

Estudos de Séries Temporais de Vento Utilizando Análises Estatísticas e Agrupamento de Dados

Cínthia Sales Malta

PROJETO SUBMETIDO AO CORPO DOCENTE DO DEPARTAMENTO DE ENGENHARIA ELÉTRICA DA ESCOLA POLITÉCNICA DA UNIVERSIDADE FEDERAL DO RIO DE JANEIRO, COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE ENGENHEIRO ELETRICISTA.

Aprovado por:

Prof. Carmen Lucia Tancredo Borges, D. Sc.
(Orientador)

Prof. Sergio Sami Hazan, Ph.D.

Eng. Júlio Alberto Silva Dias, M.Sc.

RIO DE JANEIRO, RJ - BRASIL

FEVEREIRO DE 2009

Agradecimentos

Primeiramente agradeço a Deus, sem o qual eu jamais terminaria esta importante etapa da minha vida.

Agradeço ao Paulo, meu noivo e companheiro de todas as horas, que me apoiou e suportou (em todos os sentidos da palavra) nos momentos mais difíceis dos nossos nove anos, sendo os últimos cinco, com certeza, os mais críticos.

Agradeço especialmente aos meus pais, Magno e Rejane, por permitirem que eu saísse cedo de casa e me dedicasse exclusivamente aos meus estudos, e aos meus irmãos Flávia e Fábio por estarem sempre dispostos a me ouvir.

Agradeço às minhas tias Rosane e Ângela e a minha madrinha Lúcia por me ajudarem sempre, e em todos os aspectos que eu necessitasse, em uma cidade até então desconhecida para mim.

Agradeço também aos vários amigos que fiz durante a faculdade, que me apoiaram e me ajudaram muitas vezes com estudos em grupo, boas conversas e saídas. Em especial a Vanessa, Janaína, Fellipe Lessa e Alessandra meus companheiros de todo o curso, e a Carla, Bruno Sodré, Samara, Ubaldo, todos estes grandes companheiros de final de faculdade, sendo os dois últimos “péssimas influências”. Ao amigo Renato Haddad, por toda ajuda e orientação, principalmente no último período da graduação, e no projeto final.

Agradeço a professora Carmen por toda orientação no projeto, e por me influenciar positivamente em relação à escolha da área que pretendo prosseguir.

A todos os professores do departamento de engenharia elétrica da UFRJ, que contribuíram para a minha formação como engenheira.

Resumo

Este trabalho tem por objetivo estudar séries temporais de ventos em três sítios localizados na Alemanha, a fim de avaliar a viabilidade de instalação e a correlação entre usinas eólicas.

Para análise da viabilidade de instalação de usinas eólicas, alguns parâmetros estatísticos são considerados e verificados em gráficos e histogramas gerados, sendo o principal deles a velocidade média dos ventos locais.

Após conhecidas as principais características individuais destes ventos, são realizadas novas análises, no intuito de conhecer suas correlações, o que permite planejar a interligação de sistemas eólicos. Desta forma, poder-se-ia estimar além da potência gerada em cada localidade, a potência simultânea em todas as usinas.

Ainda na análise de correlação, novas séries temporais são geradas. Porém agora com redução do número de estados de vento, utilizando para tal a técnica de agrupamento de dados conhecida como *k-means*. Assim, são encontradas as frequências e probabilidades dos novos estados coincidirem entre as localidades, o que é de suma importância para conhecimento da disponibilidade de energia proveniente de geração eólica em sistemas interligados ao sistema elétrico.

Sumário

Agradecimentos	ii
Resumo	iii
Índice de Tabelas.....	vi
Índice de Figuras	vii
1. Introdução.....	1
1.1 Motivação.....	1
1.2 Objetivo.....	2
1.3 Estrutura	3
2. Geração Eólica.....	4
2.1 Mecanismo de Geração dos Ventos	4
2.2 Histórico da Geração Eólica.....	5
2.3 Instalação de Usinas Eólicas	6
2.2.1 Impactos Sócio-Econômicos	7
2.4 Componentes do Sistema Eólico.....	9
2.5 Energia e Potência Extraídas do Vento	11
2.6 Características Elétricas do Aerogerador	13
2.7 Potencial Eólico.....	14
2.7.1 Potencial Eólico no Mundo	14
2.7.2 Potencial Eólico no Brasil	18

3. Séries Temporais, Análise Estatística e Agrupamento de Dados.....	22
3.1 Séries Temporais	22
3.1.1 Objetivo da Análise das Séries Temporais.....	23
3.1.2 Variações Presentes nas Séries Temporais.....	24
3.1.3 Processos Estocásticos e Estacionários	26
3.2 Análise Estatística	27
3.2.1 Descrição Numérica de Dados	27
3.3 Agrupamento de Dados	30
3.3.1 Técnica de Agrupamento de Dados - <i>K- Means</i>	34
4. Estudos das Séries Temporais de Vento.....	36
4.1 Gráficos e Histogramas	37
4.1.2 Meses com melhores médias de velocidade	40
4.2 Agrupamento dos Estados de Vento de Falkenhagen e Wittstock.....	46
4.2.1 Erros Absolutos – Falkenhagen e Wittstock	49
4.3 Probabilidades de Ocorrência dos Estados de Vento	52
4.3.1 Freqüências e Probabilidades Independentes – Falkenhagen.....	53
4.3.2 Freqüências e Probabilidades Independentes - Wittstock	55
4.3.3 Freqüências e Probabilidades Relacionadas – Falkenhagen x Wittstock	57
4.3.4 Freqüências e Probabilidades Relacionadas -Wittstock x Falkenhagen	61
4.4 Conclusões do Estudo das Séries Temporais	65
5. Conclusões.....	67
6. Referências Bibliográficas	68
Anexo A - Macro Freqüência e Correlação dos Estados de Velocidade de Vento	70

Índice de Tabelas

Tabela 4.1- Dados estatísticos das séries de velocidade de vento.....	39
Tabela 4.2- Falkenhagen anual.....	40
Tabela 4.3- Wittstock anual	41
Tabela 4.4 - Jabel Jan/Jul	42
Tabela 4.5 - Resumo dos meses com melhores velocidades médias.....	43
Tabela 4.6 - Correlações entre os dados das séries.	45
Tabela 4.7 - Centróides representativos das séries.....	46
Tabela 4.8 - Dados estatísticos da série agrupada - Falkenhagen.	47
Tabela 4.9 - Dados estatísticos da série agrupada – Wittstock.....	48
Tabela 4.10 - Principais parâmetros do erro de agrupamento de Falkenhagen.....	50
Tabela 4.11 - Principais parâmetros do erro de agrupamento de Wittstock.....	50
Tabela 4.12 - Centróides representativos das séries.....	52
Tabela 4.13 - Frequência de ocorrência dos estados - Falkenhagen	53
Tabela 4.14 - Probabilidade de ocorrência dos estados - Falkenhagen.....	53
Tabela 4.15 - Frequência de ocorrência dos estados - Wittstock	55
Tabela 4.16 - Probabilidade de ocorrência dos estados - Wittstock.....	55
Tabela 4.17 - Frequência de interação dos estados de vento Falkenhagen e Wittstock.	57
Tabela 4.18 - Probabilidades de interação dos estados de vento Falkenhagen e Wittstock	57
Tabela 4.19 - Frequência de interação dos estados de vento Wittstock e Falkenhagen.	61
Tabela 4.20 - Probabilidades de interação dos estados de vento Wittstock e Falkenhagen.	61

Índice de Figuras

Figura 2.1 - Formação dos Ventos.	4
Figura 2.2- Representação de um sistema de aproveitamento eólico.....	9
Figura 2.3- Evolução das dimensões e potências de turbinas eólicas	10
Figura 2.4 – Curva de rendimento de uma turbina eólica típica.	12
Figura 2.5 – Característica de desempenho de uma turbina eólica.	13
Figura 2.6- Distribuição de área por continente segundo a velocidade média do vento.	14
Figura 2.7 – Potência eólica instalada no mundo (1996-2007).	15
Figura 2.8 - Potência eólica instalada no mundo.....	15
Figura 2.9- Mapa das velocidades médias dos ventos no Brasil.	19
Figura 2.10 - Séries temporais utilizadas para análise na referência [7].....	20
Figura 4.1 - Série de ventos Falkenhagen Jan/Dez 2006.	37
Figura 4.2 - Histograma Falkenhagen Jan/Dez 2006.	37
Figura 4.3 - Série de ventos Wittstock Jan/Dez 2003.	38
Figura 4.4 - Histograma Wittstock Jan/Dez 2003.	38
Figura 4.5 - Série de ventos Jabel Jan/Jul 2004.	38
Figura 4.6 - Histograma Jabel Jan/Jul 2004.	39
Figura 4.7 - Série de ventos Falkenhagen - Dez 2006.....	40
Figura 4.8 - Histograma Falkenhagen Dez - 2006.	41
Figura 4.9 - Série de ventos Wittstock – Dez 2003.....	41
Figura 4.10 - Histograma Wittstock Dez - 2003.	41
Figura 4.11 - Série de ventos Jabel - Dez 2004.....	42
Figura 4.12 - Série de ventos Falkenhagen - Dez 2006.....	42
Figura 4.13 - Série de dados de Jabel e Falkenhagen sobrepostas.	44
Figura 4.14 - Série de dados de Jabel e Wittstock sobrepostas.....	44
Figura 4.15 - Série de dados de Falkenhagen e Wittstock sobrepostas.....	44

Figura 4.16 - Centróides representativos das séries a serem obtidas por agrupamento.	46
Figura 4.17 - Série de dados Falkenhagen agrupada.....	47
Figura 4.18 - Série de dados de Wittstock agrupada.....	47
Figura 4.19 - Porcentagem dos centróides nas séries agrupadas – Falkenhagen.	48
Figura 4.20 - Porcentagem dos centróides nas séries agrupadas - Wittstock.....	49
Figura 4.21 - Erro do agrupamento de Falkenhagen.....	49
Figura 4.22 - Erro do agrupamento de Wittstock.....	50
Figura 4.23 - CDF independente – Falkenhagen.....	54
Figura 4.24 - CDF independente – Wittstock.	55
Figura 4.25 - CDF primeiro estado vento - Falkenhagen.....	58
Figura 4.26 - CDF segundo estado vento - Falkenhagen.	58
Figura 4.27 - CDF terceiro estado vento - Falkenhagen.	59
Figura 4.28 - CDF quarto estado vento - Falkenhagen.	59
Figura 4.29 - CDF quinto estado vento - Falkenhagen.	60
Figura 4.30 - CDF primeiro estado vento - Wittstock.....	62
Figura 4.31 - CDF segundo estado vento - Wittstock.....	62
Figura 4.32 - CDF terceiro estado vento - Wittstock.	63
Figura 4.33 - CDF quarto estado vento - Wittstock.	63
Figura 4.34 - CDF quinto estado vento - Wittstock.	64

1. Introdução

1.1 Motivação

O aumento da demanda por energia elétrica no cenário mundial vem gerando contínua preocupação por partes dos agentes que buscam encontrar novas fontes renováveis de energia, gerando assim a necessidade que estas sejam economicamente viáveis e principalmente menos agressivas ao meio ambiente, quando comparadas aos principais recursos energéticos utilizados atualmente.

Desta forma, existe um contexto de dificuldade cada vez maior na implantação de novas unidades geradoras que utilizem recursos convencionais, tais como a energia hidráulica e combustíveis fósseis, associando a estes fatores técnicos e econômicos e, principalmente ambientais.

Os combustíveis fósseis, principais fontes de energia atual, geram alguns impactos ambientais, dentre eles a emissão de gases poluentes, degradando a atmosfera e provocando o agravamento do aquecimento global. Além disso, os preços dos derivados do petróleo são muito instáveis e estão constantemente em alta, o que pode ser explicado em parte pelos conflitos internos ou internacionais, associados ao cenário de seu esgotamento próximo. No tangente ao aproveitamento hídrico, verifica-se que os locais potenciais ainda não explorados estão cada vez mais escassos e mais afastados dos centros de consumo, acarretando altos custos para sua exploração.

Verifica-se hoje que, além dos fatores ambientais existem outros que objetivam a utilização de novos recursos de geração originada de fontes alternativas e renováveis. O aprimoramento tecnológico é um deles. Pode ser citado como tal o desenvolvimento de novos materiais. Assim como as novas e mais eficientes técnicas de controle, fundamentais para melhor desempenho e confiabilidade de dispositivos, as quais vêm se estabelecendo no mercado de produção de energia. Desta forma as inovadoras técnicas vêm provocando mudanças e tornando as fontes alternativas competitivas em relação às tradicionais [1].

Muitas são as energias alternativas em uso atualmente no mundo. Sua escolha se deve principalmente a fatores como o domínio de determinada tecnologia e principalmente a disponibilidade de tal recurso. De forma geral, podem-se considerar relevantes no mercado de energia mundial, as provenientes da biomassa, dos pequenos aproveitamentos hidroelétricos e da energia solar.

Em relação à energia produzida pelo sol, destaca-se a energia eólica, que nos últimos anos obteve expressivo crescimento. Este aproveitamento tem sido uma das alternativas mais bem sucedidas, com tendência de crescimento acelerado e sustentado em muitos países [4].

Deve-se ressaltar que a utilização da energia eólica pelo homem data de pelo menos 5000 anos, quando os Egípcios já utilizavam técnicas para navegação. No entanto, sua utilização em maior escala veio com a crise do petróleo vivenciada nos anos 70, o que forçou diversos setores da sociedade a investir recursos em busca de sistemas alternativos mais baratos, eficientes e confiáveis [1].

Os fatores mais atraentes nesta forma energética é o domínio na fabricação de turbinas eólicas, e seu processo de fabricação de energia considerado “limpo”, diferentemente dos principais recursos energéticos utilizados atualmente.

Para avaliar a disponibilidade de recursos eólicos de uma região, são necessários sistemáticos trabalhos de coleta e análise de dados sobre a velocidade e regime dos ventos. Para uma avaliação rigorosa, estudos específicos devem ser feitos, porém, para uma análise bruta, dados coletados em aeroportos, estações meteorológicas, podem fornecer uma primeira estimativa [2].

A partir dos dados encontrados deve ser feita uma avaliação técnica, para definir a viabilidade de geração de energia a partir dos ventos desta localidade. Assim, os *trade winds* podem ser identificados. Estes são definidos por ventos que possuem altas velocidades médias, pouca variação em suas direções e pouca turbulência ao longo do ano [3].

1.2 Objetivo

Este trabalho tem por objetivo estudar o potencial eólico, e conseqüentemente a viabilidade de instalação de usinas eólicas em específicas localidades na Alemanha. Isto será feito a partir da análise de dados reais de velocidades médias de ventos, obtidos nos locais pretendidos.

Os dados estudados são caracterizados por séries temporais, possuindo muitas características comuns. Para tanto, um estudo da teoria de séries temporais, análises estatísticas e agrupamento de dados será realizado, e desta forma o inter-relacionamento das localidades será conhecido.

Em um segundo momento, após o levantamento das principais características dos ventos estudados, uma nova série será obtida a partir do agrupamento dos dados das séries originais. Este estudo visa diminuir o número de estados de vento de cada série, sendo que isto pode ser realizado por várias técnicas de agrupamento de dados, porém a utilizada neste trabalho é o *k-means*, a qual associa os dados a partir de suas similaridades.

Por fim será feito um levantamento das frequências e probabilidades de ocorrência dos novos estados de vento, obtidos a partir do agrupamento de dados. Neste contexto serão feitas duas novas análises. A primeira no intuito de avaliar individualmente as frequências e correlações dos estados dentro de sua própria série agrupada, e na segunda as análises serão feitas avaliando as frequências e correlações dos novos estados de uma localidade referência com os estados agrupados do outro local de estudo.

As análises de frequência e correlação são de suma importância para o estudo de sistemas de energia eólica interligados ao sistema elétrico, quando é preciso estudar a disponibilidade de energia proveniente da geração eólica.

1.3 Estrutura

O capítulo 2 do presente trabalho está dedicado ao estudo do sistema eólico, conceituando-o e apresentando seus principais aspectos históricos, características construtivas, é apresentado os principais dados de potencial eólico instalado no mundo e no Brasil.

O capítulo 3 tem por objetivo fornecer base para o estudo das séries temporais, apresentando as principais análises estatísticas e técnicas de agrupamento de dados, onde as primeiras são utilizadas para estudo das séries originais e as últimas para geração de novas séries representativas das séries originais.

O objetivo do capítulo 4 é aplicar os conceitos vistos no capítulo 3. Assim, cada uma das séries estudadas tem seus parâmetros estatísticos e análise do comportamento das séries originais e das obtidas pelos agrupamentos, apresentadas. No intuito de conhecer o comportamento individual e inter-relacionado das localidades em que foram medidos os dados de vento, são encontradas as frequências e correlações entre os novos estados de vento.

E por fim, o capítulo 5 apresenta as principais considerações referentes ao trabalho, resumindo os principais resultados encontrados no capítulo 4.

2. Geração Eólica

2.1 Mecanismo de Geração dos Ventos

A energia eólica é uma das formas de energia provenientes da radiação solar, uma vez que os ventos são gerados pelo aquecimento irregular da superfície terrestre. A atmosfera terrestre possui papel fundamental na formação dos ventos, pois é nela que ocorrem as mudanças que influenciam o clima e os ventos, causando também os aquecimentos não homogêneos da superfície terrestre [4].

As regiões tropicais que recebem os raios solares quase perpendicularmente, são mais aquecidas do que as regiões polares. O ar quente localizado nas baixas altitudes nas regiões tropicais tende a subir, sendo substituído por uma massa de ar mais frio que se desloca das regiões polares. Os deslocamentos de massas de ar determinam a formação dos ventos [4].

Assim, devido ao mecanismo descrito anteriormente, existem locais os quais os ventos jamais cessam. Estes são chamados de ventos planetários ou constantes, e podem ser classificados como: ventos alísios, que se deslocam dos trópicos para o equador, os contra alísios, sopram do equador para os pólos, os ventos de oeste, que se deslocam dos trópicos para os pólos, e os ventos polares que sopram dos pólos para as zonas temperadas.

Devido à inclinação do eixo da Terra em relação a sua órbita, aparecem também os ventos continentais ou periódicos, que compreendem as monções ou as brisas. Estes são formados devido às variações sazonais na distribuição de radiação recebida na superfície da Terra, provocando variações sazonais na intensidade e duração dos ventos.

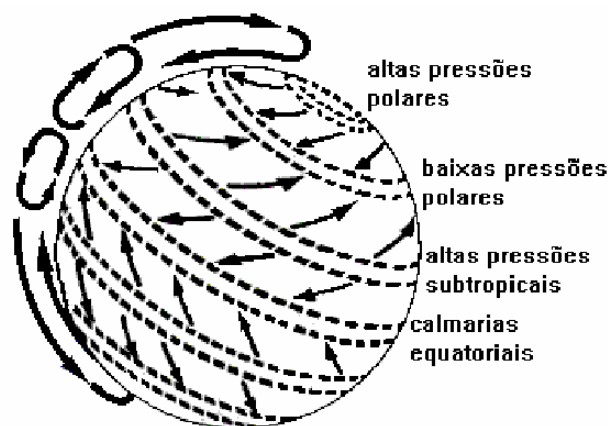


Figura 2.1 - Formação dos Ventos [3].

Fonte: Energia Eólica e Meio Ambiente, site: www.meioambienteenergiaeolica.com.br.

É importante conhecer a direção predominante dos ventos locais, tanto para os fabricantes de turbinas eólicas, quanto para a implantação de usinas eólicas. Os primeiros precisam de informações que diminuam seus custos e otimizem as turbinas produzidas, enquanto as empresas encarregadas da geração de energia local necessitam estimar a potência e o rendimento que poderão ser obtidos [7].

2.2 Histórico da Geração Eólica

Os primeiros indícios de aproveitamento da força dos ventos pelo homem datam de aproximadamente milhares de anos no oriente. Os Egípcios a mais de 5.000 anos já utilizavam a força dos ventos para a navegação. Na idade média já era possível utilizar em larga escala força aerodinâmica dos ventos para melhor aproveitamento das máquinas eólicas e principalmente para as grandes navegações. A partir do século XIV os holandeses já alcançaram a liderança na melhoria contínua dos moinhos de vento, apresentando grande potência e aplicação como fonte de energia, principalmente na moagem de grãos e bombeamento de água.

As máquinas eólicas também foram utilizadas largamente durante a invenção da imprensa, onde eram usadas para fabricação de papel. Nos Estados Unidos do século XIX, após a abolição da escravidão, muitos cataventos multipás eram utilizados, contribuindo para a fixação às partes áridas e o acesso a água, no então desabitado oeste, possibilitando sua exploração e dominação. Porém, principalmente na Europa, por volta do século XIX, com a introdução das máquinas a vapor oriundas da revolução industrial, houve um desuso gradual desse tipo de máquina [1].

As primeiras utilizações da energia eólica para produção de energia elétrica datam de 1930, com o uso de pequenos aerogeradores, principalmente para carregamento de baterias, o que facilitou o acesso a energia elétrica aos moradores do meio rural.

As turbinas eólicas de grande porte, usadas para geração de eletricidade em grande escala como as existentes hoje, tiveram início na Alemanha nos anos 50 [1]. Foi principalmente a partir da década de 70 com o primeiro choque do petróleo, que a obtenção de energia elétrica via sistemas eólicos passou a ser uma atraente forma de geração energética.

Contudo foi a partir de grandes estímulos governamentais nas décadas de 80 e 90, em alguns países como Estados Unidos, Dinamarca e Alemanha, que a energia eólica passou a contribuir de forma mais significativa ao sistema elétrico, com sistemas mais eficientes e com custos mais baixos [1].

2.3 Instalação de Usinas Eólicas

Conforme mencionado anteriormente, a direção dos ventos está constantemente variando. Para avaliar o potencial energético de um local, é necessário realizar estudos por um longo período de tempo, para assim analisar seu comportamento e evitar resultados imprecisos.

O vento pode ter sua representação no domínio do tempo e no domínio da frequência [8]. Esta última forma é feita através de representação espectral. Para correta análise dos dados registrados, é necessário que haja uma quantidade significativa destes, preferencialmente um período de pelo menos um ano.

Para que a energia eólica seja considerada tecnicamente aproveitável, a densidade dos ventos deve ser maior ou igual a 500 W/m^2 , a uma altura de 50 m, isso equivale a ventos com velocidades médias mínimas de 7 a 8 m/s. Porém costuma-se iniciar a geração eólica com ventos a partir de 2,5 a 3 m/s, e ventos com velocidades acima de 12 a 15 m/s ativam o sistema automático de limitação de potência da máquina, sendo este controle definido pelo tipo de turbina. Em ventos muito fortes, como por exemplo, acima de 25 m/s, atua o sistema automático de proteção. Estes ventos possuem alta turbulência e podem danificar a estrutura da máquina. Em sistemas conectados ao sistema elétrico, a rotação das pás é reduzida e o sistema do gerador é desconectado da rede [2].

Outro fato de importância a ser mencionado para cálculo de projeto de instalação de uma usina eólica é a turbulência atmosférica. Esta afeta a conversão de energia, principalmente devido às variações na direção dos ventos, submetendo a turbina eólica a maiores esforços mecânicos.

Quando se pretende instalar o aproveitamento eólico, outros fatores podem influenciar o escoamento do vento e devem ser avaliados, são eles [7]:

- **Obstáculos:** possuem grande influência na diminuição da velocidade dos ventos, além de serem fontes de turbulência nas vizinhanças. São considerados obstáculos: edifícios, árvores, formações rochosas.
- **Sombreamento da torre e Efeito Esteira:** uma das possíveis causas da variação da potência ativa, o que pode provocar variações na tensão. Efeito Esteira: o vento que sai da turbina eólica possui conteúdo energético inferior ao vento de entrada. Isto se deve a produção de energia mecânica a partir da energia cinética incidente. E ainda atrás da turbina se forma uma esteira de vento turbulento devido à ação das pás, e com velocidade reduzida.

Para evitar este efeito, instalam-se as turbinas distanciando entre 5 a 9 vezes o diâmetro do rotor da turbina eólica, na direção predominante do vento, e entre 3 a 5 vezes na direção perpendicular a direção predominante do vento.

2.2.1 Impactos Sócio-Econômicos

Os aproveitamentos eólicos são alternativas não apenas a grandes níveis de demanda, mas também para suprir pequenas localidades distantes da rede, universalizando o abastecimento energético. As usinas de grande porte, mais conhecidas como fazendas eólicas, estão interligadas ao sistema elétrico geral [1].

Possui vantagens significativas como a diminuição da emissão de poluentes, liberados pelas usinas térmicas, diminuindo a necessidade de construção de grandes reservatórios, e a sensibilidade a sazonalidade hidrológica sentida pelas usinas hidráulicas.

Em suma, a energia eólica demonstra ser uma forma limpa e eficiente na produção de energia elétrica. Apesar de inúmeras vantagens em instalação de usinas deste tipo, estudos rigorosos devem ser efetuados visando avaliar os possíveis impactos ao meio ambiente e ao homem. Seguem algumas de suas principais características [3]:

- Parques eólicos não limitam sua utilização da área ocupada, podendo ser usadas também para agricultura, pecuária etc.
- A área utilizada para instalação das fazendas eólicas não é excessiva quando comparada, por exemplo, à construção de uma usina hidrelétrica.
- As turbinas das fazendas eólicas devem ser construídas com distâncias mínimas de segurança para o funcionamento aerodinâmico de turbinas próximas.
- Sistemas de conversão eólica emitem ruídos provenientes das partes mecânicas, e também de efeitos aerodinâmicos. Em sistemas com velocidade variável possuem geradores de baixa rotação, dispensando a caixa de velocidade. Já em sistemas de grande porte, os efeitos aerodinâmicos são mais acentuados emitem maiores ruídos, tornando-se desconfortáveis para a população residente nas proximidades.

- A grande altura atingida pelas torres de alguns modelos de aerogeradores (podem ter alturas superiores a 100 m de altura) provoca alteração visual na paisagem. Isto pode ser amenizado com a pintura das torres de geração em tons parecidos com a paisagem.
- Algumas vezes também podem ser observadas aves mortas devido ao choque com as pás dos aerogeradores. Por isso deve ser considerado a possibilidade desta localidade ser rota de aves migratórias, mesmo que isto não represente grande problema à construção destas usinas.
- Os aerogeradores algumas vezes podem refletir as ondas eletromagnéticas do local. Isto pode provocar interferência aos sistemas de telecomunicações. Por isso, deve-se fazer um estudo mais criterioso caso se queira instalar um parque eólico próximo a aeroportos, ou sistemas de transmissão de sinais.
- Quanto ao risco ao ser humano, poucos foram os casos relatados de acidentes por desprendimentos de pedaços de pás, ou gelo que estejam presos à estrutura dos aerogeradores. Portanto trata-se de um dos sistemas mais seguros de geração de energia elétrica.

2.4 Componentes do Sistema Eólico

A figura abaixo representa os principais componentes do sistema conversor de energia eólica.

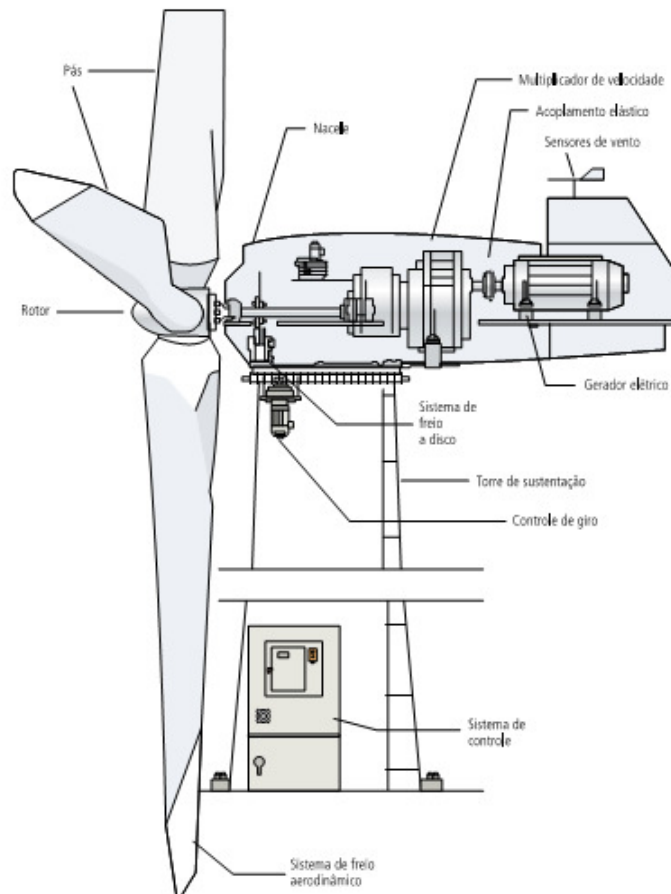


Figura 2.2- Representação de um sistema de aproveitamento eólico [2].

Fonte: Atlas de Energia Elétrica do Brasil, Energia Eólica, 2ª Edição.

Considera-se que os principais componentes do sistema de conversão eólico são: rotor, transmissão, controle, geradores ou conversores e sistema de armazenamento [4].

- **Rotor**

O rotor de um aerogerador é responsável por transformar a energia cinética do vento em energia mecânica através da rotação do eixo, transmitindo-a por intermédio das pás ao eixo do gerador elétrico. Estes podem ser classificados em rotores de eixo vertical (Savonius, Darrieus etc.) e eixo horizontal (hélice, holandês, multipás etc.). As turbinas se desenvolveram rapidamente em tamanho e tecnologia principalmente durante as duas últimas décadas, e esse aumento de tamanho possui relação direta com a potência que pode ser extraída do vento.

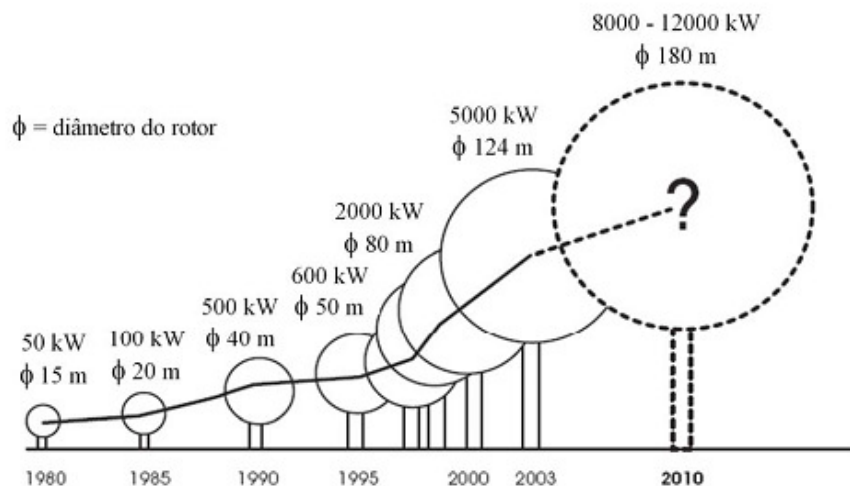


Figura 2.3- Evolução das dimensões e potências de turbinas eólicas

Fonte: www.ewea.org/documents/facts_volume%201.pdf

- **Transmissão**

Não aparece em todos os sistemas eólicos, mas tem como função transmitir a energia mecânica entregue pelo eixo do rotor até a carga. Quando estes componentes não são utilizados o eixo do rotor é acoplado diretamente à carga. A transmissão mais usada é a do tipo por engrenagens, que multiplica a velocidade angular para melhor aproveitamento do gerador.

- **Conversor e Gerador**

Responsável pela conversão da energia mecânica do eixo em energia elétrica. Apesar de modelos existentes no mercado de alta confiabilidade, nestas aplicações possuem algumas limitações. O principal problema é a variação de velocidade de vento, o que acarreta o funcionamento do gerador em uma extensa faixa de rotações, variações nos controles de torque, tensão e controle de frequência.

- **Mecanismo de Controle**

Trata-se do controle dos componentes do sistema, responsável pela orientação do rotor, controle de velocidade, controle da carga etc.. Este sistema é constituído por sensores, como de vento, rotor, bateria e outros, permitem que haja um controle e armazenamento de dados que levam ao melhor funcionamento do sistema como um todo.

- **Sistemas de Armazenamento**

Armazenam a energia para produção de energia firme a partir de uma fonte intermitente.

- **Suporte Estrutural**

Formado principalmente pela torre, estas suportam o rotor em alturas nas quais a velocidade do vento é maior e menos sujeita a perturbações, já que o rotor próximo ao solo sofre influência de obstáculos. As torres modernas chegam a ter alturas superiores a 100 m, o que necessita de bons projetos, para que estas possam resistir a intempéries ao longo de sua vida útil, que é de aproximadamente 20 anos.

2.5 Energia e Potência Extraídas do Vento

Para o aproveitamento da energia elétrica a partir da força dos ventos, a condição primordial é que estes possuam fluxo contínuo e razoavelmente forte. Como já fora dito anteriormente, a velocidade média de operação das turbinas eólicas deve estar preferencialmente entre 7 e 12 m/s, para velocidades maiores que estas o sistema manterá a potência de saída num valor máximo ou nominal. A escolha do valor máximo da potência de saída depende do tipo de conversor eólico utilizado, se é do tipo de velocidade fixa ou variável [3].

A partir da energia disponível no vento pode-se estipular a correspondente potência mecânica (P_{disp}), mostrada abaixo [2]:

$$P_{disp} = \frac{1}{2} \rho A v^3 \quad (2.1)$$

Porém nem toda energia cinética dos ventos é aproveitada pelas turbinas eólicas. Para levar em consideração este fato, é introduzido um coeficiente chamado coeficiente de aproveitamento aerodinâmico, C_p , que é dado por:

$$C_p = \frac{P_{mec}}{P_{disp}} \quad (2.2)$$

A partir das equações 2.1 e 2.2 obtêm-se a seguinte expressão que representa a potência na turbina eólica:

$$P = \frac{1}{2} \rho A_r v^3 C_p \quad (2.3)$$

Sendo:

ρ = densidade do ar em kg/m.

$A_r = \pi D^2 / 4$, em que D é o diâmetro do rotor.

C_p = coeficiente aerodinâmico de potência do rotor.

v = velocidade cinética do vento.

A figura representa o desempenho do coeficiente de potência C_p em função da razão da velocidade do vento. Observa-se que o coeficiente de potência cresce até que seja atingida a velocidade nominal da turbina, a partir daí ele decresce com o aumento da velocidade, o valor máximo ocorre a aproximadamente 44%.

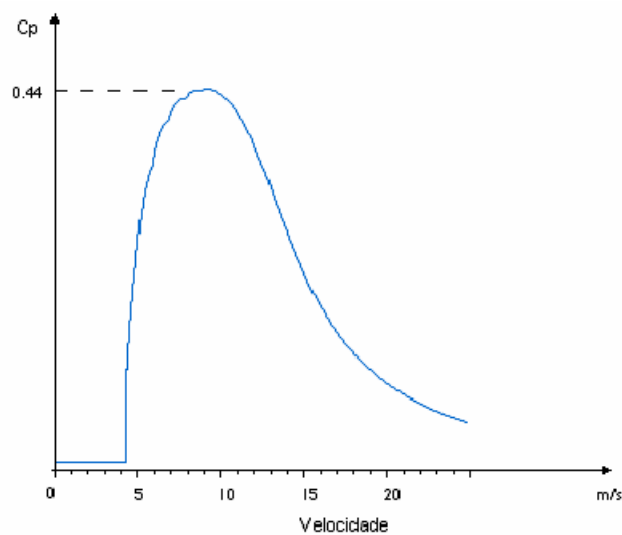


Figura 2.4 – Curva de rendimento de uma turbina eólica típica [7].

2.6 Características Elétricas do Aerogerador

As turbinas eólicas são projetadas para gerar a máxima potência para uma determinada velocidade de vento. Esta potência é chamada de potência nominal, e a velocidade do vento na qual esta é atingida, é conhecida como velocidade nominal do vento. Esta velocidade é ajustada de acordo com a média de vento do local, sendo comuns valores de 12 m/s a 15 m/s [2].

Para iniciar a operação, as turbinas eólicas necessitam de uma velocidade mínima, conhecida como velocidade de acionamento ou de partida (V_{in} - *cut-in wind speed*), como citado anteriormente estes valores encontram-se entre 2.5 a 3 m/s. Por outro lado, a turbina é retirada de operação quando a velocidade do vento atinge outro limite, denominado vento de corte ou de saída (V_{out} - *cut-out wind speed*), com valores aproximados de 25 m/s, podendo caso não haja interrupção do equipamento, gerar danos físicos ao aerogerador devido às cargas mecânicas que surgirão [7].

A característica de desempenho das turbinas eólicas é dada por gráficos que correlacionam o vento com a potência, recendo o nome de curvas de “*velocidade de vento x potência de saída*”.

Para exemplificar, a figura 2.5 representa uma curva para um sistema eólico típico. Portanto, para uma exata determinação da característica do aerogerador, é necessário conhecer o regime de ventos no local da instalação do equipamento.

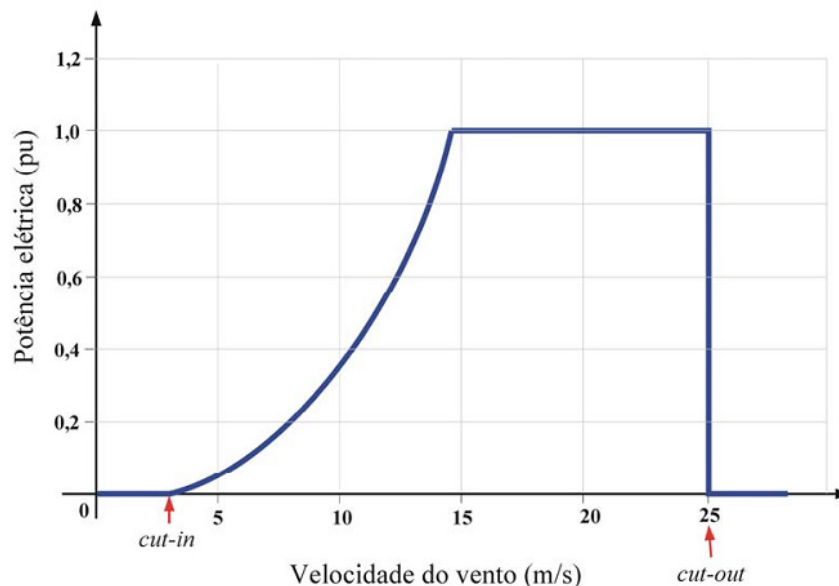


Figura 2.5 – Característica de desempenho de uma turbina eólica.

Observando a figura 2.5 nota-se que para velocidades inferiores a aproximadamente 3 m/s a potência gerada é nula. Como a potência gerada é proporcional a velocidade do vento ao cubo, não é economicamente atrativo gerar energia para ventos com valores abaixo das velocidades de acionamento, por volta de 2.5 a 3m/s, como citado anteriormente.

Analogamente, para valores muito altos, não é econômico aumentar a potência das turbinas, já que isso exigiria maior robustez do sistema, acarretando maiores custos. Muitas vezes estes valores muito altos de velocidades de ventos ocorrem poucas vezes ao ano, o que não compensa a adaptação do sistema eólico.

2.7 Potencial Eólico

2.7.1 Potencial Eólico no Mundo

Segundo a Organização Mundial de Meteorologia, apenas 13% da superfície terrestre tem velocidade média superior a 7 m/s, a uma altura de 50 m, que é a principal condição de viabilidade de instalação de sistemas eólicos. Esta distribuição é demonstrada na figura 2.6 abaixo:

Região/Continentes	Velocidade do Vento (m/s) a 50 m de Altura					
	6,4 a 7,0		7,0 a 7,5		7,5 a 11,9	
	(10 ³ km ²)	(%)	(10 ³ km ²)	(%)	(10 ³ km ²)	(%)
África	3.750	12	3.350	11	200	1
Austrália	850	8	400	4	550	5
América do Norte	2.550	12	1.750	8	3.350	15
América Latina	1.400	8	850	5	950	5
Europa Ocidental	345	8,6	416	10	371	22
Europa Ocidental & ex-URSS	3.377	15	2.260	10	1.146	5
Ásia (excluindo ex-URSS)	1.550	6	450	2	200	5
Mundo	13.650	10	9.550	7	8.350	6

Figura 2.6- Distribuição de área por continente segundo a velocidade média do vento [2].

Fonte: Atlas de Energia Elétrica do Brasil, Energia Eólica, 2ª Edição.

Em 2007 foram instalados aproximadamente 20000 MW em energia eólica pelo mundo. Sendo os principais contribuintes deste significativo aumento países como EUA, Espanha e China [5]. Os países com maiores capacidades instaladas são: a Alemanha, com 22,3 GW; os Estados Unidos, com 18,8 GW; a Espanha, com 15,1 GW; a Índia, com 7,8 GW e a China, com 5,9 GW [5].

Mesmo que o maior crescimento do mercado de energia eólica em 2007 não tenha sido na Europa, a União Européia continua a ser o maior mercado de energia eólica do mundo, com mais de 8,5 GW instalados em 2007, totalizando 56,5 GW. Os principais países europeus em capacidade eólica utilizada são a Alemanha, Espanha, França, Itália e Ucrânia [5].

A figura 2.7 mostra a potência eólica acumulada no mundo nos anos de 1996 a 2007.

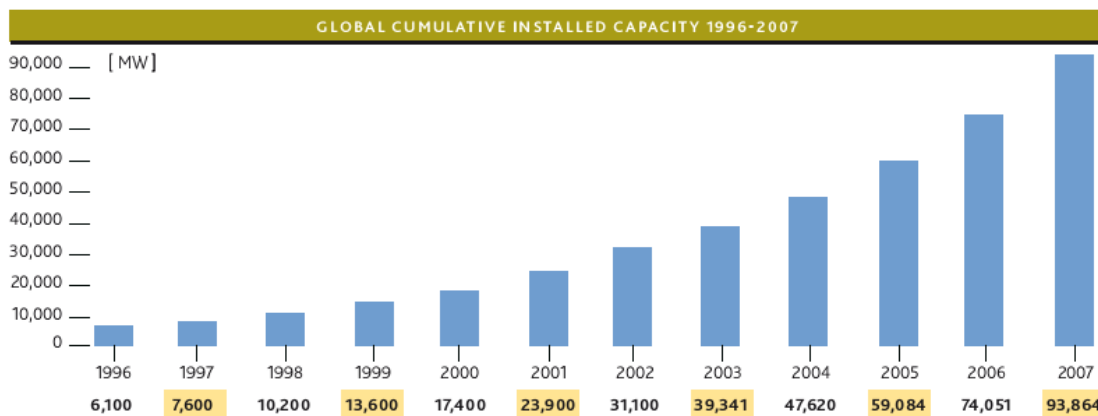


Figura 2.7 – Potência eólica instalada no mundo (1996-2007) [5].

Fonte: Global Wind Energy Outlook 2008.

A figura 2.8 mostra como a potência instalada mundial está dividida, considerando os principais países inseridos no mercado de energia eólica.

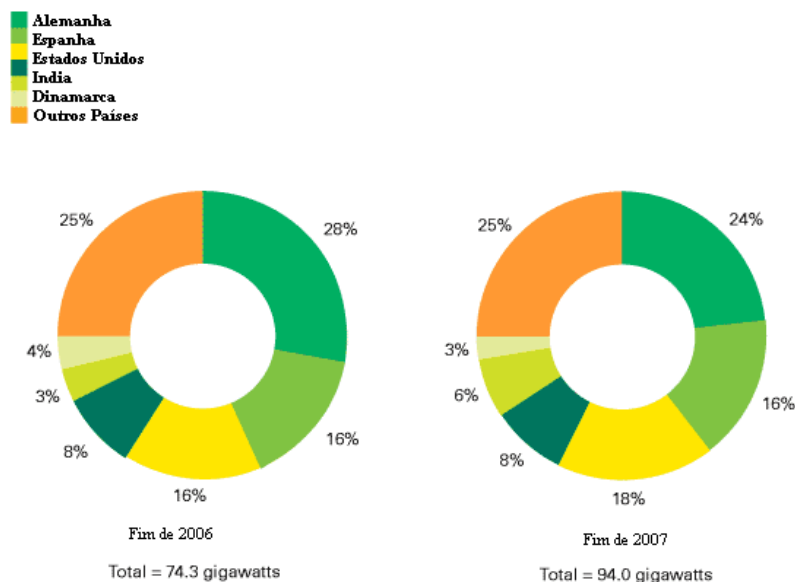


Figura 2.8 - Potência eólica instalada no mundo [2].

Fonte: <http://www.bp.com>.

Os tópicos abaixo sintetizam as mais significativas características dos principais mercados de energia eólica mundiais.

- **Alemanha**

A capacidade de energia eólica instalada na Alemanha é de 22,3 GW, maior que qualquer outro país no mundo, provendo cerca de 7% de toda a energia elétrica consumida no país. Além de instalados importantes fabricantes de turbinas eólicas no país. Desde meados dos anos 90, quando recebeu apoio de importantes políticas nacionais e regionais, o rápido crescimento do mercado eólico alemão tem servido de exemplo para outros países europeus e mundiais [1].

Os fortes incentivos financeiros governamentais permitiram que o aproveitamento eólico na Alemanha tivesse dois principais aspectos. O primeiro foi desenvolver áreas não obviamente vantajosas, como ao longo do Mar do Norte. Assim a indústria respondeu produzindo turbinas mais eficientes para trabalhar em lugares com média de ventos mais baixa [1].

O segundo efeito foi a abertura dos mercados eólicos, para uma grande faixa de pessoas. Assim parques eólicos de diversos tamanhos puderam ser instalados, onde pequenos empresários e empresas adquiriram significativas cotas, se beneficiando com abatimento de impostos.

Em conjunto com estes aspectos, observa-se também a influência das políticas verdes, muito fortes na Alemanha. Com isso, o governo alemão estabeleceu como objetivo que, até 2025 25% de toda a eletricidade do país seja proveniente da geração eólica. Grande parte disso produzido em parques eólicos instalados no Mar Báltico e na praia do Norte [1].

- **Espanha**

A Espanha teve expressivo crescimento em geração eólica em 2007. Com capacidade total instalada de 15,1 GW, suprimindo 10% de toda a energia elétrica demandada no país [5], ratificando as previsões de anos anteriores, que colocavam o mercado espanhol como um dos mais promissores e prósperos de toda a Europa.

As áreas rurais escassamente povoadas, somadas a fortes incentivos governamentais constituem excelentes oportunidades para novos investidores e fabricantes de turbinas. O aproveitamento eólico na Espanha iniciou-se na região de Tarifa, chegando ao litoral Atlântico a noroeste, as Montanhas de Navarre e nas planícies de Castilla de La Mancha [1].

O sucesso do aproveitamento eólico espanhol pode ser associado a vários fatores: um ótimo regime de ventos, política de desenvolvimento regional focalizada e esquemas de apoio nacionais fortes e diretos.

Em 1998 o governo confirmou o objetivo de compromisso com as fontes de energia renováveis, estabelecendo que até 2010 pelo menos 12% de toda a eletricidade fosse proveniente destas fontes [1]. Esta meta é perfeitamente possível, já que em 2008, apenas a geração de energia eólica supre 10% de toda a demanda por eletricidade, conforme citado anteriormente.

O modelo espanhol difere-se também da maioria dos outros modelos europeus, principalmente no referente aos tipos de fazendas eólicas construídas, que são de grande porte. Os investimentos para tais construções provêm de consórcios que unem concessionárias, governo regional e fabricantes de turbinas.

- **Dinamarca**

A indústria de energia eólica na Dinamarca possui seu principal sucesso no comércio de turbinas eólicas. Estes equipamentos dominam o mercado global, com uma das maiores taxas de crescimento mundial. Uns dos motivos para sucesso da indústria eólica dinamarquesa são os incentivos energéticos nacionais para diminuir a dependência do combustível importado, a fim de promover melhorias no meio ambiente e maior sustentabilidade ambiental.

Outro importante aspecto no sucesso da indústria eólica dinamarquesa foi a inovação tecnológica. Procuram ao invés de produzir grandes turbinas, desenvolver máquinas menores e de melhor desempenho em contrapartida as tendências mundiais, que produziam maiores turbinas, acreditando ser essa a melhor tecnologia [1].

- **Índia**

A energia eólica está em constante crescimento neste país, com aproximadamente 25% ao ano, sendo justificado por seu grande potencial litorâneo. Em 2007, houve um crescimento expressivo em potência eólica instalada, 1,5 GW, totalizando ao final de 2007 7,9 GW [5].

O Ministério de Fontes de Energia Não-Convencionais da Índia (MNES) incentivou o desenvolvimento de novas fontes de energia a fim de diminuir a demanda por carvão, óleo e gás, necessários para o rápido crescimento do país [1].

Houve uma parceria entre o setor privado e o poder público para investimentos no setor, e assim fortalecendo a base industrial. Assim algumas empresas produzem até 70% dos componentes da turbina eólica, ao invés de importar dos principais fabricantes europeus [1].

- **China**

No fim de 2007, o total de potência eólica instalada na China foi de 5,9 GW, com adição de 3,3 GW neste mesmo ano, o que significa um crescimento de 146% em relação a 2006. Especialistas acreditam que isso é apenas o começo, visto o potencial energético eólico chinês, e o grande crescimento ainda está por vir. As regiões com maiores possibilidades de desenvolvimento de energia eólica são a costa sudeste, o norte e o oeste do país [5].

No passado as turbinas importadas dominavam o mercado, porém hoje a realidade não é mais essa. Há um grande incentivo por parte do governo no tocante ao desenvolvimento de turbinas em localidades chinesas. Este acredita que isto trará benefícios para as populações locais, além de diminuir os custos de instalações de usinas eólicas.

2.7.2 Potencial Eólico no Brasil

A prioridade de desenvolvimento em fontes alternativas de energia no Brasil nos últimos anos tem sido a biomassa. Porém, é esperado um aumento considerável de investimento em energia eólica nos próximos anos. Em dezembro de 2001, foi elaborado o PROINFA – Programa de Incentivo às Fontes Alternativas de Energia. Com o objetivo, a partir de 2002 de aumentar a participação de novas fontes de energia por Produtores Independentes Autônomos, em biomassa, pequenas centrais hidrelétricas e energia eólica [5].

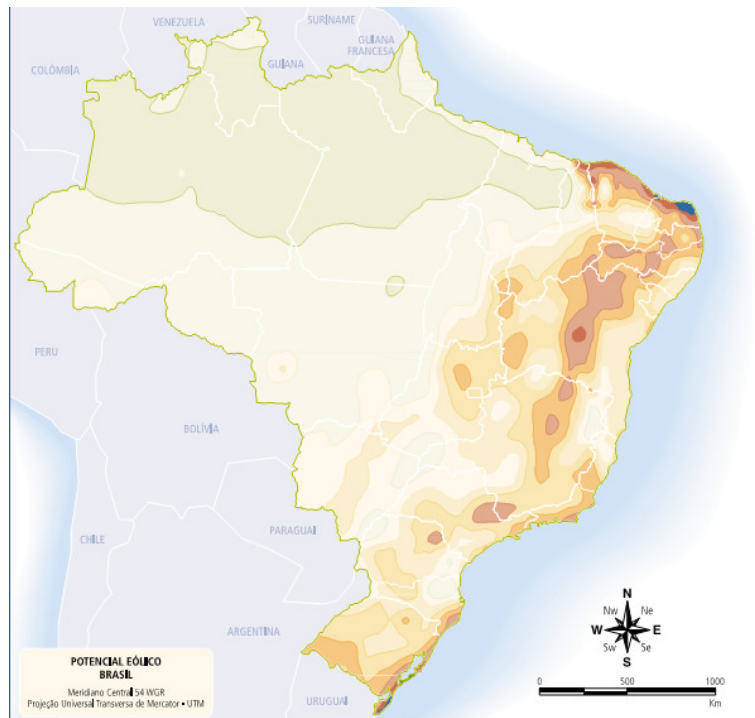
Estão em construção 14 fazendas eólicas financiadas pelo PROINFA em 2008, com uma capacidade total instalada de 103 MW. E para 2009 está prevista a construção de 27 novas usinas eólicas, com capacidade total de 901 MW, que devem ser inseridas no sistema elétrico [5].

Mais de 5 GW de potência eólica já estão registrados na agência regulatória de energia elétrica, a ANEL, esperando a aprovação de recursos para construção e planejamento de novas usinas.

No Brasil a região pioneira em instalação eólica é a Nordeste, e é também onde se localizam a maioria dos projetos instalados no Brasil. Hoje a carga instalada é de aproximadamente de 247 MW, o que corresponde a 0,2 % da potência instalada no Brasil. Estando prevista para até o fim de 2008 uma potência total instalada de 1400 MW [6].

Os tópicos abaixo resumem regionalmente os principais potenciais eólicos no Brasil.

- **Região Sudeste:** norte fluminense, Espírito Santo e elevações em São Paulo.
- **Região Nordeste:** Piauí, Rio Grande do Norte, Ceará, faixa litorânea do Maranhão, chapadas e *off shore*.
- **Região Norte:** faixa litorânea do Amapá e Pará.
- **Região Centro-Oeste:** áreas fronteiriças com o Paraguai.
- **Região Sul:** regiões litorâneas.



		Velocidade média do vento (m/s) 50 m acima do nível da superfície				
		Mata	Campo Aberto	Zona Costeira	Morro	Montanha
Classes de energia	4	> 6,0	> 7,0	> 8,0	> 9,0	> 11,0
	3	4,5 - 6,0	6,0 - 7,0	6,0 - 7,0	7,5 - 9,0	8,5 - 11,0
	2	3,0 - 4,5	4,5 - 6,0	4,5 - 6,0	6,0 - 7,5	7,0 - 8,5
	1	< 3,0	< 4,5	< 4,5	< 6,0	< 7,0

NOTAS:
Mata indica áreas de vegetação nativa, com arbustos e árvores altas.
Campo aberto refere-se a áreas planas de pastagens, plantações e/ou vegetação baixa, sem muitas árvores altas.
Zonas costeiras são áreas de praia, normalmente com larga faixa de areia, onde o vento incide predominantemente no sentido mar-terra.
Morros são áreas de relevo levemente ondulado, relativamente complexo e de pouca vegetação ou pasto.
Montanhas representam áreas de relevo complexo com altas montanhas. O potencial eólico é dado para locais nos topos das montanhas em condições favoráveis para o fluxo de vento.

Figura 2.9- Mapa das velocidades médias dos ventos no Brasil [2].

Fonte: Atlas de Energia Elétrica do Brasil, Energia Eólica.

2.7.2.1 Algumas séries de vento no Brasil

A referência [7] analisa séries de vento coletadas em localidades com grande potencial eólico no ano de 2002. As séries nomeadas Nordeste I, Nordeste II, Sul e Sudeste, são mostradas e comparadas, conforme verifica-se na figura 2.10.

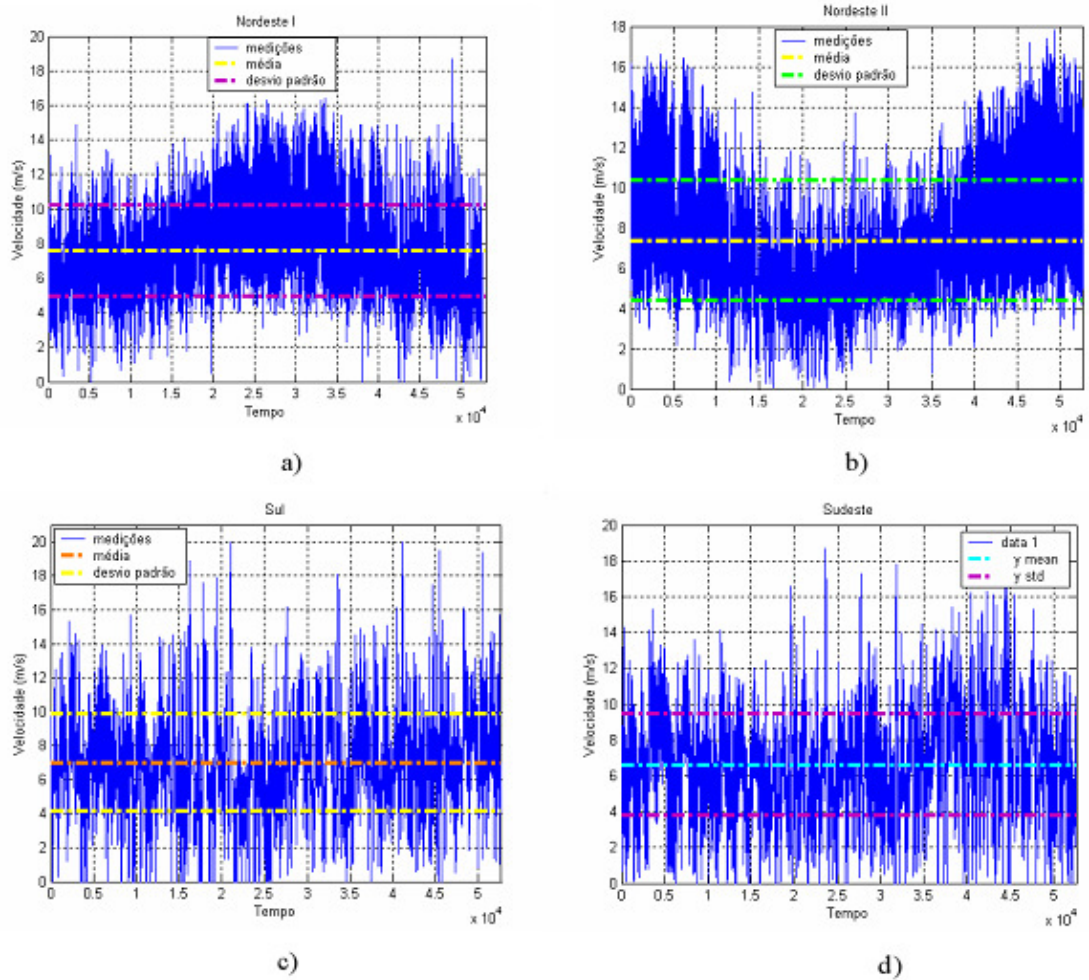


Figura 2.10 - Séries temporais utilizadas para análise na referência [7], Cap. 5.

a) Nordeste I, b) Nordeste II, c) Sul, d) Sudeste.

Verifica-se que as médias anuais dos ventos variam de 6,59 a 7,58 m/s, caracterizando de uma forma simplificada estas localidades como viáveis para instalações de usinas eólicas. Com velocidades máximas variando de 17,8 a 20 m/s nos sítios analisados, o que está sintetizado na tabela 2.1.

Em seguida foram encontrados os principais parâmetros estatísticos neste mesmo trabalho, referência [7], tais como média, desvio padrão, valores máximos e mínimos etc., os quais estão resumidos também na tabela 2.1.

Tabela 2.1 - Principais dados estatísticos dos ventos estudados na referência [7], Cap. [5].

	Nordeste I	Nordeste II	Sul	Sudeste
Número de estados diferentes	165	175	196	179
Média (m/s)	7,58	7,36	7,00	6,59
Mediana	7,2	6,6	6,9	6,5
Valor mínimo (m/s)	0	0,098	0	0
Valor máximo (m/s)	18,7	17,79	20	18,7
Desvio padrão	2,610	2,962	2,882	2,844

A partir da tabela 2.1 pode-se perceber que os quatro sítios analisados possuem boas velocidades médias anuais o que justifica o aproveitamento eólico destas regiões.

3. Séries Temporais, Análise Estatística e Agrupamento de Dados

Este capítulo tem por premissa expor os métodos de tratamento estatísticos de dados utilizados no presente trabalho. Estes são velocidades médias de ventos de três localidades diferentes, obtidos a partir de medições e organizados em séries chamadas temporais, foco deste projeto.

A idéia básica de uma série temporal é uma coleção de observações feitas em diferentes instantes de tempo, sujeitas a variações aleatórias. Um dos objetivos da análise das séries temporais é a construção de um modelo que explique o comportamento aleatório e determinístico da série [8].

Para a construção de um modelo representativo de séries temporais, deve haver inicialmente um tratamento de dados, para que desta forma estes possam ser organizados.

Este trabalho utiliza a análise estatística e espectral para este fim, associada à técnica de agrupamento de dados conhecida como *k-means*, a qual é utilizada para reduzir o número de estados de vento das séries, associando-os a partir de suas similaridades sem comprometimento de suas principais características.

3.1 Séries Temporais

Uma série temporal é uma coleção de observações feitas seqüencialmente através do tempo. As séries podem ser ditas contínuas ou discretas. São ditas contínuas quando seus dados são observados continuamente através do tempo, ou classificadas como discretas quando as observações são realizadas em tempos específicos, normalmente igualmente espaçados.

De acordo com a natureza do tipo de variáveis observadas, as séries temporais podem ser classificadas como:

- Discreta, quando T é um conjunto finito de pontos, $[X(t), t \in T], T = [1, 2, 3...]$.
- Contínua, quando T é um intervalo finito, $[X(t), t \in T], T = [t: 0 < t < T]$.
- Multivariada (discreta ou contínua). $[X_1(t), \dots, X_k(t), t \in T]$.
- Multidimensional, quando se tem $[X(t), \tilde{t} \in T]$, onde \tilde{t} é um vetor.

Um relevante fator para a análise das séries temporais é o fato de, normalmente, sucessivas observações não serem independentes. Assim, com valores de séries observados no passado, dados futuros podem ser previstos [9].

Caso a série possa ser prevista exatamente, diz-se que esta é determinística. Entretanto, grande parte das séries pode ser determinada apenas em parte por valores passados, estas são ditas estocásticas. Desta forma, pode-se apenas obter uma idéia do comportamento das novas séries, que passam a ser representadas por funções de densidade de probabilidade, condicionadas pelo conhecimento dos valores passados.

3.1.1 Objetivo da Análise das Séries Temporais

Há dois aspectos fundamentais a serem considerados no estudo de séries temporais, análise e modelagem. A primeira tem como objetivo sumarizar as propriedades características das séries e caracterizá-las quanto ao seu comportamento, identificando um modelo apropriado. Já a modelagem visa tornar a previsão de valores futuros possível, bem como a geração de seqüências sintéticas e prováveis [8].

Algumas características relevantes da análise e modelagem de sistemas são descritos nos tópicos relacionados a seguir [9].

- **Descrição**

O primeiro passo na análise das séries temporais é plotar os dados em função do tempo. Assim as principais propriedades das séries podem ser obtidas através de medidas descritivas simples. Para tal utiliza-se para descrever o comportamento das séries ferramentas como construção de gráficos, histogramas, diagramas de dispersão, obtenção de estatísticas descritivas, verificação de tendências, pontos de mudanças etc.

- **Explicação**

Quando as observações possuem duas ou mais variáveis, é possível usar a variação em uma das séries temporais para explicar a variação da outra série. Isto pode facilitar o entendimento do mecanismo gerador das séries de tempo. Para tal, muitas vezes são usadas técnicas de regressão linear.

- **Predição**

Consiste em prever o comportamento futuro da série temporal. Utiliza-se este tipo de análise principalmente para, a partir de valores passados, procurar periodicidade relevante dos dados, podendo assim fazer previsões de valores futuros.

- **Controle**

As séries temporais muitas vezes são coletadas e analisadas para promover melhorias em muitos sistemas, como por exemplo, sistemas físicos e econômicos. Os processos de controle variam em estilo e sofisticação, afetando assim a precisão dos sistemas. Em muitas situações, os problemas de controle estão fortemente ligados às questões de predição.

3.1.2 Variações Presentes nas Séries Temporais

Técnicas estatísticas simples, como os métodos descritivos, ou técnicas estatísticas mais sofisticadas, como as técnicas de inferência, podem ser utilizadas para análise de séries temporais aleatórias. Os métodos descritivos geralmente são utilizados antes de tentativas com procedimentos mais sofisticados. Isto se deve ao fato destes serem de extrema importância para a organização e “limpeza” dos dados, levando ao conhecimento do comportamento das séries, evitando obter modelos que as represente inadequadamente [8].

Para a correta análise das séries de tempo a primeira providência é o conhecimento do contexto de ocorrência dos dados, assegurando-se assim, que os dados apropriados foram, ou serão coletados. Por exemplo, caso as séries sejam muito curtas, ou as variáveis erradas tenham sido coletadas, isto pode impossibilitar a solução do problema [9].

Além dos usuais parâmetros estatísticos analisados em séries, como médias, moda, desvio padrão etc., alguns outros aspectos devem ser considerados. Isto se deve principalmente ao fato de componentes sistemáticos estarem inseridos neste tipo de série, o que pode comprometer a veracidade dos valores encontrados por análises estatísticas comuns.

Portanto, as séries temporais podem ser descritas de acordo com seus principais componentes sistemáticos, os quais podem ser definidos como as principais características presentes nas séries, classificadas como: variações sazonais, variações cíclicas, tendências e variações irregulares.

Os tópicos relacionados a seguir descrevem com maiores detalhes as principais fontes de variações que podem estar presentes nas séries temporais [8].

- Sazonal

As séries sazonais possuem variação periódica, como por exemplo, diariamente, semanalmente, mensalmente etc. Uma variação anual pode ser facilmente percebida, e facilmente estimada, caso a sazonalidade seja de interesse direto da análise. Alternativamente, a sazonalidade pode ser removida dos dados, gerando assim uma série dessazonalizada, caso não esta não seja de interesse de estudo.

- Cíclica

As séries cíclicas, além de possuírem características sazonais, apresentam também variações em outros períodos fixos, podendo ser devido a outras causas físicas. Pode ser citada como exemplo, a variação diária da temperatura. Outras séries que também podem ser classificadas como cíclicas são as que não possuem períodos fixos, porém podem ser previstas até determinado ponto. Neste caso podem ser citados como exemplo, ciclos de dados econômicos, que normalmente variam a cada 3 ou 4 anos, ou podem demorar até mais de 10 anos, até que ocorra a variação.

- Tendência

Pode ser vagamente definida como “efeitos a longo prazo na média”. Neste caso o fator de maior dificuldade é a definição de “longo prazo”. Por exemplo, variáveis climáticas exibem variações cíclicas por períodos maiores que 50 anos. Caso sejam avaliados dados de apenas 20 anos, parecerá tendência, porém se houver amostras de centenas de anos, ficará evidente a variação cíclica. Assim não se devem classificar como afetadas por tendência, amostras de curtos períodos, que possuam poucas observações.

- Irregular

Apenas devem ser consideradas séries temporais irregulares, após a impossibilidade de classificá-la como qualquer uma das variações descritas acima, o que gera uma série composta por resíduos que podem ser ou não aleatórios. Assim obtém-se uma série que não possui qualquer variação regular.

Desta forma, com o entendimento dos principais fatores que podem exercer influência às séries, o próximo passo é a plotagem dos dados obtidos em relação ao tempo, e assim identificar os componentes sistemáticos presentes, e tratá-las adequadamente.

3.1.3 Processos Estocásticos e Estacionários

A maioria dos processos físicos encontrados envolve elementos aleatórios em sua estrutura, os processos estocásticos, definidos como “fenômenos estatísticos desenvolvidos ao longo do tempo, de acordo com as leis probabilísticas” [9]. Assim, o diferencial entre as séries temporais e outras análises estatísticas é o reconhecimento da importância da ordem em que as observações são feitas, ou seja, a inter-relação entre os valores obtidos.

Muitas vezes os processos estocásticos também são chamados também de processos aleatórios. Alguns exemplos podem ser citados, como o comprimento de filas, número de acidentes em alguma cidade etc.

Matematicamente, um processo estocástico pode ser definido como uma coleção de variáveis aleatórias ordenadas no tempo, definidas como um conjunto de observações $[X(t), t \in T]$, as quais podem ser contínuas ou discretas. Assim para cada $t \in T$, $X(t)$ é uma variável aleatória [8].

Os processos estacionários são uma importante classe dos processos estocásticos. Uma série temporal pode ser considerada estacionária, caso não haja mudança na média (não seja afetada por nenhuma tendência), não exista mudança sistemática na variância e variações bruscas nos períodos tenham sido removidas.

Desta forma, as teorias probabilísticas das séries temporais se concentram nas séries de tempo estacionárias. Por este motivo, para analisar séries não estacionárias, primeiramente estas devem ser transformadas em séries estacionárias, retirando as tendências ou as variações sazonais presentes [9].

As séries que são apresentadas nas próximas seções, alvo de estudo no presente trabalho, são séries temporais de vento observadas em três diferentes sítios localizados na Alemanha. Sendo esta a motivação para apresentação dos conceitos de séries temporais estacionárias. Assim no capítulo 4, são apresentados estes dados, onde são plotados e analisados, conforme foi citado nas seções anteriores deste capítulo.

3.2 Análise Estatística

A estatística é uma parte da matemática aplicada que utiliza as teorias de probabilidade para explicar as frequências de eventos, tanto em estudos observados, quanto em dados obtidos experimentalmente [8].

Os dados podem ser obtidos a partir da observação humana, ou podem ser também coletados e armazenados eletronicamente. Os formatos diferem, dependendo do observador ou do processo de observação, e reflete as preferências pessoais e facilidades de registro [12].

Seu objetivo é proporcionar meios de coletar, organizar, descrever os dados e assim analisá-los, para que desta forma possam ser encontradas correlações e situações de causa e efeito. Em suma, a estatística visa modelar a aleatoriedade e a incerteza, para que assim possa estimar ou prever fenômenos futuros [8].

A parte da estatística mais utilizada no presente trabalho é a chamada *Estatística Descritiva*, na qual são utilizadas técnicas para descrever e sumarizar o conjunto de dados avaliados. As principais técnicas utilizadas são a construção de gráficos descritivos para sumarizar os dados, tais como gráficos temporais e histogramas; descrição tabular, onde se utilizam tabelas para o mesmo fim, como as tabelas de frequências e por fim são usadas ainda descrições paramétricas, como a média e outros importantes parâmetros estatísticos de determinado conjunto de dados.

A próxima seção, 3.2.1, tem por objetivo expor os principais parâmetros estatísticos que serão utilizados no trabalho. Assim, no próximo capítulo estes parâmetros serão aplicados a fim de tratar da melhor maneira os dados de velocidades de vento obtidos através de observações.

3.2.1 Descrição Numérica de Dados

A descrição numérica de dados é de grande utilidade para análise dos dados, principalmente quando associadas aos gráficos e tabelas descritivas.

Os tópicos a seguir apresentam as principais medidas numéricas, ou parâmetros estatísticos, importantes para a descrição das características dos dados [10], [12].

- **Média Aritmética**

A média aritmética é a medida mais comum de tendência central, ou de localização dos dados. Se as observações em uma amostra de tamanho N são x_1, x_2, \dots, x_n , então a média amostral é dada por:

$$\bar{x} = \frac{1}{N} \sum_i^N x_i \quad (3.1)$$

- **Mediana**

A mediana de N dados ordenados em ordem de grandeza, é o valor do ponto central, caso N seja ímpar, ou a média aritmética dos dois valores centrais, caso N seja par. Definida matematicamente por:

$$\tilde{x} = \begin{cases} x_{(\frac{n+1}{2})} & n \text{ ímpar} \\ \frac{x_{(\frac{n}{2})} + x_{(\frac{n}{2}+1)}}{2} & n \text{ par} \end{cases} \quad (3.2)$$

- **Moda**

Moda é o valor que aparece em maior frequência em um conjunto com N valores numéricos. Esta pode nem sempre existir e caso exista não ser única.

- **Variância**

A variância é a medida de dispersão mais importante. Se as observações em uma amostra de tamanho N são x_1, x_2, \dots, x_n , então a variância amostral é dada por:

$$\sigma^2 = \frac{\sum_i^N (x_i - \bar{x})^2}{N} \quad (3.3)$$

- **Desvio Padrão**

O desvio padrão mede o grau de dispersão de dados numéricos em torno de um valor médio. É dada pela raiz quadrada da variância, assim matematicamente o desvio padrão é dado por:

$$\sigma = \sqrt{\frac{\sum_i^N (x_i - \bar{x})^2}{N}} \quad (3.4)$$

- **Correlação**

É o coeficiente usado para estudo do comportamento de duas variáveis diferentes. O coeficiente de determinação, r^2 , é uma medida de proporção da variabilidade de uma variável que é explicada pela variabilidade da outra. Ou seja, é uma medida que descreve o grau de associação entre duas variáveis, X e Y , e é dada matematicamente por:

$$r = \frac{\sum_i^N (X_i - \bar{X}) \cdot \sum_i^N (Y_i - \bar{Y})}{\sqrt{\sum_i^N (X_i - \bar{X})^2 \cdot \sum_i^N (Y_i - \bar{Y})^2}} \quad (3.5)$$

O coeficiente de correlação é uma quantidade adimensional.

- **Covariância**

A covariância é uma medida estatística usada para capturar como duas séries se movem juntas. Assim como a correlação, descreve o grau de associação entre duas variáveis, X e Y . Quando a covariância é positiva significa que ambas se movem para o mesmo sentido, e quando esta é negativa significa que se movem em sentidos contrários. É dada matematicamente por:

$$\mathbf{covar} = \frac{\sum_i^N (X_i - \bar{X}) \cdot (Y_i - \bar{Y})}{n} \quad (3.6)$$

3.3 Agrupamento de Dados

De forma geral, obter a solução para o problema de clusterização (agrupamento de dados) significa agrupar elementos de um conjunto de dados, de forma que os novos grupos formados, denominados *clusters*, representem uma configuração em que cada elemento possua uma maior similaridade com qualquer elemento do mesmo cluster do que com elementos de outros clusters [11].

Os dados podem ser classificados de acordo com sua natureza como quantitativos ou qualitativos. Os primeiros podem ser valores contínuos, como a velocidade dos ventos; discretos, como número de grupos ou intervalos, como a duração de determinado evento [6].

De maneira mais formal, o problema da clusterização de dados pode ser dado por:

Dado um conjunto com n elementos $X = \{X_1, X_2, \dots, X_n\}$, a solução do problema de clusterização consiste na obtenção de um conjunto de k clusters $C = \{C_1, C_2, \dots, C_k\}$, tal que os elementos contidos em *cluster* possuam maior similaridade entre si do que os elementos de qualquer um dos demais *clusters* do conjunto C [11].

O valor de k pode ser conhecido ou não. Caso k seja fornecido para a solução, o problema é dito como “*problema de k-clusterização*”. Caso contrário, com k desconhecido, o problema é referenciado como “*problema de clusterização automática*”, e a obtenção do valor de k faz parte da solução do problema.

Em uma k -clusterização, o número total de diferentes formas de agrupamento de n elementos de um conjunto de k clusters, equivale à função, 3.7, $N(n,k)$ apresentada abaixo:

$$N(n, k) = \frac{1}{k!} \sum_{i=0}^k (-1)^i \binom{k}{i} (k-i)^n \quad (3.7)$$

Para o problema de clusterização automática o número total de combinações sofre alguns incrementos significativos, sendo definido da seguinte forma:

$$N(n) = \sum_{k=1}^n \frac{1}{k!} \sum_{i=0}^k (-1)^i \binom{k}{i} (k-i)^n \quad (3.8)$$

Um importante critério utilizado para identificar a similaridade entre dois elementos é a distância entre eles, que trabalha com as diferenças entre os valores de cada atributo dos elementos. Assim, quanto menor for a distância entre um par de elementos maior é a similaridade entre eles [11].

Existem vários tipos de medidas de proximidade, como a medida de Minkowski, mostrada abaixo:

$$d_p(x_i, x_j) = \left[\sum_{k=1}^d \|x_{i,k} - x_{j,k}\|^p \right]^{1/p} \quad (3.9)$$

Um caso especial da equação (3.9) é a distância Euclidiana, uma das principais medidas de distância utilizadas. Esta considera a distância d entre dois elementos x_i e x_j , dada por:

$$d_2(x_i, x_j) = \left[\sum_{k=1}^d (x_{i,k} - x_{j,k})^2 \right]^{1/2} \quad (3.10)$$

Outra importante medida de proximidade utilizada é distância *city-block*, a qual corresponde à soma das diferenças entre todos os p atributos de dois elementos, x_i e x_j , não sendo indicada para quando existe correlação entre tais atributos. Mostrada abaixo:

$$d(X_i, X_j) = \sum_{l=1}^p |x_{il} - x_{jl}| \quad (3.11)$$

Os problemas de clusterização podem ser divididos em dois grupos: *hard clustering*, e *soft clustering*. No primeiro caso os dados são divididos em grupos, e no segundo cada dado está associado a um grupo por um grau de pertinência [6], os quais são explicados melhor nos tópicos abaixo.

- ***Hard Clustering*** [6]

As técnicas que realizam este tipo de agrupamento de dados podem ser divididas em cinco categorias: hierárquicas, particionais, baseadas em densidades, baseadas em uma grade, baseadas num modelo.

As hierárquicas produzem um dendograma, a partir de um critério de união ou quebra, baseado na similaridade dos dados, onde cada nível é correspondente a uma partição. Ex. AGNES, DIANA.

As particionais identificam a partição dos dados, minimizando uma função objetivo, responsável pelo cálculo das distâncias entre os grupos. Ex. K-Means, ISODATA, CLARANS.

As baseadas em densidades assumem que os dados correspondem às distribuições de probabilidades, onde cada grupo é representado por uma distribuição e tem como objetivo identificar os parâmetros de cada distribuição. Ex. DBSCAN, OPTICS, EM.

As baseadas em grade utilizam dados em grades de multi-resolução. Considerados distribuições Gaussianas. Ex. STING, WaveCluster.

E as baseadas em um modelo, um modelo é associado a cada grupo, tendo como objetivo o melhor ajuste do modelo para cada grupo. Ex. SOM, AutoClass.

- ***Soft Clustering*** [6]

O *Fuzzy C-Means* é a técnica mais conhecida do *soft clustering*. Esta considera que os dados podem ser associados a mais de um grupo, a partir de graus de pertinência situados em um intervalo entre 0 e 1. Os centróides são calculados a partir destes graus. A cada nova iteração são atualizados os centróides e os graus de pertinência, até que a diferença dos valores das funções objetivos, atual e anterior, se encaixe num valor de tolerância pré- estabelecido.

Existem outras, tais como *Possibilistic C-Means*, *Fuzzy Possibilistic C-Means*, *Possibilistic Fuzzy C-Means* e *Fast Fuzzy Clustering Algorithm*.

Os passos mostrados abaixo resumem a técnica de agrupamento de dados [6].

i. Representação dos Dados

Nesta etapa deve-se escolher o número de variáveis e grupos a serem utilizados no agrupamento. Apenas devem ser usadas variáveis relevantes para tal. Dependendo do algoritmo utilizado, pode-se também extrair variáveis, a fim de gerar outras que serão utilizadas pelo algoritmo.

ii. Definição da Medida de Proximidade

Geralmente, uma função de distância é utilizada para o cálculo da distância entre os dados. Como citado anteriormente, existem vários tipos de medidas de distância, cabe ao usuário a escolha do mais apropriado.

iii. Escolha da Técnica

Compreende a seleção técnica de agrupamento a ser utilizada, e depende de fatores como: tipo de dados, objetivo do agrupamento etc.

iv. Abstração dos Resultados

Trata-se da extração da nova representação dos dados, assim cada grupo obtido é descrito de forma compacta. O centróide do grupo é uma forma bastante utilizada.

v. Avaliação dos Resultados

Segundo a referência [6], existem três formas de avaliar os resultados de um agrupamento. A primeira, a avaliação externa, a qual utiliza a informação previamente conhecida sobre os dados para avaliar o agrupamento obtido. A segunda, a avaliação interna, tem por objetivo verificar se o agrupamento encontrado se adéqua aos dados. E a terceira compara vários agrupamentos obtidos por técnicas diferentes, para o mesmo número de grupos.

O presente trabalho tem como um dos objetivos, a diminuição do número de estados das três séries de vento estudadas. Adicionalmente à determinação dos parâmetros estatísticos das séries, para complementação das análises requeridas. Para tal, é utilizada a técnica de agrupamento de dados conhecida como *K-Means*, uma das técnicas mais simples e rápidas.

A próxima seção, 3.3.1, fornece os principais conceitos e utilizações desta técnica.

3.3.1 Técnica de Agrupamento de Dados - *K-Means*

O método de clusterização utilizado neste trabalho é o *k-means*, conforme citado anteriormente. Este utiliza um método de particionamento, onde o conjunto de dados é dividido em k subconjuntos, com um valor de k conhecido. No *k-means*, o elemento representativo do *cluster*, é o seu centróide, que possui um valor médio para os atributos considerados, relativos a todos os elementos do *cluster* [11].

Esta técnica minimiza a seguinte função objetivo:

$$J = \sum_{j=1}^k \sum_{i=1}^n \|x_i^j - c_j\|^2 \quad (3.12)$$

O valor encontrado pela função objetivo representada pela equação (3.12) é o somatório entre as distâncias dos dados e os centróides representativos de seus grupos. A minimização desta função significa aumentar o grau de similaridade existente em cada grupo.

Após a definição da localização dos centróides representativos, cada dado da série original é associado ao seu centróide mais próximo. Quando este procedimento já estiver sido feito com todos os dados, os centróides são recalculados. Isto é feito repetidas vezes, até que o centróide obtido seja igual ao da iteração anterior.

O algoritmo do *k-means* segue os seguintes passos [6]:

- i. Escolha dos centróides iniciais
- ii. Associação de cada dado ao centróide mais próximo. Cada dado só pode ser associado a um centróide.
- iii. Atualização da posição dos centróides, através do cálculo da média dos dados associados a cada grupo.
- iv. Comparação entre os centróides desta iteração com a iteração anterior, caso sejam diferentes, volta-se ao segundo passo até que não haja mais modificação nos centróides.

As principais vantagens na utilização desta técnica são a simplicidade e sua rapidez, características que tornam fácil sua aplicação no agrupamento de dados, sendo apropriada para utilização de grandes bases de dados.

Pode ser citada como desvantagem a dependência do número de grupos a ser escolhidos, o k . Além disso, a qualidade do agrupamento é dependente da inicialização dos centróides. Uma medida para minimizar esses problemas, é executar o algoritmo com diferentes tamanhos de k e inicializações [6].

No caso de séries temporais de vento, de acordo com a referência [14], uma boa aproximação é agrupar as séries em cinco grupos, com $k=5$. Assim, cada dado da série de velocidade de vento ficará associado por similaridade a um dos cinco centróides encontrados.

O k-means neste trabalho é uma técnica bem apropriada, visto que as séries temporais em estudo neste trabalho possuem uma extensa base de dados. Outro aspecto relevante para sua utilização é o de não haver o inconveniente de não se conhecer uma boa aproximação para o número de grupos, visto que para o caso de séries de vento $k=5$ é uma boa aproximação [14]. Portanto as novas séries de vento obtidas terão 5 centróides representativos.

4. Estudos das Séries Temporais de Vento

Conforme já foi citado anteriormente, o objetivo deste trabalho é estudar o comportamento de determinadas séries temporais de vento, para conhecer o potencial eólico das localidades referentes às bases de dados e a viabilidade de instalação de usinas eólicas. A base de dados é composta por velocidades médias de vento obtidas em três sítios localizados na Alemanha, são eles: Falkenhagen, Wittstock e Jabel [13]. Estes dados foram obtidos por medições realizadas em intervalos de 15.

Assim, este capítulo aplica os conceitos vistos no capítulo 3. Primeiramente é feita uma descrição dos dados, a partir de gráficos temporais e histogramas. Estes são gerados em períodos anuais e mensais com maiores potenciais eólicos, para cada uma das três localidades. A partir destas informações, suas principais características são resumidas em tabelas, mostradas ao longo deste capítulo.

Em seguida, no intuito de diminuir o número de estados de vento das séries, utiliza-se o agrupamento de dados estudado no capítulo 3, o *k-means*, e desta forma são encontrados novos cinco estados representativos da série original. A partir disso, cada dado da série original é associado ao seu centróide mais próximo, gerando uma nova série, da qual são analisados os mesmos parâmetros estatísticos das séries originais. São encontrados também os erros absolutos entre as séries geradas e as séries originais, obtidos a partir da diferença de seus dados.

A partir de uma rotina criada em Excel, são encontradas as probabilidades individuais e correlacionadas dos novos estados de vento. Estas representam respectivamente as probabilidades de ocorrência de cada novo estado gerado dentro de sua própria série, e as probabilidades de um novo estado de uma série, coincidir com determinado novo estado de outra série. Para tal, são utilizadas curvas de distribuição de probabilidade acumulada (*CDF*), obtendo-se assim *CDF*'s individuais e correlacionadas, as quais são de grande utilidade nos estudos de viabilidade de instalação de usinas eólicas.

Por fim, as principais conclusões a respeito dos ventos destas localidades e da viabilidade de instalação de usinas eólicas nestes sítios, são apresentadas.

4.1 Gráficos e Histogramas

A base de dados é composta por três séries de velocidade de vento coletadas em localidades distintas situadas na Alemanha, são elas: Falkenhagen, Wittstock e Jabel. As medições foram realizadas em intervalos de 15 minutos. Apenas as duas primeiras possuem dados por um período de um ano, a última é composta por dados observados em apenas seis meses [13]. Assim, os gráficos 4.1 a 4.6, mostram os gráficos gerados pelo MatLab, a partir da base de dados das três localidades.

As figuras 4.1 e 4.3 reproduzem o comportamento anual do vento nas localidades de Falkenhagen e Wittstock. A figura 4.5 representa o comportamento dos ventos em Jabel, com dados observados por um período de aproximadamente seis meses. E logo em seguida a cada gráfico temporal, são mostrados seus histogramas, nas figuras 4.2, 4.4, 4.6.

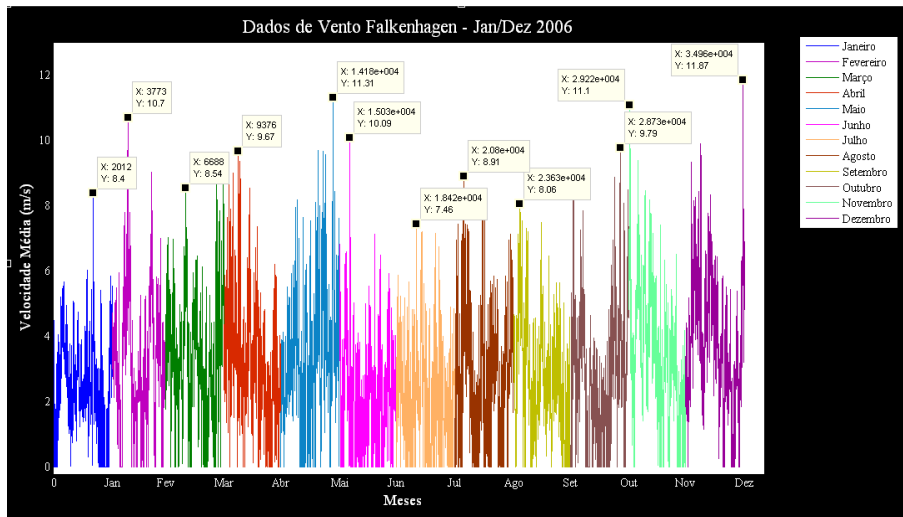


Figura 4.1 - Série de ventos Falkenhagen Jan/Dez 2006.

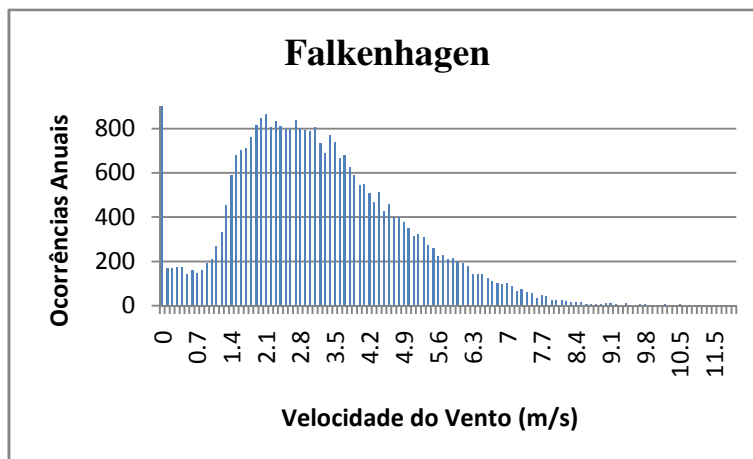


Figura 4.2 - Histograma Falkenhagen Jan/Dez 2006.

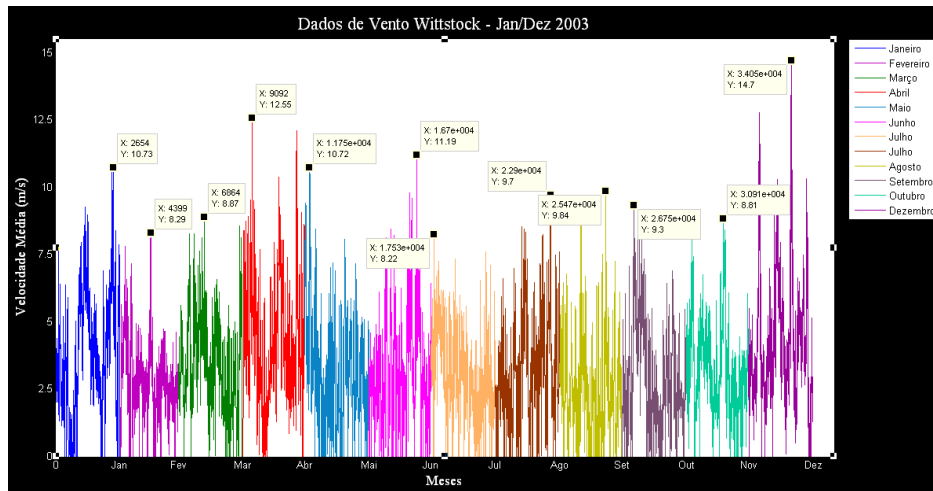


Figura 4.3 - Série de ventos Wittstock Jan/Dez 2003.

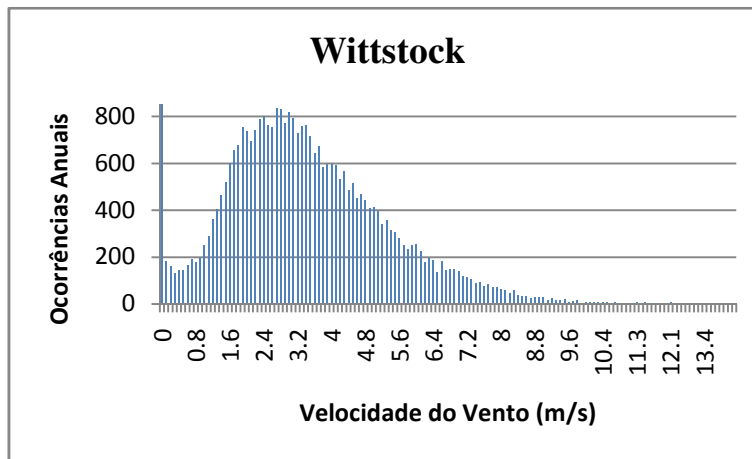


Figura 4.4 - Histograma Wittstock Jan/Dez 2003.

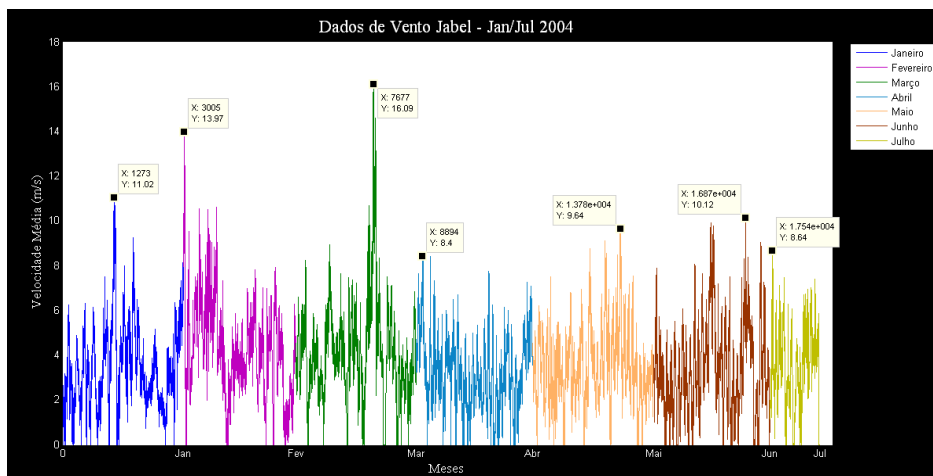


Figura 4.5 - Série de ventos Jabel Jan/Jul 2004.

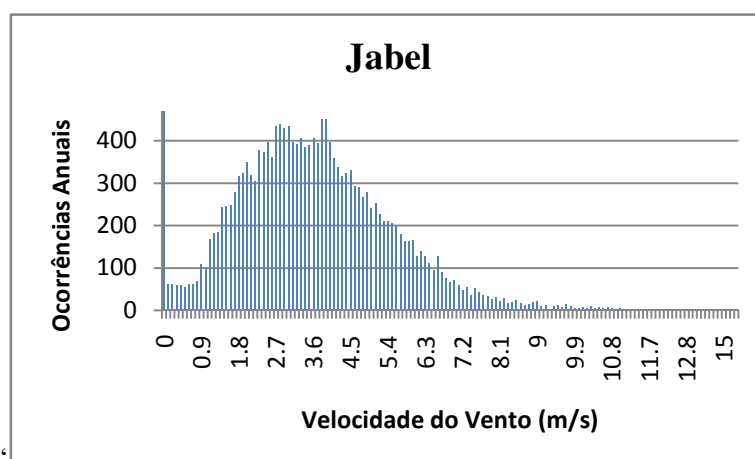


Figura 4.6 - Histograma Jabel Jan/Jul 2004.

Analisando-se os gráficos, percebe-se que duas das três localidades, Falkenhagen e Wittstock, possuem suas maiores velocidades médias em dezembro. Entretanto, Jabel possui suas maiores velocidades médias em março. Há uma particularidade com os dados desta última, devido à falta das observações no restante do ano, no período de julho a dezembro. Desta forma, a comparação direta dos dados estatísticos de Jabel com as outras duas localidades, não parece o mais adequado.

A tabela 4.1 abaixo resume os principais parâmetros estatísticos, os quais identificam o comportamento do vento em cada localidade.

Tabela 4.1- Dados estatísticos das séries de velocidade de vento em m/s.

	Falkenhagen	Wittstock	Jabel
Velocidade Mínima	0	0	0
Velocidade Máxima	14,7	11,87	16,09
Média	2,97	2,975	3,584
Mediana	3,78	2,84	3,44
Desvio Padrão	2,26	1,823	1,93
Variância	5,11	3,323	3,725
Número de Medições	35041	35041	18721

Os valores encontrados com base nas análises estatísticas de média demonstram nos três casos valores bem baixos, apesar de possuírem picos bem altos em determinadas épocas do ano. Como citado anteriormente, para que haja viabilidade e bom aproveitamento dos ventos para geração de energia elétrica, sua velocidade média mínima deve estar entre valores que variem de 7 m/s a 8 m/s, o que não ocorre em nenhum dos três casos.

4.1.2 Meses com melhores médias de velocidade

As tabelas 4.2, 4.3 e 4.4, resumem os dados anuais de Falkenhagen e Wittstock, e os dados de janeiro a julho de Jabel. E abaixo destes dados, podem-se verificar os respectivos gráficos dos meses com melhores velocidades médias, nas figuras 4.7, 4.9 e 4.11, seguidos de seus histogramas, nas figuras 4.8, 4.10, 4.12.

Tabela 4.2- Falkenhagen anual em m/s.

	Jan	Fev	Mar	Abr	Mai	Jun	Jul	Ago	Set	Out	Nov	Dez
Velocidade Máxima	8,4	10,7	9,43	9,67	11,31	10,09	7,46	8,91	8,06	9,79	11,1	11,87
Média	2,42	3,14	3,33	3,24	3,43	2,21	2,25	2,79	2,78	2,82	3,47	3,84
Mediana	2,35	3,03	3,36	3,03	3,30	2,15	2,11	2,82	2,64	2,68	3,30	3,63
Variância	1,90	3,78	3,12	2,87	3,68	2,18	2,16	3,31	2,91	3,68	3,81	3,59
Desvio Padrão	1,38	1,95	1,77	1,70	1,92	1,48	1,47	1,82	1,71	1,92	1,95	1,90

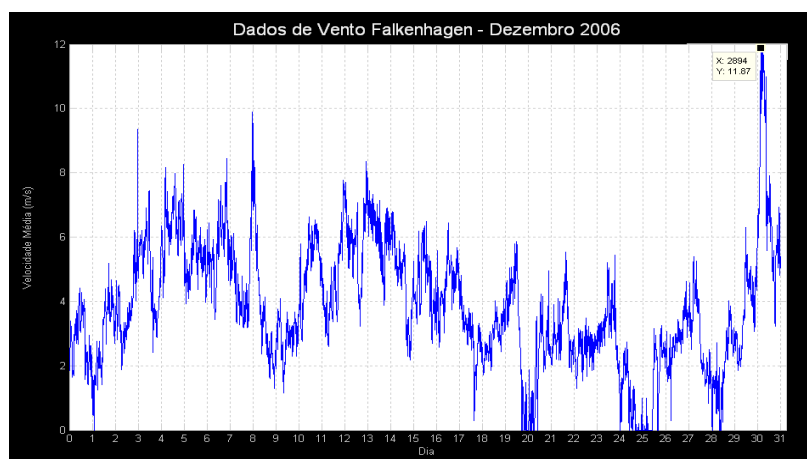


Figura 4.7 - Série de ventos Falkenhagen - Dez 2006.

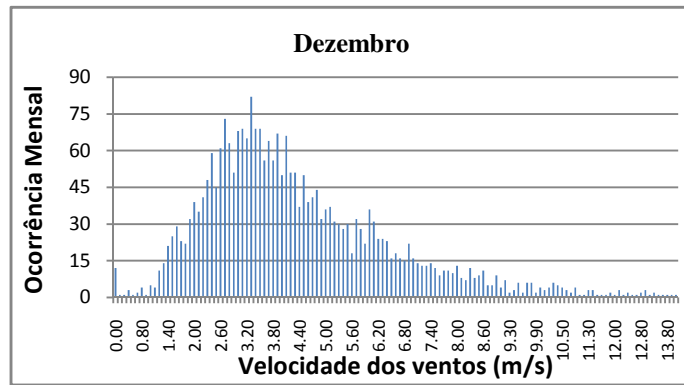


Figura 4.8 - Histograma Falkenhagen Dez - 2006.

Tabela 4.3- Wittstock anual em m/s.

	Jan	Fev	Mar	Abr	Mai	Jun	Jul	Ago	Set	Out	Nov	Dez
Velocidade Máxima	10,73	8,29	8,87	12,55	10,72	10,09	8,22	9,70	9,84	9,30	8,81	14,7
Média	3,78	2,77	3,40	4,25	3,10	2,21	2,94	3,15	2,51	2,85	3,22	4,22
Mediana	3,67	2,70	3,28	3,98	2,96	2,15	2,96	2,95	2,44	2,53	3,05	3,78
Variância	4,79	1,80	3,00	4,77	3,79	2,18	2,50	3,68	3,26	3,69	2,76	5,11
Desvio Padrão	2,19	1,34	1,73	2,18	1,95	1,48	1,58	1,92	1,81	1,92	1,66	2,26

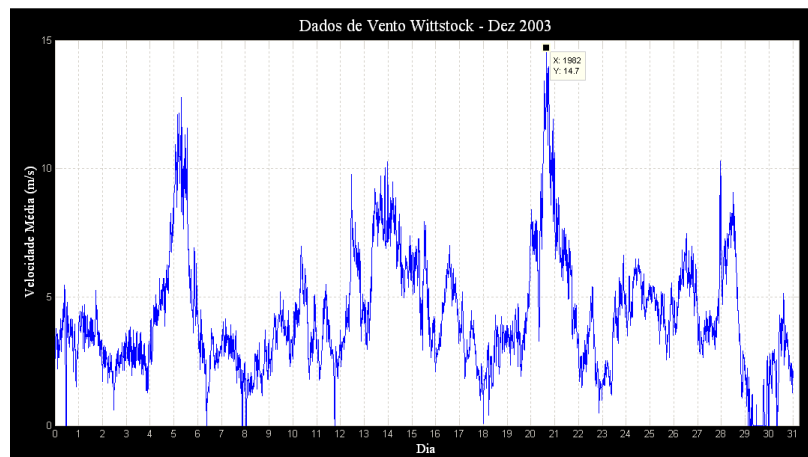


Figura 4.9 - Série de ventos Wittstock – Dez 2003.

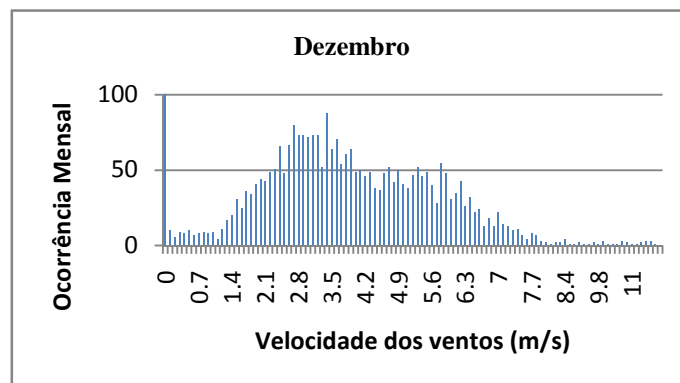


Figura 4.10 - Histograma Wittstock Dez - 2003.

Tabela 4.4 – Jabel Jan/Jul em m/s.

	Jan	Fev	Mar	Abr	Mai	Jun	Jul
Velocidade Máxima	11,02	13,97	16,09	8,40	9,64	10,12	8,64
Média	3,37	4,22	4,11	3,11	3,45	3,34	3,38
Mediana	3,08	4,04	3,89	3,05	3,31	3,08	3,52
Variância	3,35	4,34	4,60	2,52	2,76	4,06	3,04
Desvio Padrão	1,83	2,08	2,15	1,59	1,66	2,02	1,74

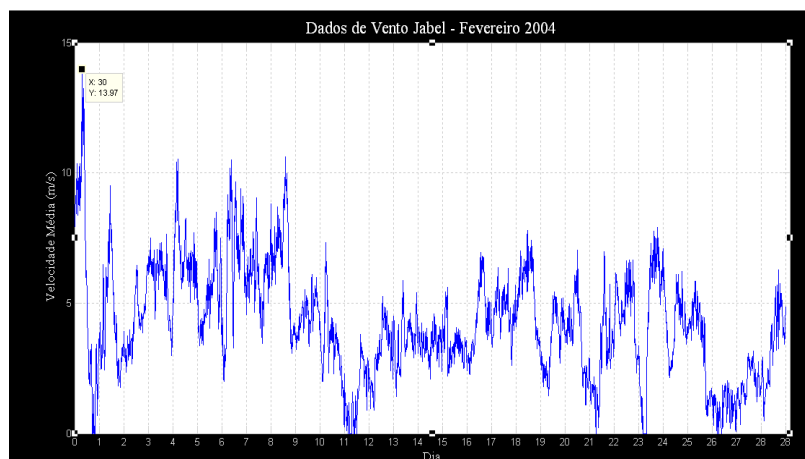


Figura 4.11 - Série de ventos Jabel - Dez 2004.

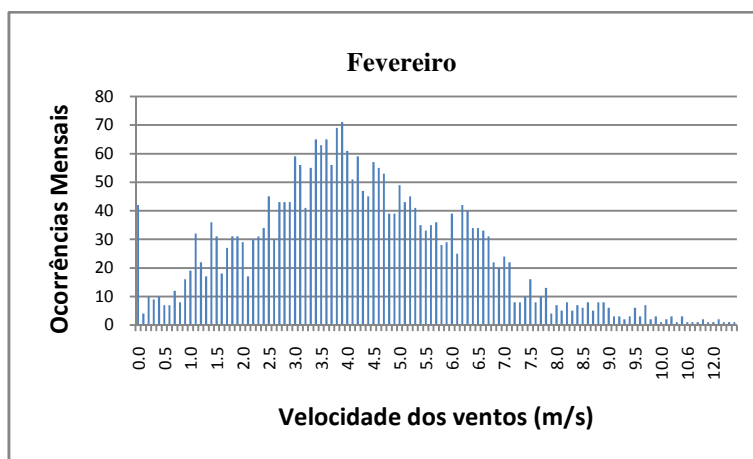


Figura 4.12 - Série de ventos Falkenhagen - Dez 2006.

A partir dos dados obtidos e organizados nas tabelas 4.2, 4.3 e 4.4, mais uma análise pode ser feita, porém agora mensal. Os dados demonstraram que apesar de meses como dezembro, para o caso de Falkenhagen e Wittstock, e fevereiro para Jabel, possuírem maiores velocidades, estas não variam muito ao longo do ano.

Em Wittstock e Falkenhagen, os meses com maiores velocidades médias de vento, também são os meses que possuem a velocidade máxima anual, porém o mesmo não ocorre em Jabel. Nesta, o mês com maior velocidade média mensal é fevereiro, mas o pico de velocidade máxima se dá em março, o que significa que seu desvio padrão de março é maior do que o de fevereiro. Apesar do pico de velocidade ocorrer em março, as velocidades de vento variaram mais neste período. Desta forma, como ambos os meses possuem velocidades máximas altas e próximas, uma maior variabilidade das velocidades fez com que a média em fevereiro fosse maior, mesmo que este mês não possuísse a velocidade máxima de pico.

Segue uma tabela resumo com os dados referentes às figuras 4.7, 4.9 e 4.11, a fim de sintetizar o texto anterior.

Tabela 4.5 - Resumo dos meses com melhores velocidades médias.

	Dezembro	Dezembro	Fevereiro
	Falkenhagen	Wittstock	Jabel
Velocidade Mínima	0	0	0
Velocidade Máxima	11,87	14,7	13,97
Média	3,84	4,22	4,22
Mediana	3,63	3,78	4,04
Moda	0	0	0
Variância	3,592	5,11	4,338
Desvio Padrão	1,895	2,26	2,083

Mais uma vez deve ser colocado que os ventos medidos nestas localidades podem ser usados para geração de energia elétrica, pois possuem velocidades médias acima de 2,5m/s ou 3 m/s. Porém, observa-se uma variabilidade das velocidades dos ventos, verificadas em seus desvios padrão, os quais são significativos incrementos nos valores de média das velocidades. Desta forma, a geração de energia elétrica a partir destes ventos esbarraria apenas no problema da viabilidade econômica, já que para que isso ocorresse os ventos deveriam possuir velocidades médias mínimas de 7m/s.

A análise de correlação entre duas séries, também é um importante parâmetro a ser extraído dos dados, pois descreve o grau de associação entre duas variáveis. Assim para verificar o comportamento dos ventos entre duas localidades será utilizada a correlação, que deve possuir o mesmo número de dados. No entanto, duas análises de correlação são feitas, com horizontes de tempo distintos, a primeira entre Falkenhagen e Wittstock, utilizando horizonte anual, e a segunda entre Falkenhagen ou Wittstock e Jabel, com horizonte temporal de seis meses.

Desta forma, novos gráficos foram gerados, cada um com duas das séries estudadas sobrepostas, que se seguem nas figuras 4.13, 4.14 e 4.15.

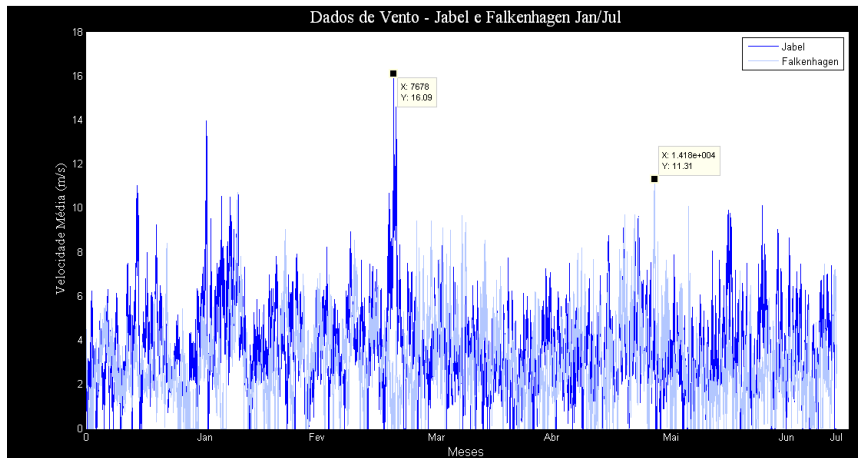


Figura 4.13 - Série de dados de Jabel e Falkenhagen sobrepostas.

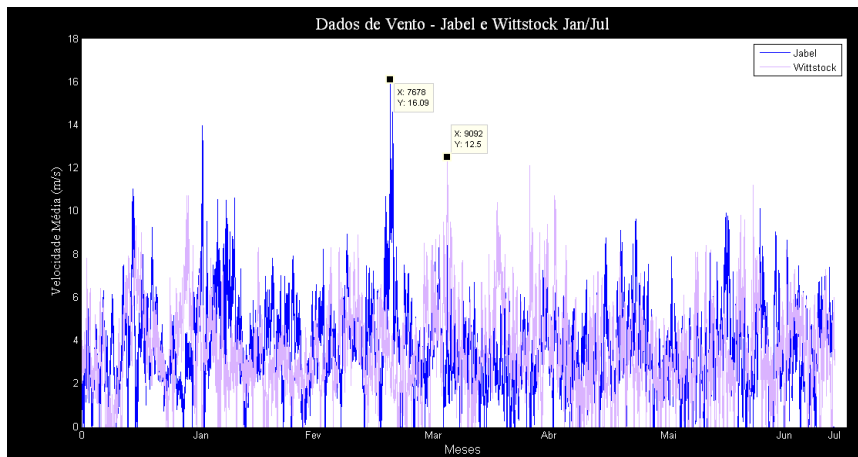


Figura 4.14 - Série de dados de Jabel e Wittstock sobrepostas.

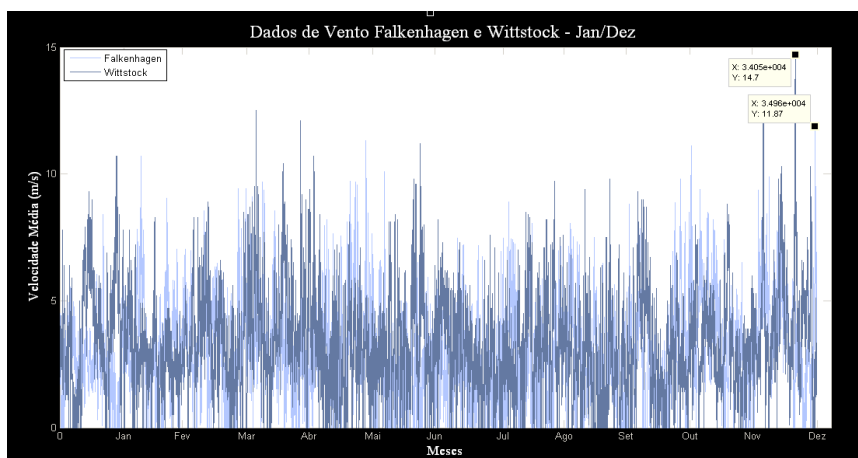


Figura 4.15 - Série de dados de Falkenhagen e Wittstock sobrepostas.

A tabela 4.6, resume os dados de correlação entre os dados das séries.

Tabela 4.6 – Correlações entre os dados das séries.

	Correlação dos Dados
Falkenhagen x Wittstock	0,1245
Wittstock x Jabel	0,1063
Falkenhagen x Jabel	0,1585

A tabela 4.6 mostra que todas as séries estão pouco relacionadas entre si, o que significa que seu comportamento não é muito parecido. As correlações foram feitas utilizando-se dados coletados em suas respectivas localidades, porém em anos diferentes. Assim acredita-se que, caso os dados fossem obtidos em um mesmo ano, os resultados poderiam ser um pouco diferentes, visto que existem variações climáticas ano a ano.

Os mais correlatos são os dados de Falkenhagen e Jabel. Porém, a análise dos dados de correlação com Jabel é restrita, visto que além dos anos serem diferentes, a correlação não é feita em dados de um período grande de tempo, apenas por aproximadamente seis meses.

Assim, correlacionar Falkenhagen ou Wittstock com Jabel não é um bom parâmetro de comparação, e o principal dado a ser utilizado em simulações futuras, será a correlação encontrada em Falkenhagen e Wittstock, que apesar de baixa ainda pode ser utilizada.

Como apenas Falkenhagen e Wittstock são séries obtidas a partir de medições de um ano completo, e Jabel por aproximadamente seis meses, as duas primeiras farão parte de um novo estudo, a fim de obter a partir de agrupamento de dados uma nova série que as represente.

4.2 Agrupamento dos Estados de Vento de Falkenhagen e Wittstock

A técnica de agrupamento de dados conhecida como *k-means*, é utilizada no presente trabalho com o objetivo de reduzir o número de estados de vento necessários para a modelagem do comportamento do vento, sem que haja perda das suas características originais.

Como explicado no capítulo 3, é necessário estabelecer o número k que se deseja obter de estados de vento, chamados neste tipo de agrupamento de centróides, para obter a nova série agrupada posteriormente. Para este tipo de série temporal de velocidade de vento, é razoável a divisão da série em cinco grupos, portanto um $k = 5$ [14].

Para Falkenhagen e Wittstock, os centróides representativos obtidos a partir do agrupamento de dados *k-means*, são mostrados na tabela 4.7:

Tabela 4.7 – Centróides representativos das séries.

Grupos	Centróides	
	Falkenhagen (m/s)	Wittstock (m/s)
1	0,20	0,26
2	1,90	1,91
3	3,22	3,26
4	4,70	4,89
5	6,70	7,28

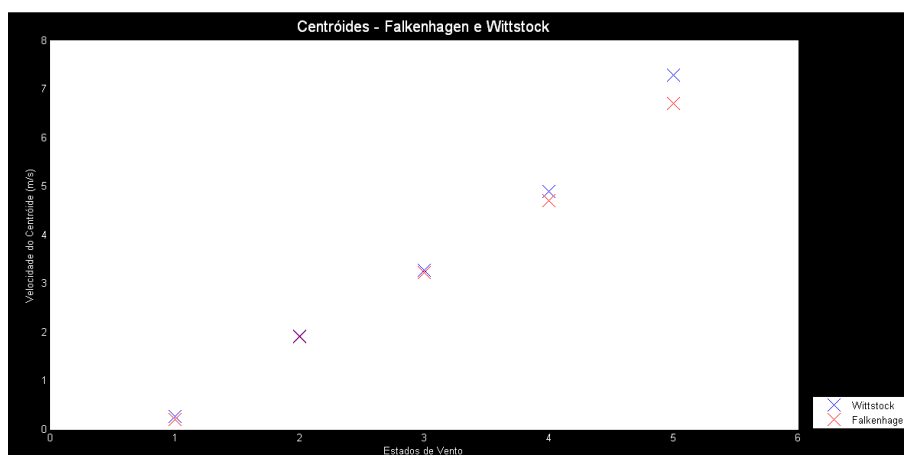


Figura 4.16 - Centróides representativos das séries a serem obtidas por agrupamento.

Após obtenção dos centróides de cada série de vento pelo MatLab, cada dado é associado ao seu centróide mais próximo, e novas séries de dados podem então ser plotadas, como mostram as figuras 4.17 e 4.18. Para melhor visualização, as séries agrupadas foram geradas em sobreposição às suas séries originais, para que assim sejam mais facilmente comparadas.

As tabelas 4.8 e 4.9 resumem as principais medidas estatísticas da série agrupada de Falkenhagen e Wittstock.

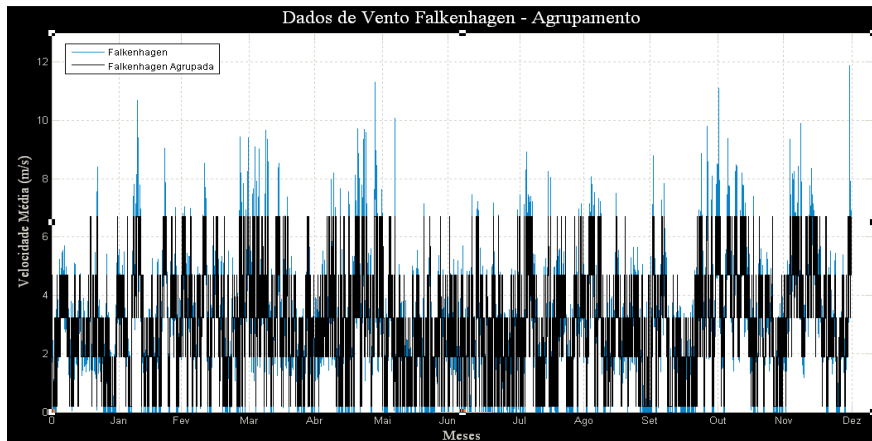


Figura 4.17 - Série de dados Falkenhagen agrupada.

Tabela 4.8 - Dados estatísticos da série agrupada - Falkenhagen.

Eixo Velocidade (m/s)	
Velocidade Mínima	0,20
Velocidade Máxima	6,70
Média	2,98
Mediana	3,22
Moda	1,90
Variância	3,10
Desvio Padrão	1,76

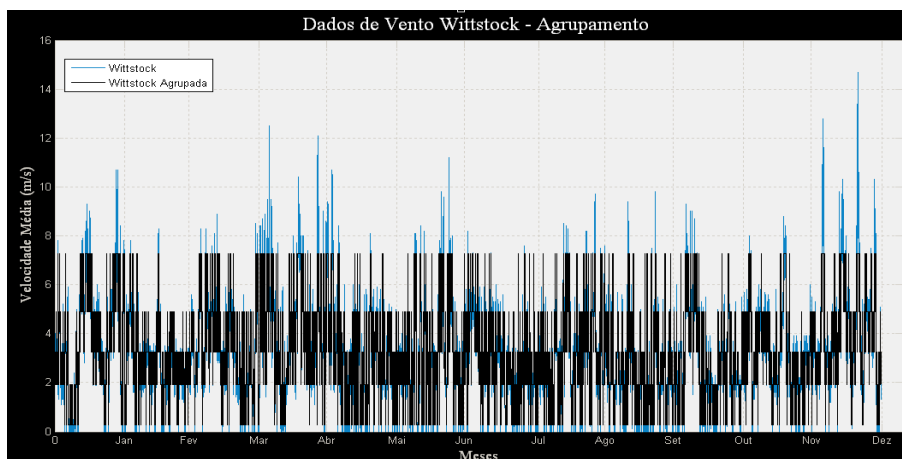


Figura 4.18 - Série de dados de Wittstock agrupada.

Tabela 4.9 - Dados estatísticos da série agrupada – Wittstock.

Eixo Velocidade (m/s)	
Velocidade Mínima	0,26
Velocidade Máxima	7,28
Média	3,28
Mediana	3,26
Moda	3,26
Variância	3,55
Desvio Padrão	1,88

Os gráficos 4.17 e 4.18, mostram que as novas séries geradas, tanto para Falkenhagen quanto para Wittstock, são boas representações, visto que o intuito neste estudo é diminuir o número de estados de vento. Assim, posteriormente, seria possível obter um modelo, por exemplo, de cálculo de potência gerada a partir da velocidade desses ventos.

As modificações nos valores de suas médias anuais foram pequenas, a de Falkenhagen praticamente se manteve, e a Wittstock foi aumentada. Uma alternativa para aproximar ainda mais os parâmetros encontrados nas novas séries dos valores reais, seria aumentar o número de estados representativos, o k da técnica de agrupamento do *k-means*. Desta forma, para que as velocidades mais altas que aparecem nas séries ficassem representadas por um centróide com valor maior de velocidade de vento diminuindo assim o desvio padrão e mantendo as médias mais próximas da realidade. As figuras 4.19 e 4.20 mostram a proporção dos centróides encontrados em cada uma das séries formadas pelo agrupamento dos dados.

Proporção dos Centróides no Agrupamento - Falkenhagen

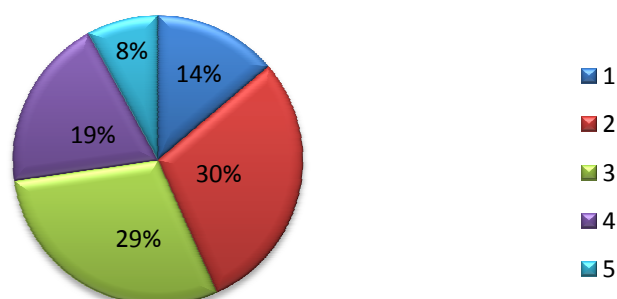


Figura 4.19 - Porcentagem dos centróides nas séries agrupadas – Falkenhagen.

Proporção dos Centróides no Agrupamento - Wittstock

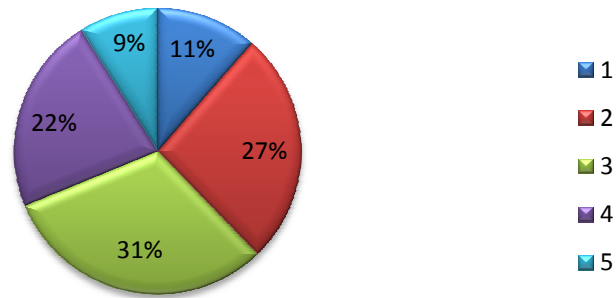


Figura 4.20 - Porcentagem dos centróides nas séries agrupadas - Wittstock.

4.2.1 Erros Absolutos – Falkenhagen e Wittstock

Os erros absolutos entre as novas séries e as séries originais, também é um importante parâmetro a ser considerado. Estes são calculados a partir da diferença entre as velocidades reais e as velocidades representativas que formaram as novas séries agrupadas.

Este erro absoluto entre a série real e a série agrupada é dado da seguinte forma:

$$Erro\ absoluto = \sum_{i=1}^n \|x_i - y_i\| \quad (4.1)$$

sendo n , o número de realizações da série temporal e $\|x_i - y_i\|$ o módulo da diferença entre a velocidade x_i e y_i da série agrupada.

O primeiro erro absoluto a ser plotado é de Falkenhagen, como mostra a figura 4.21.

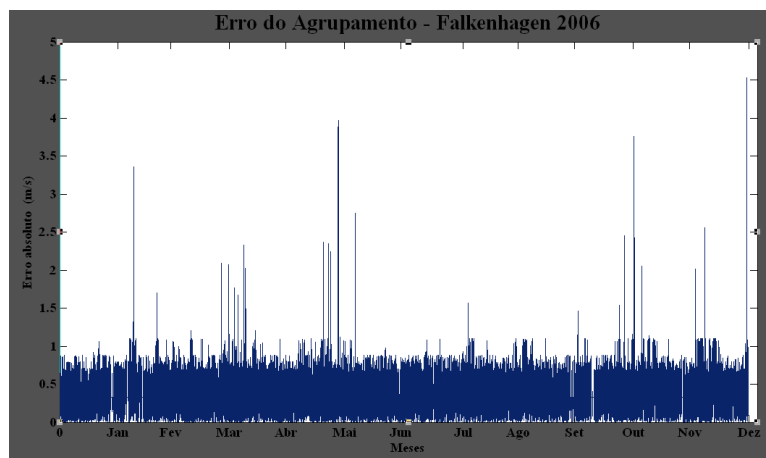


Figura 4.21 - Erro do agrupamento de Falkenhagen.

Na tabela 4.10 estão resumidos os principais parâmetros estatísticos do erro absoluto de Falkenhagen..

Tabela 4.10 – Principais parâmetros do erro de agrupamento de Falkenhagen.

Eixo Velocidade (m/s)	
Erro Mínimo	0,0005
Erro Máximo	4,531
Erro Médio	0,402
Mediana	0,347
Moda	0,329
Desvio Padrão	0,270

A seguir é mostrado na figura 4.22, o erro absoluto da outra série estudada, Wittstock.

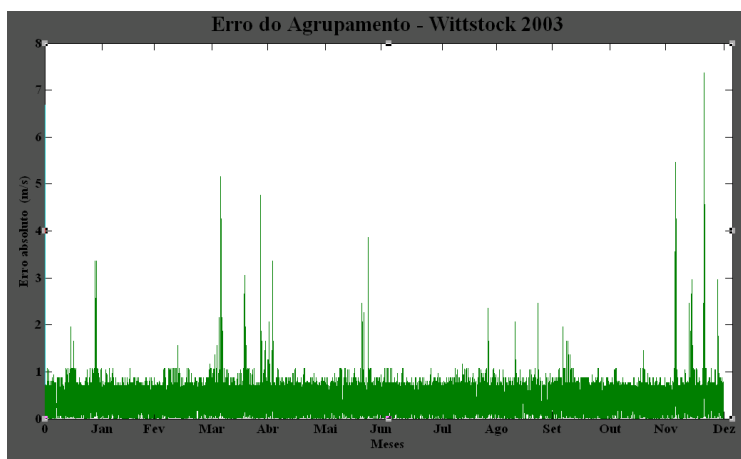


Figura 4.22 - Erro do agrupamento de Wittstock.

Seguido ao erro absoluto de Wittstock, na tabela 4.9, estão seus principais parâmetros estatísticos.

Tabela 4.11 - Principais parâmetros do erro de agrupamento de Wittstock.

Eixo Velocidade (m/s)	
Erro Mínimo	0,0031
Erro Máximo	7,361
Erro Médio	0,4261
Mediana	0,3767
Moda	0,3294
Desvio Padrão	0,3385

A partir dos dados estatísticos observados para Falkenhagen, na tabela 4.10, e para Wittstock na tabela 4.11, pode-se avaliar mais uma vez a eficiência dos valores encontrados pelo *k-means*, para as novas séries representativas.

Os erros médios de Falkenhagen e Wittstock são bem pequenos, ratificando a idéia colocada no fim da seção 4.2, de que os centróides obtidos pela técnica de agrupamento do *k-means* são boas aproximações. O desvio padrão para Falkenhagen é de 0,27 m/s e de Wittstock de 0,3385 m/s, valores bem baixos, demonstrando que os valores dos centróides ficam bem próximos aos valores das velocidades reais.

Os dados estatísticos também demonstram erros máximos elevados, os quais poderiam ser diminuídos caso a quantidade de centróides representativos fosse aumentada. Desta forma, a precisão do agrupamento seria aumentada, e os valores finais poderiam ser agrupados em torno de algum centróide de maior valor de velocidade, diminuindo o erro máximo.

4.3 Probabilidades de Ocorrência dos Estados de Vento

O próximo e último estudo tem por finalidade conhecer as ocorrências dos estados dos ventos. Primeiro são estudadas as frequências e probabilidades de ocorrências dos estados individualmente, ou seja, dentro de sua própria série. Depois é analisado o comportamento correlacionado das séries, verificando a probabilidade de ocorrência dos estados entre séries distintas.

Para tal, são avaliadas as curvas de distribuição de probabilidades acumuladas (CDF-cumulative distribution functions). Em seguida, as probabilidades de ocorrência dos estados de vento entre as duas séries são analisadas. Por exemplo, a probabilidade de ocorrência de uma velocidade representada pelo primeiro estado de Falkenhagen coincidir com o primeiro estado de Wittstock, e assim por diante.

A utilização das CDF's individuais e correlacionadas é de grande importância nas análises de viabilidade de instalação de usinas eólicas, pois permitem conhecer a probabilidade da potência gerada na usina ser menor que um determinado valor P_i de potência, podendo ser usadas em estudos de confiabilidade baseado em simulações Monte Carlo.

Assim as distribuições de probabilidades individuais e correlacionadas representam as probabilidades de ocorrência de cada estado de vento das localidades e as probabilidades de coincidências dos estados de vento dos dois sítios respectivos.

Desta forma foram encontradas as probabilidades de ocorrência dos estados de vento das localidades individualmente, ou seja, primeiro os estados de Falkenhagen e posteriormente os de Wittstock. Para tal, como já foi mencionado, utiliza-se um macro em Excel, que gera o número de frequência de cada estado de vento, proporcionando então que as probabilidades de ocorrência sejam encontradas.

Os estados de vento representativos já foram encontrados na seção 4.2, e são resumidos novamente na tabela 4.12:

Tabela 4.12 - Centróides representativos das séries.

Grupos	Centróides	
	Falkenhagen (m/s)	Wittstock (m/s)
1	0,20	0,26
2	1,90	1,91
3	3,22	3,26
4	4,70	4,89
5	6,70	7,28

4.3.1 Freqüências e Probabilidades Independentes – Falkenhagen

Assim as freqüências de ocorrência dos estados na primeira série, Falkenhagen, são mostradas na tabela 4.13, e suas probabilidades na tabela 4.14.

Tabela 4.13 – Freqüência de ocorrência dos estados - Falkenhagen

Estado	Freqüência
1	4826
2	10360
3	10299
4	6763
5	2792

Tabela 4.14 - Probabilidade de ocorrência dos estados - Falkenhagen

Estado	Probabilidade (%)
1	13,77
2	29,57
3	29,39
4	19,30
5	7,97

De posse desses valores, pode ser gerada a curva de distribuição acumulada independente de Falkenhagen, CDF, que descreve a distribuição da probabilidade de uma variável aleatória de valor real qualquer, como é o caso dos valores de freqüência encontrados anteriormente.

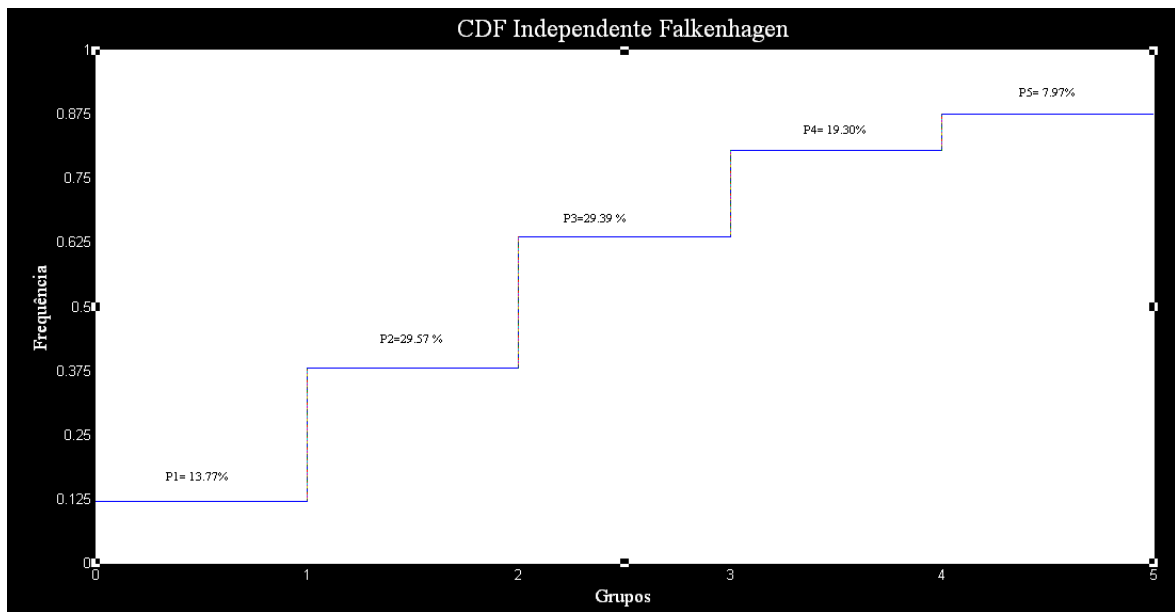


Figura 4.23 - CDF independente – Falkenhagen.

Em Falkenhagen, o estado de maior ocorrência na nova série obtida por agrupamento, é o segundo, 1,90 m/s, aparecendo 10360 vezes, dos 35040 dados, o que totaliza 29,57%. Estes dados vêm reforçar a idéia pressuposta na seção 4.1.2 de não viabilidade econômica na instalação de uma usina eólica neste local.

Na análise da nova série, representada pelos centróides associados aos dados da série original, verifica-se que a média anual de Falkenhagen é praticamente a mesma, com valor de 2,98 m/s, enquanto a série original possui velocidade média anual de 2,97 m/s. O que reforça a idéia de que os centróides representativos são boas aproximações.

4.3.2 Freqüências e Probabilidades Independentes - Wittstock

O mesmo procedimento pode ser aplicado ao estudo das freqüências e probabilidades de ocorrência dos estados para a segunda série estudada, Wittstock. As tabelas 4.15 e 4.16 trazem respectivamente esses dados.

Tabela 4.15 - Freqüência de ocorrência dos estados - Wittstock

Estado	Freqüência
1	3984
2	9255
3	10857
4	7812
5	3132

Tabela 4.16- Probabilidade de ocorrência dos estados - Wittstock

Estado	Probabilidade (%)
1	11,37
2	26,41
3	30,98
4	22,30
5	8,94

A figura 4.24 refere-se à CDF dos valores de probabilidades encontrados na tabela 4.14.

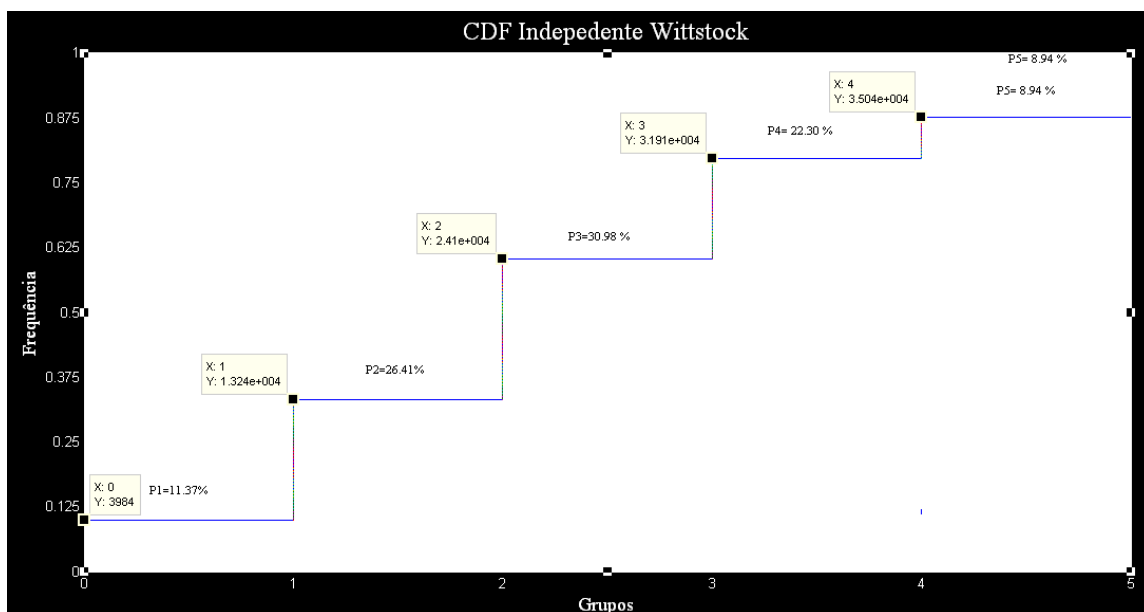


Figura 4.24 - CDF independente – Wittstock.

Em Wittstock, o estado de maior ocorrência na nova série obtida por agrupamento, é o terceiro, 3,26 m/s, aparecendo 10857 vezes, dos 35040 dados, o que totaliza 30,98 %. Estes dados vêm reforçar a idéia pressuposta na seção 4.1.2 de não viabilidade econômica na instalação de uma usina eólica neste local.

Na análise da nova série, representada pelos centróides associados aos dados da série original, verifica-se que a média anual de Wittstock obteve um relativo aumento, com valor de 3,28 m/s, enquanto a série original possui velocidade média anual de 2,978 m/s.

4.3.3 Freqüências e Probabilidades Relacionadas – Falkenhagen x Wittstock

Inicialmente foram avaliadas as probabilidades de ocorrência de determinados estados de vento em sua própria série, agora o objetivo é estudar a correlação entre os estados das duas séries de vento. Um exemplo disto seria a avaliação de qual a probabilidade de um determinado estado de vento de Falkenhagen coincidir com determinado estado de Wittstock, ou o inverso, realizado com raciocínio análogo.

Assim a tabela 4.17 em formato de matriz, resume freqüências de acontecimentos dos estados de vento utilizando a primeira referência em Falkenhagen, e a tabela 4.18, estas probabilidades.

Tabela 4.17 – Freqüência de interação dos estados de vento Falkenhagen e Wittstock.

Estados Falkenhagen	Estados Wittstock				
	1	2	3	4	5
1	848	1557	1215	832	374
2	1450	2870	3227	2019	794
3	1001	2670	3195	2511	922
4	505	1568	2269	1708	713
5	180	590	951	742	329

Tabela 4.18 – Probabilidades de interação dos estados de vento Falkenhagen e Wittstock.

Estados Falkenhagen	Estados Wittstock				
	1	2	3	4	5
1	2,42	4,44	3,47	2,37	1,07
2	4,14	8,19	9,21	5,76	2,27
3	2,86	7,62	9,12	7,17	2,63
4	1,44	4,47	6,48	4,87	2,03
5	0,51	1,68	2,71	2,12	0,94

A soma de todas as probabilidades da tabela 4.18 totaliza 100%, o que caracteriza todos os possíveis estados coincidentes entre as duas séries. A seguir, para cada estado de vento, ou seja, para os cinco valores obtidos a partir dos centróides representativos da localidade escolhida como referência (Falkenhagen), serão geradas cinco curvas de distribuição de probabilidades.

A primeira CDF, mostrada na figura 4.25, apresenta as probabilidades acumuladas do primeiro estado de Falkenhagen coincidir com cada um dos cinco estados de Wittstock. Este estado, com velocidade de 0.20 m/s, possui probabilidade total de ocorrer de 13,77% .

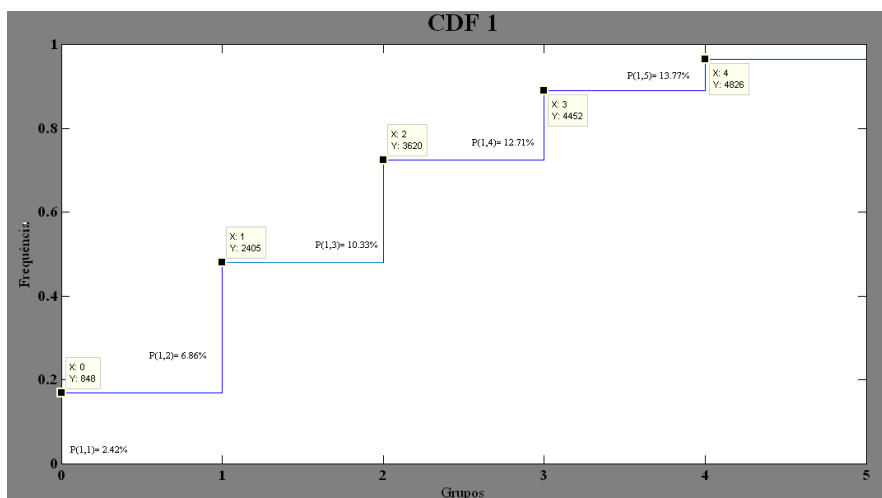


Figura 4.25 - CDF primeiro estado vento - Falkenhagen.

A segunda CDF, mostrada na figura 4.26, apresenta as probabilidades acumuladas do segundo estado de Falkenhagen coincidir com cada um dos cinco estados de Wittstock. Este estado, com velocidade de 1,90 m/s, possui probabilidade total de ocorrer de 29,57% .

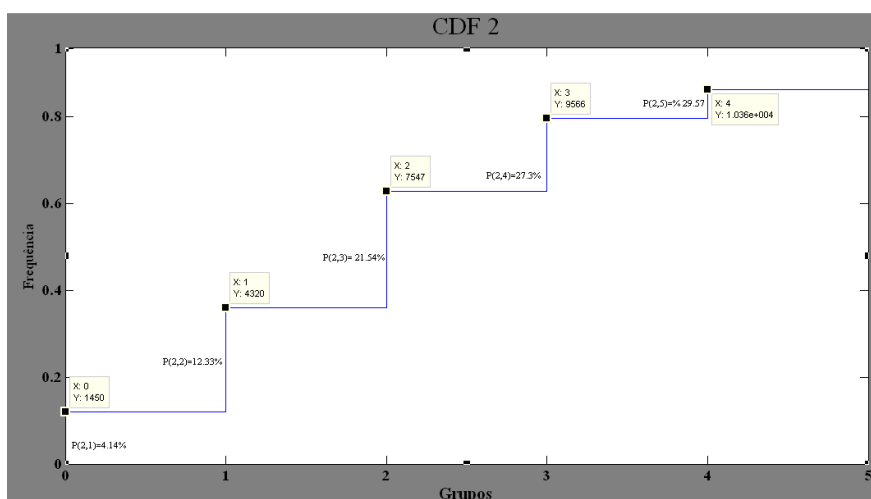


Figura 4.26 - CDF segundo estado vento - Falkenhagen.

A terceira CDF, mostrada na figura 4.27, apresenta as probabilidades acumuladas do terceiro estado de Falkenhagen coincidir com cada um dos cinco estados de Wittstock. Este estado, com velocidade de 3,22 m/s, possui probabilidade total de ocorrer de 29,39% .

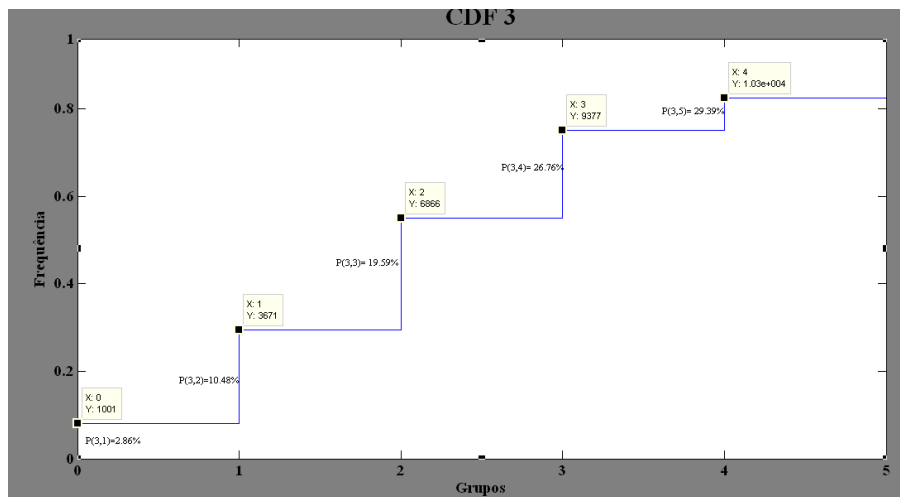


Figura 4.27 - CDF terceiro estado vento - Falkenhagen.

A quarta CDF, mostrada na figura 4.28, apresenta as probabilidades acumuladas do quarto estado de Falkenhagen coincidir com cada um dos cinco estados de Wittstock. Este estado, com velocidade de 4,7 m/s, possui probabilidade total de ocorrer de 19,3% .

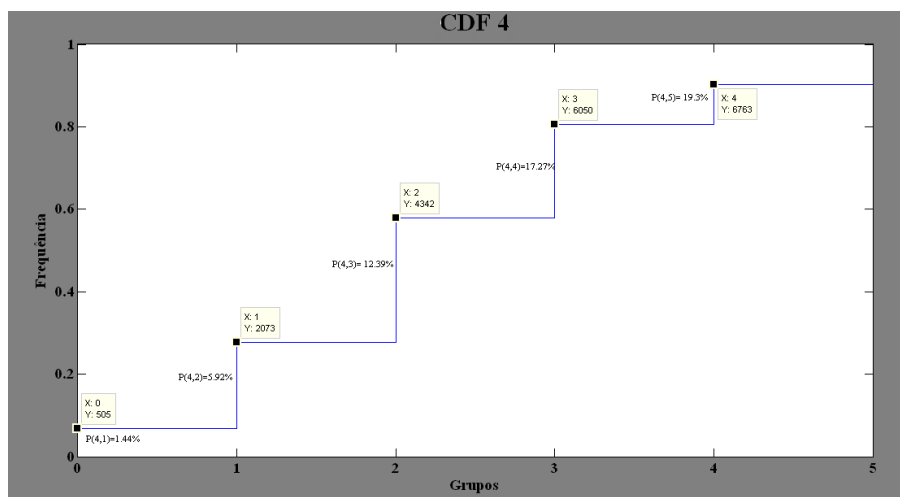


Figura 4.28 - CDF quarto estado vento - Falkenhagen.

A quinta CDF, mostrada na figura 4.29, apresenta as probabilidades acumuladas do quinto estado de Falkenhagen coincidir com cada um dos cinco estados de Wittstock. Este estado, com velocidade de 6,7 m/s, possui probabilidade total de ocorrer de 7,97% .

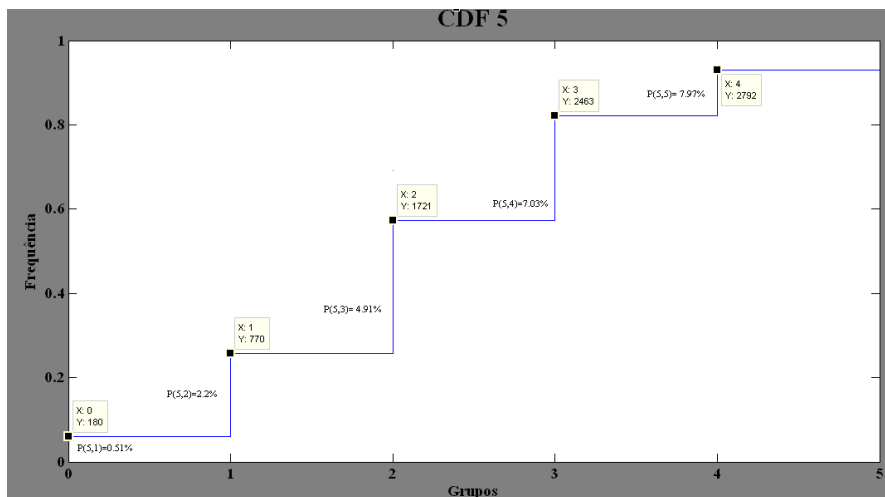


Figura 4.29 - CDF quinto estado vento - Falkenhagen.

O sistema demonstra que as maiores ocorrências dos estados seriam o estado 2 de Falkenhagen, 1,90 m/s, ocorrer em simultaneidade ao estado 3 de Wittstock, 3,26 m/s, sendo esta probabilidade de 9,21%. Desta forma, a CDF com maior probabilidade acumulada é a referente ao segundo estado, CDF 2, com 29,57%, muito próxima a CDF 3, com 29,39%. Esta proximidade se deve ao fato destes dois estados estarem muito presentes em suas séries individuais.

Com este sistema podem-se obter as probabilidades de ocorrências dos estados desejados, programando a interação das duas localidades, lembrando que a referência tomada aqui é Falkenhagen.

4.3.4 Freqüências e Probabilidades Relacionadas -Wittstock x Falkenhagen

As tabelas 4.19 e 4.20 demonstram os dados de freqüência e probabilidades dos estados de vento encontradas utilizando a mesma técnica, porém agora a referência não é mais Falkenhagen, e sim, Wittstock.

Tabela 4.19- Freqüência de interação dos estados de vento Wittstock e Falkenhagen.

		Estados Falkengahen				
		1	2	3	4	5
Estados Wittstock	1	848	1450	1001	505	180
	2	1557	2870	2670	1568	590
	3	1215	3227	3195	2269	951
	4	832	2019	2511	1708	742
	5	374	794	922	713	329

Tabela 4.20 - Probabilidades de interação dos estados de vento Wittstock e Falkenhagen.

		Estados Falkengahen				
		1	2	3	4	5
Estados Wittstock	1	2,42	4,14	2,86	1,44	0,51
	2	4,44	8,19	7,62	4,47	1,68
	3	3,47	9,21	9,12	6,48	2,71
	4	2,37	5,76	7,17	4,87	2,12
	5	1,07	2,27	2,63	2,03	0,94

Agora como no caso anterior, as CDF's a seguir analisam a correlação das duas localidades, tomando como referência Wittstock. Da mesma maneira que foi realizado na seção anterior, são mostradas abaixo 5 CDF's, uma para cada estado de vento da localidade tomada referência.

A primeira CDF, mostrada na figura 4.30, apresenta as probabilidades acumuladas do primeiro estado de Wittstock coincidir com cada um dos cinco estados de Falkenhagen. Este estado, com velocidade de 0,26 m/s, possui probabilidade total de ocorrer de 11,37% .

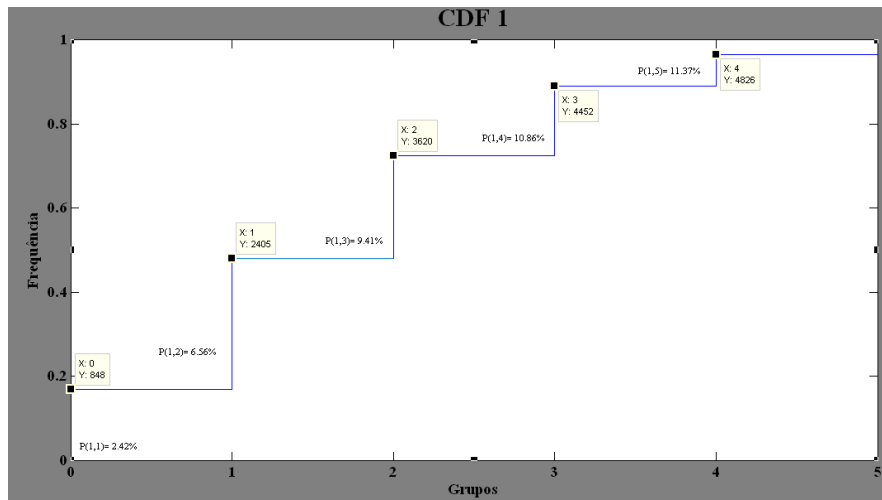


Figura 4.30 - CDF primeiro estado vento - Wittstock.

A segunda CDF, mostrada na figura 4.31, apresenta as probabilidades acumuladas do segundo estado de Wittstock coincidir com cada um dos cinco estados de Falkenhagen. Este estado, com velocidade de 1,91 m/s, possui probabilidade total de ocorrer de 26,41% .

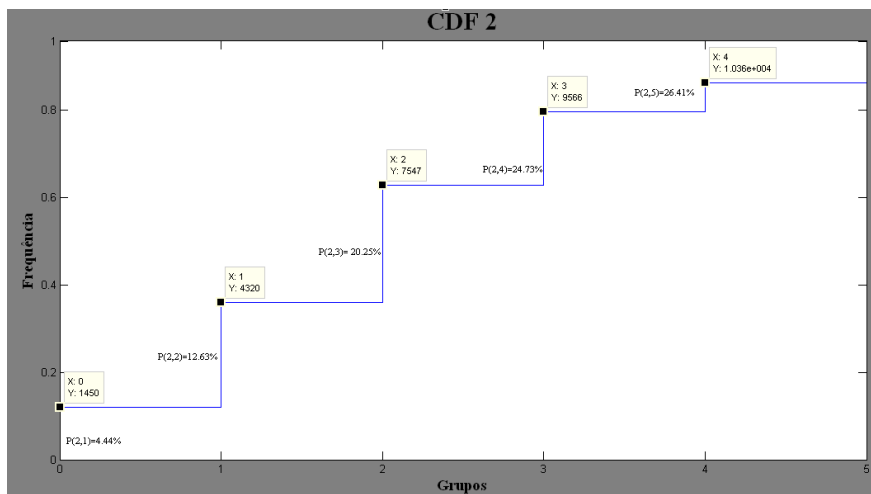


Figura 4.31 - CDF segundo estado vento - Wittstock.

A terceira CDF, mostrada na figura 4.32, apresenta as probabilidades acumuladas do terceiro estado de Wittstock coincidir com cada um dos cinco estados de Falkenhagen. Este estado, com velocidade de 3,26 m/s, possui probabilidade total de ocorrer de 30,98% .

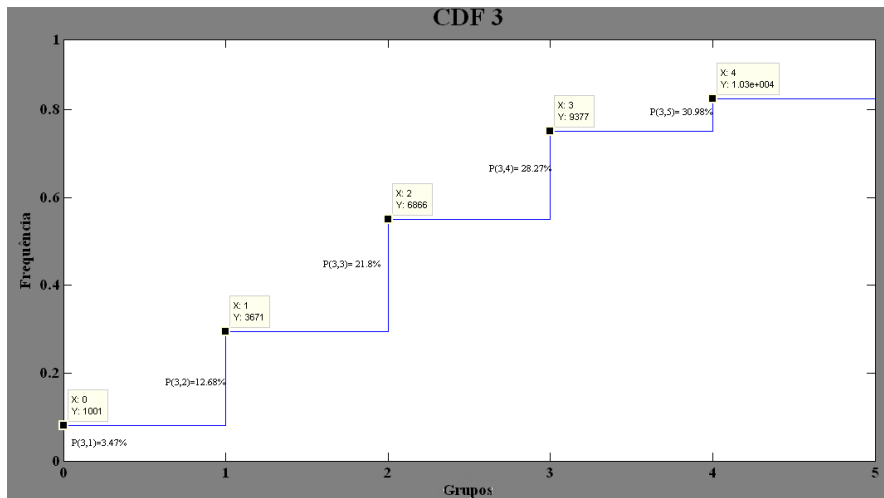


Figura 4.32 - CDF terceiro estado vento - Wittstock.

A quarta CDF, mostrada na figura 4.33, apresenta as probabilidades acumuladas do quarto estado de Wittstock coincidir com cada um dos cinco estados de Falkenhagen. Este estado, com velocidade de 4,89 m/s, possui probabilidade total de ocorrer de 19,3% .

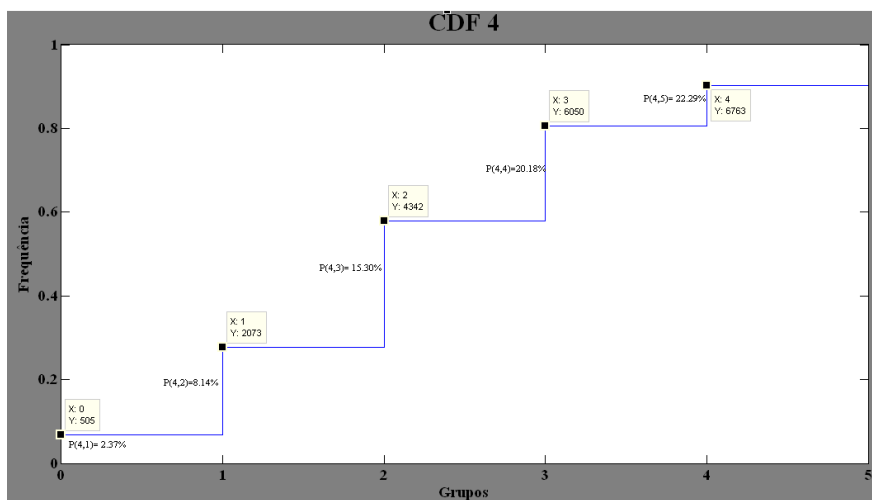


Figura 4.33 - CDF quarto estado vento - Wittstock.

A quinta CDF, mostrada na figura 4.34, apresenta as probabilidades acumuladas do quinto estado de Falkenhagen coincidir com cada um dos cinco estados de Wittstock. Este estado, com velocidade de 7,28 m/s, possui probabilidade total de ocorrer de 7,97% .

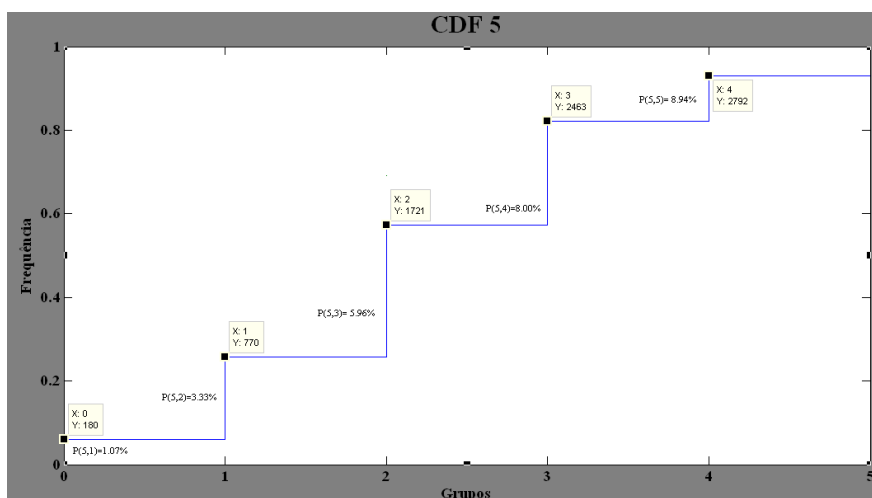


Figura 4.34 - CDF quinto estado vento - Wittstock.

Analisando as figuras 4.25 a 4.34, verifica-se que as CDF's com maiores probabilidades acumuladas em Falkenhagen e Wittstock são respectivamente CDF2-F e CDF3-W. Analisando as tabelas 4.18 e 4.20, percebe-se que estas são matrizes transpostas, e que os estados que mais coincidem são: o segundo estado de Falkenhagen com o terceiro estado de Wittstock.

Esta alta coincidência é bem razoável e poderia ser prevista, já que analisando as tabelas 4.13 e 4.16, estes são os estados que mais ocorrem em suas séries. Portanto, qualquer sorteio de probabilidades de estados pode ser feito utilizando qualquer uma das duas matrizes encontradas, bastando apenas escolher qual localidade deseja-se adotar como referência.

4.4 Conclusões do Estudo das Séries Temporais

Na primeira parte da análise deste trabalho, foram apresentados os gráficos e histogramas gerados a partir dos dados obtidos por medição das localidades alemãs. O mais alto valor de média anual de velocidade de vento foi encontrado em Jabel, com 3,58 m/s. Porém outro fato de extrema relevância é que Jabel não estava com os dados completos, ou seja, tinha apenas dados relativos a aproximadamente seis meses. Desta forma, é bem possível que, caso Jabel estivesse completa, seus valores de velocidades fossem superiores em relação às outras duas, e fosse viável a instalação de usinas eólicas

As duas séries completas, Falkenhagen e Wittstock, possuem dezembro como o mês com as melhores velocidades de vento, e em fevereiro e março, meses que Jabel apresentou maiores velocidades de vento. Essas duas possuem baixos valores de velocidade, e mesmo assim Jabel obteve a melhor média no período considerado. Caso o comportamento dos ventos em Jabel no restante do ano, seja próximo ao de Falkenhagen ou de Wittstock, esta ainda terá velocidades de ventos ainda maiores no fim do ano, o que talvez compense economicamente a instalação de usinas eólicas no local.

As velocidades encontradas nas localidades demonstram que para geração de energia elétrica a partir dos ventos destes locais, a viabilidade econômica do projeto estaria comprometida, já que, conforme fora aqui comentado algumas vezes, a velocidade deveria estar por volta dos 7 m/s pelo menos, o que não ocorre, ficando por volta dos 3 m/s nestas unidades.

No tocante às novas séries obtidas a partir do agrupamento de dados similares representados por centróides, percebe-se nas figuras 4.19 e 4.20 que os estados mais centrais das séries estão melhor representados, sendo que as velocidades mais altas ficaram mais excluídas. Mesmo assim, os valores de médias anuais sofreram poucas alterações, no caso de Falkenhagen a média manteve-se praticamente a mesma e de Wittstock sofreu um pequeno aumento. Esta pequena diferença se deve principalmente a associação dos dados originais a um reduzido número de centróides, o que seria minimizado caso a nova série fosse representada por um maior número de centróides.

Os erros encontrados nas associações dos dados, também podem ser considerados bons, já que possuíam média de erros e desvios padrão com valores bem baixos. O problema encontrado na análise dos erros absolutos também esbarra nos picos de velocidades das séries originais. Estes foram associados a centróides com valores de velocidades médias muito menores que seus valores reais, produzindo assim altos erros.

Nas análises das frequências de ocorrência dos estados de vento representativos, verifica-se que em ambas as localidades estudadas, Falkenhagen e Wittstock, os que mais aparecem são de baixo valor de velocidade média, na primeira é o segundo estado com 1,90m/s, seguido do terceiro estado, 3,22 m/s, com frequências extremamente próximas.

Nesta situação além da inviabilidade econômica, há também comprometimento da geração de energia elétrica em grande parte do ano, já que há muitos ventos com valores de velocidade inferiores a 2,5 m/s e 3 m/s, por volta de 44%.

No segundo caso, Wittstock, a maior frequência ocorre com o terceiro estado, que possui velocidade de 3.26 m/s, seguido também muito proximamente pelo segundo estado, com 1.91 m/s de velocidade. Assim como Falkenhagen, não há viabilidade econômica na geração nesta localidade, porém há um menor comprometimento na geração, já que grande parte de suas velocidades possuem velocidades de vento maiores que 2,5 m/s e 3m/s, por volta de 62%.

Na interação dos estados de vento, os dados demonstram que as maiores probabilidades são do estado 2 de Falkenhagen, 1,90 m/s, ocorrer simultaneamente ao estado 3 de Wittstock, 3,26%, podendo verificar nas tabelas 4.18 e 4.20. Assim, é possível conhecer as coincidências entre todos os estados obtidos por agrupamento das duas localidades estudadas, devendo-se apenas adotar uma referência. Sobre as maiores ocorrências serem dos estados 2 e 3 de Falkenhagen e Wittstock, isto é bem razoável, visto que ao analisar as duas localidades individualmente estes são os dados de velocidade que mais aparecem, portanto é natural que se encontrem mais.

Por fim, o principal comentário que deve ser feito é sobre a inviabilidade de instalação de usinas eólicas, devido às suas baixas velocidades médias encontradas. A Alemanha, no entanto, é o país líder em tecnologia de aerogeradores, e já consegue gerar energia a partir de velocidades de vento mais baixas [5]. Desta forma podem ser utilizados para instalação de pequenas centrais eólicas, para geração e abastecimento locais.

5. Conclusões

Este trabalho apresentou um estudo sobre velocidades de vento em sítios localizados na Alemanha, tendo como objetivo a análise de viabilidade de instalação de usinas eólicas. Para tal, é necessário um minucioso conhecimento das características do comportamento dos ventos locais. Assim, foram utilizadas principalmente descrições numéricas, análises estatísticas, e a técnica de agrupamento de dados conhecida como k-means, a fim de encontrar as frequências e probabilidades dos estados de velocidades de vento.

A técnica de agrupamento teve como objetivo a redução do número de estados de vento que pudessem representar os dados das séries originais. Isto permitiu que fossem estabelecidas as probabilidades e frequências individuais e correlacionadas das séries geradas, e desta forma conhecer o comportamento conjunto das localidades estudadas.

As novas séries obtidas, a partir de associação dos dados originais aos centróides representativos encontrados pelo k-means, foram consideradas boas aproximações, porém caso fosse requerido um grau maior de precisão, o número de centróides, deveria ser maior.

As velocidades de vento destas localidades são bem baixas, visto que algumas localidades no Brasil, como as citadas na seção 2.7.2, estudadas nas referências [6] e [7], estas chegam a valores médios anuais de 7,6 m/s, em comparação a média de 3 m/s, encontradas nas localidades alemãs estudadas.

Uma boa utilização dos conceitos vistos neste trabalho seria para o planejamento da interligação de sistemas eólicos. Visto que, conhecendo-se o comportamento do vento das localidades seria possível não só estimar a potência gerada em cada localidade, mas também a potência gerada simultaneamente por todas as usinas, a partir da correlação entre os ventos dos diferentes sítios.

6. Referências Bibliográficas

- [1] Tolmasquim, M. T., “*Fontes Renováveis de Energia*”, 1ª Edição, Editora Interciência, Rio de Janeiro, 2003.
- [2] ANEEL, “*Atlas da Energia Elétrica do Brasil*”, 2ª Edição, Brasília, 2005.
- [3] CBEE, “*Panorama da Energia Elétrica*”, site: www.eolica.org.br.
- [4] São Francisco, “*Energia Eólica e Meio Ambiente*”, site: www.colegiosaofrancisco.com.br/alfa/meio-ambiente-energia-eolica
- [5] GWEC, “*Global Wind Energy Outlook 2008*”, Bélgica, 2008.
- [6] Souza, T. A., “*Agrupamento de Séries Temporais de Vento para Avaliação da Disponibilidade de Geração de Usinas Eólicas*”, Tese de Mestrado, COPPE, UFRJ, 2008.
- [7] Leite, A. P., “*Modelagem de Fazendas Eólicas para Estudo de Confiabilidade*”, Tese de Mestrado, COPPE, UFRJ, 2005.
- [8] Pereira, B.; Pais, M.; Sales, P., *Análise Espectral de Séries Temporais – Uma Introdução para Engenharia, Economia e Estatística*, 1ª Edição, Editora Arte Final, Eletrobrás 1986.
- [9] Chatfield, C., *The Analyses of Time Series- An Introduction*, 6ª Edição, Editora LLC, Flórida, Estados Unidos, 2000.
- [10] Spiegel, M., *Estatística*, 3ª Edição, Editora MacGrowHill, 1998.
- [11] Ochi, L. S.; Dias, C. R.; Soares, S. F. S., “*Clusterização em Mineração de Dados*”, Universidade Federal Fluminense, Niterói, 2005.
- [12] Hines, W. W.; Montgomery, D. C.; Goldsman, D. M.; Borror C. M., “*Probabilidade e Estatística na Engenharia*”, 4ª Edição, Editora LTC, Rio de Janeiro, 2006.

[13] Silva, M. S.; Schawarz, H.; Pfeiffer, K.; Roman, H.; Dorendorf, S., “*Analysis of Dynamic Load Flow Behaviour for the East Germany Winds Parks Integration*”, 2006.

[14] Billinton, R.; Yi Gao, “*Multistate Wind Energy Conversion System Models for Adequacy Assessment of Generating Systems Incorporating Wind Energy*”, Energy Conversion, IEE, Volume 23, março 2008.

Anexo A - Macro Frequência e Correlação dos Estados de Velocidade de Vento

Abaixo, segue a rotina criada em Excel para cálculo das probabilidades de ocorrência e correlação, dos estados de vento das três séries de vento estudadas neste trabalho.

```

Sub Con()
Range("J2:N6").ClearContents
For i = 2 To 35041
    If ActiveSheet.Cells(i, 2) = 1 Then
        For j = 2 To 35041
            If ActiveSheet.Cells(j, 6) = 1 Then
                ActiveSheet.Cells(2, 10) =
                ActiveSheet.Cells(2, 10) + 1
            End If
            If ActiveSheet.Cells(j, 6) = 2 Then
                ActiveSheet.Cells(2, 11) =
                ActiveSheet.Cells(2, 11) + 1
            End If
            If ActiveSheet.Cells(j, 6) = 3 Then
                ActiveSheet.Cells(2, 12) =
                ActiveSheet.Cells(2, 12) + 1
            End If
            If ActiveSheet.Cells(j, 6) = 4 Then
                ActiveSheet.Cells(2, 13) =
                ActiveSheet.Cells(2, 13) + 1
            End If
            If ActiveSheet.Cells(j, 6) = 5 Then
                ActiveSheet.Cells(2, 14) =
                ActiveSheet.Cells(2, 14) + 1
            End If
        Next
    End If
    If ActiveSheet.Cells(i, 2) = 2 Then
        For j = 2 To 35041
            If ActiveSheet.Cells(j, 6) = 1 Then
                ActiveSheet.Cells(3, 10) =
                ActiveSheet.Cells(3, 10) + 1
            End If
            If ActiveSheet.Cells(j, 6) = 2 Then
                ActiveSheet.Cells(3, 11) =
                ActiveSheet.Cells(3, 11) + 1
            End If
            If ActiveSheet.Cells(j, 6) = 3 Then
                ActiveSheet.Cells(3, 12) =
                ActiveSheet.Cells(3, 12) + 1
            End If
            If ActiveSheet.Cells(j, 6) = 4 Then
                ActiveSheet.Cells(3, 13) =
                ActiveSheet.Cells(3, 13) + 1
            End If
            If ActiveSheet.Cells(j, 6) = 5 Then
                ActiveSheet.Cells(3, 14) =
                ActiveSheet.Cells(3, 14) + 1
            End If
        Next
    End If
    If ActiveSheet.Cells(i, 2) = 3 Then
        For j = 2 To 35041
            If ActiveSheet.Cells(j, 6) = 1 Then
                ActiveSheet.Cells(4, 10) =
                ActiveSheet.Cells(4, 10) + 1
            End If
        Next
    End If
End Sub

```

```

        If ActiveSheet.Cells(j, 6) = 2 Then
            ActiveSheet.Cells(4, 11) =
ActiveSheet.Cells(4, 11) + 1
        End If

        If ActiveSheet.Cells(j, 6) = 3 Then
            ActiveSheet.Cells(4, 12) =
ActiveSheet.Cells(4, 12) + 1
        End If

        If ActiveSheet.Cells(j, 6) = 4 Then
            ActiveSheet.Cells(4, 13) =
ActiveSheet.Cells(4, 13) + 1
        End If

        If ActiveSheet.Cells(j, 6) = 5 Then
            ActiveSheet.Cells(4, 14) =
ActiveSheet.Cells(4, 14) + 1
        End If

        Next
    End If

    If ActiveSheet.Cells(i, 2) = 4 Then
        For j = 2 To 35041

            If ActiveSheet.Cells(j, 6) = 1 Then
                ActiveSheet.Cells(5, 10) =
ActiveSheet.Cells(5, 10) + 1
            End If

            If ActiveSheet.Cells(j, 6) = 2 Then
                ActiveSheet.Cells(5, 11) =
ActiveSheet.Cells(5, 11) + 1
            End If

            If ActiveSheet.Cells(j, 6) = 3 Then
                ActiveSheet.Cells(5, 12) =
ActiveSheet.Cells(5, 12) + 1
            End If

            If ActiveSheet.Cells(j, 6) = 4 Then
                ActiveSheet.Cells(5, 13) =
ActiveSheet.Cells(5, 13) + 1
            End If

            End If

        End If

        If ActiveSheet.Cells(i, 2) = 5 Then
            For j = 2 To 35041

                If ActiveSheet.Cells(j, 6) = 1 Then
                    ActiveSheet.Cells(6, 10) =
ActiveSheet.Cells(6, 10) + 1
                End If

                If ActiveSheet.Cells(j, 6) = 2 Then
                    ActiveSheet.Cells(6, 11) =
ActiveSheet.Cells(6, 11) + 1
                End If

                If ActiveSheet.Cells(j, 6) = 3 Then
                    ActiveSheet.Cells(6, 12) =
ActiveSheet.Cells(6, 12) + 1
                End If

                If ActiveSheet.Cells(j, 6) = 4 Then
                    ActiveSheet.Cells(6, 13) =
ActiveSheet.Cells(6, 13) + 1
                End If

                If ActiveSheet.Cells(j, 6) = 5 Then
                    ActiveSheet.Cells(6, 14) =
ActiveSheet.Cells(6, 14) + 1
                End If

            Next
        End If

    Next
End Sub

```

