trabalhos Futuros

A diminuição do erro entre a potência prevista e a real produzida será sempre um objetivo na investigação e desenvolvimento de sistemas de previsão. No entanto, existirão inevitáveis imprecisões. Assim, importa, cada vez mais, desenvolver abordagens que tiram partido de previsões probabilísticas para tomar decisões baseadas em análise de risco.

Os benefícios económicos introduzidos pela modelização da incerteza necessitam, ainda, de maior demonstração e comprovação, uma vez que, paralelamente, possuem aspetos negativos suscetíveis de levantar dúvidas. Concretamente, aos níveis dos custos adicionais ou, mesmo, da complexidade de compreensão e avaliação [5].

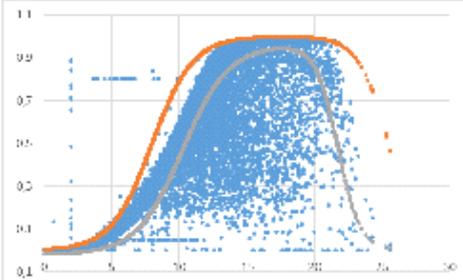
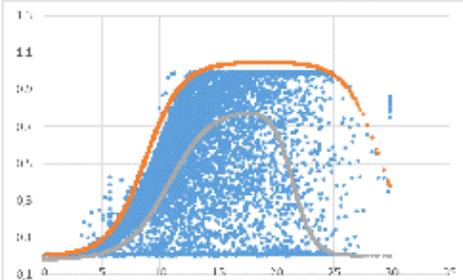
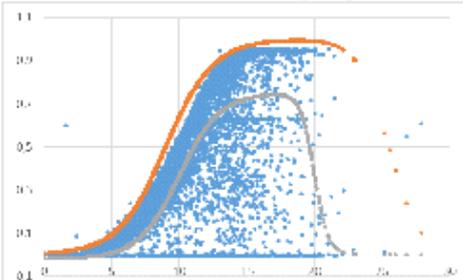
Ao longo do desenvolvimento desta dissertação, tornou-se clara a necessidade de criação de um sistema que tenha em conta as preferências dos utilizadores finais, adequando o modelo a cada caso. Também, uma melhoria a ser introduzida no *software* KDE, consiste na integração de métodos heurísticos capazes de otimizar, automaticamente, as parametrizações das variáveis explicativas utilizadas. Atualmente, o modelo implementado utiliza previsões meteorológicas com maior qualidade, apresentando melhor desempenho. No entanto, estas variáveis não chegaram a tempo de ser incluídas no presente estudo.

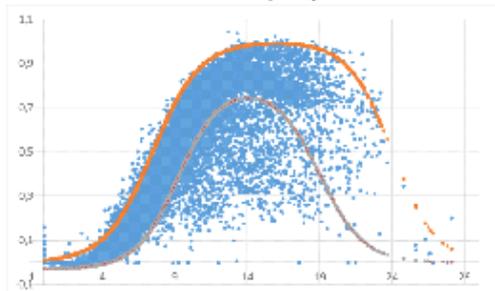
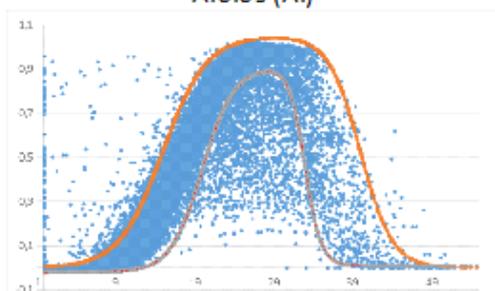
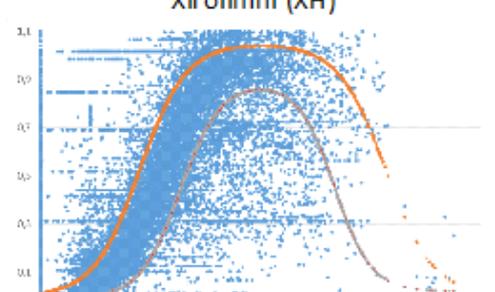
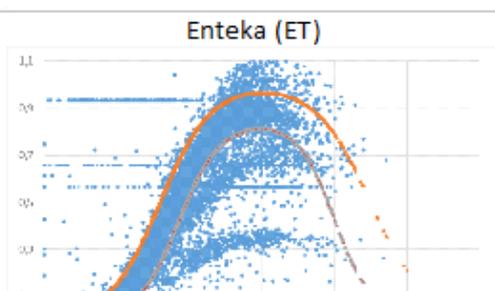
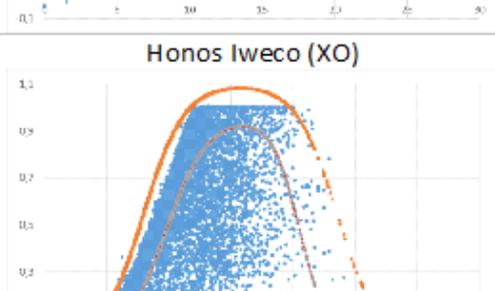
Existe, ainda, a hipótese de desenvolver modelos de segunda ordem, cujas variáveis de entrada integram previsões da variação de produção. No entanto, a previsão da variação tem-se verificado muito complexa, devido ao carácter volátil do recurso eólico, necessitando, por isso, de maior investigação.

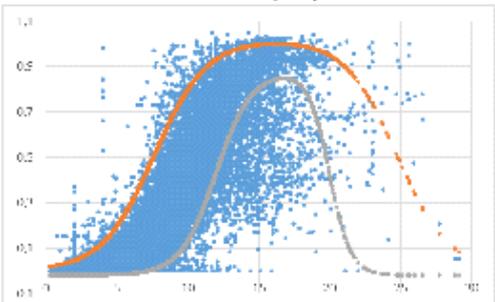
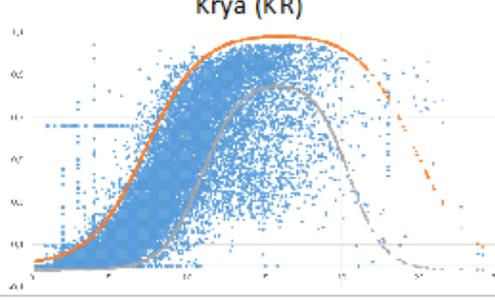
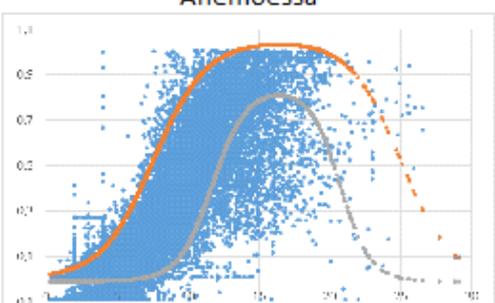
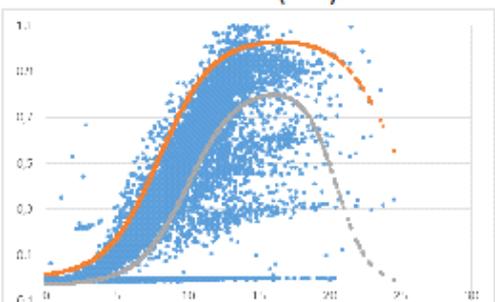
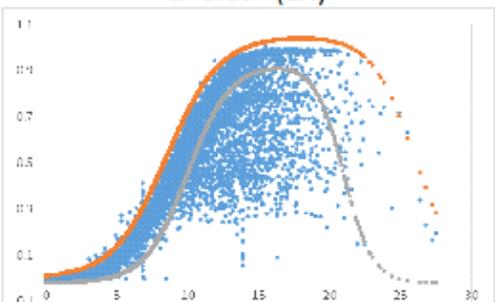
A título de conclusão, para além das melhorias a implementar nos sistemas de previsão, existe, ainda, muito trabalho a ser realizado no desenvolvimento de ferramentas e procedimentos que permitam fazer um uso mais eficiente da informação fornecida pelas previsões de produção eólica e pela incerteza a estas associada.

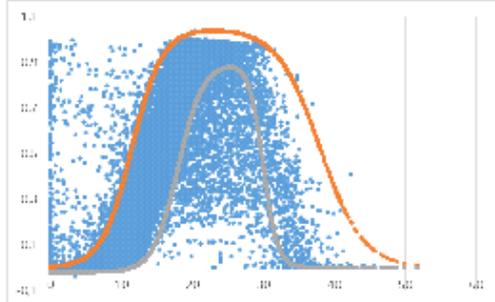
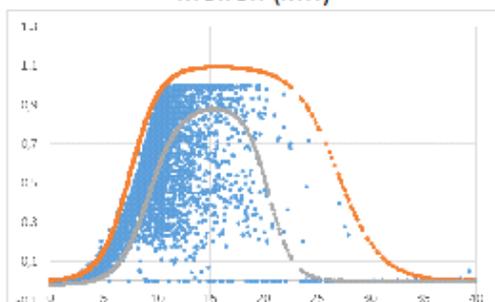
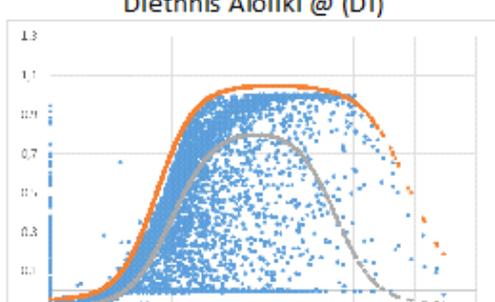
Anexo A

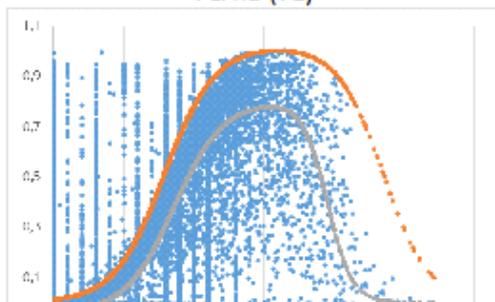
Filtragem de Erros

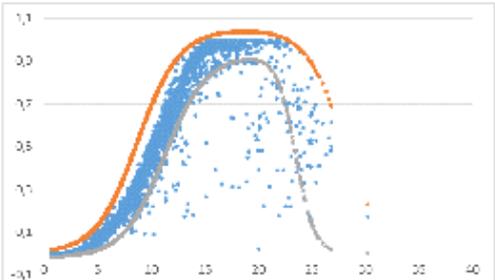
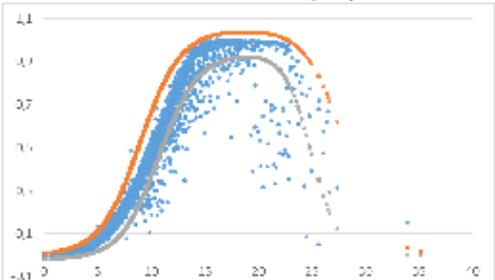
SUBSTATION KASTELIOU			
Park Name	Type	Power	Notes
<p>Ydroaioliki (YD)</p> 	Normal	Installed and Observed: 9,35 MW	
<p>Anemos Alkionis @ (AL)</p> 	Control.	Installed and Observed: 6,30 MW	
<p>Mousouron @ (MS)</p> 	Control.	Installed and Observed: 2,55 MW	There is only data from 09/07/2011 to 31/12/2012

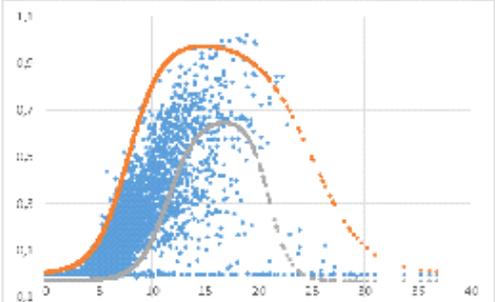
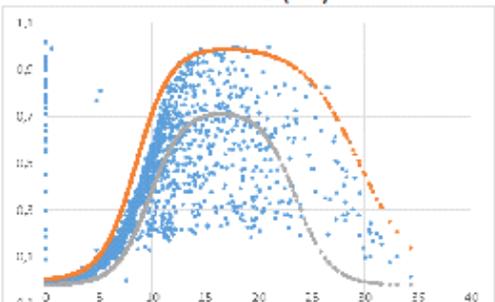
SUBSTATION SITEIAS			
Park Name	Tip	Power	Notes
<p>Rokas (RO)</p> 	Normal	<p>Installed: 12,90 MW Observed: 11,20 MW</p>	
<p>Aiolos (AI)</p> 	Normal	<p>Installed: 10,00 MW Observed: 8,20 MW</p>	<p>The curve had negative values, starting and ending in -0,1. They were corrected, so that the minimum is zero.</p>
<p>Xirolimni (XH)</p> 	Normal	<p>Installed: 10,20 MW Observed: 7,00 MW</p>	
<p>Enteka (ET)</p> 	Normal	<p>Installed: 2,70 MW Observed: 2,50 MW</p>	
<p>Honos Iweco (XO)</p> 	Normal	<p>Installed and Observed: 4,50 MW</p>	

SUBSTATION MARONIAS			
Park Name	Type	Power	Notes
<p>Axladia (AX)</p> 	Normal	Installed and Observed: 10,00 MW	The curve had negative values. They were corrected, so that the minimum is zero.
<p>Krya (KR)</p> 	Normal	Installed and Observed: 10,00 MW	The curve had negative values. They were corrected, so that the minimum is zero.
<p>Anemoessa</p> 	Normal	Installed: 5,00 MW Observed: 5,50 MW	
<p>Platuvola (WR)</p> 	Normal	Installed: 3,00 MW Observed: 2,30 MW	
<p>Enercon (EN)</p> 	Normal	Installed: 2,50 MW Observed: 3,00 MW	The curve had negative values. They were corrected, so that the minimum is zero.

SUBSTATION MOIRON			
Park Name	Tip	Power	Notes
<p>Iweco (Megali Vrysi) (IW)</p> 	Normal	Installed: 4,95 MW Observed: 4,30 MW	The curve had negative values. They were corrected, so that the minimum is zero.
<p>Moiron (MR)</p> 	Normal	Installed and Observed: 5,25 MW	
<p>Diethnis Aioliki @ (DI)</p> 	Control.	Installed and Observed: 7,20 MW	

SUBSTATION AG. VARVARAS			
Park Name	Tip	Power	Notes
<p>Terna (TE)</p> 	Normal	Installed: 14,45 MW Observed: 14,00 MW	

SUBSTATION AGYIAS			
Park Name	Type	Power	Notes
<p>Envitec Vardia (EV)</p> 	Normal	Installed and Observed: 5,40 MW	
<p>Envitec Vatali (BT)</p> 	Normal	Installed and Observed: 5,40 MW	

SUBSTATION RETHIMNOU			
Park Name	Type	Power	Notes
<p>Kouloukonas @ (KO)</p> 	Control.	Installed and Observed: 4,80 MW	There is only data from 13/10/2011 to 31/12/2012
<p>Akoumia (AK)</p> 	Normal	Installed and Observed: 7,20 MW	There is only data from 13/10/2011 to 31/12/2012

Anexo B

Erros da Previsão Probabilística

	PROBABILIDADE DESEJADA	PROBABILIDADE OBTIDA	ERRO	MAPE
MODELO I	90%	92,19%	2,19%	4,70%
	80%	81,00%	1,00%	
	70%	69,47%	0,53%	
	60%	57,50%	2,50%	
	50%	46,56%	3,44%	
	40%	36,12%	3,88%	
	30%	26,42%	3,58%	
	20%	17,41%	2,59%	
	10%	8,57%	1,43%	

Figura B.1: Resultados detalhados da previsão probabilística obtida para todo o ano de 2013, no caso do Modelo I

	PROBABILIDADE DESEJADA	PROBABILIDADE OBTIDA	ERRO	MAPE
MODELO II	90%	91,87%	1,87%	5,88%
	80%	82,74%	2,74%	
	70%	73,66%	3,66%	
	60%	64,13%	4,13%	
	50%	53,93%	3,93%	
	40%	43,99%	3,99%	
	30%	33,18%	3,18%	
	20%	21,75%	1,75%	
	10%	11,21%	1,21%	

Figura B.2: Resultados detalhados da previsão probabilística obtida para todo o ano de 2013, no caso do Modelo II

	PROBABILIDADE DESEJADA	PROBABILIDADE OBTIDA	 ERRO 	MAPE
MODELO III	90%	91,19%	1,19%	6,82%
	80%	82,79%	2,79%	
	70%	74,75%	4,75%	
	60%	65,24%	5,24%	
	50%	55,17%	5,17%	
	40%	44,91%	4,91%	
	30%	33,78%	3,78%	
	20%	21,78%	1,78%	
	10%	11,09%	1,09%	

Figura B.3: Resultados detalhados da previsão probabilística obtida para todo o ano de 2013, no caso do Modelo III

	PROBABILIDADE DESEJADA	PROBABILIDADE OBTIDA	 ERRO 	MAPE
MODELO IV	90%	92,47%	2,47%	8,51%
	80%	84,36%	4,36%	
	70%	75,76%	5,76%	
	60%	66,17%	6,17%	
	50%	55,93%	5,93%	
	40%	45,26%	5,26%	
	30%	34,13%	4,13%	
	20%	22,67%	2,67%	
	10%	11,54%	1,54%	

Figura B.4: Resultados detalhados da previsão probabilística obtida para todo o ano de 2013, no caso do Modelo IV

	PROBABILIDADE DESEJADA	PROBABILIDADE OBTIDA	 ERRO 	MAPE
MODELO V	90%	91,32%	1,32%	6,57%
	80%	82,98%	2,98%	
	70%	74,44%	4,44%	
	60%	65,40%	5,40%	
	50%	54,85%	4,85%	
	40%	44,86%	4,86%	
	30%	33,23%	3,23%	
	20%	21,69%	1,69%	
	10%	10,80%	0,80%	

Figura B.5: Resultados detalhados da previsão probabilística obtida para todo o ano de 2013, no caso do Modelo V

	PROBABILIDADE DESEJADA	PROBABILIDADE OBTIDA	 ERRO 	MAPE
MODELO VI	90%	92,61%	2,61%	8,67%
	80%	84,75%	4,75%	
	70%	76,32%	6,32%	
	60%	66,35%	6,35%	
	50%	56,01%	6,01%	
	40%	45,46%	5,46%	
	30%	34,02%	4,02%	
	20%	22,57%	2,57%	
	10%	10,94%	0,94%	

Figura B.6: Resultados detalhados da previsão probabilística obtida para todo o ano de 2013, no caso do Modelo VI

	PROBABILIDADE DESEJADA	PROBABILIDADE OBTIDA	 ERRO 	MAPE
MODELO FINAL	90%	89,96%	0,04%	2,98%
	80%	81,11%	1,11%	
	70%	71,99%	1,99%	
	60%	61,99%	1,99%	
	50%	52,12%	2,12%	
	40%	41,88%	1,88%	
	30%	31,51%	1,51%	
	20%	21,59%	1,59%	
	10%	11,17%	1,17%	

Figura B.7: Resultados detalhados da previsão probabilística obtida para todo o ano de 2013, no caso do modelo final

Referências

- [1] REN. Dados técnicos/technical data 2012, 2013.
- [2] Adelino J. C. Pereira e João Tomé Saraiva. Long term impact of wind power generation in the iberian day-ahead electricity market price. *Energy*, páginas 1159–1171, 2013.
- [3] Luís Manuel Azevedo Costa. Simulação e otimização da gestão de portfólio de agentes de mercado agregadores. master's thesis, Faculdade de Engenharia da Universidade do Porto, 2012.
- [4] Alexandre Costa, Antonio Crespo, Jorge Navarro, Gil Lizcano, Henrik Madsen, e Everaldo Feitosa. A review on the young history of the wind power short-term prediction. *Renewable and Sustainable Energy Reviews*, 12(6):1725–1744, 2008.
- [5] P McSharry, P Pinson, e R Gerard. Methodology for the evaluation of probabilistic forecasts. *SafeWind Report*, 2009.
- [6] Gregor Giebel, Richard Brownsword, George Kariniotakis, Michael Denhard, e Caroline Draxl. The state-of-the-art in short-term prediction of wind power: A literature overview. Relatório técnico, ANEMOS. plus, 2011.
- [7] J Mendes, RJ Bessa, H Keko, J Sumaili, V Miranda, C Ferreira, J Gama, A Botterud, Zhi Zhou, e Jianhui Wang. Development and testing of improved statistical wind power forecasting methods. Relatório técnico, Argonne National Laboratory (ANL), 2011.
- [8] W Glassley, J Kleissl, C.P. "Case" van Dam, H Shiu, J Huang, G Braun, e R Holland. 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.
- [9] Bruno Santos, Cláudio Monteiro, Carina Soares, João Catalão, Gerardo Osório, Juan Rojas, Rafael Nebot Medina, Gianmaria Sannino, Adriana Carillo, Carlos González García, e Javier Contreras Sanz. Report on the mathematical formulation of the different forecasting models and data analysis. Relatório técnico, SiNGULAR, 2014.
- [10] HMI Pousinho, VMF Mendes, e JPS Catalão. A stochastic programming approach for the development of offering strategies for a wind power producer. *Electric Power Systems Research*, 89:45–53, 2012.
- [11] C Monteiro, R Bessa, V Miranda, A Botterud, J Wang, e G Conzelmann. Wind power forecasting: state-of-the-art 2009. Relatório técnico, Argonne National Laboratory (ANL), 2009.

- [12] Matthias Lange e Ulrich Focken. Second workshop of international feed-in cooperation. Em *State-of-the-art in wind power prediction in germany and international developments*, volume 94, 2005.
- [13] Georges Kariniotakis, Pierre Pinson, Nils Siebert, Gregor Giebel, Rebecca Barthelmie, et al. The state of the art in short-term prediction of wind power-from an offshore perspective. Em *Proceedings of*, páginas 1–13, 2004.
- [14] Keith Parks, Yih-Huei Wan, Gerry Wiener, e Yubao Liu. Wind energy forecasting: a collaboration of the national center for atmospheric research (ncar) and xcel energy. *Contract*, 303:275–3000, 2011.
- [15] NOAA National Climatic Data Center. Global forecast system (gfs), 2014. <http://www.ncdc.noaa.gov/data-access/model-data/model-datasets/global-forecast-system-gfs>.
- [16] Instituto Português do Mar e da Atmosfera. Enciclopédia ipma.pt - previsão numérica do tempo, 2014. <http://www.ipma.pt/pt/enciclopedia/otempo/previsao.numerica/index.html>.
- [17] European Center for Medium-Range Weather Forecasts. Ensemble prediction system, 2012. http://www.ecmwf.int/services/dissemination/3.1/Ensemble_Prediction_System.html.
- [18] Xiaochen Wang, Peng Guo, e Xiaobin Huang. A review of wind power forecasting models. *Energy Procedia*, páginas 770–778, 2011.
- [19] João PS Catalão, Hugo MI Pousinho, e Victor MF Mendes. Optimal offering strategies for wind power producers considering uncertainty and risk. *Systems Journal, IEEE*, 6(2):270–277, 2012.
- [20] Walter Zucchini. Applied smoothing techniques, part 1: Kernel density estimation. *Temple University, Philadelphia, PA*, 2003.
- [21] António dos Santos Pires. Improving electrical production forecasting using gpu. master's thesis, Faculdade de Engenharia da Universidade do Porto, 2013.
- [22] Tinghui Ouyang, Xiaoming Zha, Liang Qin, et al. A survey of wind power ramp forecasting. *Energy and Power Engineering*, 5(04):368, 2013.
- [23] C Ferreira, J Gama, L Matias, A Botterud, e J Wang. A survey on wind power ramp forecasting. Relatório técnico, Argonne National Laboratory (ANL), 2011.
- [24] SiNGULAR. About singular, 2012. <http://www.singular-fp7.eu/home/index.php/abouts>.
- [25] Pierre Pinson. *Estimation of the uncertainty in wind power forecasting*. phd's thesis, École des Mines de Paris, 2006.
- [26] Peng Xiong. *Unit commitment under Uncertainty*. Tese de doutoramento, National University of Singapore, 2012.
- [27] Ali Ahmadi-Khatir, Mokhtar Bozorg, e Rachid Cherkaoui. Probabilistic spinning reserve provision model in multi-control zone power system. *Power Systems, IEEE Transactions on*, 28(3):2819–2829, 2013.

- [28] Audun Botterud, Zhi Zhou, Jianhui Wang, Ricardo J Bessa, Hrvoje Keko, Joana Mendes, Jean Sumaili, e Vladimiro Miranda. Use of wind power forecasting in operational decisions. *Argonne National Laboratory*, 2011.