



**NOTA TÉCNICA**

# **METODOLOGIA DE AGREGAÇÃO DE DADOS ANEMOMÉTRICOS**

*SISTEMA AMA – Disponibilização  
das séries de dados anemométricos e  
climatológicos*

**AGOSTO DE 2025**

## ■ Colaboradores

NOTA TÉCNICA EPE/DEE/SGR/051/2025

### **Coordenação Geral**

Thiago Guilherme Ferreira Prado

Reinaldo da Cruz Garcia

### **Coordenação Executiva**

Gustavo Pires da Ponte

Caio Monteiro Leocadio

### **Coordenação Técnica**

Gustavo Pires da Ponte

### **Equipe Técnica**

Charles Egberto Guedes Vonnegut Vieira de Mello

Paula Monteiro Pereira

Rafael Pinho Furtado

Bernardo Albuquerque Domingues da Silva (estagiário)



epe



## VALOR PÚBLICO

A EMPRESA DE PESQUISA ENERGÉTICA – EPE TEM POR FINALIDADE PRESTAR SERVIÇOS AO MINISTÉRIO DE MINAS E ENERGIA (MME) NA ÁREA DE ESTUDOS E PESQUISAS DESTINADAS A SUBSIDIAR O PLANEJAMENTO DO SETOR ENERGÉTICO, COBRINDO ENERGIA ELÉTRICA, PETRÓLEO E GÁS NATURAL E SEUS DERIVADOS E BIOCOMBUSTÍVEIS. SOMOS UMA EMPRESA PÚBLICA FEDERAL, DEPENDENTE DO ORÇAMENTO GERAL DA UNIÃO.

A EPE FOI CRIADA COM O OBJETIVO DE RESGATAR A RESPONSABILIDADE CONSTITUCIONAL DO ESTADO NACIONAL EM ASSEGURAR AS BASES PARA O DESENVOLVIMENTO SUSTENTÁVEL DA INFRAESTRUTURA ENERGÉTICA DO PAÍS. A PARTIR DE SUA CRIAÇÃO, A ATUAÇÃO DA EPE CONSOLIDOU-SE COMO PARTE FUNDAMENTAL DE UM CICLO DE ATIVIDADES QUE SE INICIA COM AS DEFINIÇÕES DE POLÍTICAS E DIRETRIZES NO ÂMBITO DO CNPE – CONSELHO NACIONAL DE POLÍTICA ENERGÉTICA E DO MME.

COM ESTAS DEFINIÇÕES, MATERIALIZAM-SE OS ESTUDOS E AS PESQUISAS QUE EFETIVAMENTE ORIENTAM O DESENVOLVIMENTO DO SETOR ENERGÉTICO BRASILEIRO.

NESTE CONTEXTO, O SISTEMA AMA FOI DESENVOLVIDO PARA OBTENÇÃO DE UMA BASE DE DADOS ANEMOMÉTRICOS E CLIMATOLÓGICOS COM OBJETIVO DE MANTER UM HISTÓRICO DE LONGO PRAZO DE DADOS CONFIÁVEIS.

NESTA NOTA TÉCNICA, A EPE TRAZ TRANSPARÊNCIA E DIMINUI A ASSIMETRIA DE INFORMAÇÕES SOBRE DADOS ANEMOMÉTRICOS, DISPONIBILIZANDO, DE FORMA AGREGADA, OS DADOS DO SISTEMA AMA, QUE PODERÃO SER UTILIZADOS PARA ELABORAÇÃO DE INÚMEROS ESTUDOS ACADÊMICOS E CORPORATIVOS PARA O SETOR ELÉTRICO NACIONAL.

DESTACA-SE MAIS UMA VEZ, A IMPORTÂNCIA DA TRANSPARÊNCIA E DISPONIBILIZAÇÃO DE DADOS PARA A PROMOÇÃO DE UMA TRANSIÇÃO ENERGÉTICA ADERENTE ÀS POTENCIALIDADES, SEGURANÇA ENERGÉTICA, COMPETITIVIDADE, SUSTENTABILIDADE E VANTAGENS VISLUMBRADAS NO CENÁRIO DO SETOR ELÉTRICO BRASILEIRO.

## Histórico de Revisões

Rev.	Data	Descrição
0	08/08/2025	Emissão Inicial

MINISTÉRIO DE  
MINAS E ENERGIA



**Ministro de Estado**  
Alexandre Silveira de Oliveira

**Secretário-Executivo**  
Arthur Cerqueira Valério

**Secretário de Planejamento e Desenvolvimento Energético**  
Gustavo Cerqueira Ataíde

**Secretário de Energia Elétrica**  
Gentil Nogueira de Sá Júnior

**Secretário de Petróleo, Gás Natural e Combustíveis Renováveis**  
Pietro Adamo Sampaio Mendes

**Secretário de Geologia, Mineração e Transformação Mineral**  
Vítor Eduardo de Almeida Saback



**Presidente**

Thiago Guilherme Ferreira Prado

**Diretor de Estudos Econômico-Energéticos e Ambientais**

Thiago Ivanoski Teixeira

**Diretor de Estudos de Energia Elétrica**  
Reinaldo da Cruz Garcia

**Diretor de Estudos do Petróleo, Gás e Biocombustíveis**

Heloísa Borges Bastos Esteves

**Diretor de Gestão Corporativa**  
Carlos Eduardo Cabral Carvalho

<http://www.epe.gov.br>

# Sumário

<b>1</b>	<b>Introdução</b>	<b>8</b>
1.1	Dados Disponíveis e Histórico de Agregações	9
1.2	Motivação	9
<b>2</b>	<b>Metodologia e Referencial Teórico</b>	<b>10</b>
2.1	Metodologia e Detalhamento das Bibliotecas Utilizadas	10
2.2	Tratamento dos Dados e Obtenção do Ano Médio	11
2.3	Matriz de Correlação entre Anos Típicos	13
2.4	Metodologia de Clusterização de Dados	14
2.5	Aplicação da Metodologia nos Dados do Sistema AMA	16
<b>3</b>	<b>Apresentação dos Dados</b>	<b>21</b>
3.1	Dados e Grupamentos	21
3.1.1.	Clusters na Região Litoral Equatorial	22
3.1.2.	Clusters na Região Rio Grande do Norte	24
3.1.3.	Clusters na Região Centro NE	26
3.1.4.	Clusters na Região Interior NE	28
3.1.5.	Clusters na Região Bahia	30
3.1.6.	Clusters na Região Sul	32
3.1.7.	Comentários Gerais	34
<b>4</b>	<b>Conclusão</b>	<b>34</b>
<b>5</b>	<b>Referências Bibliográficas</b>	<b>35</b>

## ■ Lista de Figuras

Figura 1 - Localização das estações anemométricas do Sistema AMA.....	8
Figura 2 - Fluxograma do processo .....	11
Figura 3 - Procedimento de obtenção de ano típico para cada estação anemométrica.....	12
Figura 4 - Matriz de correlação .....	13
Figura 5 - Exemplo do cálculo de distâncias pelo método de Ward.....	15
Figura 6 - Dendrograma das clusterizações dos dados exemplificados na Figura 4 .....	16
Figura 7 - Procedimento de amostragem dos clusters .....	17
Figura 8 - Visualização dos clusters obtidos a partir da amostragem vertical.....	17
Figura 9 - Representação em árvore dos clusters.....	18
Figura 10 - Exemplo de análise individual por árvore.....	18
Figura 11 – Apresentação geográfica do processo de agregação de clusters .....	19
Figura 12 - Clusters finais após análise dos resultados do modelo .....	20
Figura 13 - Distribuição de ocorrências e estimativa de Weibull dos clusters da Região Litoral Equatorial.....	22
Figura 14 – Rosa dos ventos dos clusters da Região Litoral Equatorial.....	22
Figura 15 - Dia médio e ano médio da velocidade média do vento nos clusters da Região Litoral Equatorial.....	23
Figura 16 - Distribuição horária e mensal da direção média do vento nos clusters da Região Litoral Equatorial.....	23
Figura 17 - Distribuição de ocorrências e estimativa de Weibull dos clusters da Região Rio Grande do Norte .....	24
Figura 18 – Rosa dos ventos dos clusters da Região Rio Grande do Norte .....	24
Figura 19 - Dia médio e ano médio da velocidade média do vento nos clusters da Região Rio Grande do Norte .....	25
Figura 20 - Distribuição horária e mensal da direção média do vento nos clusters da Região Rio Grande do Norte .....	25
Figura 21 - Distribuição de ocorrências e estimativa de Weibull dos clusters da Região Centro NE .....	26
Figura 22 – Rosa dos ventos dos clusters da Região Centro NE .....	26
Figura 23 - Dia médio e ano médio da velocidade média do vento nos clusters da Região Centro NE .....	27
Figura 24 - Distribuição horária e mensal da direção média do vento nos clusters da Região Centro NE .....	27
Figura 25 - Distribuição de ocorrências e estimativa de Weibull dos clusters da Região Interior NE .....	28
Figura 26 – Rosa dos ventos dos clusters da Região Interior NE .....	28
Figura 27 - Dia médio e ano médio da velocidade média do vento nos clusters da Região Interior NE .....	29
Figura 28 - Distribuição horária e mensal da direção média do vento nos clusters da Região Interior NE .....	29
Figura 29 - Distribuição de ocorrências e estimativa de Weibull dos clusters Região Bahia .....	30
Figura 30 - Rosa dos ventos dos clusters da Região Bahia .....	30
Figura 31 - Dia médio e ano médio da velocidade média do vento nos clusters da Região Bahia ...	31
Figura 32 - Distribuição horária e mensal da direção média do vento nos clusters da Região Bahia .....	31
Figura 33 - Distribuição de ocorrências e estimativa de Weibull dos clusters da Região Sul.....	32
Figura 34 – Rosa dos ventos dos clusters da Região Sul.....	32
Figura 35 - Dia médio e ano médio da velocidade média do vento nos clusters da Região Sul .....	33
Figura 36 - Distribuição horária e mensal da direção média do vento nos clusters da Região Sul ...	33

## ■ Lista de Tabelas

Tabela 1 - Características dos clusters da Região Litoral Equatorial .....	22
Tabela 2 - Características dos clusters da Região Rio Grande do Norte.....	24
Tabela 3 - Características dos clusters da Região Centro NE.....	26
Tabela 4 - Características dos clusters da Região Interior NE.....	28
Tabela 5 - Características dos clusters da Região Bahia .....	30
Tabela 6 - Características dos clusters da Região Sul .....	32

# 1 Introdução

Esta Nota Técnica tem por objetivo apresentar a metodologia e os resultados da agregação de dados anemométricos e climatológicos, formando séries históricas (vento e direção, temperatura média, pressão média e umidade relativa média) a partir da base de dados do Sistema AMA.

O AMA – Sistema de Acompanhamento de Medições Anemométricas - conta atualmente com recebimento de dados de mais de 650 estações anemométricas, formando, uma das maiores base de dados meteorológicos de estações de alta qualidade a nível mundial, com histórico desde 2011.

A base de dados do Sistema AMA contém dados de temperatura, pressão, umidade relativa, direção e velocidade do vento, registrados em intervalos de 10 minutos. Para garantir a qualidade dos dados coletados, as estações anemométricas pertencentes ao Sistema AMA devem seguir requisitos definidos pela EPE, como: medições contínuas e permanentes, utilização de instrumentos de medição aceitos, instalação de quantidade mínima de instrumentos de medição, altura de medição referente à altura do rotor do aerogerador, taxa mínima de indisponibilidade de dados e outros requisitos, conforme disponível na NT EPE-DEE-RE-057/2016 – r4 “[Instruções para medições meteorológicas em parques eólicos](#)”.

A Figura 1 apresenta a localização das estações anemométricas do Sistema AMA. As medições anemométricas são realizadas em nove estados brasileiros (BA, CE, MA, PB, PE, PI, RN, RS e SE), no entanto, destaca-se que mais da metade das estações estão concentradas na Bahia e no Rio Grande do Norte.

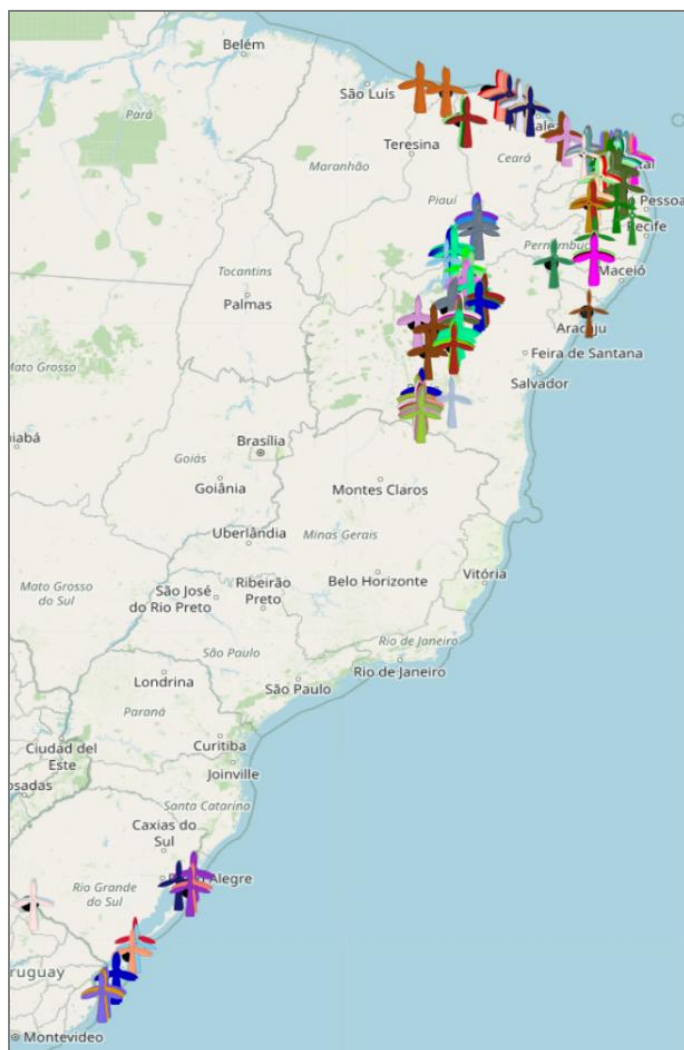


Figura 1 - Localização das estações anemométricas do Sistema AMA

## 1.1 Dados Disponíveis e Histórico de Agregações

A divulgação pela EPE dos dados anemométricos e climatológicos do Sistema AMA, vem sendo realizada das seguintes formas:

- [Dados Públicos Consolidados por Estado](#): de forma agregada por estado e por ano, e em intervalo horário, a partir de 2012, compreendendo as informações de temperatura, umidade relativa, velocidade média, permanência de velocidade e direção predominante. O público geral tem acesso a essas informações por meio de planilhas, encaminhadas via e-mail, e através de gráficos. Desta forma, o usuário tem acesso a valores médios horários e mínimos e máximos dessas variáveis, que agregam todas as torres de cada estado. Apesar do nível de agregação, que agrupa dados de locais com diferentes perfis de vento com base na divisão geopolítica/administrativa (por estados), entende-se que tais dados podem ser úteis para estudos acadêmicos e prospectivos iniciais, uma vez que compilam informações dos locais de maior potencial em cada estado, uma vez que as torres estão nas regiões dos parques eólicos em operação.
- [Boletim da Energia Eólica](#): são divulgados índices que descrevem a disponibilidade do recurso eólico (índice de energia) e o potencial de transformação desse recurso em energia elétrica (índice de produção). Esses índices são calculados com base nas medições de velocidade do vento, temperatura, pressão e umidade relativa das estações anemométricas do Sistema AMA. Os índices são divulgados para três regiões que possuem sazonalidade semelhante: Interior Nordeste, Litoral Nordeste e Rio Grande do Sul.
- [Nota Técnica “Dados de entrada para modelos energéticos: metodologias e premissas”](#): apresenta a metodologia de obtenção de séries de longo prazo de geração de energia. A divisão dos parques eólicos em regiões foi realizada em duas etapas. Na primeira etapa foi avaliada a sazonalidade dos parques, ou seja, suas características mensais. Essa análise foi usada para separar os parques em regiões. Na segunda etapa, cada uma das regiões foi dividida em sub-regiões, de acordo com suas características horárias. Como complemento à Nota Técnica, são disponibilizadas séries horárias de geração referentes a cada região.
- [Nota Técnica “Caracterização do Recurso Eólico e Resultados Preliminares de sua Aplicação no Sistema Elétrico”](#): apresenta uma análise das estações anemométricas pertencentes ao AMA entre 2011 e 2013, detalhando as características observadas e comparando a estimativa de geração realizada no escopo da nota com energia negociada nos leilões de energia além de apresentar os índices de confiabilidade e *Capacity Credit* da fonte.

## 1.2 Motivação

A Nota Técnica [“Sistema AMA - 13 anos de acompanhamento de medições anemométricas”](#) apresentada no final de 2024, discutiu a possibilidade de divulgar os dados com maior nível de detalhamento do que apresentado em outros estudos e ferramentas da EPE. Foi realizada uma consulta dirigida aos agentes envolvidos no Sistema AMA, na qual foram levantados os desafios para a divulgação desses dados.

Como resultado da consulta, observou-se que não há um consenso entre os agentes, sendo alguns favoráveis à divulgação dos dados por parque e em intervalos de 10 minutos, enquanto

outros foram contra, apontando riscos de exposição à concorrência e possível mau uso dos dados. Diante disso, concluiu-se que na divulgação dos dados anemométricos por parque e em intervalos de 10 minutos, existiria o risco de judicialização por parte dos agentes que se opuseram a essa proposta.

Um meio-termo encontrado para divulgação dos dados, em maior nível de detalhamento, foi a agregação das estações anemométricas por clusters com comportamento similar de vento, utilizando intervalos horários. Os agrupamentos foram definidos por meio de clusterização hierárquica, cuja metodologia é detalhada nesta Nota Técnica.

A partir da aplicação desta metodologia, a EPE passa a disponibilizar dados mais detalhados, permitindo usos diversos pela sociedade e contribuindo para novos estudos e o avanço das pesquisas sobre a energia eólica a nível nacional.

## 2 Metodologia e Referencial Teórico

---

Os dados meteorológicos, assim como mencionado na Nota Técnica [“Sistema AMA - 13 anos de acompanhamento de medições anemométricas”](#) [1], são influenciados por diversos fatores, tais como: características do terreno (local, regional ou continental), fenômenos meteorológicos e climatológicos, características construtivas dos instrumentos, entre outros.

Com a interação de tantos fenômenos sobre a medição, torna-se uma tarefa complexa determinar, a priori, as características representativas que permitam a definição de grupos (clusters) de medições coerentes entre si.

Adicionada a tal complexidade, nota-se a possibilidade de micro características, como alterações de relevo, rugosidade ou altura de medição, serem dominantes com relação ao comportamento de um grupo de medições, levando a possibilidade de dois comportamentos se sobreporem regionalmente, o que tende a aumentar a complexidade da agregação de dados.

Desta forma, foi então selecionada uma metodologia de agregação de dados utilizando apenas características do recurso medido de velocidade de vento, para definição dos agrupamentos, sem considerar quaisquer características locais, construtivas ou ambientais. Esta definição permite que os resultados das medições das velocidades de vento definam os agrupamentos coerentes, de forma independente das demais grandezas de medições anemométricas e climatológicas disponíveis, evitando enviesamento dos dados por premissas inadequadas.

### 2.1 Metodologia e Detalhamento das Bibliotecas Utilizadas

A clusterização de dados é um tipo de procedimento que apresenta grande desenvolvimento e utilização [2] [3] [4] [5] nos últimos anos, sendo debatido em diversas áreas científicas e técnicas. Há, com esse desenvolvimento, uma grande quantidade de metodologias para clusterização de dados, desde clusterizações com metadados e informações qualitativas, até opções para dados quantitativos e séries temporais.

Cada modelo apresenta diferentes resultados, que podem atender a diferentes objetivos. A seleção de modelos e parâmetros que garantam a correta agregação de acordo com o objetivo a ser estudado se mostra uma das principais dificuldades deste tipo de processo.

Desta forma, é importante adequar os modelos aos dados de entrada e aos objetivos. Nota-se que as diferentes classes de dados de entrada, como: séries temporais, dados qualitativos ou dados geográficos podem demandar diferentes modelos e apresentar resultados distintos.

No tema tratado nesta Nota Técnica é possível identificar a influência de diversos fatores que podem ser considerados como parâmetros, como: posição geográfica, altura da torre, altitude de instalação, período de medição, condições ambientais locais entre outras.

Objetivando simplificar o processo e realizar uma clusterização focada nas relações finais entre os dados, foi determinado que o modelo deve considerar apenas os dados principais (velocidade do vento), de maneira a desconsiderar a variabilidade de longo prazo e as velocidades médias, gerando grupos que representam as variações sazonais e diárias.

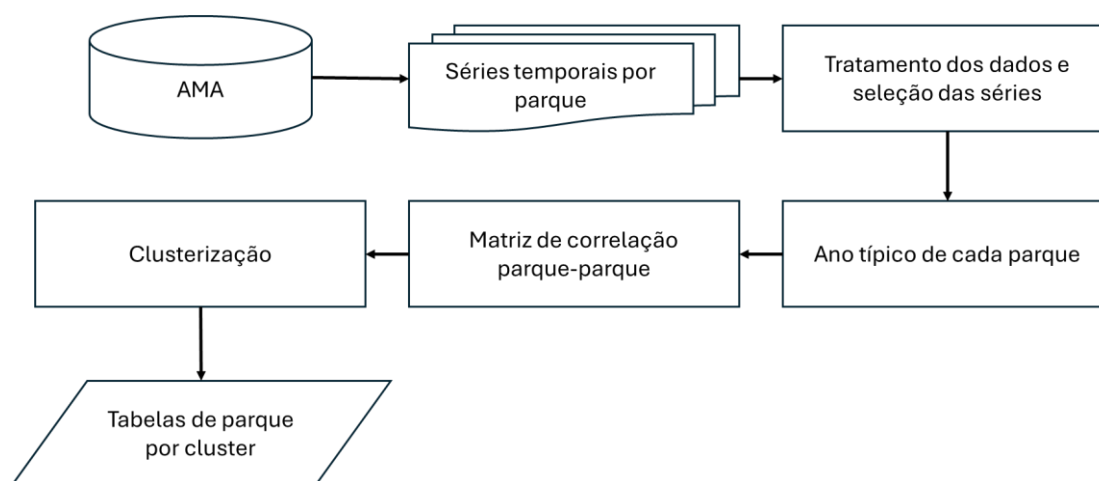
Para obter este tipo de resultado é necessário garantir a qualidade dos dados de acordo com o objetivo especificado (clusterizar de acordo com as variações do recurso eólico sazonais e diárias), evitando possíveis influências de falhas de dados ou períodos de início e fim de operação. Para tal, foi definido que seriam utilizados no processo, anos médios (séries temporais horárias com período total de um ano) de cada uma das estações cadastradas no Sistema AMA. Esse processo visa reduzir a influência das falhas de dados e períodos de medição distintos nos resultados das clusterizações.

Este tratamento permite também o foco no recurso disponível em cada localidade. Nota-se, ainda, que a depender das metodologias de clusterização a serem utilizadas, as velocidades médias podem influenciar o processo de clusterização, tendendo a eliminar torres que apresentam valores médios muito diferentes de velocidade. Com o objetivo de caracterizar regiões de acordo com a variabilidade dos dados, foi calculada então a matriz de correlação entre as torres de medição, garantindo que as variações relativas as medições fossem o parâmetro a ser considerado na análise.

A partir da matriz de correlações, obtidas a partir dos anos médios das estações, foram testados os diferentes modelos de clusterização e metodologias de agregação. Foi selecionado, após alguns testes, o método de Ward disponível na biblioteca *scipy.cluster.hierarchy* [2] para a versão final das clusterizações.

Destaca-se que a definição de anos típicos é utilizada somente para permitir o agrupamento (clusterização) de torres, enquanto para a disponibilização de dados, pretende-se considerar as séries históricas, que permitem análises mais detalhadas de comportamento e variabilidade interanual. Uma vez utilizadas para definição de clusters, é possível então, elaborar processos para agregação de suas séries originais, análise de metadados e diversos outros processos.

O fluxograma do processo é apresentado na **Figura 2**, desde a origem dos dados até a saída final, que será utilizada para tratamento e agregação das informações consolidadas por cluster, que serão disponibilizadas à sociedade a partir de um [dashboard online](#).



**Figura 2 - Fluxograma do processo**

## 2.2 Tratamento dos Dados e Obtenção do Ano Médio

Para o estudo aqui apresentado foram utilizados os dados do AMA disponíveis até 2024. Tais dados foram validados e tratados, de acordo com os padrões no Sistema AMA (marcação de erros e preenchimento de falhas), com a finalidade de garantir a consistência das informações.

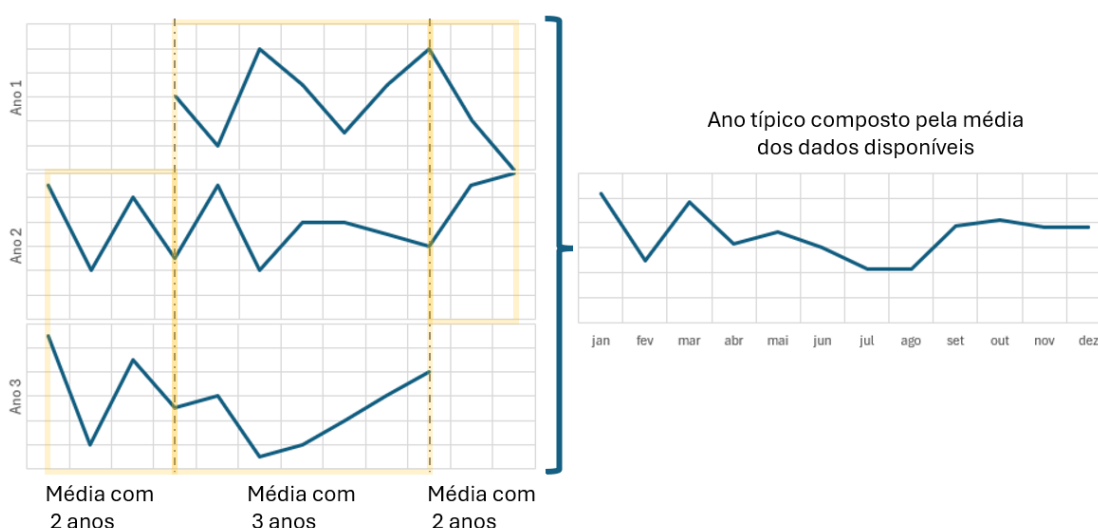
Após obtidos os dados tratados de cada estação, é necessário definir os critérios mínimos das séries temporais a serem selecionadas. As informações utilizadas para a seleção das séries são compostas por: data de início e término de operação da torre anemométrica, caso tenha sido desmobilizada; indisponibilidade de medição e continuidade das medições disponíveis. A partir da análise destes parâmetros é importante garantir a comparabilidade entre diferentes estações, tendo em vista também a máxima utilização das informações disponíveis.

Para permitir o uso do número máximo possível de dados e torres, considerando que as campanhas de medição têm períodos distintos, com início, perdas e finalização de dados diferentes, é desejável a definição de um período comum entre as estações. Este processo, no entanto, resulta na eliminação de uma quantidade grande de dados e pode gerar descontinuidade de informações, uma vez que em um grande banco de dados, como o AMA, são comuns amostras falhas ao longo do tempo.

Outra solução identificada é a obtenção de anos típicos para cada uma das estações. Tal processo é justificável por ser um período cíclico do recurso eólico e por representar uma duração útil no ponto de vista de diversas aplicações. Nota-se que a elaboração de um ano típico pode ser realizada de diferentes formas, sendo idealmente aplicada com uma grande quantidade de dados para cada ponto de amostragem.

Objetivando utilizar o máximo número de parques disponíveis e considerando que a ocorrência de mais de um ano de medição anemométrica tende a cobrir possíveis falhas e demais irregularidades nos dados medidos, a partir do cálculo de um ano típico, foi definido que somente torres com períodos superiores há 2 anos de medição anemométrica, neste caso para velocidades de vento conforme indicado anteriormente, serão utilizadas e o processo de obtenção do ano típico realizado a partir da média das amostras horárias no período de um ano.

Assim, os anos de medição disponíveis para cada uma das torres foram alinhados, considerando a informação de horário, dia e mês. O ano típico é obtido pela média dos valores de cada hora no ano, obtendo um ano único representativo de toda a campanha de medição para cada estação. Desta forma, além de problemas relacionados a falhas de dados, minimizam-se os riscos de medições exógenas ao padrão ou caracterização de séries temporais não alinhadas. Tal procedimento é apresentado na Figura 3.



**Figura 3 - Procedimento de obtenção de ano típico para cada estação anemométrica**

A partir do ano típico, diversas incoerências e falhas passíveis de estarem presentes no dado utilizado são suprimidas, resultando em um dado bem-comportado para análise. Nota-se que, apesar das vantagens da utilização de anos típicos, compostos por períodos diferentes de medição, verifica-se tendência em gerar diferenças nos valores médios obtidos pelas séries, uma vez que é notada a variabilidade sazonal do recurso eólico [3] [8] [9].

Para corrigir este comportamento, é essencial considerar apenas as variações relativas na velocidade das séries com relação a suas médias. A normalização das séries permite também analisar sobre mesmo aspecto, dados de torres com características parecidas, porém diferentes velocidades médias, que podem representar de forma equivalente uma região em comum, apresentada no item a seguir.

### 2.3 Matriz de Correlação entre Anos Típicos

Para evitar a influência dos valores médios de vento das séries foi calculada a matriz de correlação entre os anos típicos das torres anemométricas. Tal procedimento pondera a variação conjunta de dois objetos, com relação às suas médias, retirando a dependência dos valores médios no seu processo.

Seu resultado é um índice, com valores entre (-1) e (1), para cada combinação (2 a 2) de torres. Os valores (-1) do índice representam variações simetricamente opostos com relação às suas médias e valores (1), variações simetricamente alinhadas com relação a suas médias. Correlações nulas indicam dados que não apresentam qualquer tipo de padrão de variação conjunta. [10]

Por criar um valor associado a um par de objetos, o cálculo de correlações entre múltiplos objetos resulta em uma matriz de correlação de objetos onde cada relacionamento (1 a 1) é a correlação entre as torres, neste caso. A representação desta matriz de correlação pode ser feita a partir de mapas de cor, identificando de forma visual os padrões entre as estações. A Figura 4 mostra um exemplo deste procedimento, onde cada quadrado apresenta, a partir da escala de cores, um valor de correlação obtido entre a estação no eixo vertical e a estação no eixo horizontal. Nota-se que os valores obtidos na diagonal principal são sempre iguais à 1, uma vez que é resultado da correlação do ano típico da torre com ele próprio.

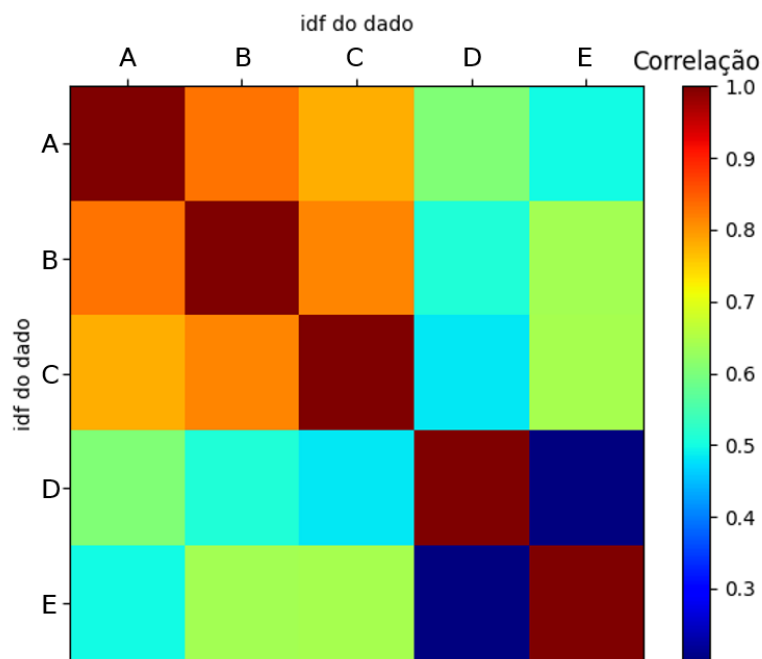


Figura 4 - Matriz de correlação

A Figura 4 permite observar que entre os objetos A, B e C há uma similaridade acentuada, quando comparado com os demais dados do exemplo, identificando uma tendência de agrupá-los. Adicionalmente, nota-se que os demais objetos, D e E, não apresentam correlações altas entre si, uma vez que obtêm a menor correlação da amostra, não tendendo a formar um grupo adicional.

Verifica-se, no entanto, que a definição de valores limites para a adição de objetos a um dos agrupamentos e a existência de iterações de múltiplos objetos com diferentes valores nestes grupos torna a clusterização de amostras grandes um problema complexo.

Como exemplo prático: a amostra C apresenta valores de correlação de aproximadamente 0,85 com relação a amostra B e 0,75 com relação a amostra A. Há diversas métricas que consideram a influência de ambas as amostras, como por exemplo: a média dos valores das amostras, o valor mínimo com relação as amostras etc. Caso o valor limite para agregação fosse definido como 0,8 o valor médio agregaria as três estações e o valor mínimo apresentaria um outro resultado. Assim sendo, é necessário definir um procedimento padronizado para a aplicação em matrizes mais complexas, com métricas e limites definidos.

## 2.4 Metodologia de Clusterização de Dados

Para a definição da métrica e limites para agregação das séries, foi selecionada a metodologia da Clusterização Hierárquica como procedimento. A clusterização hierárquica pode ser realizada a partir de métodos “top-down” ou “bottom-up”. Neste tipo de procedimento a clusterização é apresentada de maneira sequencial, isto é: a partir de incrementos ou reduções sequenciais nos limites de agregação, para uma definida métrica. Com isso, os modelos “bottom-up” tendem a sequencialmente unir clusters diversos e as metodologias “top-down” tendem a segregar clusters de maneira sequencial.

Considerando as características dos dados de entrada, onde as regiões tendem a se distanciar entre si, e as diversas possibilidades de iterações entre os agrupamentos, a metodologia “bottom-up” foi identificada como preferencial, uma vez que apresenta menor risco de separação de estações próximas entre si em grupamentos distintos.

A métrica para o cálculo da clusterização utiliza a “posição” do objeto em um espaço de “n” dimensões e calcula a distância para os demais objetos existentes. Após o cálculo das distâncias, as que apresentam valores abaixo dos limites definidos são agregadas em um cluster único. Existem diversas formas de cálculo de distância sendo a mais comum a Distância Euclidiana que é dada pela Equação 1.

$$d = \sqrt{\sum_{i=1}^n (b_i - a_i)^2} \quad \text{Eq 1}$$

Sendo:  $d$  a distância entre dois objetos em um espaço “n-dimensional”;  $n$  o número total de dimensões;  $a_i$  o valor da  $i$ -ésima coordenada do objeto A;  $b_i$  o valor da  $i$ -ésima coordenada do objeto B.

No cenário desta Nota Técnica as “posições” das torres meteorológicas são as séries ordenadas de valores de correlação obtidos entre as torres, que serão comparados aos valores obtidos pelas demais torres. No exemplo da Figura 4, o ponto do objeto A ocupa a posição [ $a_1 = 1,0$   $a_2 = 0,83$   $a_3 = 0,78$   $a_4 = 0,6$   $a_5 = 0,5$ ], que é equivalente à sua série de correlação, isto é, a primeira linha da matriz de correlação. A distância é obtida com relação posições definidas nas séries de correlação das demais torres.

Uma vez definidas as posições, a ponderação pelos clusters é realizada a partir do método de Ward [11] [12]. Neste método, o objetivo da clusterização é realizar a unificação de dois clusters, selecionando as agregações a partir da minimização do incremento da variância das distâncias dos objetos com relação aos centroides.

Na Figura 5 é apresentado o exemplo deste processo. No primeiro esquemático, são calculadas as distâncias dos objetos dos três clusters ( $\alpha$ ,  $\beta$ ,  $\gamma$ ) em relação ao centroide de cada cluster. A partir destas distâncias, é possível então calcular a variância a partir da soma de seus quadrados, isto é:

$$\sigma^2_c = \sum_{i \in C} (P_i - \bar{P}_c)^2 \quad \text{Eq. 2}$$

Sendo:  $P_i$  é a posição do objeto  $i$  pertencente ao cluster  $C$ ;  $\bar{P}_c$  a posição do centroide do cluster  $C$

Após calcular as variâncias originais de cada cluster, o método identifica três opções de agregação a partir da união dois a dois destes conjuntos, que possuem novos centroides com os objetos de ambos os conjuntos originais. As distâncias de cada um dos objetos até o novo centroide são calculadas e, o que apresentar o menor incremento de variância é selecionado.

No exemplo da Figura 5, é possível observar visualmente que o menor incremento de distância é obtido na união dos clusters  $\alpha$  e  $\beta$ , que tem a distância dos objetos até o cluster final apresentado com linhas pontilhadas. Ao fim desta iteração serão obtidos dois clusters, um formado pela união do  $\alpha$  e  $\beta$  e outro composto pelo  $\gamma$ .

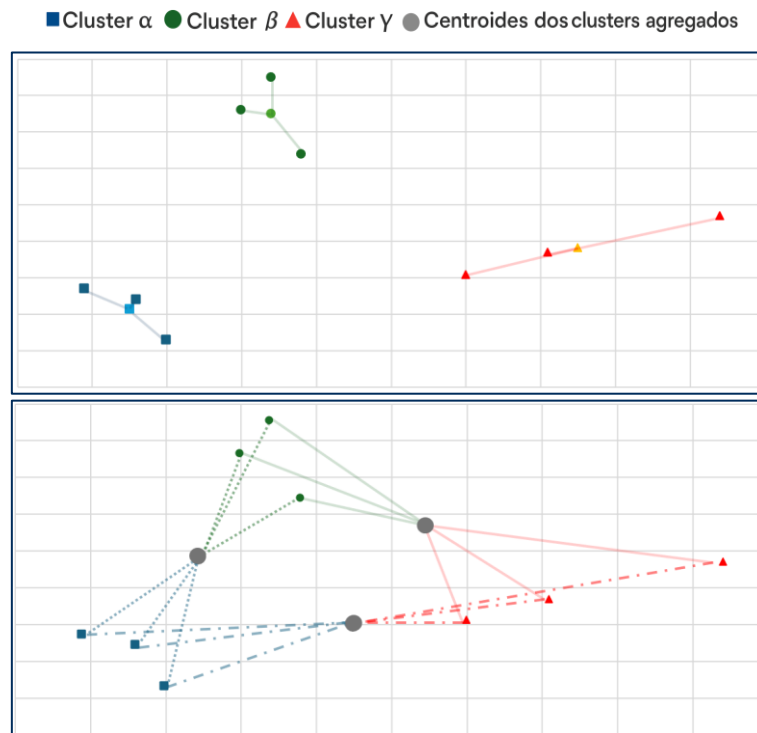


Figura 5 - Exemplo do cálculo de distâncias pelo método de Ward

Essa operação é realizada sequencialmente, agregando os clusters a partir dos menores incrementos de variância, até o momento que seja obtido como resultado apenas um cluster. Este procedimento também pode ser observado a partir de um Dendrograma, que apresenta no eixo x os objetos a serem agregados e no eixo y é apresentada a distância. Ao agregar dois objetos, uma linha horizontal, com o valor da distância do novo cluster, conecta duas linhas verticais, que indicam os dados agregados no processo.

Aplicando o método de Ward aos dados apresentados na Figura 4 obtém-se a sequência de agregações apresentadas na Figura 6.

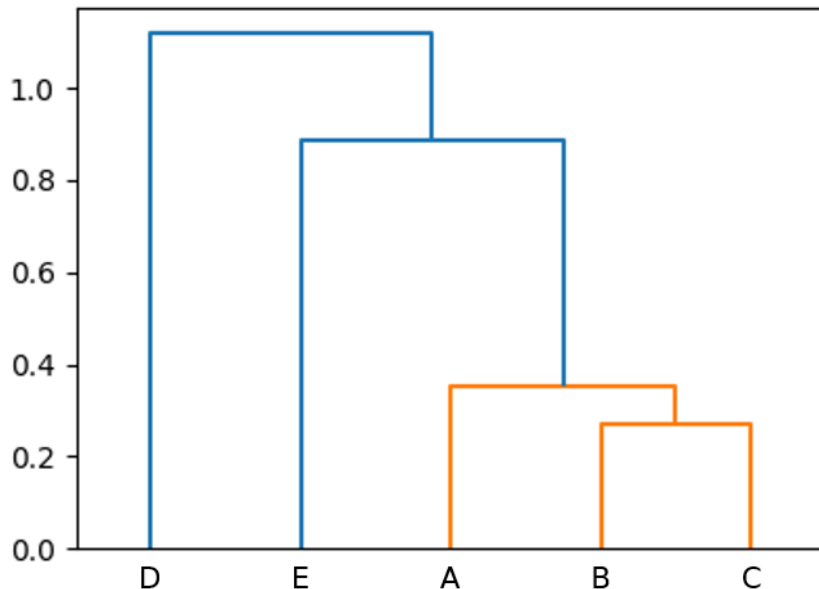


Figura 6 - Dendrograma das clusterizações dos dados exemplificados na Figura 4

A partir da observação do Dendrograma, é possível identificar que: apesar da alta correlação entre os objetos A e B, os menores desvios totais (variância) obtidos pelo método de Ward foram entre o objeto B e C, resultando na primeira agregação observada, seguido pela agregação dos objetos A, E e D ao agrupamento.

Esta observação da agregação inicial dos objetos B e C mostra a importância da metodologia em permitir a agregação de regiões a partir de uma visão sistêmica dos dados disponíveis, tendo cada correlação existente na amostra, uma influência na formação de cada um dos grupos.

## 2.5 Aplicação da Metodologia nos Dados do Sistema AMA

A metodologia apresentada permite a identificação das prioridades da agregação de torres meteorológicas em função da similaridade das correlações observadas pelos anos típicos obtidos a partir das medições nos anemômetros superiores das estações cadastradas no Sistema AMA.

Nota-se, no entanto, que a metodologia apresentada não indica um valor fixo a ser considerado para a agregação, sendo uma escolha do usuário a definição do(s) ponto(s) ótimo(s) de distância a serem considerados.

Os níveis de agregação a serem considerados se mostram então uma etapa principal da decisão dos clusters finais. Considerando uma árvore de clusterização complexa, como a apresentada neste estudo, que possui mais de 650 amostras a serem agregadas, é possível que o nível esperado de coerência entre os dados retorne um número muito grande de clusters finais, enquanto um valor razoável de clusters retorne agregações muito grosseiras.

Por um lado, quanto maior o número de clusters, mais detalhados e representativos são os dados de cada grupo, o que é desejável. Por outro, é possível chegar a clusters tão discretizados quanto um complexo de somente dois parques eólicos, o que deve ser evitado em função dos riscos discutidos na seção 1.2.

Facilitando o processo de decisão e a análise dos resultados obtidos, foram selecionados 10 níveis de amostragem na metodologia, onde cada nível apresenta uma agregação maior dos dados observados. Um exemplo de diagrama simplificado está apresentado na Figura 7. Neste diagrama é possível identificar oito pontos de amostragem, resultando em 9, 8, 7, 6, 5, 4, 3, 2 clusters, respectivamente.

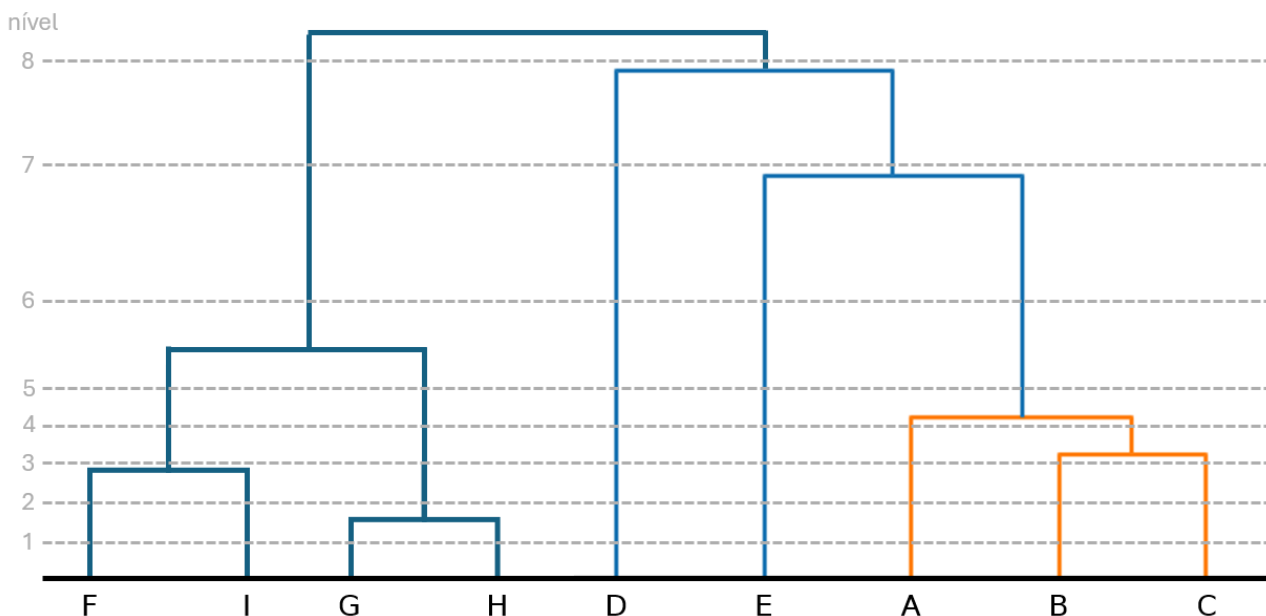


Figura 7 - Procedimento de amostragem dos clusters

Para realizar a seleção dos melhores níveis de agregação é importante a comparação entre os diferentes níveis obtidos para o mesmo cluster final. No exemplo da Figura 7 é possível, a partir dos 2 clusters finais, realizar uma árvore de agregações mostrando a evolução dos clusters, que seriam representados individualmente.

De forma ilustrativa, o cluster à esquerda na Figura 7 está detalhado abaixo. É possível identificar que os oito níveis de amostragem resultaram em diferentes configurações de clusters. No nível 1, nível de amostragem menos agregado, existem 4 clusters resultantes que foram enumerados de 1 a 4 da esquerda para direita, resultando no identificador (1-x) para este nível.

Os níveis 3, 4 e 5 e níveis 6, 7 e 8 da amostragem apresentam as mesmas configurações de clusters, uma vez que não há agregação de objetos para o incremento de “distância” definido. Com isso, os Clusters 3-2, 4-2 e 5-2 são iguais, assim como o 6-1, 7-1 e 8-1.

Por praticidade, a apresentação do dendrograma pode ser substituída pela representação em árvore, proposta na Figura 8.

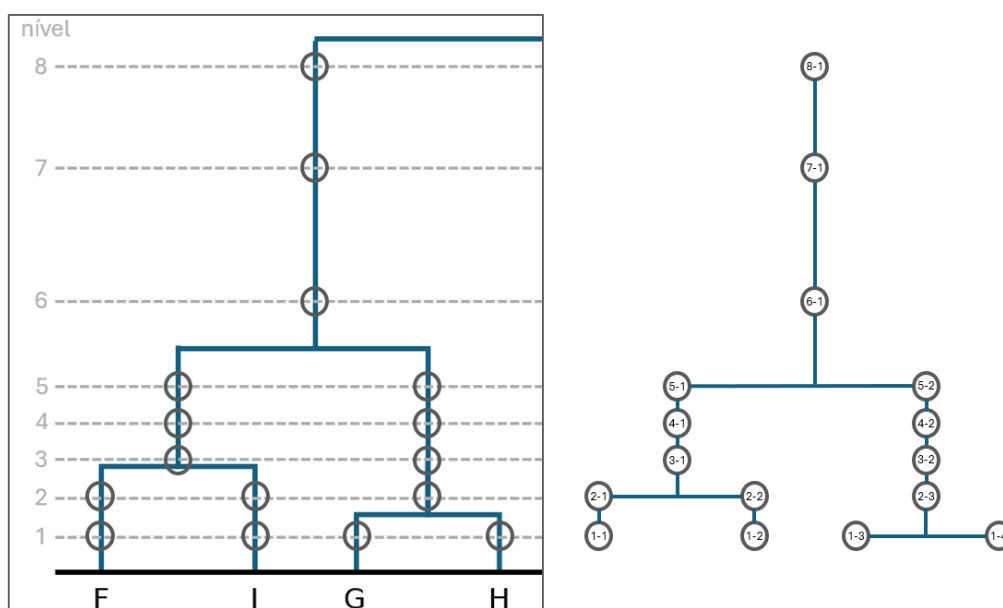


Figura 8 - Visualização dos clusters obtidos a partir da amostragem vertical

Um exemplo de representação das agregações observadas na aplicação dos dados do Sistema AMA está apresentado na Figura 9.

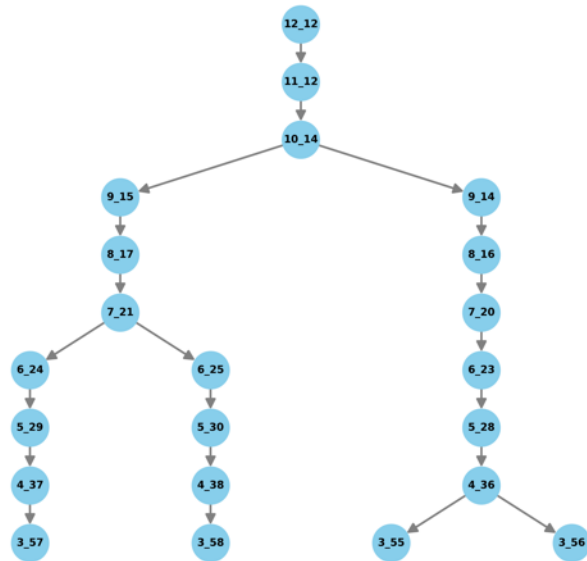


Figura 9 - Representação em árvore dos clusters

A partir deste processo, é possível analisar os dados que compõem cada um dos clusters individualmente, realizando comparações entre a qualidade dos dados resultantes dos níveis existentes no modelo e selecionando o ponto ideal a ser considerado para cada árvore.

Para a aplicação com dados do Sistema AMA, a quantidade de clusters nos pontos de amostragem variou entre 18 clusters, no nível mais agregado, e 75 clusters, no nível mais desagregado.

Considerando a diversidade de características dos parques eólicos pertencentes ao Sistema AMA, foi observada que a seleção de um valor único de corte para todas as árvores de clusterização resultaria na redução da qualidade dos dados disponibilizados a sociedade. Com isso, cada uma das 18 árvores de clusterização obtidas foi analisada individualmente, permitindo definições dos pontos mais adequados e considerando o contraponto entre a quantidade de dados e a qualidade das informações.

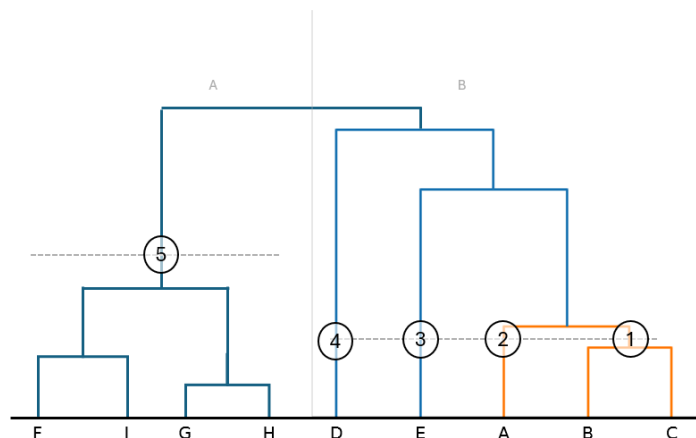


Figura 10 - Exemplo de análise individual por árvore

Na Figura 11 é possível observar geograficamente o processo de agregação ocorrendo de acordo com a redução da quantidade de clusters.

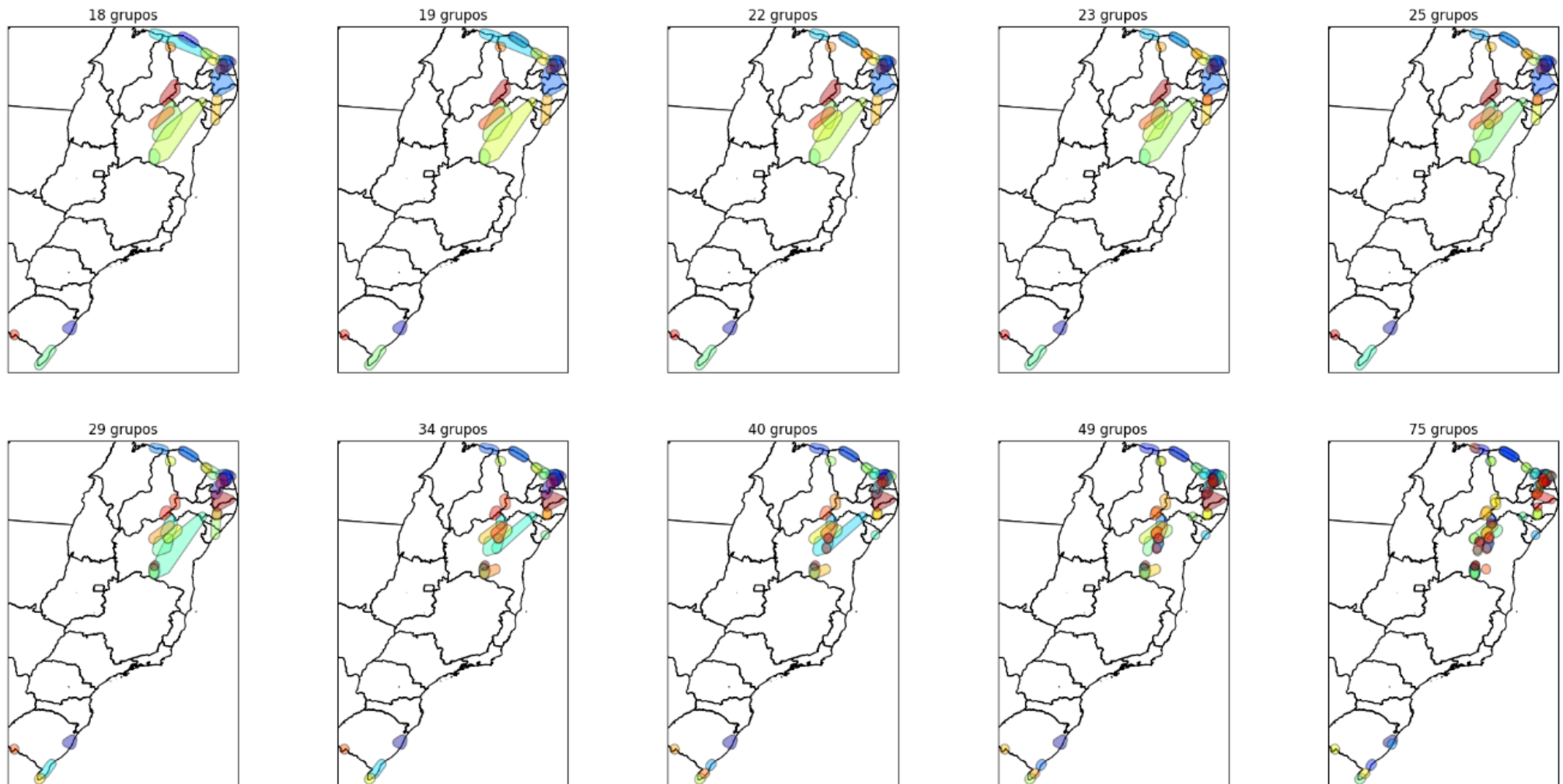


Figura 11 – Apresentação geográfica do processo de agregação de clusters

O procedimento de análise individual das 18 árvores de agregação e a definição dos pontos de corte individualizados resultou em 27 clusters finais, apresentados com maior detalhamento geográfico na Figura 12.

O processo de definição dos pontos de corte contou com análises específicas para cada uma das opções de clusters que compõem as árvores de agregação. Neste processo os dados foram analisados, de maneira sistemática, para avaliar os impactos da redução da quantidade de clusters na qualidade final dos dados, além de observações das características locais de cada uma das torres de medição.

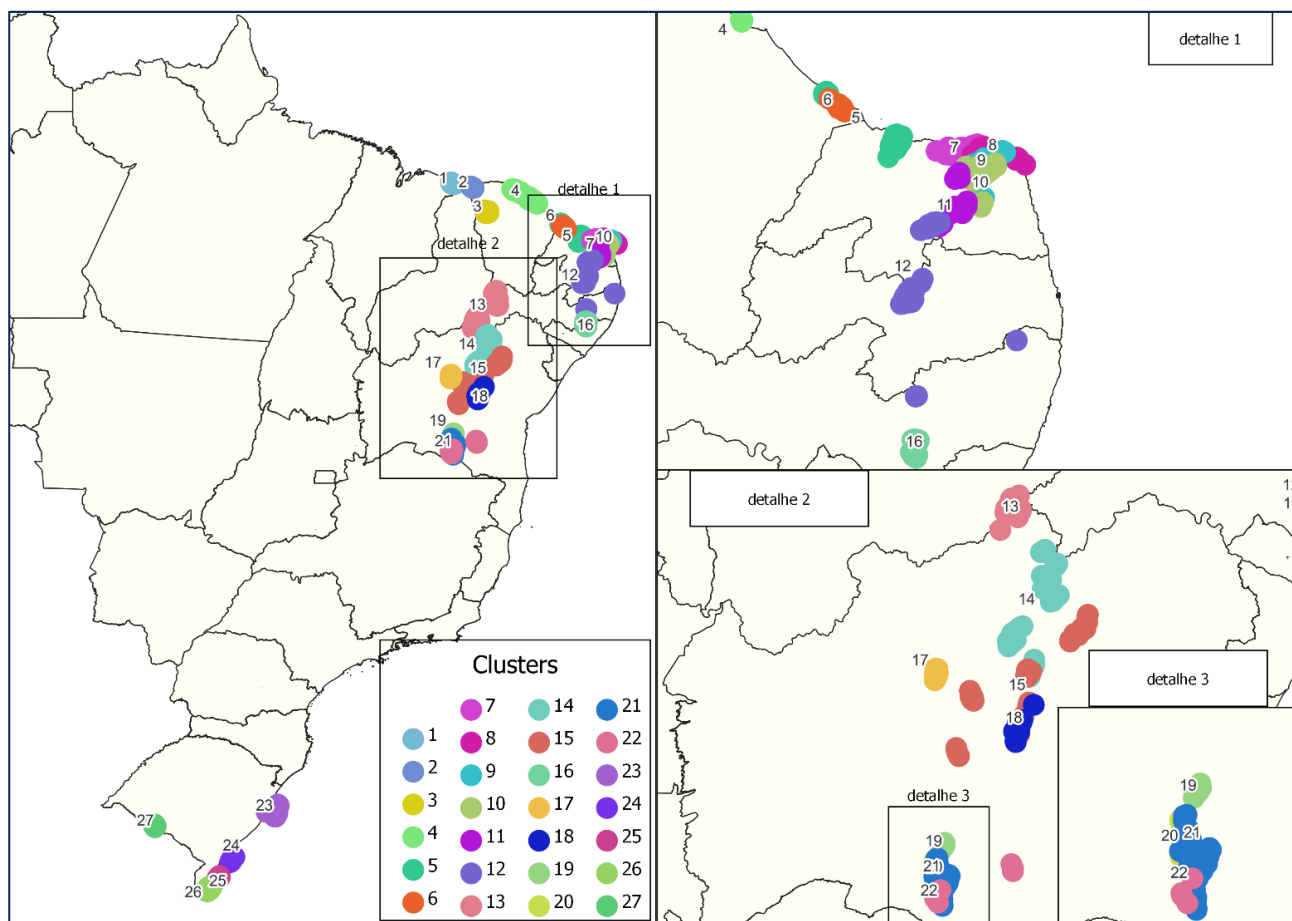


Figura 12 - Clusters finais após análise dos resultados do modelo

A partir da Figura 12 é possível observar que os clusters finais apresentam sobreposições geográficas. Com isso é possível, a partir dos resultados desta metodologia, supor que uma mesma área apresente mais de um perfil de vento, em função de suas características locais, técnicas ou ambientais. Tal comportamento é esperado, uma vez que as variáveis que influenciam as medições de vento nas torres são diversas, heterogeneamente espaçadas e não contínuas. Com isso, fenômenos distintos que afetam o comportamento do vento podem ocorrer em duas regiões próximas, sobrepostas a uma terceira região com um comportamento diferente.

Tal concomitância de diferentes clusters numa mesma região se mostra vantajosa, uma vez que permite análises mais detalhadas das possibilidades de comportamento das regiões, estudos com relação às características específicas do local e identificação da melhor opção de uso para a mesma região.

## 3 Apresentação dos Dados

---

### 3.1 Dados e Grupamentos

Conforme apresentado no item 2, foram definidos 27 clusters a partir da aplicação da metodologia e análises apresentadas nesta Nota. Estes clusters estão apresentados agrupados em regiões na seção atual, com suas características climatológicas e locais detalhadas.

Todos os dados apresentados a seguir estão consolidados de acordo com a disponibilidade de informações na data de elaboração deste estudo, sendo atualizáveis na [plataforma online](#) de disponibilização de informações elaborada pela EPE em seu *site*. Nessa plataforma estão disponíveis para download: séries horárias de velocidade, temperatura, umidade relativa, pressão e a porcentagem de ocorrências de direção do vento em cada setor, por cluster.

Considerando a variabilidade, dinamicidade e complexidade dos dados e das informações disponíveis, os resultados a seguir indicados devem ser considerados apenas para fins consultivos, uma vez que podem ser atualizados de acordo com a inclusão de novas medições e torres anemométricas.

Cada uma das regiões está compilada em uma ficha informativa apresentando as seguintes características: identificação do cluster, intervalos dos valores de temperatura e pressão médias observadas além das alturas e altitudes do anemômetro de referência, distribuição de ocorrência da velocidade média do vento e distribuição de Weibull minimizada; rosa dos ventos e distribuição da direção do vento ao longo do mês e do ano além dos gráficos com o dia médio e ano médio de cada cluster.

### 3.1.1. Clusters na Região Litoral Equatorial

Os clusters da Região **Litoral Equatorial** encontram-se localizados no litoral e em regiões até 50 km da costa entre os estados do RN e MA. Nota-se a variação da direção principal do vento do setor NE, nos clusters mais a noroeste desta região, para direções predominantes mais próximas do setor E-SE, conforme se aproximam do extremo Leste do continente. Observa-se baixa variabilidade diária para maior parte dos clusters observados, com ressalva para o comportamento mais marcante do Cluster 4, com valores máximos de velocidade próximos das 13 horas (Figura 15).

A Região **Litoral Equatorial** apresenta como principais características: variação sazonal da velocidade média, na ordem de 5 m/s, apresentada entre o primeiro e segundo trimestres, períodos estes com menores velocidades médias; e direções predominantes entre NE, E, SE em todo o período.

### Apresentação Geral da Região Litoral Equatorial

Tabela 1 - Características dos clusters da Região Litoral Equatorial

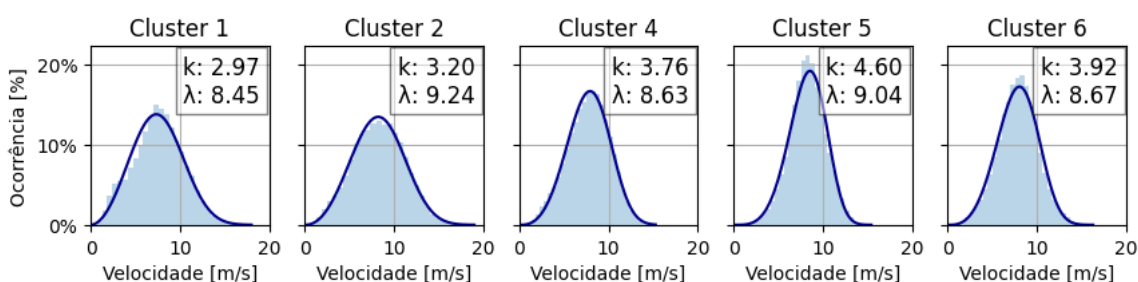
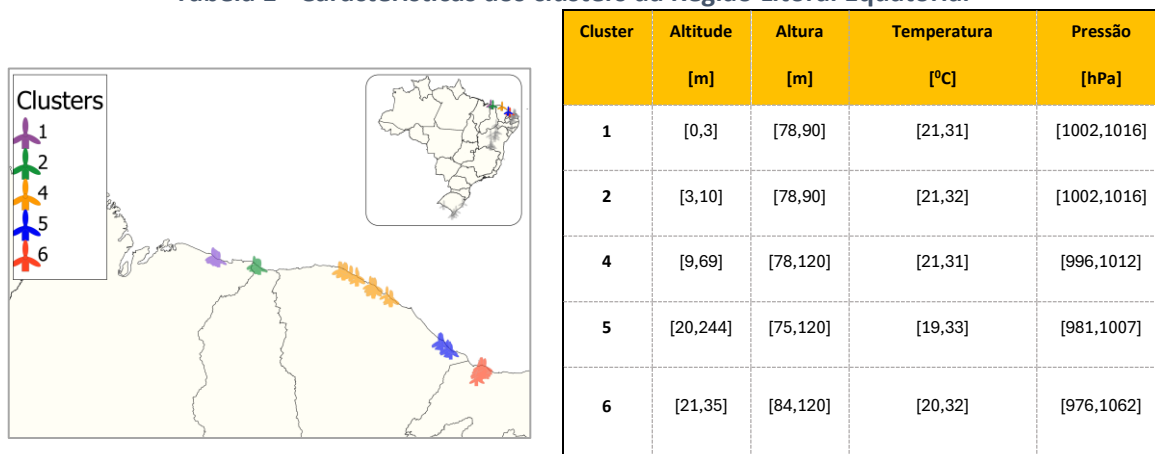


Figura 13 - Distribuição de ocorrências e estimativa de Weibull dos clusters da Região Litoral Equatorial

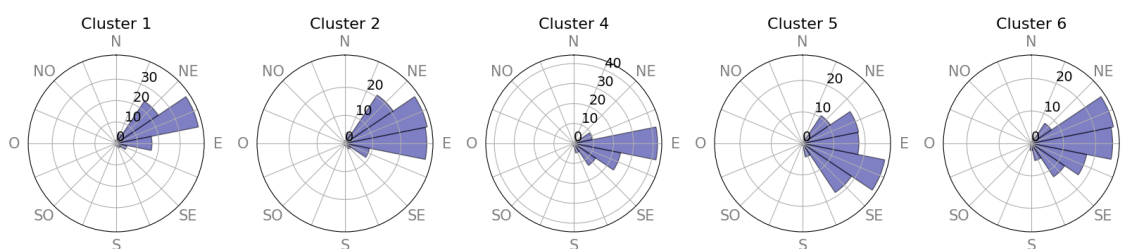


Figura 14 – Rosa dos ventos dos clusters da Região Litoral Equatorial

## Descrição Temporal da Região Litoral Equatorial

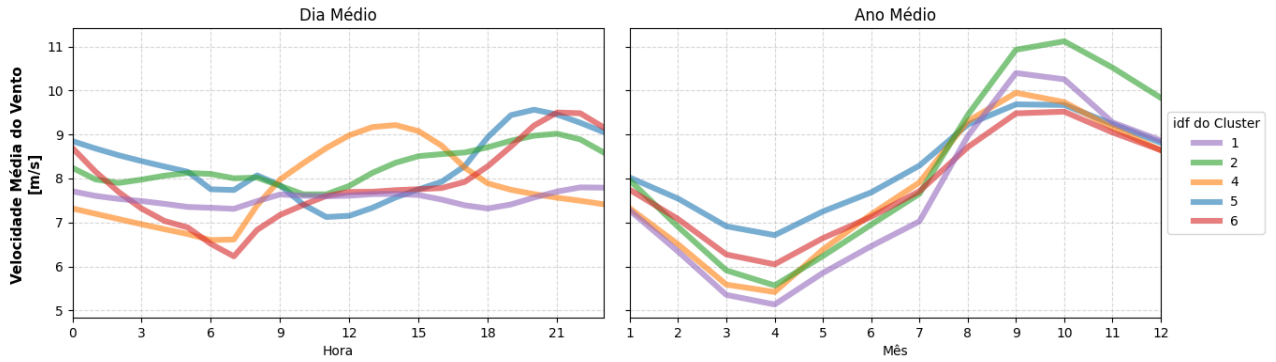


Figura 15 - Dia médio e ano médio da velocidade média do vento nos clusters da Região Litoral Equatorial

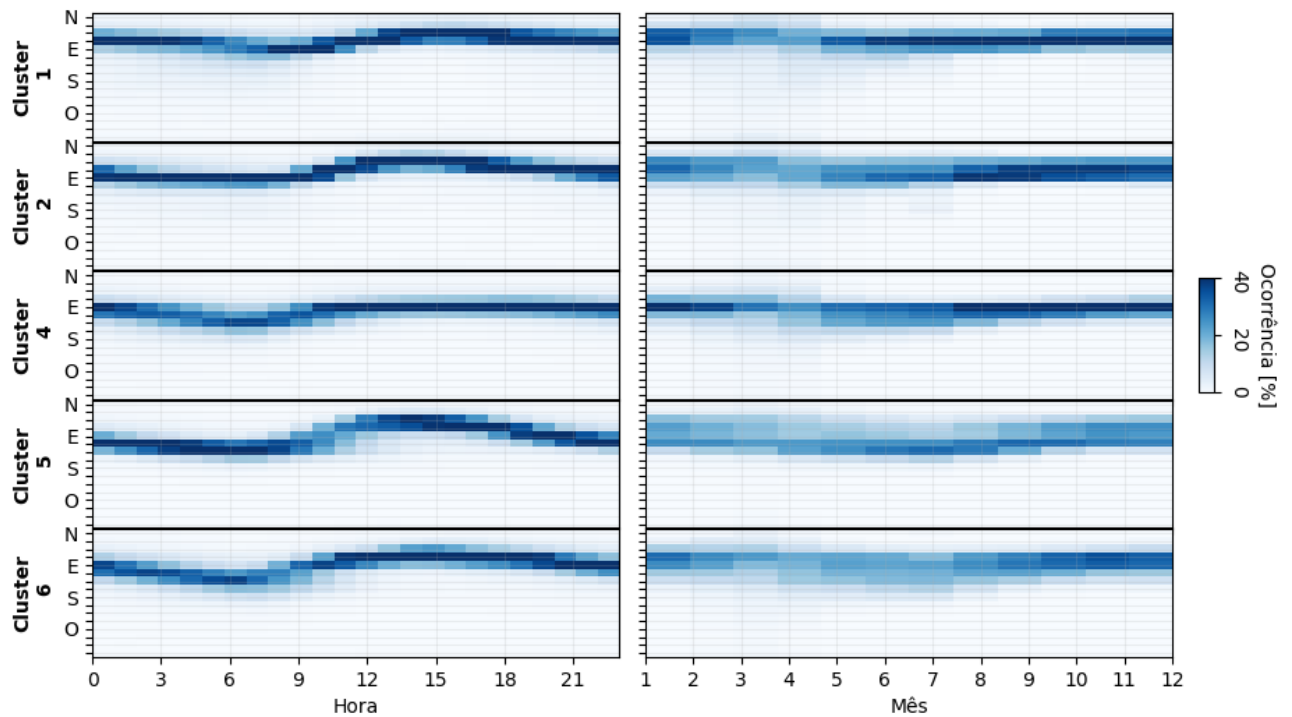


Figura 16 - Distribuição horária e mensal da direção média do vento nos clusters da Região Litoral Equatorial

### 3.1.2. Clusters na Região Rio Grande do Norte

A região denominada **Rio Grande do Norte** apresenta torres nas regiões litoral norte do estado do Rio Grande do Norte e central do estado. Os clusters desta região apresentam distribuições de velocidade e direção coerentes entre si. Há baixa variabilidade de direção, sendo os setores mais marcantes entre o E e SE, que se mantêm estáveis ao longo do dia e do ano, e máximos de velocidade observados no segundo semestre do ano. O Cluster 7 apresenta a maior variação de velocidade média no perfil diário da região, em torno de 3 m/s.

#### Apresentação geral da Região Rio Grande do Norte

Tabela 2 - Características dos clusters da Região Rio Grande do Norte

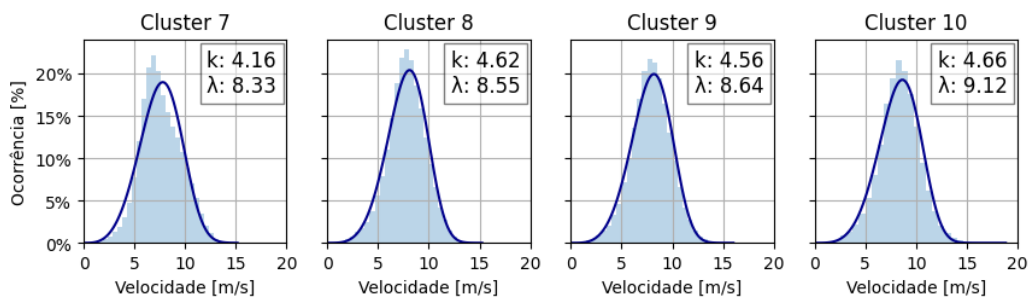
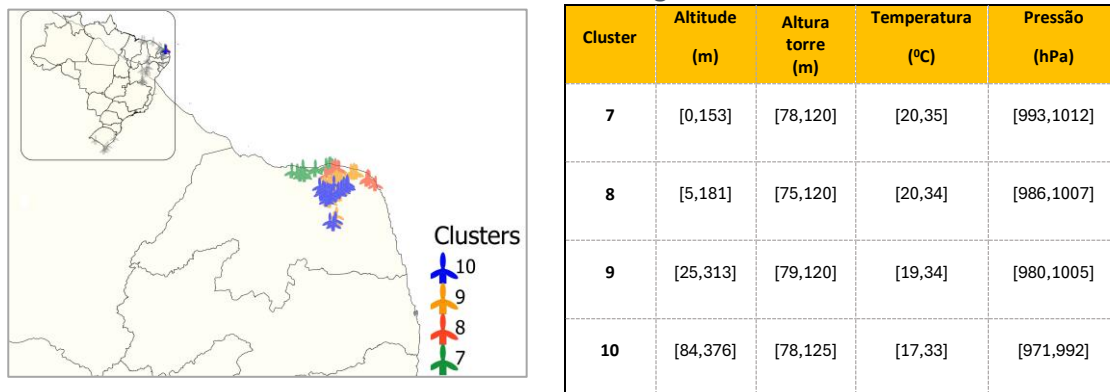


Figura 17 - Distribuição de ocorrências e estimativa de Weibull dos clusters da Região Rio Grande do Norte

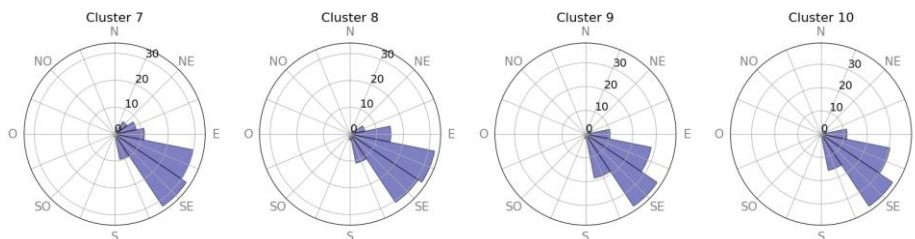


Figura 18 – Rosa dos ventos dos clusters da Região Rio Grande do Norte

## Descrição temporal da Região Rio Grande do Norte

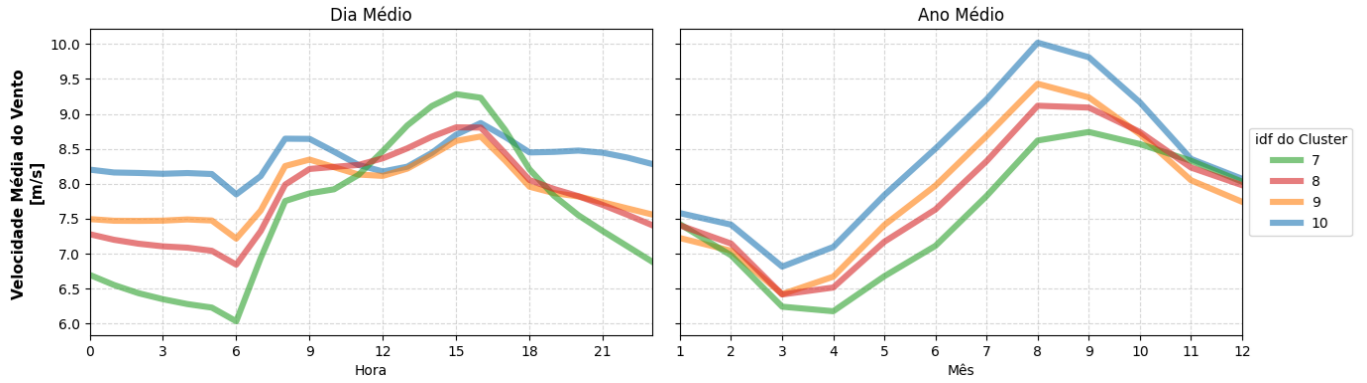


Figura 19 - Dia médio e ano médio da velocidade média do vento nos clusters da Região Rio Grande do Norte

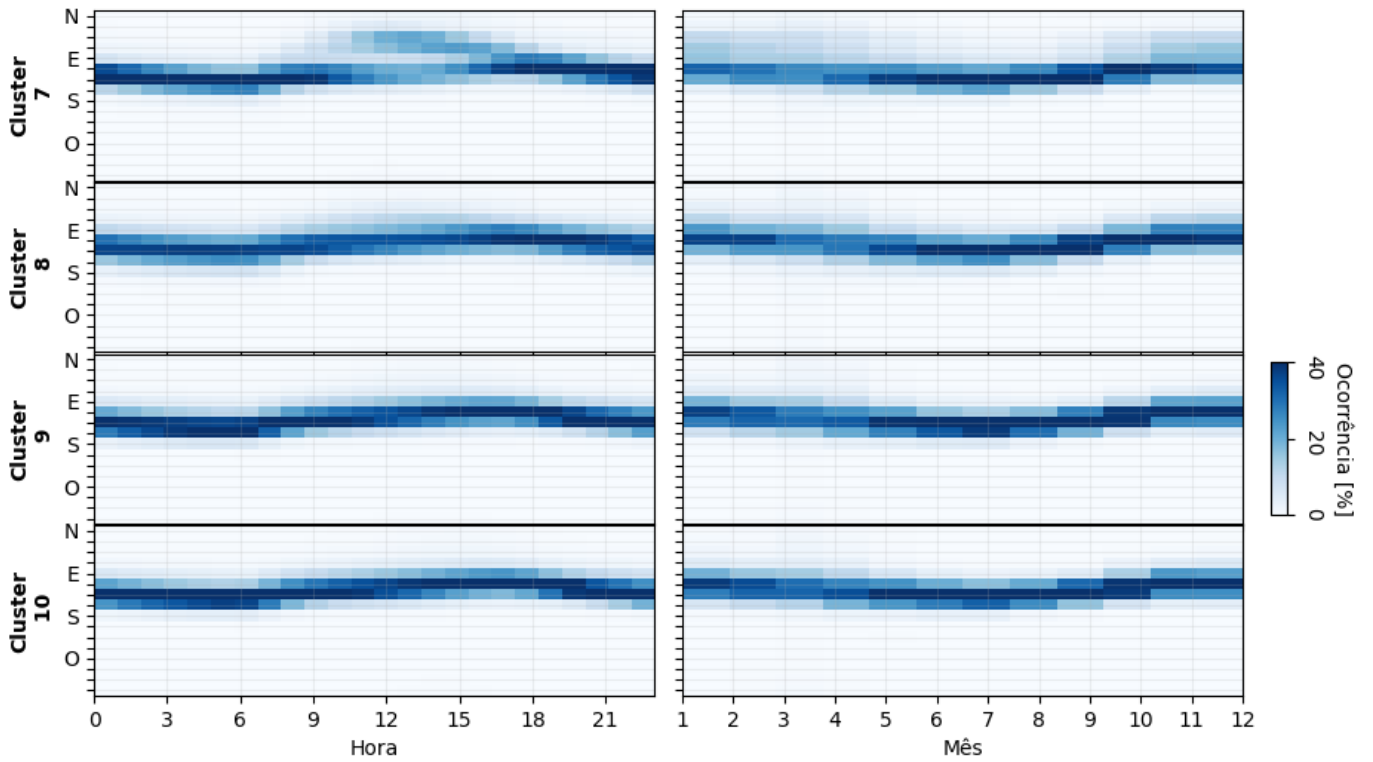


Figura 20 - Distribuição horária e mensal da direção média do vento nos clusters da Região Rio Grande do Norte

### 3.1.3. Clusters na Região Centro NE

Os Clusters na Região denominada como **Centro NE** estão localizados entre Pernambuco e Rio Grande do Norte, além do Cluster 3 que se encontra no Ceará, próximo da fronteira com Piauí. Os clusters desta região apresentam grande variabilidade de altitudes, sendo o caso mais marcante as torres do Cluster 16, que varia entre 4 e 974 metros.

Nota-se que as direções de vento são bem-marcadas e o comportamento diário das estações próximas entre si, apresentam redução da velocidade média a partir de 8hs e incremento entre 13hs e 15hs. Com exceção do Cluster 16, que apresenta variabilidade próxima a 2 m/s, a variabilidade sazonal das estações é aproximadamente 3,5 m/s. Os Clusters desta região, assim como apresentado nos Clusters da Região Rio Grande do Norte, possuem direções marcantes, entre os setores E-SE.

### Apresentação geral da Região Centro NE

Tabela 3 - Características dos clusters da Região Centro NE

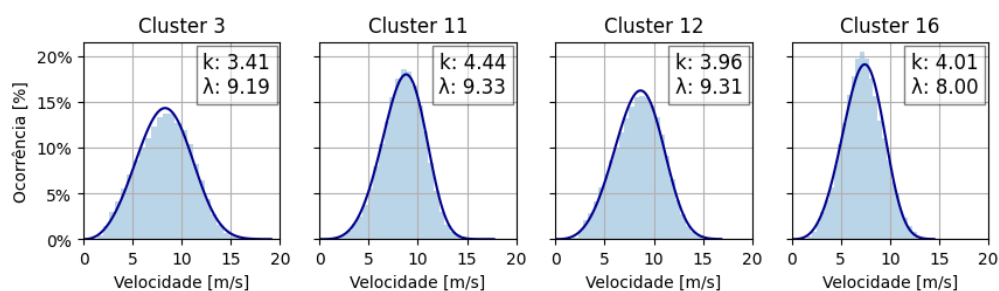
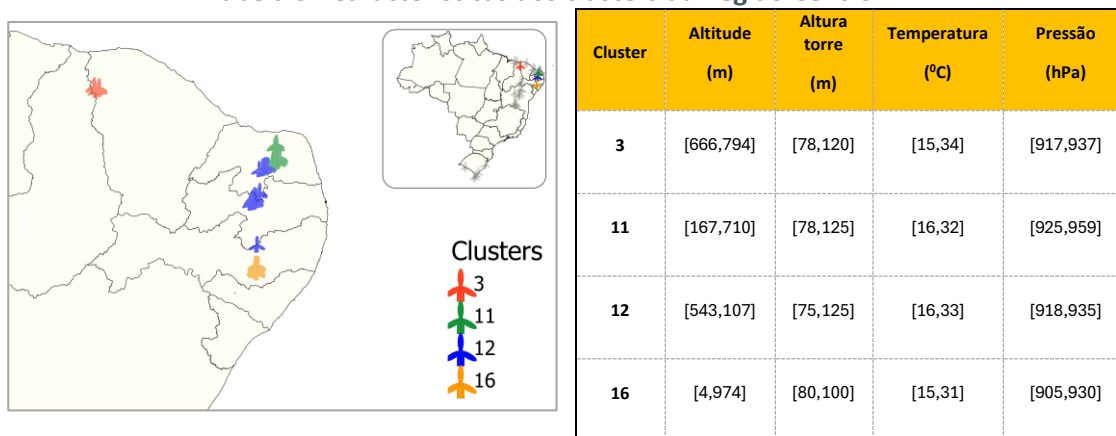


Figura 21 - Distribuição de ocorrências e estimativa de Weibull dos clusters da Região Centro NE

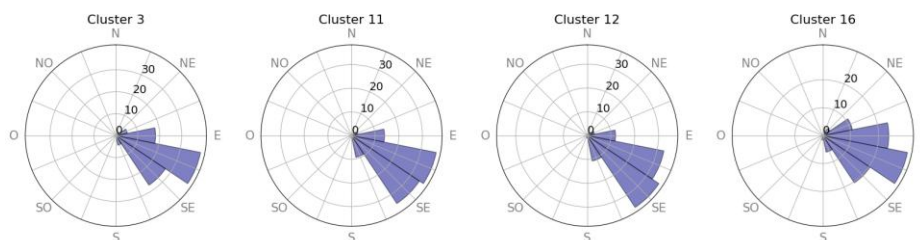


Figura 22 – Rosa dos ventos dos clusters da Região Centro NE

## Descrição temporal da Região Centro NE

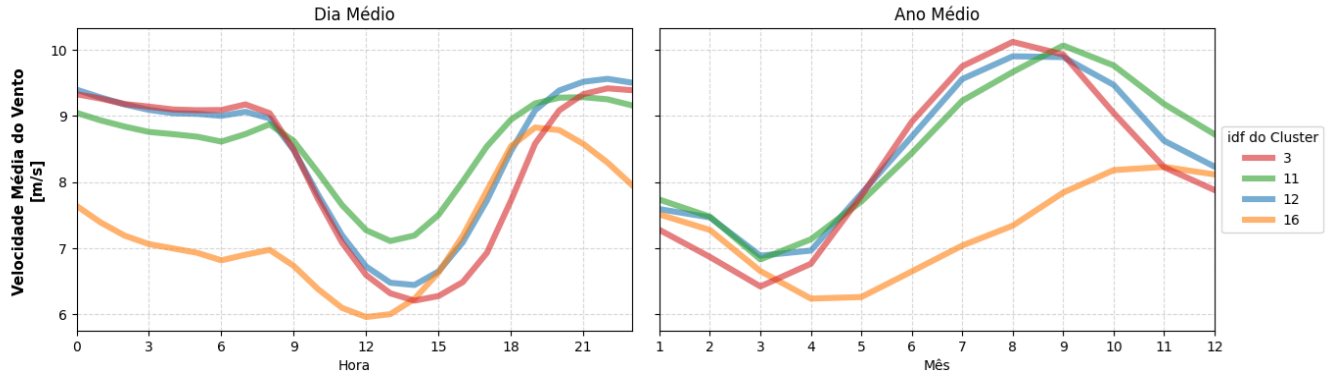


Figura 23 - Dia médio e ano médio da velocidade média do vento nos clusters da Região Centro NE

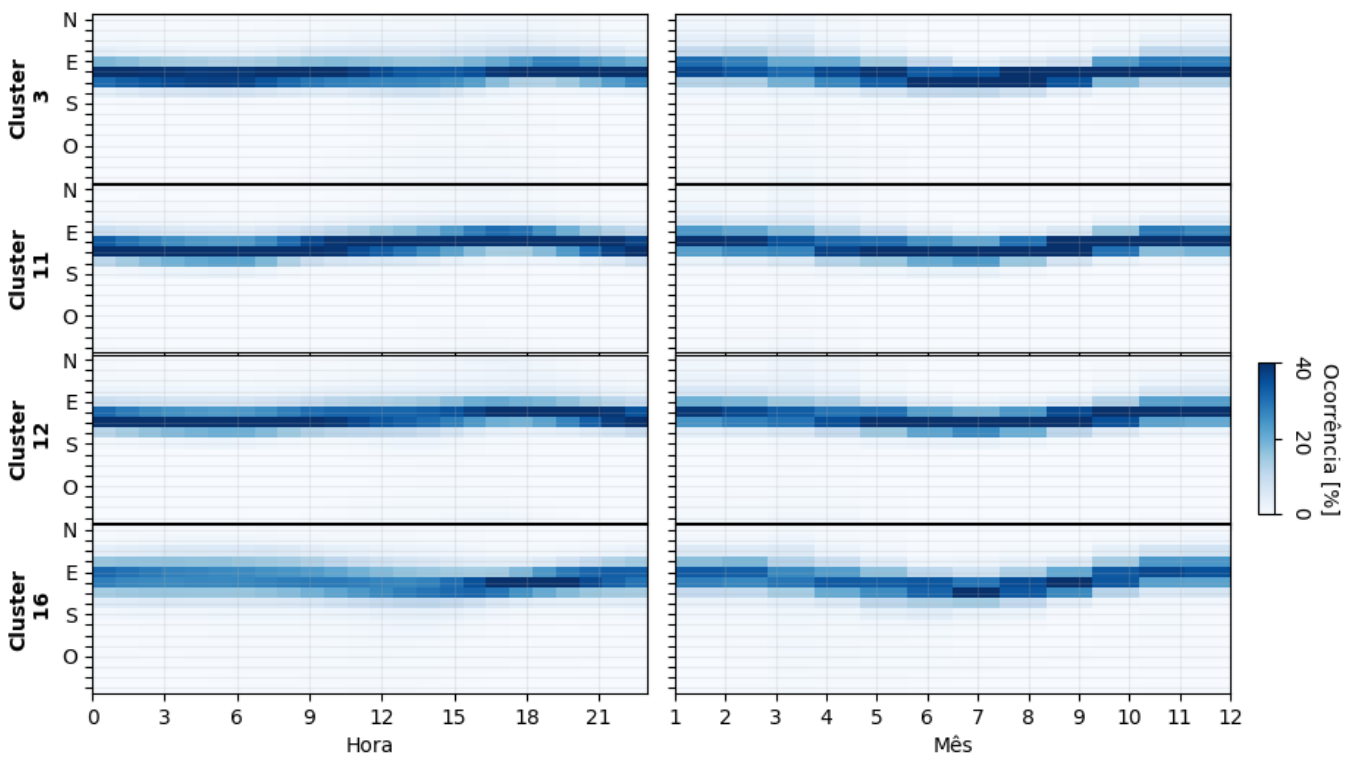


Figura 24 - Distribuição horária e mensal da direção média do vento nos clusters da Região Centro NE

### 3.1.4. Clusters na Região Interior NE

Os clusters da região denominada **Interior NE** estão situados entre os estados da Bahia e Piauí, instalados em altitudes superiores à 500 metros. Nesta região, as direções dos ventos são bastante marcadas, predominantemente nos setores E-SE e o Cluster situado no Piauí direções E-SE e SE. Ressalte-se a grande diferença entre distribuição de ocorrências de velocidade nos quatro clusters (Figura 25), com picos de velocidade próximos a 5 m/s no Cluster 13 e próximos a 10 m/s no Cluster 21. Em nenhum dos clusters observados é identificado um ajuste eficiente da curva de Weibull, sendo possível cogitar a possibilidade de aplicação de outras distribuições de probabilidade que melhor se adequem às distribuições de ocorrência.

Trata-se de uma região com alta variabilidade da velocidade média ao longo do ano e dia, com valores mínimos próximos das 15 horas. No perfil anual, assim como observado na maior parte dos Clusters da região Nordeste, os ápices de velocidade no segundo semestre.

### Apresentação geral da Região Interior NE

Tabela 4 - Características dos clusters da Região Interior NE

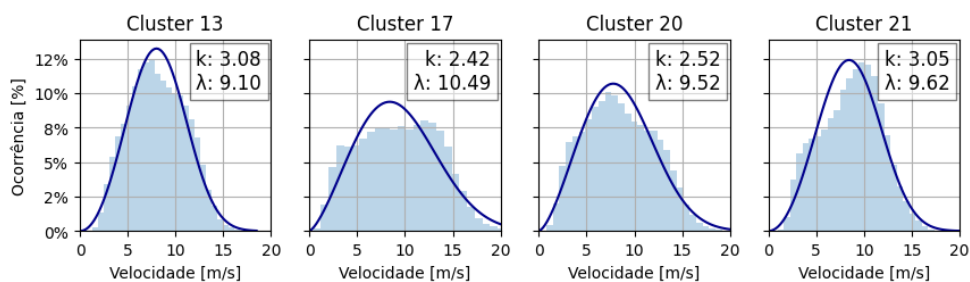
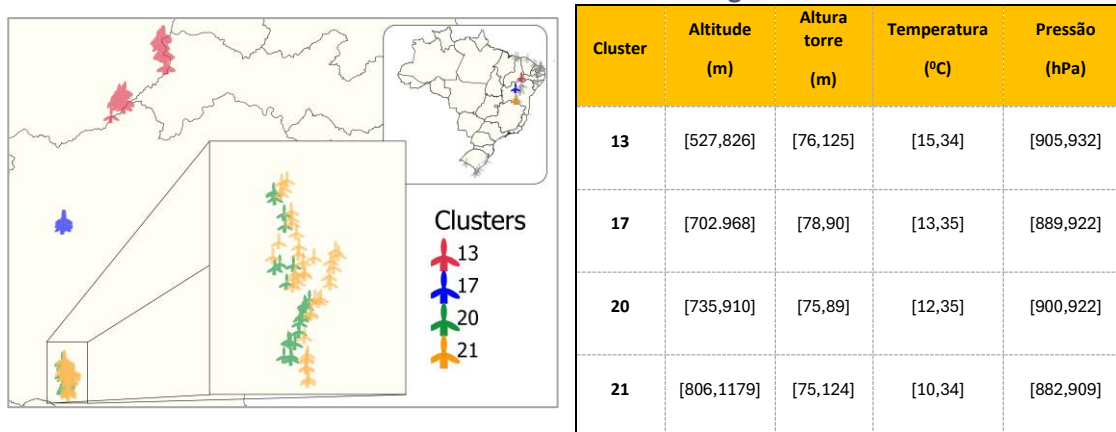


Figura 25 - Distribuição de ocorrências e estimativa de Weibull dos clusters da Região Interior NE

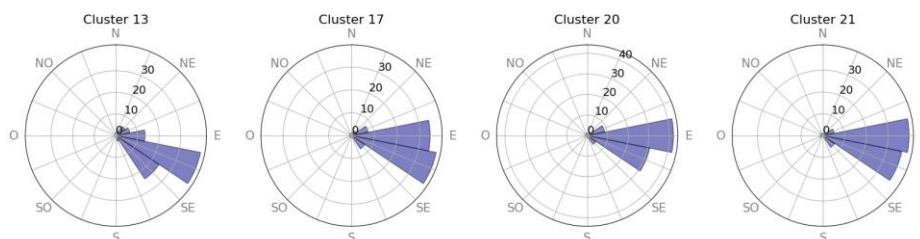


Figura 26 – Rosa dos ventos dos clusters da Região Interior NE

## Descrição temporal da Região Interior NE

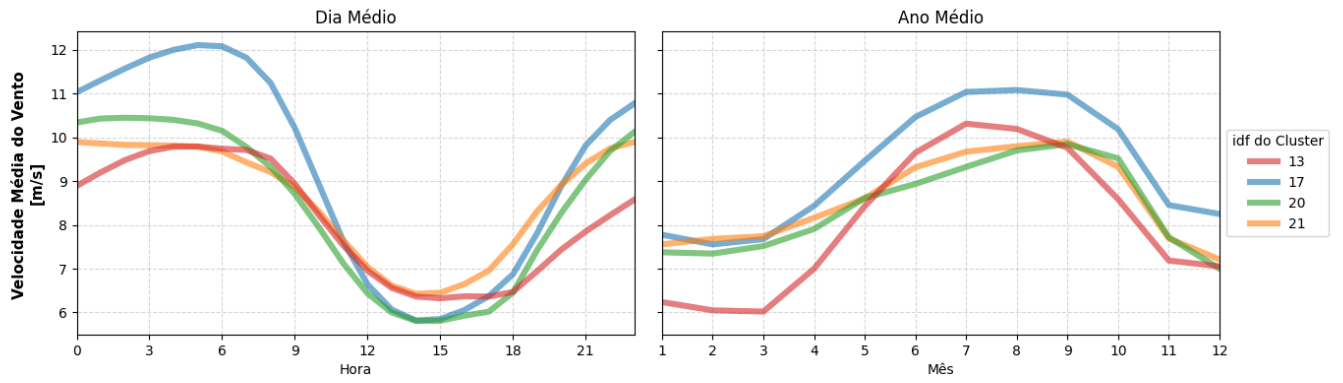


Figura 27 - Dia médio e ano médio da velocidade média do vento nos clusters da Região Interior NE

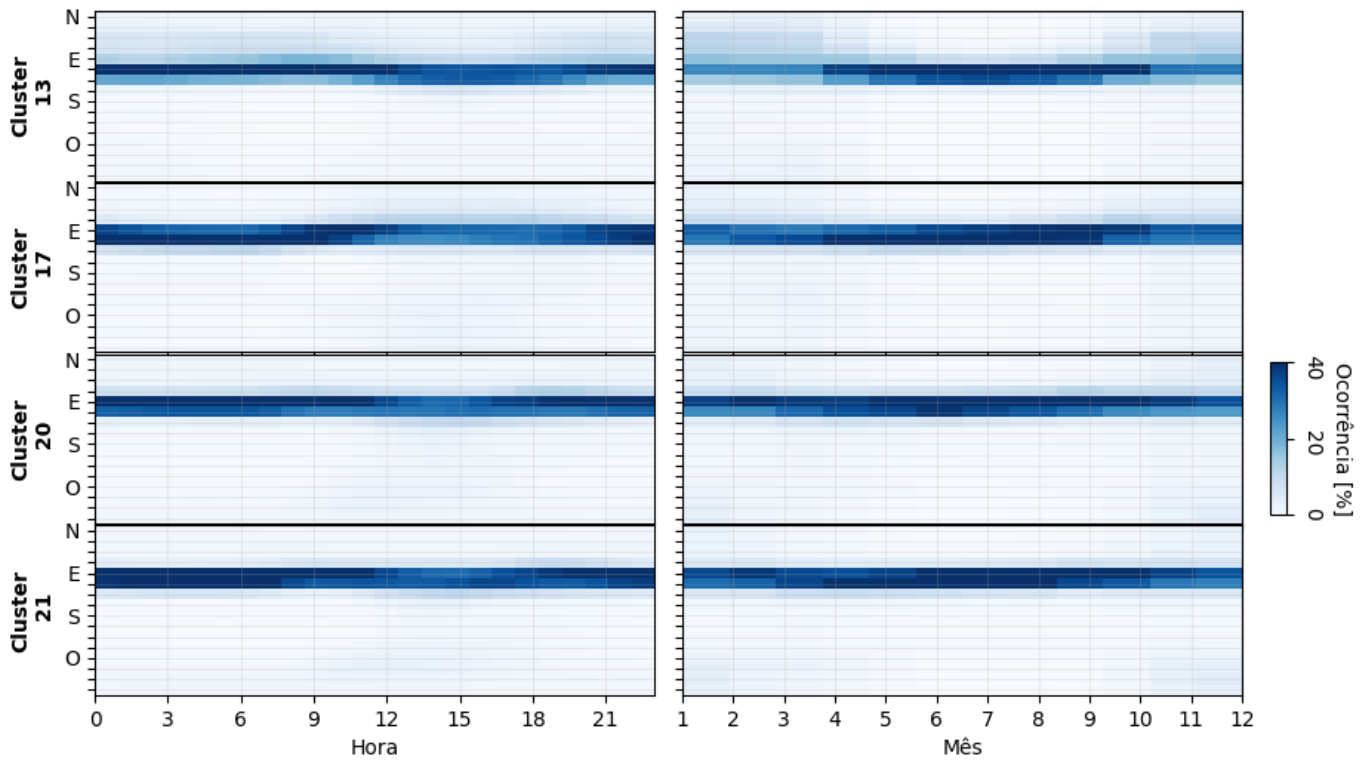


Figura 28 - Distribuição horária e mensal da direção média do vento nos clusters da Região Interior NE

### 3.1.5. Clusters na Região Bahia

Os Clusters da região denominada Bahia apresentam grande quantidade de estações com altitudes superiores a 900 metros, tendo apenas o Cluster 14 estações instaladas em altitudes inferiores a 500 metros. Apesar da proximidade dos Clusters com os da região Interior NE, observa-se que existem distribuições mais bem comportadas para alguns agrupamentos.

O Cluster 22 se caracteriza por apresentar uma distribuição mais comportada que os demais na direção predominante no setor Este. As direções predominantes dos demais clusters estão concentrados nos setores E-SE. Os ápices diários da velocidade média tendem a estar concentrados entre 21hs e 6hs e entre junho e outubro. Nota-se que a variação da direção dos ventos é mínima para todos os Clusters nos períodos diário e mensal.

### Descrição geral da Região Bahia

Tabela 5 - Características dos clusters da Região Bahia

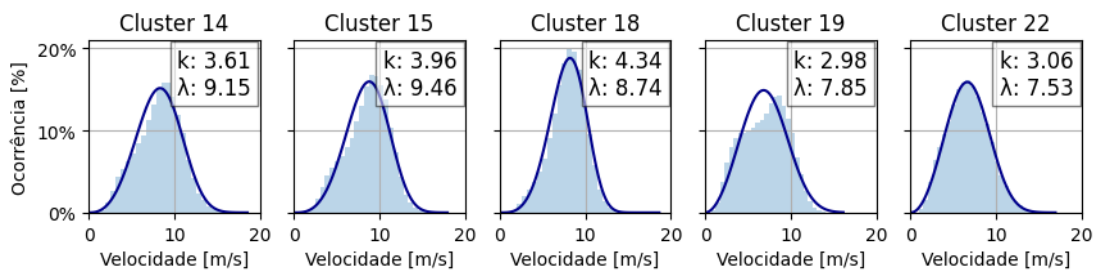
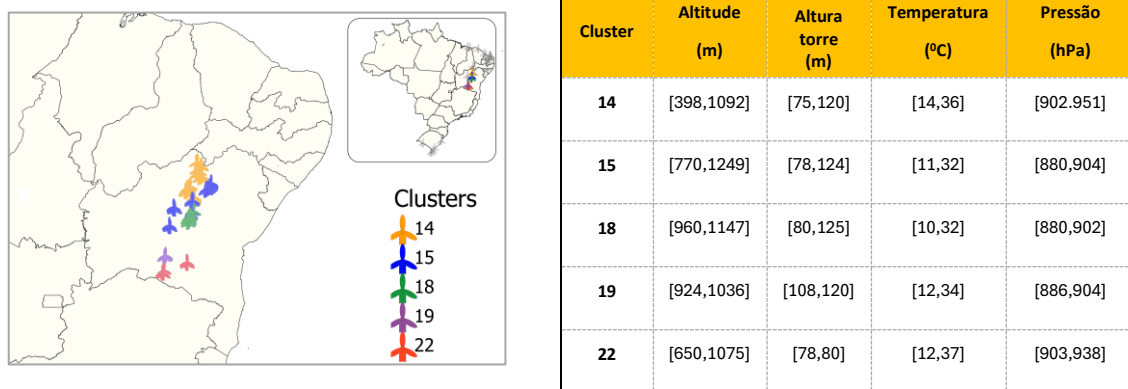


Figura 29 - Distribuição de ocorrências e estimativa de Weibull dos clusters Região Bahia

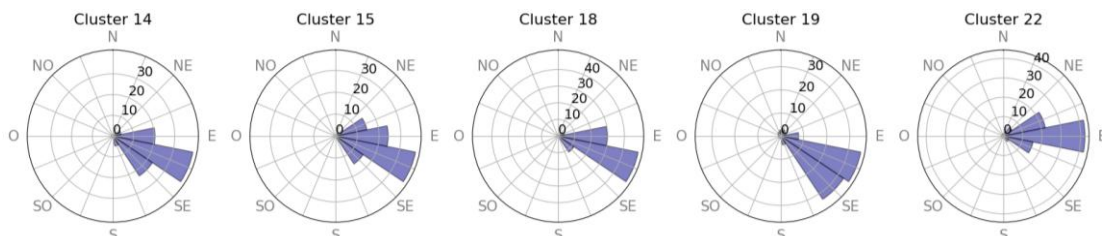


Figura 30 - Rosa dos ventos dos clusters da Região Bahia

## Descrição temporal da Região Bahia

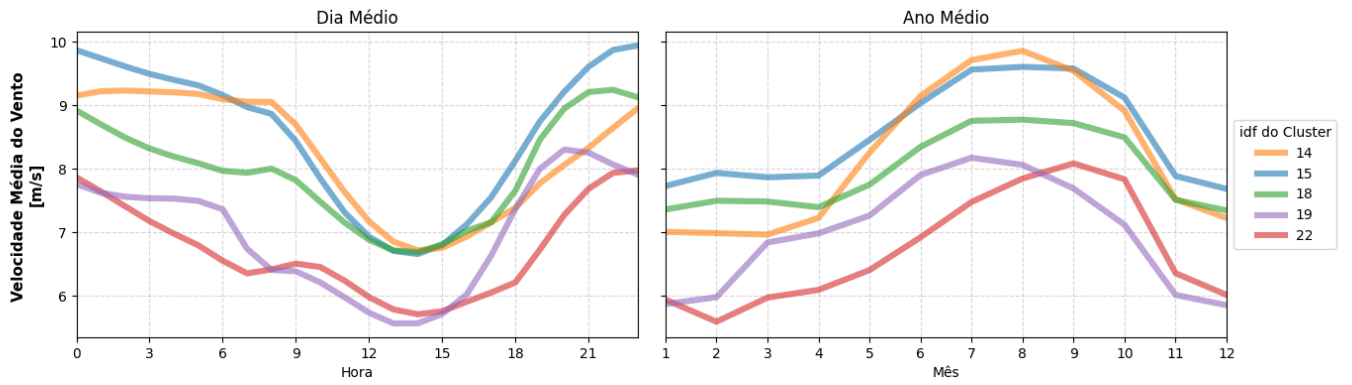


Figura 31 - Dia médio e ano médio da velocidade média do vento nos clusters da Região Bahia

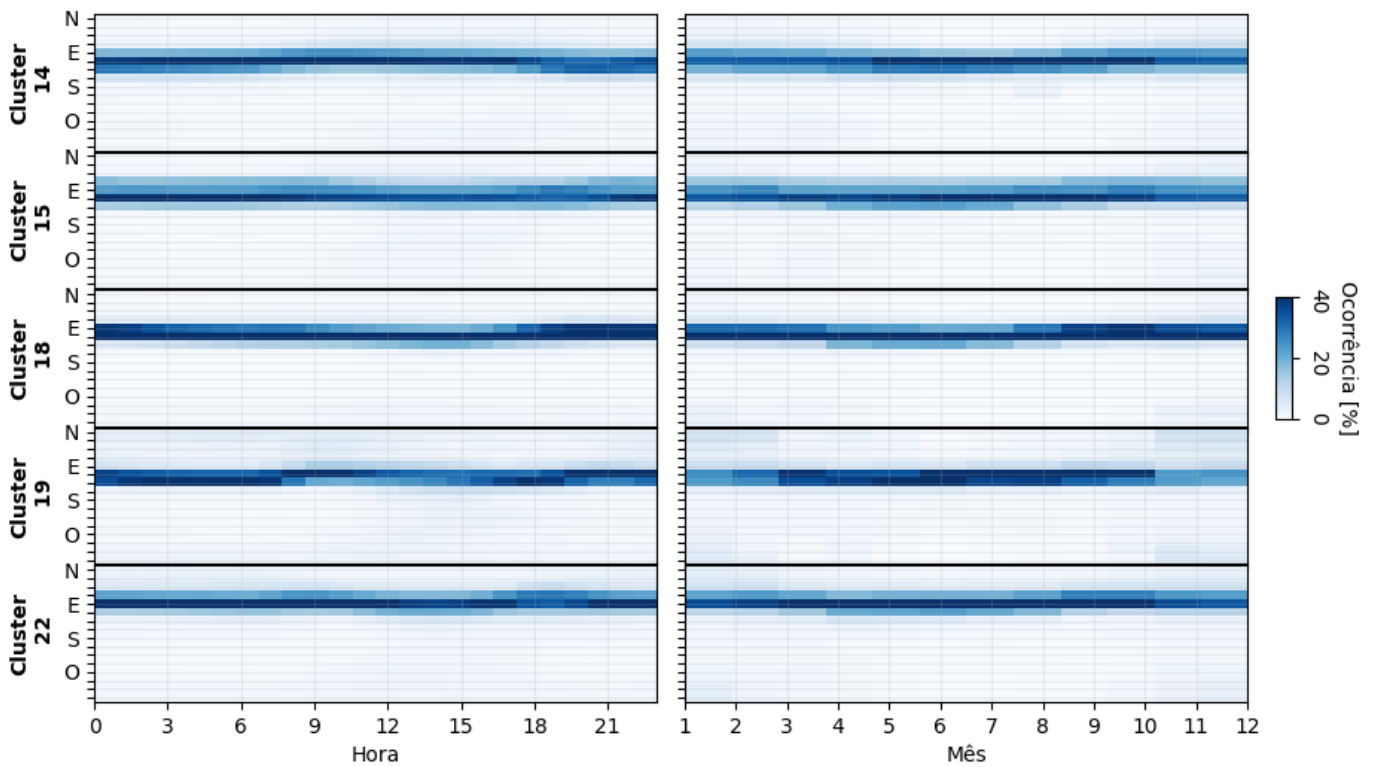


Figura 32 - Distribuição horária e mensal da direção média do vento nos clusters da Região Bahia

### 3.1.6. Clusters na Região Sul

Nos clusters da região Sul, por sua distância geográfica dos demais projetos do país, espera-se que esta região tenha a maior diferenciação com relação as características dos demais clusters do país. Esta premissa é observada graficamente ao comparar a Figura 36, que apresenta a variabilidade da direção do vento ao longo do dia e do ano. Esta variabilidade é mais visível na rosa dos ventos da região, que apresenta direções espalhadas com ocorrências significativas em vários setores, especialmente entre Sul e Nordeste. Nota-se que esta região apresenta também a menor variabilidade de velocidades de vento médias ao longo do ano, com variações na ordem de 2 m/s, padrão que traz maior estabilidade ao operador e ao sistema no ponto de vista sazonal.

É relevante também identificar que na região Sul, o Cluster 27, que é o único localizado na parte interior da região, apresenta comportamento diário diferente das demais, com forte redução da velocidade média no período entre 9hs e 18hs, com valor mínimo obtido com aproximadamente 6 horas de defasagem em relação aos demais clusters da região sul.

#### Apresentação geral da Região Sul

Tabela 6 - Características dos clusters da Região Sul

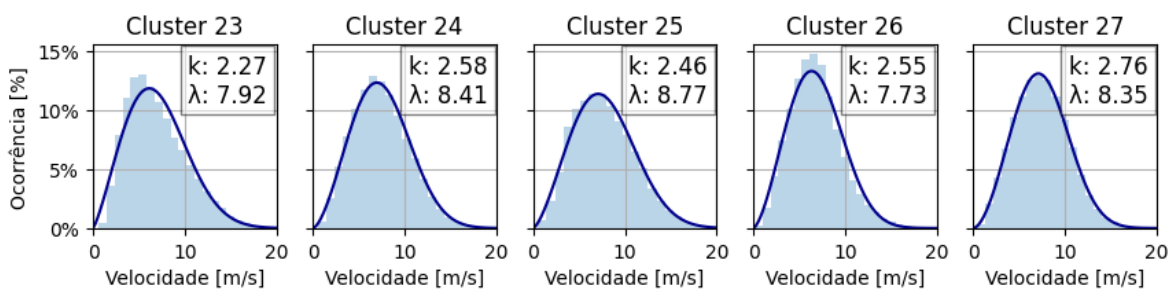
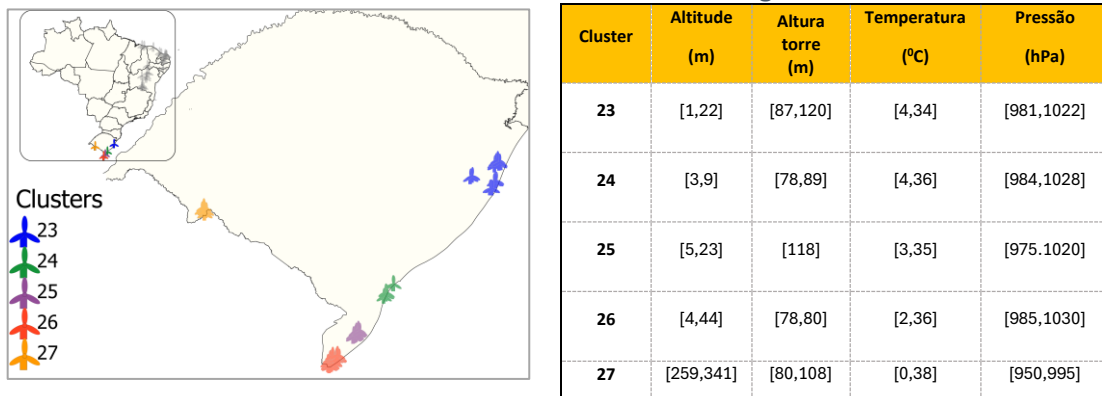


Figura 33 - Distribuição de ocorrências e estimativa de Weibull dos clusters da Região Sul

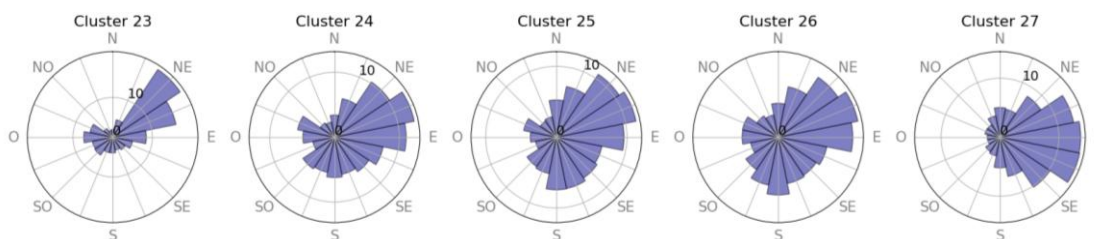


Figura 34 – Rosa dos ventos dos clusters da Região Sul

## Descrição temporal da Região Sul

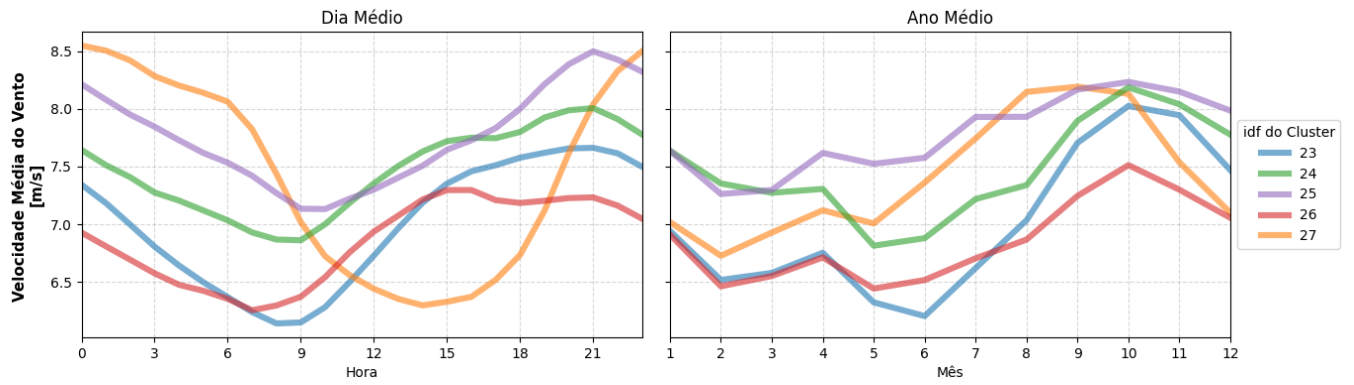


Figura 35 - Dia médio e ano médio da velocidade média do vento nos clusters da Região Sul

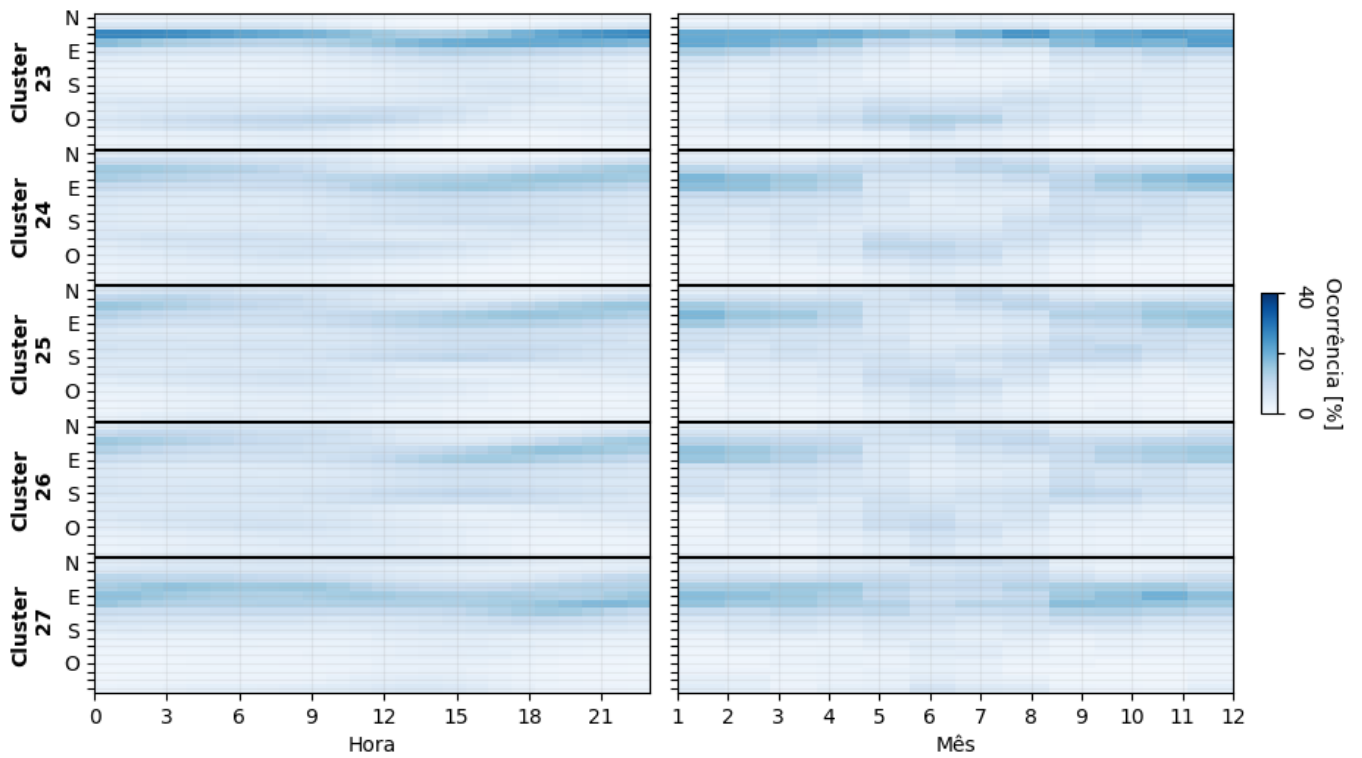


Figura 36 - Distribuição horária e mensal da direção média do vento nos clusters da Região Sul

### 3.1.7. Comentários Gerais

Observa-se a partir da análise conjunta de todos os Clusters que as regiões com maiores altitudes de instalação, tanto na região Nordeste quanto Sul, tendem a apresentar vales durante o dia, nos perfis de dia médio de velocidade, com os menores valores próximos das 15 horas.

Diversas regiões apresentam ocorrências de velocidade média que não são bem ajustadas pela distribuição de Weibull, sendo importante a análise de outras opções de avaliação de distribuições para estas, como identificação de distribuição separando os dados em períodos diferentes ou outras possibilidades.

Quase a totalidade dos clusters apresenta os mínimos de velocidade média entre os meses de fevereiro e maio e ápices entre julho e outubro, sendo estes os meses com maior dispersão de direção para a maioria das observações.

Considerando o comportamento da direção, é possível perceber que os setores próximos a Sudeste apresentam a maior ocorrência, não havendo, porém, nenhum padrão claro ao se observar a distribuição das rosas do vento ao longo do território brasileiro, mostrando que alterações de ocorrência da direção do vento no quadrante sudeste tende a ser bastante influenciada por fatores locais.

## 4 Conclusão

---

Com a aplicação da metodologia de clusterização proposta é possível padronizar um procedimento de agregação de torres meteorológicas em grupos e com isso permitir a disponibilização de uma massa de dados de diversas variáveis de alta qualidade à sociedade, fornecendo uma importante fonte de informação.

Os tratamentos aqui apresentados objetivam garantir o anonimato dos dados individualizados, a partir de um processo de representação com dados médios das regiões, sem mostrar características de cada parque ou detalhes de sua operação mas, ainda assim, representar de forma fidedigna as características do recurso eólico de cada região (ou cluster).

Os dados disponibilizados no dashboard produzido pela EPE<sup>1</sup> visam incentivar o avanço técnico e científico, não apenas no contexto do planejamento energético, mas também a melhoria contínua dos empreendimentos em operação, o desenvolvimento da fonte eólica no país e o ganho da qualidade de informações disponíveis aos diversos setores da economia e sociedade que podem, por estes dados, serem positivamente impactados.

Ressalta-se que melhorias nas metodologias são possíveis e desejáveis conforme o avanço técnico e tecnológico, e que estas atualizações podem refletir modificações tanto nos resultados quanto nas premissas. Desta forma, a EPE se resguarda no direito de atualizar periodicamente a base de dados, contextos e variáveis disponibilizadas e não se responsabiliza pelo uso destes dados por terceiros.

Ressalta-se que estas informações são de caráter consultivo, não devendo ser utilizadas individualmente na elaboração de projetos, certificações de energia e outras atividades, uma vez que os dados aqui disponibilizados não são suficientes ou adequados para a elaboração de empreendimentos eólicos. Informações essenciais para estes projetos devem ser obtidas a partir de campanhas de medição e estudos específicos para cada localidade, sendo as informações aqui indicadas em caráter complementar às medições locais e a outros estudos.

---

<sup>1</sup> Disponível em: <https://dashboard.epe.gov.br/apps/divulgacaoAMA>

## 5 Referências Bibliográficas

---

- [1] Empresa de Pesquisa Energética, “Sistema AMA - 13 anos de acompanhamento de medições anemométricas,” Novembro 2024. [Online]. Available: <https://www.epe.gov.br/sites-pt/publicacoes-dados-abertos/publicacoes/PublicacoesArquivos/publicacao-854/NT-EPE-DEE-SGR-030-2024-R1.pdf>.
- [2] R. R. CABRAL, T. P. TEIXEIRA e T. C. RIZZOTTO, “ESTIMAÇÃO DE PONTOS DE CONEXÃO PARA NOVAS SOLUÇÕES DE TRANSMISSÃO NO CONTEXTO DO PLANEJAMENTO DA EXPANSÃO CONSIDERANDO INCERTEZAS,” XXVI SNTPEE. [Online]. Available: [https://www.epe.gov.br/sites-pt/publicacoes-dados-abertos/publicacoes/PublicacoesArquivos/publicacao-458/topico-666/GPL\\_0931%20ESTIMA%C3%87%C3%83O%20DE%20PONTOS%20DE%20CONEX%C3%83O.pdf](https://www.epe.gov.br/sites-pt/publicacoes-dados-abertos/publicacoes/PublicacoesArquivos/publicacao-458/topico-666/GPL_0931%20ESTIMA%C3%87%C3%83O%20DE%20PONTOS%20DE%20CONEX%C3%83O.pdf).
- [3] A. S. S. T. Y. W. Saeed Aghabozorgi, “Time-series clustering – A decade review,” *Information Systems*, pp. 16-38, 2015.
- [4] G. e. a. Liu, “Big data-informed energy efficiency assessment of China industry sectors based on K-means clustering.,” *Journal of cleaner production*, pp. 304-314, 2018.
- [5] F. M. Gonçalves, “METODOLOGIA GEORREFERENCIA PARA AGRUPAMENTO DE USINAS EÓLICAS E FOTOVOLTAICAS,” em *XXVII Seminário Nacional de Produção e Transporte de Energia Elétrica*, Brasília, 2023. [Online]. Available: [https://www.epe.gov.br/sites-pt/publicacoes-dados-abertos/publicacoes/PublicacoesArquivos/publicacao-458/topico-707/GPL\\_0174.pdf](https://www.epe.gov.br/sites-pt/publicacoes-dados-abertos/publicacoes/PublicacoesArquivos/publicacao-458/topico-707/GPL_0174.pdf)
- [6] Pauli Virtanen e outros, “SciPy 1.0: Algoritmos fundamentais para computação científica em Python,” [Online]. Available: <https://docs.scipy.org/doc/scipy/reference/cluster.hierarchy.html>. [Acesso em dezembro 2024].
- [7] Empresa de Pesquisa Energética, “Energia Eólica no Nordeste, O papel do planejamento energético em períodos de menor geração eólica,” Empresa de Pesquisa Energética, Junho 2020. [Online]. Available: [https://www.epe.gov.br/sites-pt/sala-de-imprensa/noticias/Documents/EPE\\_FactSheet\\_Eolica\\_br.pdf](https://www.epe.gov.br/sites-pt/sala-de-imprensa/noticias/Documents/EPE_FactSheet_Eolica_br.pdf).
- [8] D. S. Christopher Jung, “The annual cycle and intra-annual variability of the global wind power distribution estimated by the system of wind speed,” *Sustainable Energy Technologies and Assessments*, 2020.
- [9] F. T. Leonardo Barriato, “Wind Speed Inter-Annual Variability - Monthly Samples Method,” *Brazil Windpower 2017*, 2017.
- [10] N. University, “Types of Correlation,” [Online]. Available: <https://www.ncl.ac.uk/webtemplate/ask-assets/external/maths-resources/statistics/regression-and-correlation/types-of-correlation.html>. [Acesso em 11 junho 2025].
- [11] F. L. P. Murtagh, “Ward’s Hierarchical Agglomerative Clustering Method: Which Algorithms Implement Ward’s Criterion?,” em *Journal of Classification*, Springer Nature, 2014, pp. 274-295.
- [12] Pauli Virtanen e outros, “Scipy.cluster.hierarchy.ward,” [Online]. Available: <https://docs.scipy.org/doc/scipy/reference/generated/scipy.cluster.hierarchy.ward.html>. [Acesso em 11 junho 2025].