

UNIVERSIDADE FEDERAL DO CEARÁ
DEE - DEPARTAMENTO DE ENGENHARIA ELÉTRICA
CURSO DE ENGENHARIA ELÉTRICA

DANIELLE BALTAZAR CAVALCANTE

**ANÁLISE DE DESEMPENHO DE PARQUES EÓLICOS POR MEIO
DE CLUSTERIZAÇÃO DE AEROGERADORES**

TRABALHO DE CONCLUSÃO DE CURSO

FORTALEZA
2020

DANIELLE BALTAZAR CAVALCANTE

**ANÁLISE DE DESEMPENHO DE PARQUES EÓLICOS POR MEIO
DE CLUSTERIZAÇÃO DE AEROGERADORES**

Trabalho de Conclusão de Curso apresentado ao Curso de Engenharia Elétrica da Universidade Federal do Ceará, como requisito parcial para a obtenção do título de Bacharelado.

Orientadora: Prof^a. Ph.D. Ruth P.S. Leão
Universidade Federal do Ceará

FORTALEZA
2020

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Biblioteca Universitária
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

C364a Cavalcante, Danielle Baltazar.
Análise de desempenho de parques eólicos por meio de clusterização de aerogeradores / Danielle
Baltazar Cavalcante. – 2020.
53 f. : il. color.

Trabalho de Conclusão de Curso (graduação) – Universidade Federal do Ceará, Centro de Tecnologia,
Curso de Engenharia Elétrica, Fortaleza, 2020.
Orientação: Profa. Dra. Ruth Pastôra Saraiva Leão.

1. Clusterização. 2. Aerogeradores. 3. K-means. 4. Análise de desempenho. 5. Parques eólicos. I. Título.
CDD 621.3

DANIELLE BALTAZAR CAVALCANTE

ANÁLISE DE DESEMPENHO DE PARQUES EÓLICOS POR MEIO DE
CLUSTERIZAÇÃO DE AEROGERADORES

Trabalho de Conclusão de Curso apresentado ao Curso de Engenharia Elétrica da Universidade Federal do Ceará, como requisito parcial à obtenção do título de Bacharelado em Engenharia Elétrica.
Área de concentração: Engenharia Elétrica.

Aprovada em: ___/___/_____.

BANCA EXAMINADORA

Profa. PhD. Ruth Pastôra Saraiva Leão (Orientadora)
Universidade Federal do Ceará (UFC)

Prof. Dr. Raimundo Furtado Sampaio
Universidade Federal do Ceará (UFC)

Eng. Jean Davi Simplício de Queiroz

À minha mãe, Telma Maria e minha avó, Maria do Socorro.

AGRADECIMENTOS

À minha família Telma Maria Baltazar Cavalcante, José Ribamar Ferreira Cavalcante e Arthur Victor Baltazar Cavalcante por todo suporte e paciência dados ao longo de minha vida.

Aos meus avós Maria do Socorro Ferreira Cavalcante e Francisco Ferreira Cavalcante por terem contribuído na minha educação e por sempre terem acreditado em mim.

Aos meus tios Geraldo Ferreira Cavalcante e Maria das Graças e às minhas primas Carolina Valente e Aline dos Santos por serem fontes de estímulo e por terem me ajudado a prosseguir em momentos de dificuldade.

À professora Ruth Pastôra Saraiva Leão por toda atenção e compartilhamento de ideias.

Aos amigos de graduação, com os quais compartilhei diversos momentos de adversidades e de alegria, sem vocês teria sido impossível.

*“Existe um tempo certo para cada coisa, momento oportuno para cada propósito debaixo do Sol: Tempo de nascer, tempo de morrer; tempo de plantar, tempo de colher.”
(Eclesiastes, Cap. 3).*

RESUMO

BALTAZAR, Danielle. Análise de Desempenho de Parques Eólicos por meio de Clusterização de Aerogeradores. 2020. 42 f. Trabalho de Conclusão de Curso – Curso de Engenharia Elétrica, Universidade Federal do Ceará. Fortaleza, 2020.

De acordo com a Associação Brasileira de Energia Eólica (ABEEólica), a capacidade instalada de parques eólicos no Brasil atingiu a marca de 16 GW no primeiro semestre de 2020, com 637 parques eólicos e 7.738 aerogeradores instalados, ocupando o posto de segunda fonte na matriz de energia elétrica brasileira, atrás apenas das usinas hidrelétricas. O país chegou à 7ª posição na classificação mundial do Global Wind Energy Council (GWEC). Neste contexto, surge a necessidade de analisar o uso dos recursos energéticos, de modo a prever e melhorar o desempenho de parques eólicos, identificando com maior rapidez as turbinas com possíveis anomalias na operação. Para tanto, neste trabalho, será implementada uma técnica de *data mining* utilizada em clusterização de dados conhecida como *K-means*, visando o agrupamento de 65 turbinas eólicas, pertencentes a 5 parques eólicos, em diferentes conjuntos, compostos por máquinas que apresentam características semelhantes. Na análise mensal do desempenho dos aerogeradores foi usado o algoritmo de clusterização K-means, desenvolvido em linguagem Python, que usa como entrada de dados as variáveis produção de energia, velocidade dos ventos e disponibilidade total dos aerogeradores. Os dados de entrada foram extraídos do Sistema de Supervisão e Aquisição de Dados (SCADA) e discretizados em intervalos de 60 minutos. A metodologia permite a avaliação de desempenho por meio de agrupamentos, ao invés de análise individual, facilitando o acompanhamento da operação em fazendas eólicas com grande quantidade de turbinas, o que faz do método uma ferramenta mais abrangente e robusta. Para cada grupo de aerogeradores, comparou-se o valor de produção de energia do centroide ao valor esperado pela equação da curva de potência do fabricante, dada a velocidade dos ventos e a disponibilidade de cada cluster. Os resultados obtidos apontaram que máquinas com especificações semelhantes, dispostas em um mesmo complexo eólico, podem apresentar diferentes desempenhos. Em agrupamentos onde a performance dos aerogeradores é inferior ao estimado pelo fabricante, faz-se necessário um estudo de causas-raízes do problema. A avaliação da performance de parques eólicos utilizando-se inteligência artificial tende a ser uma prática de mercado, e a implementação de técnicas de clusterização representa uma contribuição do trabalho em desenvolvimento.

Palavras-chave: K-means. Agrupamento de Aerogeradores. Desempenho de Parques Eólicos.

ABSTRACT

BALTAZAR, Danielle. Performance Analysis of Wind Farms through Clustering of Wind Turbines, 2020. 2020. 42 f. Trabalho de Conclusão de Curso – Curso de Engenharia Elétrica, Universidade Federal do Ceará. Fortaleza, 2020.

According to the Brazilian Association of Wind Energy (Abeeólica), the installed capacity of wind farms in Brazil reached the mark of 16 GW in the first half of 2020, with 637 wind farms and 7,738 wind turbines installed, occupying the second source of the Brazilian electric energy matrix, behind only the hydroelectric plants. The country reached 7th place in the Global Wind Energy Council (GWEC) world rankings. In this context, there is a need to analyze the use of energy resources in order to predict and improve the performance of wind farms, identifying more quickly turbines with possible anomalies in the operation. To this end, in this work, a textitdata Mining technique will be implemented used in data Clusterization known as textitK-Means, aiming at the grouping of 65 wind turbines in different sets, composed of machines that have similar specifications. The result of the study is the monthly analysis of the performance of wind turbines belonging to a wind farm composed of 5 wind farms. In the performance analysis was used an algorithm of Clusterization developed in Python language, which uses as input the variables energy production, wind speed and total availability of wind turbines. The input data were extracted from the Data Acquisition and Supervision System (SCADA) and discretized at 60-minute intervals. The methodology allows the evaluation of performance through clusters, rather than individual analysis, facilitating the monitoring of operation in wind farms with large amounts of wind turbines, which makes the method a more comprehensive and robust tool. For each group of wind turbines, the centroid energy production value was compared to the value expected by the manufacturer's power curve equation, given the wind speed and the availability of each cluster. The results showed that machines with similar specifications, arranged in the same wind farm, can present different performances. In clusters where the wind turbine performance is lower than estimated by the manufacturer, a root cause study of the problem is required.

Keywords: K-means. Cluster of Wind Turbines. Wind Farm Performance.

LISTA DE FIGURAS

Figura 1 – Curva $C_p \times v$	5
Figura 2 – Curva de Potência	6
Figura 3 – Tipos de turbinas: a) eixo vertical b) eixo horizontal	6
Figura 4 – Partes de um aerogerador	7
Figura 5 – Gerador de indução com controle de escorregamento	9
Figura 6 – Gerador de indução duplamente alimentado	9
Figura 7 – MTTR (h)	13
Figura 8 – MTBF (h)	14
Figura 9 – Velocidade média dos ventos ao longo do ano de 2019	16
Figura 10 – Fator de Capacidade ao longo do ano de 2019	16
Figura 11 – Diagrama de perdas em uma turbina conectada à rede	18
Figura 12 – Diferentes agrupamentos de dados.	21
Figura 13 – Fluxograma do algoritmo k-means	22
Figura 14 – Etapas do algoritmo k-means	23
Figura 15 – Curva método Elbow	24
Figura 16 – Curva método Elbow - Escolha de K	24
Figura 17 – Fluxograma código python	25
Figura 18 – Colunas da base de dados	26
Figura 19 – Método Elbow - valor ótimo de K'	28
Figura 20 – Dados velocidade x geração	29
Figura 21 – Clusters velocidade x geração	29
Figura 22 – Dados disponibilidade x geração	30
Figura 23 – Clusters disponibilidade x geração	31
Figura 24 – Obtenção de equação com software Wolfram	32
Figura 25 – Gráfico da Equação 25 com destaque para a região utilizada para representar a curva de potência do fabricante	32

LISTA DE TABELAS

Tabela 1 – MTTR e MTBF em componentes de um parque eólico	13
Tabela 2 – Recordes de Fator de Carga e Fator de Geração 2020	17
Tabela 3 – Limites de Distorção Harmônica Individual	19
Tabela 4 – Potência Instalada	20
Tabela 5 – Centroides	30
Tabela 6 – Clusters	30
Tabela 7 – Centroides	30
Tabela 8 – Clusters	31
Tabela 9 – Geração esperada utilizando Equação 25	32
Tabela 10 – Análise de Performance	33
Tabela 11 – Curva de Potência	34

SUMÁRIO

1 – Introdução	1
1.1 Objetivos	2
1.2 Estrutura do Trabalho	2
2 – Fundamentos da Geração Eólioelétrica	3
2.1 Potência do vento	3
2.2 Turbinas eólicas	4
2.2.1 Coeficiente de Performance e Velocidade Específica	4
2.2.2 Curva de Potência	5
2.2.3 Tipos de Turbinas Eólicas	5
2.3 Partes de um Aerogerador	6
2.3.1 Caixa Multiplicadora	6
2.3.2 Gerador	7
2.3.2.1 Gerador assíncrono de velocidade fixa	8
2.3.2.2 Gerador assíncrono com controle de escorregamento	8
2.3.2.3 Gerador assíncrono duplamente alimentado	9
2.3.2.4 Gerador síncrono	10
3 – Indicadores de Performance em Parques Eólicos	12
3.1 Tempo médio para reparo	12
3.2 Tempo médio entre falhas	13
3.3 Disponibilidade	14
3.4 Fator de Capacidade	15
3.5 Perdas	17
3.6 Harmônicos	18
4 – Metodologia	20
4.1 Clusterização	20
4.2 K-means	21
4.3 Método <i>Elbow</i>	23
4.4 Algoritmo Python	24
4.5 Base de Dados	25
4.6 Análise de Performance	26
5 – Análise e Discussão dos Resultados	28
5.1 Obtenção do valor ótimo de K	28
5.2 Clusterização: Velocidade dos ventos x Energia Produzida	29

5.3	Clusterização: Disponibilidade x Energia Produzida	30
5.4	Análise de Performance	31
6	– Conclusão e Trabalhos Futuros	35
6.1	Conclusão	35
6.2	Trabalhos Futuros	35
	Referências	37
	Apêndices	39
	APÊNDICE A–Código Python K-means	40
	APÊNDICE B–Código Python Elbow	42

1 Introdução

Com o avanço da participação da geração eólioelétrica na matriz energética brasileira em paralelo com o avanço da tecnologia na era da comunicação e informação, a análise de dados de operação e manutenção de aerogeradores é uma forte aliada na otimização do desempenho de parques eólicos. O gerenciamento de ativos usando estratégias de manutenção otimizadas pode ajudar a melhorar a confiabilidade e a disponibilidade das turbinas eólicas, tornando-as mais competitivas (A.KHOSRAVIL.MACHADO; R.O.NUNES, 2018).

Em (GONZALEZ et al., 2017) o autor apresenta uma revisão dos principais indicadores de performance utilizados na operação e manutenção de parques eólicos. Essas informações geralmente podem ser obtidas por meio do Sistema de Supervisão e Aquisição de Dados dos aerogeradores e são utilizados como valores de entrada para os algoritmos de análise de dados.

Em um sistema dinâmico complexo, como um parque eólico, o agrupamento é uma forma de reduzir a complexidade do modelo e melhorar o entendimento de sua dinâmica (MA; RUNOLFSSON; JIANG, 2011). A utilização de técnicas de clusterização é uma das ferramentas possíveis na análise de desempenho de parques eólicos, sendo possível a aplicação de diversas metodologias existentes na literatura.

Foram publicados alguns trabalhos relacionados à aplicação de mineração de dados na geração eólioelétrica e na operação e manutenção de parques eólicos. Em (STETCO et al., 2019) o autor faz uma revisão literária de sobre modelos de aprendizado de máquina, que têm sido aplicados em monitoramento de operação de turbinas eólicas. Os resultados mostram que redes neurais, máquinas de vetores de suporte e árvores de decisão são as técnicas mais comumente usadas.

Em (ROMERO; GALLEGO; PAVAS, 2011), o autor utiliza a técnica K-means para a estimativa de padrões de afundamentos de tensão e agrupamento de zonas de falha em redes de média e alta tensão por meio da modelagem do sistema de 115kV e 220kV da cidade de Bogotá-Colômbia. Como resultado foram obtidos 50 clusters e áreas reais onde ocorrem falhas que causam afundamentos de tensão na rede da cidade foram identificadas.

Neste trabalho foi escolhida a técnica de clusterização K-means, sendo implementado um algoritmo em linguagem Python, utilizando-se como dados de entrada a disponibilidade, a velocidade dos ventos e a geração de energia de 65 turbinas eólicas. Para a escolha do valor ótimo de K, foi adotado o método Elbow. Além disso, aplicou-se a técnica de interpolação polinomial para obtenção de equação característica que representa a curva de potência fornecida pelo fabricante, a fim de comparar os resultados de geração de energia esperada com os obtidos para cada centroide.

1.1 Objetivos

O objetivo geral deste trabalho é analisar o desempenho de turbinas eólicas de diferentes parques eólicos através do agrupamento de dados utilizando a técnica de clusterização *K-means*.

Os objetivos específicos do trabalho são:

- a) Obter dados reais de operação do sistema SCADA de 65 aerogeradores e desenvolver uma base de dados;
- b) Desenvolver, em linguagem Python, um algoritmo de clusterização, a fim de agrupar turbinas com desempenho semelhante;
- c) Aplicar um método de otimização do número de clusters;
- d) Analisar os diferentes grupos de turbinas ao longo de um mês de operação.

1.2 Estrutura do Trabalho

O presente trabalho é dividido em seis capítulos.

Capítulo 1 apresenta a justificativa do trabalho e seus objetivos.

Capítulo 2 é dedicado a realizar uma fundamentação técnica sobre a geração eólica, com apresentação de componentes e modelos de extração de energia de um aerogerador.

Capítulo 3 traz apresenta os mais importantes indicadores de operação e performance de parques eólicos.

Capítulo 4 descreve a metodologia utilizada, assim como, explica a modelagem matemática da técnica de agrupamento escolhida e o código desenvolvido em Python.

Capítulo 5 apresenta os resultados obtidos e análises realizadas após as simulações.

Capítulo 6 traz as conclusões e recomendações para trabalhos futuros.

2 Fundamentos da Geração Eólioelétrica

O aproveitamento da energia eólica para produção de eletricidade dá-se pela conversão da energia cinética do vento em energia mecânica e, com o uso de geradores elétricos, é feita a conversão desta em energia elétrica.

Neste capítulo serão apresentadas as equações características do processo de conversão eólio-elétrico, assim como os principais elementos componentes de uma turbina eólica.

2.1 Potência do vento

Vento é ar em movimento, desta forma, pode-se associar a um corpo de massa m , movendo-se a uma velocidade v , uma determinada quantidade de energia cinética, dada por:

$$E = \frac{mv^2}{2}. \quad (1)$$

Em que:

E = energia cinética [J];

m = massa do ar [kg];

v = velocidade da massa de ar [m/s].

A potência do vento é a derivada da energia cinética no tempo:

$$P = \dot{E} = \frac{dE}{dt} = \frac{\dot{m}v^2}{2}. \quad (2)$$

Em que:

P = potência do vento [W];

\dot{E} = fluxo de energia [J/s];

E = energia cinética [J];

t = tempo [s];

\dot{m} = fluxo de massa de ar [kg/s];

O fluxo de massa através de uma área circular é:

$$\dot{m} = \rho v A \quad (3)$$

Em que:

ρ = massa específica do ar [kg/m^3];

A = área da seção transversal [m^2].

Substituindo (3) em (1), é possível obter a expressão da potência disponível na massa de ar que passa pela seção A transversal ao fluxo de ar:

$$P_{vento} = \frac{1}{2} \rho A v^3 \quad (4)$$

2.2 Turbinas eólicas

Turbinas eólicas são máquinas capazes de converter a energia cinética do vento em energia mecânica através de um rotor e, posteriormente, essa energia é convertida em eletricidade por meio de um gerador elétrico.

2.2.1 Coeficiente de Performance e Velocidade Específica

A Equação 4 descreve a potência presente no vento, contudo, nem toda potência eólica incidente em uma turbina é convertida em energia mecânica. O coeficiente de performance C_p de uma turbina eólica estabelece a relação entre a potência extraída pela turbina $P_{turbina}$ e a potência do vento P_{vento} .

$$C_p = \frac{P_{turbina}}{P_{vento}} \quad (5)$$

Portanto,

$$P_{turbina} = \frac{1}{2} C_p \rho A v^3 \quad (6)$$

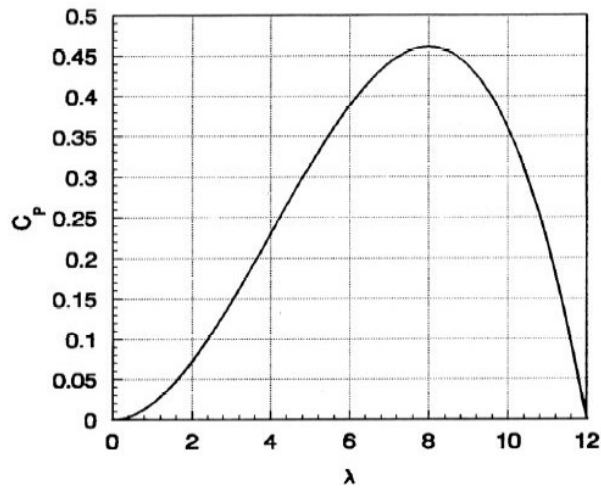
O coeficiente C_p é uma função não linear, cujo valor máximo é denominado coeficiente de Betz. O limite teórico de Betz é igual a 0,593, o que significa que o máximo de potência que uma turbina eólica é capaz de extrair é igual a 59,3% da potência disponível no vento. De acordo com (Anaya Lara; Jenkins, 2009), as turbinas comerciais possuem coeficientes de performance entre 25% e 45%.

Outro parâmetro característico de turbinas é a velocidade de ponta, definida como a relação entre a velocidade tangente à ponta da pá do rotor eólico e a velocidade do vento incidente na turbina (dos Santos Custodio, 2009):

$$\lambda = \frac{wR}{v} \quad (7)$$

A razão entre a velocidade da ponta, λ , e o coeficiente de potência, C_p , é usada para descrever a performance de rotores com diferentes raios. A curva $C_p \times \lambda$ é mostrada na figura 1.

Quando C_p é plotado em função da velocidade do vento v , pode-se observar que C_{pmax} é obtido para um único valor de v . No entanto, na curva $C_p \times \lambda$ tem-se que λ depende de v e w (velocidade de rotação). Portanto, em $C_p \times \lambda$ pode-se observar que C_{pmax} ocorre para um dado (único) valor de λ que, por sua vez, depende de v e w . Assim, o controle de w para um dado v pode fazer a turbina operar em C_{pmax} . Daí surge a tecnologia de aerogeradores de velocidade variável.

Figura 1 – Curva $C_p \times v$ Fonte: [Twente \(2020\)](#)

2.2.2 Curva de Potência

A variação da potência entregue por uma turbina é descrita por meio de sua curva de potência, como ilustrada na Figura 2. A curva é dividida em 2 regiões, uma de otimização e outra de limitação. A região de otimização está compreendida entre as velocidades de partida (*cut-in*) e a velocidade nominal da turbina. Nessa região a potência do gerador varia com o cubo da velocidade. A região de limitação está compreendida entre a velocidade nominal e a velocidade de corte ou parada (*cut-off*). Nessa região, a potência é mantida constante por ação de controle da extração de potência do vento.

Para uma velocidade média anual do vento v_m de $8m/s$, valores típicos de velocidade serão aproximadamente ([Anaya Lara; Jenkins, 2009](#)):

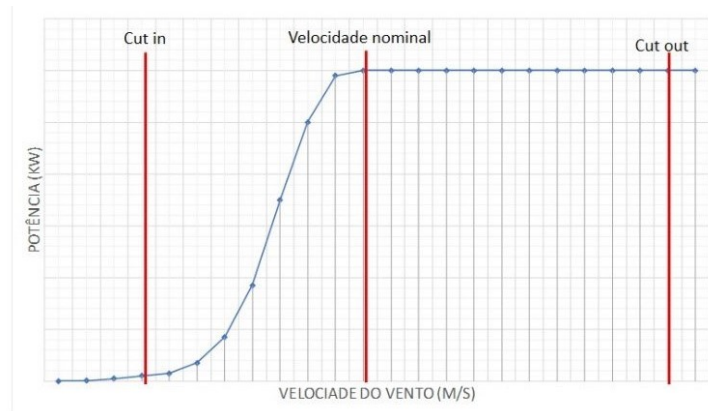
- Velocidade de partida: $5m/s = 0,6v_m$
- Velocidade nominal: $12m/s - 14m/s = 1,5v_m - 1,75v_m$
- Velocidade de corte: $25m/s = 3v_m$

2.2.3 Tipos de Turbinas Eólicas

Comercialmente existem dois tipos principais de turbinas eólicas, quanto à posição do eixo de rotação do rotor em relação ao plano terra: turbinas de eixo vertical e turbinas de eixo horizontal.

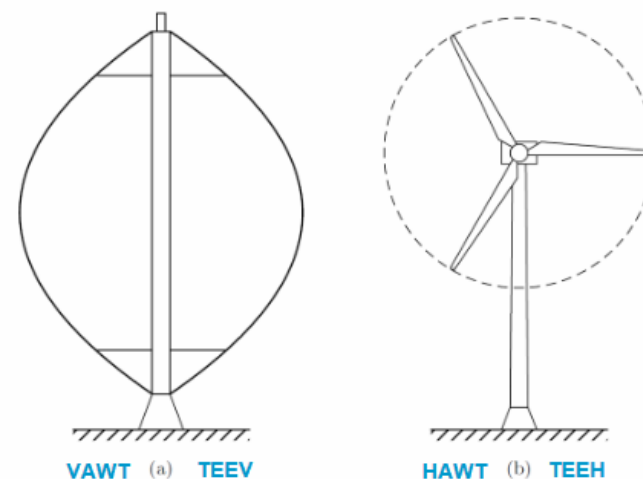
Embora as turbinas com rotores de eixo vertical possuam a vantagem de não necessitarem de mecanismos de controle de direção da nacela em função da direção do vento, existem pouquíssimas máquinas desse tipo em operação, pois as mesmas não apresentam um desempenho tão bom para a conversão eólio-elétrica. As turbinas de eixo horizontal com três pás são as mais frequentes em fazendas eólicas ([EIA, 2019](#)).

Figura 2 – Curva de Potência



Fonte: [Silva e Abreu-Harbich \(2017\)](#)

Figura 3 – Tipos de turbinas: a) eixo vertical b) eixo horizontal



Fonte: [Alé \(2012\)](#)

2.3 Partes de um Aerogerador

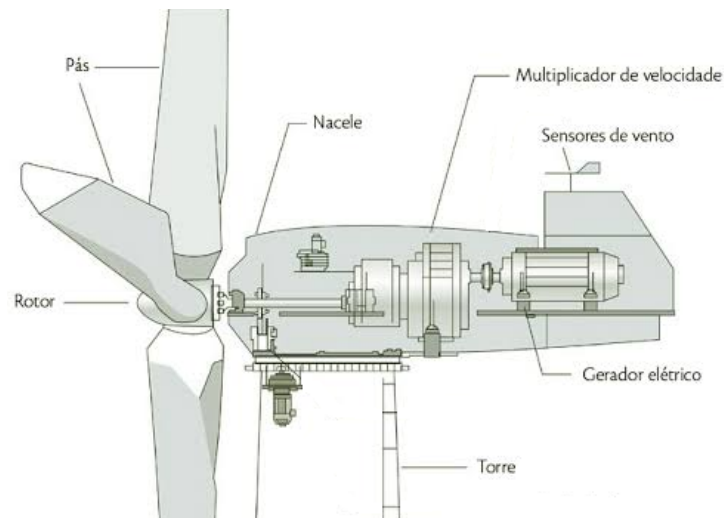
Os aerogeradores são compostos por diversos componentes, tais quais: fundação, torre, pás, eixo de transmissão, caixa multiplicadora de velocidade, gerador e sistemas de controle mecânico e eletrônico. A Figura 4 apresenta as principais partes de um aerogerador.

2.3.1 Caixa Multiplicadora

A caixa multiplicadora (*gearbox*) é utilizada para transmitir a energia mecânica entre dois pontos, adaptando as baixas velocidades com que opera o rotor da turbina às elevadas velocidades com que opera o rotor do gerador.

Os rotores em turbinas de três pás operam com velocidades entre 6 e 20 rpm, ao passo que um gerador com frequência 60 Hz opera com 1800 ou 1200 rpm, a depender do

Figura 4 – Partes de um aerogerador



Fonte: [Windbox \(2020\)](#)

número de pólos. A razão de velocidade é dada em (8):

$$r_{gb} = \frac{N_1}{N_2} \quad (8)$$

Em que:

r_{gb} = razão de velocidade da caixa de engrenagens;

N_1 = velocidade do rotor da turbina;

N_2 = velocidade do rotor do gerador;

Normalmente as turbinas eólicas operam com caixas multiplicadoras de mais de um estágio, a fim de que a relação de velocidades adequada seja mais facilmente alcançada e apresentam eficiência entre 95% e 98% ([Anaya Lara; Jenkins, 2009](#)).

Convém destacar que a caixa multiplicadora é um componente mecânico rotativo, pesado, que possui valor financeiro equivalente a cerca de 10% do custo total de uma turbina. Além disso, o equipamento é susceptível a falhas, por isso alguns fabricantes de turbinas optam por sua eliminação, aumentando o número de polos dos geradores síncronos.

Em geradores assíncronos sempre são utilizadas caixa de engrenagens, uma vez que o aumento de polos implica em maior entreferro e maior relutância, sendo necessário maior fluxo de magnetização da máquina.

2.3.2 Gerador

Os geradores utilizados na geração eólica são classificados pela capacidade de controle de velocidade, havendo quatro tipos de tecnologias predominantes:

- Geradores assíncronos de velocidade fixa;
- Geradores assíncronos com controle de escorregamento (velocidade semi variável));

- Geradores assíncronos duplamente alimentado (velocidade variável);
- Geradores síncronos (velocidade variável).

Ao estudar o comportamento do vento, indaga-se qual o tipo de gerador mais adequado para ser empregado. As máquinas síncronas são predominantemente usadas na conversão de eletricidade em usinas convencionais devido ao suporte oferecido à rede elétrica em termos de controle de frequência e magnitude de tensão, (dos Santos Custodio, 2009). Porém, devido à característica intrínseca da máquina de operar em velocidade síncrona, seu uso na conversão eólio-elétrica é complexo dada a natureza intermitente da velocidade do vento, o que dificulta seu uso para operar em velocidade síncrona. O uso de máquinas assíncronas e síncronas torna-se mais eficiente quando acionadas por conversores eletrônicos de potência.

2.3.2.1 Gerador assíncrono de velocidade fixa

Máquinas assíncronas são também conhecidas como máquinas de indução e são empregadas na maior parte dos geradores eólicos, devido a sua construção simples, custo relativamente baixo e fácil conexão e desconexão à rede. Contudo, geradores de indução normalmente operam com um baixo fator de potência, sendo necessária a adição de banco de capacitores às máquinas ou nas proximidades do ponto de conexão à rede.

O tipo mais comum de máquina assíncrona é a máquina de indução do tipo gaiola de esquilo, e o processo de funcionamento descrito pelos autores (MANWELL; MCGOWAN; ROGERS, 2014), consiste em :

- O estator possui enrolamentos dispostos de modo que as correntes com fases deslocadas produzam um campo magnético girante;
- O campo gira a uma velocidade síncrona;
- O rotor gira a uma velocidade ligeiramente diferente da velocidade síncrona;
- O campo magnético rotativo induz correntes e, portanto, um campo magnético no rotor devido à diferença de velocidades;
- A interação do campo induzido do rotor e do campo do estator causa tensão nos terminais do gerador.

Um parâmetro importante é o escorregamento s , que é a razão da diferença entre a velocidade síncrona n_s e a velocidade de operação do rotor n . Quando o escorregamento é positivo, a máquina opera como motor e quando negativo, como um gerador. Conforme descrito em (MANWELL; MCGOWAN; ROGERS, 2014), os valores típicos de escorregamento nas condições nominais são da ordem de 2%.

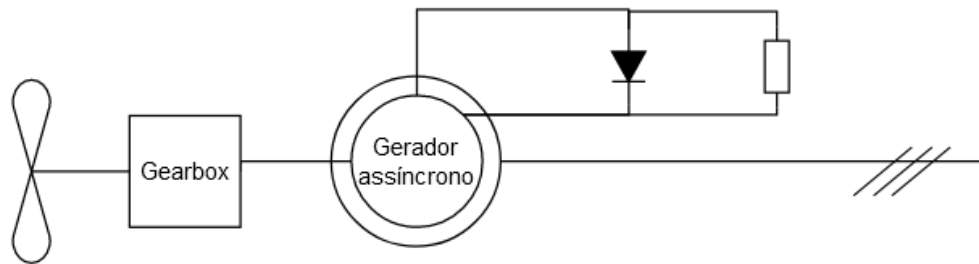
$$s = \frac{n_s - n}{n_s} \quad (9)$$

2.3.2.2 Gerador assíncrono com controle de escorregamento

A Figura 5 mostra de forma simplificada um gerador de indução do tipo rotor bobinado, ao qual pode ser acrescentada ao enrolamento do rotor uma resistência variável que permite

maior faixa de velocidade acima da velocidade síncrona (até $1,1 \times n_s$), ou seja, uma variação no escorregamento de +10%, sendo assim, são considerados por alguns autores como geradores de velocidade semi-variável.

Figura 5 – Gerador de indução com controle de escorregamento

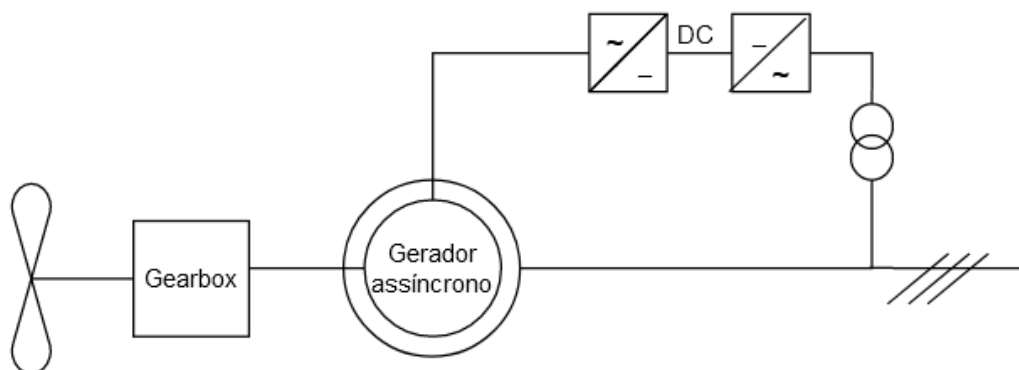


Fonte: Própria.

2.3.2.3 Gerador assíncrono duplamente alimentado

Geradores assíncronos duplamente alimentados, conhecidos em português pela sigla GIDA (Gerador de Indução de Dupla Alimentação) e no inglês por DFIG (*Doubly-Fed Induction Generator*), possuem dois conjuntos de enrolamentos com velocidades síncronas distintas. O controle da frequência rotórica é feito por meio de um conversor de frequência situado entre o rotor e o estator, permitindo a aplicação de uma tensão aos enrolamentos do rotor bobinado. A Figura 6 mostra de forma simplificada um gerador do tipo DFIG.

Figura 6 – Gerador de indução duplamente alimentado



Fonte: Própria.

2.3.2.4 Gerador síncrono

O gerador síncrono é assim nomeado pois a frequência da corrente elétrica gerada está sincronizada com a frequência de rotação do rotor. Neste tipo de gerador, normalmente a máquina primária apresenta uma rotação constante, levando a uma frequência fixa. O princípio básico de funcionamento de um gerador síncrono consiste em:

- Aplica-se uma corrente cc ao enrolamento de campo localizado no rotor, produzindo um campo magnético cc;
- O rotor é acionado por uma máquina primária, produzindo um campo magnético girante no entreferro da máquina;
- O campo girante induz um conjunto de tensões trifásicas nos enrolamentos do estator.

A relação entre a velocidade do rotor e a frequência das tensões induzidas é dada pela Equação 10:

$$n = \frac{120 \cdot f}{p} \quad (10)$$

Em que:

n = velocidade do rotor;

f = frequência elétrica;

p = número de pares de pólos.

A tensão eficaz induzida por fase é dada pela Equação 11.

$$E_{rms} = 4,44 \cdot f \cdot N \cdot \phi_{max} \cdot k_w \quad (11)$$

Em que:

N = número de espiras por fase;

f = frequência elétrica;

ϕ_{max} = fluxo máximo por pólo;

k_w = fator de enrolamento.

Na geração eólica, a turbina (máquina primária) tem sua velocidade variável em função da velocidade do vento, sendo necessários ajustes para a conexão dos aerogeradores à rede.

Quando o aerogerador é conectado diretamente à rede elétrica, a turbina deverá operar à velocidade constante, de forma que a frequência gerada se ajuste à rede. De acordo com (dos Santos Custodio, 2009), mesmo em turbinas de velocidade constante, sempre haverá uma pequena variação da rotação em função das mudanças de velocidade do vento, resultando em variações na frequência gerada, sendo recomendado o uso desta tecnologia apenas aerogeradores

de pequeno porte em redes elétricas robustas, que levem a frequência do gerador para o mesmo valor da frequência da rede.

No caso da conexão à rede ser feita por intermédio de conversores estáticos, os aerogeradores possuem velocidade variável para extração ótima da energia disponível no vento. A velocidade síncrona variável é então convertida pelo conversor eletrônico na frequência da rede.

3 Indicadores de Performance em Parques Eólicos

Os indicadores-chave de performance possuem grande influência no processo de tomada de decisões estratégicas durante a etapa de operação e manutenção de usinas eólicas.

De acordo com (GONZALEZ et al., 2017), para que um indicador seja eficaz, isto é, reflita o estado dos ativos, faz-se necessário que este seja específico, mensurável, comparável, rastreável no tempo e padronizado. Neste capítulo serão apresentadas as principais métricas utilizadas na avaliação de desempenho de parques eólicos.

3.1 Tempo médio para reparo

O tempo médio para reparo é representado pela sigla em inglês, bastante conhecida na área de confiabilidade, MTTR (Mean Time To Repair). MTTR representa o tempo médio decorrido entre a falha do componente à sua devolução ao serviço (CADWALLADER, 2012). Este indicador avalia o desempenho do serviço de manutenção, sendo diretamente afetado pela agilidade de reposição de peças em estoque.

Existem duas estratégias de manutenção normalmente adotadas em parques eólicos: intervenções frequentes e curtas e intervenções longas em períodos de ventos menos intensos, sendo essa uma estratégia que afeta diretamente na produção final de energia.

O cálculo do *MTTR* é apresentado na Equação 12:

$$MTTR = \frac{\sum(\text{tempo de indisponibilidade})_n}{\text{qtd. de falhas}} \quad (12)$$

O MTTR é o somatório dos tempos para reparo (Tempo Total de Manutenção Corretiva) dividido pelo número total de ações de manutenção corretiva durante um determinado período de tempo.

É comum na literatura o indicador ser descrito em termos da taxa de reparos (μ):

$$MTTR = \frac{1}{\mu} \quad (13)$$

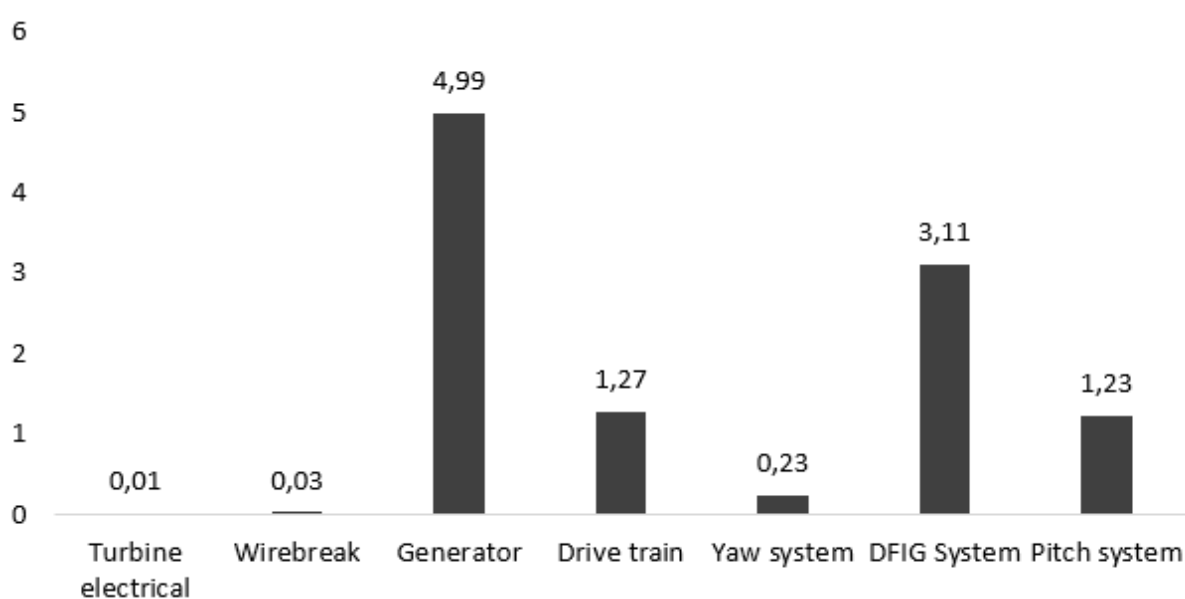
A tabela 1 e a Figura 8 apresentam valores reais do tempo médio de reparo dos componentes de aerogeradores de um fabricante indiano, operando em um parque eólico situado no litoral do Ceará durante o período de um ano.

Pode-se observar que, neste caso, as falhas devido ao sistema de Pitch possuem o maior tempo de parada, contudo, nota-se que são eventos com alta ocorrência anual, implicando em um indicador de MTTR relativamente baixo, posto que a equipe técnica de manutenção possui habilidade no procedimento necessário para a correção. Por outro lado, falhas menos recorrentes, como a do gerador, apresenta elevado tempo médio de reparo, por ser um evento de baixa recorrência em um equipamento mais complexo e que por vezes demanda a compra de peças não disponíveis em estoque.

Tabela 1 – MTTR e MTBF em componentes de um parque eólico

Componente	Tempo parado ao ano (h)	Qtd. falhas	MTTR (h)	MTBF (h)
DFIG System	12,43	4	3,11	2.877,99
Drive train	5,07	4	1,27	2.877,99
Generator	4,99	1	4,99	11.511,94
Pitch system	80,04	65	1,23	177,11
Turbine electrical	0,01	1	0,01	11.511,94
Wirebreak	0,03	1	0,03	11.511,94
Yaw system	6,79	29	0,23	396,96

Figura 7 – MTTR (h)



Fonte: Gráfico gerado a partir de dados do SCADA

3.2 Tempo médio entre falhas

O tempo médio entre falhas, cuja sigla em inglês é MTBF (*mean time between failures*) é um indicador capaz de mensurar o nível de confiabilidade de um componente ou sistema, isto é, a capacidade do sistema desempenhar uma função requerida sob condições específicas, durante um dado intervalo de tempo (ABNT, 1994). O cálculo do MTBF é apresentado na Equação 14.

$$MTBF = \frac{\sum(\text{duracao da operao})_h}{\text{qtd. de falhas}} \quad (14)$$

É comum na literatura o indicador ser descrito em termos da taxa de falhas (λ):

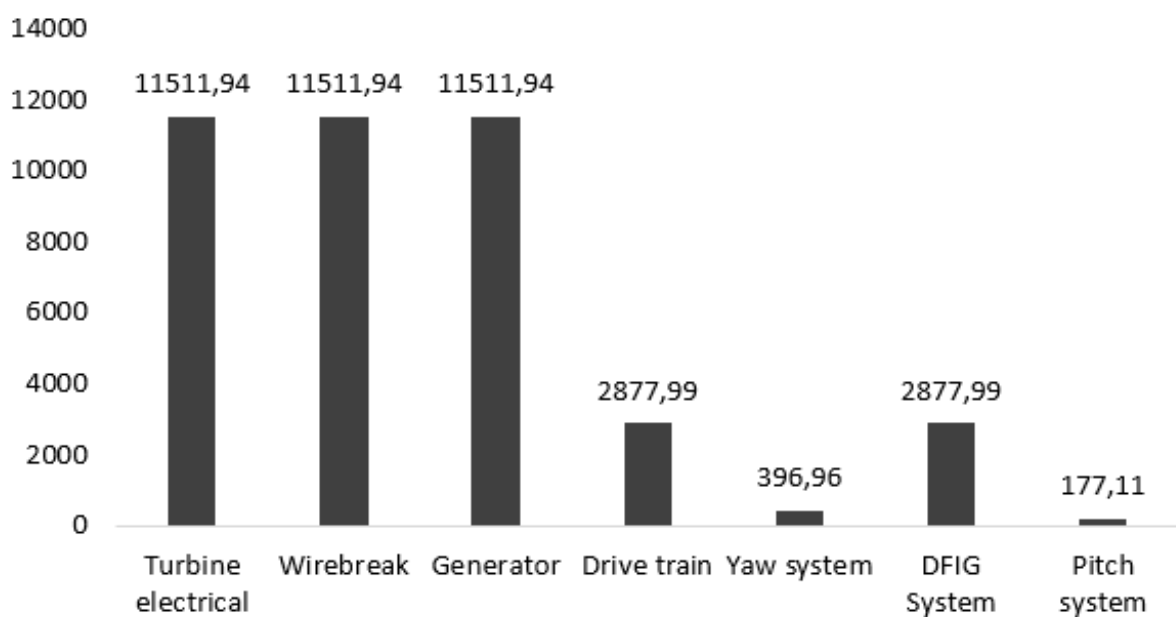
$$MTBF = \frac{1}{\lambda} \quad (15)$$

De posse do valor da taxa de falhas (λ) ou do MTBF, pode-se calcular a confiabilidade $R(t)$ do sistema por meio da Equação 16:

$$R(t) = e^{-\lambda \cdot t} \quad (16)$$

Analisando-se a tabela 1, percebe-se que as falhas devido ao sistema de Pitch e de Yaw são as mais frequentes, implicando em uma maior previsibilidade de ocorrência. Dessa forma, a manutenção pode atuar não só de forma corretiva, quando a falha ocorre, mas também de forma preditiva, a fim de mitigar os fatores de influência das mesmas.

Figura 8 – MTBF (h)



Fonte: Gráfico gerado a partir de dados do SCADA

3.3 Disponibilidade

Entende-se por disponibilidade a capacidade de um item apresentar condições de executar determinada função em um dado instante ou intervalo de tempo, supondo que os recursos externos necessários estejam assegurados (ABNT, 1994). Este indicador é o mais conhecido na avaliação de performance em parques eólicos e pode ser calculado utilizando-se a Equação 17:

$$D_{efetiva} = \frac{MDT}{(MDT + MUT)} \quad (17)$$

Em que:

MDT = tempo médio de indisponibilidade (*mean downtime*);

MUP = tempo médio de disponibilidade (*mean up time*).

Há ainda a disponibilidade do fabricante, da qual exclui-se de *MDT* a parcela referente a eventos que não são de responsabilidade do fabricante. Normalmente são estabelecidos valores contratuais assinados entre o proprietário e o produtor da turbina.

Um valor de disponibilidade contratual superior a 97% é padrão para turbinas modernas, que são normalmente projetadas para uma vida útil de 25 anos (ENERGY; TECHNOLOGY, 2020). Este indicador pode, ainda, ser expresso em termos do tempo médio entre falhas e do tempo médio para reparo (TORELL; AVELAR, 2017), de acordo com a Equação 18:

$$D_{efetiva} = \frac{MTBF}{(MTBF + MTTR)} \quad (18)$$

3.4 Fator de Capacidade

O fator de capacidade é a razão entre a energia produzida ao longo de um determinado período e a energia que teria sido produzida durante esse período se a máquina tivesse operado constantemente em sua potência nominal (MANWELL; MCGOWAN; ROGERS, 2014). O indicador pode ser calculado através da Equação 19.

$$FC = \frac{E_p}{E_n} = \frac{E_p}{P_n \times Q_t \times t} \quad (19)$$

Em que:

E_p = energia real produzida [MWh];

E_n = energia teórica produzida com turbina em operação nominal [MWh]

P_n = potência nominal da turbina [MW]

Q_t = quantidade de turbinas em operação

t = tempo de operação [h]

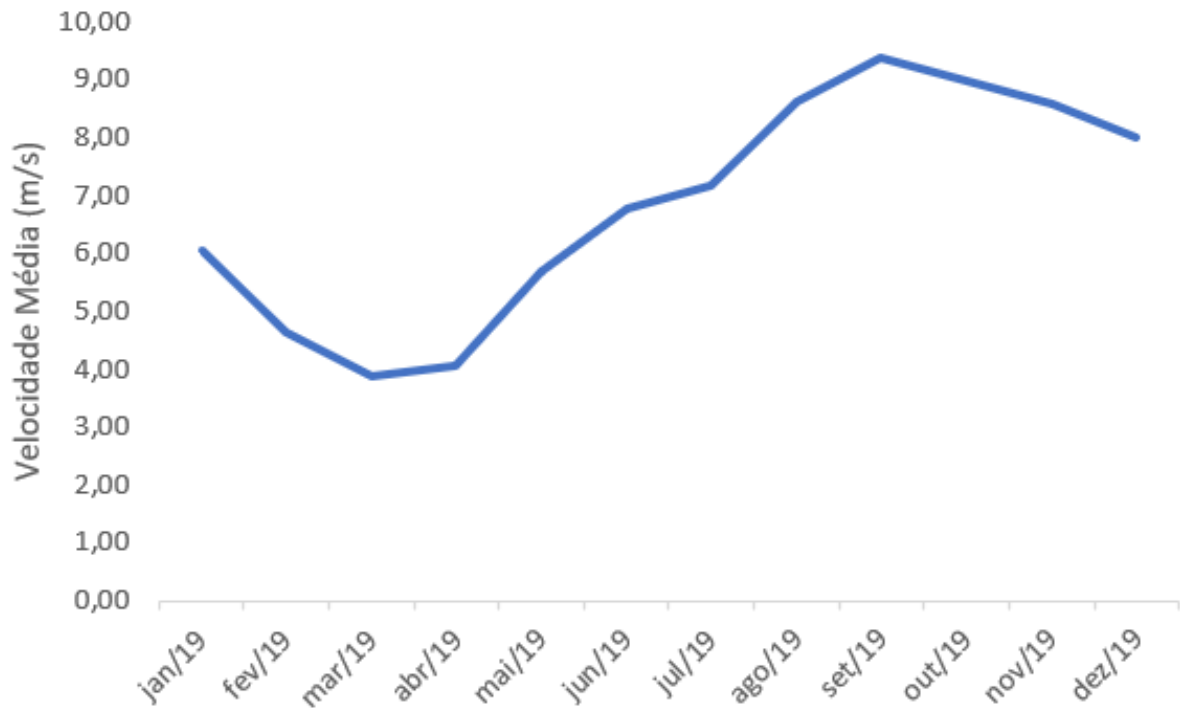
A Equação 19 pode ser reescrita em termos da velocidade do vento, utilizando-se a Equação 6, apresentada no Capítulo 1:

$$FC = \frac{\frac{1}{2}C_p\rho Av^3t}{E_n} \quad (20)$$

Sabendo-se que a velocidade dos ventos é variável ao longo do dia e ao longo dos meses do ano, como pode ser visto na Figura 9, resultante de medições em uma torre anemométrica pertencente ao mesmo parque eólico do qual foram obtidos os indicadores precedentes, o fator de capacidade, por sua vez, também apresenta intermitência, como mostra os dados reais de operação na Figura 10.

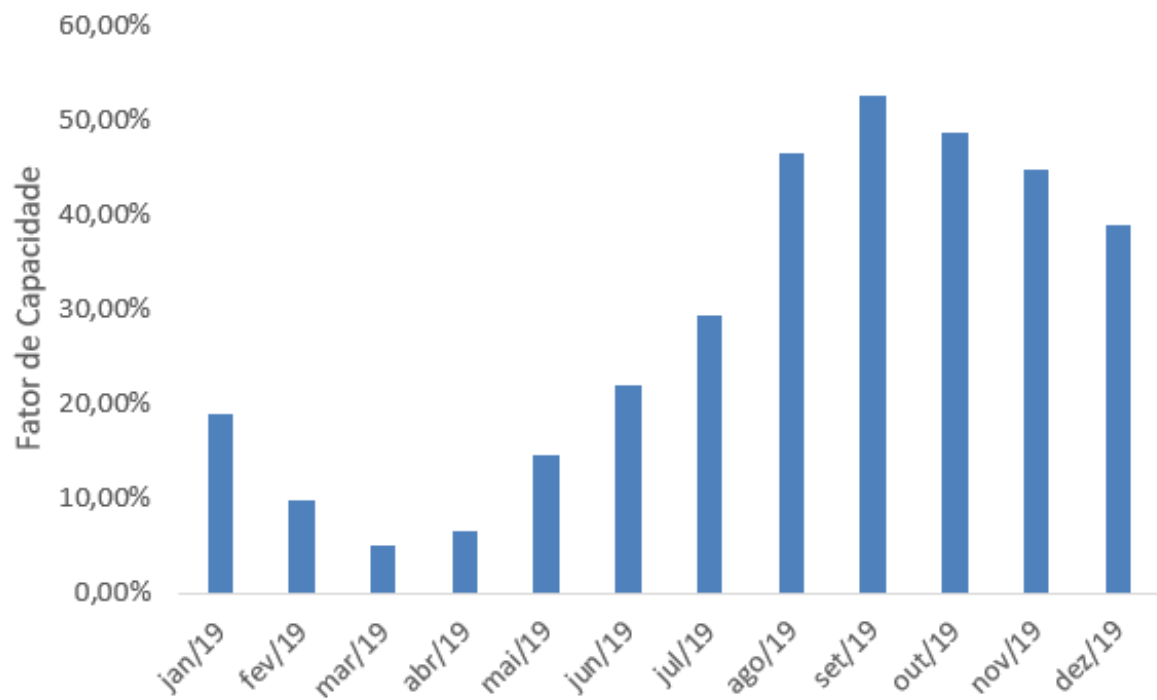
Na Tabela 2, extraída do Boletim Mensal de Geração Eólica (ONS, 2020), pode-se observar os recordes de Fator de Capacidade e Fator de Geração no Sistema Interligado Nacional (SIN) e por subsistema em base horária e em base diária (valor médio), até o mês de junho de 2020.

Figura 9 – Velocidade média dos ventos ao longo do ano de 2019



Fonte: Gráfico gerado a partir de dados da torre anemométrica

Figura 10 – Fator de Capacidade ao longo do ano de 2019



Fonte: Gráfico gerado a partir de dados do SCADA

Tabela 2 – Recordes de Fator de Carga e Fator de Geração 2020

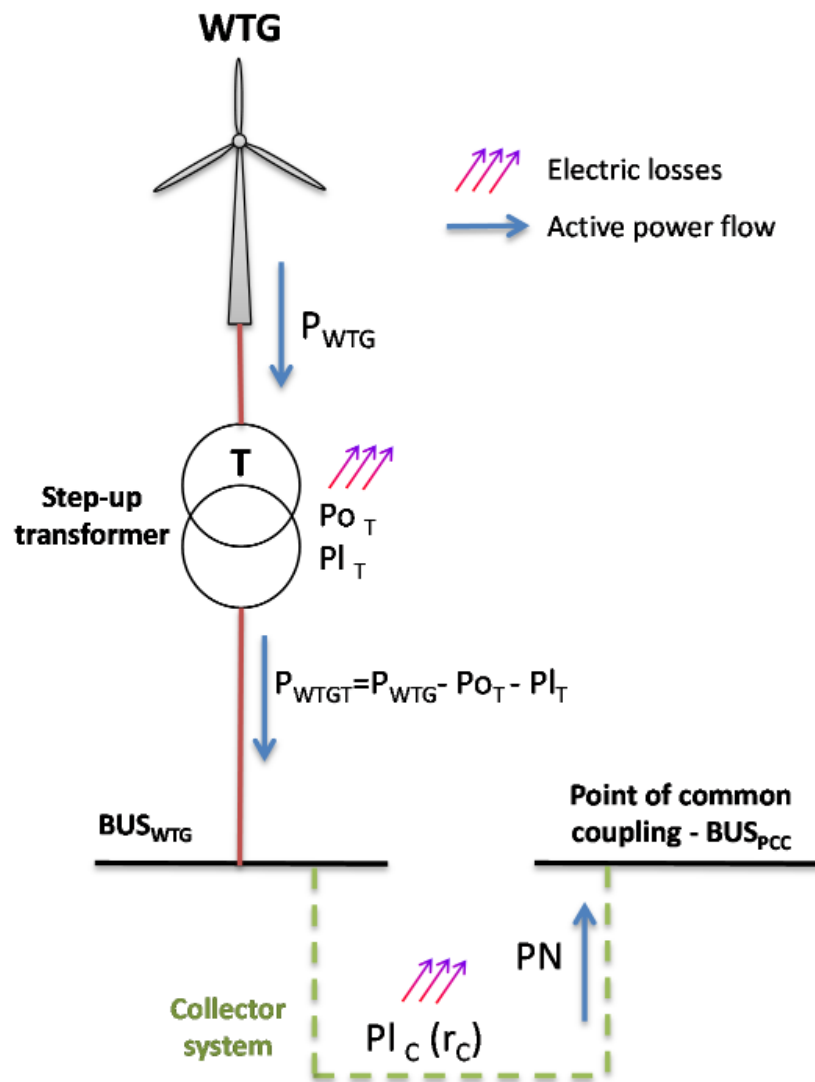
			Recorde Histórico		Recorde no Ano		Recorde no Mês	
			Data/Hora	Valor	Data/Hora	Valor	Data/Hora	Valor
SIN	Base diária	F. capacidade		75,52%		62,49%		62,49%
		Geração	06/09/2019	10.677,60	20/06/2020	9.411,07	20/06/2020	9.411,07
		% da carga		17,00%		16,92%		16,92%
	Base horária	F. capacidade		82,60%		71,97%		71,97%
		Geração	12/11/2019	11.826,92	21/06/2020	10.839,27	21/06/2020	10.839,27
		% da carga	22:00	16,36%	11:00	22,15%	11:00	22,15%
Norte	Base diária	F. capacidade		97,65%		69,27%		69,27%
		Geração	14/08/2019	321,08	29/06/2020	295,11	29/06/2020	295,11
		% da carga		5,51%		5,45%		5,45%
	Base horária	F. capacidade		99,96%		98,08%		96,92%
		Geração	25/08/2019	328,66	26/01/2020	417,83	29/06/2020	412,87
		% da carga	05:00	6,47%	17:00	8,07%	23:00	6,95%
Nordeste	Base diária	F. capacidade		76,58%		70,38%		70,38%
		Geração	13/09/2018	7.839,65	20/06/2020	8.895,61	20/06/2020	8.895,61
		% da carga		74,12%		95,18%		95,18%
	Base horária	F. capacidade		85,98%		81,29%		81,29%
		Geração	13/09/2018	8.802,45	20/06/2020	10.273,35	20/06/2020	10.273,35
		% da carga	08:00	82,34%	10:00	115,72%	10:00	115,72%
Sul	Base diária	F. capacidade		85,41%		85,41%		82,69%
		Geração	25/05/2020	1.705,09	25/05/2020	1.705,09	05/06/2020	1.650,84
		% da carga		16,90%		16,90%		14,89%
	Base horária	F. capacidade		89,99%		89,99%		88,73%
		Geração	25/05/2020	1.796,54	25/05/2020	1.796,54	05/06/2020	1.771,43
		% da carga	19:00	14,32%	19:00	14,32%	19:00	13,80%

Fonte: ONS (2020)

3.5 Perdas

As perdas de energia elétrica estão constantemente presentes na geração eólica e devem ser consideradas no cálculo da receita proveniente da venda da eletricidade produzida (COLMENAR et al., 2014). A Figura 11 ilustra de forma simplificada o fluxo de perdas em uma única turbina conectada à rede.

Figura 11 – Diagrama de perdas em uma turbina conectada à rede



Fonte: Colmenar et al. (2014)

A potência ativa bruta gerada pela turbina P_{WTG} sofre redução devido às perdas à vazio P_{O_T} e em carga P_{I_T} do transformador, resultando em P_{WTGT} . As perdas no transformador ocorrem nos enrolamentos (perdas ôhmicas) e no núcleo, por histerese e correntes parasitas (perdas magnéticas). Além das perdas internas do parque (transformadores e rede coletora), têm-se as perdas P_{I_c} devido ao sistema de transmissão, que liga a subestação coletora ao ponto de acoplamento comum (PAC), de acordo com a Figura 11. Por fim, tem-se a potência total entregue P_N , sendo esta contabilizada no faturamento da geração.

3.6 Harmônicos

Conforme descrito em (ANEEL, 2020), as distorções harmônicas são fenômenos associados a deformações nas formas de onda das tensões e correntes em relação à onda senoidal da frequência fundamental. Os principais indicadores do grau de distorção da tensão e

corrente são:

Distorção Harmônica Total de Tensão (DHT_v):

$$DHT_v = \frac{\sqrt{\sum_{h \neq 1} V_h^2}}{V_1} \quad (21)$$

Distorção Harmônica Total de Corrente (DHT_i):

$$DHT_I = \frac{\sqrt{\sum_{h \neq 1} I_h^2}}{I_1} \quad (22)$$

Em que:

h = ordem harmônica individual.

V_1, I_1 = tensão e corrente fundamentais.

De acordo com o manual de instruções para realização de estudos e medições de qualidade de energia elétrica relacionados aos acessos à rede básica (ONS, 2018), os limites individuais que devem ser atendidos por um parque eólico (ou conjunto de instalações eólicas que compartilhem de um mesmo sistema) são apresentados na Tabela 3:

Tabela 3 – Limites de Distorção Harmônica Individual

13,8 kV ≤ V < 69 kV				V ≥ 69 kV			
ÍMPARES		PARES		ÍMPARES		PARES	
ORDEM	VALOR(%)	ORDEM	VALOR(%)	ORDEM	VALOR(%)	ORDEM	VALOR(%)
3 a 25	1,5%			3 a 25	0,6%		
		todos	0,6%			todos	0,3%
≥27	0,7%			≥27	0,4%		
DTHTS95% = 3%				DTHTS95% = 1,5%			

Fonte: ONS (2018)

Na Tabela 3, DTHTS95% indica o valor de distorção harmônica total de tensão, que foi superado em apenas 5% das 1008 leituras válidas realizadas a cada 10 min no período de 7 dias consecutivos. Em situações onde os valores de distorção harmônica ultrapassem os estabelecidos na Tabela 3, o agente de geração deverá apresentar solução técnica que leve ao adequado desempenho da instalação no PAC.

4 Metodologia

Neste capítulo é descrita a metodologia utilizada no desenvolvimento do trabalho. Foi implementado o algoritmo de clusterização K-means em Python utilizando como fonte de base de dados de entrada o sistema de supervisão e aquisição de dados (SCADA) de um complexo eólico localizado no litoral do Ceará formado por cinco usinas, com um número total de 65 aerogeradores do tipo DFIG de 2,1 MW cada, como indicado na Tabela 4.

Cada aerogerador possui um conjunto de sensores de monitoramento, os quais enviam informações para o sistema responsável por agregar e armazenar tais informações, sendo possível consultas e extração de relatórios com diferentes indicadores referentes à operação das máquinas.

Tabela 4 – Potência Instalada

Parque	Qtd. Turbinas	Pot. Instalada (MW)
PQ. I	14	29,4
PQ II	13	27,3
PQ III	12	25,2
PQ IV	12	25,2
PQ V	14	29,4

Foram escolhidos os atributos geração de energia, velocidade dos ventos e disponibilidade para a clusterização das turbinas eólicas, esperando-se que o comportamento das mesmas seja condizente com o da curva de potência fornecida pelo fabricante. O algoritmo desenvolvido pode ter como dados de entrada outros indicadores, como por exemplo os descritos no Capítulo 3, contudo esse trabalho se limita a análise de performance através dos atributos escolhidos.

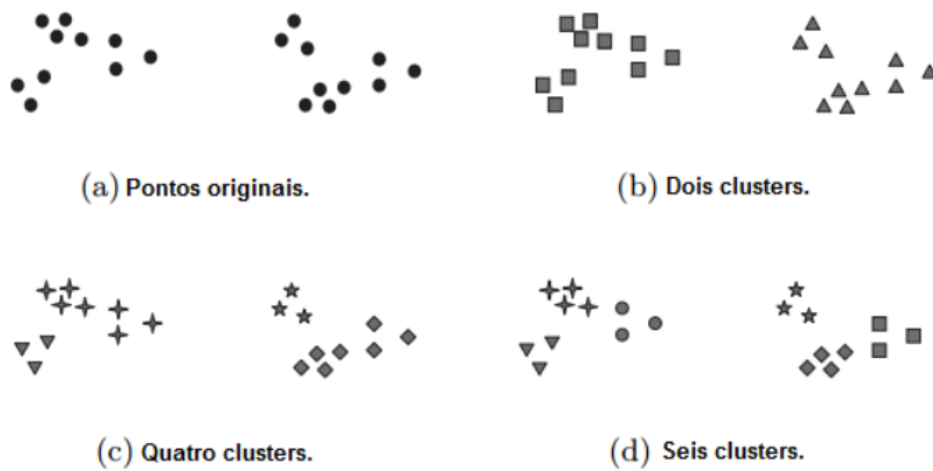
4.1 Clusterização

Em ciência de dados, clusterização diz respeito ao conjunto de técnicas utilizadas para reunir um conjunto de objetos que apresentam características semelhantes em grupos distintos, representando um papel vital na exploração de dados, na criação de previsões e na identificação de anomalias (ARORA; DEEPALI; VARSHNEY, 2015).

A idéia básica da clusterização é a de que elementos pertencentes a um mesmo grupo devem apresentar alta similaridade, contudo, devem ser deveras distintos de objetos de outros clusters. De acordo com (TAN et al., 2015), quanto maior a homogeneidade dentro de cada cluster e quanto maior a heterogeneidade entre clusters, melhor e mais distinta é a classificação.

Na figura 12, vê-se que no item (a) os dados aparecem originalmente, sem agrupamento e nos itens (b,c,d), são apresentadas as diferentes formas de classificá-los em grupos distintos, variando o número de centroides (2,4 ou 6).

Figura 12 – Diferentes agrupamentos de dados.



Fonte: Traduzido de [Tan et al. \(2015\)](#)

4.2 K-means

K-means é uma das técnicas mais comuns de agrupamento de dados ([JAIN; PRENTICE-HALL, 1988](#)). O problema de clusterização pode ser definido como um problema de otimização, onde se deseja encontrar os centros de cada conjunto, de modo que a soma das distâncias quadradas de cada objeto dentro de um agrupamento para seu centroide correspondente seja minimizada.

Em 1957, Stuart Lloyd sugeriu um modelo iterativo simples (comumente conhecido como algoritmo de Lloyd). Dado um conjunto de observações (x_1, x_2, \dots, x_n) , deve-se particionar os n dados em k conjuntos $(S = S_1, S_2, \dots, S_k)$ ($k \neq n$), minimizando a distância entre os pontos pertencentes a cada partição, conforme Equação 23.

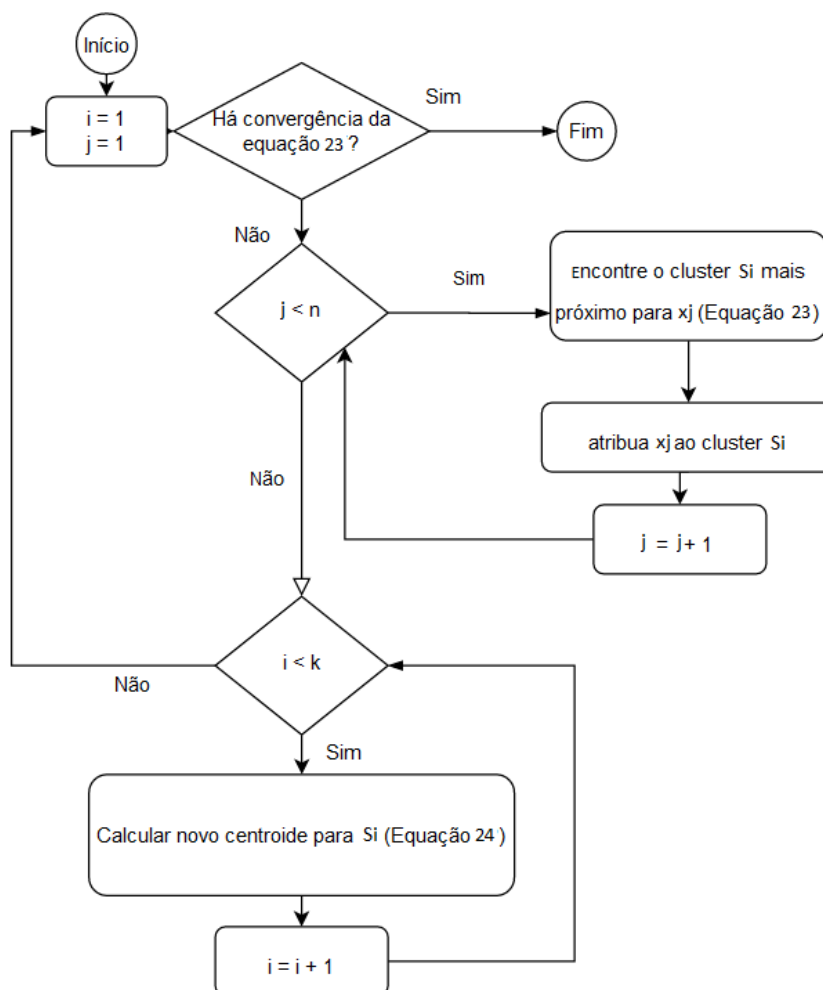
$$\operatorname{argmin} \sum_{i=1}^k \sum_{x_j \in S_i} \min(\|x_j - \mu_i\|^2). \quad (23)$$

Em que μ_i é o centroide em S_i e a expressão $\|x_j - \mu_i\|^2$ indica a distância entre os pontos de um cluster e o centroide. Para cada cluster S_i , é calculado um novo centroide μ_i , por meio do cálculo da média de todos os pontos x_j atribuídos previamente ao cluster S_i , conforme a Equação 24.

$$\mu_i = \frac{1}{|S_i|} \sum_{x_j \in S_i} x_j \quad (24)$$

Os passos necessários para a execução do algoritmo são resumidos no fluxograma da Figura 13.

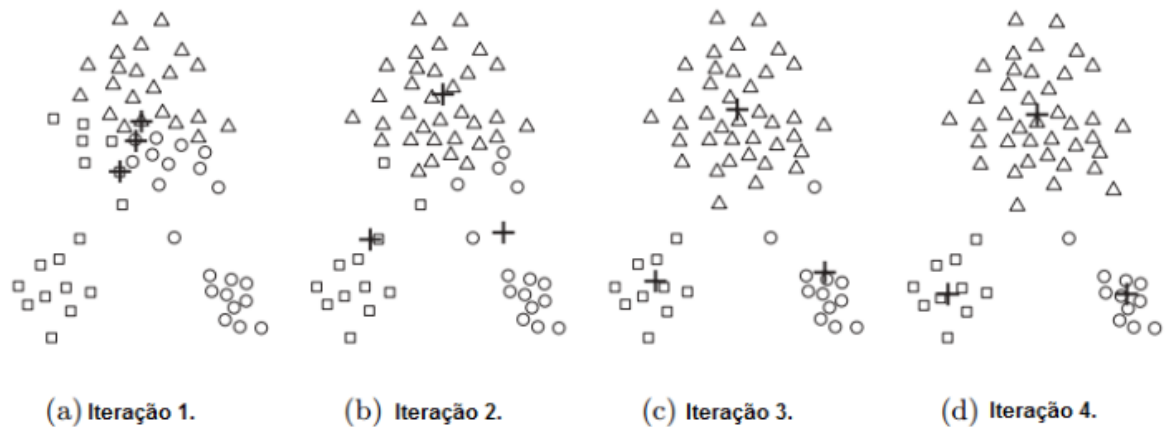
Figura 13 – Fluxograma do algoritmo k-means



Fonte: Própria

Os procedimentos são repetidos, até que haja convergência dos centroides, isto é, até que seus valores não sejam significativamente modificados após as iterações. A Figura 14 representa 4 iterações de execução do algoritmo K-means.

Figura 14 – Etapas do algoritmo k-means



Fonte: (TAN et al., 2015)

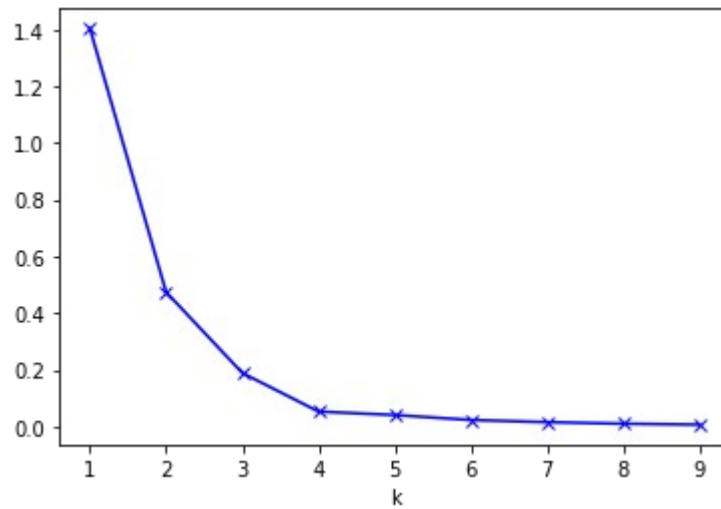
4.3 Método *Elbow*

Na escolha da quantidade de clusters, deve ser considerado que um grande valor de K pode levar ao particionamento excessivo dos dados, impossibilitando a descoberta de padrões. Por outro lado, um número muito pequeno de clusters pode gerar clusters generalistas contendo muitos dados sem semelhanças entre si.

De acordo com (MINT, 2020), uma maneira eficaz de escolher a quantidade ideal de clusters é executar o algoritmo K-Means para diferentes valores de K e calcular a soma das distâncias entre cada centróide e as demais observações contidas em um mesmo cluster. Dessa forma, busca-se encontrar um valor para K de modo que a distância intraclasse seja mínima. A variância dos clusters é calculada conforme Equação 23.

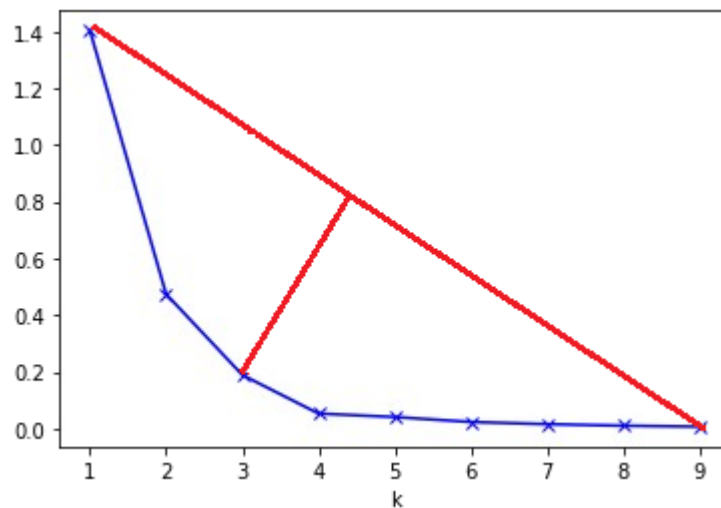
Para a curva da Figura 15, que apresenta 9 centroides e seus respectivos erros médios, traça-se uma reta entre os centroides extremos (1,1.4) e (9,0). O ponto com maior homogeneidade dentro do cluster e maior diferença entre clusters é o ponto mais distante da reta traçada, conforme representado na Figura 16.

Figura 15 – Curva método Elbow



Fonte: (MINT, 2020)

Figura 16 – Curva método Elbow - Escolha de K



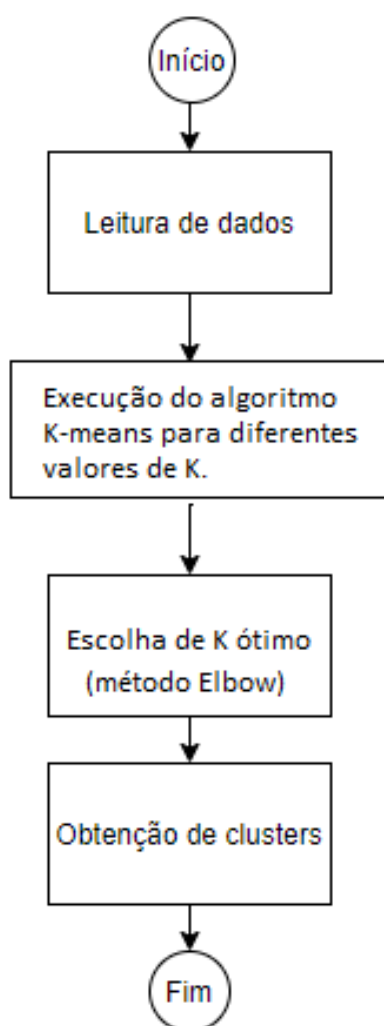
Fonte: Adaptado de (MINT, 2020)

4.4 Algoritmo Python

A implementação do algoritmo K -means foi realizada através da programação em python, utilizando-se a scikit-learn, uma biblioteca de aprendizado de máquina de código aberto que possui vários algoritmos de classificação, regressão e agrupamento de dados. Um tutorial completo da utilização dos algoritmos contidos nesta biblioteca pode ser encontrado em (SCIKIT-LEARN, 2020). O algoritmo foi executado para 10 iterações de k .

O fluxograma da Figura 17 ilustra de forma simplificada a sequência de passos do código desenvolvido. O algoritmo completo encontra-se no Apêndice B.

Figura 17 – Fluxograma código python



Fonte: Própria

4.5 Base de Dados

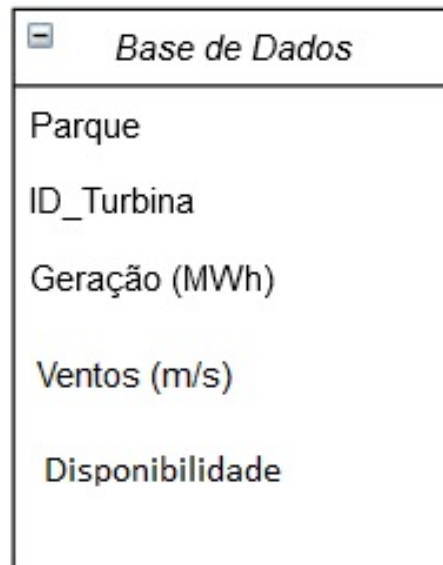
Compreender a dinâmica da produção de energia de um parque eólico é importante para a integração da energia eólico-elétrica em grande escala. Em um grande sistema dinâmico complexo, como um parque eólico, o agrupamento é uma forma de reduzir a complexidade do modelo e melhorar seu entendimento. *(Cluster analysis of wind turbines of large wind farm with diffusion distance method)

Neste trabalho foram coletados durante o mês de dezembro de 2019 dados de velocidade média dos ventos, energia gerada e disponibilidade de cada um dos 65 aerogeradores do complexo eólico. As informações presentes no SCADA têm granularidade de 10 minutos, desta forma, foi necessário um tratamento de dados para deixá-las no formato desejado. Para tanto, foram excluídas amostras com valores incoerentes, (por exemplo: velocidades negativas e valores nulos). Em seguida, calculou-se a média de cada grupo de 6 amostras de 10 minutos consecutivas

de forma a se obter um valor de geração e velocidade média do vento equivalente a cada 60 minutos (1 hora).

Por fim, as informações foram armazenadas em arquivos no formato CSV. A base de dados após ajustes resultou em 65 linhas, onde cada uma destas possui um identificador de turbina e os atributos que a caracterizam, conforme mostrado na Figura 18.

Figura 18 – Colunas da base de dados



Base de Dados
Parque
ID_Turbina
Geração (MWh)
Ventos (m/s)
Disponibilidade

Fonte: Própria

4.6 Análise de Performance

A análise de performance dos aerogeradores é realizada em seis etapas utilizando-se como entrada do algoritmo a base de dados descrita em 4.5 e a curva de potência dos aerogeradores. Os passos podem ser resumidos em:

1. O algoritmo K-means é executado para diferentes valores de K , escolhendo-se o valor ótimo ao final das execuções;
2. É feita a clusterização dos 65 aerogeradores utilizando-se a velocidade dos ventos e a energia produzida;
3. É feita a clusterização dos 65 aerogeradores utilizando-se a disponibilidade das máquinas e a energia produzida;
4. De acordo com os dados da curva de potência fornecida pelo fabricante do modelo de aerogerador utilizado, descrita em 11, é feita interpolação polinomial utilizando pontos próximos aos centroides da etapa 2, de modo a obter uma equação do terceiro grau que represente o trecho da curva;
5. Multiplica-se a disponibilidade de cada centroide da etapa 3 pelo valor de energia esperada obtida com a função descrita na etapa 4;

6. Analisa-se a performance de cada cluster com base na energia produzida e na energia esperada.

5 Análise e Discussão dos Resultados

Nesta seção serão executadas e discutidas as etapas descritas na Seção 4.6.

5.1 Obtenção do valor ótimo de K

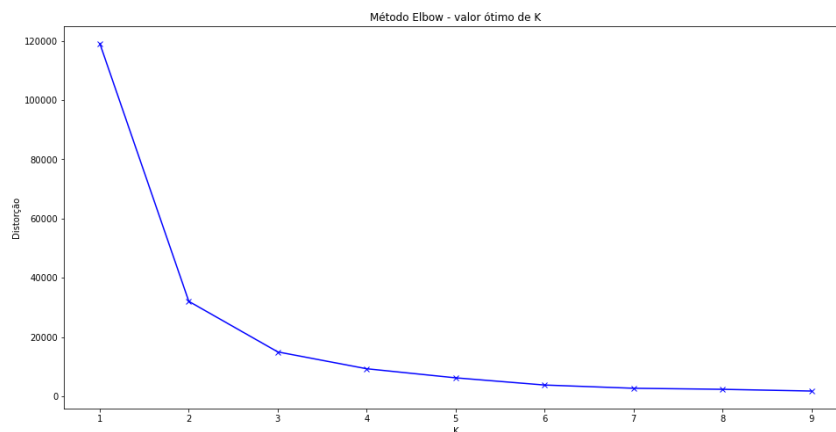
O valor ótimo de K foi obtido para 10 iterações do trecho de código descrito abaixo:

```
distortions = []
K = range(1,10)
for k in K:
    kmeanModel = KMeans(n_clusters=k)
    kmeanModel.fit(X)
    distortions.append(kmeanModel.inertia_)

plt.figure(figsize=(16,8))
plt.plot(K, distortions, 'bx-')
plt.xlabel('K')
plt.ylabel('Distorcao')
plt.title('Metodo Elbow - valor otimo de K')
plt.show()
```

De acordo com a Figura 19, nota-se que para valores de $K > 3$, não há variações significativas na distorção, desta forma, escolhe-se $K = 3$.

Figura 19 – Método Elbow - valor ótimo de K '

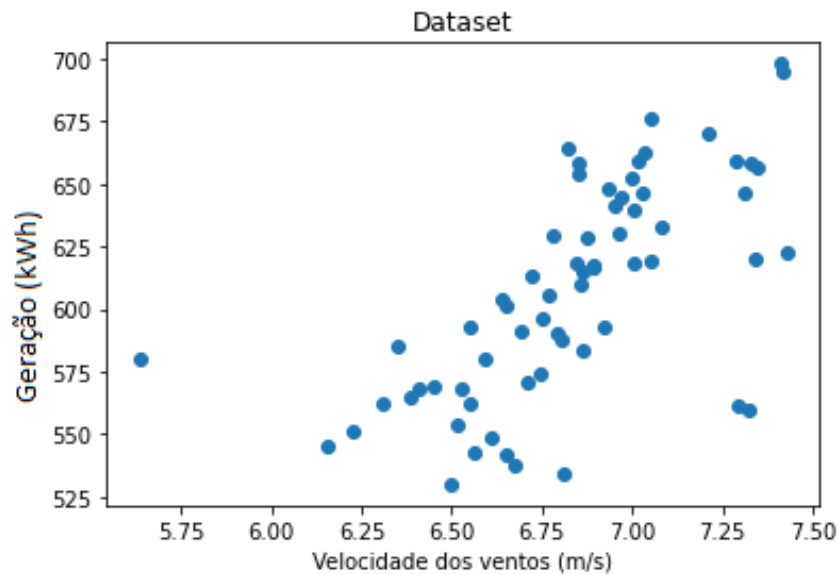


Fonte: Própria

5.2 Clusterização: Velocidade dos ventos x Energia Produzida

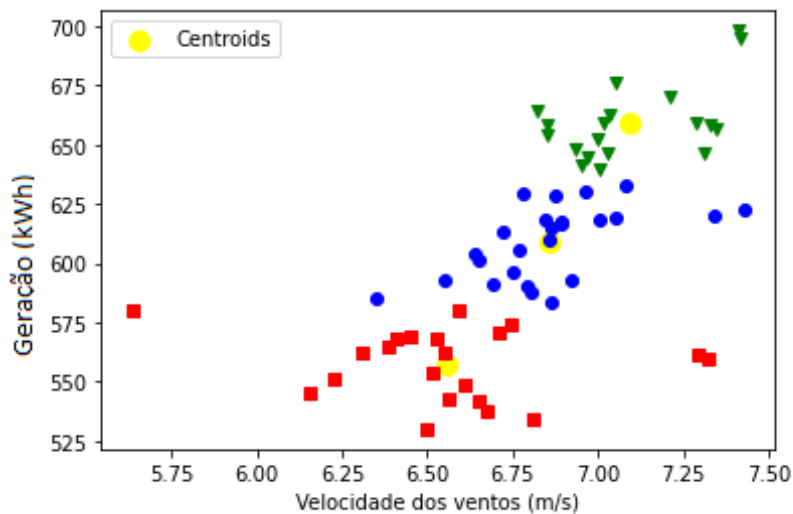
O gráfico da Figura 20, mostra as turbinas sem agrupamento, dispostas de acordo com a velocidade dos ventos e a geração de energia e a Figura 21 representa os três clusters obtidos.

Figura 20 – Dados velocidade x geração



Fonte: Própria

Figura 21 – Clusters velocidade x geração



Fonte: Própria

Tabela 5 – Centroides

Cluster	Vento (m/s)	Geração (kWh)
0	6,55	557,50
1	7,09	659,59
2	6,85	608,91

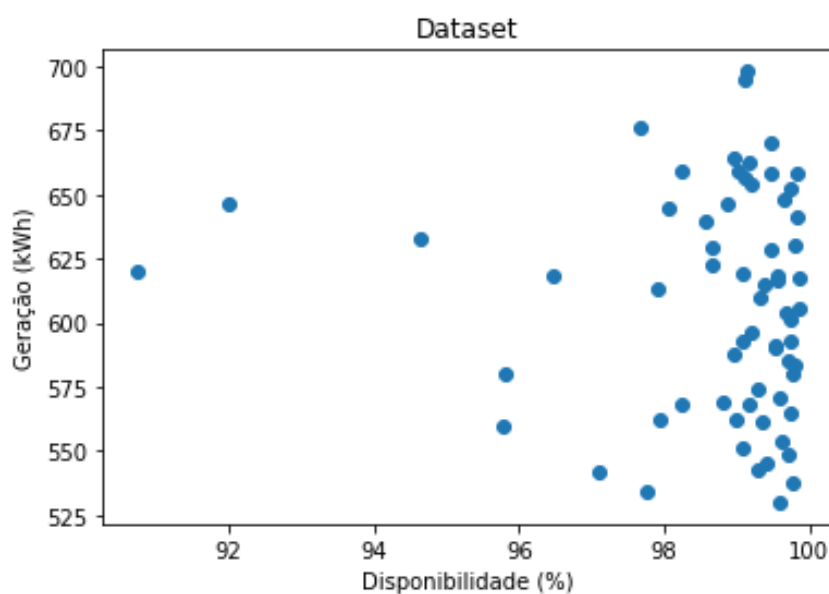
Tabela 6 – Clusters

Cluster	Turbinas	Qtd.
0	2, 3, 5, 6, 7, 8, 11, 17, 27, 28, 29, 42, 43, 44, 45, 46, 48, 50, 63, 64	20
1	14, 15, 16, 18, 19, 20, 21, 22, 23, 32, 33, 35, 36, 37, 38, 39, 51, 52, 58	19
2	0, 1, 4, 9, 10, 12, 13, 24, 25, 26, 30, 31, 34, 40, 41, 47, 53, 54, 55, 56, 57, 59, 60, 61, 62	26

5.3 Clusterização: Disponibilidade x Energia Produzida

O gráfico da Figura 22, mostra as turbinas sem agrupamento, dispostas de acordo com a disponibilidade e a geração de energia e a Figura 23 representa os três clusters obtidos. .

Figura 22 – Dados disponibilidade x geração

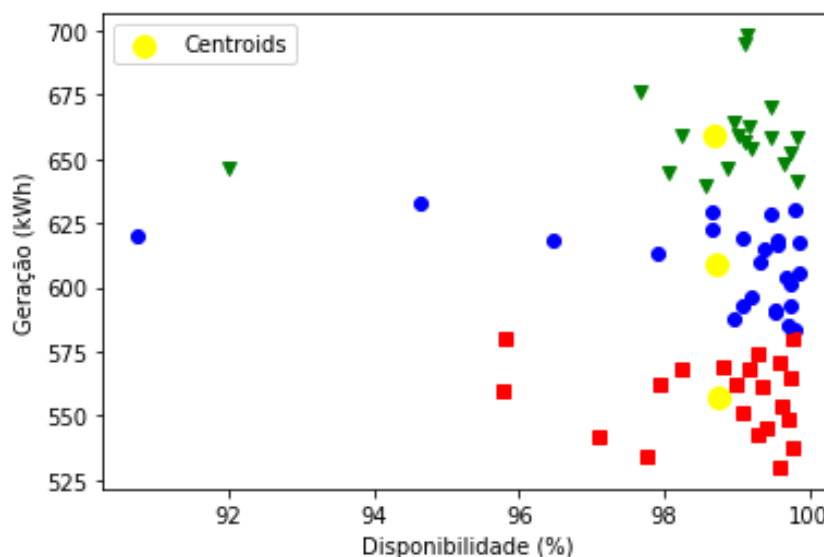


Fonte: Própria

Tabela 7 – Centroides

Cluster	Disponibilidade (%)	Geração (kWh)
0	98,74	557,50
1	98,68	659,59
2	98,71	608,91

Figura 23 – Clusters disponibilidade x geração



Fonte: Própria

Tabela 8 – Clusters

Cluster	Turbinas	Qtd.
0	2, 3, 5, 6, 7, 8, 11, 17, 27, 28, 29, 42, 43, 44, 45, 46, 48, 50, 63, 64	20
1	14, 15, 16, 18, 19, 20, 21, 22, 23, 32, 33, 35, 36, 37, 38, 39, 51, 52, 58	19
2	0, 1, 4, 9, 10, 12, 13, 24, 25, 26, 30, 31, 34, 40, 41, 47, 53, 54, 55, 56, 57, 59, 60, 61, 62	26

5.4 Análise de Performance

Para analisar a performance de cada grupo de aerogeradores foram comparados os valores da Tabela 5 com os dados da curva de potência fornecidos pelo fabricante, representada na Tabela 11. Para uma melhor precisão do estudo, foi aplicada interpolação polinomial utilizando-se quatro pontos da curva de potência correspondentes à velocidade entre 6 m/s e 7,5 m/s os quais compreendem os valores obtidos na Tabela 5, de modo a escolher a equação que melhor representa esse trecho da curva. Para tanto, optou-se pelo software *Wolfram*, conforme é mostrado na Figura 24.

Simplificando o resultado obtido na Figura 24, tem-se a Equação 25 que é uma aproximação da curva de potência para valores de velocidade entre 6 m/s e 7,5 m/s. A Figura 25 destaca em vermelho o trecho da Equação 5 que representa a curva de potência dos aerogeradores.

$$P(v) = -8.66667x^3 + 207x^2 - 1322.13x + 2767.7 \quad (25)$$

A Tabela 9 apresenta a aplicação da Equação 25 aos valores de velocidade média dos

Figura 24 – Obtenção de equação com software Wolfram

Input:

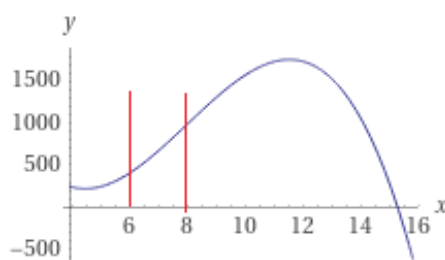
$$\text{InterpolatingPolynomial}\left[\left[\begin{array}{cc} 6 & 414.9 \\ 6.5 & 539.5 \\ 7 & 683.1 \\ 7.5 & 839.2 \end{array}\right], x\right]$$

Result:

$$(x - 7.5)((x - 6)(33.6667 - 8.66667(x - 6.5)) + 282.867) + 839.2$$

Fonte: Própria

Figura 25 – Gráfico da Equação 25 com destaque para a região utilizada para representar a curva de potência do fabricante



Fonte: Própria

clusters, contidos na Tabela 5. Na terceira coluna da Tabela 10 encontram-se os valores de geração esperada, de acordo com a Equação 25 multiplicados pelas disponibilidades da Tabela 7. Esse passo é necessário na análise, uma vez que a curva de potência do fabricante é fornecida considerando-se que o aerogerador atua com uma disponibilidade igual a 100%, desta forma, faz-se necessária a proporcionalização do nível de produção esperada pela disponibilidade do cluster. A performance é calculada pela razão entre a terceira coluna e a segunda coluna da Tabela 10.

Tabela 9 – Geração esperada utilizando Equação 25

Cluster	Velocidade (m/s)	Geração Esperada
0	6,55	553,13
1	7,09	710,49
2	6,85	638,43

A Tabela 10 apresenta o resultado final da análise de performance dos grupos de aerogeradores, mostrando que, embora o cluster 0 possua a menor velocidade média dos ventos, o desempenho do grupo de aerogeradores é superior aos demais. Pode-se constatar com esse cluster, que, conforme publicado em (FOCUS, 2012), algumas máquinas com tecnologia S9X

Tabela 10 – Análise de Performance

Cluster	Geração Cluster	Geração Esperada × Disponibilidade	Performance
0	557,50	546,16	+2%
1	659,59	701,11	-6%
2	608,91	630,19	-3%

podem operar com 102% da capacidade descrita pela curva de potência do fabricante. Em contrapartida, os clusters 1 e 3 apresentam performance abaixo do previsto pelo fabricante.

Existem muitas causas para a diferença de desempenho de turbinas eólicas pertencentes a um mesmo complexo eólico, algumas delas são: dimensionamento das máquinas para velocidades superiores aos valores prevaletentes no sítio, posicionamento das máquinas, ângulo de ataque, desgaste dos componentes, taxa de falhas durante todo o período de utilização do equipamento, entre outras. Além disso, convém ressaltar que faz-se necessária uma análise de dados históricos de pelo menos um ano para resultados mais acurados.

Tabela 11 – Curva de Potência

Velocidade do Vento (m/s)	Potência (kW)
3	0
3,5	2,6
4	52,3
4,5	122,5
5	206,4
5,5	304
6	414,9
6,5	539,5
7	683,1
7,5	839,2
8	1010
8,5	1197,1
9	1393,2
9,5	1592,4
10	1787,7
10,5	1931,2
11	2016,3
11,5	2068,3
12	2096,4
12,5	2109,2
13	2114
13,5	2116,5
14	2116,2
14,5	2116
15	2115,8
15,5	2115
16	2114,5
16,5	2114,2
17	2113,8
17,5	2112,2
18	2111,2
18,5	2110,5
19	2109,8
19,5	2107,3
20	2105,5
20,5	2104,4
21	2103,2
21,5	2102
22	2096,4
22,5	2092,8
23	2091,2
23,5	2089,7
24	2088,2
24,5	2086,6

6 Conclusão e Trabalhos Futuros

6.1 Conclusão

A operação de parques eólicos é feita em um contexto competitivo de produção de eletricidade, sendo fundamental que as empresas possuam controle de seus ativos. Desta forma, este trabalho se propôs a desenvolver uma metodologia de análise da performance de fazendas eólicas compostas por grande número de aerogeradores.

A metodologia proposta se baseia na análise de dados e aplicação da técnica de clusterização denominada *K - means*, resultando no agrupamento de aerogeradores com desempenhos semelhantes. Foram coletados dados de operação de 65 aerogeradores e aplicados em um algoritmo desenvolvido em Python, sendo a quantidade ótima de clusters K determinada através do método Elbow. A performance dos aerogeradores foi analisada com base na equação $P(v)$ obtida por interpolação polinomial do trecho da curva de potência fornecida pelo fabricante. O algoritmo de clusterização foi aplicado utilizando-se a velocidade dos ventos \times produção de energia e , em seguida, aplicando-se a disponibilidade das máquinas e produção de energia. Foram obtidos 3 grupos de aerogeradores com performances distintas.

Os resultados mostraram que, embora o cluster 0 apresente menor velocidade média do centroide, ele opera com desempenho 2% superior ao estimado pelo fabricante, fato previamente divulgado em artigo publicado pelos produtores de aerogeradores com tecnologia S9X. Os clusters 1 e 2, com maiores velocidades de ventos associadas, atuam com geração 6% e 3% inferiores, respectivamente, ao valor esperado de energia gerada, podendo ser resultado de diversos fatores, que podem ser investigados por meio dos indicadores de performance descritos no capítulo 3. Isto sugere a importância da aplicação de declusterização nas áreas de operação e manutenção como mecanismo para integrar e gerenciar de forma otimizada um grande volume de dados dos recursos energéticos.

Uma dificuldade enfrentada nesse trabalho foi o tratamento das informações obtidas no Sistema de Controle e Aquisição de Dados, pois as bases devem ser ajustadas de modo que as mesmas possam ser utilizadas com informações coerentes e sem erros de leitura, que podem comprometer o resultado da pesquisa. Convém destacar que a técnica utilizada pode ser aplicada com diversas variáveis, a depender do tipo de análise que se deseja obter. Por exemplo, pode-se agrupar os aerogeradores de acordo com os demais indicadores de performance descritos no Capítulo 3. Acredita-se que com esse trabalho foi possível difundir o conhecimento adquirido a respeito de clusterização de dados aplicada à operação e manutenção de parques eólicos.

6.2 Trabalhos Futuros

Como sugestão para trabalhos futuros, são deixados os seguintes tópicos:

- Aplicação da técnica *K-means* utilizando-se outros indicadores, tais quais: emissão de

harmônicos, mttr, mtbf e taxa de falhas.

- Comparação dos resultados utilizando-se outra técnica de *clustering*.
- Análise de causas-raiz de aerogeradores com baixa produtividade.
- Aplicação de técnicas mineração de dados na modelagem da curva de potência dos aerogeradores e no tratamento de dados tais quais descritos em (HUANG, 2019).
- Análise de performance de inversores em complexos de geração de energia solar fotovoltaica.

Referências

- A.KHOSRAVIL.MACHADO; R.O.NUNES. Time-series prediction of wind speed using machine learning algorithms: A case study osorio wind farm, brazil. **Elsevier**, 2018. Citado na página 1.
- ALÉ, J. A. V. **Small wind turbines workshop**. 2012. Disponível em: <<https://app.eventize.com.br/upload/000388/files>>. Acesso em: 21 de julho de 2020. Citado na página 6.
- Anaya Lara, O.; Jenkins, N. **Wind Energy Generation Modelling and Control**. 1. ed. UK: Wiley, 2009. Citado 3 vezes nas páginas 4, 5 e 7.
- ANEEL. **PRODIST - Módulo 8: Qualidade de Energia Elétrica**. Brasil, 2020. Citado na página 18.
- ARORA, P.; DEEPALI, D.; VARSHNEY, S. Analysis of k-means and k-medoids algorithm for big data. **International Conference on Information Security Privacy**, 2015. Citado na página 20.
- ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **ABNT NBR 5462**: Confiabilidade e manutenibilidade. Rio de Janeiro, 1994. Citado 2 vezes nas páginas 13 e 14.
- CADWALLADER, L. C. Key performance indicators for wind farm operation and maintenance. **INL - U.S. Department of Energy National Laboratory**, 2012. Citado na página 12.
- COLMENAR, A. et al. Simplified analysis of the electric power losses for on-shore wind farms considering weibull distribution parameters. **Energies - ISSN 1996-1073**, 2014. Citado 2 vezes nas páginas 17 e 18.
- dos Santos Custodio, R. **Energia eólica para produção de energia elétrica**. 1. ed. Brasil: Eletrobrás, 2009. Citado 3 vezes nas páginas 4, 8 e 10.
- EIA. **Wind explained Types of wind turbines**. 2019. Disponível em: <<https://www.eia.gov/energyexplained/wind/types-of-wind-turbines.php>>. Acesso em: 05 de maio de 2020. Citado na página 5.
- ENERGY, T. A. F. R.; TECHNOLOGY, C. **Wind**. 2020. Disponível em: <<https://www.r-e-a.net/technologies/wind/>>. Acesso em: 03 de agosto de 2020. Citado na página 15.
- FOCUS, R. E. on. **Suzlon's S95 wind turbine delivers 102%**. 2012. Disponível em: <<http://www.renewableenergyfocus.com/view/24159/suzlons-s95-wind-turbine-delivers-102/>>. Acesso em: 10 de outubro de 2020. Citado na página 32.
- GONZALEZ, E. et al. Key performance indicators for wind farm operation and maintenance. **Elsevier**, 2017. Citado 2 vezes nas páginas 1 e 12.
- HUANG, B. **Wind Turbine Performance Assessment Modeling Using Machine Learning Method for Condition Based Maintenance**. Dezembro 2018. 77 f. Dissertação (DEGREE PROJECT IN ELECTRICAL ENGINEERING) — KTH ROYAL INSTITUTE OF TECHNOLOGY SCHOOL OF ELECTRICAL ENGINEERING AND COMPUTER SCIENCE, Suécia, 2019. Citado na página 36.

JAIN, A. K.; PRENTICE-HALL, R. C. D. Introduction to data mining. ISBN 0-13-022278-X, 1988. Citado na página 21.

MA, Y.; RUNOLFSSON, T.; JIANG, J. Cluster analysis of wind turbines of large wind farm with diffusion distance method. **The Institution of Engineering and Technology**, 2011. Citado na página 1.

MANWELL, J. F.; MCGOWAN, J. G.; ROGERS, A. L. **Wind Energy Explained: Theory, Design and Application**. 2. ed. USA: Wiley, 2014. Citado 2 vezes nas páginas 8 e 15.

MINT, M. **l'algorithmme K-Means**. 2020. Disponível em: <<https://mrmint.fr/algorithmme-k-means>>. Acesso em: 01 de outubro de 2020. Citado 2 vezes nas páginas 23 e 24.

ONS. **INSTRUÇÕES PARA REALIZAÇÃO DE ESTUDOS E MEDIÇÕES DE QEE RELACIONADOS AOS ACESSOS À REDE BÁSICA OU NOS BARRAMENTOS DE FRONTEIRA COM A REDE BÁSICA PARA PARQUES EÓLICOS, SOLARES, CONSUMIDORES LIVRES E DISTRIBUIDORAS**. Brasil, 2018. Citado na página 19.

ONS. **Boletim Mensal de Geração Eólica**. Brasil, 2020. Citado 2 vezes nas páginas 15 e 17.

ROMERO, M.; GALLEGO, L.; PAVAS, A. Machine learning methods for wind turbine condition monitoring: A review. **Scielo**, 2011. Citado na página 1.

SCIKIT-LEARN. 2020. Disponível em: <<https://scikit-learn.org/stable/index.html>>. Acesso em: 26 de agosto de 2020. Citado na página 24.

SILVA, L. M. da; ABREU-HARBICH, L. V. de. Metodologia simplificada para avaliação do potencial de energia eólica em centros urbanos. **Periódicos Unicamp**, 2017. Citado na página 6.

STETCO, A. et al. Machine learning methods for wind turbine condition monitoring: A review. **Elsevier**, 2019. Citado na página 1.

TAN, P.-N. et al. **International Conference on Information Security Privacy**, 2015. Citado 3 vezes nas páginas 20, 21 e 23.

TORELL, W.; AVELAR, V. Mean time between failure: Explanation and standards. **Schneider Electric**, 2017. Citado na página 15.

TWENTE, U. of. **WIND ENERGY ONLINE READER**. 2020. Disponível em: <http://mstudioblackboard.tudelft.nl/duwind/Wind%20energy%20online%20reader/Static_pages/Cp_lambda_curve.htm>. Acesso em: 01 de outubro de 2020. Citado na página 5.

WINDBOX. **Componentes dos aerogeradores**. 2020. Disponível em: <<https://windbox.com.br/blog/componentes-dos-aerogeradores/>>. Acesso em: 01 de outubro de 2020. Citado na página 7.

Apêndices

APÊNDICE A – Código Python K-means

```

from sklearn.cluster import KMeans
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd

baseDisp = pd.read_csv('C:/Users/nf1378/Documents/tcc/disponibilidades.
                        csv', sep=';')
baseGen  = pd.read_csv('C:/Users/nf1378/Documents/tcc/geracao.csv', sep='
                        ;')
baseVentos = pd.read_csv('C:/Users/nf1378/Documents/tcc/ventos.csv', sep='
                        ;')

disp = np.array(baseDisp['dezembro'])
gen  = np.array(baseGen['dezembro'])
ventos = np.array(baseVentos['dezembro'])

plt.plot()
plt.title('Dataset')
plt.scatter(disp, gen)
plt.xlabel('Disponibilidade (%)')
plt.ylabel('Geração (kWh)')
plt.show()

# Criando o primeiro gráfico (geração x ventos)
X = np.array(list(zip(disp, gen)))
colors = ['b', 'g', 'r']
markers = ['o', 'v', 's']
plt.ylabel('Geração (kWh)')

kmeans = KMeans(n_clusters=3).fit(X)
plt.scatter(kmeans.cluster_centers_[0], kmeans.cluster_centers_[1],
            s = 100, c = 'yellow', label = '
            Centroids')

for i, l in enumerate(kmeans.labels_):
    plt.plot(disp[i], gen[i], color=colors[l], marker=markers[l])
plt.xlabel('Disponibilidade (%)')
plt.legend()
plt.show()

# Obtendo os clusters
kmeans.labels_

```

```
{i: np.where(kmeans.labels_ == i)[0] for i in range(kmeans.n_clusters)}  
{i: X[np.where(kmeans.labels_ == i)] for i in range(kmeans.n_clusters)}  
  
mydict = {i: np.where(kmeans.labels_ == i)[0] for i in range(kmeans.  
                n_clusters)}  
  
# Transform this dictionary into list (if you need a list as result)  
dictlist = []  
for key, value in mydict.items():  
    temp = [key, value]  
    dictlist.append(temp)
```

APÊNDICE B – Código Python Elbow

```
distortions = []
K = range(1,10)
for k in K:
    kmeanModel = KMeans(n_clusters=k)
    kmeanModel.fit(X)
    distortions.append(kmeanModel.inertia_)

plt.figure(figsize=(16,8))
plt.plot(K, distortions, 'bx-')
plt.xlabel('K')
plt.ylabel('Distorção')
plt.title('Método Elbow - valor ótimo de K')
plt.show()
```