

Faculdade de Engenharia da Universidade do Porto



**Sistema de Apoio à Análise de Desempenho de
Parques Eólicos**

Nuno André Ferreira Magalhães

Dissertação realizada no âmbito do
Mestrado Integrado em Engenharia Electrotécnica e de Computadores
Major Automação

Orientador: Prof. Dr. José Machado da Silva
Co-orientador: Eng.º José Carlos Matos

Março 2011

Resumo

Com o consecutivo aumento das preocupações ambientais, a implementação de medições de eficiência energética e o recurso à utilização de fontes de energia renovável tornou-se essencial para a sustentabilidade energética dos países. Esta energia é proveniente de recursos naturais e renováveis como o sol, vento, chuva, marés, calor, entre outras, perfazendo o seu conjunto o mix energético que corresponde à solução mais eficiente para diminuir a dependência do Homem dos combustíveis fósseis, diminuindo também os índices de poluição proveniente da utilização destes. Para que se entenda a dimensão que os recursos endógenos têm no sector da energia, já em 2008 cerca de 19% do consumo mundial de energia provinha de fontes de energia renovável [1].

Do conjunto de tecnologias utilizadas na transformação de energia baseada em recursos endógenos, a que tem apresentado maior crescimento anual em termos de capacidade instalada é a eólica. Só em 2009 a capacidade instalada a nível mundial era de cerca de 159 GW [1]. Este facto advém da maturidade da tecnologia e do custo de investimento ser mais baixo quando comparado com outras tecnologias. No entanto, o investimento efectuado na construção de um parque eólico é avultado, justificando por um lado que decisão sobre a sua construção seja baseada em projectos de avaliação do recursos e por outro que exista uma monitorização contínua do desempenho de operação do parque durante a fase de exploração. Um dos processos de monitorização pode ser baseado em sistemas de análise de desempenho que permitem avaliar as condições de funcionamento dos centros electroprodutores de base eólica, designados normalmente de parques eólicos.

O objectivo desta tese é o desenvolvimento de um sistema que através do recurso a uma base de dados e a um conjunto de aplicações, processe a vasta informação operacional de um determinado parque eólico. A base de dados para além de conter informação caracterizadora de cada parque eólico contém também os dados de operação relativos a cada parque. Quando a informação operacional se encontra na base de dados, serão executadas aplicações de processamento da informação, baseadas em algoritmos estatísticos e métodos heurísticos, que possibilitarão a classificação de dados em diferentes grupos identificadores de diferentes estados de operação. A classificação realizada possibilitará, identificar períodos

de indisponibilidade de produção energética ou de baixo rendimento, problemas mecânicos e electrónicos, falhas nos equipamentos de leitura de dados e condições climatéricas adversas. Com esta informação, o operador poderá regularizar e otimizar o funcionamento dos centros electroprodutores que monitoriza, facultando o seu cliente de informação vital para uma melhoria do rendimento energético da sua exploração e alertando para possíveis danos detectados em equipamentos, pela simples análise de dados do seu funcionamento.

Esta metodologia de aquisição e processamento de dados possui vantagens em relação aos métodos actualmente utilizados. O rápido processamento e a simplificação na implementação lógica dos métodos e algoritmos utilizados, quando aliados a uma maior organização no armazenamento da informação adquirida e a uma maior flexibilidade em alterações logarítmicas dos métodos, permitem uma diferenciação relativamente às metodologias actuais.

Os resultados obtidos da análise de performance para os diferentes métodos implementados para esta aplicação, revelam elevadas performances energéticas para a generalidade dos métodos e ainda erros de classificação para os mesmo a variar entre os 7 e os 15 por cento. Desta análise se conclui, que os métodos fundamentados em critérios estatísticos, que recorrem a medidas de dispersão (Desvio Padrão) e médias amostrais, permitem obter erros de classificação menores, relativamente a métodos de estimação e métodos baseados em aproximações às curvas de potência teóricas dos aerogeradores. Esta aparente vantagem, não se revela contudo suficiente, para concluir quais os métodos mais eficazes, porque testes realizados revelam igualmente que a capacidade de um método em identificar *outliers* de funcionamento, nem sempre está relacionada com o seu grau de desempenho energético global.

Uma análise pormenorizada de tudo o que foi descrito neste resumo será realizada ao longo desta dissertação, de maneira a que se entenda no que consiste um sistema de apoio à análise de desempenho de parques eólicos e de que forma este pode ser implementado.

Abstract

With the successive increase of the environment concerns, the implementation of energy efficiency measures with the recourse to the use of renewable energy sources, become essential for energy sustainability of countries. This energy comes from natural resources such as sun, wind, rain, tides, heat, among others, which are renewable and represent together, the best way for the Man reduce dependence of fossil fuels and also decrease the levels of pollution from the use thereof. To understand the extent of endogenous resources in the energy sector, in 2008 approximately 19% of global energy consumption came from renewable energy sources [1].

Of the set of technologies used in energy transformation based on indigenous resources, the one who presented the greatest annual growth in terms of installed capacity was the wind. Only in 2009 the installed capacity worldwide was about 159 GW. This fact stems from the maturity of the technology and the cost of the lowest investment compared with other technologies. However, the investment made in building a wind farm is large and the decision on its construction is not easy, justifies an evaluation of the existing resources and a constant performance monitoring operation of the park during the exploration. One of the processes of monitoring systems can be based on performance analysis for assessing the operating conditions of the basic wind power plants, commonly called wind farms.

The aim of this thesis is to develop a system that through the use of a database and a set of applications, process the vast information of a wind farm. The database, beyond the information for characterizing each wind park, also contains the operating data for each park. When information is in the operational database, data processing applications based on statistical algorithms and heuristic methods will run, which will enable the classification of data in different groups of different states of operation. The classification performed allows the identification of periods of unavailability of energy production or low yield, mechanical and electronic problems, equipment problems reading data and adverse weather conditions. With this information, the operator can regulate and optimize the operation of power plants that it monitors, providing vital information to his customer, to improve the energy of operation efficiency and warn possible damage detected in equipment, just by the analyzing of the operation data.

This methodology of acquisition and data processing has advantages over methods currently used. The processing speed and the implementation of methods and algorithms more simple, when combined with a larger organization in information storage and better

flexibility in changes of logarithmic methods, allows a differentiation in relation to existing methodologies.

The analysis performance of the results for the different methods implemented for this application shows a high performance to the generality of the methods and classification errors between the range of 7 and 15 percent. This analysis concludes that the methods based on statistical criteria, which use measures of dispersion (standard deviation) and sample average, achieve lower classification errors with respect to methods based on estimation and methods based on approximations of the theoretical power curve of the wind turbines. This apparent advantage, however it does not appear sufficient to conclude which methods are most effective, because realized tests also show that the ability to identify outliers in a method of operation is not always related with their degree of overall energy performance.

A detailed analysis of what has been described in this summary will be conducted throughout this dissertation, so as to understand how is constituted a support system to analysis performance of wind turbines and how this can be implemented.

x

x

Agradecimentos

Servem estas curtas palavras para agradecer a todos que me fazem sorrir a cada dia, de forma especial à minha família, à Monika e aos meus amigos.

Queria também deixar o meu especial agradecimento ao João Sousa e João Rio, pelo apoio, paciência e entrega que sempre disponibilizaram. Agradeço também ao meu orientador e co-orientador e a todos que contribuíram para a minha educação.

Índice

Resumo	iii
Abstract.....	vii
Agradecimentos	xi
Índice.....	xiii
Lista de figuras	xvii
Lista de tabelas	xxi
Abreviaturas e Símbolos	xxiii
Capítulo 1	1
Introdução.....	1
1.1 - Enquadramento	1
1.2 - A Energia Eólica	5
1.3 - Enquadramento da Tese	12
1.4 - Objectivos.....	12
1.5 - Organização da Dissertação	13
Capítulo 2	15
Estado da Arte.....	15
2.1 - A classificação de dados na produção de energia eólica.....	15
2.2 - Métodos de classificação de dados utilizados em diversas áreas.....	26
2.3 - Comparação entre linguagens de programação passíveis de serem utilizados.....	32
Capítulo 3	36
Metodologias Desenvolvidas.....	37
3.1 - Metodologia de classificação de dados baseada no desvio padrão.....	38
3.2 - Metodologia de classificação de dados baseada na aplicação do critério de Chauvenet	39
3.3 - Metodologia de classificação de dados baseada numa aproximação à curva de Potência de um aerogerador.....	41
3.4 - Metodologia de classificação de dados baseada num modelo Auto-regressivo	43
3.5 - Análise Global sobre as Metodologias desenvolvidas	47
Capítulo 4	48
Sistema de classificação desenvolvido	49

4.1 - Especificação da Linguagem de Programação Utilizada	49
4.2 - Especificação do Sistema de Gestão de Base de Dados	50
4.3 - Especificação do sistema desenvolvido para Análise de Desempenho	51
4.4 - Especificação da Aplicação - Python	56
Capítulo 5	73
Análise de Resultados	73
5.1 - Resultado da Análise efectuada sobre o Aerogerador λ para um intervalo de tempo T1	74
5.2 - Resultado da Análise efectuada sobre o Aerogerador ω para um intervalo de tempo T2	82
5.3 - Resultado da Análise efectuada sobre diferentes aerogeradores	88
5.4 - Considerações Finais sobre os resultados.....	90
Capítulo 6	93
Conclusão	93
Desenvolvimento de trabalhos futuros.....	95
Referências	96
Anexos	98
A - Exposição das características de uma Base de Dados.....	98
B - Exposição de conceitos presentes nas tabelas da Base de Dados.....	101
C - Características das Linguagens de Programação.....	105

Lista de figuras

Figura 1.1 - Concentração de dióxido de carbono na atmosfera desde 1958 a 1991 [2].	2
Figura 1.2 - Variação da temperatura global do planeta de 1880 a 2000 [2].	3
Figura 1.3 - Crescimento da exploração da energia eólica desde 1993 até 2009 [6].	5
Figura 1.4 - Consumo [GWh] para o ano de 2010 em Portugal [7].	6
Figura 1.5 - Representação de um aerogerador de rotor com eixo vertical [8].	8
Figura 1.6 - Representação de um aerogerador de rotor com eixo horizontal [8].	9
Figura 1.7 - Representação das partes constituintes de aerogerador de eixo horizontal [8].	9
Figura 1.8 - Constituição de uma nacele de gerador convencional [8].	10
Figura 1.9 - Constituição de uma nacele de gerador multi-pólos [8].	10
Figura 1.10 - Vista interior de um gerador[8].	11
Figura 1.11 - Forma típica da curva de potência de um aerogerador com controlo de <i>pitch</i> [8].	11
Figura 2.1 - Animação representativa de um conjunto de turbinas eólicas analisadas pelo método <i>Time Lapse Animation</i> [10].	16
Figura 2.2 - Exemplo de Turbina com limites de desempenho em potência [10].	17
Figura 2.3 - Representação esquemático do processo metodológico implementado pela MegaJoule [11].	18
Figura 2.4 - Representação da correlação cruzada de potências entre duas turbinas vizinhas [11].	19
Figura 2.5 - Representação filtrada da correlação cruzada de potências entre duas turbinas vizinhas e padrão linear da correlação entre potências [11].	20
Figura 2.6 - Representação da relação entre rácio de potência e direcção do vento [11].	21
Figura 2.7 - Representação da relação entre rácio de potência e direcção do vento, com filtragem de dados e uma representação padrão e linear da mesma relação [11].	22

Figura 2.8 - Esquema proposto para análise do tempo de inatividade do sistema [12].	24
Figura 2.9 - Interface pós-processamento para avaliação de dados por parte do operador [12].	25
Figura 2.10 - Interface do <i>software Casas</i> utilizado neste estudo [13].	28
Figura 2.11 - Pseudo código do logaritmo estatístico para encontrar <i>outliers</i> [13].	29
Figura 2.12 - Pseudo código dos logaritmos KNN e DTW para detecção de <i>outliers</i> [13].	30
Figura 3.1 - Representação teórica da Relação Potência-Velocidade do Vento fornecida por um fabricante [17].	39
Figura 3.2 - Representação da curva de Potência-Velocidade do Vento.	42
Figura 3.3 - Variação da velocidade do vento num curto intervalo de tempo [19].	44
Figura 3.4 - Auto correlação para a potência medida pelo SCADA [19].	45
Figura 3.5 - Auto correlação parcial para a potência medida pelo SCADA [19].	45
Figura 4.1 - Esquema representativo do funcionamento do sistema de análise e classificação de dados.	52
Figura 4.2 - Diagrama de classes da Base de Dados do Sistema	53
Figura 4.3 - Fluxograma geral do funcionamento da aplicação desenvolvida	56
Figura 4.4 - Interface da Aplicação.	57
Figura 4.5 - Fluxograma do processamento realizado para análise da relação Potência-Velocidade do Vento.	58
Figura 4.6 - Análise da relação Potência-Velocidade do vento	59
Figura 4.7 - Identificação de limites de potência pela análise da relação Potência-Velocidade do vento.	60
Figura 4.8 - Identificação de problemas de controlo do pitch pela análise da relação Potência-Velocidade do vento.	61
Figura 4.9 - Identificação de problemas de leitura do anemómetro pela análise da relação Potência-Velocidade do vento	62
Figura 4.10 - Fluxograma do processamento realizado para análise da relação Coeficiente de Potência - Velocidade do Vento	63
Figura 4.11 - Análise da relação Coeficiente de Potência-Velocidade do vento, sem aplicação do limite de Betz	64
Figura 4.12 - Ilustração da Teoria de Betz [25].	64
Figura 4.13 - Identificação de limites de potência pela análise da relação Coeficiente de Potência-Velocidade do vento.	65
Figura 4.14 - Fluxograma do processamento realizado para análise da relação Rotações por Minuto [RPM] - Velocidade do Vento.	66

Figura 4.15 - Rotações por minuto - Velocidade do vento.....	67
Figura 4.16 - Identificação de limites de rotação pela análise da relação Rotações por minuto - Velocidade do vento.....	68
Figura 4.17 - Análise temporal do funcionamento de um aerogerador com recurso a código de cores.	68
Figura 4.18 - Análise comparativa de registos efectuados pelo SCADA por bin de velocidade de vento.	69
Figura 4.19 - Análise da potência e velocidade de vento em função do tempo.	70
Figura 5.1 - Análise gráfica das relações de Potência-Velocidade do vento para os métodos implementados, no intervalo de tempo T1.	74
Figura 5.2 - Saídas gráficas da aplicação, para aerogerador λ no intervalo de tempo T1	76
Figura 5.3 - Coeficiente de Potência-Velocidade do vento, com a aplicação do limite de Betz para o método 1.....	77
Figura 5.4 - Outras saídas gráficas da aplicação, para aerogerador λ no intervalo de tempo T1	78
Figura 5.5 - Saídas gráficas da aplicação, para análise do rendimento energético do aerogerador λ no intervalo de tempo T1.....	79
Figura 5.6 - Análise gráfica das relações Potência-Velocidade do vento para os métodos implementados no intervalo de tempo T2	82
Figura 5.7 - Curva de Potência ampliada	83
Figura 5.8 - Saídas gráficas da aplicação, para aerogerador ω no intervalo de tempo T2	84
Figura 5.9 - Outras saídas gráficas da aplicação para o aerogerador ω no intervalo de tempo T2	85
Figura 5.10 - Saídas gráficas da aplicação, para análise do rendimento energético do aerogerador ω no intervalo de tempo T2	86
Figura 6.1 - Modo de funcionamento de uma transaction num SGBD.....	99

Lista de tabelas

Tabela 1.1 – Potência instalada obtida através do vento em diferentes países europeus em MW [6].	7
Tabela 2.1 – Detalhes sobre os dados utilizados no estudo da Universidade de Washington [13].	27
Tabela 2.2 – Comparação das técnicas de identificação de <i>outliers</i> para o conjunto de dados Reais e Sintéticos.	31
Tabela 2.3 – Comparação das técnicas de identificação de <i>outliers</i> para o conjunto de dados Sintéticos, com variação de <i>outliers</i>	31
Tabela 2.4 – Comparação entre diferentes tipos de linguagens de programação.	33
Tabela 2.5 – Comparação de características entre diferentes tipos de SGBD.	35
Tabela 3.1 – Tabela de referência do critério de Chauvenet, para rejeição do valor medido [18].	40
Tabela 3.2 – Resumo das metodologias desenvolvidas.	47
Tabela 5.1 – Resultado da análise numérica para o Aerogerador λ no intervalo de tempo T1	80
Tabela 5.2 – Resultado da análise numérica para o Aerogerador ω no intervalo de tempo T2	87
Tabela 5.3 – Resultado da análise numérica para diferentes aerogeradores relativamente ao erro absoluto médio.....	89
Tabela 5.4 – Resultado da análise numérica para diferentes aerogeradores relativamente ao erro absoluto médio.....	90
Tabela 6.1 – Exemplos puros e híbridos de linguagens de programação orientadas a objectos.....	105

Abreviaturas e Símbolos

Lista de abreviaturas

CO ₂	Símbolo químico do dióxido carbono
DTW	<i>Discrete Time Warping</i>
HC	Hidrocarbonetos
HTML	<i>HyperText Markup Language</i> , que significa Linguagem de Marcação de Hipertexto
KNN	<i>k-Nearest Neighbor</i>
NO _x	Óxidos de Azoto
NSA	<i>National Security Agency</i> , agência de criptologia inteligente do departamento de defesa dos Estados Unidos
PDF	<i>Portable Document Format</i>
PIB	Produto Interno Bruto
SCADA	<i>Supervisory Control And Data Acquisition</i>
SGBD	Sistema de Gestão de Base de Dados
SO ₂	Dióxidos de enxofre
SQL	<i>Structured Query Language</i>

Lista de símbolos

σ	Desvio padrão
ϕ, β	Coefficientes constantes reais

Capítulo 1

Introdução

1.1 - Enquadramento

Com o desenvolvimento económico, científico e social, as dependências energéticas necessárias a nível mundial aumentam a cada dia que passa. Devido às preocupações ambientais e sustentabilidade da nossa sociedade, exige-se uma diminuição do consumo de energia através de medidas que promovam a eficiência energética e simultaneamente um aumento da produção de energia com base em fontes de energia renovável implicando uma diminuição na produção de energia proveniente dos combustíveis fósseis. Este novo paradigma energético requer que nos domínios científicos e tecnológicos sejam desenvolvidas soluções de monitorização e controlo fundamentais ao funcionamento do sistema eléctrico. Actualmente, a ideia de um Mundo sem uma organização social, económica e política baseada num auxílio computacional seria inconcebível. Parece claro que sem este desenvolvimento, para o qual contribuiu em muito a revolução industrial, seria impossível conhecer o Mundo como hoje o conhecemos. Ciente de tudo isto, a preocupação do Homem, tem cada vez mais vindo a ser redireccionada no sentido de otimizar e desenvolver novos os meios e fontes de exploração energética actuais, utilizando para isso o recurso a fontes energéticas renováveis e tentando reduzir a dependência de energias não renováveis. Consideram-se fontes de energia não renováveis os combustíveis fósseis (carvão, petróleo bruto e gás natural, entre outras) e o urânio, sendo este último a matéria-prima necessária para obter a energia resultante dos processos de fissão ou fusão nuclear. Todas estas fontes de energia têm reservas finitas, uma vez que é necessário muito tempo para as repor, e a sua distribuição geográfica não é homogénea, ao contrário das fontes de energia renováveis que têm origem num ciclo de renovação mais curto na natureza.

Geralmente, as fontes de energia não renováveis são denominadas de energias “sujas”, sendo a sua utilização causa directa de importantes danos para o meio ambiente e para a sociedade, tais como: destruição de ecossistemas, doenças, redução da produtividade agrícola, corrosão de infra-estruturas, deterioração da camada de ozono ou chuvas ácidas,

entre outros efeitos. Estas são também causas de efeitos indirectos como acidentes em sondagens petrolíferas e minas de carvão ou a contaminação por derramamentos químicos e de combustível.

Actualmente, um dos problemas ambientais mais graves, resultante de um sistema energético que privilegia a utilização de fontes de energia não renováveis é denominado de “efeito de estufa”. Os combustíveis fósseis não produzem apenas energia, mas também grandes quantidades de dióxido de carbono (CO_2), que é um dos principais responsáveis pelo efeito de estufa do planeta. São ainda emitidos para a atmosfera outros gases nocivos como os óxidos de azoto (NO_x), dióxidos de enxofre (SO_2) e os hidrocarbonetos (HC). Na figura 1.1 pode ser verificado o sucessivo aumento da concentração de CO_2 na atmosfera ao longo dos anos [2].

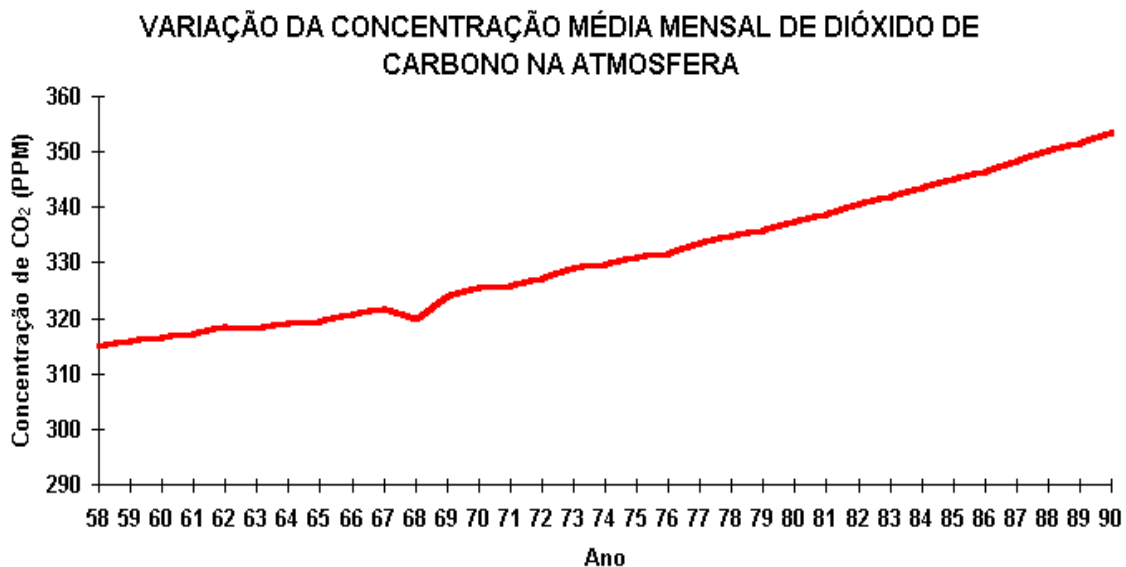


Figura 1.1 - Concentração de dióxido de carbono na atmosfera desde 1958 a 1991 [2].

É também possível verificar na figura 1.2 a associação directa da concentração de CO_2 na atmosfera, com o aumento gradual da temperatura na Terra ao longo do século passado.

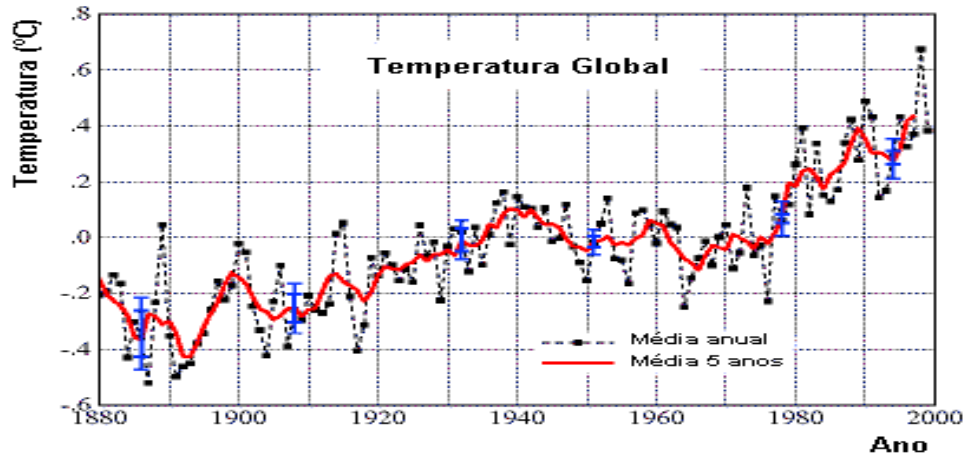


Figura 1.2 - Variação da temperatura global do planeta de 1880 a 2000 [2].

Outro problema que resulta de um sistema energético baseado na utilização de combustíveis fósseis é a dependência económica dos países não produtores das matérias-primas, que sendo obrigados a comprá-las externamente agravam significativamente as suas balanças de transacções. Em contrapartida, as energias renováveis são geralmente consumidas no local onde são geradas, isto é, são fontes de energia autóctones. Desta forma, é possível diminuir a dependência dos fornecimentos externos e contribuir ainda para o equilíbrio interterritorial e para a criação de postos de trabalho em zonas mais deficitárias. Neste sentido, estima-se que as energias renováveis sejam responsáveis pela criação de cinco vezes mais postos de trabalho do que as convencionais, que geram reduzidas oportunidades de emprego, atendendo ao seu volume de negócio.

O rápido crescimento observado das necessidades energéticas, aliado aos problemas do actual modelo energético representado em grande parte pelas energias não renováveis, fazem com que seja imprescindível propor um novo modelo baseado na eficiência e na poupança, fazendo recurso das energias renováveis.

Uma energia diz-se renovável quando não é possível estabelecer um fim temporal para a sua utilização. São exemplos disso mesmo, as energias obtidas por radiações solares, vento e força das águas. Com aparente vantagem sobre as energias fósseis, as energias renováveis possuem também entraves à sua utilização, normalmente relacionados com custos de instalação, limites tecnológicos e à falta de sensibilização para o seu uso, por parte dos consumidores e governos. Contudo, estes são problemas ultrapassáveis se tivermos em conta todas as vantagens que este tipo de exploração energética pode oferecer.

A comissão Europeia definiu para o ano de 2020 um conjunto de medidas com vista a combater as alterações climáticas e fazer frente aos grandes desafios energéticos em matéria de sustentabilidade. As metas estabelecidas receberam a designação de “20-20-20” [3] e pretendem incentivar a concretização dos seguintes objectivos:

1. 20% de redução na emissão de gases causadores do efeito de estufa, relativamente aos níveis de 1990;
2. 20% de aumento da utilização de energias renováveis na produção energética total;
3. 20% de redução do consumo de energia, como resultado de uma maior eficiência energética

A aplicação destas medidas, provocará um crescimento considerável no uso das energias renováveis e um acréscimo significativo no grau de importância que estas representam, não só pelas características naturais que já fazem destas uma melhor solução, mas também pela importância que estas terão para a sustentabilidade energética mundial.

No caso particular de Portugal, cuja abundância em recursos energéticos fósseis é diminuta, o aproveitamento das fontes de energia renováveis deve ser um dos objectivos primordiais da política energética nacional [4]. O programa de Governo do XVIII Governo Constitucional [5] estabelece como objectivo nacional colocar Portugal na vanguarda Europeia, no que em matéria de energias renováveis diz respeito. Nos últimos quatro anos, uma política de promoção das energias renováveis posicionou o país entre os líderes reconhecidos no uso dessas energias para produção de electricidade. Tendo em vista as metas Europeias estabelecidas para 2020, o governo português estabeleceu um conjunto de objectivos [5], que visam:

1. Reduzir a dependência energética do País face ao exterior para 74% em 2020;
2. Garantir o cumprimento dos compromissos assumidos por Portugal no contexto das políticas europeias de combate às alterações climáticas, permitindo que em 2020, 60% da electricidade produzida, e 31% do consumo de energia final, tenha origem em fontes renováveis e uma redução de 20% do consumo de energia final;
3. Reduzir em 25% a importação energética, com a energia produzida a partir de fontes endógenas;
4. Criar riqueza e consolidar um *cluster* energético no sector das energias renováveis em Portugal. Dos 135.000 postos de trabalho do sector, 45.000 serão directos e 90.000 indirectos. O impacto no PIB passará de 0,8% para 1,7% até 2020;
5. Desenvolver um *cluster* industrial associado à promoção da eficiência energética assegurando a criação de 21.000 postos de trabalho anuais, gerando um investimento previsível de 13.000 M€ até 2020 e proporcionando exportações equivalentes a 400 M€;
6. Promover o desenvolvimento sustentável, criando condições para o cumprimento das metas de redução de emissões assumidas por Portugal no quadro europeu.

A nova estratégia nacional para a energia promove a competitividade, o crescimento e a independência energética e financeira do País. Recorrendo às energias renováveis pretende-se promover uma sustentabilidade económica e energética, que ao mesmo tempo garanta a redução de emissões de CO₂, gerando benefícios para a sociedade.

1.2 - A Energia Eólica

Dentro do grupo das energias renováveis encontramos a energia eólica, que se entende como a energia resultante da exploração da força do vento. O vento tem origem nas diferenças de pressão causadas pelo aquecimento diferencial da superfície terrestre. A sua passagem pelo sistema de pás de um aerogerador permite o seu deslocamento, originado a conversão de energia cinética em energia eléctrica pela turbina eólica.

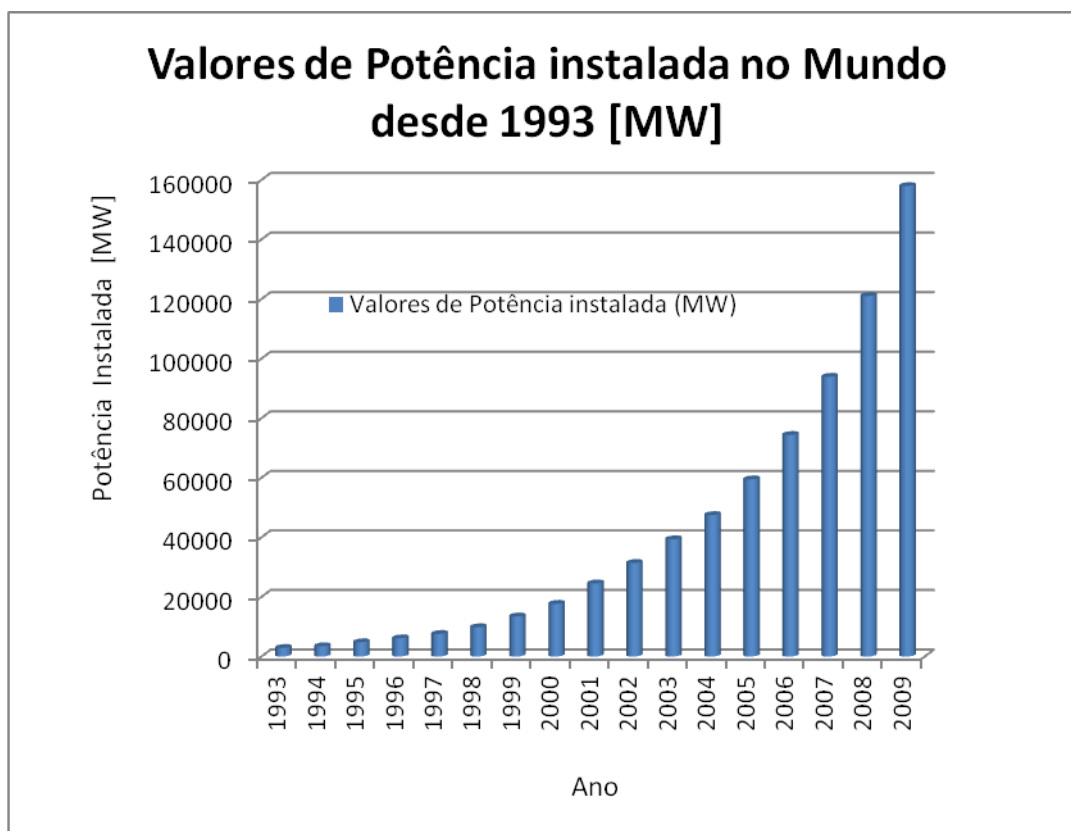


Figura 1.3 - Crescimento da exploração da energia eólica desde 1993 até 2009 [6].

A energia eólica possui um potencial de exploração elevado, que tem vindo a sofrer um crescimento substancial nos últimos 20 anos à escala mundial. Em Portugal, apesar de este não ser um dos países mais ventosos da Europa, existem condições e zonas, que apresentam grande potencial de aproveitamento energético deste tipo de energia.

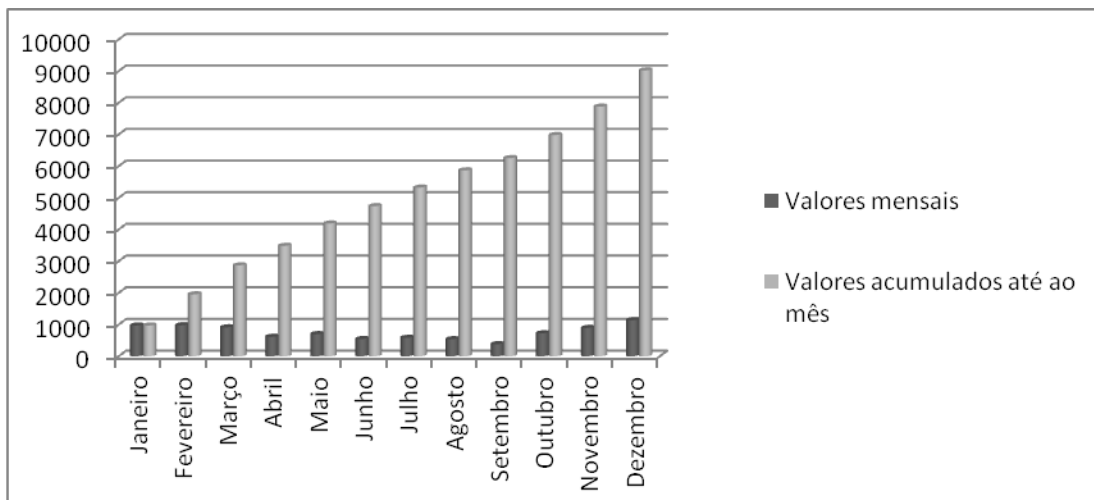


Figura 1.4 - Consumo [GWh] para o ano de 2010 em Portugal [7].

Tabela 1.1 – Potência instalada obtida através do vento em diferentes países europeus em MW [6].

Países	2008	2009	Capacidade Instalada em 2009	Desactivadas em 2009
Alemanha	23896,9	25777	1916,8	36,7
Espanha	16689,4	19148,8	2459,4	0
Itália	3736,5	4850	1113,5	0
França	3542	4521	979	0
Reino Unido	3406,2	4050	645,2	0
Portugal	2862	3535	673	0
Dinamarca	3162,8	3480,6	350,7	32,8
Holanda	2216	2220,7	39,1	34,4
Suécia	1048	1560	512	0
Irlanda	1027	1260	233	0
Grécia	985	1087	102	0
Áustria	994,9	994,9	0	0
Polónia	451,1	705,3	259,4	5,2
Bélgica	392,5	563	171,1	0,6
Hungria	127	201	74	0
Bélgica	150	193,3	43,3	0
Hungria	120	177	57	0
República				
Checa	85,2	149,2	64	0
Bulgária	143	146	4	1
Estónia	54	91	37	0
Finlândia	43,3	43,3	0	0
Lituânia	27	28	1	0
Luxemburgo	9,5	14	4,5	0
Letónia	3,1	3,1	0	0
Roménia	0	0	0	0
Eslováquia	0	0	0	0
Eslovénia	0	0	0	0
Chipre	0	0	0	0
Malta	0	0	0	0
Total EU 27	65172,3	74800,2	9739,1	11,2

Actualmente, Portugal está muito bem posicionado relativamente a outros países e as perspectivas actuais apontam para um crescimento acentuado neste sector, tal como podemos verificar pela análise da tabela 1.1 e da figura 1.3.

A tendência crescente da exploração eólica é verificada de uma forma generalizada a nível Europeu. A sua exploração é realizada um pouco por todos os países, sendo que, se verifica em maior escala naqueles que possuem um grau de maturidade tecnológico suficiente para o realizarem, condições ambientais favoráveis e ainda carência de outras fontes energéticas.

Apesar de todas as vantagens inerentes à exploração eólica, alguns problemas são também identificados, que podem dificultar iniciativas menos estudadas de implementação explorativa. Exemplo disso mesmo são por exemplo as dificuldades encontradas à hora de instalar os aerogeradores em locais com regime de vento favorável, que se encontram geralmente, em montanhas e em zonas remotas, frequentemente servidas por redes eléctricas de fraca capacidade, o que dificulta o transporte da energia produzida. Também a nível ambiental existem limitações que põem em causa a viabilização de alguns projectos, como o ruído, o impacto visual e a influência na avifauna. Contudo, e quer através da consciencialização dos promotores para os cuidados a adoptar na fase de construção, quer ainda pelas inovações tecnológicas que vão sendo incorporadas, a evolução é claramente, no sentido da compatibilização do ambiente, com a tecnologia.

1.2.1 -Turbinas eólicas

Existem, dois tipos de rotor de turbinas eólicas modernas, que se classificam de acordo com o tipo de eixo [8]:

- Eixo horizontal;
- Eixo vertical.



Figura 1.5 - Representação de um aerogerador de rotor com eixo vertical [8].

Em análise, os aerogeradores com rotor de eixo vertical dispensam mecanismos de apoio para as variações da direcção do vento, captando-o em qualquer direcção. Constituem normalmente sistemas de complexidade de projecto menor.



Figura 1.6 - Representação de um aerogerador de rotor com eixo horizontal [8].

Os aerogeradores com rotor de eixo horizontal são os mais comuns. Constituídos por uma estrutura sólida, com duas ou três pás aerodinâmicas, movem-se por acção de forças aerodinâmicas, designadas de forças de sustentação e de arrasto. Devem possuir mecanismos que possibilitem que o disco varrido pelas pás se encontre sempre em posição perpendicular ao vento.

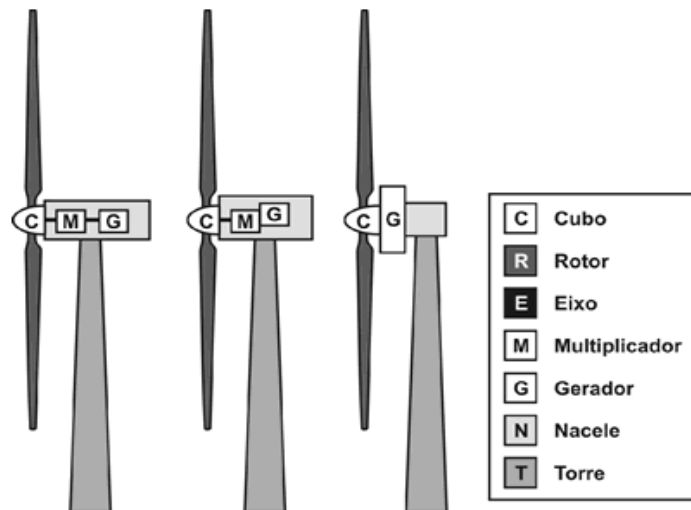


Figura 1.7 - Representação das partes constituintes de aerogerador de eixo horizontal [8].

Na figura 1.5 são apresentados os componentes de um aerogerador de eixo horizontal. Estes aerogeradores distinguem-se pelo tamanho e formato da nacele, pela presença ou não de uma caixa multiplicadora e pelo tipo de gerador utilizado (convencional ou multi-pólos).

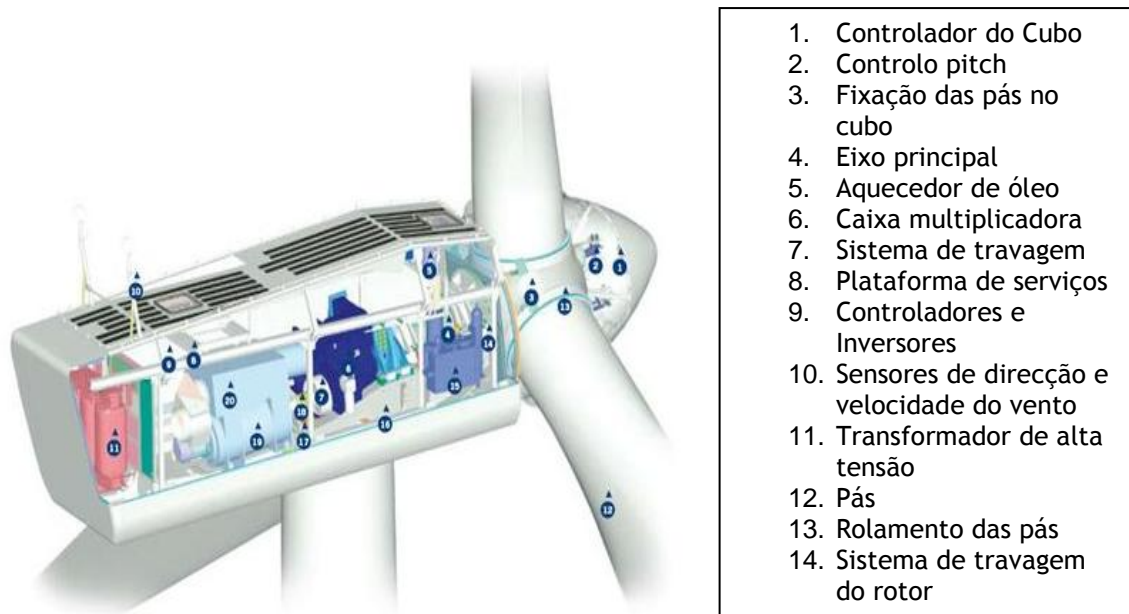


Figura 1.8 - Constituição de uma nacela de gerador convencional [8].

A nacela é um dos componentes principais de um aerogerador e representa o local aonde se encontram alojados o gerador, caixa de engrenagens, o sistema de controlo, sensores de medição do vento e mecanismos de rotação para que seja sempre possível um posicionamento favorável relativamente á direcção do vento.



Figura 1.9 - Constituição de uma nacela de gerador multi-pólos [8].

Existem dois tipos de nacelle usados, que se diferenciam pelo tipo de gerador utilizado, podendo ser do tipo convencional ou multi-pólos.

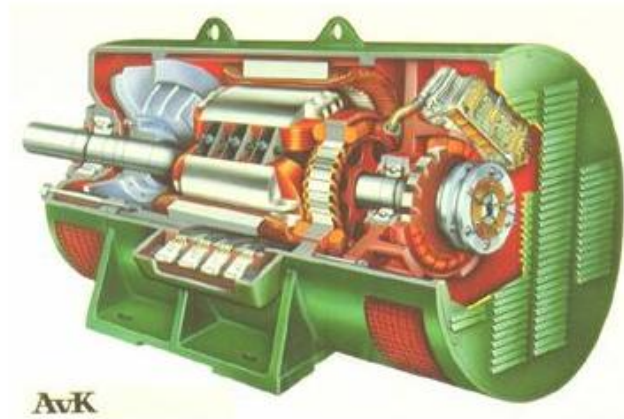


Figura 1.10 - Vista interior de um gerador [8].

Outros dos componentes mais importantes da constituição de um aerogerador, é o gerador e a caixa multiplicadora. Esta permite transmitir a energia mecânica entregue pelo eixo do rotor ao gerador e este realiza a conversão de energia mecânica em energia eléctrica. Alguns problemas decorrem da integração de geradores em aerogeradores para a realização de uma conversão electromecânica, as constantes variações da velocidade do vento e a força mecânica variável que é exercida à entrada do sistema, são apenas dois exemplos de dificuldades encontradas, para que o sistema electroprodutor funcione.

Para o correcto funcionamento de um aerogerador, é necessário que este possua sistemas de controlo. O difícil acesso aos mesmos quando em funcionamento, implica uma necessidade de autonomia gerida por processos de controlo, devido às funções que realizam e à necessidade de controlo sobre processos mecânicos e electrónicos.

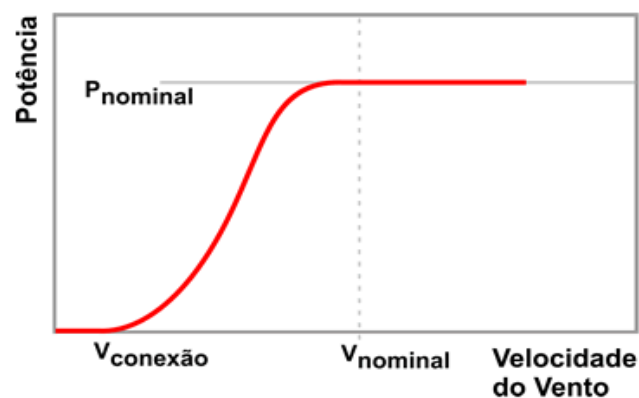


Figura 1.11 - Forma típica da curva de potência de um aerogerador com controlo de *pitch* [8].

Actualmente para a realização de funções de controlo em aerogeradores, é usado o controlo de *pitch*. Este tipo de controlo é usado como estratégia de segurança que permite regular a potência eléctrica do aerogerador através de aumentos do ângulo de calagem das pás e das consequentes diminuições do ângulo de ataque e do coeficiente de sustentação. Sempre que a potência nominal de um aerogerador é ultrapassada as pás giram em torno do seu eixo longitudinal de forma a reduzir as forças aerodinâmicas actuantes e a consequente conversão de potência através do vento. Na figura 1.9 podemos observar a típica curva de potência de um aerogerador com controlo de *pitch*, aonde se pode verificar um comportamento estável a partir do momento em que o aerogerador atinge o valor de potência nominal. Este comportamento deve-se à quebra aerodinâmica proporcionada pelo controlo de rotação das pás do aerogerador, aumentando e diminuído a área de exposição ao vento.

1.3 - Enquadramento da Tese

Dentro das energias renováveis actualmente conhecidas e exploradas, a energia eólica, é a que maior dimensão tem a nível mundial. Em Portugal, a obtenção de energia, recorrendo à força do vento, possui cada vez mais um papel importante relativamente à percentagem total da energia produzida pelo país. Mais exploração energética e novos meios para o fazer, implicam novos sistemas que monitorizem sistemas geradores de electricidade e tratem informação recolhida destes. Assim sendo, de forma natural se entende, que o grau de complexidade e nível de exigência tanto em relação aos níveis de tecnologia do *hardware* utilizado, como em processos de controlo, comunicação e armazenamento de dados, venham sendo melhorados e tratados como ferramentas essenciais para o sucesso.

Actualmente, são instalados em parques eólicos sistemas *Supervisory Control and Data Acquisition (SCADA)* que procedem ao registo de parâmetros essenciais [9] para permitir controlar os equipamentos e retirar conclusões sobre o desempenho da operação dos aerogeradores. Isto permite, entre outras coisas, avaliar rendimentos energéticos e anomalias de funcionamento no sistema produtivo, garantido desta forma melhorias de produtividade.

Nesta dissertação é realizado um estudo sobre metodologias de tratamento e classificação de dados de forma a proporcionar um apoio à análise do desempenho de aerogeradores em parques eólicos.

1.4 - Objectivos

O objectivo desta tese é o desenvolvimento de um sistema que permita proceder à análise e classificação de dados adquiridos por sistemas *SCADA* em parques eólicos. Com esta análise e classificação pretende-se detectar anomalias no sistema, de forma a serem reportadas ao cliente, para melhoramento de desempenho.

Simultaneamente pretende-se realizar o armazenamento dos dados adquiridos, recorrendo a uma base de dados que armazene e imponha um conjunto de relações de integridade na manipulação dos mesmos.

Outro dos objectivos é a implementação de uma interface interactiva e gráfica, que permita especificar a informação que pretendemos analisar e as condições dessa mesma análise. De igual modo, pretende-se realizar a exposição gráfica e textual dos resultados de saída da aplicação e a geração de relatórios de estado de funcionamento, em formato *PDF* e *HTML*.

No que diz respeito à análise e classificação de informação, são estudados diferentes métodos tendo em vista a minimização do erro na formulação da curva de interpolação dos dados e realizada a sua implementação recorrendo a uma linguagem de programação que garanta um rápido processamento e uma flexibilidade adaptativa relativamente a desenvolvimentos futuros que impliquem alterações ao logaritmo desenvolvido.

1.5 - Organização da Dissertação

Esta dissertação é organizada em seis capítulos. No primeiro pretende-se realizar um enquadramento da dissertação relativamente à área temática em que esta está inserida e ao mesmo tempo esclarecer quais os objectivos que se pretende com a elaboração da mesma, fornecendo uma base teórica sobre alguns aspectos mais generalistas.

No capítulo 2 apresenta-se um resumo sobre os desenvolvimentos que têm vindo a surgir na área da aplicação desta dissertação. Projectos e estudos realizados de interesse teórico ou prático para a elaboração desta dissertação.

No capítulo 3 é apresentado o conjunto de fundamentos teóricos em que as metodologias desenvolvidas nesta dissertação se basearam, explicando em que consistem e de que forma é possível a sua aplicação para o desenvolvimento desta dissertação.

O capítulo 4 apresenta o sistema desenvolvido, desde o seu funcionamento às ferramentas utilizadas para o seu desenvolvimento e também a forma como este está organizado, focando as suas características essenciais e a sua aplicabilidade.

No capítulo 5 apresenta-se o conjunto de testes efectuados sobre a aplicação desenvolvida, utilizando um conjunto de dados para comparar os diferentes métodos anteriormente desenvolvidos.

Capítulo 2

Estado da Arte

2.1 - A classificação de dados na produção de energia eólica

A geração de energia através da energia eólica a uma escala industrial é um conceito relativamente novo. É por isso natural que o estudo do desempenho deste tipo de exploração não tenha ainda sido estudado de forma exaustiva, muito embora, já existam um conjunto de metodologias fundamentadas para o fazer. Uma análise pormenorizada do desempenho de um parque eólico por parte da empresa que o explora é, sem dúvida, um factor determinante para o sucesso do projecto.

Para uma empresa que gere um ou vários parques eólicos é importante tomar conhecimento de diversos factores cuja pequena variação ou incumprimento podem levar a prejuízos monetariamente assinaláveis. Para uma empresa que explora este tipo de energia é vital compreender alguns aspectos funcionais, para que saiba como actuar sobre estes. Assim sendo, uma análise rigorosa sobre o desempenho de operação dos seus parques, de forma a compreender e quantificar os erros cometidos em operações anteriores, minimizar a incerteza relativamente a operações futuras e maximizar o *output* energético de saída, são factores fulcrais para que esta funcione a máximo rendimento.

Num passado recente, conjuntos de métodos foram criados, para que, operando sobre dados obtidos por sistemas *SCADA* capazes de adquirir dados operativos do funcionamento de turbinas eólicas, permitissem uma análise por parte das empresas do funcionamento dos seus equipamentos, quando sujeitos a diferentes condições meteorológicas inerentes ao próprio local de instalação dos parques monitorizados.

Ao longo da década 1990-2000 a empresa Garrad Hassan (GH) desenvolveu um conjunto de técnicas de análise de desempenho operacional, procedendo à sua aplicação dentro dos seguintes contextos:

- Identificação de áreas onde o desempenho poderia ser aumentado;
- Diagnóstico de quebras de produção relativamente aos valores esperados;
- Avaliação/reavaliação de projectos durante financiamento/refinanciamento;
- Previsão energética depois de alterações à potência instalada.

Com estes desenvolvimentos, esta empresa procurou sobretudo retirar informação sobre um conjunto de factores técnicos e naturais que podem afectar o correcto funcionamento de um parque eólico. As análises realizadas focam-se essencialmente sobre os factores técnicos do funcionamento das turbinas eólicas, relativamente ao seu desempenho e operacionalidade. Já as análises sobre causas naturais avaliam as características do vento, sobretudo a sua presença no local de instalação dos parques e a estimativa da sua velocidade média a longo tempo. Seguindo uma lógica de análise baseada nos contextos e factores acima descritos, duas técnicas foram desenvolvidas recorrendo à informação registada pelos sistemas *SCADA* instalados nos parques eólicos.

Time Lapse Animation

Técnica baseada numa animação gráfica ilustrativa do desempenho geral de um parque eólico durante o seu período de operação. A animação apresenta a média a cada intervalo de dez minutos dos valores de potência lidos pelo sistema *SCADA* instalado num parque eólico.

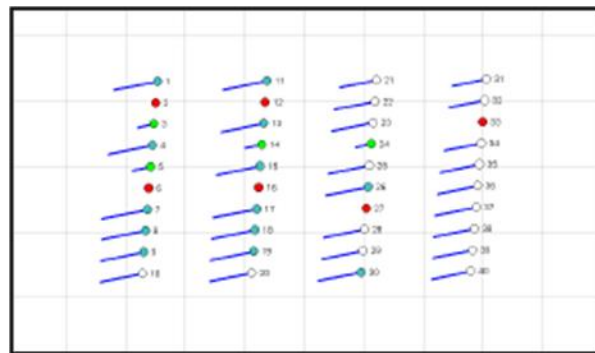


Figura 2.1 - Animação representativa de um conjunto de turbinas eólicas analisadas pelo método *Time Lapse Animation* [10].

Na figura 2.1 ilustra-se um exemplo da animação gráfica desenvolvida para a análise de desempenho de um parque eólico. Cada um dos círculos representa uma turbina eólica na disposição real em que esta se encontra no parque e os traços correspondentes a cada um dos círculos associados, indicam a direcção e a potência referentes a cada turbina. As cores diferentes atribuídas a cada turbina são representativas do seu estado de funcionamento registado a cada intervalo de dez minutos. Branco identifica um funcionamento abaixo de valores nominais, azul-claro uma operação a valores nominais, verde-claro quando é limitado

a valores abaixo do nominal e vermelho quando não está disponível para normal operação. Esta técnica reveste-se de importância pela visão geral que oferece de todo o sistema produtivo, ilustrando os desempenhos individuais de cada turbina.

Power performance and availability tracking

Para a análise do desempenho em termos de produção, torna-se necessário o recurso a uma análise individual dos dados que nos chegam pelo sistema *SCADA* de cada turbina. As causas para um mau desempenho podem ser várias, tais como: deficiente controlo do *pitch*, danos nas pás ou incrustação, redefinição do ângulo de pás, problemas de controlo e limites operativos. Uma vez que as causas para um fraco desempenho, podem ser várias e com diferentes origens, como mencionado anteriormente, estes testes sobre dados que identifiquem a sua ocorrência são necessários para evitar quebras de desempenho e mesmo danos em equipamentos. Por exemplo, para um caso em que se tenha imposto um limite de potência de forma deliberada, por forma a proteger os componentes das turbinas eólicas, os membros da empresa GH desenvolveram uma técnica de análise que permite identificar períodos de baixo rendimento em potência de uma forma gráfica. Pela análise da figura 2.2 pode-se claramente identificar a verde, os períodos de baixo rendimento em potência, devido à acção limitadora imposta, que faz com que valores registados sejam inferiores aos esperados para a velocidade de vento registada.

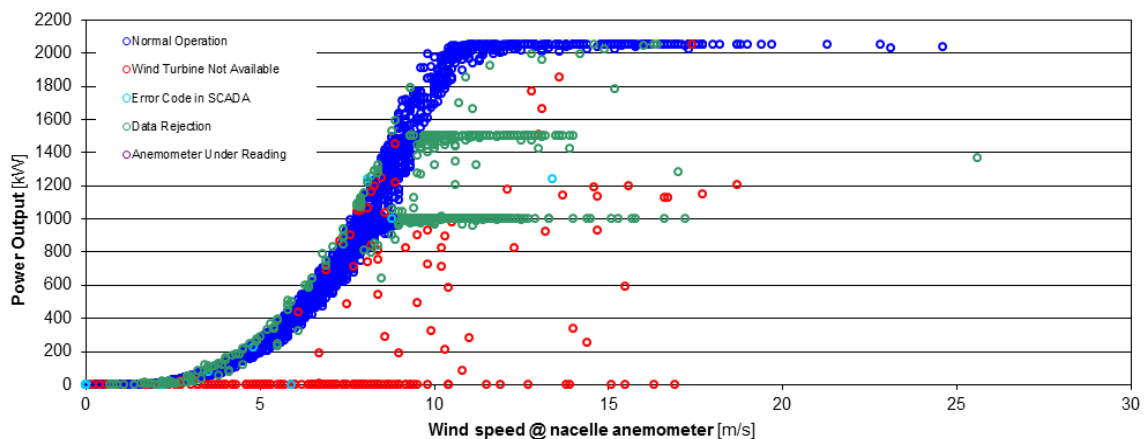


Figura 2.2 - Exemplo de Turbina com limites de desempenho em potência [10].

Outro factor importante a ser considerado na classificação do desempenho de um parque eólico é a disponibilidade de operação de cada turbina eólica. Também para estes períodos uma outra técnica foi desenvolvida, que consiste na verificação dos valores registados de potência. Estas situações são facilmente identificáveis, procurando por períodos de valor de potência a zero ou inferiores e se para esses mesmos períodos, a velocidade registada é

superior à velocidade mínima de corte. Pela análise da figura 2.2 esses períodos encontram-se identificados a cor vermelha [10].

Uma metodologia usada para identificação e classificação de inconsistências oriundas de dados obtidos por um sistema SCADA, pode ser realizada desenvolvendo um plano de estudos baseado nas seguintes tarefas [10].

- Estudo de períodos operativos de um parque eólico e verificação do histórico de manutenção;
- Realização de mineração de dados e verificações de consistência;
- Análise de dados filtrados e aplicação de relações paramétricas;
- Identificação de cenários de referência, padrões e índices;
- Estudo de indicadores e tendências.

Realizando um estudo de relação entre parâmetros de operação e partes constituintes de um parque eólico, torna-se possível avaliar o desempenho em termos de produção do mesmo. A análise efectuada pela MegaJoule para medição de desempenho apoia-se em três grupos fundamentais: relações entre torres de monitorização meteorológica e turbinas eólicas, relações entre turbinas eólicas e relações entre torres de monitorização meteorológica.

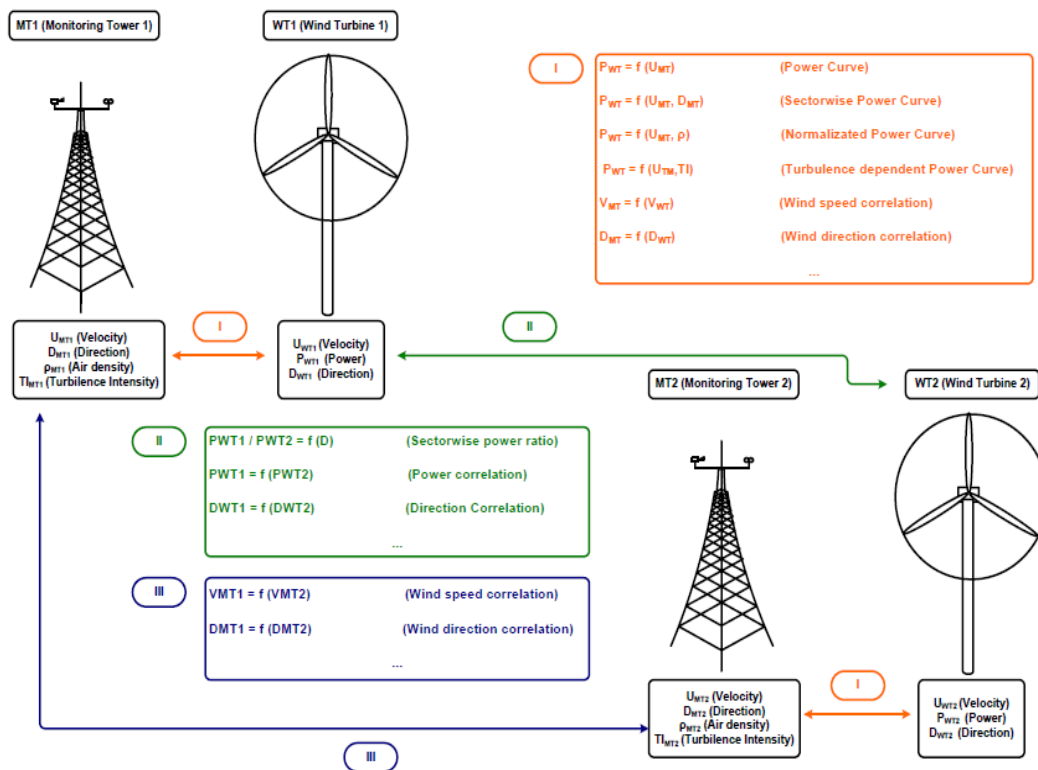


Figura 2.3 - Representação esquemático do processo metodológico implementado pela MegaJoule [11].

A simples análise dos valores obtidos por um sistema *SCADA* não é suficiente para perceber o modo operativo de um parque eólico. Problemas de consistência e clareza de valores podem-se verificar e devem ser ultrapassados antes de qualquer análise dos dados. Algumas das maiores falhas encontradas e que devem ser alvo de reparação antes de qualquer tipo de análise são:

- Ausência de registos de manutenção;
- Inconsistência nos registos de estado das turbinas eólicas;
- Taxa de aquisição inconstante;
- Registos fora de alcance.

Algumas análises preliminares podem ser realizadas, relacionando parâmetros registados pelo sistema *SCADA*. Recorrendo a dados de potência, previamente filtrados de inconsistências e erros, obtidos através das saídas de duas turbinas vizinhas de um parque eólico com potência nominal de cada uma de 600 kW, é possível identificar, tal como ilustra a figura 2.4, uma zona de boa correlação entre a potência das duas turbinas.

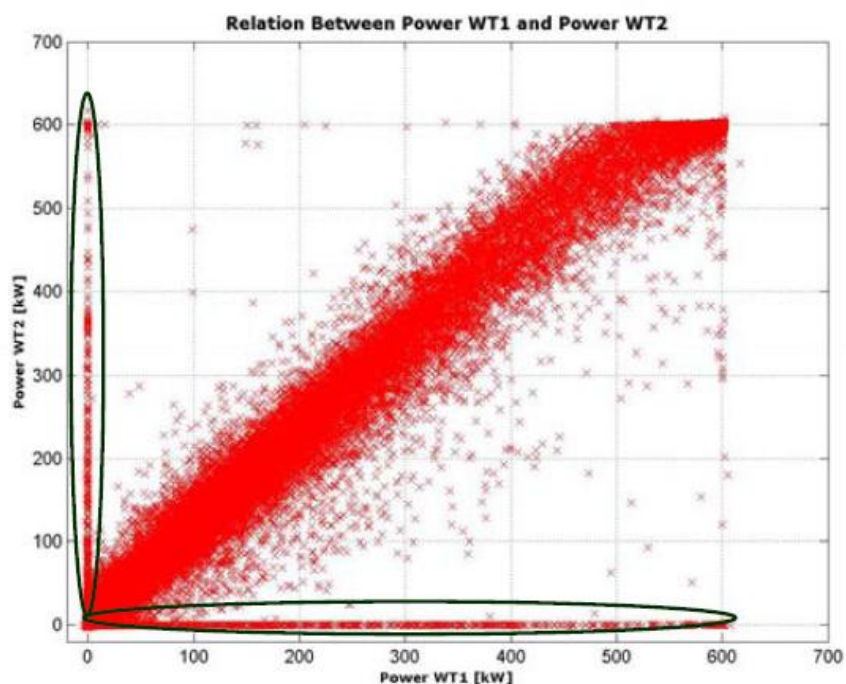


Figura 2.4 - Representação da correlação cruzada de potências entre duas turbinas vizinhas [11].

Obtendo-se a correlação entre potências, é possível aplicar posteriormente um filtro para os valores registados em períodos de não funcionamento ou períodos fora do alcance dos limites superiores e inferiores de funcionamento. A figura 2.5 representa os dados obtidos

após aplicação de este mesmo filtro e o padrão linear da correlação entre potências das duas turbinas após filtragem de dados.

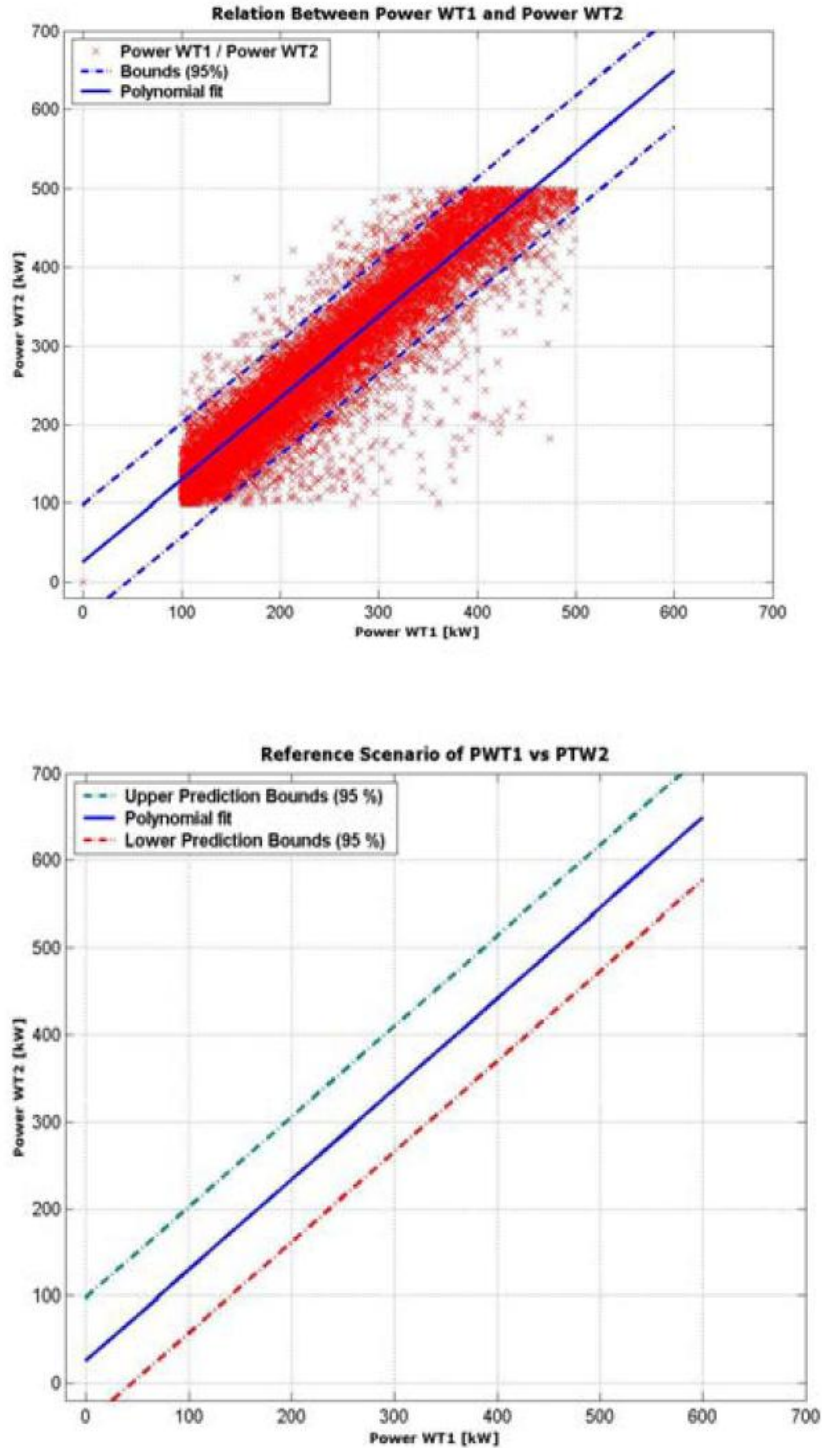


Figura 2.5 - Representação filtrada da correlação cruzada de potências entre duas turbinas vizinhas e padrão linear da correlação entre potências [11]

A partir deste processamento e obtidos os resultados finais, uma boa correlação entre a produção realizada pelas duas turbinas eólicas é identificada.

A empresa MegaJoule emprega ainda uma outra metodologia para perceber de que forma se comportam as turbinas eólicas em função da direcção do vento, relacionando rácios de potência com a direcção do vento. As figuras 2.6 e 2.7 são representativas da metodologia acima descrita para o mesmo conjunto de turbinas eólicas e para o mesmo conjunto de dados colectados pelo sistema *SCADA*. Em particular na figura 2.7, é ilustrado uma relação entre parâmetros após filtragem de dados referentes a períodos de não funcionamento ou períodos fora do alcance dos limites superiores e inferiores de funcionamento, assim como, um padrão linear dessa mesma relação.

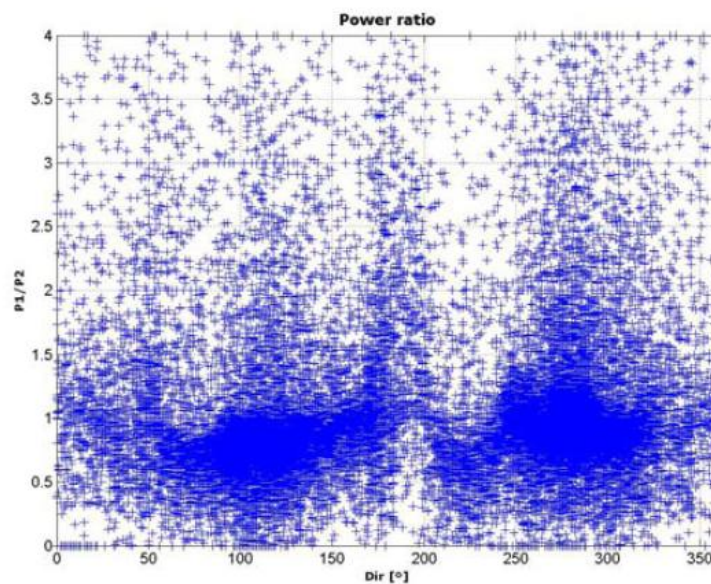


Figura 2.6 - Representação da relação entre rácio de potência e direcção do vento [11].

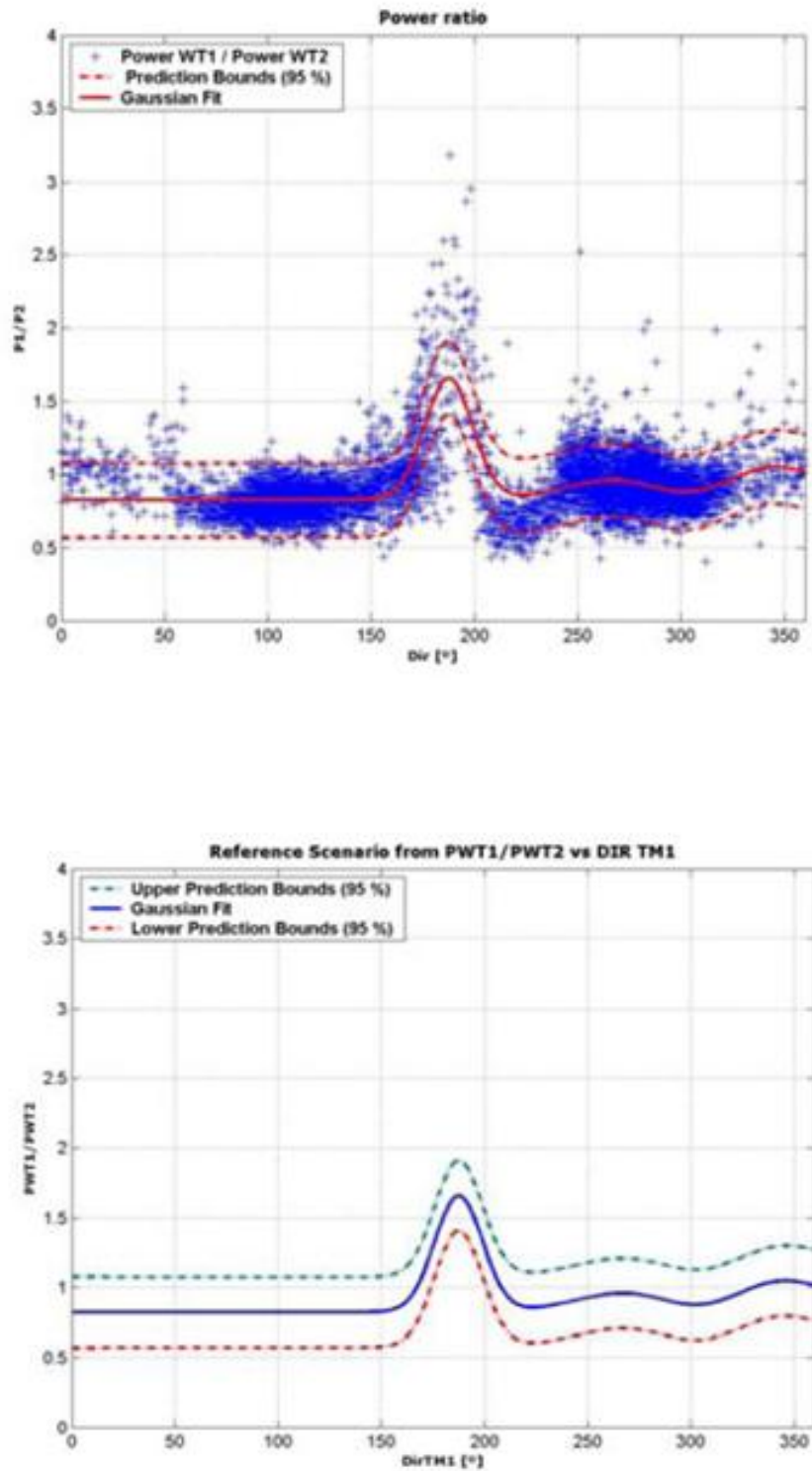


Figura 2.7 - Representação da relação entre rácio de potência e direcção do vento, com filtragem de dados e uma representação padrão e linear da mesma relação [11].

Realizando uma análise final sobre os testes desenvolvidos pela empresa MegaJoule, verifica-se que relações paramétricas podem ajudar a monitorizar a performance de parques eólicos. Avaliando tendências ou padrões de comportamento é possível avaliar o funcionamento de parques, alertando responsáveis de gestão e facultando informações úteis de como estes devem operar para obtenção de elevados desempenhos energéticos para os seus parques.

Os sistemas *SCADA* instalados nos parques eólicos reportam normalmente anomalias de funcionamento das turbinas e restantes equipamentos, mas raramente são capazes de realizar uma análise da causa dessa mesma anomalia, ou avaliar esses erros dentro de um contexto financeiro de operação. A utilização da informação registada, sem que seja processada sob a forma de classificação das situações de operação, resulta em decisões erradas por parte dos operadores, o que se reflecte em custos financeiros.

A *Natural Power Consultants Ltd*, concentrou os seus esforços na criação de uma ferramenta de análise do tempo de inactividade, que permite que todos os eventos anómalos detectados sejam alocados em categorias de falhas e causas [12]. Esta classificação é realizada independentemente do modelo da turbina e do sistema *SCADA* utilizado. Um sistema de monitorização e armazenamento *SCADA* revela-se por vezes incapaz de registar de forma precisa períodos de inactividade. Isto deve-se normalmente a problemas de *software*, tais como:

- Encobrimento de dados devido a cascatas de alarmes;
- Perdas de dados;
- Estados operacionais das turbinas não registados.

Um esquema de dados para um sistema de análise de tempo de inactividade, que possibilita evitar as deficiências inerentes aos sistemas *SCADA* na classificação de informação, pode ser visto na figura 2.8.

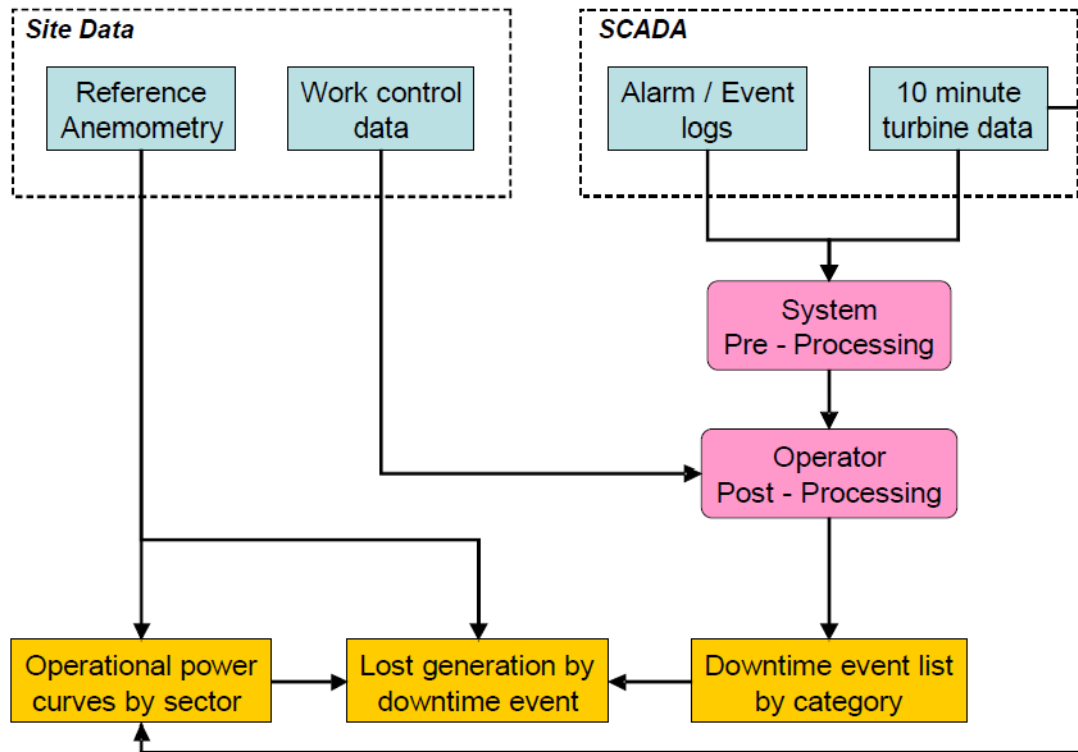


Figura 2.8 - Esquema proposto para análise do tempo de inactividade do sistema [12].

A informação registada pelo sistema *SCADA* de uma turbina pode não conter dados suficientes para realizar análises precisas do tempo de inactividade, contudo, existem mais do que uma fonte de dados disponível que o permite fazer, tais como:

- *SCADA* - registos realizados de 10 em 10 minutos de dados operacionais;
- *SCADA* - Alarmes/Eventos;
- Controlo de dados de trabalho e relatórios técnicos;
- Referências de anemometria.

Toda esta informação permite um cruzamento de dados, seja para controlar ou para substituir quando um ou mais fluxos de dados não estão disponíveis. Os dados adquiridos podem ser analisados automaticamente como uma preparação para uma análise mais profunda. Inicialmente, essa análise consistirá em identificar os períodos de tempo de inactividade a partir dos dados operacionais, bem como quaisquer outros períodos desconhecidos de operação, passando-se de seguida a uma análise detalhada de cada período de inactividade de forma individual, para que se conheçam quais as condições existentes nesses mesmos períodos. De realçar, que dessas mesmas condições fazem parte condições ambientais ou estados conhecidos da rede obtidos através de registos de alarmes.

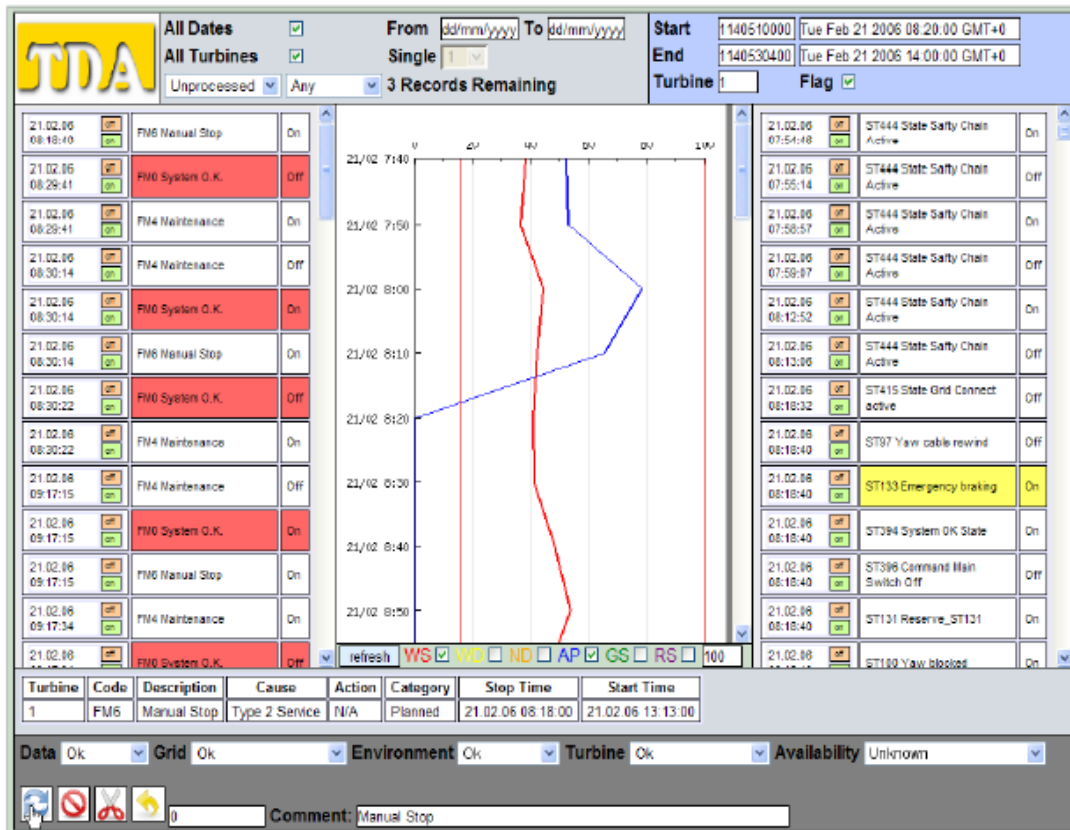


Figura 2.9 - Interface pós-processamento para avaliação de dados por parte do operador [12].

O pré-processamento de períodos de inactividade pode ser apresentado por intermédio de um interface a um operador, de forma, a que seja possível a este realizar uma visualização do fluxo de dados correspondente a períodos de inactividade e actuar sobre diferentes tipos de anomalias detectadas. Como resultado final, destas medidas interventivas, seguem-se períodos de manutenção para correcção e reavaliação do sistema.

Os custos relativos à operação e manutenção de parques eólicos são factores de extrema importância. Além das quebras de rendimento energético provocadas pelas características do vento, outras existem que são provocadas por falhas de sistema e que implicam custos elevados de reparação. Devido aos custos associados às falhas de sistema, é de extrema importância a implementação de técnicas de detecção e prevenção que possam além de prevenir falhas em componentes, reduzir custos de manutenção e fornecer informações detalhadas de outros parâmetros operacionais úteis ao bom funcionamento.

Três das falhas mais comuns que podemos encontrar no funcionamento de um parque eólico são: falhas eléctricas, electrónicas e mecânicas. Estas falhas estão associadas a componentes das turbinas eólicas, como o sistema de pás, a caixa de velocidades e o gerador, que são de difícil reparação dado o seu custo elevado e o tempo necessário para a sua realização. De forma a evitar que estas falhas se dêem, foram desenvolvidas técnicas e

ferramentas no sentido de evitar avarias, através da análise das condições de funcionamento dos mais diversos componentes constituintes de uma turbina eólica. Muitas dessas técnicas e ferramentas têm como base a análise das correntes circulantes, cuja discrepância de valores em relação aos valores normais de funcionamento, poderá significar existência de falhas.

Análise da Corrente do Estátor

Esta é uma técnica que resulta de uma análise do espectro de correntes do estator para detectar possíveis falhas. Fazendo uso da mesma é possível detectar barras quebradas pela simples monitorização da corrente espectral que circula nos componentes que sofre variações provocadas pelas barras quebradas.

Análise de Velocidade do Vento vs Potência produzida

Os dados obtidos através de sensores instalados em aerogeradores e posteriormente colectados por sistemas SCADA, podem dar informações úteis das condições de funcionamento dos aerogeradores de um parque eólico. A comparação entre valores reais e esperados da relação velocidade do vento e potência, permite tirar conclusões sobre o estado de “saúde” de um aerogerador, podendo uma diferença acentuada entre esses dois valores significar uma degradação do sistema de pás e levar à necessidade de aplicação de limites sobre o seu funcionamento, por forma a evitar problemas futuros.

Análise de Temperatura vs Potência

Fazendo recurso de redes neuronais para procurar correlações entre potência e temperatura num aerogerador, é possível prever aquilo que se considera serem as condições de normal funcionamento para os seus componentes. Desta forma, comparando valores reais com os valores previstos calculados através das redes, consegue-se identificar falhas de operação e evitar problemas graves.

2.2 - Métodos de classificação de dados utilizados em diversas áreas

Metodologias que nos permitam entender informação para saber como actuar correctamente perante problemas, sempre fizeram parte da vida do Homem devido à necessidade que este possui de ultrapassar problemas e de desenvolver métodos para o fazer. Com a revolução industrial e a evolução tecnológica da sociedade, a quantidade de informação e a sua complexidade incrementaram, levando à necessidade de desenvolver metodologias sofisticadas para tratamento, descodificação e classificação de dados em áreas diversas. O consumo de energia eléctrica não é uma excepção, no que à necessidade de estudo de parâmetros e condições de funcionamento de equipamentos diz respeito.

Num estudo realizado pela Universidade de Washington, foram utilizados dados recolhidos do sistema eléctrico de uma casa por um período de três meses para estudar possíveis irregularidades de consumo energético. Foram também utilizados dados sintetizados e gerados aleatoriamente para um período de doze meses, nos quais foram injectados *outliers* representativos de diferentes cenários. Estes *outliers* podem ter origem nas mais diversificadas situações, representando situações anormais de gastos de energia [13].

Tabela 2.1 – Detalhes sobre os dados utilizados no estudo da Universidade de Washington [13].

Tipo de Dados	Conjunto total de dados	Período de aquisição/geração
Reais	1644	3 meses
Sintetizados	8748	12 meses

Para obtenção destes conjuntos de dados foi utilizado um *software* designado de Casas, que além de proporcionar dados relativamente ao consumo eléctrico de uma casa e possibilitar o gerar de dados sintetizados, possibilita ainda a aquisição de dados como temperatura e consumo de água.

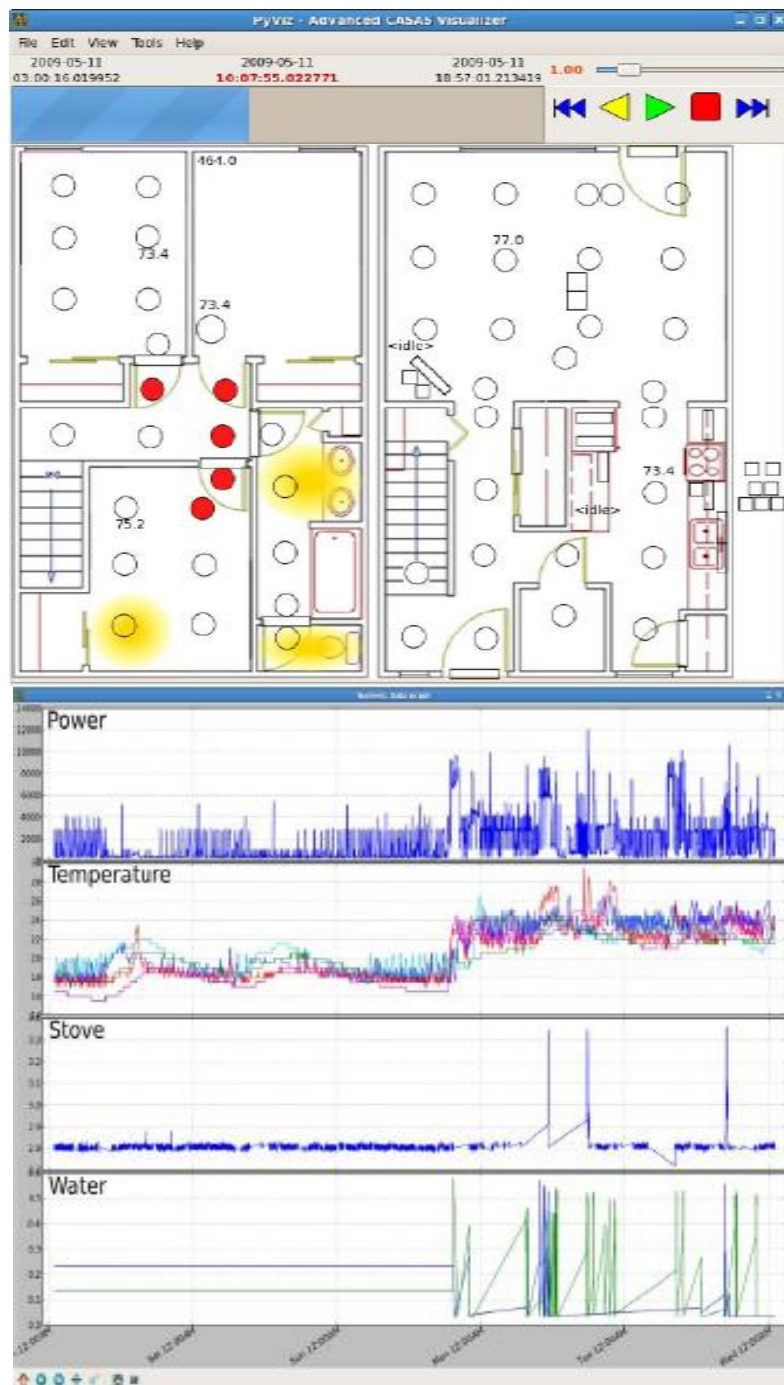


Figura 2.10 - Interface do *software Casas* utilizado neste estudo [13].

Com os conjuntos de dados reais e sintéticos obtidos, dois métodos foram implementados para detecção de outliers de forma, a que com este conhecimento a otimização do consumo energético da casa seja conseguido.

Detecção de *outliers* - Abordagem Estatística

Por forma, a realizar uma análise sobre os dados de consumo energético obtidos, a implementação de algoritmos estatísticos para esse fim, pode ser um método de grande utilidade. Nesta experiência um algoritmo estatístico foi desenvolvido, para identificar *outliers*, procurando situações irregulares de consumo energético e classificando-as por ordem de gravidade. Primeiramente este algoritmo calcula a média e desvio padrão de um conjunto de dados em análise, usando-os de seguida para estabelecer um critério de selecção entre valores, classificando-os como *outliers* ou não. De seguida realiza ainda uma ordenação dos *outliers* identificados de acordo com o seu grau de importância.

```

Step 1:
Define Set of outlier = {} // start with empty.

Step 2:
For i = 0 to winSize // winSize is defined by the experimenter, helps as
moving window.
    Compute average (A) of elements in set X
    Compute Standard Deviation(S) of Elements in set X
    If S>0
        Find i extreme in X // we can use standard deviation based distance
measures here.
        Calculate standardized element:  $R(i) = \frac{x(i) - A}{S}$ 
        Calculate critical value of i: alpha
         $\alpha(i) = (n - \text{no of outliers}) * t\text{-score of limited sample} / ((n - \text{outliers} - 1) * t\text{-score of limited sample})^{1/2}$ 
        If  $R(i) > \alpha(i)$  then outlier
        Remove i from set; i is added to the outlier set
We continue till end is reached.

Step 3: Ranking (set of outliers);

```

Figura 2.11 - Pseudo código do algoritmo estatístico para encontrar *outliers* [13].

Detecção de *outliers* - Abordagem por aglomeração

Numa segunda experiência realizada sobre o mesmo conjunto de dados, um outro método foi utilizado para detecção de *outliers*. Os algoritmos *k-Nearest Neighbor* (KNN) e *Discrete*

Time Warping (DTW) foram implementados de forma a possibilitar a distinção entre dados normais e anômalos, através da aglomeração de dados.

```

For each element X in the data do
  If X has unknown distance or missing data
    X is abnormal
  Else then
    For each element in data
      Calculate distance
    Find the biggest score of X in distance measure
    Calculate the average distance for k nearest neighbors
    If average distance is greater than threshold then
      X is normal
    Else then
      X is abnormal

Declare DTW [0..n, 0..m]
Declare i, j, cost
  For i: = 1 to m
    DTW [0, i]:= infinity
  For i: = 1 to n
    DTW [i, 0]:= infinity
  DTW [0, 0]:= 0
  For i: = 1 to n
    For j: = 1 to m
      Cost: = d(s[i], t[j])
    DTW [i, j]:= cost + minimum (DTW [i-1, j], // insertion
    DTW [i, j-1], // deletion
    DTW [i-1, j-1]) // match
Return DTW [n, m]

```

Figura 2.12 - Pseudo código dos algoritmos KNN e DTW para detecção de *outliers* [13].

Com estes dois métodos, duas experiências distintas são aplicadas. A primeira consiste em comparar a análise de *outliers* nos dados reais e sintéticos utilizando para isso, as duas técnicas de detecção apresentadas anteriormente. A segunda experiência consiste em variar os *outliers* no conjunto de dados sintéticos e verificar de que forma as técnicas se comportam com essas variações, ou seja, se os identificam correctamente.

Tabela 2.2 – Comparação das técnicas de identificação de *outliers* para o conjunto de dados Reais e Sintéticos.

Conjunto de Dados	Conjunto total de dados	<i>Outliers</i> actuais	<i>Outliers</i> identificados por via estatística	<i>Outliers</i> identificados por agrupamento
Reais	1644	N/A	86	2
Sintetizados	8748	16	104	16

Facilmente se verifica por análise da tabela 2.2 que as duas técnicas quando aplicadas sobre os dois conjuntos de dados chegam a resultados diferentes, sendo que a técnica de abordagem por aglomeração supera a técnica por abordagem estatística, uma vez que é mais precisa na detecção de *outliers*, detectando-os quando estes realmente existem e obtendo um erro mínimo de classificação, o que já não acontece com os métodos estatísticos.

Tabela 2.3 – Comparação das técnicas de identificação de *outliers* para o conjunto de dados Sintéticos, com variação de *outliers*.

Processo Experimental	<i>Outliers</i> nos dados sintetizados	Identificados	Correctamente Identificados	Incorrectamente Identificados	Precisão
	2	2	1	1	50%
<i>Outliers</i> identificados por via estatística	5	3	3	0	100%
	50	39	12	17	30,70%
	80	52	28	24	35%
	100	69	34	35	34%
<i>Outliers</i> identificados por agrupamento	2	2	2	0	100%
	5	5	5	0	100%
	50	47	47	0	94%
	80	82	69	13	92,5%
	100	110	86	24	92%

Dos resultados, da tabela 2.3, concluímos que a técnica por abordagem estatística para um número limitado de dados, apresenta um comportamento bastante satisfatório para pequenos conjuntos de outliers, perdendo desempenho quando este número aumenta. Já a abordagem por aglomeração de dados se mantém com um desempenho constante, fazendo variar o número de outliers.

2.3 - Comparação entre linguagens de programação passíveis de serem utilizados

Uma linguagem de programação serve para que um programador expresse instruções a um computador, actuando sobre dados e definindo acções a serem executadas. A escrita destas instruções permite que de uma forma relativamente simples ao raciocínio humano, o programador expresse aquilo que pretende que o computador realize, que será posteriormente traduzido para código máquina e executado pelo processador.

Existem diferentes aspectos, que devem ser considerados numa análise comparativa entre diferentes linguagens de programação, por forma a eleger qual delas, se adapta melhor ao propósito que se pretende. Como estes aspectos nem sempre são de simples compreensão, apresenta-se uma breve descrição de alguns deles e uma comparação entre diferentes linguagens de programação baseadas nestes mesmos aspectos [14]. Os dados registados na tabela permitem concluir qual a escolha a ser realizada para a linguagem de programação a utilizar no desenvolvimento da aplicação que se pretende implementar com esta dissertação. Não existe uma linguagem perfeita, sendo que cada uma possui diferentes características que as podem valorizar para um determinado uso. Para os objectivos desta dissertação nem todas as características mencionadas na tabela 2.4 são consideradas importantes, aspectos como a possibilidade de conexão com sistemas de gestão de bases de dados, características da sintaxe da linguagem, integração com outras linguagens, a sua eficiência de processamento e ser uma linguagem orientada a objectos, são factores a ter em conta para sua eleição.

Tabela 2.4 – Comparação entre diferentes tipos de linguagens de programação.

Características da Linguagem	Eiffel	Ruby	Java	C++	Python	Perl	Visual Basic
Object-Orientated	Pura	Pura	Híbrida	Híbrida / Multi-Paradigma	Híbrida	Add-On / Híbrida	Suporte Parcial
Static / Dynamic Typing	Estática	Dinâmica	Estática	Estática	Dinâmica	Dinâmica	Estática
Generic Classes	Sim	N/A	Não	Sim	N/A	N/A	Não
Inheritance	Múltipla	Classe única, múltiplos "mixins"	Classe única, múltiplos interfaces	Múltipla	Múltipla	Múltipla	Nenhum
Feature Renaming	Sim	Sim	Não	Não	Não	Não	Não
Method Overloading	Não	Não	Sim	Sim	Não	Não	Não
Higher Order Functions	Agentes com versão 5	Blocos	Não	Não	Expressões Lambda	Sim	Não
Lexical Closures	Sim	Sim (blocos)	Não	Não	Sim	Sim	Não
Garbage Collection	Marca e Limpa ou Geracional	Marca e Limpa	Marca e Limpa ou Geracional	Nenhum	Contagem de referência	Contagem de referência	Contagem de referência
Uniform Access	Sim	Sim	Não	Não	Não	Não	Sim
Class Variables / Methods	Não	Sim	Sim	Sim	Não	Não	Não
Access Control	Exportação Selectiva	Publico, Protegido, Privado	Publico, Protegido, "package", Privado	Publico, Protegido, Privado	Nome Desconfigurado	Nenhum	Público e privado
Multithreading	Implementação Dependente	Sim	Sim	Livrarias	Sim	Não	Não
Regular Expressions	Não	Built-in	Standard Livraria	Não	Standard Livraria	Built-in	Não
Pointer Arithmetic	Não	Não	Não	Sim	Não	Não	Não
Language Integration	C, C++, Java	C, C++, Java	C, some C++	C, Assembler	C, C++, Java	C, C++	C (via DCOM)
Built-In Security	Não	Sim	Sim	Não	Não	Sim	Não

1

¹ Explicação detalhada do conteúdo da tabela 2.4 em anexo na secção C

De forma a proceder ao armazenamento de dados oriundos de um qualquer sistema, para monitorização e controlo, é de boa prática recorrer a bases de dados de forma a colocar toda a informação recolhida de forma organizada e pronta para qualquer manipulação que se pretenda exercer sobre a mesma.

As bases de dados são projectadas para oferecer um mecanismo organizado para armazenar, gerir e recuperar informações, o que é conseguido, fazendo recurso a tabelas, constituídas por colunas e linhas. Cada coluna contém um tipo diferente de atributo e cada linha corresponde a um único registo.

Dentro do grupo de acções que normalmente se podem executar recorrendo a uma base de dados temos:

- Recuperação de registos que correspondem a determinados critérios;
- Actualização de registos em massa;
- Referenciação cruzada, de registos em tabelas diferentes;
- Realização de cálculos complexos de agregação;

Há uma grande variedade de soluções em gestão de bases de dados existentes no mercado, pelo que o processo de selecção da melhor solução a utilizar para um dado projecto, deve por razões económicas e práticas, ser objecto de estudo antes de uma possível compra de *software* ou utilização do mesmo.

De entre os sistemas de gestão de bases de dados mais utilizados temos: *Oracle*, *SQL Server* e *MySQL*. Estes são alvo de uma breve análise nesta tese, de forma a estudar vantagens e desvantagens de cada opção para a criação de um sistema de apoio e análise de desempenho de parques eólicos. Apresenta-se de seguida uma análise comparativa entre diferentes tipos de sistemas de gestão de Bases de Dados, tendo em conta, alguns dos seus aspectos mais importantes.

ACID (*Atomicity, Consistency, Isolation and Durability*)

Numa base de dados denomina-se ACID ao conjunto de características necessárias para que uma série de instruções possam ser consideradas como uma transacção. Para que um SGBD (Sistema de Gestão de Base de Dados) seja considerado ACID necessita de verificar 4 condições:

- **Atomicity** - é a propriedade que assegura que uma operação se realizou ou não;
- **Consistency** - é a propriedade que assegura que só se começa aquilo que se pode acabar. Portanto executam-se aquelas operações que não vão romper as regras e directrizes de integridade da base de dados;
- **Isolation** - é a propriedade que assegura que uma operação não pode afectar a outras. Assegura que as realizações de duas transacções sobre a mesma informação sejam independentes e não gerem nenhum tipo de erro;
- **Durability** - é a propriedade que assegura que uma vez realizada a operação, esta persistirá e não poder-se-á desfazer ainda que o sistema falhe.

Tabela 2.5 – Comparação de características entre diferentes tipos de SGBD.

SGBDs	ACID	Referential Integrity	Transactions	Unicode	Interface	Materialized View	Temporary Table
<i>Firebird</i>	Sim	Sim	Sim	Sim	SQL	Não (só vistas comuns)	Sim
<i>Microsoft Access</i>	Sim	Sim	Sim	Sim	GUI & SQL	Não	Sim
<i>Microsoft SQL Server</i>	Sim	Sim	Sim	Sim	GUI & SQL	Sim	Sim
<i>MySQL</i>	Sim	Sim	Sim	Sim	SQL	Não	Sim
<i>Oracle</i>	Sim	Sim	Sim	Sim	API & GUI & SQL	Sim	Sim
<i>PostgreSQL</i>	Sim	Sim	Sim	Sim	GUI & SQL	Não	Sim
<i>SQLite</i>	Sim	Sim	Sim	Opcional	SQL	Não	Sim
<i>Firebird</i>	Sim	Sim	Sim	Sim	SQL	Não (só vistas comuns)	Sim
<i>Microsoft Access</i>	Sim	Sim	Sim	Sim	GUI & SQL	Não	Sim
<i>Microsoft SQL Server</i>	Sim	Sim	Sim	Sim	GUI & SQL	Sim	Sim

2

² Explicação detalhada do conteúdo da tabela 2.5 em anexo na secção B

Capítulo 3

Metodologias Desenvolvidas

Técnicas e metodologias para classificação de dados, são ferramentas amplamente usadas em diversas áreas. Na sua implementação é necessário recorrer a regras que as validem e que permitam a obtenção de resultados coerentes e úteis para o domínio em que as pretendemos aplicar. Para a avaliação do desempenho de uma regra de classificação é necessário avaliar a sua aplicação sobre um conjunto de exemplos tomados como padrão e cujo valor de saída é conhecido. Para um bom desempenho devem verificar medidas como [15]:

- Aplicabilidade - proporção de exemplos aos quais o método é aplicável;
- Exactidão - proporção de exemplos para os quais a regra retorna valores coincidentes com o original, ou seja, as condições impostas por este são verdadeiras;
- Cobertura - proporção de exemplos classificados correctamente pela regra, no conjunto de exemplos aos quais a regra é aplicável.

Existem algoritmos que permitem gerar regras de classificação que podem ser usadas na implementação de diferentes metodologias de classificação de dados. Um exemplo deste tipo de algoritmos é o algoritmo 1R, que é de implementação muito simples e se baseia na implementação de um conjunto de regras que testam um único atributo em particular. A ideia consiste em criar um conjunto de regras por cada atributo particular e testá-las seleccionando o conjunto que apresentar menor erro.

No capítulo anterior apresentou-se uma análise sobre o estado da arte, referindo o que nos últimos anos foi desenvolvido na área da análise e classificação de dados, com especial foco na área da energia eólica. No âmbito da presente dissertação, a realização de um sistema de apoio para análise de desempenho em parques eólicos e respectiva implementação assenta num conjunto de metodologias, que quando aplicadas sobre conjuntos de dados fornecidos por clientes, permitam obter uma informação detalhada sobre o desempenho de funcionamento de cada parque.

As metodologias implementadas recorrem a critérios estatísticos, modelos de previsão e a aproximações à curva teórica dos aerogeradores, para proceder à classificação de dados, identificado assim registos *outliers* de funcionamento.

3.1 - Metodologia de classificação de dados baseada no desvio padrão

Esta é uma metodologia que recorre ao desvio padrão para uma medição da dispersão de modo a filtrar períodos de funcionamento anómalo de um conjunto de dados recebidos do cliente [16]. O desvio padrão (σ) mede a variabilidade dos valores em torno da sua média e é dado por [16]:

$$\sqrt{\frac{(X_1 - \bar{X})^2 + (X_2 - \bar{X})^2 + \dots + (X_n - \bar{X})^2}{n - 1}} = \sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n - 1}} \quad (1.1)$$

onde \bar{X} representa a média e n o número de valores da amostra.

Fazendo relacionar variáveis como Potência, Coeficiente de Potência e Rotação por minuto (RPM) da Turbina Eólica em função da Velocidade do Vento, é possível retirar conclusões sobre o funcionamento de um aerogerador. Para aplicação desta metodologia torna-se necessário primeiramente obter as relações dessas variáveis e ordenar a informação em função da velocidade do vento, agrupando-a por barras (*bins*) de largura correspondente a incrementos de velocidade de 0,5 m/s numa escala variável, dependendo do tipo de máquina e da velocidade do vento, mas que normalmente varia entre os 0 e 25 m/s.

$$|\text{Variável}(x) - \text{Variável Média (bin)}| \leq 2 \times \sigma (\text{bin}) \quad (1.2)$$

onde σ o desvio padrão.

Fazendo uma análise bin a bin e procedendo ao cálculo do desvio padrão e média dos registos em análise (Potência, Coeficiente de Potência ou Rotação por minuto da Turbina Eólica) é possível proceder à identificação de *outliers* de funcionamento, pela verificação ou não da inequação 1.2. Esta inequação verifica se os registos analisados apresentam uma discrepância considerada elevada em relação aos valores esperados. Considera os registos como sendo *outliers*, se a diferença do valor a analisar num determinado bin pela média desse mesmo *bin*, é superior a 2 desvios padrões. A opção pela medida de dois desvios padrões, usada na inequação 1.2, como medida de referência para teste de comparação, era

já anteriormente aplicada pelo INEGI, nas metodologias anteriormente desenvolvidas, que devolve resultados considerados satisfatórios para uma boa análise de desempenho.

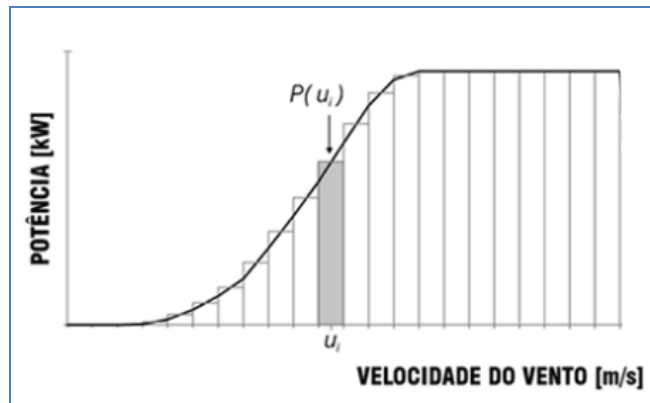


Figura 3.1 - Representação teórica da Relação Potência-Velocidade do Vento fornecida por um fabricante [17].

3.2 - Metodologia de classificação de dados baseada na aplicação do critério de Chauvenet

A implementação deste método [18]recorre aos mesmos parâmetros utilizados pelo método anterior (potência, velocidade do vento, coeficiente de potência e rotações por minuto da turbina eólica) e também às mesmas relações entre parâmetros, para realizar a classificação de dados. Tal como no método baseado na aplicação de um desvio padrão, é definida uma restrição que permite classificar todos os valores registados que não a cumpram como *outliers*.

Para obter essa restrição, utiliza-se o critério de Chauvenet [18]. Este é um teste que quando aplicado sobre um conjunto de valores medidos, poderá eventualmente eliminar alguns deles. É de natureza bastante restritiva e define que se a probabilidade de desvio em relação à média de um dado valor medido for inferior a $1/2n$, então esse valor pode ser excluído. Este critério é de extrema utilidade no processo de decisão em diversas áreas, permitindo no caso de estudo em causa a classificação de dados de uma forma simples e eficaz.

Para que seja possível a sua aplicação, é necessário obter a média e o desvio padrão por cada *bin* de velocidade, efectuando de seguida os seguintes cálculos:

$$\Delta R = \frac{d_i}{\sigma}, \quad (1.3)$$

$$x_m = \frac{\sum_{i=1}^m x_i}{n}, \quad (1.4)$$

onde $d_i = x_i - x_m$, n é o número de valores medidos por *bin*, x_i é um valor de índice i do grupo de dados presentes num determinado *bin*, x_m é a média dos valores de um *bin*, σ é o desvio padrão por *bin* e ΔR é a razão do desvio individual para o desvio padrão

Os valores obtidos de ΔR são comparados com valores de referência ΔR_0 tais como os apresentados na tabela 3.1. Se resulta que $\Delta R > \Delta R_0$, onde $\Delta R_0 = \frac{d_{MAX}}{\sigma}$, então o valor x_i correspondente é rejeitado e considerado como *outlier*.

Tabela 3.1 – Tabela de referência do critério de Chauvenet, para rejeição do valor medido [18].

Número de leituras n	Razão entre o máximo desvio aceitável e o desvio padrão, d_{max} / σ
3	1,38
4	1,54
5	1,65
6	1,73
7	1,8
10	1,96
15	2,13
25	2,33
50	2,57
100	2,81
300	3,14
500	3,29

Para aplicação do critério de Chauvenet ao caso particular da análise de dados do funcionamento de um aerogerador, facilmente se percebe que os valores presentes na tabela de referência (3.1) acima apresentada são insuficientes. Por exemplo, ao aplicar esta metodologia para um conjunto de dados colectados durante um período de um mês, cujo período de amostragem de dados do sistema SCADA é de 10 minutos, o mais provável é que número de valores presentes num determinado bin, não seja nenhum dos referenciados na tabela 2.1. Para resolver este problema, efectua-se uma regressão linear sobre os valores tabelados, para que a restrição $\Delta R > \Delta R_0$ apresente sempre valores válidos para ΔR_0 .

3.3 - Metodologia de classificação de dados baseada numa aproximação à curva de Potência de um aerogerador

Esta é uma metodologia de recurso, ou seja, seria a metodologia a ser utilizada se nenhum dos outros métodos aplicados retornasse resultados satisfatórios. A sua aplicação é usada somente para análise e classificação de dados, através do estudo da relação Potência-Velocidade do Vento e actua simultaneamente sobre a totalidade dos dados a serem analisados, não se efectuando neste caso uma análise bin a bin. Neste caso obtém-se a curva que melhor aproxima a característica Potência=f (Velocidade do Vento) a partir da função

$$Y = \frac{P_{MAX}}{1 + e^{-(a \cdot x + b)}} \quad , \quad (1.4)$$

onde P_{max} corresponde à potência nominal do aerogerador em análise, x é a variável velocidade do vento e “a” e “b” os coeficientes reais que melhor permitirão uma aproximação á curva teórica de potência.

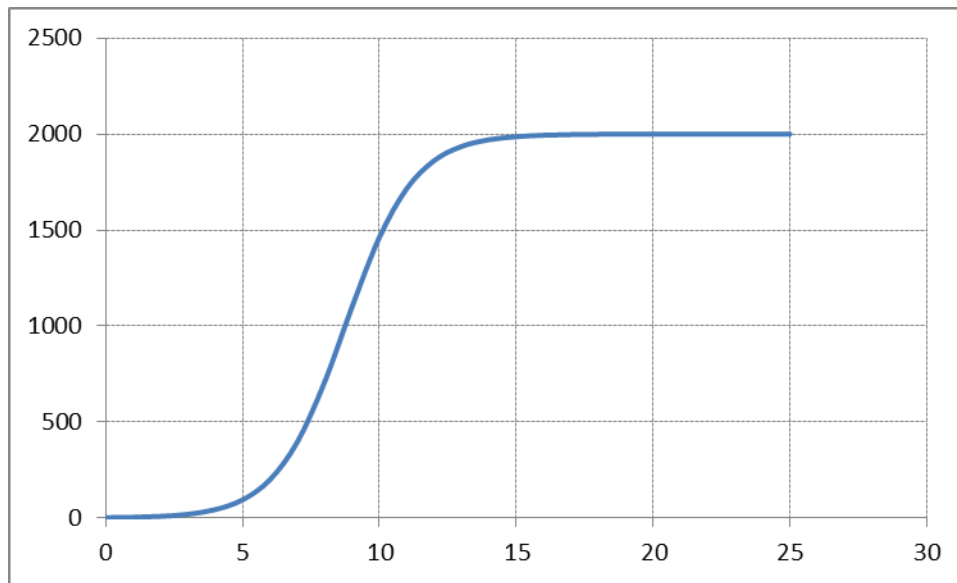


Figura 3.2 - Representação da curva de Potência-Velocidade do Vento.

Para realizar uma aproximação à curva de potência teórica, é possível recorrer à equação 1.4 que, quando lhe atribuirmos parâmetros a e b adequados, retorna para diferentes valores de velocidade, valores de potência aproximados aos que obteríamos para a curva teórica de potência, como se pode verificar pela análise comparativa entre as figuras 3.1 e 3.2.

$$|P_{Calculada} - P_{Medida}| \leq (\beta \times P_{Calculada}), \quad (1.5)$$

onde $P_{Calculado}$ representa um valor estimado pelo modelo AR, P_{Medida} um valor registado na base de dados e β uma percentagem a aplicar sobre a condição, por forma, a que esta retorne resultados coerentes.

Estes valores serão posteriormente comparados com os valores reais medidos segundo a inequação 1.5, cuja verificação determina se um determinado valor de P_{Medida} é rejeitado ou não.

Para que isto seja possível, será necessário aplicar o método dos mínimos quadrados sobre a equação 1.4, procurando o melhor ajustamento para o conjunto de dados em análise. Daí resulta a obtenção de valores ajustados para a e b a definir na equação 1.4, que melhor garantem uma correcta aproximação ao que se considera serem as condições de normal funcionamento de um aerogerador.

3.4 - Metodologia de classificação de dados baseada num modelo Auto-regressivo

Os valores adquiridos pelo SCADA podem ser interpretados como séries temporais. Estas séries são constituídas por um conjunto de valores relativos a uma determinada variável, que se encontram ordenados segundo o instante em que ocorrem, normalmente registados em intervalos de tempo regulares. Um exemplo da estrutura de uma série temporal pode ser encontrado na equação seguinte:

$$\mathbf{x}_{t+1} = f(x_t, x_{t-1}, \dots, x_{t-d}) \quad (1.6)$$

É possível a utilização destas séries para determinação de valores futuros através de valores passados, recorrendo a técnicas de previsão para prever séries temporais.

Uma das técnicas mais utilizadas na previsão de valores futuros em séries temporais, é o modelo ARIMA [19]. Este modelo possui três componentes: auto-regressivo (AR), integrador (I) e média móvel (MA). No modelo auto-regressivo (AR), o valor presente na série x_t está relacionado linearmente com valores passados e com a perturbação aleatória a_t relativamente ao instante t . Este modelo pode ser representado por:

$$\mathbf{x}_t = \phi_0 + \phi_1 \mathbf{x}_{t-1} + \dots + \phi_p x_{t-p} + a_t, \quad (1.7)$$

onde a_t ruído (normalmente considerado branco) e os restantes termos a inércia do modelo. Ou seja, os coeficientes ϕ_i são constantes reais, ao passo que a_t apresenta uma distribuição normal de valores independentes. O valor do termo a_t ao longo do tempo, apresenta uma média nula, uma variância constante e uma co-variância nula entre valores de instantes de tempo diferentes. O que possibilita a aplicação do modelo AR mediante a equação 1.8, onde \hat{X} representa a previsão para a variável X no período t .

$$\hat{x} = \phi_0 + \phi_1 \mathbf{x}_{t-1} + \dots + \phi_p x_{t-p} \quad (1.8)$$

No campo da energia eólica, as séries temporais de velocidade de vento podem apresentar variabilidade em períodos longos de tempo, contudo, o mesmo não se passa para intervalos de tempo reduzidos, como intervalos de dez minutos.

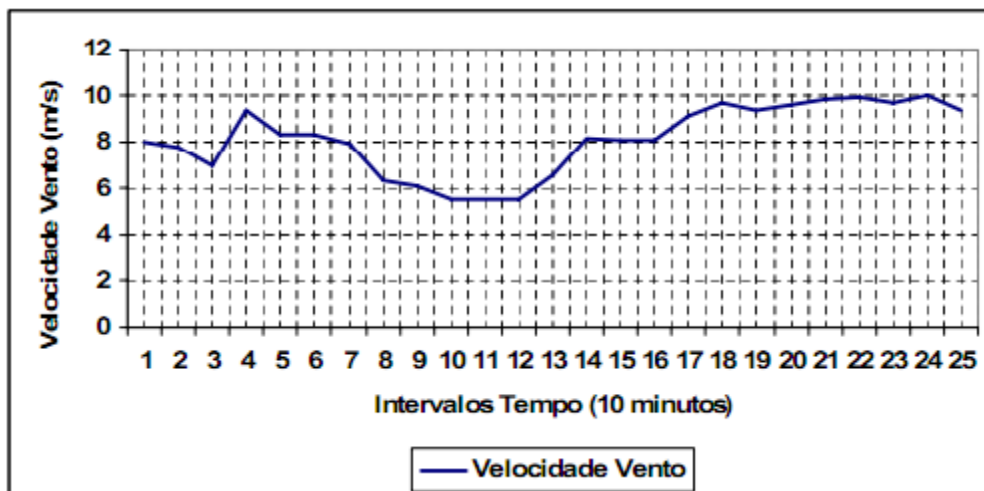


Figura 3.3 - Variação da velocidade do vento num curto intervalo de tempo [19].

Para determinar que tipo de modelo a aplicar e a sua ordem para a realização de previsões nesta área de aplicação, é possível recorrer a uma análise gráfica dos coeficientes de autocorrelação total (ACF) e parcial (PACF), que permitem identificar o modelo da série temporal e avaliar a ordem do modelo a utilizar.

Avaliar a autocorrelação de uma série temporal, não é mais que avaliar a relação entre valores dessa série e outros valores presentes na mesma série com um certo atraso. Avalia-se de seguida, para um período de amostragem de trinta minutos, o gráfico da função de auto correlação para uma série de dados de produção de um parque eólico.

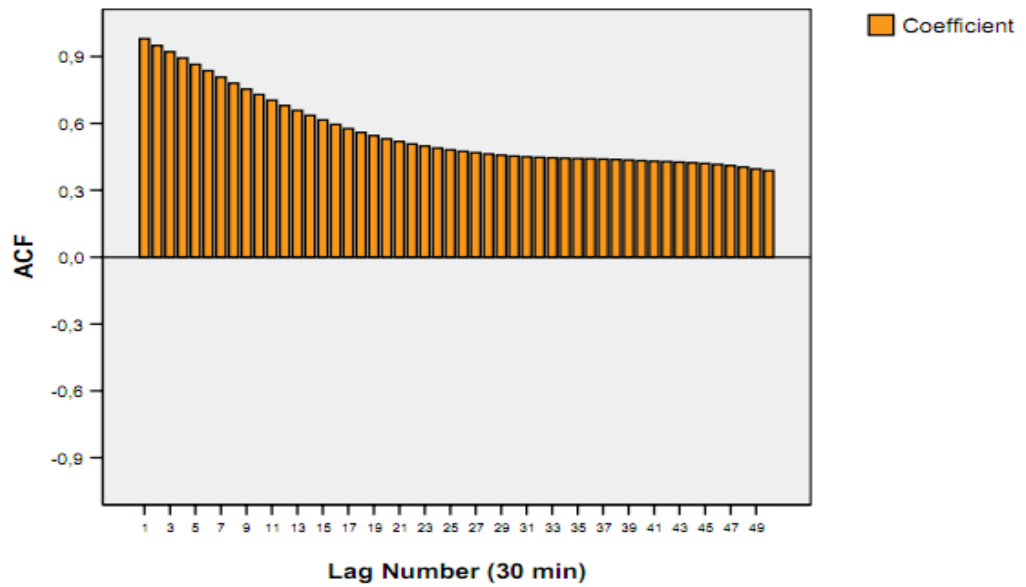


Figura 3.4 - Auto correlação para a potência medida pelo SCADA [19].

A autocorrelação parcial tem em consideração que as autocorrelações para intervalos sucessivos são dependentes, ou seja, a auto correlação parcial de atraso k corresponde à autocorrelação entre X e $X-k$.

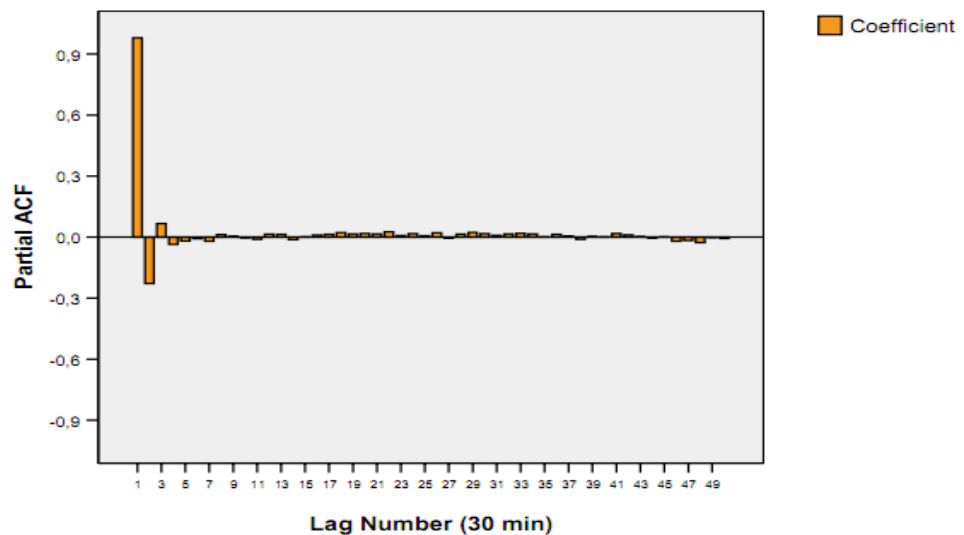


Figura 3.5 - Auto correlação parcial para a potência medida pelo SCADA [19].

A figura 3.2 representa a função de auto correlação parcial para os mesmos dados medidos através de um sistema SCADA, sugerindo um modelo de primeira ordem para o modelo AR independentemente do período de amostragem.

O objectivo neste método é o de aproveitar o modelo AR tal como se encontra representado na equação 1.8, aplicando o teorema dos mínimos quadrados para obter os coeficientes ϕ que melhor permitem a realização da previsão representada. Desta forma, torna-se possível realizar uma estimação de valores futuros de potência, através de valores medidos em intervalos de tempo anteriores, comparando-os posteriormente para os mesmos registos de velocidade, com os valores de potência da curva teórica (figura 3.2) do tipo de aerogerador em análise. Esta comparação possibilita avaliar se os valores armazenados devem ser considerados *outliers* de funcionamento ou valores representativos de um bom funcionamento do elemento electroprodutor em análise. Este filtro de dados é realizado utilizando a condição representada pela inequação 1.9.

$$|P_{\text{Estimado}} - P_{\text{Teórica}}| \leq (\beta \times P_{\text{Teórico}}) , \quad (1.9)$$

onde P_{Estimado} representa um valor estimado pelo modelo AR, $P_{\text{Teórico}}$ um valor da curva teórica do aerogerador em análise e β uma percentagem a aplicar sobre a condição, por forma a que esta realize o filtro de dados adequadamente.

Breve explicação do método mínimos quadrados utilizado neste método [15]

O método dos mínimos quadrados é um método de optimização matemática, que permite encontrar o melhor ajustamento para um conjunto de dados, tentando minimizar a soma dos quadrados das diferenças entre o valor estimado e os dados observados.

Para o caso do nosso método, aplicando o modelo AR de previsão, estamos na presença de um modelo auto-regressivo de primeira ordem, através do qual se pretende estimar valores de potência, minimizando a soma dos quadrados dos resíduos da regressão de forma a maximizar o grau de ajuste do modelo aos dados observados. Isto consegue-se recorrendo à equação seguinte:

$$\hat{x}_t = \phi_0 + \phi_1 x_{t-1} + e_t , \quad (1.10)$$

onde \hat{x}_t representa o valor a ser estimado, x_{t-1} o valor de potência registado no intervalo de tempo imediatamente anterior ao que pretendemos estimar e e_t os resíduos da regressão.

Minimizando o valor residual $\sum_{t=1}^n e_t^2$ é possível encontrar ϕ_0 e ϕ_1 para obter a menor diferença entre a \hat{x}_t e x_{t-1} .

3.5 - Análise Global sobre as Metodologias desenvolvidas

Mesmo sem ainda ter realizado qualquer análise de desempenho sobre as metodologias implementadas, é já possível nesta fase avaliar as suas características, por forma a verificar as semelhanças e diferenças de implementação e ao mesmo tempo realizar um resumo de toda a informação presente neste capítulo, que pode ser observado pela análise da tabela 3.2.

Tabela 3.2 – Resumo das metodologias desenvolvidas.

Metodologias Desenvolvidas	Fundamentos Teóricos	Condição Imposta para detecção de outliers
Sub-Capítulo (3.1)	Desvio Padrão	$ Variável(x) - Variável Média (bin) \leq 2 \times \sigma (bin)$
Sub-Capítulo (3.2)	Critério Chauvenet	$\Delta R > \Delta R_0$
Sub-Capítulo (3.3)	Aproximação á curva Teórica	$ P_{Calculada} - P_{Medida} \leq (\beta \times P_{Calculada})$
Sub-Capítulo (3.4)	Modelo Auto-regressivo	$ P_{Estimado} - P_{Teórica} \leq (\beta \times P_{Teórico})$

Capítulo 4

Sistema de classificação desenvolvido

4.1 - Especificação da Linguagem de Programação Utilizada

Após uma análise entre as diferentes linguagens de programação mais conhecidas de domínio público, já apresentada no capítulo 2, a opção de escolha recaiu sobre a linguagem de programação Python. Esta opção foi tomada com base numa análise sobre algumas das características desta linguagem, relevantes para a implementação bem-sucedida da aplicação a ser desenvolvida no âmbito desta dissertação. De acordo com dados estatísticos, esta linguagem tem vindo a ter uma crescente adesão por parte de grandes referências mundiais, das quais constam nomes como [20]:

- Air Canada: uso de Python no sistema de gestão de reservas;
- Google: utiliza Python no seu algoritmo de busca;
- Intel, Cisco, HP, Seagate e IBM: utilizam Python para testes nos seus equipamentos e dispositivos;
- IRobot: usado na elaboração de robôs comerciais, para inteligência artificial;
- NASA e Fermelab: uso em aplicações de computação científica e simulações;
- NSA: utilizado para análises de criptologia e inteligência artificial;
- Pixar: usado na implementação de animações computacionais;
- Youtube: escrito na sua maior parte em Python.

O python foi criado por Guido Van Rossum [20] em 1991, com o objectivo de disponibilizar uma linguagem que ao mesmo tempo fosse:

- Fácil e intuitiva, ainda que ao mesmo tempo mantivesse um nível de competitividade elevado;
- De código aberto, para que qualquer um possa contribuir para o seu desenvolvimento;
- De código inteligível;
- Adaptável para tarefas diárias, permitindo menor tempo de desenvolvimento.

Ao longo dos anos, grande parte destes objectivos foram sendo conseguidos e a linguagem tornou-se popular, com um acréscimo de adesão considerável no meio da comunidade científica.

As razões que nos podem guiar à utilização do Python podem ser muitas. Se alguns programadores já começam a ver o domínio desta linguagem como um conhecimento rentável para futuro, devido à sua crescente aplicação nas mais diversificadas áreas, outros focam a sua atenção nas suas características e vêem no Python uma linguagem diferencial em relação às linguagens competidoras.

Características diferenciais da linguagem de programação Python [21]:

- O código em Python realiza-se em poucas linhas de código. Mesmo operações que envolvam já um nível elevado de complexidade logarítmica são desenvolvidas em Python recorrendo a poucas linhas de código fonte, o que possibilita o desenvolvimento mais rápido daquilo que se pretende implementar;
- Uso de indentação para marcar blocos;
- É interpretada, ou seja, o código fonte é interpretado por uma máquina virtual, possibilitando a portabilidade do código;
- Permite um rápido processamento.

É de salientar ainda, que além de todas estas características, que influenciam a opção por Python, também a possibilidade de valorização pessoal em futuros trabalhos, por conhecimento adquirido, influenciou em pequena parte, a sua escolha.

4.2 - Especificação do Sistema de Gestão de Base de Dados

Do conjunto de possibilidades estudadas no capítulo 2, para a escolha de um sistema de gestão de base de dados (SGBD) a opção de escolha recaiu sobre o MySQL. Esta é uma opção

baseada nas necessidades intrínsecas ao projecto a desenvolver, procurando encontrar uma solução simples e gratuita que cumpra com as necessidades. O MySQL é usado por diversas companhias mundialmente conhecidas, entre as quais constam nomes como [22]:

- Alcatel;
- Cisco;
- Dataprev;
- Friendster;
- Google;
- HP Nokia;
- Lufthansa;
- NASA;
- Sony;
- U.S. Army;
- U.S. Federal Reserve Bank.

Em relação às suas características gerais, este sistema de gestão de base de dados além de possuir quase todas as características que os seus concorrentes directos possuem, consegue aliar a isso as vantagens de ser gratuito, possuir código fonte aberto, ser leve, rápido, seguro e prático [23].

O elevado desempenho (entendido como um rápido acesso e fácil manipulação de dados), é uma das características que melhor o caracteriza. Outro ponto importante a ser realçado é a segurança. Com os seus diversos tipos de tabelas, que são característica exclusiva do MySQL, é possível ter uma base de dados segura e estável, contando com integridade referencial, *backup* e *restore*, controlo de utilizadores e se necessário verificação e correcção de tabelas.

A sua aplicabilidade é diversificada e de uso frequente para aplicações de internet, contudo, não se remete somente a isso, mas pode ser utilizado em aplicações corporativas, fornecendo suporte a diversas linguagens de programação incluindo Delphi, Java, Python e muitas outras [24].

4.3 - Especificação do sistema desenvolvido para Análise de Desempenho

O objectivo nuclear desta tese compreende a elaboração de um sistema para análise e classificação de dados. Este sistema está repartido em duas partes essenciais, uma base de

dados gerida por um SGBD (MySQL) e uma aplicação implementada em Python que permite além da implementação de métodos de classificação de dados, o acesso a estes e a sua manipulação.

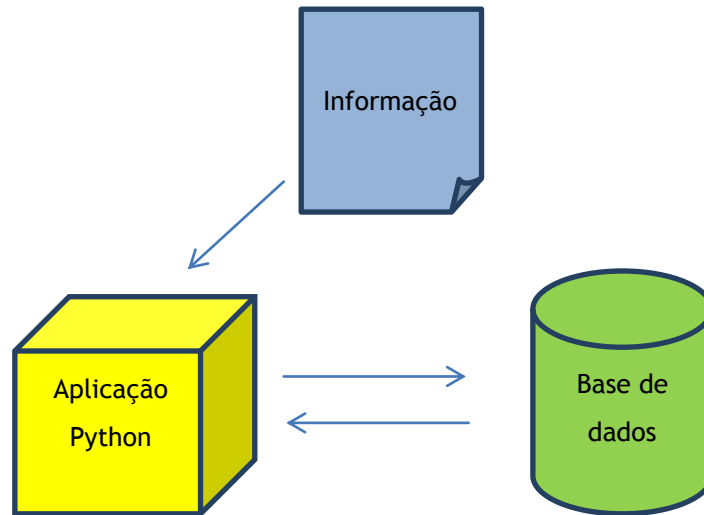


Figura 4.1 - Esquema representativo do funcionamento do sistema de análise e classificação de dados.

Estas duas unidades que compõem o sistema interagem entre si, realizando trocas de informação, com o objectivo de armazenar dados e classificá-los. Com esta estrutura funcional, consegue-se realizar a classificação de novos dados que chegam ao sistema, e ao mesmo tempo, obter um repositório de informação sempre activo, com informação já tratada e constantemente disponível para consulta.

Sistema de Gestão de Base de Dados - MySQL

A base de dados implementada para este sistema, encontra-se dividida em 8 tabelas, que armazenam informação e a organizam em diferentes categorias, para que seja consultada e manipulada. Estas oito tabelas estão relacionadas entre si e mantêm relações de integridade para acesso e manipulação de dados, garantindo desta forma uma segurança sobre a informação nelas contidas. A figura 4.2 apresenta o diagrama de classes da base de dados, que como se pode verificar se centra nas tabelas windfarm e windturbine. De facto, são estas as tabelas que identificam e caracterizam o parque eólico (windfarm) e os seus aerogeradores constituintes (windturbine).

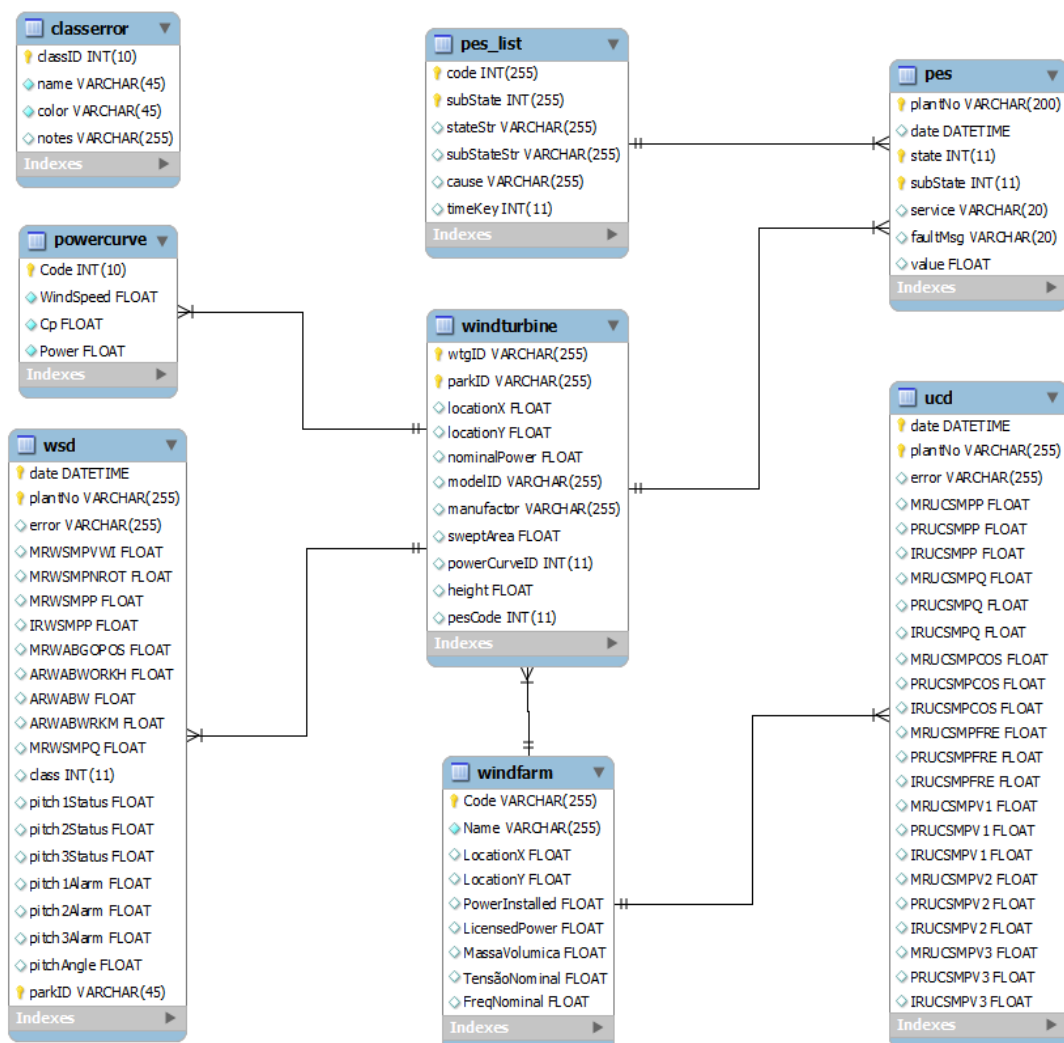


Figura 4.2 - Diagrama de classes da Base de Dados do Sistema

As restantes tabelas para além de armazenar a informação de operação do parque e dos seus aerogeradores (ucd e wsd respectivamente), contêm um conjunto de informação sobre o

estado de operação de cada turbina (*pes_list*). Durante o período de operação o sistema SCADA regista um conjunto de eventos (*pes*) que correspondem a um determinado estado de operação que consta desta lista. Por fim, existem duas tabelas que servem para armazenar informação referente às curvas de potência de cada máquina (*powercurve*) e uma tabela que indica todos os tipos de classificação possíveis de serem utilizados no processo de classificação. De seguida, descreve-se cada uma das tabelas.

Tabela UCD

Esta é uma tabela que contém informações do funcionamento global de um parque eólico, apresentando valores médios, máximos e mínimos de potências e frequências. Num parque eólico a informação adquirida pelos aerogeradores, é posteriormente enviada a uma unidade central de controlo e monitorização, que além de a armazenar, actua sobre esta, gerando nova informação. Esta é uma tabela que se encontra numa relação de um para muitos (1-N) com a tabela *windfarm*.

Tabela WSD

Contém informação referente ao funcionamento de cada aerogerador pertencente a um determinado parque, armazenando dados registados em intervalos de 10 em 10 minutos.

Desta tabela constam parâmetros como potências e velocidades médias, coeficientes de potência e rotações por minuto. Estes parâmetros medidos revelam-se de extrema importância, pois como já anteriormente foi dito no capítulo 3, estes são parâmetros utilizados para implementação das metodologias de análise e classificação de dados.

Tabela PES

A informação contida nesta tabela está associada à ocorrência de eventos que alteram o estado de funcionamento de um aerogerador. Estes eventos são registados pelo SCADA e preenchidos nesta tabela, por forma, a que sejam consultados mais tarde para classificação de dados fazendo recurso aos métodos anteriormente mencionados.

Tabela classerror

Nesta tabela são registados os tipos de classificação que se podem atribuir aos dados fornecidos pelo sistema SCADA, que representem períodos de anormal funcionamento. Estes são detectados alguns pelo próprio sistema SCADA do parque eólico e outros apenas

evidenciados mais tarde, quando os dados adquiridos, são objecto de análise pelos métodos de classificação implementados.

Tabela powercurve

Aqui estão registados dados fornecidos pelos diferentes fabricantes de aerogeradores, que nos permitem retirar conclusões sobre a curva de potência teórica referente ao funcionamento dos mesmos. Os parâmetros de potência e velocidade do vento, são extremamente importantes, pois através destes, é possível traçar a curva teórica de potência e obter uma ideia generalizada do funcionamento de cada aerogerador.

Tabela pes_list

Tabela auxiliar á tabela Pes que contém uma lista de todos os estados de funcionamento do aerogerador que o SCADA consegue identificar. Esta tabela encontra-se numa relação de um para muitos (1-N) com a tabela Pes, permitindo desta forma, que os eventos registados em Pes, sejam classificados.

Tabela windfarm

Nesta tabela, encontram-se dados mais generalistas identificadores de cada parque eólico monitorizado. Parâmetros como nome de parque, código do mesmo, localização, tensão nominal, potência instalada e licenciada, são exemplo de alguma informação que é registada nesta tabela.

Tabela windturbine

Tabela que contém dados generalistas sobre os aerogeradores, de diferentes fabricantes. Contém informações de potências, localizações, modelos, fabricantes e encontra-se relacionada com quase todas as tabelas que compõe o sistema.

4.4 - Especificação da Aplicação - Python

A aplicação desenvolvida foi realizada com o intuito de permitir uma manipulação sobre os dados históricos de operação. Com esta é garantido um acesso à base de dados, para realização de operações de inserção e actualização, que permitam classificar informação e colocá-la disponível, para que sejam tomadas decisões de actuação por parte de quem a analisa.

Da sua constituição fazem parte um interface gráfico no qual o utilizador coloca os dados da operação, um conjunto de algoritmos desenvolvidos por forma a actuar automaticamente sobre o conjunto de dados que o utilizador pretende analisar e ainda um conjunto de saídas gráficas e textuais da classificação efectuada sobre os dados analisados.

A par da aplicação desenvolvida, foi utilizado o recurso a uma outra aplicação designada *PyReport* desenvolvida por Gaël Varoquaux, que executa códigos *python* e gera com os resultados de saída, relatórios *PDF* e arquivos *HTML*. Além disso, permite processar comentários e inclusão do código *python* nos *PDFs* e arquivos *HTML* gerados.

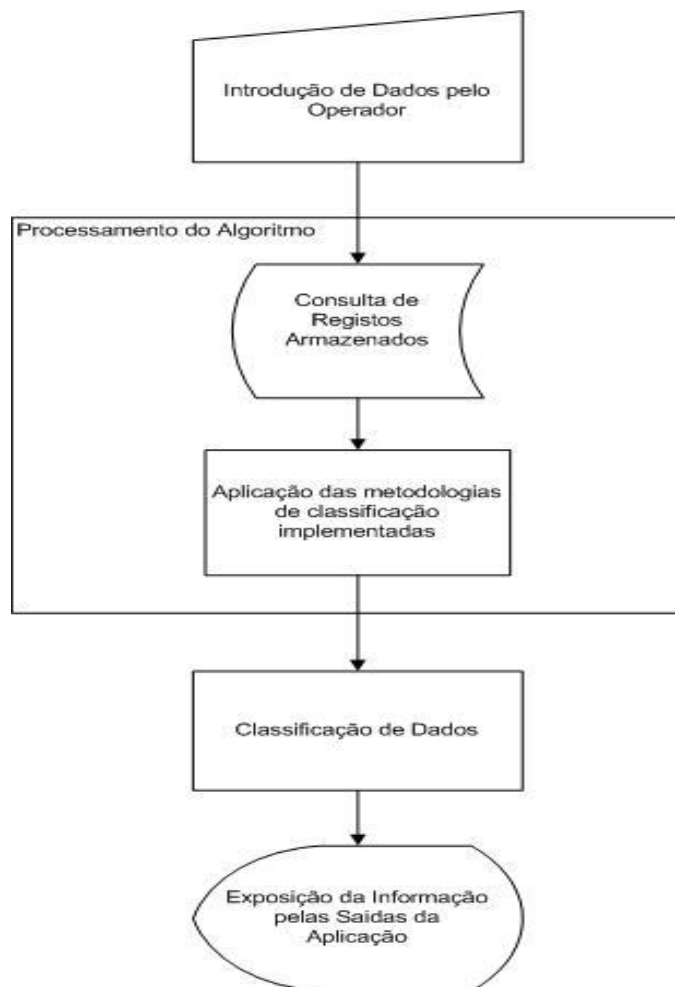


Figura 4.3 - Fluxograma geral do funcionamento da aplicação desenvolvida

Na figura 4.3 encontra-se descrita de forma esquemática, uma visão geral da aplicação desenvolvida, com todos os seus processos, subprocessos, entradas e saídas. As etapas representadas neste diagrama, são sempre executadas em todo e qualquer processamento efectuado através desta aplicação, contudo, processamentos diferentes serão aplicados consoante os valores inseridos nas entradas da aplicação, resultando dessa escolha, a exposição de diferentes tipos de resultados nas saídas.

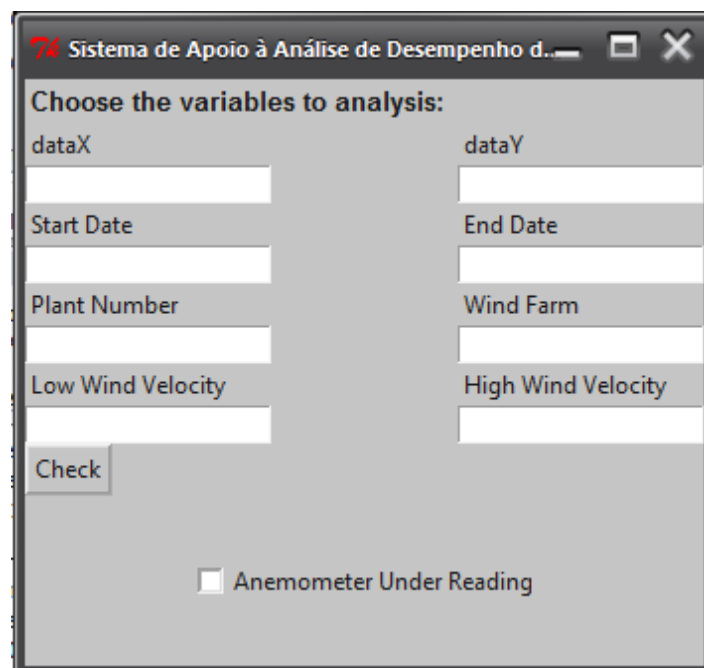


Figura 4.4 - Interface da Aplicação

A interface com o utilizador da aplicação desenvolvida apresenta uma série de campos de preenchimento obrigatório, que permitem ao utilizador definir a análise a realizar sobre os dados armazenados na base dados. Os dados a especificar dizem respeito a aspectos como: tipos de parâmetros em análise, parque eólico e aerogerador em análise, data/hora de início e fim referentes aos períodos em que os dados foram registados pelo *SCADA*, intervalo de velocidade do vento em análise (parâmetro presente em toda e qualquer análise efectuada) e verificação da existência de dados referentes a períodos de anemómetro em leitura, por activação da *checkBox Anemometer Under Reading*.

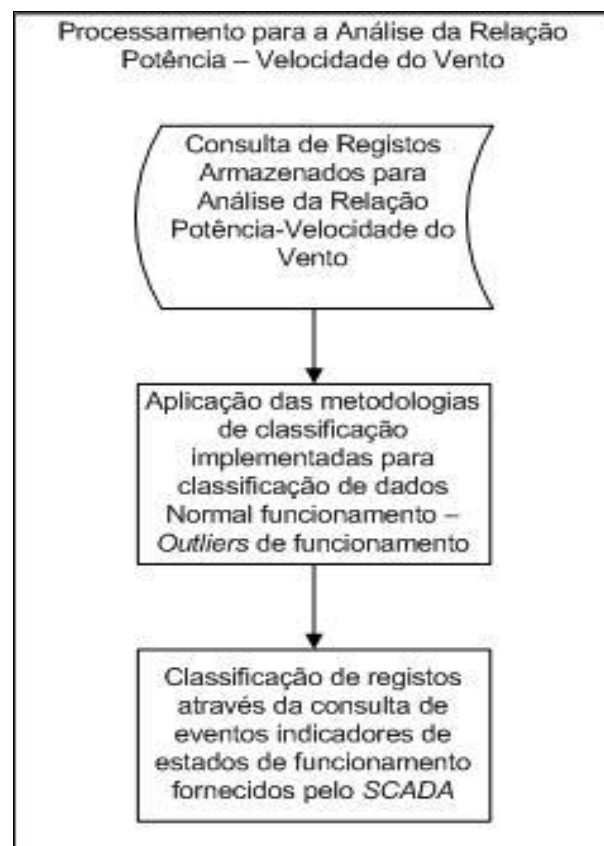


Figura 4.5 - Fluxograma do processamento realizado para análise da relação Potência-Velocidade do Vento

Quando um operador especifica os dados e parâmetros dos quais pretende realizar uma análise, recebe como resultado do processamento efectuado um conjunto de saídas gráficas e textuais. Estas permitem-lhe retirar ilações sobre o funcionamento de um determinado aerogerador durante o intervalo de tempo e condições de funcionamento que este definiu para a análise.

Uma das possíveis saídas gráficas da aplicação é uma análise da relação potência e velocidade média registadas em intervalos de dez em dez minutos num aerogerador, durante um determinado período de tempo. Para a realização desta análise, uma série de processamentos lógicos têm de ser realizados de maneira a operar uma classificação sobre os dados inicialmente indistinguíveis. Uma representação de todo este processo é representada na figura 4.5.

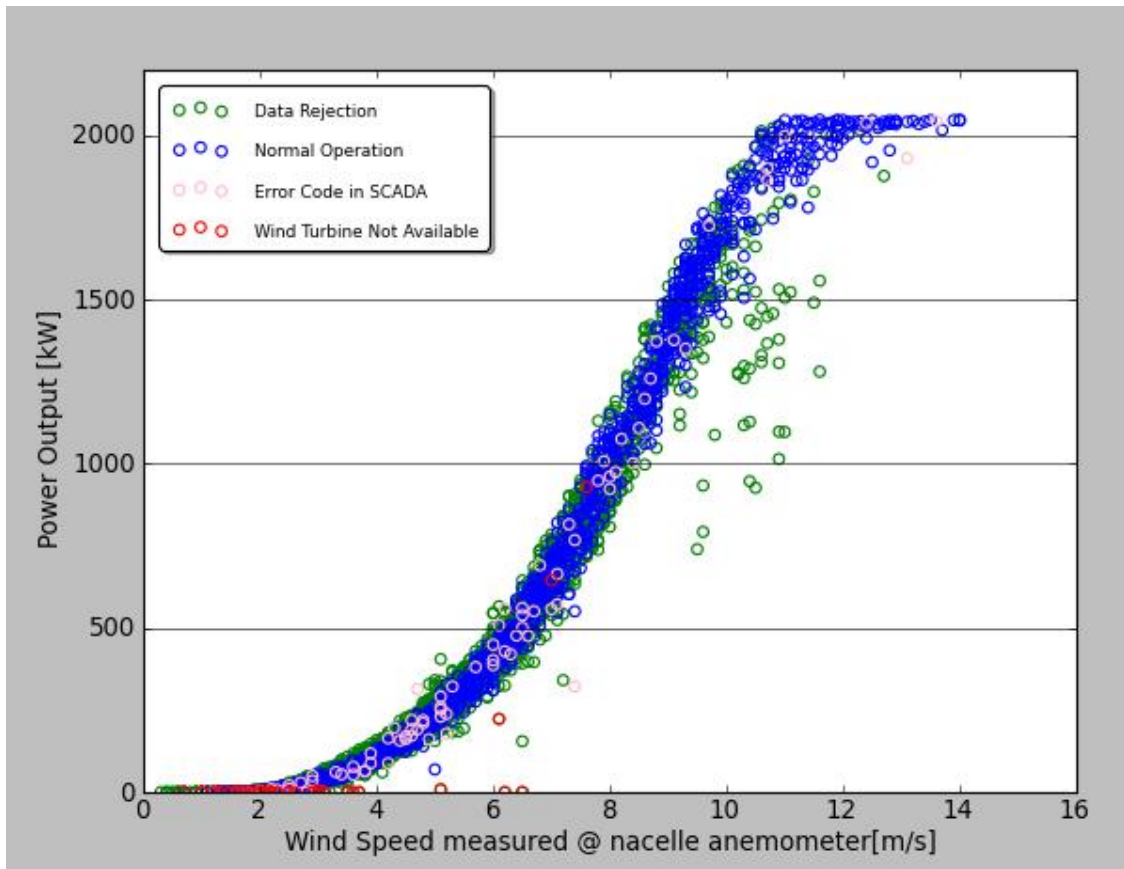


Figura 4.6 - Análise da relação Potência-Velocidade do vento

A primeira classificação a ser realizada utiliza os métodos anteriormente descritos no capítulo 3, distinguindo dados representativos de um normal funcionamento de um aerogerador e dados considerados *outliers* a essa mesma condição de funcionamento. Por observação da figura 4.4 pode-se observar essa classificação, com a representação a azul dos dados de operação a normal funcionamento e a verde os dados rejeitados pelos métodos. Posteriormente a esta primeira classificação, outras referentes a períodos de erro de leitura de SCADA, indisponibilidade de turbina eólica e anemómetro sobre leitura, são realizadas analisando outros dados de funcionamento registados. Situações de ocorrência de erro de leitura de SCADA (designados erro de código SCADA) são registadas pelo próprio e representadas pela aplicação a rosa, como representado na figura 4.4. Períodos de indisponibilidade da turbina são também identificados avaliando os eventos indicadores dessa situação e calculando posteriormente os períodos em que a turbina esteve indisponível até voltar a um normal funcionamento. Estes períodos são representados a vermelho pela saída da aplicação, tal como se observa na figura 4.6.

Pela análise da relação de Potência-Velocidade do vento, é possível ainda identificar outras anomalias de funcionamento. Limitações de potência aplicadas aos aerogeradores,

situações de má calibração de anemómetros e problemas de controlo do *pitch* são normalmente passíveis de serem identificados por uma análise cuidada desta saída de aplicação.

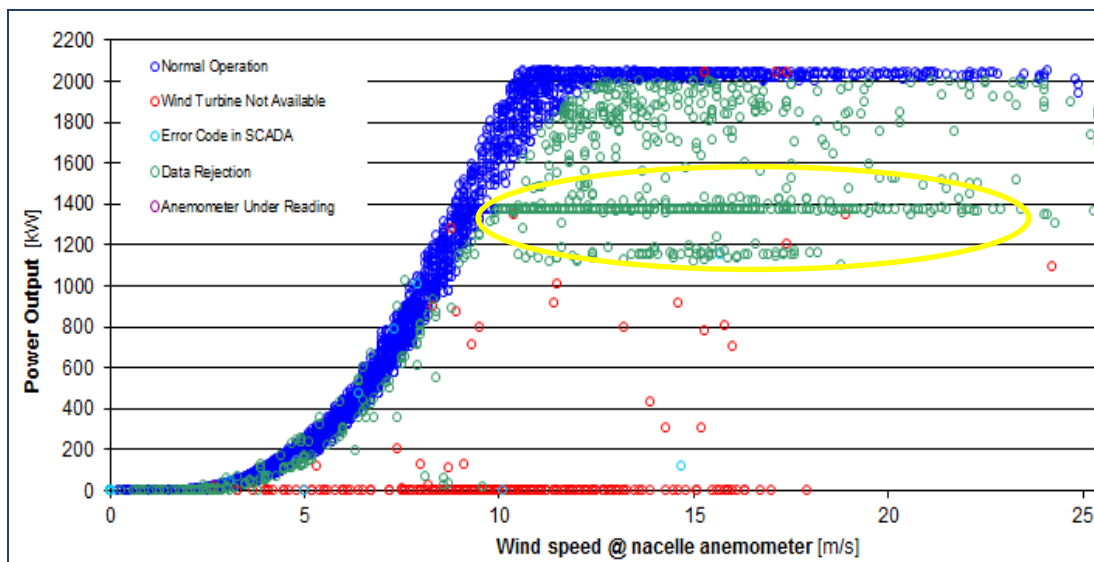


Figura 4.7 - Identificação de limites de potência pela análise da relação Potência-Velocidade do vento.

Os limites de potência aplicados a aerogeradores, podem dever-se a diversas razões, sendo que, normalmente o seu objectivo é obter um controlo sobre o funcionamento do aerogerador. Do conjunto de razões, normalmente associadas a este procedimento, fazem parte:

- Limitações impostas pelo próprio gestor do parque;
- Limitações impostas pela rede, relativamente à potência que lhe é entregue;
- Limitações impostas após períodos de inactividade, de forma a controlar o arranque do gerador;
- Limitações impostas devido a problemas identificados no equipamento dos aerogeradores, que impossibilitem um funcionamento destes a condições nominais.

Na figura 4.7, está representado um limite de potência aplicado a um aerogerador, identificado pela zona delimitada a amarelo, em que se verifica facilmente que os valores de potência se encontram limitados para alguns dos registos efectuados a velocidades de vento, que em normal funcionamento permitiriam obter outros valores de potência.

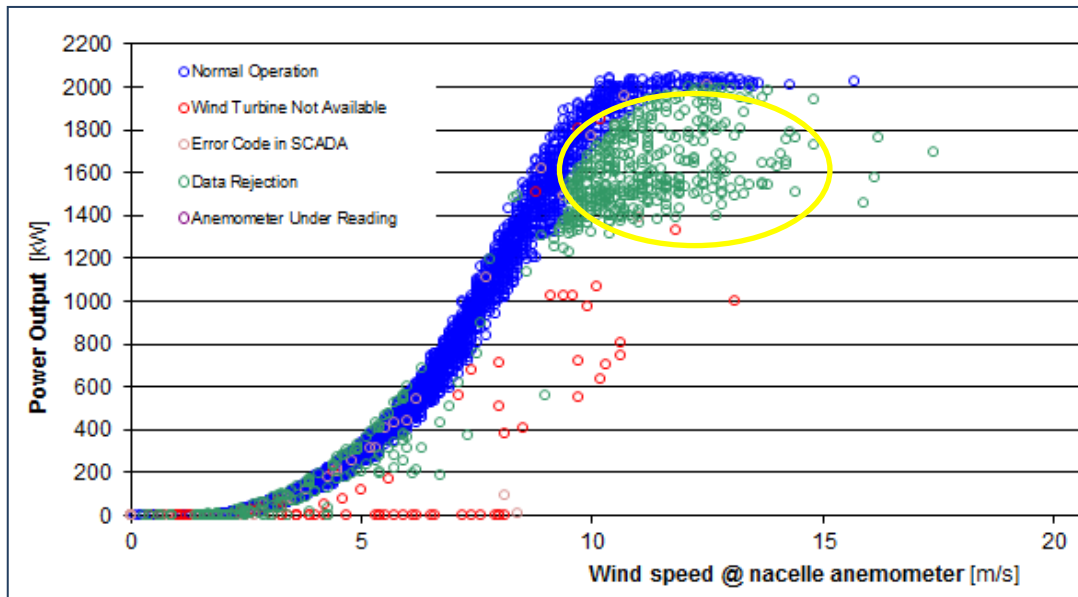


Figura 4.8 - Identificação de problemas de controlo do pitch pela análise da relação Potência-Velocidade do vento.

O controlo do *pitch*, é o controlo efectuado por forma a garantir que o ângulo das pás de um aerogerador se ajuste de forma adequada a proporcionar um bom arranque do mesmo e que posteriormente se adapte de acordo com a velocidade do vento que vai sendo registada durante o seu funcionamento. Na figura 4.8 é possível identificar uma anomalia de controlo, na área delimitada a amarelo. Esta é uma zona de funcionamento em que o ângulo das pás tem de ser definido de forma, a que se proporcione uma estabilização mecânica do funcionamento do aerogerador, que se traduzirá numa perda aerodinâmica por parte do mesmo.

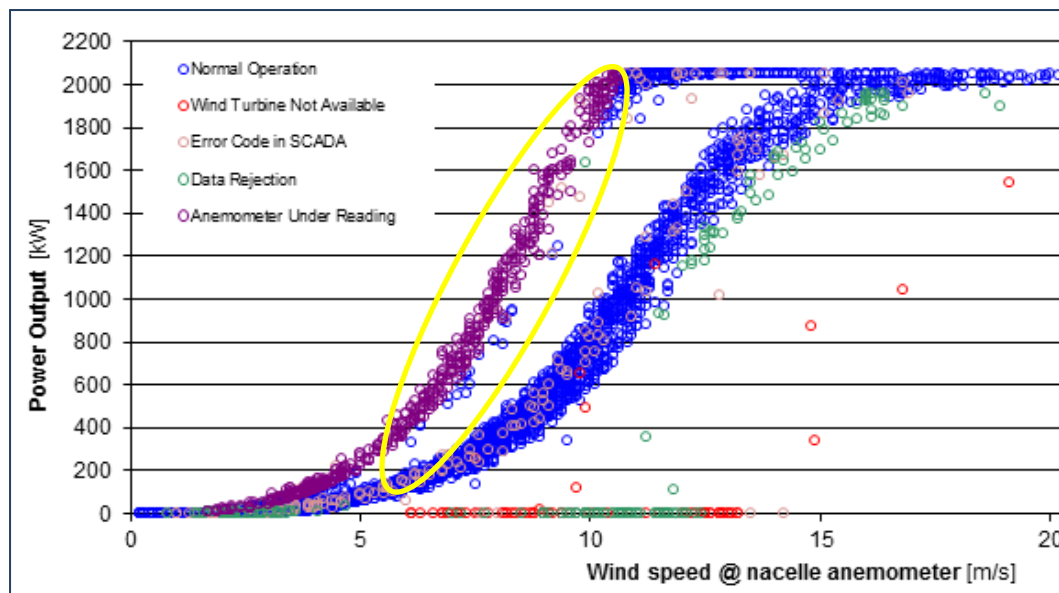


Figura 4.9 - Identificação de problemas de leitura do anemómetro pela análise da relação Potência-Velocidade do vento

Problemas relacionados com o mau funcionamento de equipamentos de medição instalados nos aerogeradores, como anemómetros, podem também ser identificados, recorrendo a esta saída de aplicação. Estes equipamentos, estão sujeitos a intervenções de ajuste de funcionamento consoante o local em que são instalados e ainda por muitas vezes a condições climáticas adversas, que no seu conjunto podem levar a que os registos de funcionamento por estes efectuados, não sejam os que melhor retratem o real funcionamento do elemento electroprodutor aonde foram instalados. A figura 4.9 representa na zona delimitada a amarelo, um exemplo de um mau ajustamento da função de transferência, que rege o funcionamento dos anemómetros instalados num aerogerador.

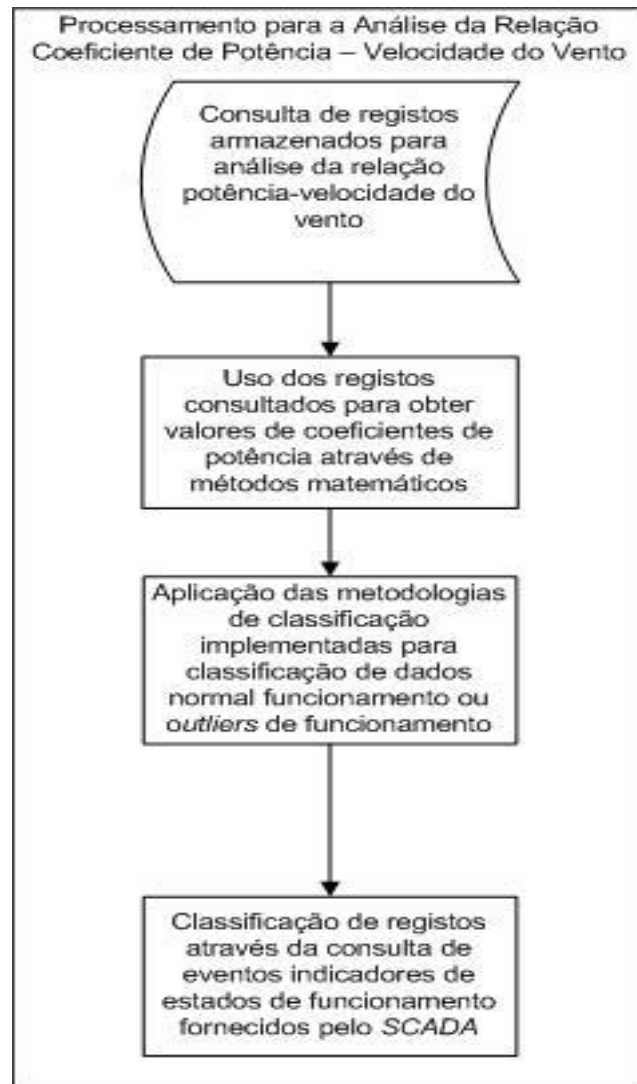


Figura 4.10 - Fluxograma do processamento realizado para análise da relação Coeficiente de Potência - Velocidade do Vento

Um operador, quando elege os parâmetros em análise para um determinado período de tempo, poderá optar por uma análise do coeficiente de potência em função da velocidade do vento. Esta análise envolve um conjunto de processamentos, que se encontram representados na figura 4.10.

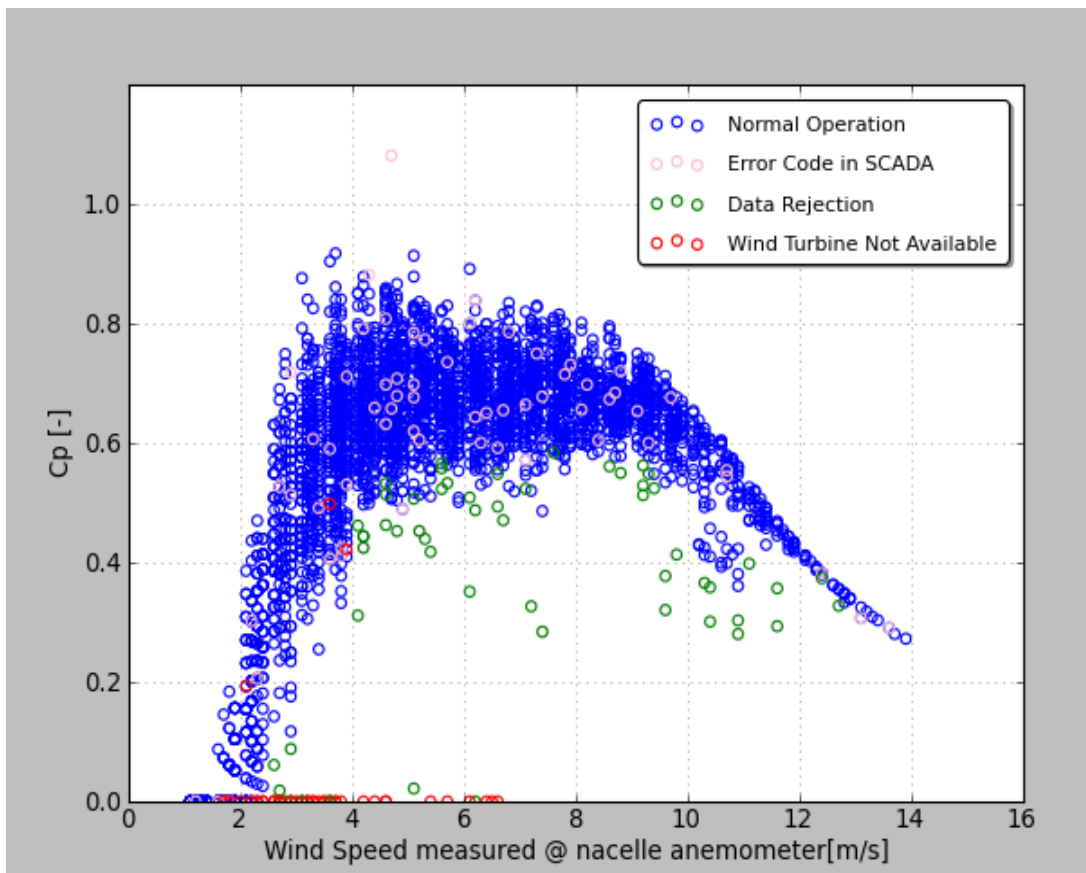


Figura 4.11 - Análise da relação Coeficiente de Potência-Velocidade do vento, sem aplicação do limite de Betz

O uso desta relação revela-se útil não só para identificar anomalias de funcionamento que já se poderiam identificar pela análise da saída da relação Potência-Velocidade do vento (limites de potência e problemas de controlo de *pitch*), assim como realizar uma análise sobre os registos armazenados, verificando se se encontram em conformidade com o limite teórico de coeficiente de potência (C_p) imposto pelo limite de Betz.

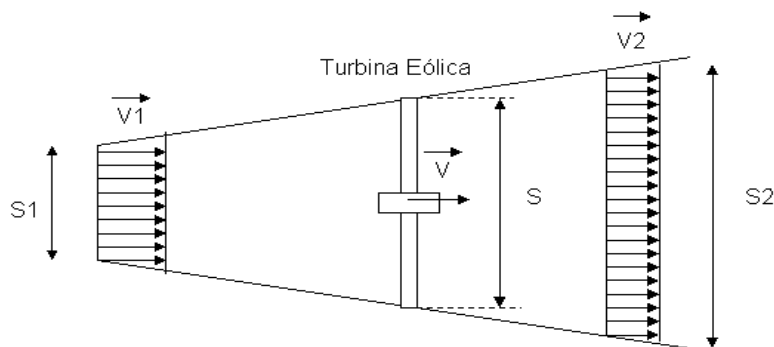


Figura 4.12 - Ilustração da Teoria de Betz [25].

onde V_1 é o vento antes da turbina, V é o vento na turbina, V_2 é o vento depois da turbina. Verifica-se $V_1 > V > V_2$, sendo estas velocidades paralelas ao eixo do rotor.

O limite de Betz [25] representado na figura 4.12 diz-nos que mesmo para os melhores sistemas electroprodutores eólicos, apenas um máximo de 59% da energia do vento é aproveitada, o que significa que o C_p máximo (teórico) é 0,59. Para uma aplicação real, este coeficiente é ainda mais baixo, da ordem de 0,3 a 0,4 no máximo. A teoria de Betz modela a passagem do ar antes e após a turbina, por um tubo de corrente onde se verifica se existem valores superiores aos impostos pelo limite.

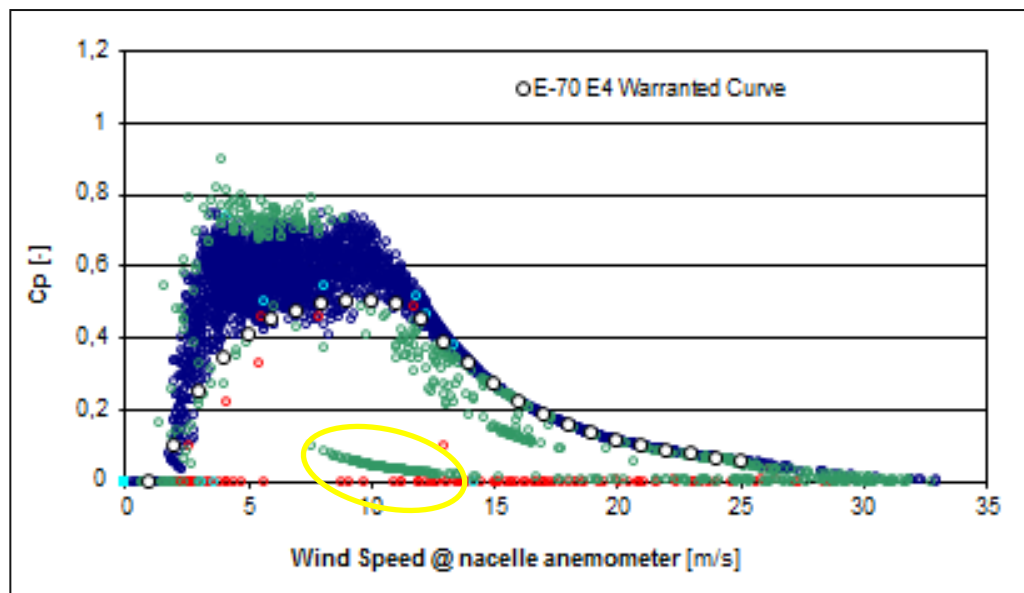


Figura 4.13 - Identificação de limites de potência pela análise da relação Coeficiente de Potência-Velocidade do vento.

Pela análise da figura 4.13 podemos identificar na área delimitada a amarelo um limite de potência aplicado sobre um aerogerador, que tal como na relação Potência-Velocidade do vento, é perfeitamente perceptível. Neste caso, uma simples análise sobre o nível de rendimento de uma turbina eólica, realizada através dos valores de C_p registados, permite identificar períodos de baixo rendimento para valores de velocidade do vento registados que implicariam níveis de rendimento superiores.

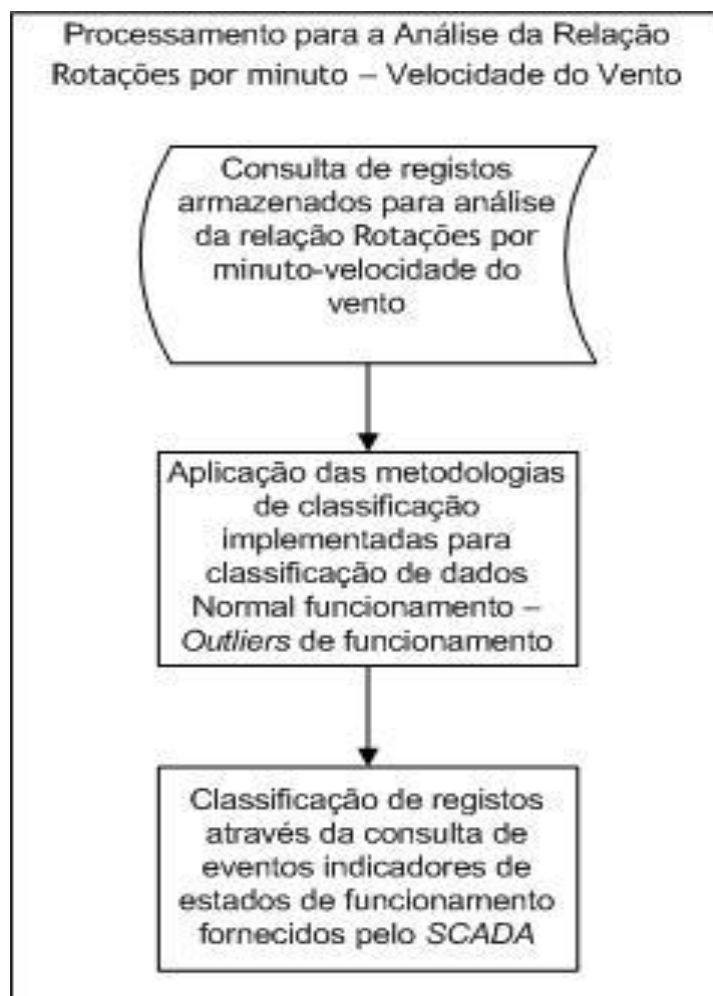


Figura 4.14 - Fluxograma do processamento realizado para análise da relação Rotações por Minuto [RPM] - Velocidade do Vento.

Também uma análise das rotações por minuto (RPM) registadas pelo SCADA, em função da velocidade do vento é possível ser realizada. Esta possui como principal vantagem relativamente às outras análises efectuadas o facto de permitir avaliar aspectos mecânicos do funcionamento dos aerogeradores e daí retirar conclusões não só dos rendimentos energéticos e intervenções efectuadas sobre os mesmos, mas também do estado dos equipamentos que os constituem. Uma representação do processamento efectuado para esta análise, encontra-se representada na figura 4.14

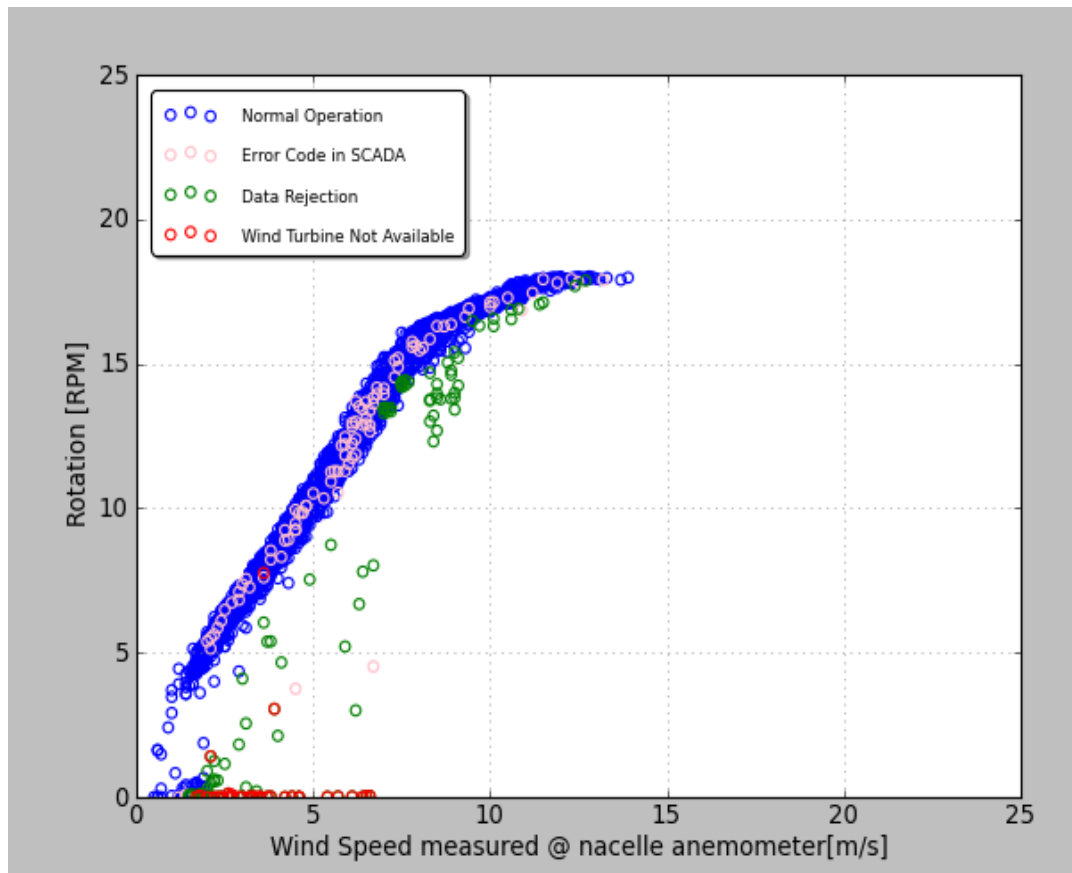


Figura 4.15 - Rotações por minuto - Velocidade do vento.

Desta forma, é possível retirar ilações sobre a existência ou não de limites de rotação impostos pela empresa que realiza a gestão do parque eólico, de maneira a controlar o desgaste provocado pelo vento no sistema de pás dos seus aerogeradores. A identificação, destes limites permite justificar períodos de tempo superiores aos que normalmente seriam necessários a um aerogerador para trabalhar a valores de potência nominais.

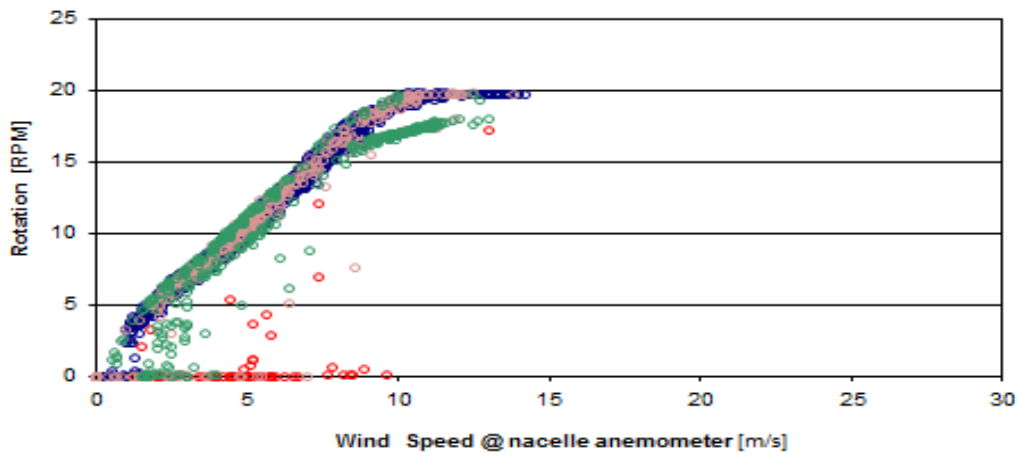


Figura 4.16 - Identificação de limites de rotação pela análise da relação Rotações por minuto - Velocidade do vento.

Por esta análise é ainda possível detectar limites de potência, contudo, dado a sua semelhança com os limites de rotação impostos, esta não será a melhor maneira de os analisar. Na figura 4.16 pode-se observar na zona delimitada a amarelo um exemplo, de como esta saída pode ser útil na identificação de limites de rotação impostos a um aerogerador.

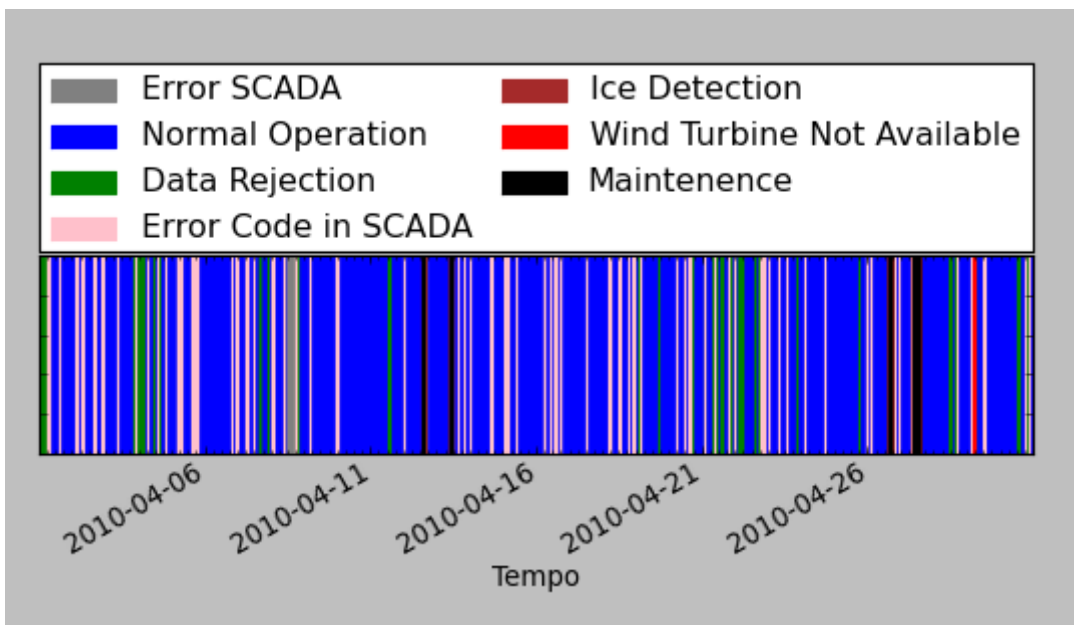


Figura 4.17 - Análise temporal do funcionamento de um aerogerador com recurso a código de cores.

Uma das saídas geradas pela aplicação, para análise do funcionamento de um aerogerador em diferentes períodos de tempo, é efectuada recorrendo a uma representação por cores, em que cada uma representa um estado de funcionamento diferente. Uma representação dessa mesma saída pode ser encontrada na figura 4.17, que para além dos estados de funcionamento já anteriormente mencionados na descrição da relação Potência-Velocidade do vento, mais três são representados recorrendo a informação disponibilizada pelo SCADA. Assim sendo, períodos de manutenção do aerogerador, detecção de gelo nos equipamentos de sensorização e períodos em que simplesmente não se verifica nenhum dos estados conhecidos de operação (erro de SCADA), são representados a preto, castanho-alaranjado e cinzento respectivamente.

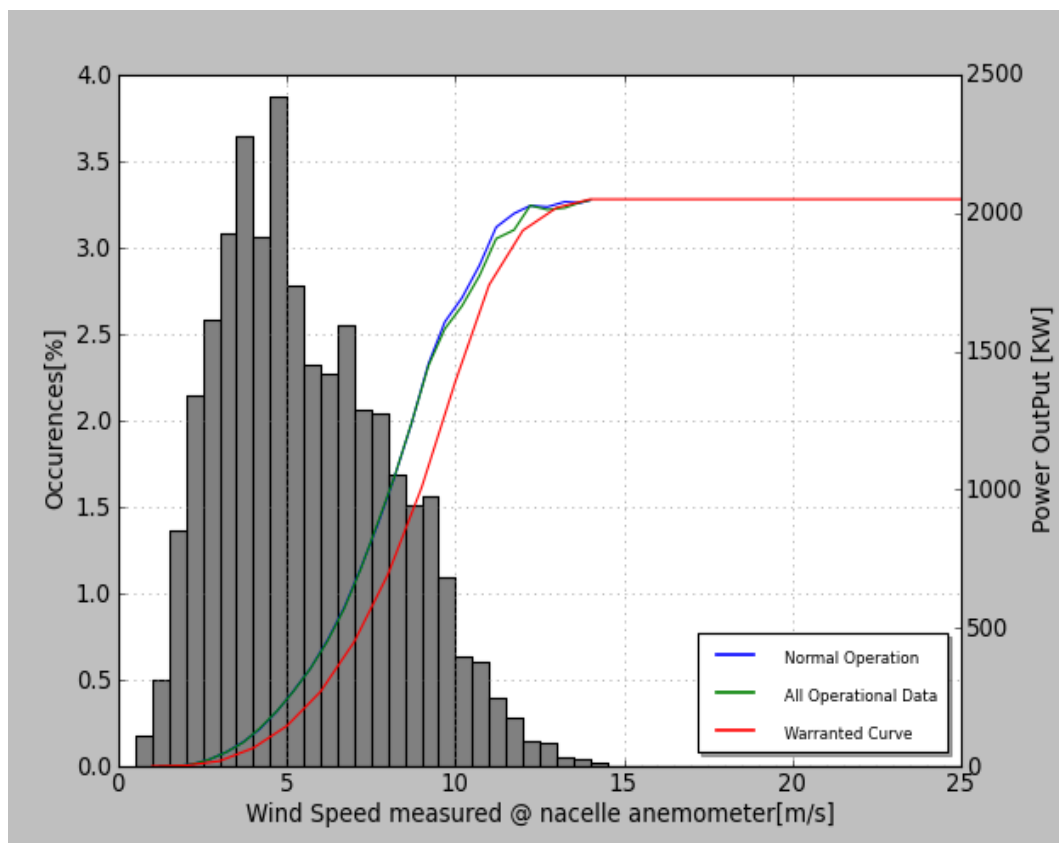


Figura 4.18 - Análise comparativa de registos efectuados pelo SCADA por bin de velocidade de vento.

Por forma a complementar a análise da relação Potência-Velocidade do vento, uma outra é efectuada para verificar o regime de ventos durante um determinado período. Esta análise permite perceber irregularidades de funcionamento, como indisponibilidade do aerogerador,

situações de baixo rendimento, obter uma ideia generalizada da velocidade do vento da região e ainda, verificar a consistência de dados registados por bin.

Para além da análise acima referida, também uma média por bin de potência para os dados em condição de normal funcionamento e para a totalidade dos dados analisados é realizada, permitindo verificar qual a diferença de comportamento em rendimento energético para os diferentes bins analisados, com ou sem a presença de outliers de funcionamento.

Por fim, uma representação da curva teórica do aerogerador em análise dada pelo fabricante é também representada a vermelho na figura 4.18, para que se possa realizar uma análise comparativa entre valores de potência medidos (reais) e o comportamento ideal de funcionamento de um aerogerador.

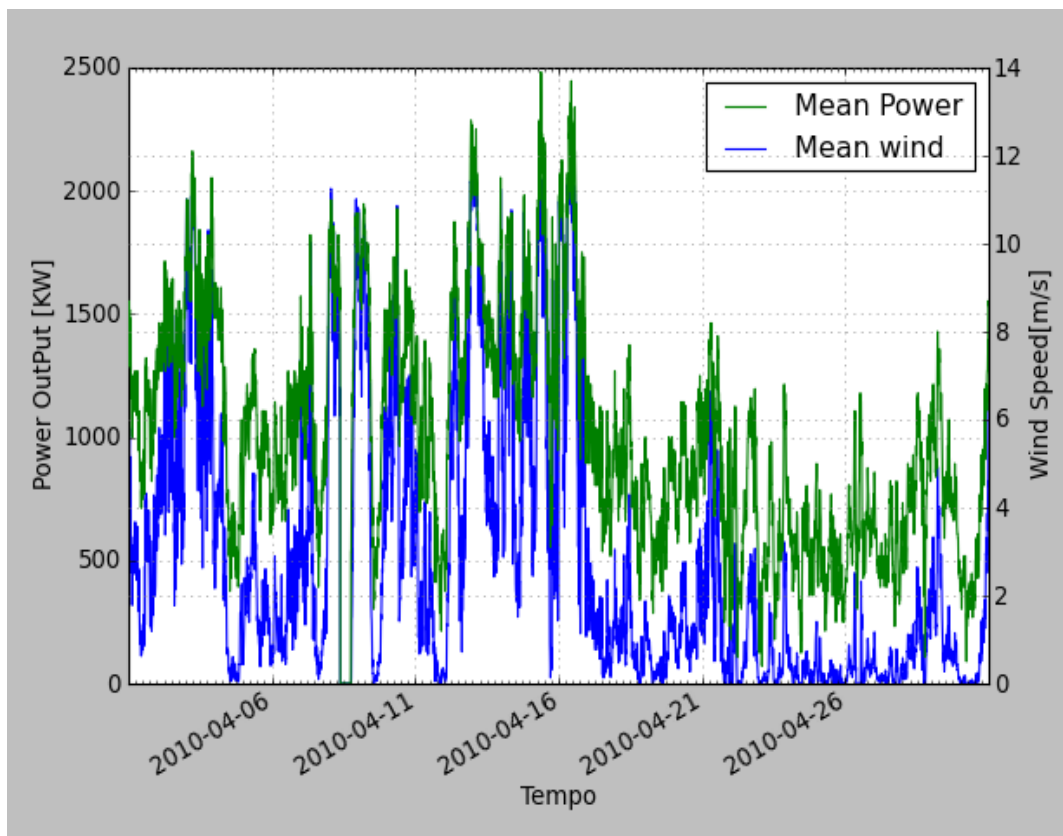


Figura 4.19 - Análise da potência e velocidade de vento em função do tempo.

Outra das análises realizada pela aplicação desenvolvida coloca no mesmo gráfico os valores registados de potência e de velocidade do vento para um período de tempo seleccionado pelo operador. Esta análise é de grande utilidade pra detecção de outliers de funcionamento, verificando-se discrepâncias existentes entre valores de potência e valores de velocidade do vento por uma simples análise gráfica, como a representada na figura 4.19.

Períodos de falha de *SCADA*, podem também ser identificados por observação directa deste tipo de gráficos, correspondendo a situações de total ausência de registos.

Também saídas de cunho não gráfico, são apresentadas como resultado do processamento desta aplicação. Dados como:

- Erro PECC, valor correspondente à percentagem de dados não representativos de um normal funcionamento de um aerogerador;
- Energia esperada e medida, em MWh para um determinado período em análise;
- Desempenho energético verificado, como resultado da divisão da energia medida pela energia esperada;
- Erro absoluto médio e quadrático em relação á curva teórica do aerogerador em análise;
- Tempo correspondente ao período em análise elegido pelo operador;
- Média da velocidade do vento para o período em análise;
- Média das potências para os dados em condição de normal funcionamento e para a totalidade dos dados em análise de acordo com as especificações do operador;
- Percentagem sobre a totalidade dos dados em análise dos períodos de operação normal, dos rejeitados pelo método, de turbina indisponível, de manutenção, de detecção de gelo nos sensores de medição e de erro no *SCADA*.

Capítulo 5

Análise de Resultados

Para testar os diferentes métodos desenvolvidos de classificação de dados realizou-se uma análise sobre o desempenho individual de cada um perante o mesmo conjunto de dados referentes a um determinado aerogerador e para o mesmo período de tempo de funcionamento.

Os testes empregues para a realização desta análise, foram desenvolvidos em conjunto com o próprio algoritmo da aplicação, por forma, a servirem de ferramenta de ajuda para decisões operativas e para a detecção de irregularidades de funcionamento, através da análise da informação facultada. A Análise do erro PECC (razão entre os registos representativos de normal funcionamento e a totalidade de registos em análise), comparação entre a energia esperada e medida, desempenho energético e erro médio absoluto e quadrático em relação à curva teórica do aerogerador em análise servem de indicadores que permitem realizar uma comparação entre estes.

Nesta análise foram escolhidos dois aerogeradores, analisando os seus registos de funcionamento em períodos de tempo diferentes e previamente determinados. De seguida são apresentadas o conjunto de saídas obtidas para os diferentes métodos, numeradas pela seguinte ordem de maneira a facilitar a sua identificação gráfica:

1. Método baseado no desvio padrão;
2. Método baseado na aplicação do critério de Chauvenet;
3. Método baseado num modelo auto regressivo;
4. Método baseado numa aproximação à curva teórica de um aerogerador.

5.1 - Resultado da Análise efectuada sobre o Aerogerador λ para um intervalo de tempo T1

Esta primeira análise foi realizada para um aerogerador a funcionar durante um período de tempo T1, em que não se verificaram grandes desvios relativamente aquilo que se considera ser um normal funcionamento, ou seja, não se verificaram períodos significativos de funcionamento anómalo, que poderiam comprometer o rendimento energético do aerogerador em análise.

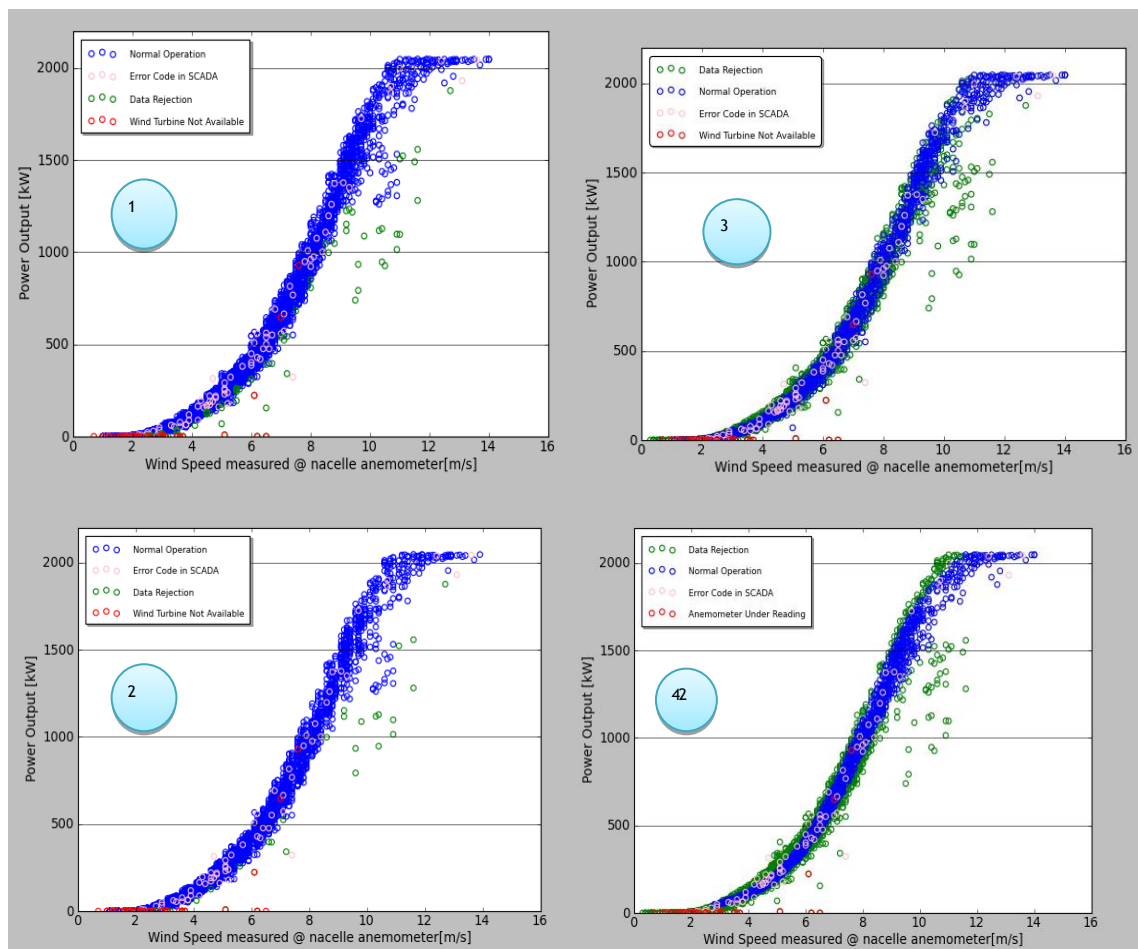


Figura 5.1 - Análise gráfica das relações de Potência-Velocidade do vento para os métodos implementados, no intervalo de tempo T1.

Na figura 5.1 estão representadas as quatro representações gráficas de Potência-Velocidade do vento, obtidas para os quatro diferentes métodos de classificação implementados.

Nas saídas identificadas com os números um e quatro estão representados os resultados gráficos obtidos pelos dois métodos estatísticos elaborados. Destas duas saídas obtêm-se resultados de classificação bastantes semelhantes, sendo que os registos *outliers* de funcionamento mais evidentes são identificados na sua maioria, à excepção de alguns registos compreendidos entre os 10 e os 12 m/s. Esta falha de classificação poderá dever-se ao facto de estes dois métodos se basearem em médias e desvios para identificação de *outliers* e de neste intervalo de velocidade o grau de incerteza para este tipo de classificação ser elevado, devido à necessidade de um controlo *pitch* que influencia as amostras recolhidas.

Os gráficos representados na figura 5.1 pelos números 3 e 4, apresentam uma melhoria em relação às saídas 1 e 2, resolvendo o problema de identificação de *outliers* anteriormente verificado. Os dois métodos usados na sua elaboração utilizam como base comparativa de classificação a curva de potência teórica do aerogerador, relacionando-a com os valores medidos a ser analisados. Isto permite a obtenção de bons resultados, embora os critérios de classificação por estes utilizados sejam bastante distintos, como já foi referido no capítulo 4.

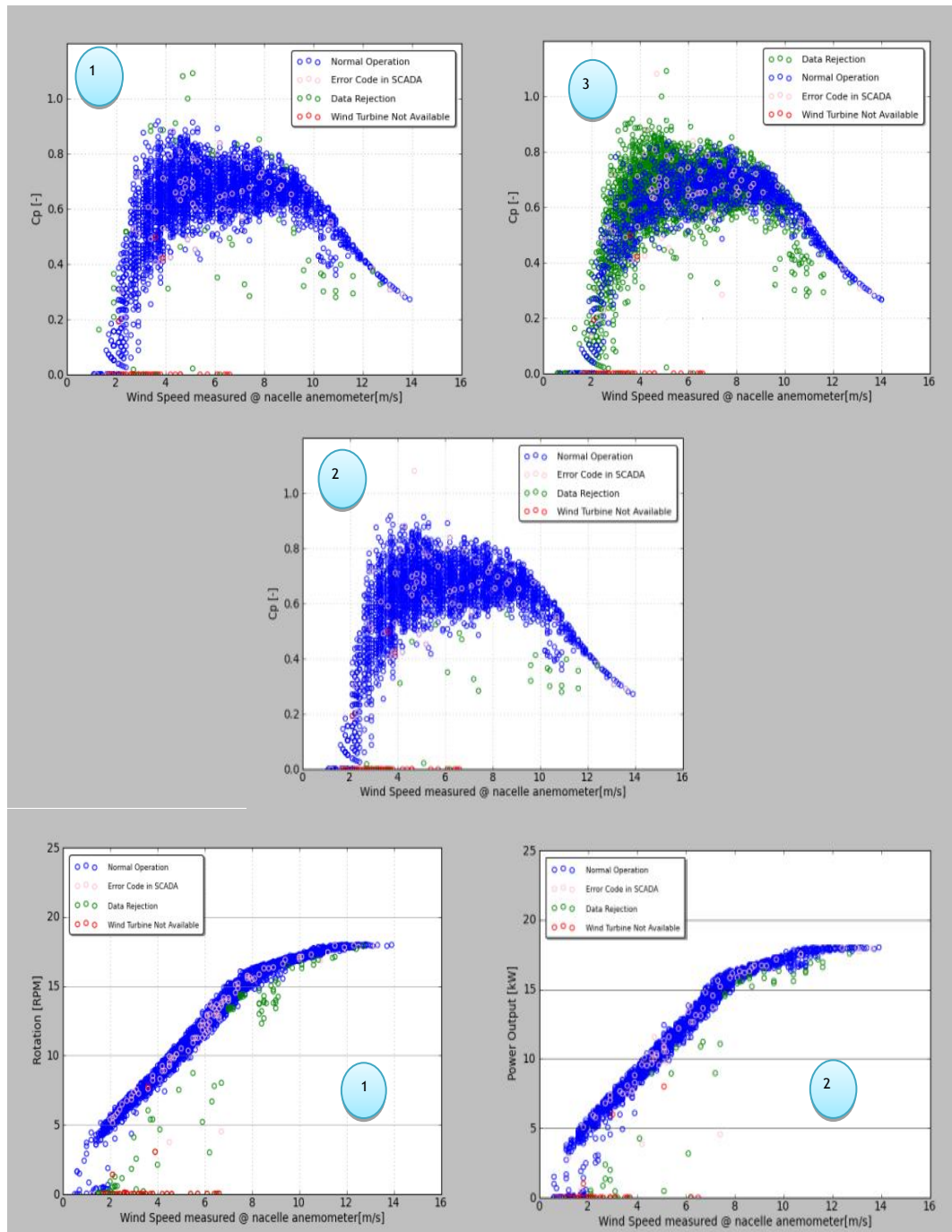


Figura 5.2 - Saídas gráficas da aplicação, para aerogerador λ no intervalo de tempo T1

As análises gráficas para as saídas das relações de Coeficiente de Potência-Velocidade do Vento e Rotações por Minuto-Velocidade do Vento são também saídas da aplicação desenvolvida e são representadas tal como se pode observar pela figura 5.2. Contudo, para os casos de teste usados, estes apenas permitem identificar, aspectos de funcionamento já identificados pela análise das saídas da figura 5.1, não deixando por isso de ser recursos de análise válidos. As relações de coeficiente de potência - velocidade do vento, são

implementadas pela aplicação desenvolvida, a partir, das relações de potência-velocidade do vento, sendo utilizados estes últimos dois parâmetros para obter os valores de coeficiente de potência por processos de cálculo. Desta relação entre saídas, resulta a possibilidade de identificação de *outliers* comuns às duas saídas gráficas. A grande vantagem referente à relação do coeficiente de potência-velocidade do vento representada pela figura 5.2, consiste em verificar se os dados de funcionamento do aerogerador em análise são indicadores do seu funcionamento dentro do limite teórico estabelecido pelo limite de Betz.

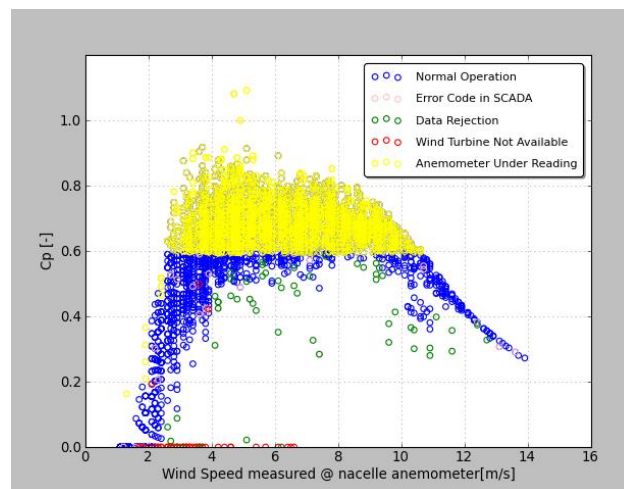


Figura 5.3 - Coeficiente de Potência-Velocidade do vento, com a aplicação do limite de Betz para o método 1.

No caso em análise, essa verificação não é realizada, porque a sua aplicação, reduz substancialmente os registos em análise, que influenciam o resultados de outras saídas da aplicação. Contudo, para que seja possível demonstrar a sua aplicabilidade, dado que esta análise é uma opção disponibilizada pela aplicação desenvolvida, é testado para o método 1 a aplicação do limite de Betz para a análise de registos de funcionamento, podendo estes resultados ser observados na figura 5.3, com os registos rejeitados pelo limite teórico de Betz devidamente representados a amarelo.

Relativamente às saídas de aplicação representativas da relação rotações por minuto da turbina eólica-velocidade do vento, apenas são realizadas para os métodos 1 e 3, porque os restantes utilizam a curva teórica de potência para a realização de uma análise comparativa na classificação de dados. Este tipo de análise permite-nos perceber limitações referentes à rotação da turbina do aerogerador, sendo que para este caso de funcionamento, não se verificam qualquer tipo de limitações dessa ordem.

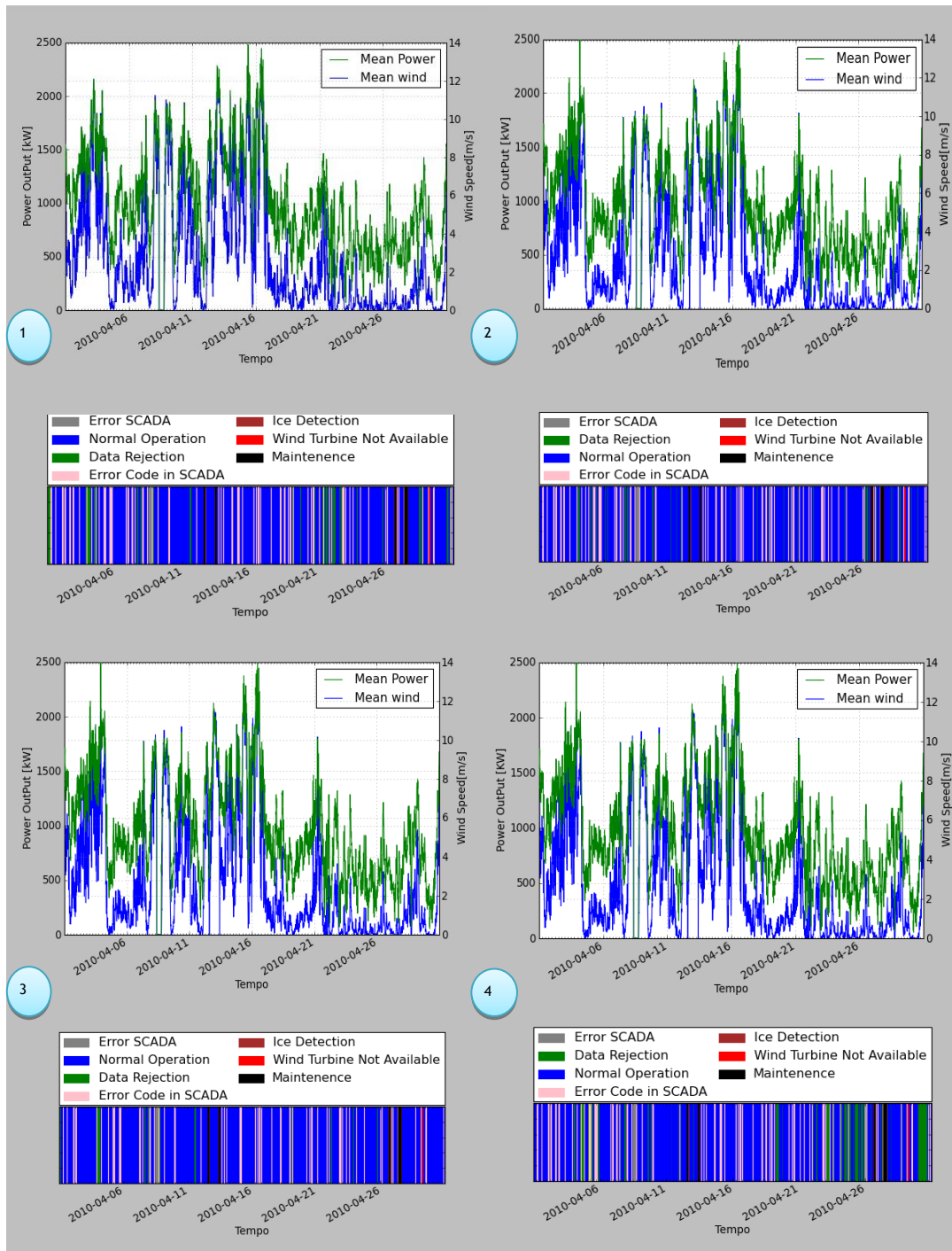


Figura 5.4 - Outras saídas gráficas da aplicação, para aerogerador λ no intervalo de tempo T1

Da figura 5.4 é possível analisar um conjunto de saídas fornecidas pela aplicação indicadoras de diferentes estados de funcionamento de um aerogerador durante o tempo de funcionamento em análise. O caso teste da figura 5.4 permite verificar o que já por outras saídas da aplicação se consegue identificar, ou seja, que estamos na presença de um período bastante regular de funcionamento, detectando-se apenas alguns períodos de erro de leitura

SCADA e de manutenção, mas que na generalidade não influenciam de todo o desempenho energético do aerogerador no período em análise.

Nesta análise os diferentes métodos apresentam resultados bastante semelhantes, dado que grande parte da informação apresentada provém do SCADA e posteriormente não sofre qualquer tipo de cálculo, sendo apenas disponibilizada para consulta. Contudo existem registos apresentados que dependem da classificação realizada por cada método sobre os dados analisados, definindo-os como *outliers*, ou como indicadores de um funcionamento normal do aerogerador. Pela observação da figura 5.4 é ainda possível verificar uma relação da potência com a velocidade do vento ao longo do período em análise. Esta análise pode revelar-se também ela bastante útil, porque permite analisar directamente o comportamento destas duas variáveis, identificando períodos de possível ausência de leitura de Scada e períodos em que o aerogerador não se encontra em produção, que são confirmados pelo complemento gráfico que indica por diferentes cores o estado do funcionamento do aerogerador para um determinado período de tempo.

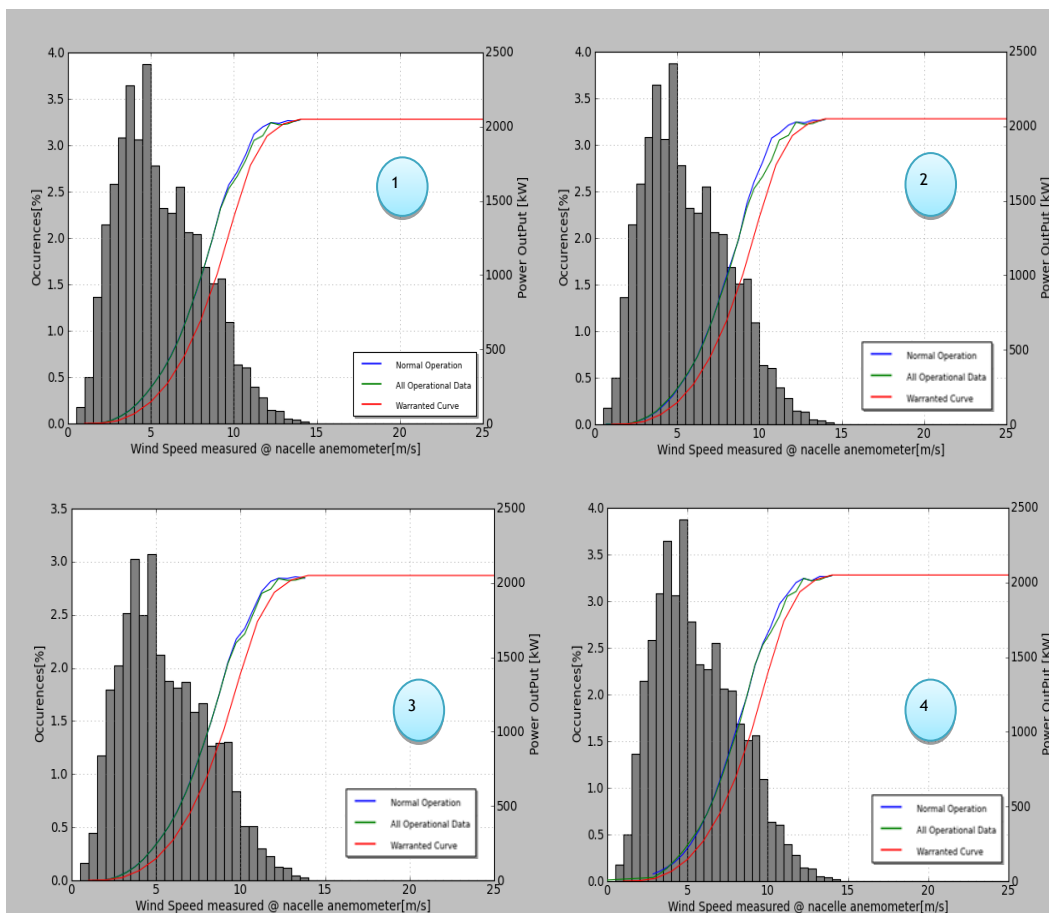


Figura 5.5 - Saídas gráficas da aplicação, para análise do rendimento energético do aerogerador λ no intervalo de tempo T1

O desempenho energético de um aerogerador é uma das razões fulcrais, que motiva a que se proceda à monitorização dos mesmos. Assim sendo, a aplicação desenvolvida realiza análises específicas, por forma a obter um conhecimento do valor energético obtido por cada método, permitindo a realização de uma comparação de desempenho entre estes.

Uma das análises realizadas neste sentido é representada na figura 5.4, em que para cada método as curvas de energia esperada, medida e teórica são representadas, possibilitando uma comparação gráfica do rendimento entre métodos. Estas curvas são elaboradas realizando-se uma análise *bin a bin* do valor de potência médio dos registos presentes em cada *bin*. No caso em análise, verifica-se que os métodos possuem desempenhos bastante semelhantes, notando-se uma ligeira superioridade de rendimento relativamente ao método 4, justificado pelo bom comportamento da curva de potência representativa dos registos classificados como em normal funcionamento, na proximidade da zona de controlo do *pitch*.

Tabela 5.1 – Resultado da análise numérica para o Aerogerador λ no intervalo de tempo T1

Análise de Eficiência	Metodologia baseada no desvio padrão (1)	Metodologia baseada na aplicação de um modelo auto regressivo (3)	Metodologia baseada na aplicação do critério de Chauvenet (2)	Metodologia baseada numa aproximação à curva de potência de um aerogerador (4)
Erro PECC	6,42%	71,94%	4,29%	54,37%
Erro absoluto médio	7,42%	6,01%	7,43%	3,57%
Erro absoluto quadrático	9,36%	8,89%	9,38%	5,34%
% Registos em normal funcionamento	92,35%	27,71%	75,89%	45,07%
% Registos rejeitados pelo método (<i>outliers</i>)	1,85%	69,11%	0,856%	51,77%
Energia Esperada [MWh]	355,33	357,27	355,40	360,03
Energia Medida [MWh]	353,27	353,27	353,27	353,27
Desempenho Energético	99,42%	98,88%	99,40%	98,12%

Tal como descrito no capítulo 4, a aplicação desenvolvida possui um conjunto de saídas não gráficas, que permitem retirar informações importantes do funcionamento dos aerogeradores, complementando uma análise gráfica, que pode por vezes revelar-se insuficiente para perceber em detalhe o comportamento do sistema. Na tabela 5.1 são apresentados alguns cálculos realizados para cada método implementado, de forma a conseguir uma análise sobre a sua eficiência e a proceder a uma comparação entre os mesmos.

O erro de PECC é o resultado da razão entre os registos representativos de normal funcionamento e a totalidade de registos em análise, servindo apenas para que se tenha uma informação relativa à quantidade de registos que cada método classifica como *outliers* de funcionamento. Esta informação poderá ser útil em alguns casos, mas por si só não é suficiente para se concluir algo acerca do desempenho de um determinado método. No caso apresentado, os valores percentuais reduzidos apresentados pelas saídas 1 e 3 para o erro de PECC, apoiados pela análise gráfica realizada anteriormente, podem fazer denotar que esses mesmos métodos poderão não estar a realizar uma classificação eficaz, pela quantidade reduzida de registos que classificam como *outliers*, principalmente no caso da saída 1. Por outro lado, uma observação atenta sobre a saída 2 indica que grande parte dos registos é classificada como *outliers* de funcionamento, o que poderá deixar também algumas dúvidas sobre a boa análise classificativa do referido método.

De forma a perceber, qual o erro dos métodos por comparação com a curva teórica dos aerogeradores, foram calculados os erros absolutos médio e quadrático para cada um relativamente a essa mesma curva, que podem ser observados na tabela 5.1. Os valores obtidos para cada método, são bastante semelhantes, sendo os calculados pelos métodos 1 e 3 (métodos estatísticos) ligeiramente piores. Isto poderá ser justificado pelo facto dos referidos métodos se basearem em médias e desvios padrão para a realização de classificações, ao passo que os outros dois métodos implementados se baseiam na curva teórica de potência dos aerogeradores para a classificação de dados, conferindo um erro inferior ao seu processo classificativo.

Para testar o desempenho obtido pelos diferentes métodos, de forma a complementar a análise gráfica anteriormente efectuada na figura 5.4, foi calculado o valor energético esperado para o funcionamento do aerogerador, tendo em conta a totalidade dos registos de potência referentes ao período em análise e o valor energético medido, obtido através do conjunto de registos classificados como registos de normal funcionamento. A razão entre estes cálculos energéticos possibilitou obter o desempenho obtido por cada método, para o caso em análise, confirmando a análise gráfica efectuada pela figura 5.4 e indicando o método 4 como sendo o que melhor desempenho possui, para o período de tempo em análise.

5.2 - Resultado da Análise efectuada sobre o Aerogerador ω para um intervalo de tempo T2

No conjunto de testes, realizados para um período de tempo T2, procurou-se realizar uma análise sobre registos de funcionamento de um aerogerador, que apresenta registos menos comuns de funcionamento, devido a condições inesperadas ou impostas deliberadamente.

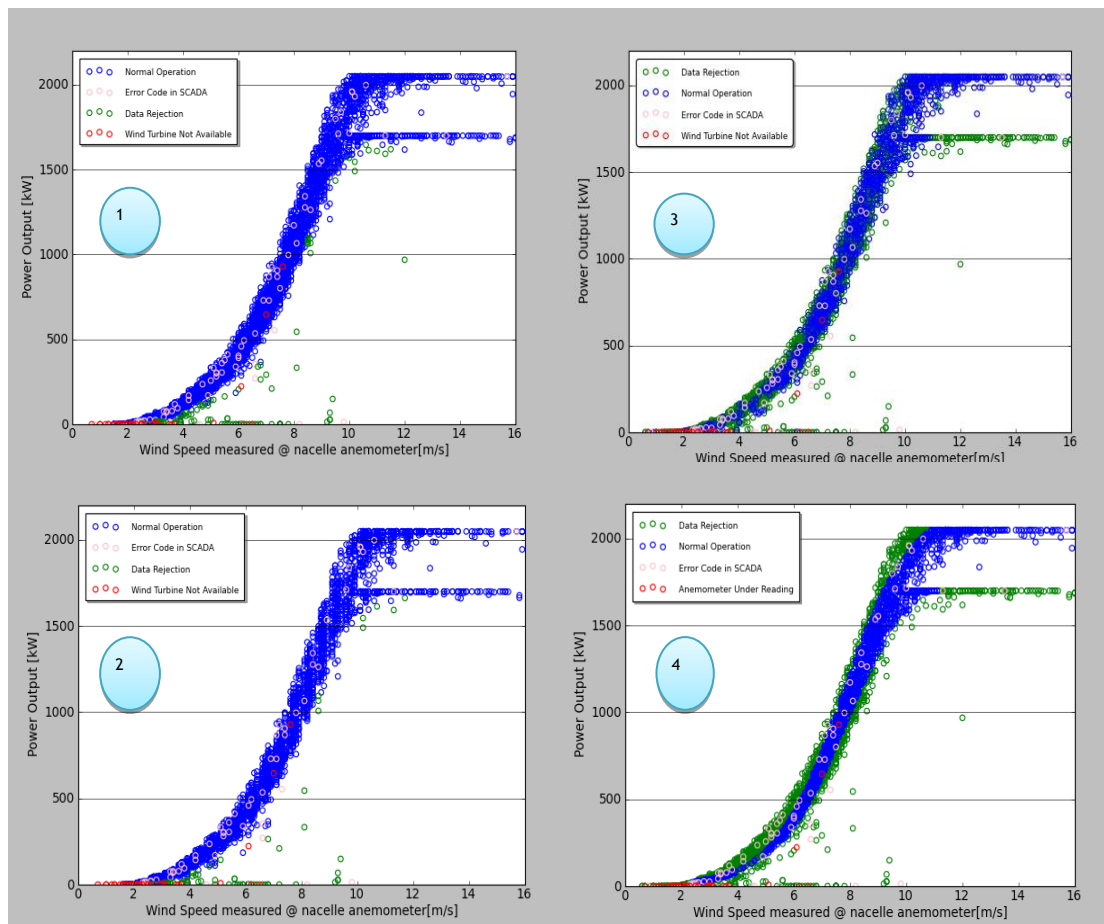


Figura 5.6 - Análise gráfica das relações Potência-Velocidade do vento para os métodos implementados no intervalo de tempo T2

Na figura 5.5, pode-se observar as curvas da saída da relação potência-velocidade do vento, para a análise efectuada, que demonstra de forma clara, que para este período ocorreu uma limitação de potência, ainda que se desconheça o motivo dessa ocorrência.

Nas saídas 1 e 2, é possível verificar que os métodos estatísticos revelam-se incapazes de identificar alguns registos *outliers* de funcionamento. No caso em análise esse facto fica ainda mais evidenciado pela limitação de potência verificada, cujos registos correspondentes, não são classificados como *outliers* de funcionamento pelos métodos das saídas 1 e 2. Este facto é justificado pelas mesmas razões já esclarecidas no subcapítulo 5.1, que explicam que este tipo de métodos realiza uma análise por *bins* de velocidade, calculando as médias e desvios padrão de potência para cada *bin*, estabelecendo de seguida, quais os valores que se encontram como *outliers* ou não.

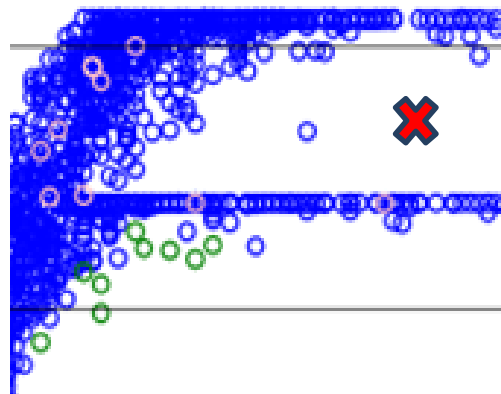


Figura 5.7 - Curva de Potência ampliada

Passa-se que em situações de funcionamento como a representada na figura 5.5, este sistema de classificação não é suficientemente eficaz, dado que a média calculado para os *bins* na zona representada pela figura 5.6 (representação ampliada de parte do método 1) se encontra perto da zona indicada com um “x” possibilitando que o intervalo definido para classificação de dados nesta zona, inclua valores de normal funcionamento e valores *outliers* de funcionamento.

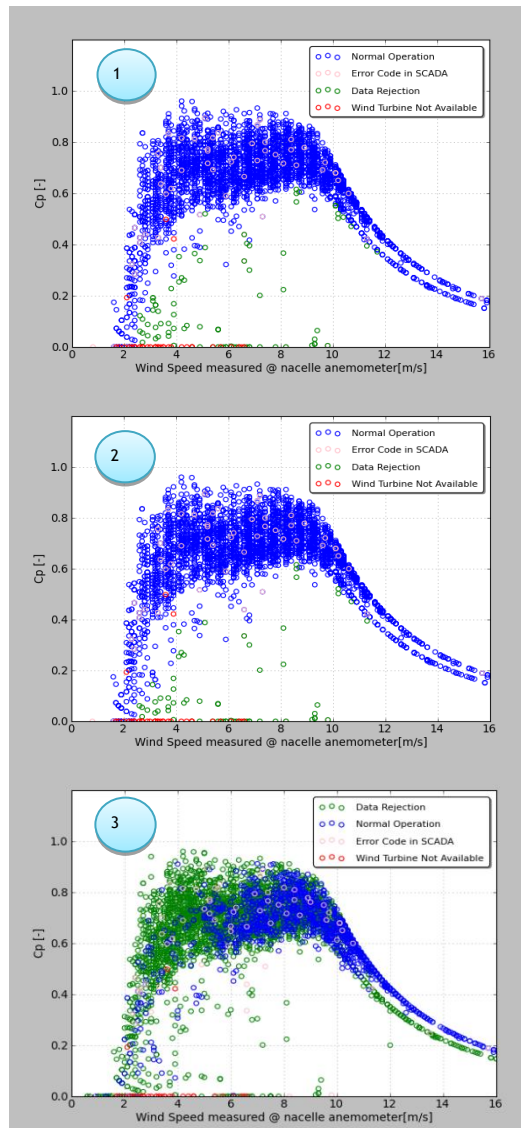


Figura 5.8 - Saídas gráficas da aplicação, para aerogerador ω no intervalo de tempo T2

Outras saídas da aplicação como as representadas pela figura 5.6 são esclarecedoras relativamente a aspectos de funcionamento já analisados na figura 5.5. Da análise da figura 5.6, se distingue facilmente a limitação de potência ocorrida durante o período de funcionamento e a capacidade dos diferentes métodos em classificar essa ocorrência de forma adequada. É possível ainda observar que o método 3, baseado num método de previsão auto-regressivo, consegue obter melhores resultados na classificação destes *outliers* de funcionamento, identificando a verde a área representativa do limite de potência ocorrido, o que não acontece para os métodos 1 e 2.

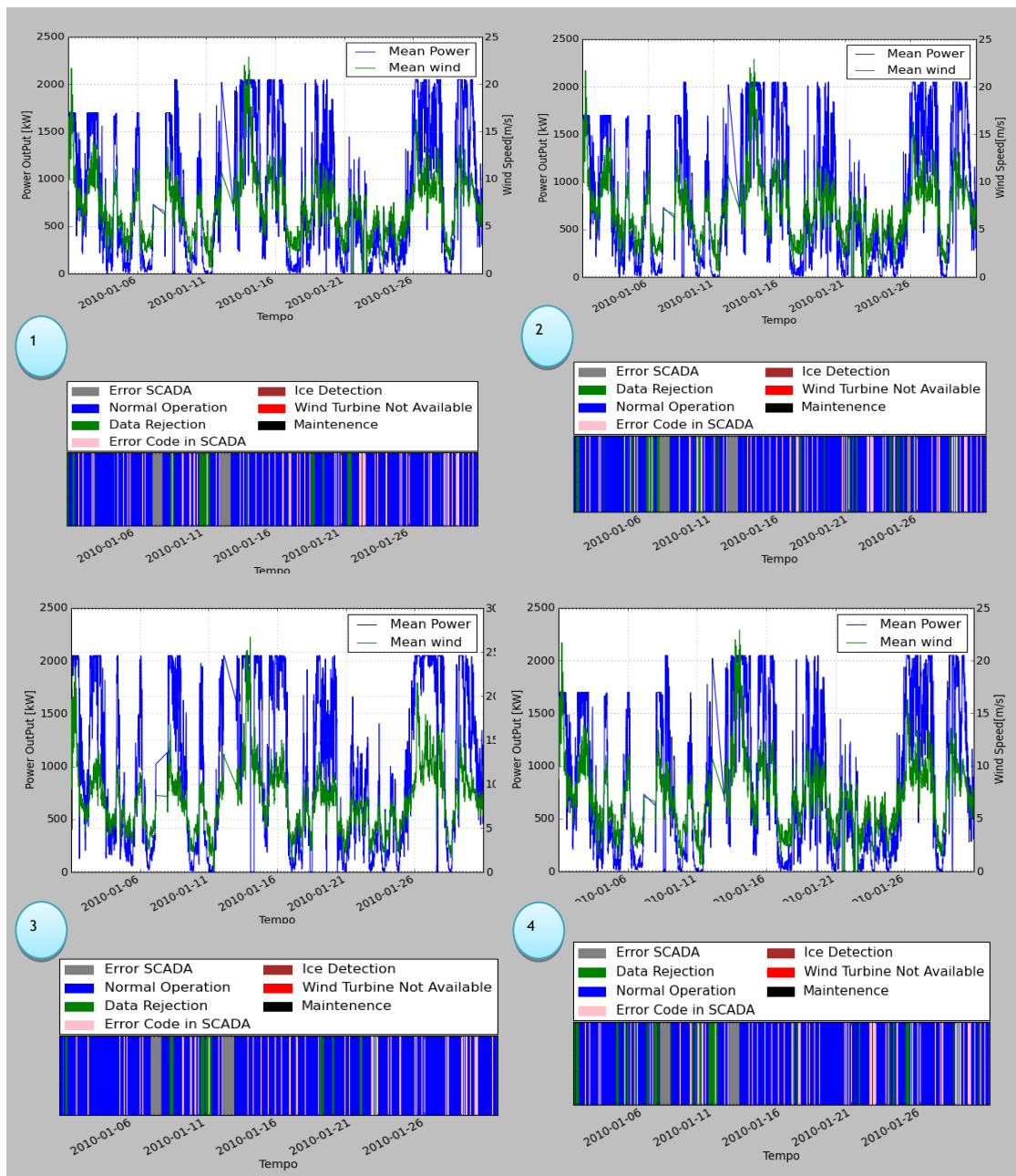


Figura 5.9 - Outras saídas gráficas da aplicação para o aerogerador ω no intervalo de tempo T2

A informação disponibilizada pelo SCADA e apresentada pela figura 5.7, do funcionamento deste aerogerador é relativamente diferente da analisada para o aerogerador λ . Períodos de falha de sistema, representados a cinzento, ocorrem por períodos de tempo significativos e são identificados por todos os métodos. Estes estados de funcionamento, não permitem perceber a origem do problema e verificar se de facto algum problema existiu que

influenciado o desempenho energético do elemento electroprodutor. É ainda comum a todas as saídas períodos de erro de leitura de SCADA e períodos de potência nula, para valores de velocidade não nulos, que representam paragens de produção de energia.

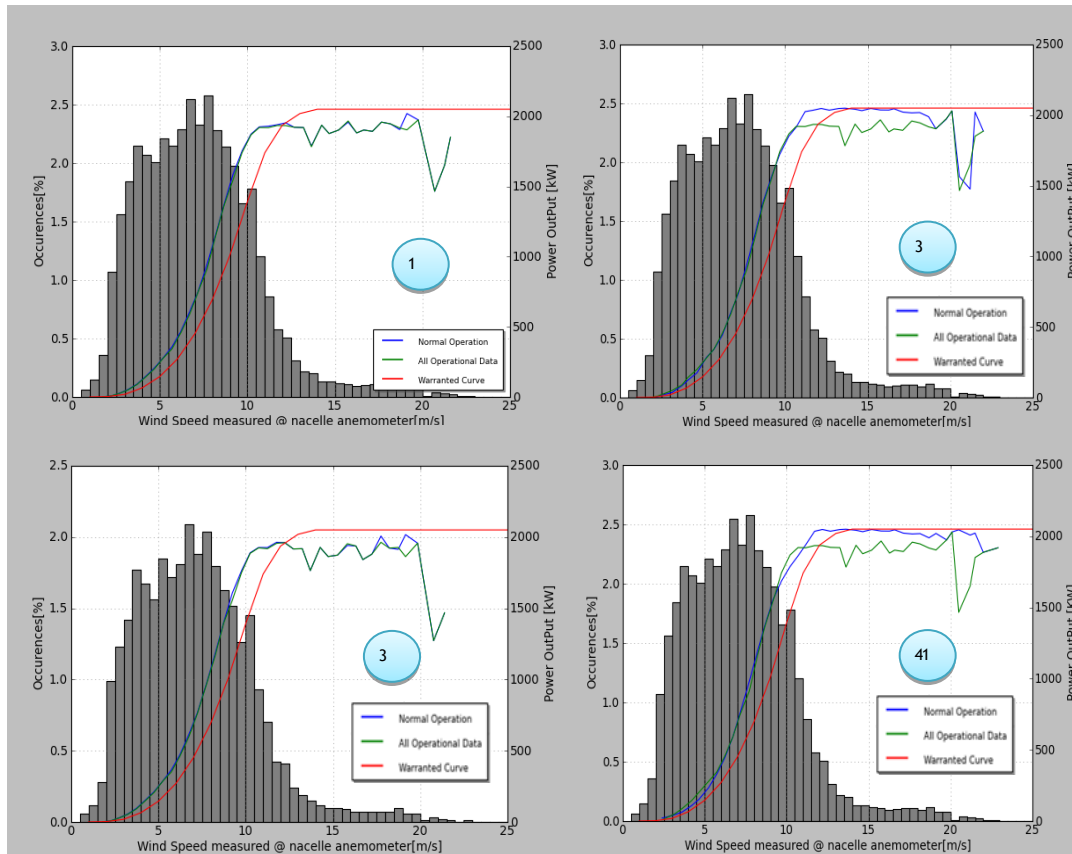


Figura 5.10 - Saídas gráficas da aplicação, para análise do rendimento energético do aerogerador ω no intervalo de tempo T2

A análise gráfica das saídas representativas do balanço energético do aerogerador ω para o período de tempo T2, revela-se de extrema importância para que se possa verificar o comportamento dos diferentes métodos, perante a limitação de potência imposta durante este tempo de funcionamento.

Observando a figura 5.8, verificamos de forma clara que o melhor desempenho energético registado, refere-se ao método 4, que obtém um comportamento para a sua curva representativa de dados em normal operação, muito próxima da curva teórica para o aerogerador em estudo. Os restantes métodos não apresentam resultados tão positivos, verificando para o método 2 um bom comportamento durante alguns períodos de

funcionamento mais críticos, que os métodos 1 e 3 não conseguem resolver, mas no entanto não consegue manter esse comportamento para toda análise classificativa realizada.

Tabela 5.2 – Resultado da análise numérica para o Aerogerador ω no intervalo de tempo T2

Análise de Eficiência	Metodologia baseada no desvio padrão	Metodologia baseada na aplicação de um modelo auto regressivo	Metodologia baseada na aplicação do critério de Chauvenet	Metodologia baseada numa aproximação à curva de potência de um aerogerador
Erro PECC	5,68%	69,62%	4,62%	55,05%
Erro absoluto médio	11,53%	8,43%	11,51%	4,67%
Erro absoluto quadrático	13,98%	10,23%	13,97%	6,78%
% Registos em normal funcionamento	93,60%	30,14%	75,39%	44,60%
% Registos rejeitados pelo método (<i>outliers</i>)	6,93%	67,25%	1,54%	52,81%
Energia Esperada [MWh]	609,67	613,75	608,49	617,27
Energia Medida [MWh]	603,57	603,57	603,57	603,57
Desempenho Energético	99,00%	98,34%	99,19%	97,78%

Para o aerogerador ω , os métodos 1 e 3, apresentam uma percentagem de valores classificados como *outliers* inferior aos outros dois métodos, registando-se contudo uma redução substancial do número de registos classificados como *outliers* pelo método 2.

Uma análise sobre os erros absoluto médio e quadrático permite verificar que os métodos 2 e 4 possuem um menor erro relativamente à curva teórica do que os métodos 1 e 3, isso poderá dever-se como já foi reflectido anteriormente, ao facto de estes últimos, recorrerem a desvios e médias calculados através dos registos em análise, para a classificação de dados, o que não acontece com os outros métodos, que são fundamentados em aproximações (4) e métodos de previsão (2), que têm como base comparativa a curva teórica de potência para classificação, diminuindo desta forma, o erro de classificação.

As percentagens de desempenho, são elevadas para o aerogerador ω , confirmando a informação fornecida pela figura 5.9. É de salientar ainda que para o aerogerador ω , verifica-se uma melhoria de rendimento em quase todos os métodos, possivelmente devido ao limite de potência aplicado sobre o aerogerador durante o período de funcionamento. É excepção a esta tendência o método 4, que em ambos os casos apresenta sempre elevados níveis de desempenho.

5.3 - Resultado da Análise efectuada sobre diferentes aerogeradores

Quando se pretende realizar testes sobre algoritmos ou metodologias, por forma a retirar conclusões sobre quais os melhores para serem utilizados para um determinado fim, é de boa prática, seja qual for a área de aplicação, calcular o erro inerente ao conjunto de metodologias ou algoritmos em análise, para determinado grupo de dados de teste.

Desta forma, para efeitos de teste entre métodos são calculados o erro absoluto médio e o erro absoluto quadrático, sobre diversos conjuntos de dados referentes ao funcionamento de vários aerogeradores. Nas tabelas 5.3 e 5.4, são representados os resultados obtidos para o erro absoluto médio e para o erro absoluto quadrático respectivamente.

Tabela 5.3 – Resultado da análise numérica para diferentes aerogeradores relativamente ao erro absoluto médio.

Aerogeradores	Metodologia baseada no desvio padrão (1)	Metodologia baseada na aplicação de um modelo auto regressivo (2)	Metodologia baseada na aplicação do critério de Chauvenet (3)	Metodologia baseada numa aproximação à curva de potência de um aerogerador (4)
Aerogerador 1	11,00%	8,98%	11,12%	6,78%
Aerogerador 2	7,76%	8,90%	9,33%	6,77%
Aerogerador 3	10,47%	7,61%	10,41%	5,56%
Aerogerador 4	10%	7,12%	9,18%	5,23%
Aerogerador 5	7,53%	8,16%	9,38%	6,03%
Aerogerador 6	8,23%	8,98%	8,34%	6,24%
Aerogerador 7	6,98%	9,34%	9,45%	7,56%
Aerogerador 8	10,72%	9,04%	10,05%	7,40%
Aerogerador 9	11,45%	8,37%	10,77%	7,03%

Das tabelas 5.3 e 5.4 apresentadas, verificamos que os valores percentuais obtidos para os dois tipos de erro calculado, não variam muito quando os calculámos para diferentes métodos ou para diferentes conjuntos de registos de funcionamento. Para o erro absoluto médio os valores obtidos variam entre os 7 e os 14 por cento, enquanto os valores de erro absoluto quadrado, são ligeiramente superiores variando entre os 9 e os 15 %.

Existe uma ligeira tendência para os valores referentes ao método baseado num modelo auto-regressivo e ao baseado numa aproximação à curva teórica de potência, apresentarem erros superiores, aos métodos baseados em critérios estatísticos, como o desvio padrão e critério de Chauvenet. As razões para essa tendência deverão estar relacionadas, como já antes mencionado com o facto de estes últimos recorrerem a desvios e médias calculados

através dos registos em análise para a classificar dados, enquanto os restantes métodos, fundamentam-se em aproximações e métodos de previsão.

Tabela 5.4 – Resultado da análise numérica para diferentes aerogeradores relativamente ao erro absoluto médio.

Aerogeradores	Metodologia baseada no desvio padrão (1)	Metodologia baseada na aplicação de um modelo auto regressivo (2)	Metodologia baseada na aplicação do critério de Chauvenet (3)	Metodologia baseada numa aproximação à curva de potência de um aerogerador (4)
Aerogerador 1	13,40%	10,51%	13,38%	8,89%
Aerogerador 2	9,35%	6,83%	7,75%	9,58%
Aerogerador 3	12,24%	9,77%	12,24%	12,29%
Aerogerador 4	12,2%	9,63%	11,48%	12,07%
Aerogerador 5	9,76%	6,45%	11,94%	9,18%
Aerogerador 6	10,7%	7,89%	10,12%	110,94%
Aerogerador 7	9,23%	6,92%	12,17%	10,64%
Aerogerador 8	12,66%	9,73%	11,23%	13,09%
Aerogerador 9	13,11%	10,08%	12,61%	14,23%

5.4 - Considerações Finais sobre os resultados

A avaliação efectuada sobre os métodos de classificação, não permite definir a existência de um método perfeito. Os métodos desenvolvidos possuem características diferentes, que os permitem obter melhores resultados algumas análises do que em outras. Os métodos 2 e 4, apresentam normalmente erros absolutos médios e quadráticos inferiores aos restantes dois

métodos, possuindo uma capacidade para detecção de *outliers* em algumas situações mais incaracterísticas de funcionamento dos aerogeradores, tal como ficou demonstrado pelo exemplo de limitação de potência do subcapítulo 5.2.

Além dos contrastes existentes entre métodos, semelhanças também existem, como é o caso dos valores obtidos para os desempenhos energéticos dos métodos que sofrem muito poucas variações de método para método e que permitem afirmar que todos estes são válidos e podem ser aplicados, para a classificação de registos de funcionamento de aerogeradores.

Capítulo 6

Conclusão

A realização de uma dissertação cujo tema se insere numa área tão vasta e em voga como é o caso das energias renováveis, permite colaborar para o crescimento científico, contribuindo assim para a constante evolução nesta área e para um crescimento sustentável da Sociedade.

O estudo de enquadramento desta dissertação permite reflectir sobre a importância que as energias renováveis terão nesse futuro que se anseia construir. Através do recurso a meios não poluentes como a exploração da energia proveniente do vento, é possível colmatar necessidades energéticas que até aos dias de hoje eram resolvidas pelo recurso a combustíveis fósseis. Esta exploração é realizada, recorrendo a elementos electroprodutores designados de aerogeradores, cuja correcta monitorização é essencial para que se garanta uma maior optimização do processo de transformação de energia. Diversas causas externas e internas podem originar o seu incorrecto funcionamento implicando a necessidade de recorrer a métodos que realizem a sua identificação e eventual previsão ou estimação. Esta informação pode ser utilizada na elaboração dos planos de manutenção de cada máquina, evidenciando as acções correctivas e preventivas de manutenção

Do pré-estudo realizado conclui-se que diversificadas ferramentas de análise foram desenvolvidas no para apoiar a gestão de operação dos parques eólicos. Na análise de funcionamento de um parque eólico, diversos estudos podem ser realizados, quando se pretende incrementar o seu rendimento energético. Os estudos até ao momento realizados apoiam-se substancialmente em métodos estatísticos para solucionar problemas de classificação de dados de funcionamento de aerogeradores e em modelos de previsão baseados em redes neuronais e modelos ARIMA (Auto-regressivo integrado de média móvel) para modelar séries temporais apoiando assim as decisões de gestão de operação. Avaliações sobre os equipamentos constituintes dos aerogeradores e classificações realizadas sobre o seu estado de funcionamento, são apenas alguns dos exemplos de monitorização que existem e podem ser aplicados para auxílio ao bom funcionamento de um parque eólico.

Orientando o estudo realizado para esta dissertação para a área da classificação de registos operativos extraídos do funcionamento de um parque eólico, foi possível concluir o

que seria essencial analisar desse funcionamento e de como se pode obter, ou seja, que parâmetros seriam necessários avaliar para que fosse possível extrair informação útil ao gestor de um parque. Analisando o conjunto de dados fornecidos pelos sistemas SCADA foi possível verificar que relações existem entre alguns dos parâmetros por estes monitorizados, que permitem discernir informações importantes de funcionamento com influência directa sobre o desempenho energético obtido. Alguns exemplos desta informação obtida referem-se a limitações de potência registadas durante o tempo de produção, problemas de controlo aerodinâmico das pás constituintes dos aerogeradores (designados normalmente de problemas de controlo de *pitch*), falta de calibração dos anemómetros instalados nos aerogeradores e limites de rotação dos aerogeradores.

No desenvolvimento desta dissertação, foram desenvolvidos um conjunto de métodos que permitem realizar uma classificação de dados, baseada em análises de funcionamento de aerogeradores, através de registos de informação adquiridos por sistemas SCADA. Com a abordagem realizada sobre o problema pretendia-se não só implementar métodos de classificação já anteriormente utilizados de índole estatística, mas também a elaboração de novos métodos que permitissem verificar a possibilidade de classificação de dados recorrendo a outros critérios. Com uma revisão efectuada sobre conceitos e teoremas anteriormente estudados, foi possível concluir que o desenvolvimento de outros métodos seria possível recorrendo a modelos auto regressivos e através de aproximações à curva teórica de potência dos aerogeradores em análise.

Após implementação dos métodos acima referidos, conjuntos de testes foram realizados sobre os mesmos, por forma a concluir o seu comportamento sobre diferentes tipos de dados, referentes a diferentes condições de funcionamento e a diferentes aerogeradores. Com as análises gráficas e numéricas comparativas efectuadas, pretendeu-se verificar níveis os de desempenho energético dos métodos, erros absolutos médio e quadrático em relação à curva teórica dos aerogeradores em análise e ainda a verificação de outros parâmetros que de ajuda para uma análise comparativa.

Dos testes realizados, foi possível concluir que não existe um método perfeito para classificação de dados, mas que estes se complementam. Se é verdade que alguns métodos de índole estatística apresentam normalmente erros absolutos médios e quadráticos inferiores, também é verdade que não possuem por vezes a mesma capacidade para detecção de *outliers* em algumas situações mais incaracterísticas que outros possuem.

Assim sendo, e como conclusão final de todo o trabalho desenvolvido para esta dissertação, chega-se à conclusão que todos os métodos desenvolvidos são válidos e que o facto de se comportarem melhor ou pior para diferentes situações, pode ser ultrapassado, se estes forem usados simultaneamente para o objectivo nuclear para o qual foram implementados, proceder classificação de dados, de forma a extrair informação útil dos mesmos.

Desenvolvimento de trabalhos futuros

Desenvolvida a aplicação para classificação de dados no âmbito da presente dissertação, desde já se identificam alguns trabalhos futuros que podem e trazer alguma melhoria ao sistema desenvolvido. Um dos aspectos de maior relevo que pode ser visto como o passo seguinte a ser realizado, é a adaptação da aplicação desenvolvida ao ambiente industrial. Para que fosse possível uma avaliação comparativa entre os diferentes métodos desenvolvidos para classificação de dados, estes foram implementados isoladamente, sendo que, para a sua aplicação a uma situação real, seria necessário a realização de um novo conjunto de testes, por forma a verificar que junções se teriam de realizar entre estes, de forma a avaliar situações específicas de funcionamento. Apesar de na presente dissertação se ter identificado o método que melhores resultados médios obteve para o problema em causa, em determinadas situações esse mesmo método pode não conseguir classificar correctamente alguns registos que deveriam ser considerados de outliers. Sendo assim, a utilização de outro método que apesar de apresentar um desempenho médio inferior pode comportar-se melhor na situação em causa. A associação de metodologias permitiria fortalecer o sistema desenvolvido através da assimilação de modelos de classificação.

Outros novos desenvolvimentos também poderão contribuir para uma melhoria significativa do desempenho do sistema. O recurso a redes neuronais, de forma a tirar partido do seu funcionamento como estimador de estados, pode revelar-se bastante útil na estimação de valores registados, através de valores medidos, criando uma base comparativa, que possibilitaria, uma classificação de registos de informação armazenados em diferentes categorias de funcionamento. Uma opção não menos válida seria o uso de lógica difusa para a classificação da informação SCADA adquirida. A lógica difusa é uma extensão da lógica *booleana*, que permite a classificação de valores recorrendo a processos estatísticos. As implementações da lógica difusa permitem que estados indeterminados possam ser tratados por dispositivos de controlo, através da aplicação de um conjunto de regras definidas pelo programador, que permitem actuar sobre um conjunto de dados de entrada e destes obter saídas coerentes com as regras definidas. A utilização de conjuntos de métodos numéricos para análise e reconhecimento de padrões para aplicação na classificação de valores registados do funcionamento de aerogeradores, é também uma outra implementação que pode ser desenvolvida, como por exemplo SVM (Support Vector Machines). Estes métodos actuam sobre um conjunto de valores de entrada, definindo a qual das duas classes de saída possíveis, pertencem esses mesmos valores, possibilitando a sua classificação.

Referências

- [1] Century, "Renewables 2010 - Global Status Report," 2010, Disponível em: <http://www.ren21.net/REN21ProductsServices/Publications/GlobalStatusReport/Renewables2010GlobalStatusReport/tabid/5824/Default.aspx>
- [2] Getulina/SP, *EFEITO ESTUFA*. Disponível em: <http://www.fc.unesp.br/~lavarda/procie/dez14/luciana/index.htm>
- [3] Climaconforto. *Metas "20-20-20" da Comissão Europeia para 2020*. Disponível em: http://www.climaconforto.pt/?PRODUTOS:Bombas_calor_Climatizacao_alta_temperatura_energia_renovavel:Energia_Sustentavel_Metas_20-20-20_reducoes_emissoes_CO2
- [4] APREN, Energias Renováveis, Disponível em: <http://www.apren.pt>
- [5] RE.NEW.ABLE. *Novas Energias a inspirar Portugal*," Lisboa.
- [6] EurObserv'ER. *wind energy barometer*. Disponível em: <http://www.euroobserver.org/pdf/baro195.pdf>
- [7] REN. *Estatística Mensal - SEN, 2010*. Disponível em: <http://www.centrodeinformacao.ren.pt/PT/InformacaoExploracao/Paginas/EstatisticaMensual.aspx>
- [8] CRESESB. *Tipos de Aerogeradores para Geração de Energia Elétrica*. Disponível em: http://www.cresesb.cepel.br/index.php?link=/tutorial/tutorial_eolica.htm
- [9] IEC. *INTERNATIONAL STANDARD IEC 61400-25-1 - Wind Turbines*
- [10] HARMAN, Analytical Techniques for understanding the performance of operating wind farms, EWEC 2003
- [11] MegaJoule ,Helder Carvalho-Paulo Pinto-Ricarxo Guedes-Miguel Ferreira, "Monitoring the Performance of a Wind Farm Production, EWEC 2006"
- [12] Natural Power Consultants Ltd, "Downtime Analysis for Improved Decision-Making," ed, 2007.
- [13] Vikramaditya Jakkul-Diane Cook, "Outlier Detection in Smart Environment Structured Power Datasets", Washington University
- [14] J. Voegele, (27-12-2010). *Programming Language Comparison*. Disponível em: <http://www.jvoegele.com/software/langcomp.html>
- [15] Miguel Rocha, Paulo Cortez, José Maia Neves, *Análise inteligente de dados*. Lisboa: FCA - Editora de Informatica Lda, 2008.
- [16] MathWorld, *StandardDeviation*, Disponível em: <http://mathworld.wolfram.com/StandardDeviation.html>
- [17] Portal Energia, "Curva Teórica de um Aerogerador.", Disponível em : http://www.portal-energia.com/downloads/aulas/Aula04_Energia%20Produzida.pdf
- [18] Curva Teórica de Potência, Alfaiate, John R. *Uma introdução à análise do erro*. 2a edição. Sausalito, Califórnia: Ciência Livro da universidade, 1997. pp 166-8.
- [19] João Nuno Monteiro da Silva e Sousa, "Previsão da Produção Eléctrica em Parques Eólicos," Master, Engenharia Electrotécnica e de Computadores (Área de

- Especialização em Energias Renováveis), FEUP - Faculdade de Engenharia da Universidade do Porto, Porto, 2007.
- [20] John Zelle, Python as a First Language. Disponível em: <http://mcs.wartburg.edu/zelle/python/python-first.html>
- [21] Alex Martelli, *Python in a Nutshell*: O`Reilly, March 2003.
- [22] MySQL, "About MySQL. Disponível em : <http://www.mysql.com/about/>"
- [23] MySQL General Information. Disponível em: <http://www-css.fnal.gov/dsg/external/freeware/mysqlInfo.html>
- [24] MySQL. *Why MySQL?* Disponível em: <http://www.mysql.com/why-mysql/>
- [25] Instituto Superior Técnico - IST, *Limite de Betz*. Disponível em: <http://elee.ist.utl.pt/realisations/EnergiesRenouvelables/FiliereEolienne/Generalites/Etudeappliquee/EtudeappliqueeEolien3.htm>

Anexos

A - Exposição das características de uma Base de Dados

Referential Integrity

Permite a uma base de dados que se mantenha consistente e válida, impondo um conjunto de regras sobre a interacção do utilizador com a base de dados que impede que este realize acções como:

- Introduzir um valor no campo de chave externa da tabela relacionada, que não exista na chave primária da tabela primária;
- Eliminar um registo de uma tabela primária se existirem registos coincidentes numa tabela relacionada;
- Alterar um valor de uma chave primária na tabela primária se esse registo tiver registos relacionados.

Transactions

É um conceito fundamental de todos os sistemas de gestão de bases de dados. Delimita um conjunto de operações da base de dados que vão operar como um todo, fazendo com que, desta forma, operações incluídas dentro de uma *Transaction* sejam validadas ou anuladas através de uma operação única.

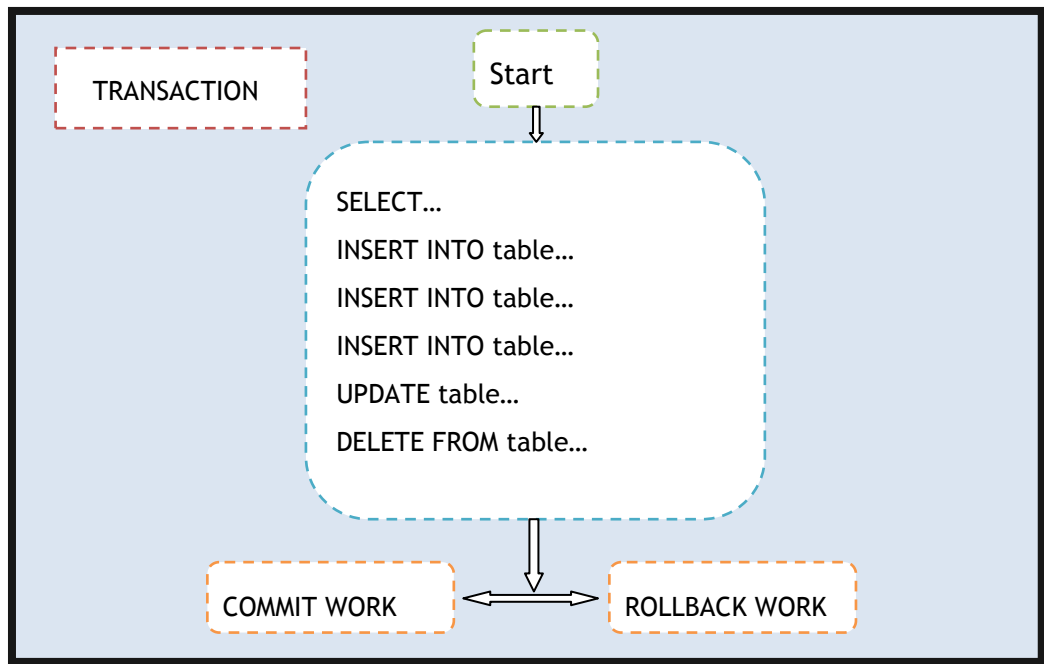


Figura 6.1 - Modo de funcionamento de uma transaction num SGBD

Unicode

O Unicode é um padrão para mapear pontos de código para caracteres. É projectado para cobrir todos os caracteres de todos os idiomas do mundo, não necessitando de páginas de código diferentes para lidar com diferentes conjuntos de caracteres.

Interface

Um SGBD disponibiliza sempre uma interface definida de forma a possibilitar a manipulação de dados de forma directa pelos utilizadores, de modo a que estes possam incluir, alterar, consultar ou eliminar dados numa base de dados. Em bases de dados relacionais, a interface é constituída pelas APIs ou drivers do próprio SGBD, que executam comandos na linguagem SQL.

Materialized View

É um objecto de base de dados que contém os resultados de uma consulta. São cópias locais de dados localizados remotamente e podem ser usados para criar resumos de tabelas com base em agregações de dados de uma tabela.

Numa *Materialized View*, o resultado da consulta é armazenado em cache como uma tabela concreta que pode ser actualizada a partir das tabelas originais de tempo a tempo.

Desta forma consegue-se um acesso mais eficiente, com a desvantagem de existir a possibilidade de alguns dados estarem desactualizados.

Temporary Table

Uma Temporary Table é uma tabela com vida útil de uma sessão ou transacção. Ela está vazia quando a sessão ou transacção começa e descarta os dados ao fim da sessão ou transacção. Com este tipo de tabela não temos necessidade de remover os dados ao término da transacção.

B - Exposição de conceitos presentes nas tabelas da Base de Dados

Tabela UCD

Esta é uma tabela que contém informações do funcionamento global de um parque eólico, ou seja, representa a unidade central de controlo e monitorização e contém informação descritiva de parâmetros como:

- date: referentes à data e hora de registo de valores de funcionamento;
- plantNo: código de identificação da central do parque eólico;
- error: erro que o SCADA atribui a cada registo;
- MRUCSMPP: Potência média [kW];
- PRUCSMPP: Potência máxima [kW];
- IRUCSMPP: Potência mínima [kW];
- MRUCSMPQ: Potência média reactiva [kW];
- PRUCSMPQ: Potência máxima reactiva [kW];
- IRUCSMPQ: Potência mínima reactiva [kW];
- MRUCSMPCOS: Ângulo entre as potências médias;
- PRUCSMPCOS: Ângulo entre as potências máximas;
- IRUCSMPCOS: Ângulo entre as potências médias;
- MRUCSMPFRE: Frequência média [Hz];
- PRUCSMPFRE: Frequência máxima [Hz];
- IRUCSMPFRE: Frequência mínima [Hz];
- MRUCSMPV1: Tensão média fase 1 [kV];
- PRUCSMPV1: Tensão máxima fase 1 [kV];
- IRUCSMPV1: Tensão mínima fase 1 [kV];
- MRUCSMPV2: Tensão média fase 2 [kV];
- PRUCSMPV2: Tensão máxima fase 2 [kV];
- IRUCSMPV2: Tensão mínima fase 2 [kV];
- MRUCSMPV3: Tensão média fase 3 [kV];
- PRUCSMPV3: Tensão máxima fase 3 [kV];
- IRUCSMPV3: Tensão mínima fase 3 [kV].

Tabela WSD

Contém informação referente ao funcionamento de cada aerogerador, armazenando dados registados em intervalos de 10 em 10 minutos de parâmetros como:

- date: referentes à data e hora de registo de valores de funcionamento;
- plantNo: Número do aerogerador num determinado parque eólico;
- class: código de classificação atribuído a cada entrada de informação, após passar pela análise dos diferentes métodos;
- MRWSMPVWI: Velocidade média do vento registada na nacelle;
- MRWSMPNROT: Rotação média registada num aerogerador [RPM];
- MRWSMPP: Potência média registada num aerogerador;
- MRWABGOPOS: Posição da nacelle (varia entre -760 e + 760; -2,5 voltas a +2,5 voltas);
- ARWABWORKH: Horas de trabalho;
- ARWABW: kWh produzidas em cada instante (acumulado);
- ARWABWRKM: Minutos de trabalho;
- MRWSMPQ: Potência média reactiva.

Tabela PES

A informação contida nesta tabela está associada à ocorrência de eventos que alteram o estado de funcionamento de um aerogerador. Estes eventos são registados pelo SCADA e preenchidos nesta tabela, por forma, a que sejam consultados mais tarde para classificação de períodos de funcionamento. Os parâmetros registados nesta tabela são:

- date: referentes à data e hora de registo de ocorrência de erro;
- plantNo: Número do aerogerador em que se verificou o erro registado;
- state: código identificador do estado geral de funcionamento de um aerogerador;
- subState: código complementar ao state, que quando associado ao mesmo, indica o estado de funcionamento pormenorizado de um aerogerador;
- Service;

- FaultMsg;
- Value.

Tabela classerror

Nesta tabela são registados os tipos de erros possíveis para classificação de dados pelo sistema. É útil do ponto de vista de consulta e de auxílio à programação, contendo informação como:

- classID: código identificativo de um determinado tipo de erro;
- name: representa o nome do tipo de erro;
- color: cor a ser usada graficamente para um determinado tipo de erro;
- notes: notas adicionais, a serem inseridas para um tipo de erro.

Tabela powercurve

Aqui estão registados alguns dados fornecidos pelos diferentes fabricantes de aerogeradores, dos quais constam:

- Code: código identificador de entrada de dados;
- WindSpeed : registos de velocidade do vento da curva teórica;
- Cp: registos de factor potência da curva teórica;
- Power: registos de potência da curva teórica.

Tabela pes_list

Tabela auxiliar á tabela Pes que contém uma lista de todos os estados de funcionamento do aerogerador que o SCADA consegue identificar. Desta tabela constam os seguintes parâmetros:

- code: código de entrada de informação na tabela (sequencial);
- state: código identificador do estado geral de funcionamento de um aerogerador;
- subState: código complementar ao state, que quando associado ao mesmo, indica o estado de funcionamento pormenorizado de um aerogerador;

- stateStr: descrição textual do código state;
- subStateStr: descrição textual do código subState;
- cause: campo de descrição auxiliar, se necessária relativamente a uma entrada de dados;
- timeKey.

Tabela windfarm

Nesta tabela, encontram-se dados identificadores mais generalistas dos parques eólicos. Aqui podem-se encontrar parâmetros do tipo:

- Code: número de código do parque eólico;
- Name: nome do parque eólico;
- LocationX: coordenada geográfica de latitude;
- LocationY: coordenada geográfica de longitude;
- PowerInstalled: refere-se ao valor total de potência instalada em cada parque monitorizado;
- LicensedPower: refere-se ao valor de potência para o qual, cada parque está licenciado para operar;
- MassaVolumica: representa a massa volúmica de ar que passa pelo sistema de pás de um aerogerador;
- TensãoNominal: tensão nominal de um aerogerador;
- FreqNominal: frequência nominal de um aerogerador.

C - Características das Linguagens de Programação

Object-Orientation

Dentro do grupo das linguagens de programação, existem as que são orientadas a objectos, que representam uma forma especial de programar, mais próxima, da forma como nos expressaríamos na vida real, do que em outros tipos de programação.

Para que uma linguagem pertença a este grupo específico, ela deve conter as seguintes qualidades:

- Encapsulamento / Ocultação de informação;
- Herança;
- Polimorfismo/ligação dinâmica;
- Todos os tipos pré-definidos são objectos;
- Todas as operações são mensagens para objectos;
- Todos os tipos definidos pelo usuário são objectos.

Tabela 6.1 – Exemplos puros e híbridos de linguagens de programação orientadas a objectos.

Características da Linguagem	Eiffel	Smalltalk	Ruby	Java	C#	C++	Python	Perl	Visual Basic
Encapsulamento/ Ocultação de informação	Sim	Sim	Sim	Sim	Sim	Sim	Não	Sim	Sim
Herança	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Não
Polimorfismo / Ligação dinâmica	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Todos os tipos pré-definidos são objectos	Sim	Sim	Sim	Não	Não	Não	Sim	Não	Não
Todas as operações são mensagens para objectos	Sim	Sim	Sim	Não	Não	Não	Não	Não	Não
Todos os tipos definidos pelo usuário são objectos	Sim	Sim	Sim	Sim	Sim	Não	Sim	Não	Não

Static vs. Dynamic Typing

As linguagens de programação são igualmente qualificadas em função de serem dinâmicas ou estáticas, diferenciando-se estes dois modelos, devido a alguns aspectos que caracterizam cada um. No caso dinâmico, temos uma maior produtividade e flexibilidade, enquanto que, no caso estático factores como, segurança e fiabilidade são os mais importantes.

- **Estáticas** - As linguagens estáticas são de entendimento e aprendizagem fácil. Numa linguagem de programação estática, como C ou Pascal, o programador deve declarar o tipo de cada variável antes de o código ser compilado, tornando o código menos flexível, mas também menos propenso a erros. Para desenvolvimento web, as linguagens estáticas, possuem funções que possibilitam o directo processamento no navegador de internet, e desta forma, não dependem de recursos de processamento de servidor, sendo que o mesmo deve apenas manter os arquivos armazenados e acessíveis;

- **Dinâmicas** - Numa linguagem dinâmica, como Perl ou LISP, o programador pode criar variáveis, sem especificar o seu tipo. Esta característica, possibilita criar aplicações mais flexíveis, embora menos seguras e propensas a erros. É ainda possível fazer alterações em diversos elementos de um site, quando este é construído com recurso a este tipo de linguagem. Além disso, as linguagens dinâmicas trazem a possibilidade de gravar e resgatar dados de bases de dados, o que se revela numa enorme vantagem relativamente às linguagens estáticas.

Generic Classes

São classes que podem conter objectos de diferentes classes e ser invocadas por diversas aplicações. Normalmente fazem parte de forma indirecta do contexto de uma aplicação com o objectivo de complementar um determinado processo.

Inheritance

Permite que as classes partilhem atributos e métodos para que se proceda a um reaproveitamento de código. É um mecanismo, que permite que diversas características comuns a diferentes classes sejam colocadas numa só classe (classe base), a partir da qual serão especificadas outras, que são designadas de classes derivadas. As classes derivadas apresentam características da classe base e ainda características particulares referentes a cada uma delas.

Feature renaming

Capacidade de uma classe de renomear as características que herdou de uma superclasse.

Method overloading

Conhecido como polimorfismo paramétrico, é a capacidade de uma classe ou módulo de ter dois ou mais métodos com o mesmo nome.

Operator overloading

Refere-se à definição de novas tarefas para o mesmo operador. Ao definirmos novas funções para os operadores padrão de uma linguagem, na verdade não substituímos a sua função, apenas adicionamos mais uma função ao mesmo operador.

Higher Order Functions & Lexical Closures

Higher Order Functions são funções que podem ser tratadas como se fossem objectos de dados. Quando vinculadas a variáveis podem ser passadas para outras funções como parâmetros, e retornadas como resultado de outras funções. Devido a esta capacidade, podem ser vistas como uma forma de execução diferida, em que uma função pode ser definida em um contexto, passar a um outro contexto e posteriormente, ser invocada pelo segundo contexto.

Dizer que uma linguagem possui, *Lexical Closures*, significa que funções internas podem ter acesso aos *nomes* das variáveis locais das funções de inclusão e não somente apenas a cópias desses mesmos *valores*.

Garbage Collection

É um processo para gestão de memória, com o qual é possível recuperar memória inutilizada, evitando problemas de vazamento de memória, que resultariam no esgotar da memória livre.

Class variables and methods

São variáveis e métodos, que pertencem a uma classe e não a uma qualquer instância específica de uma classe. Isto significa que por mais instâncias de classe que existam, em qualquer momento, apenas uma cópia de cada variável de classe/método existe e é compartilhada por todas as instâncias dessa mesma classe.

Reflection

Entende-se como a característica das linguagens que define a capacidade de um programa para determinar várias informações sobre um objecto em tempo de execução.

Isto inclui a capacidade de determinar o tipo de objecto, a sua estrutura de herança, os métodos que este contém, os tipos de parâmetros, retornos e atributos.

Access Control

Refere-se à capacidade na implementação de módulos, de permanecer oculto por trás da interface pública. O controlo de acesso está intimamente relacionado com o princípio de encapsulamento/Ocultação de informação das linguagens orientadas a objectos. A maioria das linguagens orientadas a objectos têm pelo menos dois níveis de controlo de acesso: público e protegido. As características que se encontram protegidas, não estão disponíveis fora da classe em que estão inseridas, excepto para subclasses. Algumas linguagens de programação, nomeadamente o Java e o C++, fornecem ainda um terceiro nível de controlo de acesso conhecido como “privado”. Os recursos que se encontram definidos como privados, não estão disponíveis fora da classe em que foram declarados, mesmo para as subclasses.

Multithreading

Entende-se como a capacidade para um único processo de processar duas ou mais tarefas simultaneamente (em intervalos de tempo muito curtos, no caso da inexistência de múltiplos processadores).

Regular Expressions

É um método formal de se especificar um padrão de texto. Ou seja, é uma composição de símbolos ou caracteres com funções especiais, que quando agrupados entre si e com caracteres literais, formam uma expressão. Essa expressão é interpretada como uma regra, que indicará sucesso se uma entrada de dados combinar com essa regra, obedecendo exactamente a todas as suas condições.

Pointer Arithmetic

Pointer Arithmetic é a capacidade de uma linguagem para manipular directamente os endereços de memória e os seus conteúdos. Devido à insegurança inerente à manipulação directa de memória, esta capacidade não é muitas vezes considerada adequada para linguagens de alto nível, mas favorável para aplicações em sistemas de baixo nível.

Language Integration

Capacidade que resulta da necessidade de integração de sistemas existentes. Esta necessidade é importante para que uma linguagem de alto nível seja capaz de se integrar perfeitamente com outras linguagens.

Built-In Security

Refere-se à capacidade existente numa implementação de linguagem, para determinar se uma amostra de código provém de uma fonte fidedigna ou não, limitando as permissões do código se isso não acontece.